ESSENTIAL MATHEMATICAL METHODS
for the Physical Sciences

STUDENT SOLUTION MANUAL

This *Student Solution Manual* provides complete solutions to all the odd-numbered problems in *Essential Mathematical Methods for the Physical Sciences*. It takes students through each problem step-by-step, so they can clearly see how the solution is reached, and understand any mistakes in their own working. Students will learn by example how to select an appropriate method and improve their problem-solving skills.

CAMBRIDGE
UNIVERSITY PRESS
www.cambridge.org

ISBN 978-0-521-14102-4

9 780521 141024

RILEY | HOBSON

ESSENTIAL MATHEMATICAL METHODS FOR THE PHYSICAL SCIENCES

STUDENT SOLUTION MANUAL

CAMBRIDGE

ESSENTIAL MATHEMATICAL METHODS
for the Physical Sciences

STUDENT SOLUTION MANUAL

K. F. RILEY | M. P. HOBSON

Student Solution Manual for
Essential Mathematical Methods for the Physical Sciences

This *Student Solution Manual* provides complete solutions to all the odd-numbered problems in *Essential Mathematical Methods for the Physical Sciences*. It takes students through each problem step by step, so they can clearly see how the solution is reached, and understand any mistakes in their own working. Students will learn by example how to select an appropriate method, improving their problem-solving skills.

K. F. Riley read mathematics at the University of Cambridge and proceeded to a Ph.D. there in theoretical and experimental nuclear physics. He became a Research Associate in elementary particle physics at Brookhaven, and then, having taken up a lectureship at the Cavendish Laboratory, Cambridge, continued this research at the Rutherford Laboratory and Stanford; in particular he was involved in the experimental discovery of a number of the early baryonic resonances. As well as having been Senior Tutor at Clare College, where he has taught physics and mathematics for over 40 years, he has served on many committees concerned with the teaching and examining of these subjects at all levels of tertiary and undergraduate education. He is also one of the authors of *200 Puzzling Physics Problems* (Cambridge University Press, 2001).

M. P. Hobson read natural sciences at the University of Cambridge, specializing in theoretical physics, and remained at the Cavendish Laboratory to complete a Ph.D. in the physics of star formation. As a Research Fellow at Trinity Hall, Cambridge, and subsequently an Advanced Fellow of the Particle Physics and Astronomy Research Council, he developed an interest in cosmology, and in particular in the study of fluctuations in the cosmic microwave background. He was involved in the first detection of these fluctuations using a ground-based interferometer. Currently a University Reader at the Cavendish Laboratory, his research interests include both theoretical and observational aspects of cosmology, and he is the principal author of *General Relativity: An Introduction for Physicists* (Cambridge University Press, 2006). He is also a Director of Studies in Natural Sciences at Trinity Hall and enjoys an active role in the teaching of undergraduate physics and mathematics.

Essential Mathematical Methods for the Physical Sciences

Student Solution Manual

K. F. RILEY
University of Cambridge

M. P. HOBSON
University of Cambridge

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,
São Paulo, Delhi, Dubai, Tokyo, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521141024

First published 2011

Printed in the United Kingdom at the University Press, Cambridge

A catalogue record for this publication is available from the British Library

ISBN 978-0-521-14102-4 Paperback

Contents

Preface

For reasons that are explained in the preface to *Essential Mathematical Methods for the Physical Sciences* the text of the third edition of *Mathematical Methods for Physics and Engineering* (*MMPE*) (Cambridge: Cambridge University Press, 2006) by Riley, Hobson and Bence, after a number of additions and omissions, has been republished as two slightly overlapping texts. *Essential Mathematical Methods for the Physical Sciences* (*EMMPS*) contains most of the more advanced material, and specifically develops mathematical *methods* that can be applied throughout the physical sciences; an augmented version of the more introductory material, principally concerned with mathematical *tools* rather than methods, is available as *Foundation Mathematics for the Physical Sciences*. The full text of *MMPE*, including all of the more specialized and advanced topics, is still available under its original title.

As in the third edition of *MMPE*, the penultimate subsection of each chapter of *EMMPS* consists of a significant number of problems, nearly all of which are based on topics drawn from several sections of that chapter. Also as in the third edition, hints and outline answers are given in the final subsection, but only to the odd-numbered problems, leaving all even-numbered problems free to be set as unaided homework.

This book is the solutions manual for the problems in *EMMPS*. For the 230 plus *odd-numbered* problems it contains, complete solutions are available, to both students and their teachers, in the form of this manual; these are in addition to the hints and outline answers given in the main text. For each problem, the original question is reproduced and then followed by a fully worked solution. For those original problems that make internal reference to the main text or to other (even-numbered) problems not included in this solutions manual, the questions have been reworded, usually by including additional information, so that the questions can stand alone. Some further minor rewording has been included to improve the page layout.

In many cases the solution given is even fuller than one that might be expected of a good student who has understood the material. This is because we have aimed to make the solutions instructional as well as utilitarian. To this end, we have included comments that are intended to show how the plan for the solution is formulated and have provided the justifications for particular intermediate steps (something not always done, even by the best of students). We have also tried to write each individual substituted formula in the form that best indicates how it was obtained, before simplifying it at the next or a subsequent stage. Where several lines of algebraic manipulation or calculus are needed to obtain a final result, they are normally included in full; this should enable the student to determine whether an incorrect answer is due to a misunderstanding of principles or to a technical error.

As noted above, the original questions are reproduced in full, or in a suitably modified stand-alone form, at the start of each problem. Reference to the main text is not needed provided that standard formulae are known (and a set of tables is available for a few of the statistical and numerical problems). This means that, although it is not its prime purpose, this manual could be used as a test or quiz book by a student who has learned, or thinks that he or she has learned, the material covered in the main text.

1 Matrices and vector spaces

1.1 Which of the following statements about linear vector spaces are true? Where a statement is false, give a counter-example to demonstrate this.

(a) Non-singular $N \times N$ matrices form a vector space of dimension N^2.
(b) Singular $N \times N$ matrices form a vector space of dimension N^2.
(c) Complex numbers form a vector space of dimension 2.
(d) Polynomial functions of x form an infinite-dimensional vector space.
(e) Series $\{a_0, a_1, a_2, \dots, a_N\}$ for which $\sum_{n=0}^{N} |a_n|^2 = 1$ form an N-dimensional vector space.
(f) Absolutely convergent series form an infinite-dimensional vector space.
(g) Convergent series with terms of alternating sign form an infinite-dimensional vector space.

We first remind ourselves that for a set of entities to form a vector space, they must pass five tests: (i) closure under commutative and associative addition; (ii) closure under multiplication by a scalar; (iii) the existence of a null vector in the set; (iv) multiplication by unity leaves any vector unchanged; (v) each vector has a corresponding negative vector.

(a) False. The matrix $\mathbf{0}_N$, the $N \times N$ null matrix, required by (iii) is *not* non-singular and is therefore not in the set.

(b) Consider the sum of $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. The sum is the unit matrix which is not singular and so the set is not closed; this violates requirement (i). The statement is false.

(c) The space is closed under addition and multiplication by a scalar; multiplication by unity leaves a complex number unchanged; there is a null vector ($= 0 + i0$) and a negative complex number for each vector. All the necessary conditions are satisfied and the statement is true.

(d) As in the previous case, all the conditions are satisfied and the statement is true.

(e) This statement is false. To see why, consider $b_n = a_n + a_n$ for which $\sum_{n=0}^{N} |b_n|^2 = 4 \neq 1$, i.e. the set is not closed (violating (i)), or note that there is no zero vector with unit norm (violating (iii)).

(f) True. Note that an absolutely convergent series remains absolutely convergent when the signs of all of its terms are reversed.

(g) False. Consider the two series defined by

$$a_0 = \tfrac{1}{2}, \quad a_n = 2\left(-\tfrac{1}{2}\right)^n \text{ for } n \geq 1; \quad b_n = -\left(-\tfrac{1}{2}\right)^n \text{ for } n \geq 0.$$

The series that is the sum of $\{a_n\}$ and $\{b_n\}$ does not have alternating signs and so closure (required by (i)) does not hold.

1.3 By considering the matrices

$$\mathsf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \mathsf{B} = \begin{pmatrix} 0 & 0 \\ 3 & 4 \end{pmatrix},$$

show that $\mathsf{AB} = \mathsf{0}$ does *not* imply that either A or B is the zero matrix but that it does imply that at least one of them is singular.

We have

$$\mathsf{AB} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Thus AB is the zero matrix $\mathbf{0}$ without either $\mathsf{A} = \mathbf{0}$ or $\mathsf{B} = \mathbf{0}$.

However, $\mathsf{AB} = \mathbf{0} \Rightarrow |\mathsf{A}||\mathsf{B}| = |\mathbf{0}| = 0$ and therefore either $|\mathsf{A}| = 0$ or $|\mathsf{B}| = 0$ (or both).

1.5 Using the properties of determinants, solve with a minimum of calculation the following equations for x:

$$\text{(a)} \quad \begin{vmatrix} x & a & a & 1 \\ a & x & b & 1 \\ a & b & x & 1 \\ a & b & c & 1 \end{vmatrix} = 0, \qquad \text{(b)} \quad \begin{vmatrix} x+2 & x+4 & x-3 \\ x+3 & x & x+5 \\ x-2 & x-1 & x+1 \end{vmatrix} = 0.$$

(a) In view of the similarities between some rows and some columns, the property most likely to be useful here is that if a determinant has two rows/columns equal (or multiples of each other) then its value is zero.

(i) We note that setting $x = a$ makes the first and fourth columns multiples of each other and hence makes the value of the determinant 0; thus $x = a$ is one solution to the equation.

(ii) Setting $x = b$ makes the second and third rows equal, and again the determinant vanishes; thus b is another root of the equation.

(iii) Setting $x = c$ makes the third and fourth rows equal, and yet again the determinant vanishes; thus c is also a root of the equation.

Since the determinant contains no x in its final column, it is a cubic polynomial in x and there will be exactly three roots to the equation. We have already found all three!

(b) Here, the presence of x multiplied by unity in every entry means that subtracting rows/columns will lead to a simplification. After (i) subtracting the first column from each of the others, and then (ii) subtracting the first row from each of the others, the determinant becomes

$$\begin{aligned} \begin{vmatrix} x+2 & 2 & -5 \\ x+3 & -3 & 2 \\ x-2 & 1 & 3 \end{vmatrix} &= \begin{vmatrix} x+2 & 2 & -5 \\ 1 & -5 & 7 \\ -4 & -1 & 8 \end{vmatrix} \\ &= (x+2)(-40+7) + 2(-28-8) - 5(-1-20) \\ &= -33(x+2) - 72 + 105 \\ &= -33x - 33. \end{aligned}$$

Thus $x = -1$ is the only solution to the original (linear!) equation.

1.7 Prove the following results involving Hermitian matrices.

(a) If A is Hermitian and U is unitary then $\mathsf{U}^{-1}\mathsf{A}\mathsf{U}$ is Hermitian.
(b) If A is anti-Hermitian then $i\mathsf{A}$ is Hermitian.
(c) The product of two Hermitian matrices A and B is Hermitian if and only if A and B commute.
(d) If S is a real antisymmetric matrix then $\mathsf{A} = (\mathsf{I} - \mathsf{S})(\mathsf{I} + \mathsf{S})^{-1}$ is orthogonal. If A is given by

$$\mathsf{A} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

then find the matrix S that is needed to express A in the above form.
(e) If K is skew-Hermitian, i.e. $\mathsf{K}^\dagger = -\mathsf{K}$, then $\mathsf{V} = (\mathsf{I} + \mathsf{K})(\mathsf{I} - \mathsf{K})^{-1}$ is unitary.

The general properties of matrices that we will need are $(\mathsf{A}^\dagger)^{-1} = (\mathsf{A}^{-1})^\dagger$ and

$$(\mathsf{A}\mathsf{B}\cdots\mathsf{C})^{\mathrm{T}} = \mathsf{C}^{\mathrm{T}}\cdots\mathsf{B}^{\mathrm{T}}\mathsf{A}^{\mathrm{T}}, \qquad (\mathsf{A}\mathsf{B}\cdots\mathsf{C})^\dagger = \mathsf{C}^\dagger\cdots\mathsf{B}^\dagger\mathsf{A}^\dagger.$$

(a) Given that $\mathsf{A} = \mathsf{A}^\dagger$ and $\mathsf{U}^\dagger\mathsf{U} = \mathsf{I}$, consider

$$(\mathsf{U}^{-1}\mathsf{A}\mathsf{U})^\dagger = \mathsf{U}^\dagger\mathsf{A}^\dagger(\mathsf{U}^{-1})^\dagger = \mathsf{U}^{-1}\mathsf{A}(\mathsf{U}^\dagger)^{-1} = \mathsf{U}^{-1}\mathsf{A}(\mathsf{U}^{-1})^{-1} = \mathsf{U}^{-1}\mathsf{A}\mathsf{U},$$

i.e. $\mathsf{U}^{-1}\mathsf{A}\mathsf{U}$ is Hermitian.

(b) Given $\mathsf{A}^\dagger = -\mathsf{A}$, consider

$$(i\mathsf{A})^\dagger = -i\mathsf{A}^\dagger = -i(-\mathsf{A}) = i\mathsf{A},$$

i.e. $i\mathsf{A}$ is Hermitian.

(c) Given $\mathsf{A} = \mathsf{A}^\dagger$ and $\mathsf{B} = \mathsf{B}^\dagger$.

(i) Suppose $\mathsf{A}\mathsf{B} = \mathsf{B}\mathsf{A}$, then

$$(\mathsf{A}\mathsf{B})^\dagger = \mathsf{B}^\dagger\mathsf{A}^\dagger = \mathsf{B}\mathsf{A} = \mathsf{A}\mathsf{B},$$

i.e. $\mathsf{A}\mathsf{B}$ is Hermitian.

(ii) Now suppose that $(\mathsf{A}\mathsf{B})^\dagger = \mathsf{A}\mathsf{B}$. Then

$$\mathsf{B}\mathsf{A} = \mathsf{B}^\dagger\mathsf{A}^\dagger = (\mathsf{A}\mathsf{B})^\dagger = \mathsf{A}\mathsf{B},$$

i.e. A and B commute.

Thus, $\mathsf{A}\mathsf{B}$ is Hermitian $\Longleftrightarrow$ A and B commute.

(d) Given that S is real and $\mathsf{S}^{\mathrm{T}} = -\mathsf{S}$ with $\mathsf{A} = (\mathsf{I} - \mathsf{S})(\mathsf{I} + \mathsf{S})^{-1}$, consider

$$\begin{aligned}
\mathsf{A}^{\mathrm{T}}\mathsf{A} &= [(\mathsf{I} - \mathsf{S})(\mathsf{I} + \mathsf{S})^{-1}]^{\mathrm{T}}[(\mathsf{I} - \mathsf{S})(\mathsf{I} + \mathsf{S})^{-1}] \\
&= [(\mathsf{I} + \mathsf{S})^{-1}]^{\mathrm{T}}(\mathsf{I} + \mathsf{S})(\mathsf{I} - \mathsf{S})(\mathsf{I} + \mathsf{S})^{-1} \\
&= (\mathsf{I} - \mathsf{S})^{-1}(\mathsf{I} + \mathsf{S} - \mathsf{S} - \mathsf{S}^2)(\mathsf{I} + \mathsf{S})^{-1} \\
&= (\mathsf{I} - \mathsf{S})^{-1}(\mathsf{I} - \mathsf{S})(\mathsf{I} + \mathsf{S})(\mathsf{I} + \mathsf{S})^{-1} \\
&= \mathsf{I}\,\mathsf{I} = \mathsf{I},
\end{aligned}$$

i.e. A is orthogonal.

If $\mathsf{A} = (\mathsf{I} - \mathsf{S})(\mathsf{I} + \mathsf{S})^{-1}$, then $\mathsf{A} + \mathsf{AS} = \mathsf{I} - \mathsf{S}$ and $(\mathsf{A} + \mathsf{I})\mathsf{S} = \mathsf{I} - \mathsf{A}$, giving

$$\begin{aligned}
\mathsf{S} &= (\mathsf{A} + \mathsf{I})^{-1}(\mathsf{I} - \mathsf{A}) \\
&= \begin{pmatrix} 1 + \cos\theta & \sin\theta \\ -\sin\theta & 1 + \cos\theta \end{pmatrix}^{-1} \begin{pmatrix} 1 - \cos\theta & -\sin\theta \\ \sin\theta & 1 - \cos\theta \end{pmatrix} \\
&= \frac{1}{2 + 2\cos\theta} \begin{pmatrix} 1 + \cos\theta & -\sin\theta \\ \sin\theta & 1 + \cos\theta \end{pmatrix} \begin{pmatrix} 1 - \cos\theta & -\sin\theta \\ \sin\theta & 1 - \cos\theta \end{pmatrix} \\
&= \frac{1}{4\cos^2(\theta/2)} \begin{pmatrix} 0 & -2\sin\theta \\ 2\sin\theta & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & -\tan(\theta/2) \\ \tan(\theta/2) & 0 \end{pmatrix}.
\end{aligned}$$

(e) This proof is almost identical to the first section of part (d) but with S replaced by $-\mathsf{K}$ and transposed matrices replaced by Hermitian conjugate matrices.

1.9 The *commutator* $[\mathsf{X}, \mathsf{Y}]$ of two matrices is defined by the equation

$$[\mathsf{X}, \mathsf{Y}] = \mathsf{XY} - \mathsf{YX}.$$

Two anticommuting matrices A and B satisfy

$$\mathsf{A}^2 = \mathsf{I}, \qquad \mathsf{B}^2 = \mathsf{I}, \qquad [\mathsf{A}, \mathsf{B}] = 2i\mathsf{C}.$$

(a) Prove that $\mathsf{C}^2 = \mathsf{I}$ and that $[\mathsf{B}, \mathsf{C}] = 2i\mathsf{A}$.
(b) Evaluate $[[[\mathsf{A}, \mathsf{B}], [\mathsf{B}, \mathsf{C}]], [\mathsf{A}, \mathsf{B}]]$.

(a) From $\mathsf{AB} - \mathsf{BA} = 2i\mathsf{C}$ and $\mathsf{AB} = -\mathsf{BA}$ it follows that $\mathsf{AB} = i\mathsf{C}$. Thus,

$$-\mathsf{C}^2 = i\mathsf{C}i\mathsf{C} = \mathsf{ABAB} = \mathsf{A}(-\mathsf{AB})\mathsf{B} = -(\mathsf{AA})(\mathsf{BB}) = -\mathsf{I}\mathsf{I} = -\mathsf{I},$$

i.e. $\mathsf{C}^2 = \mathsf{I}$. In deriving the above result we have used the associativity of matrix multiplication.

For the commutator of B and C,

$$\begin{aligned}
[\mathsf{B}, \mathsf{C}] &= \mathsf{BC} - \mathsf{CB} \\
&= \mathsf{B}(-i\mathsf{AB}) - (-i)\mathsf{ABB} \\
&= -i(\mathsf{BA})\mathsf{B} + i\mathsf{AI} \\
&= -i(-\mathsf{AB})\mathsf{B} + i\mathsf{A} \\
&= i\mathsf{A} + i\mathsf{A} = 2i\mathsf{A}.
\end{aligned}$$

(b) To evaluate this multiple-commutator expression we must work outwards from the innermost "explicit" commutators. There are three such commutators at the first stage. We also need the result that $[\mathsf{C}, \mathsf{A}] = 2i\mathsf{B}$; this can be proved in the same way as that for $[\mathsf{B}, \mathsf{C}]$ in part (a), or by making the cyclic replacements $A \to B \to C \to A$ in the assumptions and their consequences, as proved in part (a). Then we have

$$\begin{aligned}\big[\big[\,[\,\mathsf{A},\mathsf{B}\,],[\,\mathsf{B},\mathsf{C}\,]\big],[\,\mathsf{A},\mathsf{B}\,]\big] &= [\,[\,2i\mathsf{C},2i\mathsf{A}\,],2i\mathsf{C}\,]\\ &= -4[\,[\,\mathsf{C},\mathsf{A}\,],2i\mathsf{C}\,]\\ &= -4[\,2i\mathsf{B},2i\mathsf{C}\,]\\ &= (-4)(-4)[\,\mathsf{B},\mathsf{C}\,] = 32i\mathsf{A}.\end{aligned}$$

1.11 A general triangle has angles α, β and γ and corresponding opposite sides a, b and c. Express the length of each side in terms of the lengths of the other two sides and the relevant cosines, writing the relationships in matrix and vector form, using the vectors having components a, b, c and $\cos\alpha$, $\cos\beta$, $\cos\gamma$. Invert the matrix and hence deduce the cosine-law expressions involving α, β and γ.

By considering each side of the triangle as the sum of the projections onto it of the other two sides, we have the three simultaneous equations:

$$\begin{aligned} a &= b\cos\gamma + c\cos\beta,\\ b &= c\cos\alpha + a\cos\gamma,\\ c &= b\cos\alpha + a\cos\beta.\end{aligned}$$

Written in matrix and vector form, $\mathsf{A}\mathbf{x} = \mathbf{y}$, they become

$$\begin{pmatrix} 0 & c & b\\ c & 0 & a\\ b & a & 0\end{pmatrix}\begin{pmatrix}\cos\alpha\\ \cos\beta\\ \cos\gamma\end{pmatrix} = \begin{pmatrix} a\\ b\\ c\end{pmatrix}.$$

The matrix A is non-singular, since $|\,\mathsf{A}\,| = 2abc \neq 0$, and therefore has an inverse given by

$$\mathsf{A}^{-1} = \frac{1}{2abc}\begin{pmatrix} -a^2 & ab & ac\\ ab & -b^2 & bc\\ ac & bc & -c^2\end{pmatrix}.$$

And so, writing $\mathbf{x} = \mathsf{A}^{-1}\mathbf{y}$, we have

$$\begin{pmatrix}\cos\alpha\\ \cos\beta\\ \cos\gamma\end{pmatrix} = \frac{1}{2abc}\begin{pmatrix} -a^2 & ab & ac\\ ab & -b^2 & bc\\ ac & bc & -c^2\end{pmatrix}\begin{pmatrix} a\\ b\\ c\end{pmatrix}.$$

From this we can read off the cosine-law equation

$$\cos\alpha = \frac{1}{2abc}(-a^3 + ab^2 + ac^2) = \frac{b^2 + c^2 - a^2}{2bc},$$

and the corresponding expressions for $\cos\beta$ and $\cos\gamma$.

1.13 Determine which of the matrices below are mutually commuting, and, for those that are, demonstrate that they have a complete set of eigenvectors in common:

$$\mathsf{A} = \begin{pmatrix} 6 & -2 \\ -2 & 9 \end{pmatrix}, \quad \mathsf{B} = \begin{pmatrix} 1 & 8 \\ 8 & -11 \end{pmatrix},$$
$$\mathsf{C} = \begin{pmatrix} -9 & -10 \\ -10 & 5 \end{pmatrix}, \quad \mathsf{D} = \begin{pmatrix} 14 & 2 \\ 2 & 11 \end{pmatrix}.$$

To establish the result we need to examine all pairs of products.

$$\begin{aligned}
\mathsf{AB} &= \begin{pmatrix} 6 & -2 \\ -2 & 9 \end{pmatrix}\begin{pmatrix} 1 & 8 \\ 8 & -11 \end{pmatrix} \\
&= \begin{pmatrix} -10 & 70 \\ 70 & -115 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 8 \\ 8 & -11 \end{pmatrix}\begin{pmatrix} 6 & -2 \\ -2 & 9 \end{pmatrix} = \mathsf{BA}. \\
\mathsf{AC} &= \begin{pmatrix} 6 & -2 \\ -2 & 9 \end{pmatrix}\begin{pmatrix} -9 & -10 \\ -10 & 5 \end{pmatrix} \\
&= \begin{pmatrix} -34 & -70 \\ -72 & 65 \end{pmatrix} \neq \begin{pmatrix} -34 & -72 \\ -70 & 65 \end{pmatrix} \\
&= \begin{pmatrix} -9 & -10 \\ -10 & 5 \end{pmatrix}\begin{pmatrix} 6 & -2 \\ -2 & 9 \end{pmatrix} = \mathsf{CA}.
\end{aligned}$$

Continuing in this way, we find:

$$\begin{aligned}
\mathsf{AD} &= \begin{pmatrix} 80 & -10 \\ -10 & 95 \end{pmatrix} = \mathsf{DA}. \\
\mathsf{BC} &= \begin{pmatrix} -89 & 30 \\ 38 & -135 \end{pmatrix} \neq \begin{pmatrix} -89 & 38 \\ 30 & -135 \end{pmatrix} = \mathsf{CB}. \\
\mathsf{BD} &= \begin{pmatrix} 30 & 90 \\ 90 & -105 \end{pmatrix} = \mathsf{DB}. \\
\mathsf{CD} &= \begin{pmatrix} -146 & -128 \\ -130 & 35 \end{pmatrix} \neq \begin{pmatrix} -146 & -130 \\ -128 & 35 \end{pmatrix} = \mathsf{DC}.
\end{aligned}$$

These results show that whilst A, B and D are mutually commuting, none of them commutes with C.

We could use any of the three mutually commuting matrices to find the common set (actually a pair, as they are 2×2 matrices) of eigenvectors. We arbitrarily choose A. The eigenvalues of A satisfy

$$\begin{aligned}
\begin{vmatrix} 6-\lambda & -2 \\ -2 & 9-\lambda \end{vmatrix} &= 0, \\
\lambda^2 - 15\lambda + 50 &= 0, \\
(\lambda - 5)(\lambda - 10) &= 0.
\end{aligned}$$

For $\lambda = 5$, an eigenvector $(x \quad y)^{\mathrm{T}}$ must satisfy $x - 2y = 0$, whilst, for $\lambda = 10$, $4x + 2y = 0$. Thus a pair of independent eigenvectors of A are $(2 \quad 1)^{\mathrm{T}}$ and $(1 \quad -2)^{\mathrm{T}}$. Direct substitution verifies that they are also eigenvectors of B and D with pairs of eigenvalues $5, -15$ and $15, 10$, respectively.

1.15 Solve the simultaneous equations

$$\begin{aligned} 2x + 3y + z &= 11, \\ x + y + z &= 6, \\ 5x - y + 10z &= 34. \end{aligned}$$

To eliminate z, (i) subtract the second equation from the first and (ii) subtract 10 times the second equation from the third.

$$\begin{aligned} x + 2y &= 5, \\ -5x - 11y &= -26. \end{aligned}$$

To eliminate x add 5 times the first equation to the second

$$-y = -1.$$

Thus $y = 1$ and, by resubstitution, $x = 3$ and $z = 2$.

1.17 Show that the following equations have solutions only if $\eta = 1$ or 2, and find them in these cases:

$$\begin{aligned} x + y + z &= 1, & \text{(i)} \\ x + 2y + 4z &= \eta, & \text{(ii)} \\ x + 4y + 10z &= \eta^2. & \text{(iii)} \end{aligned}$$

Expressing the equations in the form $\mathsf{Ax} = \mathsf{b}$, we first need to evaluate $|\mathsf{A}|$ as a preliminary to determining A^{-1}. However, we find that $|\mathsf{A}| = 1(20 - 16) + 1(4 - 10) + 1(4 - 2) = 0$. This result implies both that A is singular and has no inverse, and that the equations must be linearly dependent.

Either by observation or by solving for the combination coefficients, we see that for the LHS this linear dependence is expressed by

$$2 \times \text{(i)} + 1 \times \text{(iii)} - 3 \times \text{(ii)} = 0.$$

For a consistent solution, this must also be true for the RHSs, i.e.

$$2 + \eta^2 - 3\eta = 0.$$

This quadratic equation has solutions $\eta = 1$ and $\eta = 2$, which are therefore the only values of η for which the original equations have a solution. As the equations are linearly dependent, we may use any two to find these allowed solutions; for simplicity we use the first two in each case.

For $\eta = 1$,

$$x + y + z = 1, \quad x + 2y + 4z = 1 \Rightarrow \mathsf{x}^1 = (1 + 2\alpha \quad -3\alpha \quad \alpha)^{\mathrm{T}}.$$

For $\eta = 2$,

$$x + y + z = 1, \quad x + 2y + 4z = 2 \Rightarrow \mathsf{x}^2 = (2\alpha \quad 1 - 3\alpha \quad \alpha)^{\mathrm{T}}.$$

In both cases there is an infinity of solutions as α may take any finite value.

1.19 Make an LU decomposition of the matrix

$$\mathsf{A} = \begin{pmatrix} 3 & 6 & 9 \\ 1 & 0 & 5 \\ 2 & -2 & 16 \end{pmatrix}$$

and hence solve $\mathsf{Ax} = \mathsf{b}$, where (i) $\mathsf{b} = (21 \quad 9 \quad 28)^{\mathrm{T}}$, (ii) $\mathsf{b} = (21 \quad 7 \quad 22)^{\mathrm{T}}$.

Using the notation

$$\mathsf{A} = \begin{pmatrix} 1 & 0 & 0 \\ L_{21} & 1 & 0 \\ L_{31} & L_{32} & 1 \end{pmatrix} \begin{pmatrix} U_{11} & U_{12} & U_{13} \\ 0 & U_{22} & U_{23} \\ 0 & 0 & U_{33} \end{pmatrix},$$

and considering rows and columns alternately in the usual way for an LU decomposition, we require the following to be satisfied.

1st row: $U_{11} = 3, \quad U_{12} = 6, \quad U_{13} = 9.$
1st col: $L_{21}U_{11} = 1, \quad L_{31}U_{11} = 2 \Rightarrow L_{21} = \frac{1}{3}, \quad L_{31} = \frac{2}{3}.$
2nd row: $L_{21}U_{12} + U_{22} = 0, \quad L_{21}U_{13} + U_{23} = 5 \Rightarrow U_{22} = -2, \quad U_{23} = 2.$
2nd col: $L_{31}U_{12} + L_{32}U_{22} = -2 \Rightarrow L_{32} = 3.$
3rd row: $L_{31}U_{13} + L_{32}U_{23} + U_{33} = 16 \Rightarrow U_{33} = 4.$

Thus

$$\mathsf{L} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ \frac{2}{3} & 3 & 1 \end{pmatrix} \quad \text{and} \quad \mathsf{U} = \begin{pmatrix} 3 & 6 & 9 \\ 0 & -2 & 2 \\ 0 & 0 & 4 \end{pmatrix}.$$

To solve $\mathsf{Ax} = \mathsf{b}$ with $\mathsf{A} = \mathsf{LU}$, we first determine y from $\mathsf{Ly} = \mathsf{b}$ and then solve $\mathsf{Ux} = \mathsf{y}$ for x.

(i) For $\mathsf{Ax} = (21 \quad 9 \quad 28)^{\mathrm{T}}$, we first solve

$$\begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ \frac{2}{3} & 3 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 21 \\ 9 \\ 28 \end{pmatrix}.$$

This can be done, almost by inspection, to give $\mathsf{y} = (21 \quad 2 \quad 8)^{\mathrm{T}}$.

We can now write $\mathsf{Ux} = \mathsf{y}$ explicitly as

$$\begin{pmatrix} 3 & 6 & 9 \\ 0 & -2 & 2 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 21 \\ 2 \\ 8 \end{pmatrix}$$

to give, equally easily, that the solution to the original matrix equation is $\mathsf{x} = (-1 \quad 1 \quad 2)^{\mathrm{T}}$.

(ii) To solve $\mathsf{A}\mathsf{x} = (21 \quad 7 \quad 22)^{\mathrm{T}}$ we use exactly the same forms for L and U, but the new values for the components of b, to obtain $\mathsf{y} = (21 \quad 0 \quad 8)^{\mathrm{T}}$ leading to the solution $\mathsf{x} = (-3 \quad 2 \quad 2)^{\mathrm{T}}$.

1.21 Use the Cholesky decomposition method to determine whether the following matrices are positive definite. For each that is, determine the corresponding lower diagonal matrix L :

$$\mathsf{A} = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 3 & -1 \\ 3 & -1 & 1 \end{pmatrix}, \qquad \mathsf{B} = \begin{pmatrix} 5 & 0 & \sqrt{3} \\ 0 & 3 & 0 \\ \sqrt{3} & 0 & 3 \end{pmatrix}.$$

The matrix A is real and so we seek a real lower-diagonal matrix L such that $\mathsf{L}\mathsf{L}^{\mathrm{T}} = \mathsf{A}$. In order to avoid a lot of subscripts, we use lower-case letters as the non-zero elements of L:

$$\begin{pmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{pmatrix} \begin{pmatrix} a & b & d \\ 0 & c & e \\ 0 & 0 & f \end{pmatrix} = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 3 & -1 \\ 3 & -1 & 1 \end{pmatrix}.$$

Firstly, from A_{11}, $a^2 = 2$. Since an overall negative sign multiplying the elements of L is irrelevant, we may choose $a = +\sqrt{2}$. Next, $ba = A_{12} = 1$, implying that $b = 1/\sqrt{2}$. Similarly, $d = 3/\sqrt{2}$.

From the second row of A we have

$$b^2 + c^2 = 3 \quad \Rightarrow \quad c = \sqrt{\tfrac{5}{2}},$$

$$bd + ce = -1 \quad \Rightarrow \quad e = \sqrt{\tfrac{2}{5}}\left(-1 - \tfrac{3}{2}\right) = -\sqrt{\tfrac{5}{2}}.$$

And, from the final row,

$$d^2 + e^2 + f^2 = 1 \quad \Rightarrow \quad f = \left(1 - \tfrac{9}{2} - \tfrac{5}{2}\right)^{1/2} = \sqrt{-6}.$$

That f is imaginary shows that A is not a positive definite matrix.

The corresponding argument (keeping the same symbols but with different numerical values) for the matrix B is as follows.

Firstly, from A_{11}, $a^2 = 5$. Since an overall negative sign multiplying the elements of L is irrelevant, we may choose $a = +\sqrt{5}$. Next, $ba = B_{12} = 0$, implying that $b = 0$. Similarly, $d = \sqrt{3}/\sqrt{5}$.

From the second row of B we have

$$b^2 + c^2 = 3 \quad \Rightarrow \quad c = \sqrt{3},$$

$$bd + ce = 0 \quad \Rightarrow \quad e = \sqrt{\tfrac{1}{3}}(0 - 0) = 0.$$

And, from the final row,

$$d^2 + e^2 + f^2 = 3 \quad \Rightarrow \quad f = (3 - \tfrac{3}{5} - 0)^{1/2} = \sqrt{\tfrac{12}{5}}.$$

Thus all the elements of L have been calculated and found to be real and, in summary,

$$\mathsf{L} = \begin{pmatrix} \sqrt{5} & 0 & 0 \\ 0 & \sqrt{3} & 0 \\ \sqrt{\frac{3}{5}} & 0 & \sqrt{\frac{12}{5}} \end{pmatrix}.$$

That $\mathsf{LL}^{\mathrm{T}} = \mathsf{B}$ can be confirmed by substitution.

1.23 Find three real orthogonal column matrices, each of which is a simultaneous eigenvector of

$$\mathsf{A} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathsf{B} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

We first note that

$$\mathsf{AB} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \mathsf{BA}.$$

The two matrices commute and so they *will* have a common set of eigenvectors.

The eigenvalues of A are given by

$$\begin{vmatrix} -\lambda & 0 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 0 & -\lambda \end{vmatrix} = (1-\lambda)(\lambda^2 - 1) = 0,$$

i.e. $\lambda = 1$, $\lambda = 1$ and $\lambda = -1$, with corresponding eigenvectors $\mathbf{e}^1 = (1 \quad y_1 \quad 1)^{\mathrm{T}}$, $\mathbf{e}^2 = (1 \quad y_2 \quad 1)^{\mathrm{T}}$ and $\mathbf{e}^3 = (1 \quad 0 \quad -1)^{\mathrm{T}}$. For these to be mutually orthogonal requires that $y_1 y_2 = -2$.

The third vector, $\mathbf{e}^3$, is clearly an eigenvector of B with eigenvalue $\mu_3 = -1$. For $\mathbf{e}^1$ or $\mathbf{e}^2$ to be an eigenvector of B with eigenvalue μ requires

$$\begin{pmatrix} 0-\mu & 1 & 1 \\ 1 & 0-\mu & 1 \\ 1 & 1 & 0-\mu \end{pmatrix} \begin{pmatrix} 1 \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix};$$

$$\begin{aligned} &\text{i.e.} && -\mu + y + 1 = 0, \\ &\text{and} && 1 - \mu y + 1 = 0, \\ &\text{giving} && -\frac{2}{y} + y + 1 = 0, \\ &\Rightarrow && y^2 + y - 2 = 0, \\ &&\Rightarrow\quad & y = 1 \quad \text{or} \quad -2. \end{aligned}$$

Thus, $y_1 = 1$ with $\mu_1 = 2$, whilst $y_2 = -2$ with $\mu_2 = -1$.

The common eigenvectors are thus

$$\mathbf{e}^1 = (1 \quad 1 \quad 1)^{\mathrm{T}}, \quad \mathbf{e}^2 = (1 \quad -2 \quad 1)^{\mathrm{T}}, \quad \mathbf{e}^3 = (1 \quad 0 \quad -1)^{\mathrm{T}}.$$

We note, as a check, that $\sum_i \mu_i = 2 + (-1) + (-1) = 0 = \mathrm{Tr}\,\mathsf{B}$.

1.25 Given that A is a real symmetric matrix with normalized eigenvectors e^i, obtain the coefficients α_i involved when column matrix x, which is the solution of

$$\mathsf{Ax} - \mu\mathsf{x} = \mathsf{v}, \qquad (*)$$

is expanded as $\mathsf{x} = \sum_i \alpha_i \mathsf{e}^i$. Here μ is a given constant and v is a given column matrix.

(a) Solve $(*)$ when

$$\mathsf{A} = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix},$$

$\mu = 2$ and $\mathsf{v} = (1 \quad 2 \quad 3)^{\mathrm{T}}$.

(b) Would $(*)$ have a solution if (i) $\mu = 1$ and $\mathsf{v} = (1 \quad 2 \quad 3)^{\mathrm{T}}$, (ii) $\mathsf{v} = (2 \quad 2 \quad 3)^{\mathrm{T}}$? Where it does, find it.

Let $\mathsf{x} = \sum_i \alpha_i \mathsf{e}^i$, where $\mathsf{Ae}^i = \lambda_i \mathsf{e}^i$. Then

$$\begin{aligned} \mathsf{Ax} - \mu\mathsf{x} &= \mathsf{v}, \\ \sum_i \mathsf{A}\alpha_i \mathsf{e}^i - \sum_i \mu\alpha_i \mathsf{e}^i &= \mathsf{v}, \\ \sum_i \left(\lambda_i \alpha_i \mathsf{e}^i - \mu\alpha_i \mathsf{e}^i\right) &= \mathsf{v}, \\ \alpha_j &= \frac{(\mathsf{e}^j)^\dagger \mathsf{v}}{\lambda_j - \mu}. \end{aligned}$$

To obtain the last line we have used the mutual orthogonality of the eigenvectors. We note, in passing, that if $\mu = \lambda_j$ for any j there is no solution unless $(\mathsf{e}^j)^\dagger \mathsf{v} = 0$.

(a) To obtain the eigenvalues of the given matrix A, consider

$$0 = |\mathsf{A} - \lambda\mathsf{I}| = (3-\lambda)(4 - 4\lambda + \lambda^2 - 1) = (3-\lambda)(3-\lambda)(1-\lambda).$$

The eigenvalues, and a possible set of corresponding normalized eigenvectors, are therefore,

$$\begin{aligned} &\text{for} \quad \lambda = 3, \quad \mathsf{e}^1 = (0 \quad 0 \quad 1)^{\mathrm{T}}; \\ &\text{for} \quad \lambda = 3, \quad \mathsf{e}^2 = 2^{-1/2}(1 \quad 1 \quad 0)^{\mathrm{T}}; \\ &\text{for} \quad \lambda = 1, \quad \mathsf{e}^3 = 2^{-1/2}(1 \quad -1 \quad 0)^{\mathrm{T}}. \end{aligned}$$

Since $\lambda = 3$ is a degenerate eigenvalue, there are infinitely many acceptable pairs of orthogonal eigenvectors corresponding to it; any pair of vectors of the form (a_i, a_i, b_i) with $2a_1a_2 + b_1b_2 = 0$ will suffice. The pair given is just about the simplest choice possible.

With $\mu = 2$ and $\mathsf{v} = (1 \quad 2 \quad 3)^{\mathrm{T}}$,

$$\alpha_1 = \frac{3}{3-2}, \qquad \alpha_2 = \frac{3/\sqrt{2}}{3-2}, \qquad \alpha_3 = \frac{-1/\sqrt{2}}{1-2}.$$

Thus the solution vector is

$$\mathsf{x} = 3\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \frac{3}{\sqrt{2}}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}.$$

(b) If $\mu = 1$ then it is equal to the third eigenvalue and a solution is only possible if $(\mathsf{e}^3)^\dagger \mathsf{v} = 0$.

For (i) $\mathsf{v} = (1 \quad 2 \quad 3)^{\mathrm{T}}$, $(\mathsf{e}^3)^\dagger \mathsf{v} = -1/\sqrt{2}$ and so no solution is possible.

For (ii) $\mathsf{v} = (2 \quad 2 \quad 3)^{\mathrm{T}}$, $(\mathsf{e}^3)^\dagger \mathsf{v} = 0$, and so a solution is possible. The other scalar products needed are $(\mathsf{e}^1)^\dagger \mathsf{v} = 3$ and $(\mathsf{e}^2)^\dagger \mathsf{v} = 2\sqrt{2}$. For this vector v the solution to the equation is

$$\mathsf{x} = \frac{3}{3-1}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \frac{2\sqrt{2}}{3-1}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \frac{3}{2} \end{pmatrix}.$$

[The solutions to both parts can be checked by resubstitution.]

1.27 By finding the eigenvectors of the Hermitian matrix

$$\mathsf{H} = \begin{pmatrix} 10 & 3i \\ -3i & 2 \end{pmatrix},$$

construct a unitary matrix U such that $\mathsf{U}^\dagger \mathsf{H} \mathsf{U} = \Lambda$, where Λ is a real diagonal matrix.

We start by finding the eigenvalues of H using

$$\begin{vmatrix} 10-\lambda & 3i \\ -3i & 2-\lambda \end{vmatrix} = 0,$$

$$20 - 12\lambda + \lambda^2 - 3 = 0,$$

$$\lambda = 1 \quad \text{or} \quad 11.$$

As expected for an Hermitian matrix, the eigenvalues are real.

For $\lambda = 1$ and normalized eigenvector $(x \quad y)^{\mathrm{T}}$,

$$9x + 3iy = 0 \quad \Rightarrow \quad \mathsf{x}^1 = (10)^{-1/2}(1 \quad 3i)^{\mathrm{T}}.$$

For $\lambda = 11$ and normalized eigenvector $(x \quad y)^{\mathrm{T}}$,

$$-x + 3iy = 0 \quad \Rightarrow \quad \mathsf{x}^2 = (10)^{-1/2}(3i \quad 1)^{\mathrm{T}}.$$

Again as expected, $(\mathsf{x}^1)^\dagger \mathsf{x}^2 = 0$, thus verifying the mutual orthogonality of the eigenvectors. It should be noted that the normalization factor is determined by $(\mathbf{x}^i)^\dagger \mathbf{x}^i = 1$ (and *not* by $(\mathbf{x}^i)^{\mathrm{T}} \mathbf{x}^i = 1$).

We now use these normalized eigenvectors of H as the columns of the matrix U and check that it is unitary:

$$\mathsf{U} = \frac{1}{\sqrt{10}}\begin{pmatrix} 1 & 3i \\ 3i & 1 \end{pmatrix}, \quad \mathsf{U}^\dagger = \frac{1}{\sqrt{10}}\begin{pmatrix} 1 & -3i \\ -3i & 1 \end{pmatrix},$$

$$\mathsf{U}\mathsf{U}^\dagger = \frac{1}{10}\begin{pmatrix} 1 & 3i \\ 3i & 1 \end{pmatrix}\begin{pmatrix} 1 & -3i \\ -3i & 1 \end{pmatrix} = \frac{1}{10}\begin{pmatrix} 10 & 0 \\ 0 & 10 \end{pmatrix} = \mathsf{I}.$$

U has the further property that

$$\begin{aligned}
\mathsf{U}^\dagger\mathsf{H}\mathsf{U} &= \frac{1}{\sqrt{10}}\begin{pmatrix} 1 & -3i \\ -3i & 1 \end{pmatrix}\begin{pmatrix} 10 & 3i \\ -3i & 2 \end{pmatrix}\frac{1}{\sqrt{10}}\begin{pmatrix} 1 & 3i \\ 3i & 1 \end{pmatrix} \\
&= \frac{1}{10}\begin{pmatrix} 1 & -3i \\ -3i & 1 \end{pmatrix}\begin{pmatrix} 1 & 33i \\ 3i & 11 \end{pmatrix} \\
&= \frac{1}{10}\begin{pmatrix} 10 & 0 \\ 0 & 110 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 11 \end{pmatrix} = \Lambda.
\end{aligned}$$

That the diagonal entries of Λ are the eigenvalues of H is in accord with the general theory of normal matrices.

1.29 Given that the matrix

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

has two eigenvectors of the form $(1 \quad y \quad 1)^{\mathrm{T}}$, use the stationary property of the expression $J(\mathsf{x}) = \mathsf{x}^{\mathrm{T}}\mathsf{A}\mathsf{x}/(\mathsf{x}^{\mathrm{T}}\mathsf{x})$ to obtain the corresponding eigenvalues. Deduce the third eigenvalue.

Since A is real and symmetric, each eigenvalue λ is real. Further, from the first component of $\mathsf{A}\mathbf{x} = \lambda\mathbf{x}$, we have that $2 - y = \lambda$, showing that y is also real. Considered as a function of a general vector of the form $(1 \quad y \quad 1)^{\mathrm{T}}$, the quadratic form $\mathsf{x}^{\mathrm{T}}\mathsf{A}\mathsf{x}$ can be written explicitly as

$$\begin{aligned}
\mathsf{x}^{\mathrm{T}}\mathsf{A}\mathsf{x} &= (1\ y\ 1)\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ y \\ 1 \end{pmatrix} \\
&= (1\ y\ 1)\begin{pmatrix} 2-y \\ 2y-2 \\ 2-y \end{pmatrix} \\
&= 2y^2 - 4y + 4.
\end{aligned}$$

The scalar product $\mathsf{x}^{\mathrm{T}}\mathsf{x}$ has the value $2 + y^2$, and so we need to find the stationary values of

$$I = \frac{2y^2 - 4y + 4}{2 + y^2}.$$

These are given by

$$0 = \frac{dI}{dy} = \frac{(2+y^2)(4y-4)-(2y^2-4y+4)2y}{(2+y^2)^2}$$
$$0 = 4y^2 - 8,$$
$$y = \pm\sqrt{2}.$$

The corresponding eigenvalues are the values of I at the stationary points:

$$\text{for} \quad y = \sqrt{2}, \qquad \lambda_1 = \frac{2(2) - 4\sqrt{2} + 4}{2+2} = 2 - \sqrt{2};$$
$$\text{for} \quad y = -\sqrt{2}, \qquad \lambda_2 = \frac{2(2) + 4\sqrt{2} + 4}{2+2} = 2 + \sqrt{2}.$$

The final eigenvalue can be found using the fact that the sum of the eigenvalues is equal to the trace of the matrix; so

$$\lambda_3 = (2+2+2) - (2-\sqrt{2}) - (2+\sqrt{2}) = 2.$$

1.31 The equation of a particular conic section is

$$Q \equiv 8x_1^2 + 8x_2^2 - 6x_1x_2 = 110.$$

Determine the type of conic section this represents, the orientation of its principal axes, and relevant lengths in the directions of these axes.

The eigenvalues of the matrix $\begin{pmatrix} 8 & -3 \\ -3 & 8 \end{pmatrix}$ associated with the quadratic form on the LHS (without any prior scaling) are given by

$$0 = \begin{vmatrix} 8-\lambda & -3 \\ -3 & 8-\lambda \end{vmatrix}$$
$$= \lambda^2 - 16\lambda + 55$$
$$= (\lambda - 5)(\lambda - 11).$$

Referred to the corresponding eigenvectors as axes, the conic section (an ellipse since both eigenvalues are positive) will take the form

$$5y_1^2 + 11y_2^2 = 110 \quad \text{or, in standard form,} \quad \frac{y_1^2}{22} + \frac{y_2^2}{10} = 1.$$

Thus the semi-axes are of lengths $\sqrt{22}$ and $\sqrt{10}$; the former is in the direction of the vector $(x_1 \quad x_2)^{\mathrm{T}}$ given by $(8-5)x_1 - 3x_2 = 0$, i.e. it is the line $x_1 = x_2$. The other principal axis will be the line at right angles to this, namely the line $x_1 = -x_2$.

1.33 Find the direction of the axis of symmetry of the quadratic surface

$$7x^2 + 7y^2 + 7z^2 - 20yz - 20xz + 20xy = 3.$$

The straightforward, but longer, solution to this problem is as follows.

Consider the characteristic polynomial of the matrix associated with the quadratic surface, namely,

$$\begin{aligned} f(\lambda) &= \begin{vmatrix} 7-\lambda & 10 & -10 \\ 10 & 7-\lambda & -10 \\ -10 & -10 & 7-\lambda \end{vmatrix} \\ &= (7-\lambda)(-51-14\lambda+\lambda^2) + 10(30+10\lambda) - 10(-30-10\lambda) \\ &= -\lambda^3 + 21\lambda^2 + 153\lambda + 243. \end{aligned}$$

If the quadratic surface has an axis of symmetry, it must have two equal major axes (perpendicular to it), and hence the characteristic equation must have a repeated root. This same root will therefore also be a root of $df/d\lambda = 0$, i.e. of

$$\begin{aligned} -3\lambda^2 + 42\lambda + 153 &= 0, \\ \lambda^2 - 14\lambda - 51 &= 0, \\ \lambda = 17 \quad \text{or} \quad &-3. \end{aligned}$$

Substitution shows that -3 is a root (and therefore a double root) of $f(\lambda) = 0$, but that 17 is not. The non-repeated root can be calculated as the trace of the matrix minus the repeated roots, i.e. $21 - (-3) - (-3) = 27$. It is the eigenvector that corresponds to this eigenvalue that gives the direction $(x \quad y \quad z)^{\mathrm{T}}$ of the axis of symmetry. Its components must satisfy

$$\begin{aligned} (7-27)x + 10y - 10z &= 0, \\ 10x + (7-27)y - 10z &= 0. \end{aligned}$$

The axis of symmetry is therefore in the direction $(1 \quad 1 \quad -1)^{\mathrm{T}}$.

A more subtle solution is obtained by noting that setting $\lambda = -3$ makes *all three* of the rows (or columns) of the determinant multiples of each other, i.e. it reduces the determinant to rank one. Thus -3 is a repeated root of the characteristic equation and the third root is $21 - 2(-3) = 27$. The rest of the analysis is as above.

We note in passing that, as two eigenvalues are negative and equal, the surface is the hyperboloid of revolution obtained by rotating a (two-branched) hyperbola about its axis of symmetry. Referred to this axis and two others forming a mutually orthogonal set, the equation of the quadratic surface takes the form $-3\chi^2 - 3\eta^2 + 27\zeta^2 = 3$ and so the tips of the two "nose cones" ($\chi = \eta = 0$) are separated by $\frac{2}{3}$ of a unit.

1.35 This problem demonstrates the reverse of the usual procedure of diagonalizing a matrix.

(a) Rearrange the result $\mathsf{A}' = \mathsf{S}^{-1}\mathsf{A}\mathsf{S}$ (which shows how to make a change of basis that diagonalizes A) so as to express the original matrix A in terms of the unitary matrix S and the diagonal matrix A'. Hence show how to construct a matrix A that has given eigenvalues and given (orthogonal) column matrices as its eigenvectors.

(b) Find the matrix that has as eigenvectors $(1 \quad 2 \quad 1)^{\mathrm{T}}$, $(1 \quad -1 \quad 1)^{\mathrm{T}}$ and $(1 \quad 0 \quad -1)^{\mathrm{T}}$ and corresponding eigenvalues λ, μ and ν.

(c) Try a particular case, say $\lambda = 3$, $\mu = -2$ and $\nu = 1$, and verify by explicit solution that the matrix so found does have these eigenvalues.

(a) Since S is unitary, we can multiply the given result on the left by S and on the right by $\mathsf{S}^\dagger$ to obtain

$$\mathsf{SA'S}^\dagger = \mathsf{SS}^{-1}\mathsf{ASS}^\dagger = (\mathsf{I})\,\mathsf{A}\,(\mathsf{I}) = \mathsf{A}.$$

More explicitly, in terms of the eigenvalues and normalized eigenvectors x^i of A,

$$\mathsf{A} = (\mathsf{x}^1 \quad \mathsf{x}^2 \quad \cdots \quad \mathsf{x}^n)\Lambda(\mathsf{x}^1 \quad \mathsf{x}^2 \quad \cdots \quad \mathsf{x}^n)^\dagger.$$

Here Λ is the diagonal matrix that has the eigenvalues of A as its diagonal elements.

Now, given normalized orthogonal column matrices and n specified values, we can use this result to construct a matrix that has the column matrices as eigenvectors and the values as eigenvalues.

(b) The normalized versions of the given column vectors are

$$\frac{1}{\sqrt{6}}(1 \quad 2 \quad 1)^{\mathrm{T}}, \quad \frac{1}{\sqrt{3}}(1 \quad -1 \quad 1)^{\mathrm{T}}, \quad \frac{1}{\sqrt{2}}(1 \quad 0 \quad -1)^{\mathrm{T}},$$

and the orthogonal matrix S can be constructed using these as its columns:

$$\mathsf{S} = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 & \sqrt{2} & \sqrt{3} \\ 2 & -\sqrt{2} & 0 \\ 1 & \sqrt{2} & -\sqrt{3} \end{pmatrix}.$$

The required matrix A can now be formed as $\mathsf{S}\Lambda\mathsf{S}^\dagger$:

$$\begin{aligned}
\mathsf{A} &= \frac{1}{6}\begin{pmatrix} 1 & \sqrt{2} & \sqrt{3} \\ 2 & -\sqrt{2} & 0 \\ 1 & \sqrt{2} & -\sqrt{3} \end{pmatrix}\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \nu \end{pmatrix}\begin{pmatrix} 1 & 2 & 1 \\ \sqrt{2} & -\sqrt{2} & \sqrt{2} \\ \sqrt{3} & 0 & -\sqrt{3} \end{pmatrix} \\
&= \frac{1}{6}\begin{pmatrix} 1 & \sqrt{2} & \sqrt{3} \\ 2 & -\sqrt{2} & 0 \\ 1 & \sqrt{2} & -\sqrt{3} \end{pmatrix}\begin{pmatrix} \lambda & 2\lambda & \lambda \\ \sqrt{2}\mu & -\sqrt{2}\mu & \sqrt{2}\mu \\ \sqrt{3}\nu & 0 & -\sqrt{3}\nu \end{pmatrix} \\
&= \frac{1}{6}\begin{pmatrix} \lambda + 2\mu + 3\nu & 2\lambda - 2\mu & \lambda + 2\mu - 3\nu \\ 2\lambda - 2\mu & 4\lambda + 2\mu & 2\lambda - 2\mu \\ \lambda + 2\mu - 3\nu & 2\lambda - 2\mu & \lambda + 2\mu + 3\nu \end{pmatrix}.
\end{aligned}$$

(c) Setting $\lambda = 3$, $\mu = -2$ and $\nu = 1$, as a particular case, gives A as

$$\mathsf{A} = \frac{1}{6}\begin{pmatrix} 2 & 10 & -4 \\ 10 & 8 & 10 \\ -4 & 10 & 2 \end{pmatrix}.$$

We complete the problem by solving for the eigenvalues of A in the usual way. To avoid working with fractions, and any confusion with the value $\lambda = 3$ used when constructing

A, we will find the eigenvalues of 6A and denote them by η.

$$\begin{aligned}
0 &= |\,6\mathsf{A} - \eta\mathsf{I}\,| \\
&= \begin{vmatrix} 2-\eta & 10 & -4 \\ 10 & 8-\eta & 10 \\ -4 & 10 & 2-\eta \end{vmatrix} \\
&= (2-\eta)(\eta^2 - 10\eta - 84) + 10(10\eta - 60) - 4(132 - 4\eta) \\
&= -\eta^3 + 12\eta^2 + 180\eta - 1296 \\
&= -(\eta - 6)(\eta^2 - 6\eta - 216) \\
&= -(\eta - 6)(\eta + 12)(\eta - 18).
\end{aligned}$$

Thus 6A has eigenvalues 6, -12 and 18; the values for A itself are 1, -2 and 3, as expected.

1.37 A more general form of expression for the determinant of a 3×3 matrix A than (1.45) is given by

$$|\mathsf{A}|\epsilon_{lmn} = A_{li}A_{mj}A_{nk}\epsilon_{ijk}. \tag{1.1}$$

The former could, as stated earlier in this chapter, have been written as

$$|\mathsf{A}| = \epsilon_{ijk}A_{i1}A_{j2}A_{k3}.$$

The more general form removes the explicit mention of 1, 2, 3 at the expense of an additional Levi–Civita symbol; the form of (1.1) can be readily extended to cover a general $N \times N$ matrix.

Use this more general form to prove properties (i), (iii), (v), (vi) and (vii) of determinants stated in Subsection 1.9.1. Property (iv) is obvious by inspection. For definiteness take $N = 3$, but convince yourself that your methods of proof would be valid for any positive integer N.

A full account of the answer to this problem is given in the *Hints and answers* section at the end of the chapter, almost as if it were part of the main text. The reader is referred there for the details.

1.39 Three coupled pendulums swing perpendicularly to the horizontal line containing their points of suspension, and the following equations of motion are satisfied:

$$\begin{aligned}
-m\ddot{x}_1 &= cmx_1 + d(x_1 - x_2), \\
-M\ddot{x}_2 &= cMx_2 + d(x_2 - x_1) + d(x_2 - x_3), \\
-m\ddot{x}_3 &= cmx_3 + d(x_3 - x_2),
\end{aligned}$$

where x_1, x_2 and x_3 are measured from the equilibrium points; m, M and m are the masses of the pendulum bobs; and c and d are positive constants. Find the normal frequencies of the system and sketch the corresponding patterns of oscillation. What happens as $d \to 0$ or $d \to \infty$?

In a normal mode all three coordinates x_i oscillate with the same frequency and with fixed relative phases. When this is represented by solutions of the form $x_i = X_i \cos \omega t$, where the X_i are fixed constants, the equations become, in matrix and vector form,

$$\begin{pmatrix} cm+d-m\omega^2 & -d & 0 \\ -d & cM+2d-M\omega^2 & -d \\ 0 & -d & cm+d-m\omega^2 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \mathbf{0}.$$

For there to be a non-trivial solution to these simultaneous homogeneous equations, we need

$$\begin{aligned} 0 &= \begin{vmatrix} (c-\omega^2)m+d & -d & 0 \\ -d & (c-\omega^2)M+2d & -d \\ 0 & -d & (c-\omega^2)m+d \end{vmatrix} \\ &= \begin{vmatrix} (c-\omega^2)m+d & 0 & -(c-\omega^2)m-d \\ -d & (c-\omega^2)M+2d & -d \\ 0 & -d & (c-\omega^2)m+d \end{vmatrix} \\ &= [(c-\omega^2)m+d]\,\{[(c-\omega^2)M+2d]\,[(c-\omega^2)m+d]-d^2-d^2\} \\ &= (cm-m\omega^2+d)(c-\omega^2)[Mm(c-\omega^2)+2dm+dM]. \end{aligned}$$

Thus, the normal (angular) frequencies are given by

$$\omega^2 = c, \quad \omega^2 = c+\frac{d}{m} \quad \text{and} \quad \omega^2 = c+\frac{2d}{M}+\frac{d}{m}.$$

If the solution column matrix is $\mathsf{X} = (X_1 \quad X_2 \quad X_3)^{\mathrm{T}}$, then
(i) for $\omega^2 = c$, the components of X must satisfy

$$\begin{aligned} dX_1 - dX_2 &= 0, \\ -dX_1 + 2dX_2 - dX_3 &= 0, \quad \Rightarrow \quad \mathsf{X}^1 = (1 \quad 1 \quad 1)^{\mathrm{T}}; \end{aligned}$$

(ii) for $\omega^2 = c + \dfrac{d}{m}$, we have

$$\begin{aligned} -dX_2 &= 0, \\ -dX_1 + \left(-\frac{dM}{m}+2d\right) X_2 - dX_3 &= 0, \quad \Rightarrow \quad \mathsf{X}^2 = (1 \quad 0 \quad -1)^{\mathrm{T}}; \end{aligned}$$

(iii) for $\omega^2 = c + \dfrac{2d}{M} + \dfrac{d}{m}$, the components must satisfy

$$\begin{aligned} \left[\left(-\frac{2d}{M}-\frac{d}{m}\right)m+d\right] X_1 - dX_2 &= 0, \\ -dX_2 + \left[\left(-\frac{2d}{M}-\frac{d}{m}\right)m+d\right] X_3 &= 0, \quad \Rightarrow \quad \mathsf{X}^3 = \left(1 \quad -\frac{2m}{M} \quad 1\right)^{\mathrm{T}}. \end{aligned}$$

The corresponding patterns are shown in Figure 1.1.

If $d \to 0$, the three oscillations decouple and each pendulum swings independently with angular frequency $\sqrt{c}$.

If $d \to \infty$, the three pendulums become rigidly coupled. The second and third modes have (theoretically) infinite frequency and therefore zero amplitude. The only sustainable

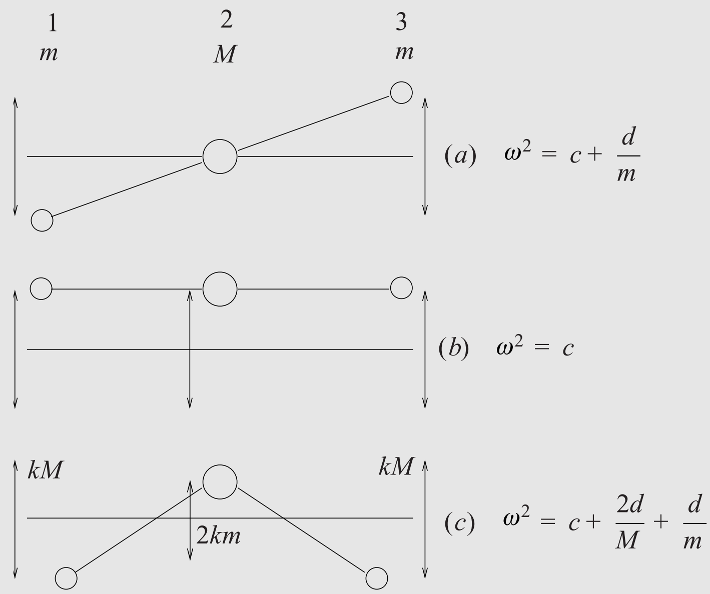

Figure 1.1 The normal modes, as viewed from above, of the coupled pendulums in Problem 1.39.

mode is the one shown as case (b) in the figure; one in which all the pendulums swing as a single entity with angular frequency $\sqrt{c}$.

1.41 Find the normal frequencies of a system consisting of three particles of masses $m_1 = m, m_2 = \mu\, m$, $m_3 = m$ connected in that order in a straight line by two equal light springs of force constant k. Describe the corresponding modes of oscillation.

Now consider the particular case in which $\mu = 2$.

(a) Show that the eigenvectors derived above have the expected orthogonality properties with respect to both the kinetic energy matrix A and the potential energy matrix B.
(b) For the situation in which the masses are released from rest with initial displacements (relative to their equilibrium positions) of $x_1 = 2\epsilon$, $x_2 = -\epsilon$ and $x_3 = 0$, determine their subsequent motions and maximum displacements.

Let the coordinates of the particles, x_1, x_2, x_3, be measured from their equilibrium positions, at which the springs are neither extended nor compressed.

The kinetic energy of the system is simply

$$T = \tfrac{1}{2}m\left(\dot{x}_1^2 + \mu\,\dot{x}_2^2 + \dot{x}_3^2\right),$$

whilst the potential energy stored in the springs takes the form

$$V = \tfrac{1}{2}k\left[(x_2 - x_1)^2 + (x_3 - x_2)^2\right].$$

The kinetic- and potential-energy symmetric matrices are thus

$$\mathsf{A} = \frac{m}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \mathsf{B} = \frac{k}{2}\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}.$$

To find the normal frequencies we have to solve $|\mathsf{B} - \omega^2\mathsf{A}| = 0$. Thus, writing $m\omega^2/k = \lambda$, we have

$$\begin{aligned} 0 &= \begin{vmatrix} 1-\lambda & -1 & 0 \\ -1 & 2-\mu\lambda & -1 \\ 0 & -1 & 1-\lambda \end{vmatrix} \\ &= (1-\lambda)(2-\mu\lambda-2\lambda+\mu\lambda^2-1)+(-1+\lambda) \\ &= (1-\lambda)\lambda(-\mu-2+\mu\lambda), \end{aligned}$$

which leads to $\lambda = 0$, 1 or $1 + 2/\mu$.

The normalized eigenvectors corresponding to the first two eigenvalues can be found by inspection and are

$$\mathsf{x}^1 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad \mathsf{x}^2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

The components of the third eigenvector must satisfy

$$-\frac{2}{\mu}x_1 - x_2 = 0 \quad \text{and} \quad x_2 - \frac{2}{\mu}x_3 = 0.$$

The normalized third eigenvector is therefore

$$\mathsf{x}^3 = \frac{1}{\sqrt{2+(4/\mu^2)}}\begin{pmatrix} 1 & -\dfrac{2}{\mu} & 1 \end{pmatrix}^{\mathrm{T}}.$$

The physical motions associated with these normal modes are as follows.

The first, with $\lambda = \omega = 0$ and all the x_i equal, merely describes bodily translation of the whole system, with no (i.e. zero-frequency) internal oscillations.

In the second solution, the central particle remains stationary, $x_2 = 0$, whilst the other two oscillate with equal amplitudes in antiphase with each other. This motion has frequency $\omega = (k/m)^{1/2}$, the same as that for the oscillations of a single mass m suspended from a single spring of force constant k.

The final and most complicated of the three normal modes has angular frequency $\omega = \{[(\mu+2)/\mu](k/m)\}^{1/2}$, and involves a motion of the central particle which is in antiphase with that of the two outer ones and which has an amplitude $2/\mu$ times as great. In this motion the two springs are compressed and extended in turn. We also note that in the second and third normal modes the center of mass of the system remains stationary.

Now setting $\mu = 2$, we have as the three normal (angular) frequencies 0, Ω and $\sqrt{2}\Omega$, where $\Omega^2 = k/m$. The corresponding (unnormalized) eigenvectors are

$$\mathsf{x}^1 = (1 \quad 1 \quad 1)^{\mathrm{T}}, \quad \mathsf{x}^2 = (1 \quad 0 \quad -1)^{\mathrm{T}}, \quad \mathsf{x}^3 = (1 \quad -1 \quad 1)^{\mathrm{T}}.$$

(a) The matrices A and B have the forms

$$\mathsf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathsf{B} = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}.$$

To verify the standard orthogonality relations we need to show that the quadratic forms $(\mathsf{x}^i)^\dagger \mathsf{A} \mathsf{x}^j$ and $(\mathsf{x}^i)^\dagger \mathsf{B} \mathsf{x}^j$ have zero value for $i \neq j$. Direct evaluation of all the separate cases is as follows:

$$\begin{aligned}
(\mathsf{x}^1)^\dagger \mathsf{A} \mathsf{x}^2 &= 1 + 0 - 1 = 0, \\
(\mathsf{x}^1)^\dagger \mathsf{A} \mathsf{x}^3 &= 1 - 2 + 1 = 0, \\
(\mathsf{x}^2)^\dagger \mathsf{A} \mathsf{x}^3 &= 1 + 0 - 1 = 0, \\
(\mathsf{x}^1)^\dagger \mathsf{B} \mathsf{x}^2 &= (\mathsf{x}^1)^\dagger \,(1 \quad 0 \quad -1)^{\mathrm{T}} = 1 + 0 - 1 = 0, \\
(\mathsf{x}^1)^\dagger \mathsf{B} \mathsf{x}^3 &= (\mathsf{x}^1)^\dagger \,(2 \quad -4 \quad 2)^{\mathrm{T}} = 2 - 4 + 2 = 0, \\
(\mathsf{x}^2)^\dagger \mathsf{B} \mathsf{x}^3 &= (\mathsf{x}^2)^\dagger \,(2 \quad -4 \quad 2)^{\mathrm{T}} = 2 + 0 - 2 = 0.
\end{aligned}$$

If $(\mathsf{x}^i)^\dagger \mathsf{A} \mathsf{x}^j$ has zero value then so does $(\mathsf{x}^j)^\dagger \mathsf{A} \mathsf{x}^i$ (and similarly for B). So there is no need to investigate the other six possibilities and the verification is complete.

(b) In order to determine the behavior of the system we need to know which modes are present in the initial configuration. Each contributory mode will subsequently oscillate with its own frequency. In order to carry out this initial decomposition we write

$$(2\epsilon \quad -\epsilon \quad 0)^{\mathrm{T}} = a\,(1 \quad 1 \quad 1)^{\mathrm{T}} + b\,(1 \quad 0 \quad -1)^{\mathrm{T}} + c\,(1 \quad -1 \quad 1)^{\mathrm{T}},$$

from which it is clear that $a = 0$, $b = \epsilon$ and $c = \epsilon$. As each mode vibrates with its own frequency, the subsequent displacements are given by

$$\begin{aligned}
x_1 &= \epsilon(\cos \Omega t + \cos \sqrt{2}\Omega t), \\
x_2 &= -\epsilon \cos \sqrt{2}\Omega t, \\
x_3 &= \epsilon(-\cos \Omega t + \cos \sqrt{2}\Omega t).
\end{aligned}$$

Since Ω and $\sqrt{2}\Omega$ are not rationally related, at some times the two modes will, for all practical purposes (but not mathematically), be in phase and, at other times, be out of phase. Thus the maximum displacements will be $x_1(\max) = 2\epsilon$, $x_2(\max) = \epsilon$ and $x_3(\max) = 2\epsilon$.

1.43 It is shown in physics and engineering textbooks that circuits containing capacitors and inductors can be analyzed by replacing a capacitor of capacitance C by a "complex impedance" $1/(i\omega C)$ and an inductor of inductance L by an impedance $i\omega L$, where ω is the angular frequency of the currents flowing and $i^2 = -1$.

Use this approach and Kirchhoff's circuit laws to analyze the circuit shown in Figure 1.2 and obtain three linear equations governing the currents I_1, I_2 and I_3. Show that the only possible frequencies of self-sustaining currents satisfy either (a) $\omega^2 LC = 1$ or (b) $3\omega^2 LC = 1$. Find the corresponding current patterns and, in each case, by identifying parts of the circuit in which no current flows, draw an equivalent circuit that contains only one capacitor and one inductor.

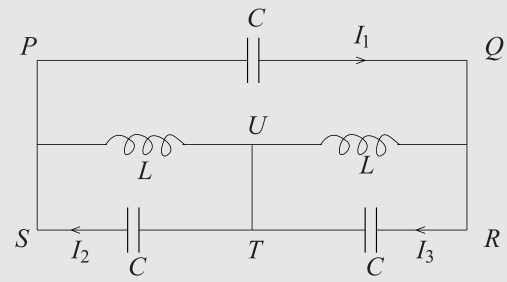

Figure 1.2 The circuit and notation for Problem 1.43.

We apply Kirchhoff's laws to the three closed loops $PQUP$, $SUTS$ and $TURT$ and obtain, respectively,

$$\begin{aligned}
\frac{1}{i\omega C}I_1 + i\omega L(I_1 - I_3) + i\omega L(I_1 - I_2) &= 0,\\
i\omega L(I_2 - I_1) + \frac{1}{i\omega C}I_2 &= 0,\\
i\omega L(I_3 - I_1) + \frac{1}{i\omega C}I_3 &= 0.
\end{aligned}$$

For these simultaneous homogeneous linear equations to be consistent, it is necessary that

$$0 = \begin{vmatrix} \frac{1}{i\omega C} + 2i\omega L & -i\omega L & -i\omega L \\ -i\omega L & \frac{1}{i\omega C} + i\omega L & 0 \\ -i\omega L & 0 & \frac{1}{i\omega C} + i\omega L \end{vmatrix} = \begin{vmatrix} \lambda - 2 & 1 & 1 \\ 1 & \lambda - 1 & 0 \\ 1 & 0 & \lambda - 1 \end{vmatrix},$$

where, after dividing all entries by $-i\omega L$, we have written the combination $(LC\omega^2)^{-1}$ as λ to save space. Expanding the determinant gives

$$\begin{aligned}
0 &= (\lambda - 2)(\lambda - 1)^2 - (\lambda - 1) - (\lambda - 1)\\
&= (\lambda - 1)(\lambda^2 - 3\lambda + 2 - 2)\\
&= \lambda(\lambda - 1)(\lambda - 3).
\end{aligned}$$

Only the non-zero roots are of practical physical interest, and these are $\lambda = 1$ and $\lambda = 3$.

(a) The first of these eigenvalues has an eigenvector $\mathsf{I}^1 = (I_1 \quad I_2 \quad I_3)^{\mathrm{T}}$ that satisfies

$$\begin{aligned}
-I_1 + I_2 + I_3 &= 0,\\
I_1 &= 0 \quad \Rightarrow \quad \mathsf{I}^1 = (0 \quad 1 \quad -1)^{\mathrm{T}}.
\end{aligned}$$

Thus there is no current in PQ and the capacitor in that link can be ignored. Equal currents circulate, in opposite directions, in the other two loops and, although the link TU carries both, there is no transfer between the two loops. Each loop is therefore equivalent to a capacitor of capacitance C in parallel with an inductor of inductance L.

(b) The second eigenvalue has an eigenvector $\mathsf{I}^2 = (I_1 \quad I_2 \quad I_3)^{\mathrm{T}}$ that satisfies

$$\begin{aligned} I_1 + I_2 + I_3 &= 0, \\ I_1 + 2I_2 &= 0 \quad \Rightarrow \quad \mathsf{I}^2 = (-2 \quad 1 \quad 1)^{\mathrm{T}}. \end{aligned}$$

In this mode there is no current in TU and the circuit is equivalent to an inductor of inductance $L + L$ in parallel with a capacitor of capacitance $3C/2$; this latter capacitance is made up of C in parallel with the capacitance equivalent to two capacitors C in series, i.e. in parallel with $\frac{1}{2}C$. Thus, the equivalent single components are an inductance of $2L$ and a capacitance of $3C/2$.

1.45 A double pendulum consists of two identical uniform rods, each of length ℓ and mass M, smoothly jointed together and suspended by attaching the free end of one rod to a fixed point. The system makes small oscillations in a vertical plane, with the angles made with the vertical by the upper and lower rods denoted by θ_1 and θ_2, respectively. The expressions for the kinetic energy T and the potential energy V of the system are (to second order in the θ_i)

$$\begin{aligned} T &\approx Ml^2\left(\tfrac{8}{3}\dot{\theta}_1^2 + 2\dot{\theta}_1\dot{\theta}_2 + \tfrac{2}{3}\dot{\theta}_2^2\right), \\ V &\approx Mgl\left(\tfrac{3}{2}\theta_1^2 + \tfrac{1}{2}\theta_2^2\right). \end{aligned}$$

Determine the normal frequencies of the system and find new variables ξ and η that will reduce these two expressions to diagonal form, i.e. to

$$a_1\dot{\xi}^2 + a_2\dot{\eta}^2 \quad \text{and} \quad b_1\xi^2 + b_2\eta^2.$$

To find the new variables we will use the following result. If the reader is not familiar with it, a standard textbook should be consulted.

If $Q_1 = \mathsf{u}^{\mathrm{T}}\mathsf{A}\mathsf{u}$ and $Q_2 = \mathsf{u}^{\mathrm{T}}\mathsf{B}\mathsf{u}$ are two real symmetric quadratic forms and u^n are those column matrices that satisfy

$$\mathsf{B}\mathsf{u}^n = \lambda_n \mathsf{A}\mathsf{u}^n,$$

then the matrix P whose columns are the vectors u^n is such that the change of variables $\mathsf{u} = \mathsf{P}\mathsf{v}$ reduces both quadratic forms simultaneously to sums of squares, i.e. $Q_1 = \mathsf{v}^{\mathrm{T}}\mathsf{C}\mathsf{v}$ and $Q_2 = \mathsf{v}^{\mathrm{T}}\mathsf{D}\mathsf{v}$, with both C and D diagonal.

Further points to note are:

(i) that for the u^i as determined above, $(\mathsf{u}^m)^{\mathrm{T}}\mathsf{A}\mathsf{u}^n = 0$ if $m \neq n$ and similarly if A is replaced by B;

(ii) that P is not in general an orthogonal matrix, even if the vectors u^n are normalized.

(iii) In the special case that A is the identity matrix I: the above procedure is the same as diagonalizing B; P *is* an orthogonal matrix if normalized vectors are used; mutual orthogonality of the eigenvectors takes on its usual form.

This problem is a physical example to which the above mathematical result can be applied, the two real symmetric (actually positive-definite) matrices being the kinetic and potential energy matrices.

$$\mathsf{A} = \begin{pmatrix} \frac{8}{3} & 1 \\ 1 & \frac{2}{3} \end{pmatrix}, \quad \mathsf{B} = \begin{pmatrix} \frac{3}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \quad \text{with} \quad \lambda_i = \frac{\omega_i^2 l}{g}.$$

We find the normal frequencies by solving

$$
\begin{aligned}
0 &= |\mathsf{B} - \lambda\mathsf{A}| \\
&= \begin{vmatrix} \frac{3}{2} - \frac{8}{3}\lambda & -\lambda \\ -\lambda & \frac{1}{2} - \frac{2}{3}\lambda \end{vmatrix} \\
&= \tfrac{3}{4} - \tfrac{7}{3}\lambda + \tfrac{16}{9}\lambda^2 - \lambda^2 \\
\Rightarrow \quad 0 &= 28\lambda^2 - 84\lambda + 27.
\end{aligned}
$$

Thus, $\lambda = 2.634$ or $\lambda = 0.3661$, and the normal frequencies are $(2.634g/l)^{1/2}$ and $(0.3661g/l)^{1/2}$.

The corresponding column vectors u^i have components that satisfy the following.
(i) For $\lambda = 0.3661$,

$$\left(\tfrac{3}{2} - \tfrac{8}{3}\, 0.3661\right)\theta_1 - 0.3661\theta_2 = 0 \quad \Rightarrow \quad \mathsf{u}^1 = (1 \quad 1.431)^{\mathrm{T}}.$$

(ii) For $\lambda = 2.634$,

$$\left(\tfrac{3}{2} - \tfrac{8}{3}\, 2.634\right)\theta_1 - 2.634\theta_2 = 0 \quad \Rightarrow \quad \mathsf{u}^2 = (1 \quad -2.097)^{\mathrm{T}}.$$

We can now construct P as

$$\mathsf{P} = \begin{pmatrix} 1 & 1 \\ 1.431 & -2.097 \end{pmatrix}$$

and define new variables (ξ, η) by $(\theta_1 \quad \theta_2)^{\mathrm{T}} = \mathsf{P}(\xi \quad \eta)^{\mathrm{T}}$. When the substitutions $\theta_1 = \xi + \eta$ and $\theta_2 = 1.431\xi - 2.097\eta \equiv \alpha\xi - \beta\eta$ are made into the expressions for T and V, they both take on diagonal forms. This can be checked by computing the coefficients of $\xi\eta$ in the two expressions. They are as follows.

$$\text{For } V\text{:} \quad 3 - \alpha\beta = 0, \qquad \text{and} \qquad \text{for } T\text{:} \quad \frac{16}{3} + 2(\alpha - \beta) - \frac{4}{3}\alpha\beta = 0.$$

As an example, the full expression for the potential energy becomes $V = Mg\ell\,(2.524\,\xi^2 + 3.699\,\eta^2)$.

1.47 Three particles each of mass m are attached to a light horizontal string having fixed ends, the string being thus divided into four equal portions, each of length a and under a tension T. Show that for small transverse vibrations the amplitudes x^i of the normal modes satisfy $\mathsf{Bx} = (ma\omega^2/T)\mathsf{x}$, where B is the matrix

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

Estimate the lowest and highest eigenfrequencies using trial vectors $(3 \quad 4 \quad 3)^{\mathrm{T}}$ and $(3 \quad -4 \quad 3)^{\mathrm{T}}$. Use also the exact vectors $\left(1 \quad \sqrt{2} \quad 1\right)^{\mathrm{T}}$ and $\left(1 \quad -\sqrt{2} \quad 1\right)^{\mathrm{T}}$ and compare the results.

For the ith mass, with displacement y_i, the force it experiences as a result of the tension in the string connecting it to the $(i+1)$th mass is the resolved component of that tension perpendicular to the equilibrium line, i.e. $f = \dfrac{y_{i+1} - y_i}{a}T$. Similarly the force due to the tension in the string connecting it to the $(i-1)$th mass is $f = \dfrac{y_{i-1} - y_i}{a}T$. Because the ends of the string are fixed the notional zeroth and fourth masses have $y_0 = y_4 = 0$.

The equations of motion are, therefore,

$$\begin{aligned}
m\ddot{x}_1 &= \frac{T}{a}[(0 - x_1) + (x_2 - x_1)],\\
m\ddot{x}_2 &= \frac{T}{a}[(x_1 - x_2) + (x_3 - x_2)],\\
m\ddot{x}_3 &= \frac{T}{a}[(x_2 - x_3) + (0 - x_3)].
\end{aligned}$$

If the displacements are written as $x_i = X_i \cos\omega t$ and $\mathsf{x} = (X_1 \quad X_2 \quad X_3)^{\mathrm{T}}$, then these equations become

$$\begin{aligned}
-\frac{ma\omega^2}{T}X_1 &= -2X_1 + X_2,\\
-\frac{ma\omega^2}{T}X_2 &= X_1 - 2X_2 + X_3,\\
-\frac{ma\omega^2}{T}X_3 &= X_2 - 2X_3.
\end{aligned}$$

This set of equations can be written as $\mathsf{Bx} = \dfrac{ma\omega^2}{T}\mathsf{x}$, with

$$\mathsf{B} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}.$$

The Rayleigh–Ritz method shows that any estimate λ of $\dfrac{\mathsf{x}^{\mathrm{T}}\mathsf{Bx}}{\mathsf{x}^{\mathrm{T}}\mathsf{x}}$ always lies between the lowest and highest possible values of $ma\omega^2/T$.

Using the suggested *trial* vectors gives the following estimates for λ.

(i) For $\mathsf{x} = (3 \quad 4 \quad 3)^{\mathrm{T}}$

$$\begin{aligned}
\lambda &= [(3,\ 4,\ 3)\mathsf{B}\,(3 \quad 4 \quad 3)^{\mathrm{T}}]/34\\
&= [(3,\ 4,\ 3)\,(2 \quad 2 \quad 2)^{\mathrm{T}}]/34\\
&= 20/34 = 0.588.
\end{aligned}$$

(ii) For $\mathsf{x} = (3 \quad -4 \quad 3)^{\mathrm{T}}$

$$\begin{aligned}
\lambda &= [(3,\ -4,\ 3)\mathsf{B}\,(3 \quad -4 \quad 3)^{\mathrm{T}}]/34\\
&= [(3,\ -4,\ 3)\,(10 \quad -14 \quad 10)^{\mathrm{T}}]/34\\
&= 116/34 = 3.412.
\end{aligned}$$

Using, instead, the *exact* vectors yields the exact values of λ as follows.

(i) For the eigenvector corresponding to the lowest eigenvalue, $\mathsf{x} = (1,\ \sqrt{2},\ 1)^{\mathrm{T}}$,

$$\begin{aligned}\lambda &= \left[(1,\ \sqrt{2},\ 1)\mathsf{B}(1,\ \sqrt{2},\ 1)^{\mathrm{T}}\right]/4 \\ &= \left[(1,\ \sqrt{2},\ 1)(2-\sqrt{2},\ 2\sqrt{2}-2,\ 2-\sqrt{2})^{\mathrm{T}}\right]/4 \\ &= 2-\sqrt{2} = 0.586.\end{aligned}$$

(ii) For the eigenvector corresponding to the highest eigenvalue, $\mathsf{x} = (1,\ -\sqrt{2},\ 1)^{\mathrm{T}}$,

$$\begin{aligned}\lambda &= \left[(1,\ -\sqrt{2},\ 1)\mathsf{B}(1,\ -\sqrt{2},\ 1)^{\mathrm{T}}\right]/4 \\ &= \left[(1,\ -\sqrt{2},\ 1)(2+\sqrt{2},\ -2\sqrt{2}-2,\ 2+\sqrt{2})^{\mathrm{T}}\right]/4 \\ &= 2+\sqrt{2} = 3.414.\end{aligned}$$

As can be seen, the (crude) trial vectors give excellent approximations to the lowest and highest eigenfrequencies.

2 Vector calculus

2.1 Evaluate the integral

$$\int \left[\mathbf{a}(\dot{\mathbf{b}} \cdot \mathbf{a} + \mathbf{b} \cdot \dot{\mathbf{a}}) + \dot{\mathbf{a}}(\mathbf{b} \cdot \mathbf{a}) - 2(\dot{\mathbf{a}} \cdot \mathbf{a})\mathbf{b} - \dot{\mathbf{b}}|\mathbf{a}|^2 \right] dt$$

in which $\dot{\mathbf{a}}$ and $\dot{\mathbf{b}}$ are the derivatives of the real vectors $\mathbf{a}$ and $\mathbf{b}$ with respect to t.

In order to evaluate this integral, we need to group the terms in the integrand so that each is a part of the total derivative of a product of factors. Clearly, the first three terms are the derivative of $\mathbf{a}(\mathbf{b} \cdot \mathbf{a})$, i.e.

$$\frac{d}{dt}[\, \mathbf{a}(\mathbf{b} \cdot \mathbf{a})\,] = \dot{\mathbf{a}}(\mathbf{b} \cdot \mathbf{a}) + \mathbf{a}(\dot{\mathbf{b}} \cdot \mathbf{a}) + \mathbf{a}(\mathbf{b} \cdot \dot{\mathbf{a}}).$$

Remembering that the scalar product is commutative, and that $|\mathbf{a}|^2 = \mathbf{a} \cdot \mathbf{a}$, we also have

$$\begin{aligned} \frac{d}{dt}[\, \mathbf{b}(\mathbf{a} \cdot \mathbf{a})\,] &= \dot{\mathbf{b}}(\mathbf{a} \cdot \mathbf{a}) + \mathbf{b}(\dot{\mathbf{a}} \cdot \mathbf{a}) + \mathbf{b}(\mathbf{a} \cdot \dot{\mathbf{a}}) \\ &= \dot{\mathbf{b}}(\mathbf{a} \cdot \mathbf{a}) + 2\mathbf{b}(\dot{\mathbf{a}} \cdot \mathbf{a}). \end{aligned}$$

Hence,

$$\begin{aligned} I &= \int \left\{ \frac{d}{dt}[\, \mathbf{a}(\mathbf{b} \cdot \mathbf{a})\,] - \frac{d}{dt}[\, \mathbf{b}(\mathbf{a} \cdot \mathbf{a})\,] \right\} dt \\ &= \mathbf{a}(\mathbf{b} \cdot \mathbf{a}) - \mathbf{b}(\mathbf{a} \cdot \mathbf{a}) + \mathbf{h} \\ &= \mathbf{a} \times (\mathbf{a} \times \mathbf{b}) + \mathbf{h}, \end{aligned}$$

where $\mathbf{h}$ is the (vector) constant of integration. To obtain the final line above, we used a special case of the expansion of a vector triple product.

2.3 The general equation of motion of a (non-relativistic) particle of mass m and charge q when it is placed in a region where there is a magnetic field $\mathbf{B}$ and an electric field $\mathbf{E}$ is

$$m\ddot{\mathbf{r}} = q(\mathbf{E} + \dot{\mathbf{r}} \times \mathbf{B});$$

here $\mathbf{r}$ is the position of the particle at time t and $\dot{\mathbf{r}} = d\mathbf{r}/dt$, etc. Write this as three separate equations in terms of the Cartesian components of the vectors involved.

For the simple case of crossed uniform fields $\mathbf{E} = E\mathbf{i}$, $\mathbf{B} = B\mathbf{j}$, in which the particle starts from the origin at $t = 0$ with $\dot{\mathbf{r}} = v_0\mathbf{k}$, find the equations of motion and show the following:

(a) if $v_0 = E/B$ then the particle continues its initial motion;
(b) if $v_0 = 0$ then the particle follows the space curve given in terms of the parameter ξ by

$$x = \frac{mE}{B^2 q}(1 - \cos\xi), \qquad y = 0, \qquad z = \frac{mE}{B^2 q}(\xi - \sin\xi).$$

Interpret this curve geometrically and relate ξ to t. Show that the total distance traveled by the particle after time t is given by

$$\frac{2E}{B}\int_0^t \left|\sin\frac{Bqt'}{2m}\right| dt'.$$

Expressed in Cartesian coordinates, the components of the vector equation read

$$\begin{aligned} m\ddot{x} &= qE_x + q(\dot{y}B_z - \dot{z}B_y), \\ m\ddot{y} &= qE_y + q(\dot{z}B_x - \dot{x}B_z), \\ m\ddot{z} &= qE_z + q(\dot{x}B_y - \dot{y}B_x). \end{aligned}$$

For $E_x = E$, $B_y = B$ and all other field components zero, the equations reduce to

$$m\ddot{x} = qE - qB\dot{z}, \qquad m\ddot{y} = 0, \qquad m\ddot{z} = qB\dot{x}.$$

The second of these, together with the initial conditions $y(0) = \dot{y}(0) = 0$, implies that $y(t) = 0$ for all t. The final equation can be integrated directly to give

$$m\dot{z} = qBx + mv_0, \qquad (*)$$

which can now be substituted into the first to give a differential equation for x:

$$\begin{aligned} m\ddot{x} &= qE - qB\left(\frac{qB}{m}x + v_0\right), \\ \Rightarrow \quad \ddot{x} + \left(\frac{qB}{m}\right)^2 x &= \frac{q}{m}(E - v_0 B). \end{aligned}$$

(i) If $v_0 = E/B$ then the equation for x is that of simple harmonic motion and

$$x(t) = A\cos\omega t + B\sin\omega t,$$

where $\omega = qB/m$. However, in the present case, the initial conditions $x(0) = \dot{x}(0) = 0$ imply that $x(t) = 0$ for all t. Thus, there is no motion in either the x- or the y-direction and, as is then shown by $(*)$, the particle continues with its initial speed v_0 in the z-direction.

(ii) If $v_0 = 0$, the equation of motion is

$$\ddot{x} + \omega^2 x = \frac{qE}{m},$$

which again has sinusoidal solutions but has a non-zero RHS. The full solution consists of the same complementary function as in part (i) together with the simplest possible particular integral, namely $x = qE/m\omega^2$. It is therefore

$$x(t) = A\cos\omega t + B\sin\omega t + \frac{qE}{m\omega^2}.$$

The initial condition $x(0) = 0$ implies that $A = -qE/(m\omega^2)$, whilst $\dot{x}(0) = 0$ requires that $B = 0$. Thus,

$$x = \frac{qE}{m\omega^2}(1 - \cos\omega t),$$
$$\Rightarrow \quad \dot{z} = \frac{qB}{m}x = \omega\frac{qE}{m\omega^2}(1 - \cos\omega t) = \frac{qE}{m\omega}(1 - \cos\omega t).$$

Since $z(0) = 0$, straightforward integration gives

$$z = \frac{qE}{m\omega}\left(t - \frac{\sin\omega t}{\omega}\right) = \frac{qE}{m\omega^2}(\omega t - \sin\omega t).$$

Thus, since $qE/m\omega^2 = mE/B^2q$, the path is of the given parametric form with $\xi = \omega t$. It is a cycloid in the plane $y = 0$; the x-coordinate varies in the restricted range $0 \le x \le 2qE/(m\omega^2)$, whilst the z-coordinate continually increases, though not at a uniform rate.

The element of path length is given by $ds^2 = dx^2 + dy^2 + dz^2$. In this case, writing $qE/(m\omega) = E/B$ as μ,

$$\begin{aligned} ds &= \left[\left(\frac{dx}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2\right]^{1/2} dt \\ &= \left[\mu^2\sin^2\omega t + \mu^2(1 - \cos\omega t)^2\right]^{1/2} dt \\ &= \left[2\mu^2(1 - \cos\omega t)\right]^{1/2} dt = 2\mu|\sin\tfrac{1}{2}\omega t|\,dt. \end{aligned}$$

Thus the total distance traveled after time t is given by

$$s = \int_0^t 2\mu|\sin\tfrac{1}{2}\omega t'|\,dt' = \frac{2E}{B}\int_0^t \left|\sin\frac{qBt'}{2m}\right| dt'.$$

2.5 If two systems of coordinates with a common origin O are rotating with respect to each other, the measured accelerations differ in the two systems. Denoting by $\mathbf{r}$ and $\mathbf{r}'$ position vectors in frames $OXYZ$ and $OX'Y'Z'$, respectively, the connection between the two is

$$\ddot{\mathbf{r}}' = \ddot{\mathbf{r}} + \dot{\boldsymbol{\omega}} \times \mathbf{r} + 2\boldsymbol{\omega} \times \dot{\mathbf{r}} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}),$$

where $\boldsymbol{\omega}$ is the angular velocity vector of the rotation of $OXYZ$ with respect to $OX'Y'Z'$ (taken as fixed). The third term on the RHS is known as the Coriolis acceleration, whilst the final term gives rise to a centrifugal force.

Consider the application of this result to the firing of a shell of mass m from a stationary ship on the steadily rotating earth, working to the first order in ω ($= 7.3 \times 10^{-5}$ rad s^{-1}). If the shell is fired with velocity $\mathbf{v}$ at time $t = 0$ and only reaches a height that is small compared with the radius of the earth, show that its acceleration, as recorded on the ship, is given approximately by

$$\ddot{\mathbf{r}} = \mathbf{g} - 2\boldsymbol{\omega} \times (\mathbf{v} + \mathbf{g}t),$$

where $m\mathbf{g}$ is the weight of the shell measured on the ship's deck.

The shell is fired at another stationary ship (a distance $\mathbf{s}$ away) and $\mathbf{v}$ is such that the shell would have hit its target had there been no Coriolis effect.

(a) Show that without the Coriolis effect the time of flight of the shell would have been $\tau = -2\mathbf{g} \cdot \mathbf{v}/g^2$.
(b) Show further that when the shell actually hits the sea it is off-target by approximately

$$\frac{2\tau}{g^2}[(\mathbf{g} \times \boldsymbol{\omega}) \cdot \mathbf{v}](\mathbf{g}\tau + \mathbf{v}) - (\boldsymbol{\omega} \times \mathbf{v})\tau^2 - \frac{1}{3}(\boldsymbol{\omega} \times \mathbf{g})\tau^3.$$

(c) Estimate the order of magnitude $\boldsymbol{\Delta}$ of this miss for a shell for which the initial speed v is $300\,\mathrm{m\,s^{-1}}$, firing close to its maximum range ($\mathbf{v}$ makes an angle of $\pi/4$ with the vertical) in a northerly direction, whilst the ship is stationed at latitude 45° North.

As the earth is rotating steadily $\dot{\boldsymbol{\omega}} = \mathbf{0}$, and for the mass at rest on the deck,

$$m\ddot{\mathbf{r}}' = m\mathbf{g} + \mathbf{0} + 2\boldsymbol{\omega} \times \dot{\mathbf{0}} + m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}).$$

This, including the centrifugal effect, defines $\mathbf{g}$ which is assumed constant throughout the trajectory.

For the moving mass ($\ddot{\mathbf{r}}'$ is unchanged),

$$m\mathbf{g} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) = m\ddot{\mathbf{r}} + 2m\boldsymbol{\omega} \times \dot{\mathbf{r}} + m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}),$$

i.e.

$$\ddot{\mathbf{r}} = \mathbf{g} - 2\boldsymbol{\omega} \times \dot{\mathbf{r}}.$$

Now, $\omega\dot{r} \ll g$ and so to zeroth order in ω

$$\ddot{\mathbf{r}} = \mathbf{g} \quad \Rightarrow \quad \dot{\mathbf{r}} = \mathbf{g}t + \mathbf{v}.$$

Resubstituting this into the Coriolis term gives, to first order in ω,

$$\ddot{\mathbf{r}} = \mathbf{g} - 2\boldsymbol{\omega} \times (\mathbf{v} + \mathbf{g}t).$$

(a) With no Coriolis force,

$$\dot{\mathbf{r}} = \mathbf{g}t + \mathbf{v} \quad \text{and} \quad \mathbf{r} = \tfrac{1}{2}\mathbf{g}t^2 + \mathbf{v}t.$$

Let $\mathbf{s} = \frac{1}{2}\mathbf{g}\tau^2 + \mathbf{v}\tau$ and use the observation that $\mathbf{s} \cdot \mathbf{g} = 0$, giving

$$\tfrac{1}{2}g^2\tau^2 + \mathbf{v} \cdot \mathbf{g}\tau = 0 \quad \Rightarrow \quad \tau = -\frac{2\mathbf{v} \cdot \mathbf{g}}{g^2}.$$

(b) With Coriolis force,

$$\begin{aligned}
\ddot{\mathbf{r}} &= \mathbf{g} - 2(\boldsymbol{\omega} \times \mathbf{g})t - 2(\boldsymbol{\omega} \times \mathbf{v}),\\
\dot{\mathbf{r}} &= \mathbf{g}t - (\boldsymbol{\omega} \times \mathbf{g})t^2 - 2(\boldsymbol{\omega} \times \mathbf{v})t + \mathbf{v},\\
\mathbf{r} &= \tfrac{1}{2}\mathbf{g}t^2 - \tfrac{1}{3}(\boldsymbol{\omega} \times \mathbf{g})t^3 - (\boldsymbol{\omega} \times \mathbf{v})t^2 + \mathbf{v}t. \qquad (*)
\end{aligned}$$

If the shell hits the sea at time T in the position $\mathbf{r} = \mathbf{s} + \boldsymbol{\Delta}$, then $(\mathbf{s} + \boldsymbol{\Delta}) \cdot \mathbf{g} = 0$, i.e.

$$\begin{aligned}
0 = (\mathbf{s} + \boldsymbol{\Delta}) \cdot \mathbf{g} &= \tfrac{1}{2}g^2T^2 - 0 - (\boldsymbol{\omega} \times \mathbf{v}) \cdot \mathbf{g}T^2 + \mathbf{v} \cdot \mathbf{g}T,\\
\Rightarrow \quad -\mathbf{v} \cdot \mathbf{g} &= T(\tfrac{1}{2}g^2 - (\boldsymbol{\omega} \times \mathbf{v}) \cdot \mathbf{g}),\\
\Rightarrow \quad T &= -\frac{\mathbf{v} \cdot \mathbf{g}}{\frac{1}{2}g^2}\left[1 - \frac{(\boldsymbol{\omega} \times \mathbf{v}) \cdot \mathbf{g}}{\frac{1}{2}g^2}\right]^{-1}\\
&\approx \tau\left(1 + \frac{2(\boldsymbol{\omega} \times \mathbf{v}) \cdot \mathbf{g}}{g^2} + \cdots\right).
\end{aligned}$$

Working to first order in ω, we may put $T = \tau$ in those terms in $(*)$ that involve another factor ω, namely $\boldsymbol{\omega} \times \mathbf{v}$ and $\boldsymbol{\omega} \times \mathbf{g}$. We then find, to this order, that

$$\begin{aligned}\mathbf{s} + \boldsymbol{\Delta} &= \frac{1}{2}\mathbf{g}\left(\tau^2 + \frac{4(\boldsymbol{\omega} \times \mathbf{v}) \cdot \mathbf{g}}{g^2}\tau^2 + \cdots\right) - \frac{1}{3}(\boldsymbol{\omega} \times \mathbf{g})\tau^3 \\ &\quad - (\boldsymbol{\omega} \times \mathbf{v})\tau^2 + \mathbf{v}\tau + 2\frac{(\boldsymbol{\omega} \times \mathbf{v}) \cdot \mathbf{g}}{g^2}\mathbf{v}\tau \\ &= \mathbf{s} + \frac{(\boldsymbol{\omega} \times \mathbf{v}) \cdot \mathbf{g}}{g^2}(2\mathbf{g}\tau^2 + 2\mathbf{v}\tau) - \frac{1}{3}(\boldsymbol{\omega} \times \mathbf{g})\tau^3 - (\boldsymbol{\omega} \times \mathbf{v})\tau^2.\end{aligned}$$

Hence, as stated in the question,

$$\boldsymbol{\Delta} = \frac{2\tau}{g^2}[\,(\mathbf{g} \times \boldsymbol{\omega}) \cdot \mathbf{v}\,](\mathbf{g}\tau + \mathbf{v}) - (\boldsymbol{\omega} \times \mathbf{v})\tau^2 - \frac{1}{3}(\boldsymbol{\omega} \times \mathbf{g})\tau^3.$$

(c) With the ship at latitude 45° and firing the shell at close to 45° to the local horizontal, $\mathbf{v}$ and $\boldsymbol{\omega}$ are almost parallel and the $\boldsymbol{\omega} \times \mathbf{v}$ term can be set to zero. Further, with $\mathbf{v}$ in a northerly direction, $(\mathbf{g} \times \boldsymbol{\omega}) \cdot \mathbf{v} = 0$.

Thus we are left with only the cubic term in τ. In this,

$$\tau = \frac{2 \times 300 \cos(\pi/4)}{9.8} = 43.3 \text{ s},$$

and $\boldsymbol{\omega} \times \mathbf{g}$ is in a westerly direction (recall that $\boldsymbol{\omega}$ is directed northwards and $\mathbf{g}$ is directed downwards, towards the origin) and of magnitude $7\,10^{-5} \times 9.8 \times \sin(\pi/4) = 4.85\,10^{-4}$ m s^{-3}. Thus the miss is by approximately

$$-\tfrac{1}{3} \times 4.85\,10^{-4} \times (43.3)^3 = -13 \text{ m},$$

i.e. some 10–15 m to the East of its intended target.

2.7 Parameterizing the hyperboloid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

by $x = a\cos\theta\sec\phi$, $y = b\sin\theta\sec\phi$, $z = c\tan\phi$, show that an area element on its surface is

$$dS = \sec^2\phi\left[c^2\sec^2\phi\left(b^2\cos^2\theta + a^2\sin^2\theta\right) + a^2b^2\tan^2\phi\right]^{1/2}\,d\theta\,d\phi.$$

Use this formula to show that the area of the curved surface $x^2 + y^2 - z^2 = a^2$ between the planes $z = 0$ and $z = 2a$ is

$$\pi a^2\left(6 + \frac{1}{\sqrt{2}}\sinh^{-1} 2\sqrt{2}\right).$$

With $x = a\cos\theta\sec\phi$, $y = b\sin\theta\sec\phi$ and $z = c\tan\phi$, the tangent vectors to the surface are given in Cartesian coordinates by

$$\begin{aligned}\frac{d\mathbf{r}}{d\theta} &= (-a\sin\theta\sec\phi,\ b\cos\theta\sec\phi,\ 0), \\ \frac{d\mathbf{r}}{d\phi} &= (a\cos\theta\sec\phi\tan\phi,\ b\sin\theta\sec\phi\tan\phi,\ c\sec^2\phi),\end{aligned}$$

and the element of area by

$$\begin{aligned}
dS &= \left| \frac{d\mathbf{r}}{d\theta} \times \frac{d\mathbf{r}}{d\phi} \right| d\theta\, d\phi \\
&= \left| (bc\cos\theta\sec^3\phi,\ ac\sin\theta\sec^3\phi,\ -ab\sec^2\phi\tan\phi) \right| \, d\theta\, d\phi \\
&= \sec^2\phi\left[c^2\sec^2\phi\left(b^2\cos^2\theta + a^2\sin^2\theta\right) + a^2b^2\tan^2\phi\right]^{1/2} d\theta\, d\phi.
\end{aligned}$$

We set $b = c = a$ and note that the plane $z = 2a$ corresponds to $\phi = \tan^{-1} 2$. The ranges of integration are therefore $0 \le \theta < 2\pi$ and $0 \le \phi \le \tan^{-1} 2$, whilst

$$dS = \sec^2\phi(a^4\sec^2\phi + a^4\tan^2\phi)^{1/2}\, d\theta\, d\phi,$$

i.e. it is independent of θ.

To evaluate the integral of dS, we set $\tan\phi = \sinh\psi/\sqrt{2}$, with

$$\sec^2\phi\, d\phi = \frac{1}{\sqrt{2}}\cosh\psi\, d\psi \quad \text{and} \quad \sec^2\phi = 1 + \tfrac{1}{2}\sinh^2\psi.$$

The upper limit for ψ will be given by $\Psi = \sinh^{-1} 2\sqrt{2}$; we note that $\cosh\Psi = 3$. Integrating over θ and making the above substitutions yields

$$\begin{aligned}
S &= 2\pi\int_0^{\Psi} \frac{1}{\sqrt{2}}\cosh\psi\, d\psi\, a^2\left(1 + \frac{1}{2}\sinh^2\psi + \frac{1}{2}\sinh^2\psi\right)^{1/2} \\
&= \sqrt{2}\pi a^2 \int_0^{\Psi} \cosh^2\psi\, d\psi \\
&= \frac{\sqrt{2}\pi a^2}{2}\int_0^{\Psi} (\cosh 2\psi + 1)\, d\psi \\
&= \frac{\sqrt{2}\pi a^2}{2}\left[\frac{\sinh 2\psi}{2} + \psi\right]_0^{\Psi} \\
&= \frac{\pi a^2}{\sqrt{2}}\left[\sinh\psi\cosh\psi + \psi\right]_0^{\Psi} \\
&= \frac{\pi a^2}{\sqrt{2}}\left[\,(2\sqrt{2})(3) + \sinh^{-1} 2\sqrt{2}\,\right] = \pi a^2\left(6 + \frac{1}{\sqrt{2}}\sinh^{-1} 2\sqrt{2}\right).
\end{aligned}$$

2.9 Verify by direct calculation that

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

The proof of this standard result for the divergence of a vector product is most easily carried out in Cartesian coordinates though, of course, the result is valid in any three-dimensional coordinate system.

$$
\begin{aligned}
\text{LHS} &= \nabla \cdot (\mathbf{a} \times \mathbf{b}) \\
&= \frac{\partial}{\partial x}(a_y b_z - a_z b_y) + \frac{\partial}{\partial y}(a_z b_x - a_x b_z) + \frac{\partial}{\partial z}(a_x b_y - a_y b_x) \\
&= a_x \left(-\frac{\partial b_z}{\partial y} + \frac{\partial b_y}{\partial z} \right) + a_y \left(\frac{\partial b_z}{\partial x} - \frac{\partial b_x}{\partial z} \right) + a_z \left(-\frac{\partial b_y}{\partial x} + \frac{\partial b_x}{\partial y} \right) \\
&\quad + b_x \left(\frac{\partial a_z}{\partial y} - \frac{\partial a_y}{\partial z} \right) + b_y \left(-\frac{\partial a_z}{\partial x} + \frac{\partial a_x}{\partial z} \right) + b_z \left(\frac{\partial a_y}{\partial x} - \frac{\partial a_x}{\partial y} \right) \\
&= -\mathbf{a} \cdot (\nabla \times \mathbf{b}) + \mathbf{b} \cdot (\nabla \times \mathbf{a}) = \text{RHS}.
\end{aligned}
$$

2.11 Evaluate the Laplacian of the function

$$\psi(x, y, z) = \frac{zx^2}{x^2 + y^2 + z^2}$$

(a) directly in Cartesian coordinates, and (b) after changing to a spherical polar coordinate system. Verify that, as they must, the two methods give the same result.

(a) In Cartesian coordinates we need to evaluate

$$\nabla^2 \psi = \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2}.$$

The required derivatives are

$$
\begin{aligned}
\frac{\partial \psi}{\partial x} &= \frac{2xz(y^2 + z^2)}{(x^2 + y^2 + z^2)^2}, & \frac{\partial^2 \psi}{\partial x^2} &= \frac{(y^2 + z^2)(2zy^2 + 2z^3 - 6x^2 z)}{(x^2 + y^2 + z^2)^3}, \\
\frac{\partial \psi}{\partial y} &= \frac{-2x^2 yz}{(x^2 + y^2 + z^2)^2}, & \frac{\partial^2 \psi}{\partial y^2} &= -\frac{2zx^2(x^2 + z^2 - 3y^2)}{(x^2 + y^2 + z^2)^3}, \\
\frac{\partial \psi}{\partial z} &= \frac{x^2(x^2 + y^2 - z^2)}{(x^2 + y^2 + z^2)^2}, & \frac{\partial^2 \psi}{\partial z^2} &= -\frac{2zx^2(3x^2 + 3y^2 - z^2)}{(x^2 + y^2 + z^2)^3}.
\end{aligned}
$$

Thus, writing $r^2 = x^2 + y^2 + z^2$,

$$
\begin{aligned}
\nabla^2 \psi &= \frac{2z[\,(y^2 + z^2)(y^2 + z^2 - 3x^2) - 4x^4\,]}{(x^2 + y^2 + z^2)^3} \\
&= \frac{2z[\,(r^2 - x^2)(r^2 - 4x^2) - 4x^4\,]}{r^6} \\
&= \frac{2z(r^2 - 5x^2)}{r^4}.
\end{aligned}
$$

(b) In spherical polar coordinates,

$$\psi(r, \theta, \phi) = \frac{r \cos\theta \, r^2 \sin^2\theta \cos^2\phi}{r^2} = r \cos\theta \sin^2\theta \cos^2\phi.$$

The three contributions to $\nabla^2\psi$ in spherical polars are

$$
\begin{aligned}
(\nabla^2\psi)_r &= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right)\\
&= \frac{2}{r}\cos\theta\sin^2\theta\cos^2\phi,\\
(\nabla^2\psi)_\theta &= \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right)\\
&= \frac{1}{r}\frac{\cos^2\phi}{\sin\theta}\frac{\partial}{\partial\theta}\left[\sin\theta\frac{\partial}{\partial\theta}(\cos\theta\sin^2\theta)\right]\\
&= \frac{\cos^2\phi}{r}(4\cos^3\theta - 8\sin^2\theta\cos\theta),\\
(\nabla^2\psi)_\phi &= \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2}\\
&= \frac{\cos\theta}{r}(-2\cos^2\phi + 2\sin^2\phi).
\end{aligned}
$$

Thus, the full Laplacian in spherical polar coordinates reads

$$
\begin{aligned}
\nabla^2\psi &= \frac{\cos\theta}{r}(2\sin^2\theta\cos^2\phi + 4\cos^2\theta\cos^2\phi\\
&\quad - 8\sin^2\theta\cos^2\phi - 2\cos^2\phi + 2\sin^2\phi)\\
&= \frac{\cos\theta}{r}(4\cos^2\phi - 10\sin^2\theta\cos^2\phi - 2\cos^2\phi + 2\sin^2\phi)\\
&= \frac{\cos\theta}{r}(2 - 10\sin^2\theta\cos^2\phi)\\
&= \frac{2r\cos\theta(r^2 - 5r^2\sin^2\theta\cos^2\phi)}{r^4}.
\end{aligned}
$$

Rewriting this last expression in terms of Cartesian coordinates, one finally obtains

$$\nabla^2\psi = \frac{2z(r^2 - 5x^2)}{r^4},$$

which establishes the equivalence of the two approaches.

2.13 The (Maxwell) relationship between a time-independent magnetic field **B** and the current density **J** (measured in SI units in A m^{-2}) producing it,

$$\nabla \times \mathbf{B} = \mu_0\mathbf{J},$$

can be applied to a long cylinder of conducting ionized gas which, in cylindrical polar coordinates, occupies the region $\rho < a$.

(a) Show that a uniform current density $(0, C, 0)$ and a magnetic field $(0, 0, B)$, with B constant $(= B_0)$ for $\rho > a$ and $B = B(\rho)$ for $\rho < a$, are consistent with this equation. Given that $B(0) = 0$ and that **B** is continuous at $\rho = a$, obtain expressions for C and $B(\rho)$ in terms of B_0 and a.

(b) The magnetic field can be expressed as $\mathbf{B} = \nabla \times \mathbf{A}$, where $\mathbf{A}$ is known as the vector potential. Show that a suitable $\mathbf{A}$ can be found which has only one non-vanishing component, $A_\phi(\rho)$, and obtain explicit expressions for $A_\phi(\rho)$ for both $\rho < a$ and $\rho > a$. Like $\mathbf{B}$, the vector potential is continuous at $\rho = a$.

(c) The gas pressure $p(\rho)$ satisfies the hydrostatic equation $\nabla p = \mathbf{J} \times \mathbf{B}$ and vanishes at the outer wall of the cylinder. Find a general expression for p.

(a) In cylindrical polars with $\mathbf{B} = (0, 0, B(\rho))$, for $\rho \le a$ we have

$$\mu_0(0,\ C,\ 0) = \nabla \times \mathbf{B} = \left(\frac{1}{\rho}\frac{\partial B}{\partial \phi},\ -\frac{\partial B}{\partial \rho},\ 0 \right).$$

As expected, $\partial B/\partial \phi = 0$. The azimuthal component of the equation gives

$$-\frac{\partial B}{\partial \rho} = \mu_0 C \quad \text{for} \quad \rho \le a \quad \Rightarrow \quad B(\rho) = B(0) - \mu_0 C \rho.$$

Since $\mathbf{B}$ has to be differentiable at the origin of ρ and have no ϕ-dependence, $B(0)$ must be zero. This, together with $B = B_0$ for $\rho > a$ requires that $C = -B_0/(a\mu_0)$ and $B(\rho) = B_0\rho/a$ for $0 \le \rho \le a$.

(b) With $\mathbf{B} = \nabla \times \mathbf{A}$, consider $\mathbf{A}$ of the form $\mathbf{A} = (0,\ A(\rho),\ 0)$. Then

$$\begin{aligned}(0,\ 0,\ B(\rho)) &= \frac{1}{\rho}\left(\frac{\partial}{\partial z}(\rho A),\ 0,\ \frac{\partial}{\partial \rho}(\rho A) \right)\\ &= \left(0,\ 0,\ \frac{1}{\rho}\frac{\partial}{\partial \rho}(\rho A) \right).\end{aligned}$$

We now equate the only non-vanishing component on each side of the above equation, treating inside and outside the cylinder separately.

For $0 < \rho \le a$,

$$\begin{aligned}\frac{1}{\rho}\frac{\partial}{\partial \rho}(\rho A) &= \frac{B_0 \rho}{a},\\ \rho A &= \frac{B_0 \rho^3}{3a} + D,\\ A(\rho) &= \frac{B_0 \rho^2}{3a} + \frac{D}{\rho}.\end{aligned}$$

Since $A(0)$ must be finite (so that A is differentiable there), $D = 0$.

For $\rho > a$,

$$\begin{aligned}\frac{1}{\rho}\frac{\partial}{\partial \rho}(\rho A) &= B_0,\\ \rho A &= \frac{B_0 \rho^2}{2} + E,\\ A(\rho) &= \frac{1}{2}B_0 \rho + \frac{E}{\rho}.\end{aligned}$$

At $\rho = a$, the continuity of $\mathbf{A}$ requires

$$\frac{B_0 a^2}{3a} = \frac{1}{2}B_0 a + \frac{E}{a} \Rightarrow E = -\frac{B_0 a^2}{6}.$$

Thus, to summarize,

$$A(\rho) = \frac{B_0 \rho^2}{3a} \quad \text{for} \quad 0 \le \rho \le a,$$

$$\text{and} \qquad A(\rho) = B_0 \left(\frac{\rho}{2} - \frac{a^2}{6\rho} \right) \quad \text{for} \quad \rho \ge a.$$

(c) For the gas pressure $p(\rho)$ in the region $0 < \rho \le a$, we have $\nabla p = \mathbf{J} \times \mathbf{B}$. In component form,

$$\left(\frac{dp}{d\rho},\ 0,\ 0 \right) = \left(0,\ -\frac{B_0}{a\mu_0},\ 0 \right) \times \left(0,\ 0,\ \frac{B_0 \rho}{a} \right),$$

with $p(a) = 0$.

$$\frac{dp}{d\rho} = -\frac{B_0^2 \rho}{\mu_0 a^2} \quad \Rightarrow \quad p(\rho) = \frac{B_0^2}{2\mu_0} \left[1 - \left(\frac{\rho}{a} \right)^2 \right].$$

2.15 Maxwell's equations for electromagnetism in free space (i.e. in the absence of charges, currents and dielectric or magnetic media) can be written

$$\text{(i)} \ \ \nabla \cdot \mathbf{B} = 0, \qquad \text{(ii)} \ \ \nabla \cdot \mathbf{E} = 0,$$
$$\text{(iii)} \ \ \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = \mathbf{0}, \qquad \text{(iv)} \ \ \nabla \times \mathbf{B} - \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t} = \mathbf{0}.$$

A vector $\mathbf{A}$ is defined by $\mathbf{B} = \nabla \times \mathbf{A}$, and a scalar ϕ by $\mathbf{E} = -\nabla\phi - \partial\mathbf{A}/\partial t$. Show that if the condition

$$\text{(v)} \ \ \nabla \cdot \mathbf{A} + \frac{1}{c^2}\frac{\partial \phi}{\partial t} = 0$$

is imposed (this is known as choosing the Lorentz gauge), then $\mathbf{A}$ and ϕ satisfy wave equations as follows.

$$\text{(vi)} \ \ \nabla^2 \phi - \frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} = 0,$$
$$\text{(vii)} \ \ \nabla^2 \mathbf{A} - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} = \mathbf{0}.$$

The reader is invited to proceed as follows.

(a) Verify that the expressions for $\mathbf{B}$ and $\mathbf{E}$ in terms of $\mathbf{A}$ and ϕ are consistent with (i) and (iii).
(b) Substitute for $\mathbf{E}$ in (ii) and use the derivative with respect to time of (v) to eliminate $\mathbf{A}$ from the resulting expression. Hence obtain (vi).
(c) Substitute for $\mathbf{B}$ and $\mathbf{E}$ in (iv) in terms of $\mathbf{A}$ and ϕ. Then use the gradient of (v) to simplify the resulting equation and so obtain (vii).

(a) Substituting for $\mathbf{B}$ in (i),

$$\nabla \cdot \mathbf{B} = \nabla \cdot (\nabla \times \mathbf{A}) = 0, \quad \text{as it is for } \textit{any} \text{ vector } \mathbf{A}.$$

Substituting for $\mathbf{E}$ and $\mathbf{B}$ in (iii),

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = -(\nabla \times \nabla\phi) - \nabla \times \frac{\partial \mathbf{A}}{\partial t} + \frac{\partial}{\partial t}(\nabla \times \mathbf{A}) = \mathbf{0}.$$

Here we have used the facts that $\nabla \times \nabla\phi = \mathbf{0}$ for any scalar, and that, since $\partial/\partial t$ and ∇ act on different variables, the order in which they are applied to $\mathbf{A}$ can be reversed. Thus (i) and (iii) are automatically satisfied if $\mathbf{E}$ and $\mathbf{B}$ are represented in terms of $\mathbf{A}$ and ϕ.

(b) Substituting for $\mathbf{E}$ in (ii) and taking the time derivative of (v),

$$0 = \nabla \cdot \mathbf{E} = -\nabla^2\phi - \frac{\partial}{\partial t}(\nabla \cdot \mathbf{A}),$$

$$0 = \frac{\partial}{\partial t}(\nabla \cdot \mathbf{A}) + \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2}.$$

Adding these equations gives

$$0 = -\nabla^2\phi + \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2}.$$

This is result (vi), the wave equation for ϕ.

(c) Substituting for $\mathbf{B}$ and $\mathbf{E}$ in (iv) and taking the gradient of (v),

$$\nabla \times (\nabla \times \mathbf{A}) - \frac{1}{c^2}\left(-\frac{\partial}{\partial t}\nabla\phi - \frac{\partial^2 \mathbf{A}}{\partial t^2}\right) = \mathbf{0},$$

$$\nabla(\nabla \cdot \mathbf{A}) - \nabla^2\mathbf{A} + \frac{1}{c^2}\frac{\partial}{\partial t}(\nabla\phi) + \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} = \mathbf{0}.$$

From (v), $$\nabla(\nabla \cdot \mathbf{A}) + \frac{1}{c^2}\frac{\partial}{\partial t}(\nabla\phi) = \mathbf{0}.$$

Subtracting these gives $$-\nabla^2\mathbf{A} + \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} = \mathbf{0}.$$

In the second line we have used the vector identity

$$\nabla^2\mathbf{F} = \nabla(\nabla \cdot \mathbf{F}) - \nabla \times (\nabla \times \mathbf{F})$$

to replace $\nabla \times (\nabla \times \mathbf{A})$. The final equation is result (vii).

2.17 Paraboloidal coordinates u, v, ϕ are defined in terms of Cartesian coordinates by

$$x = uv\cos\phi, \qquad y = uv\sin\phi, \qquad z = \tfrac{1}{2}(u^2 - v^2).$$

Identify the coordinate surfaces in the u, v, ϕ system. Verify that each coordinate surface ($u =$ constant, say) intersects every coordinate surface on which one of the other two coordinates (v, say) is constant. Show further that the system of coordinates is an orthogonal one and determine its scale factors. Prove that the u-component of $\nabla \times \mathbf{a}$ is given by

$$\frac{1}{(u^2+v^2)^{1/2}}\left(\frac{a_\phi}{v} + \frac{\partial a_\phi}{\partial v}\right) - \frac{1}{uv}\frac{\partial a_v}{\partial \phi}.$$

To find a surface of constant u we eliminate v from the given relationships:

$$x^2 + y^2 = u^2v^2 \quad \Rightarrow \quad 2z = u^2 - \frac{x^2+y^2}{u^2}.$$

This is an inverted paraboloid of revolution about the z-axis. The range of z is $-\infty < z \leq \frac{1}{2}u^2$.

Similarly, the surface of constant v is given by

$$2z = \frac{x^2 + y^2}{v^2} - v^2.$$

This is also a paraboloid of revolution about the z-axis, but this time it is not inverted. The range of z is $-\frac{1}{2}v^2 \leq z < \infty$.

Since every constant-u paraboloid has some part of its surface in the region $z > 0$ and every constant-v paraboloid has some part of its surface in the region $z < 0$, it follows that every member of the first set intersects each member of the second, and vice versa.

The surfaces of constant ϕ, $y = x\tan\phi$, are clearly (half-) planes containing the z-axis; each cuts the members of the other two sets in parabolic lines.

We now determine (the Cartesian components of) the tangential vectors and test their orthogonality:

$$\begin{aligned}
\mathbf{e}_1 &= \frac{\partial \mathbf{r}}{\partial u} = (v\cos\phi,\ v\sin\phi,\ u),\\
\mathbf{e}_2 &= \frac{\partial \mathbf{r}}{\partial v} = (u\cos\phi,\ u\sin\phi,\ -v),\\
\mathbf{e}_3 &= \frac{\partial \mathbf{r}}{\partial \phi} = (-uv\sin\phi,\ uv\cos\phi,\ 0),\\
\mathbf{e}_1\cdot\mathbf{e}_2 &= uv(\cos\phi\cos\phi + \sin\phi\sin\phi) - uv = 0,\\
\mathbf{e}_2\cdot\mathbf{e}_3 &= u^2v(-\cos\phi\sin\phi + \sin\phi\cos\phi) = 0,\\
\mathbf{e}_1\cdot\mathbf{e}_3 &= uv^2(-\cos\phi\sin\phi + \sin\phi\cos\phi) = 0.
\end{aligned}$$

This shows that all pairs of tangential vectors are orthogonal and therefore that the coordinate system is an orthogonal one. Its scale factors are given by the magnitudes of these tangential vectors:

$$\begin{aligned}
h_u^2 &= |\mathbf{e}_1|^2 = (v\cos\phi)^2 + (v\sin\phi)^2 + u^2 = u^2 + v^2,\\
h_v^2 &= |\mathbf{e}_2|^2 = (u\cos\phi)^2 + (u\sin\phi)^2 + v^2 = u^2 + v^2,\\
h_\phi^2 &= |\mathbf{e}_3|^2 = (uv\sin\phi)^2 + (uv\cos\phi)^2 = u^2v^2.
\end{aligned}$$

Thus

$$h_u = h_v = \sqrt{u^2 + v^2}, \qquad h_\phi = uv.$$

The u-component of $\nabla \times \mathbf{a}$ is given by

$$\begin{aligned}
[\nabla \times \mathbf{a}]_u &= \frac{h_u}{h_u h_v h_\phi}\left[\frac{\partial}{\partial v}(h_\phi a_\phi) - \frac{\partial}{\partial \phi}(h_v a_v)\right]\\
&= \frac{1}{uv\sqrt{u^2 + v^2}}\left[\frac{\partial}{\partial v}(uva_\phi) - \frac{\partial}{\partial \phi}(\sqrt{u^2 + v^2}a_v)\right]\\
&= \frac{1}{\sqrt{u^2 + v^2}}\left(\frac{a_\phi}{v} + \frac{\partial a_\phi}{\partial v}\right) - \frac{1}{uv}\frac{\partial a_v}{\partial \phi},
\end{aligned}$$

as stated in the question.

2.19 Hyperbolic coordinates u, v, ϕ are defined in terms of Cartesian coordinates by

$$x = \cosh u \cos v \cos\phi, \qquad y = \cosh u \cos v \sin\phi, \qquad z = \sinh u \sin v.$$

Sketch the coordinate curves in the $\phi = 0$ plane, showing that far from the origin they become concentric circles and radial lines. In particular, identify the curves $u = 0$, $v = 0$, $v = \pi/2$ and $v = \pi$. Calculate the tangent vectors at a general point, show that they are mutually orthogonal and deduce that the appropriate scale factors are

$$h_u = h_v = (\cosh^2 u - \cos^2 v)^{1/2}, \qquad h_\phi = \cosh u \cos v.$$

Find the most general function $\psi(u)$ of u only that satisfies Laplace's equation $\nabla^2\psi = 0$.

In the plane $\phi = 0$, i.e. $y = 0$, the curves $u =$ constant have x and z connected by

$$\frac{x^2}{\cosh^2 u} + \frac{z^2}{\sinh^2 u} = 1.$$

This general form is that of an ellipse, with foci at $(\pm 1, 0)$. With $u = 0$, it is the line joining the two foci (covered twice). As $u \to \infty$, and $\cosh u \approx \sinh u$ the form becomes that of a circle of very large radius.

The curves $v =$ constant are expressed by

$$\frac{x^2}{\cos^2 v} - \frac{z^2}{\sin^2 v} = 1.$$

These curves are hyperbolae that, for large x and z and fixed v, approximate $z = \pm x \tan v$, i.e. radial lines. The curve $v = 0$ is the part of the x-axis $1 \le x \le \infty$ (covered twice), whilst the curve $v = \pi$ is its reflection in the z-axis. The curve $v = \pi/2$ *is* the z-axis.

In Cartesian coordinates a general point and its derivatives with respect to u, v and ϕ are given by

$$\begin{aligned}
\mathbf{r} &= \cosh u \cos v \cos\phi\,\mathbf{i} + \cosh u \cos v \sin\phi\,\mathbf{j} + \sinh u \sin v\,\mathbf{k},\\
\mathbf{e}_1 = \frac{\partial \mathbf{r}}{\partial u} &= \sinh u \cos v \cos\phi\,\mathbf{i} + \sinh u \cos v \sin\phi\,\mathbf{j} + \cosh u \sin v\,\mathbf{k},\\
\mathbf{e}_2 = \frac{\partial \mathbf{r}}{\partial v} &= -\cosh u \sin v \cos\phi\,\mathbf{i} - \cosh u \sin v \sin\phi\,\mathbf{j} + \sinh u \cos v\,\mathbf{k},\\
\mathbf{e}_3 = \frac{\partial \mathbf{r}}{\partial \phi} &= \cosh u \cos v(-\sin\phi\,\mathbf{i} + \cos\phi\,\mathbf{j}).
\end{aligned}$$

Now consider the scalar products:

$$\begin{aligned}
\mathbf{e}_1 \cdot \mathbf{e}_2 &= \sinh u \cos v \cosh u \sin v(-\cos^2\phi - \sin^2\phi + 1) = 0,\\
\mathbf{e}_1 \cdot \mathbf{e}_3 &= \sinh u \cos^2 v \cosh u(-\sin\phi\cos\phi + \sin\phi\cos\phi) = 0,\\
\mathbf{e}_2 \cdot \mathbf{e}_3 &= \cosh^2 u \sin v \cos v(\sin\phi\cos\phi - \sin\phi\cos\phi) = 0.
\end{aligned}$$

As each is zero, the system is an orthogonal one.

The scale factors are given by $|\mathbf{e}_i|$ and are thus found from:

$$\begin{aligned}
|\mathbf{e}_1|^2 &= \sinh^2 u \cos^2 v(\cos^2\phi + \sin^2\phi) + \cosh^2 u \sin^2 v \\
&= (\cosh^2 u - 1)\cos^2 v + \cosh^2 u(1 - \cos^2 v) \\
&= \cosh^2 u - \cos^2 v; \\
|\mathbf{e}_2|^2 &= \cosh^2 u \sin^2 v(\cos^2\phi + \sin^2\phi) + \sinh^2 u \cos^2 v \\
&= \cosh^2 u(1 - \cos^2 v) + (\cosh^2 u - 1)\cos^2 v \\
&= \cosh^2 u - \cos^2 v; \\
|\mathbf{e}_3|^2 &= \cosh^2 u \cos^2 v(\sin^2\phi + \cos^2\phi) = \cosh^2 u \cos^2 v.
\end{aligned}$$

The immediate deduction is that

$$h_u = h_v = (\cosh^2 u - \cos^2 v)^{1/2}, \qquad h_\phi = \cosh u \cos v.$$

An alternative form for h_u and h_v is $(\sinh^2 u + \sin^2 v)^{1/2}$.

If a solution of Laplace's equation is to be a function, $\psi(u)$, of u only, then all differentiation with respect to v and ϕ can be ignored. The expression for $\nabla^2\psi$ reduces to

$$\begin{aligned}
\nabla^2\psi &= \frac{1}{h_u h_v h_\phi}\left[\frac{\partial}{\partial u}\left(\frac{h_v h_\phi}{h_u}\frac{\partial\psi}{\partial u}\right)\right] \\
&= \frac{1}{\cosh u \cos v(\cosh^2 u - \cos^2 v)}\left[\frac{\partial}{\partial u}\left(\cosh u \cos v \frac{\partial\psi}{\partial u}\right)\right].
\end{aligned}$$

Laplace's equation itself is even simpler and reduces to

$$\frac{\partial}{\partial u}\left(\cosh u \frac{\partial\psi}{\partial u}\right) = 0.$$

This can be rewritten as

$$\frac{\partial\psi}{\partial u} = \frac{k}{\cosh u} = \frac{2k}{e^u + e^{-u}} = \frac{2ke^u}{e^{2u} + 1},$$

$$d\psi = \frac{Ae^u\, du}{1 + (e^u)^2} \quad \Rightarrow \quad \psi = B\tan^{-1} e^u + c.$$

This is the most general function of u only that satisfies Laplace's equation.

3 Line, surface and volume integrals

3.1 The vector field $\mathbf{F}$ is defined by

$$\mathbf{F} = 2xz\mathbf{i} + 2yz^2\mathbf{j} + (x^2 + 2y^2z - 1)\mathbf{k}.$$

Calculate $\nabla \times \mathbf{F}$ and deduce that $\mathbf{F}$ can be written $\mathbf{F} = \nabla\phi$. Determine the form of ϕ.

With $\mathbf{F}$ as given, we calculate the curl of $\mathbf{F}$ to see whether or not it is the zero vector:

$$\nabla \times \mathbf{F} = (4yz - 4yz,\ 2x - 2x,\ 0 - 0) = \mathbf{0}.$$

The fact that it is implies that $\mathbf{F}$ can be written as $\nabla\phi$ for some scalar ϕ.

The form of $\phi(x, y, z)$ is found by integrating, in turn, the components of $\mathbf{F}$ until consistency is achieved, i.e. until a ϕ is found that has partial derivatives equal to the corresponding components of $\mathbf{F}$:

$$\begin{aligned}
2xz = F_x = \frac{\partial\phi}{\partial x} \quad &\Rightarrow \quad \phi(x, y, z) = x^2z + g(y, z),\\
2yz^2 = F_y = \frac{\partial}{\partial y}[\,x^2z + g(y, z)\,] \quad &\Rightarrow \quad g(y, z) = y^2z^2 + h(z),\\
x^2 + 2y^2z - 1 = F_z \quad &= \quad \frac{\partial}{\partial z}[\,x^2z + y^2z^2 + h(z)\,]\\
&\Rightarrow \quad h(z) = -z + k.
\end{aligned}$$

Hence, to within an unimportant constant, the form of ϕ is

$$\phi(x, y, z) = x^2z + y^2z^2 - z.$$

3.3 A vector field $\mathbf{F}$ is given by $\mathbf{F} = xy^2\mathbf{i} + 2\mathbf{j} + x\mathbf{k}$ and L is a path parameterized by $x = ct$, $y = c/t$, $z = d$ for the range $1 \le t \le 2$. Evaluate the three integrals

$$\text{(a)} \int_L \mathbf{F}\,dt, \qquad \text{(b)} \int_L \mathbf{F}\,dy, \qquad \text{(c)} \int_L \mathbf{F}\cdot d\mathbf{r}.$$

Although all three integrals are along the same path L, they are not necessarily of the same type. The vector or scalar nature of the integral is determined by that of the integrand when it is expressed in a form containing the infinitesimal dt.

(a) This is a vector integral and contains three separate integrations. We express each of the integrands in terms of t, according to the parameterization of the integration path

L, before integrating:

$$\begin{aligned}\int_L \mathbf{F}\,dt &= \int_1^2 \left(\frac{c^3}{t}\,\mathbf{i} + 2\,\mathbf{j} + ct\,\mathbf{k}\right) dt \\ &= \left[c^3 \ln t\,\mathbf{i} + 2t\,\mathbf{j} + \frac{1}{2}ct^2\,\mathbf{k}\right]_1^2 \\ &= c^3 \ln 2\,\mathbf{i} + 2\,\mathbf{j} + \frac{3}{2}c\,\mathbf{k}.\end{aligned}$$

(b) This is a similar vector integral but here we must also replace the infinitesimal dy by the infinitesimal $-c\,dt/t^2$ before integrating:

$$\begin{aligned}\int_L \mathbf{F}\,dy &= \int_1^2 \left(\frac{c^3}{t}\,\mathbf{i} + 2\,\mathbf{j} + ct\,\mathbf{k}\right)\left(\frac{-c}{t^2}\right) dt \\ &= \left[\frac{c^4}{2t^2}\,\mathbf{i} + \frac{2c}{t}\,\mathbf{j} - c^2 \ln t\,\mathbf{k}\right]_1^2 \\ &= -\frac{3c^4}{8}\,\mathbf{i} - c\,\mathbf{j} - c^2 \ln 2\,\mathbf{k}.\end{aligned}$$

(c) This is a scalar integral and before integrating we must take the scalar product of $\mathbf{F}$ with $d\mathbf{r} = dx\,\mathbf{i} + dy\,\mathbf{j} + dz\,\mathbf{k}$ to give a single integrand:

$$\begin{aligned}\int_L \mathbf{F}\cdot d\mathbf{r} &= \int_1^2 \left(\frac{c^3}{t}\,\mathbf{i} + 2\,\mathbf{j} + ct\,\mathbf{k}\right)\cdot\left(c\,\mathbf{i} - \frac{c}{t^2}\,\mathbf{j} + 0\,\mathbf{k}\right) dt \\ &= \int_1^2 \left(\frac{c^4}{t} - \frac{2c}{t^2}\right) dt \\ &= \left[c^4 \ln t + \frac{2c}{t}\right]_1^2 \\ &= c^4 \ln 2 - c.\end{aligned}$$

3.5 Determine the point of intersection P, in the first quadrant, of the two ellipses

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad \text{and} \quad \frac{x^2}{b^2} + \frac{y^2}{a^2} = 1.$$

Taking $b < a$, consider the contour L that bounds the area in the first quadrant that is common to the two ellipses. Show that the parts of L that lie along the coordinate axes contribute nothing to the line integral around L of $x\,dy - y\,dx$. Using a parameterization of each ellipse of the general form $x = X\cos\phi$ and $y = Y\sin\phi$, evaluate the two remaining line integrals and hence find the total area common to the two ellipses.

Note: The line integral of $x\,dy - y\,dx$ around a general closed convex contour is equal to twice the area enclosed by that contour.

From the symmetry of the equations under the interchange of x and y, the point P must have $x = y$. Thus,

$$x^2\left(\frac{1}{a^2}+\frac{1}{b^2}\right)=1 \quad \Rightarrow \quad x=\frac{ab}{(a^2+b^2)^{1/2}}.$$

Denoting as curve C_1 the part of

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

that lies on the boundary of the common region, we parameterize it by $x = a\cos\theta_1$ and $y = b\sin\theta_1$. Curve C_1 starts from P and finishes on the y-axis. At P,

$$a\cos\theta_1 = x = \frac{ab}{(a^2+b^2)^{1/2}} \quad \Rightarrow \quad \tan\theta_1 = \frac{a}{b}.$$

It follows that θ_1 lies in the range $\tan^{-1}(a/b) \le \theta_1 \le \pi/2$. Note that θ_1 is *not* the angle between the x-axis and the line joining the origin O to the corresponding point on the curve; for example, when the point is P itself then $\theta_1 = \tan^{-1} a/b$, whilst the line OP makes an angle of $\pi/4$ with the x-axis.

Similarly, referring to that part of

$$\frac{x^2}{b^2}+\frac{y^2}{a^2}=1$$

that lies on the boundary of the common region as curve C_2, we parameterize it by $x = b\cos\theta_2$ and $y = a\sin\theta_2$ with $0 \le \theta_2 \le \tan^{-1}(b/a)$.

On the x-axis, both y and dy are zero and the integrand, $x\,dy - y\,dx$, vanishes. Similarly, the integrand vanishes at all points on the y-axis. Hence,

$$\begin{aligned}
I &= \oint_L (x\,dy - y\,dx)\\
&= \int_{C_2}(x\,dy - y\,dx) + \int_{C_1}(x\,dy - y\,dx)\\
&= \int_0^{\tan^{-1}(b/a)} [\,ab(\cos\theta_2\cos\theta_2) - ab\sin\theta_2(-\sin\theta_2)\,]\,d\theta_2\\
&\quad + \int_{\tan^{-1}(a/b)}^{\pi/2} [\,ab(\cos\theta_1\cos\theta_1) - ab\sin\theta_1(-\sin\theta_1)\,]\,d\theta_1\\
&= ab\tan^{-1}\frac{b}{a} + ab\left(\frac{\pi}{2} - \tan^{-1}\frac{a}{b}\right)\\
&= 2ab\tan^{-1}\frac{b}{a}.
\end{aligned}$$

As noted in the question, the area enclosed by L is equal to $\frac{1}{2}$ of this value, i.e. the total common area in all four quadrants is

$$4 \times \frac{1}{2} \times 2ab\tan^{-1}\frac{b}{a} = 4ab\tan^{-1}\frac{b}{a}.$$

Note that if we let $b \to a$ then the two ellipses become identical circles and we recover the expected value of πa^2 for their common area.

3.7 Evaluate the line integral

$$I = \oint_C \left[y(4x^2 + y^2)\, dx + x(2x^2 + 3y^2)\, dy \right]$$

around the ellipse $x^2/a^2 + y^2/b^2 = 1$.

As it stands this integral is complicated and, in fact, it is the sum of two integrals. The form of the integrand, containing powers of x and y that can be differentiated easily, makes this problem one to which Green's theorem in a plane might usefully be applied. The theorem states that

$$\oint_C (P\, dx + Q\, dy) = \int \int_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx\, dy,$$

where C is a closed contour enclosing the convex region R.

In the notation used above,

$$P(x, y) = y(4x^2 + y^2) \quad \text{and} \quad Q(x, y) = x(2x^2 + 3y^2).$$

It follows that

$$\frac{\partial P}{\partial y} = 4x^2 + 3y^2 \quad \text{and} \quad \frac{\partial Q}{\partial x} = 6x^2 + 3y^2,$$

leading to

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 2x^2.$$

This can now be substituted into Green's theorem and the y-integration carried out immediately as the integrand does not contain y. Hence,

$$\begin{aligned} I &= \int \int_R 2x^2\, dx\, dy \\ &= \int_{-a}^{a} 2x^2\, 2b \left(1 - \frac{x^2}{a^2} \right)^{1/2} dx \\ &= 4b \int_{\pi}^{0} a^2 \cos^2\phi\ \sin\phi\ (-a \sin\phi\, d\phi), \text{ on setting } x = a\cos\phi, \\ &= -ba^3 \int_{\pi}^{0} \sin^2(2\phi)\, d\phi = \tfrac{1}{2}\pi b a^3. \end{aligned}$$

In the final line we have used the standard result for the integral of the square of a sinusoidal function.

3.9 A single-turn coil C of arbitrary shape is placed in a magnetic field $\mathbf{B}$ and carries a current I. Show that the couple acting upon the coil can be written as

$$\mathbf{M} = I \int_C (\mathbf{B} \cdot \mathbf{r})\, d\mathbf{r} - I \int_C \mathbf{B}(\mathbf{r} \cdot d\mathbf{r}).$$

For a planar rectangular coil of sides $2a$ and $2b$ placed with its plane vertical and at an angle ϕ to a uniform horizontal field $\mathbf{B}$, show that $\mathbf{M}$ is, as expected, $4abBI \cos\phi\, \mathbf{k}$.

For an arbitrarily shaped coil the total couple acting can only be found by considering that on an infinitesimal element and then integrating this over the whole coil. The force on an element $d\mathbf{r}$ of the coil is $d\mathbf{F} = I\, d\mathbf{r} \times \mathbf{B}$, and the moment of this force about the origin is $d\mathbf{M} = \mathbf{r} \times \mathbf{F}$. Thus the total moment is given by

$$\begin{aligned}\mathbf{M} &= \oint_C \mathbf{r} \times (I\, d\mathbf{r} \times \mathbf{B}) \\ &= I \oint_C (\mathbf{r} \cdot \mathbf{B})\, d\mathbf{r} - I \oint_C \mathbf{B}(\mathbf{r} \cdot d\mathbf{r}).\end{aligned}$$

To obtain this second form we have used the vector identity

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

To determine the couple acting on the rectangular coil we work in Cartesian coordinates with the z-axis vertical and choose the orientation of axes in the horizontal plane such that the edge of the rectangle of length $2a$ is in the x-direction. Then

$$\mathbf{B} = B\cos\phi\, \mathbf{i} + B\sin\phi\, \mathbf{j}.$$

Considering the first term in $\mathbf{M}$:
(i) for the horizontal sides

$$\mathbf{r} = x\,\mathbf{i} \pm b\,\mathbf{k}, \quad d\mathbf{r} = dx\,\mathbf{i}, \quad \mathbf{r} \cdot \mathbf{B} = xB\cos\phi,$$

$$\int (\mathbf{r} \cdot \mathbf{B})\, d\mathbf{r} = B\cos\phi\, \mathbf{i} \left(\int_{-a}^{a} x\, dx + \int_{a}^{-a} x\, dx \right) = \mathbf{0};$$

(ii) for the vertical sides

$$\mathbf{r} = \pm a\,\mathbf{i} + z\,\mathbf{k}, \quad d\mathbf{r} = dz\,\mathbf{k}, \quad \mathbf{r} \cdot \mathbf{B} = \pm aB\cos\phi,$$

$$\int (\mathbf{r} \cdot \mathbf{B})\, d\mathbf{r} = B\cos\phi\, \mathbf{k} \left(\int_{-b}^{b} (+a)\, dz + \int_{b}^{-b} (-a)\, dz \right) = 4abB\cos\phi\, \mathbf{k}.$$

For the second term in $\mathbf{M}$, since the field is uniform it can be taken outside the integral as a (vector) constant. On the horizontal sides the remaining integral is

$$\int \mathbf{r} \cdot d\mathbf{r} = \pm \int_{-a}^{a} x\, dx = 0.$$

Similarly, the contribution from the vertical sides vanishes and the whole of the second term contributes nothing in this particular configuration.

The total moment is thus $4abB\cos\phi\, \mathbf{k}$, as expected.

3.11 An axially symmetric solid body with its axis AB vertical is immersed in an incompressible fluid of density ρ_0. Use the following method to show that, whatever the shape of the body, for $\rho = \rho(z)$ in cylindrical polars the Archimedean upthrust is, as expected, $\rho_0 g V$, where V is the volume of the body.

Express the vertical component of the resultant force ($-\int p\, d\mathbf{S}$, where p is the pressure) on the body in terms of an integral; note that $p = -\rho_0 g z$ and that for an annular surface element of width dl, $\mathbf{n} \cdot \mathbf{n}_z\, dl = -d\rho$. Integrate by parts and use the fact that $\rho(z_A) = \rho(z_B) = 0$.

We measure z negatively from the water's surface $z = 0$ so that the hydrostatic pressure is $p = -\rho_0 g z$. By symmetry, there is no net horizontal force acting on the body.

The upward force, F, is due to the net vertical component of the hydrostatic pressure acting upon the body's surface:

$$
\begin{aligned}
F &= -\hat{\mathbf{n}}_z \cdot \int p\, d\mathbf{S} \\
&= -\hat{\mathbf{n}}_z \cdot \int (-\rho_0 g z)(2\pi\rho\, \hat{\mathbf{n}}\, dl),
\end{aligned}
$$

where $2\pi\rho\, dl$ is the area of the strip of surface lying between z and $z + dz$ and $\hat{\mathbf{n}}$ is the outward unit normal to that surface.

Now, from geometry, $\hat{\mathbf{n}}_z \cdot \hat{\mathbf{n}}$ is equal to minus the sine of the angle between dl and dz and so $\hat{\mathbf{n}}_z \cdot \hat{\mathbf{n}}\, dl$ is equal to $-d\rho$. Thus,

$$
\begin{aligned}
F &= 2\pi\rho_0 g \int_{z_A}^{z_B} \rho z(-d\rho) \\
&= -2\pi\rho_0 g \int_{z_A}^{z_B} \left(\rho \frac{\partial \rho}{\partial z}\right) z\, dz \\
&= -2\pi\rho_0 g \left\{ \left[z\frac{\rho^2}{2} \right]_{z_A}^{z_B} - \int_{z_A}^{z_B} \frac{\rho^2}{2}\, dz \right\}.
\end{aligned}
$$

But $\rho(z_A) = \rho(z_B) = 0$, and so the first contribution vanishes, leaving

$$
F = \rho_0 g \int_{z_A}^{z_B} \pi\rho^2\, dz = \rho_0 g V,
$$

where V is the volume of the solid. This is the mathematical form of Archimedes' principle. Of course, the result is also valid for a closed body of arbitrary shape, $\rho = \rho(z, \phi)$, but a different method would be needed to prove it.

3.13 A vector field $\mathbf{a}$ is given by $-zxr^{-3}\mathbf{i} - zyr^{-3}\mathbf{j} + (x^2 + y^2)r^{-3}\mathbf{k}$, where $r^2 = x^2 + y^2 + z^2$. Establish that the field is conservative (a) by showing that $\nabla \times \mathbf{a} = \mathbf{0}$, and (b) by constructing its potential function ϕ.

We are told that

$$
\mathbf{a} = -\frac{zx}{r^3}\mathbf{i} - \frac{zy}{r^3}\mathbf{j} + \frac{x^2 + y^2}{r^3}\mathbf{k},
$$

with $r^2 = x^2 + y^2 + z^2$. We will need to differentiate r^{-3} with respect to x, y and z, using the chain rule, and so note that $\partial r/\partial x = x/r$, etc.

(a) Consider $\nabla \times \mathbf{a}$, term by term:

$$
\begin{aligned}
[\nabla \times \mathbf{a}]_x &= \frac{\partial}{\partial y}\left(\frac{x^2+y^2}{r^3}\right) - \frac{\partial}{\partial z}\left(\frac{-zy}{r^3}\right) \\
&= \frac{-3(x^2+y^2)y}{r^4\,r} + \frac{2y}{r^3} + \frac{y}{r^3} - \frac{3(zy)z}{r^4\,r} \\
&= \frac{3y}{r^5}(-x^2 - y^2 + x^2 + y^2 + z^2 - z^2) = 0; \\
[\nabla \times \mathbf{a}]_y &= \frac{\partial}{\partial z}\left(\frac{-zx}{r^3}\right) - \frac{\partial}{\partial x}\left(\frac{x^2+y^2}{r^3}\right) \\
&= \frac{3(zx)z}{r^4\,r} - \frac{x}{r^3} - \frac{2x}{r^3} + \frac{3(x^2+y^2)x}{r^4\,r} \\
&= \frac{3x}{r^5}(z^2 - x^2 - y^2 - z^2 + x^2 + y^2) = 0; \\
[\nabla \times \mathbf{a}]_z &= \frac{\partial}{\partial x}\left(\frac{-zy}{r^3}\right) - \frac{\partial}{\partial y}\left(\frac{-zx}{r^3}\right) \\
&= \frac{3(zy)x}{r^4\,r} - \frac{3(zx)y}{r^4\,r} = 0.
\end{aligned}
$$

Thus all three components of $\nabla \times \mathbf{a}$ are zero, showing that $\mathbf{a}$ is a conservative field.

(b) To construct its potential function we proceed as follows:

$$
\begin{aligned}
\frac{\partial \phi}{\partial x} &= \frac{-zx}{(x^2+y^2+z^2)^{3/2}} \Rightarrow \phi = \frac{z}{(x^2+y^2+z^2)^{1/2}} + f(y,z), \\
\frac{\partial \phi}{\partial y} &= \frac{-zy}{(x^2+y^2+z^2)^{3/2}} = \frac{-zy}{(x^2+y^2+z^2)^{3/2}} + \frac{\partial f}{\partial y} \Rightarrow f(y,z) = g(z), \\
\frac{\partial \phi}{\partial z} &= \frac{x^2+y^2}{(x^2+y^2+z^2)^{3/2}} \\
&= \frac{1}{(x^2+y^2+z^2)^{1/2}} + \frac{-z\,z}{(x^2+y^2+z^2)^{3/2}} + \frac{\partial g}{\partial z} \\
&\Rightarrow g(z) = c.
\end{aligned}
$$

Thus,

$$
\phi(x, y, z) = c + \frac{z}{(x^2+y^2+z^2)^{1/2}} = c + \frac{z}{r}.
$$

The very fact that we can construct a potential function $\phi = \phi(x, y, z)$ whose derivatives are the components of the vector field shows that the field is conservative.

3.15 A force $\mathbf{F}(\mathbf{r})$ acts on a particle at $\mathbf{r}$. In which of the following cases can $\mathbf{F}$ be represented in terms of a potential? Where it can, find the potential.

(a) $\mathbf{F} = F_0 \left[\mathbf{i} - \mathbf{j} - \dfrac{2(x-y)}{a^2}\mathbf{r} \right] \exp\left(-\dfrac{r^2}{a^2} \right)$;

(b) $\mathbf{F} = \dfrac{F_0}{a} \left[z\mathbf{k} + \dfrac{(x^2+y^2-a^2)}{a^2}\mathbf{r} \right] \exp\left(-\dfrac{r^2}{a^2} \right)$;

(c) $\mathbf{F} = F_0 \left[\mathbf{k} + \dfrac{a(\mathbf{r}\times\mathbf{k})}{r^2} \right]$.

(a) We first write the field entirely in terms of the Cartesian unit vectors using $\mathbf{r} = x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}$ and then attempt to construct a suitable potential function ϕ:

$$
\begin{aligned}
\mathbf{F} &= F_0 \left[\mathbf{i} - \mathbf{j} - \frac{2(x-y)}{a^2}\mathbf{r} \right] \exp\left(-\frac{r^2}{a^2} \right) \\
&= \frac{F_0}{a^2} \big[(a^2 - 2x^2 + 2xy)\,\mathbf{i} + (-a^2 - 2xy + 2y^2)\,\mathbf{j} \\
&\qquad + (-2xz + 2yz)\,\mathbf{k} \big] \exp\left(-\frac{r^2}{a^2} \right).
\end{aligned}
$$

Since the partial derivative of $\exp(-r^2/a^2)$ with respect to any Cartesian coordinate u is $\exp(-r^2/a^2)(-2r/a^2)(u/r)$, the z-component of $\mathbf{F}$ appears to be the most straightforward to tackle first:

$$
\begin{aligned}
\frac{\partial \phi}{\partial z} &= \frac{F_0}{a^2}(-2xz + 2yz) \exp\left(-\frac{r^2}{a^2} \right) \\
&\Rightarrow \phi(x, y, z) = F_0 (x - y) \exp\left(-\frac{r^2}{a^2} \right) + f(x, y) \\
&\equiv \phi_1(x, y, z) + f(x, y).
\end{aligned}
$$

Next we examine the derivatives of $\phi = \phi_1 + f$ with respect to x and y to see how closely they generate F_x and F_y:

$$
\begin{aligned}
\frac{\partial \phi_1}{\partial x} &= F_0 \left[\exp\left(-\frac{r^2}{a^2} \right) + (x - y)\exp\left(-\frac{r^2}{a^2} \right)\left(\frac{-2x}{a^2} \right) \right] \\
&= \frac{F_0}{a^2}(a^2 - 2x^2 + 2xy)\exp(-r^2/a^2) = F_x \quad \text{(as given)},
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial \phi_1}{\partial y} &= F_0 \left[-\exp\left(-\frac{r^2}{a^2} \right) + (x - y)\exp\left(-\frac{r^2}{a^2} \right)\left(\frac{-2y}{a^2} \right) \right] \\
&= \frac{F_0}{a^2}(-a^2 - 2xy + 2y^2)\exp(-r^2/a^2) = F_y \quad \text{(as given)}.
\end{aligned}
$$

Thus, to within an arbitrary constant, $\phi_1(x, y, z) = F_0(x - y)\exp\left(-\dfrac{r^2}{a^2} \right)$ is a suitable potential function for the field, without the need for any additional function $f(x, y)$.

(b) We follow the same line of argument as in part (a). First, expressing $\mathbf{F}$ in terms of $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$,

$$\begin{aligned}\mathbf{F} &= \frac{F_0}{a}\left[z\,\mathbf{k} + \frac{x^2+y^2-a^2}{a^2}\mathbf{r}\right]\exp\left(-\frac{r^2}{a^2}\right)\\ &= \frac{F_0}{a^3}\big[x(x^2+y^2-a^2)\,\mathbf{i} + y(x^2+y^2-a^2)\,\mathbf{j}\\ &\qquad + z(x^2+y^2)\,\mathbf{k}\big]\exp\left(-\frac{r^2}{a^2}\right),\end{aligned}$$

and then constructing a possible potential function ϕ. Again starting with the z-component:

$$\begin{aligned}\frac{\partial\phi}{\partial z} &= \frac{F_0 z}{a^3}(x^2+y^2)\exp\left(-\frac{r^2}{a^2}\right),\\ \Rightarrow\quad \phi(x,y,z) &= -\frac{F_0}{2a}(x^2+y^2)\exp\left(-\frac{r^2}{a^2}\right) + f(x,y)\\ &\equiv \phi_1(x,y,z) + f(x,y).\end{aligned}$$

Then, $$\frac{\partial\phi_1}{\partial x} = -\frac{F_0}{2a}\left[2x - \frac{2x(x^2+y^2)}{a^2}\right]\exp\left(-\frac{r^2}{a^2}\right) = F_x \quad \text{(as given)},$$

and $$\frac{\partial\phi_1}{\partial y} = -\frac{F_0}{2a}\left[2y - \frac{2y(x^2+y^2)}{a^2}\right]\exp\left(-\frac{r^2}{a^2}\right) = F_y \quad \text{(as given)}.$$

Thus, $\phi_1(x,y,z) = \dfrac{F_0}{2a}(x^2+y^2)\exp\left(-\dfrac{r^2}{a^2}\right)$, as it stands, is a suitable potential function for $\mathbf{F}(\mathbf{r})$ and establishes the conservative nature of the field.

(c) Again we express F in Cartesian components:

$$\mathbf{F} = F_0\left[\mathbf{k} + \frac{a(\mathbf{r}\times\mathbf{k})}{r^2}\right] = \frac{ay}{r^2}\mathbf{i} - \frac{ax}{r^2}\mathbf{j} + \mathbf{k}.$$

That the z-component of $\mathbf{F}$ has no dependence on y whilst its y-component does depend upon z suggests that the x-component of $\nabla\times\mathbf{F}$ may not be zero. To test this out we compute

$$(\nabla\times\mathbf{F})_x = \frac{\partial(1)}{\partial y} - \frac{\partial}{\partial z}\left(\frac{-ax}{r^2}\right) = 0 - \frac{2axz}{r^4} \neq 0,$$

and find that it is not. To have even one component of $\nabla\times\mathbf{F}$ non-zero is sufficient to show that $\mathbf{F}$ is not conservative and that no potential function can be found. There is no point in searching further!

The same conclusion can be reached by considering the implication of $\mathbf{F}_z = \mathbf{k}$, namely that any possible potential function has to have the form $\phi(x,y,z) = z + f(x,y)$. However, $\partial\phi/\partial x$ is known to be $-ay/r^2 = -ay/(x^2+y^2+z^2)$. This yields a contradiction, as it requires $\partial f(x,y)/\partial x$ to depend on z, which is clearly impossible.

3.17 The vector field $\mathbf{f}$ has components $y\mathbf{i} - x\mathbf{j} + \mathbf{k}$ and γ is a curve given parametrically by

$$\mathbf{r} = (a - c + c\cos\theta)\mathbf{i} + (b + c\sin\theta)\mathbf{j} + c^2\theta\mathbf{k}, \quad 0 \le \theta \le 2\pi.$$

Describe the shape of the path γ and show that the line integral $\int_\gamma \mathbf{f} \cdot d\mathbf{r}$ vanishes. Does this result imply that $\mathbf{f}$ is a conservative field?

As θ increases from 0 to 2π, the x- and y-components of $\mathbf{r}$ vary sinusoidally and in quadrature about fixed values $a - c$ and b. Both variations have amplitude c and both return to their initial values when $\theta = 2\pi$. However, the z-component increases monotonically from 0 to a value of $2\pi c^2$. The curve γ is therefore one loop of a circular spiral of radius c and pitch $2\pi c^2$. Its axis is parallel to the z-axis and passes through the points $(a - c,\ b,\ z)$.

The line element $d\mathbf{r}$ has components $(-c\sin\theta\, d\theta, c\cos\theta\, d\theta, c^2\, d\theta)$ and so the line integral of f along γ is given by

$$\begin{aligned}
\int_\gamma \mathbf{f} \cdot d\mathbf{r} &= \int_0^{2\pi} \left[y(-c\sin\theta) - x(c\cos\theta) + c^2\right] d\theta \\
&= \int_0^{2\pi} \left[-c(b + c\sin\theta)\sin\theta - c(a - c + c\cos\theta)\cos\theta + c^2\right] d\theta \\
&= \int_0^{2\pi} \left(-bc\sin\theta - c^2\sin^2\theta - c(a - c)\cos\theta - c^2\cos^2\theta + c^2\right) d\theta \\
&= 0 - \pi c^2 - 0 - \pi c^2 + 2\pi c^2 = 0.
\end{aligned}$$

However, this does not imply that $\mathbf{f}$ is a conservative field since (i) γ is not a closed loop, and (ii) even if it were, the line integral has to vanish for *every* loop, not just for a particular one.

Further,

$$\nabla \times \mathbf{f} = (0 - 0,\ 0 - 0,\ -1 - 1) = (0,\ 0,\ -2) \ne \mathbf{0},$$

showing explicitly that $\mathbf{f}$ is not conservative.

3.19 Evaluate the surface integral $\int \mathbf{r} \cdot d\mathbf{S}$, where $\mathbf{r}$ is the position vector, over that part of the surface $z = a^2 - x^2 - y^2$ for which $z \ge 0$, by each of the following methods.

(a) Parameterize the surface as $x = a\sin\theta\, \cos\phi$, $y = a\sin\theta\, \sin\phi$, $z = a^2\, \cos^2\theta$, and show that

$$\mathbf{r} \cdot d\mathbf{S} = a^4(2\sin^3\theta\cos\theta + \cos^3\theta\sin\theta)\, d\theta\, d\phi.$$

(b) Apply the divergence theorem to the volume bounded by the surface and the plane $z = 0$.

(a) With $x = a\sin\theta\, \cos\phi$, $y = a\sin\theta\, \sin\phi$, $z = a^2\, \cos^2\theta$, we first check that this does parameterize the surface appropriately:

$$a^2 - x^2 - y^2 = a^2 - a^2\sin^2\theta(\cos^2\phi + \sin^2\phi) = a^2(1 - \sin^2\theta) = a^2\cos^2\theta = z.$$

We see that it does so for the relevant part of the surface, i.e. that which lies above the plane $z = 0$ with $0 \le \theta \le \pi/2$. It would not do so for the part with $z < 0$ for which $x^2 + y^2$ has to be greater than a^2; this is not catered for by the given parameterization.

Having carried out this check, we calculate expressions for $d\mathbf{S}$ and hence $\mathbf{r}\cdot d\mathbf{S}$ in terms of θ and ϕ as follows:

$$\mathbf{r} = a\sin\theta\cos\phi\,\mathbf{i} + a\sin\theta\sin\phi\,\mathbf{j} + a^2\cos^2\theta\,\mathbf{k},$$

and the tangent vectors at the point (θ,ϕ) on the surface are given by

$$\frac{\partial\mathbf{r}}{\partial\theta} = a\cos\theta\cos\phi\,\mathbf{i} + a\cos\theta\sin\phi\,\mathbf{j} - 2a^2\cos\theta\sin\theta\,\mathbf{k},$$
$$\frac{\partial\mathbf{r}}{\partial\phi} = -a\sin\theta\sin\phi\,\mathbf{i} + a\sin\theta\cos\phi\,\mathbf{j}.$$

The corresponding vector element of surface area is thus

$$\begin{aligned} d\mathbf{S} &= \frac{\partial\mathbf{r}}{\partial\theta}\times\frac{\partial\mathbf{r}}{\partial\phi} \\ &= 2a^3\cos\theta\sin^2\theta\cos\phi\,\mathbf{i} + 2a^3\cos\theta\sin^2\theta\sin\phi\,\mathbf{j} + a^2\cos\theta\sin\theta\,\mathbf{k}, \end{aligned}$$

giving $\mathbf{r}\cdot d\mathbf{S}$ as

$$\begin{aligned} \mathbf{r}\cdot d\mathbf{S} &= 2a^4\cos\theta\sin^3\theta\cos^2\phi + 2a^4\cos\theta\sin^3\theta\sin^2\phi + a^4\cos^3\theta\sin\theta \\ &= 2a^4\cos\theta\sin^3\theta + a^4\cos^3\theta\sin\theta. \end{aligned}$$

This is to be integrated over the ranges $0 \le \phi < 2\pi$ and $0 \le \theta \le \pi/2$ as follows:

$$\begin{aligned} \int \mathbf{r}\cdot d\mathbf{S} &= a^4\int_0^{2\pi} d\phi \int_0^{\pi/2} (2\sin^3\theta\cos\theta + \cos^3\theta\sin\theta)\,d\theta \\ &= 2\pi a^4\left(2\left[\frac{\sin^4\theta}{4}\right]_0^{\pi/2} + \left[\frac{-\cos^4\theta}{4}\right]_0^{\pi/2}\right) \\ &= 2\pi a^4\left(\frac{2}{4}+\frac{1}{4}\right) = \frac{3\pi a^4}{2}. \end{aligned}$$

(b) The divergence of the vector field $\mathbf{r}$ is 3, a constant, and so the surface integral $\int \mathbf{r}\cdot d\mathbf{S}$ taken over the complete surface Σ (including the part that lies in the plane $z = 0$) is, by the divergence theorem, equal to three times the volume V of the region bounded by Σ. Now,

$$V = \int_0^{a^2} \pi\rho^2\,dz = \int_0^{a^2} \pi(a^2 - z)\,dz = \pi(a^4 - \tfrac{1}{2}a^4) = \tfrac{1}{2}\pi a^4,$$

and so $\int_\Sigma \mathbf{r}\cdot d\mathbf{S} = 3\pi a^4/2$.

However, on the part of the surface lying in the plane $z = 0$, $\mathbf{r} = x\,\mathbf{i} + y\,\mathbf{j} + 0\,\mathbf{k}$, whilst $d\mathbf{S} = -dS\,\mathbf{k}$. Consequently the scalar product $\mathbf{r}\cdot d\mathbf{S} = 0$; in words, for any point on this face its position vector is orthogonal to the normal to the face. The surface integral over this face therefore contributes nothing to the total integral and the value obtained is that due to the curved surface alone, in agreement with the result in (a).

3.21 Use the result

$$\int_V \nabla\phi \, dV = \oint_S \phi \, d\mathbf{S},$$

together with an appropriately chosen scalar function ϕ, to prove that the position vector $\bar{\mathbf{r}}$ of the center of mass of an arbitrarily shaped body of volume V and uniform density can be written

$$\bar{\mathbf{r}} = \frac{1}{V} \oint_S \tfrac{1}{2} r^2 \, d\mathbf{S}.$$

The position vector of the center of mass is defined by

$$\bar{\mathbf{r}} \int_V \rho \, dV = \int_V \mathbf{r}\rho \, dV.$$

In order to make use of the given equation, we need to find a scalar function f that is such that $\nabla f = \mathbf{r}$; when this is substituted into the RHS of the above equation, the expression for $\bar{\mathbf{r}}$ can be transformed into a surface integral, rather than a volume integral.

A suitable function for this purpose is $f(r) = \frac{1}{2}r^2$. Writing $\mathbf{r}$ in this form and canceling the constant ρ, we have, using the given general result, that

$$\bar{\mathbf{r}}\, V = \int_V \nabla\left(\frac{1}{2}r^2\right) dV = \oint_S \tfrac{1}{2} r^2 \, d\mathbf{S}.$$

From this it follows immediately that

$$\bar{\mathbf{r}} = \frac{1}{V} \oint_S \tfrac{1}{2} r^2 \, d\mathbf{S}.$$

This result provides an alternative method of finding the center of mass $\bar{z}\mathbf{k}$ of the uniform hemisphere $r = a$, $0 \leq \theta \leq \pi/2$, $0 \leq \phi < 2\pi$. The curved surface contributes $3a/4$ to $\bar{z}$ and the plane surface contributes $-3a/8$, giving $\bar{z} = 3a/8$.

3.23 Demonstrate the validity of the divergence theorem:

(a) by calculating the flux of the vector

$$\mathbf{F} = \frac{\alpha \mathbf{r}}{(r^2 + a^2)^{3/2}}$$

through the spherical surface $|\mathbf{r}| = \sqrt{3}a$;

(b) by showing that

$$\nabla \cdot \mathbf{F} = \frac{3\alpha a^2}{(r^2 + a^2)^{5/2}}$$

and evaluating the volume integral of $\nabla \cdot \mathbf{F}$ over the interior of the sphere $|\mathbf{r}| = \sqrt{3}a$. The substitution $r = a \tan\theta$ will prove useful in carrying out the integration.

(a) The field is radial with

$$\mathbf{F} = \frac{\alpha \, \mathbf{r}}{(r^2 + a^2)^{3/2}} = \frac{\alpha \, r}{(r^2 + a^2)^{3/2}} \, \hat{\mathbf{e}}_r.$$

The total flux is therefore given by

$$\Phi = \left.\frac{4\pi r^2\,\alpha\,r}{(r^2+a^2)^{3/2}}\right|_{r=a\sqrt{3}} = \frac{4\pi a^3\,\alpha\,3\sqrt{3}}{8a^3} = \frac{3\sqrt{3}\pi\alpha}{2}.$$

(b) From the divergence theorem, the total flux over the surface of the sphere is equal to the volume integral of its divergence within the sphere. The divergence is given by

$$\begin{aligned}\nabla\cdot\mathbf{F} &= \frac{1}{r^2}\frac{\partial}{\partial r}(r^2F_r) = \frac{1}{r^2}\frac{\partial}{\partial r}\left(\frac{r^2\,\alpha\,r}{(r^2+a^2)^{3/2}}\right)\\ &= \frac{1}{r^2}\left[\frac{3\alpha r^2}{(r^2+a^2)^{3/2}} - \frac{3\alpha r^4}{(r^2+a^2)^{5/2}}\right]\\ &= \frac{3\alpha a^2}{(r^2+a^2)^{5/2}},\end{aligned}$$

and on integrating over the sphere, we have

$$\begin{aligned}\int_V \nabla\cdot\mathbf{F}\,dV &= \int_0^{\sqrt{3}a} \frac{3\alpha a^2}{(r^2+a^2)^{5/2}}4\pi r^2\,dr,\ \text{ set } r = a\tan\theta,\ 0\le\theta\le\tfrac{\pi}{3},\\ &= 12\pi\alpha a^2\int_0^{\pi/3}\frac{a^2\tan^2\theta\, a\sec^2\theta}{a^5\sec^5\theta}\,d\theta\\ &= 12\pi\alpha\int_0^{\pi/3}\sin^2\theta\cos\theta\,d\theta\\ &= 12\pi\alpha\left[\frac{\sin^3\theta}{3}\right]_0^{\pi/3} = 12\pi\alpha\frac{\sqrt{3}}{8} = \frac{3\sqrt{3}\pi\alpha}{2}, \quad \text{as in (a).}\end{aligned}$$

The equality of the results in parts (a) and (b) is in accordance with the divergence theorem.

3.25 In a uniform conducting medium with unit relative permittivity, charge density ρ, current density $\mathbf{J}$, electric field $\mathbf{E}$ and magnetic field $\mathbf{B}$, Maxwell's electromagnetic equations take the form (with $\mu_0\epsilon_0 = c^{-2}$)

$$\text{(i) } \nabla\cdot\mathbf{B} = 0, \qquad \text{(ii) } \nabla\cdot\mathbf{E} = \rho/\epsilon_0,$$

$$\text{(iii) } \nabla\times\mathbf{E} + \dot{\mathbf{B}} = \mathbf{0}, \quad \text{(iv) } \nabla\times\mathbf{B} - (\dot{\mathbf{E}}/c^2) = \mu_0\mathbf{J}.$$

The density of stored energy in the medium is given by $\frac{1}{2}(\epsilon_0E^2 + \mu_0^{-1}B^2)$. Show that the rate of change of the total stored energy in a volume V is equal to

$$-\int_V \mathbf{J}\cdot\mathbf{E}\,dV - \frac{1}{\mu_0}\oint_S(\mathbf{E}\times\mathbf{B})\cdot d\mathbf{S},$$

where S is the surface bounding V.

[The first integral gives the ohmic heating loss, whilst the second gives the electromagnetic energy flux out of the bounding surface. The vector $\mu_0^{-1}(\mathbf{E}\times\mathbf{B})$ is known as the Poynting vector.]

The total stored energy is equal to the volume integral of the energy density. Let R be its rate of change. Then, differentiating under the integral sign, we have

$$R = \frac{d}{dt}\int_V \left(\frac{\epsilon_0}{2}E^2 + \frac{1}{2\mu_0}B^2\right) dV$$
$$= \int_V \left(\epsilon_0 \mathbf{E}\cdot\dot{\mathbf{E}} + \frac{1}{\mu_0}\mathbf{B}\cdot\dot{\mathbf{B}}\right) dV.$$

Now using (iv) and (iii), we have

$$R = \int_V \left[\epsilon_0 \mathbf{E}\cdot(-\mu_0 c^2 \mathbf{J} + c^2 \nabla\times\mathbf{B}) - \frac{1}{\mu_0}\mathbf{B}\cdot(\nabla\times\mathbf{E})\right] dV$$
$$= -\int_V \mathbf{E}\cdot\mathbf{J}\, dV + \int_V \left[\epsilon_0 c^2\, \mathbf{E}\cdot(\nabla\times\mathbf{B}) - \frac{1}{\mu_0}\mathbf{B}\cdot(\nabla\times\mathbf{E})\right] dV$$
$$= -\int_V \mathbf{E}\cdot\mathbf{J}\, dV - \frac{1}{\mu_0}\int_V \nabla\cdot(\mathbf{E}\times\mathbf{B})\, dV$$
$$= -\int_V \mathbf{E}\cdot\mathbf{J}\, dV - \frac{1}{\mu_0}\oint_S (\mathbf{E}\times\mathbf{B})\cdot d\mathbf{S}, \quad \text{by the divergence theorem.}$$

To obtain the penultimate line we used the vector identity

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

3.27 The vector field $\mathbf{F}$ is given by

$$\mathbf{F} = (3x^2yz + y^3z + xe^{-x})\mathbf{i} + (3xy^2z + x^3z + ye^x)\mathbf{j} + (x^3y + y^3x + xy^2z^2)\mathbf{k}.$$

Calculate (a) directly, and (b) by using Stokes' theorem the value of the line integral $\int_L \mathbf{F}\cdot d\mathbf{r}$, where L is the (three-dimensional) closed contour $OABCDEO$ defined by the successive vertices $(0, 0, 0)$, $(1, 0, 0)$, $(1, 0, 1)$, $(1, 1, 1)$, $(1, 1, 0)$, $(0, 1, 0)$, $(0, 0, 0)$.

(a) This calculation is a piece-wise evaluation of the line integral, made up of a series of scalar products of the length of a straight piece of the contour and the component of $\mathbf{F}$ parallel to it (integrated if that component varies along the particular straight section).

On OA, $y = z = 0$ and $F_x = xe^{-x}$;

$$I_1 = \int_0^1 xe^{-x}\, dx = \left[-xe^{-x}\right]_0^1 + \int_0^1 e^{-x}\, dx = 1 - 2e^{-1}.$$

On AB, $x = 1$ and $y = 0$ and $F_z = 0$; the integral I_2 is zero.
On BC, $x = 1$ and $z = 1$ and $F_y = 3y^2 + 1 + ey$;

$$I_3 = \int_0^1 (3y^2 + 1 + ey)\, dy = 1 + 1 + \tfrac{1}{2}e.$$

On CD, $x = 1$ and $y = 1$ and $F_z = 1 + 1 + z^2$;

$$I_4 = \int_1^0 (1 + 1 + z^2)\, dz = -1 - 1 - \tfrac{1}{3}.$$

On DE, $y = 1$ and $z = 0$ and $F_x = xe^{-x}$;

$$I_5 = \int_1^0 xe^{-x}\,dx = -1 + 2e^{-1}.$$

On EO, $x = z = 0$ and $F_y = ye^0$;

$$I_6 = \int_1^0 ye^0\,dy = -\tfrac{1}{2}.$$

Adding up these six contributions shows that the complete line integral has the value $\dfrac{e}{2} - \dfrac{5}{6}$.

(b) As a simple sketch shows, the given contour is three-dimensional. However, it is equivalent to two plane square contours, one $OADEO$ (denoted by S_1) lying in the plane $z = 0$ and the other $ABCDA$ (S_2) lying in the plane $x = 1$; the latter is traversed in the negative sense. The common segment AD does not form part of the original contour but, as it is traversed in opposite senses in the two constituent contours, it (correctly) contributes nothing to the line integral.

To use Stokes' theorem we first need to calculate

$$\begin{aligned}
(\nabla \times \mathbf{F})_x &= x^3 + 3y^2x + 2yxz^2 - 3xy^2 - x^3 = 2yxz^2,\\
(\nabla \times \mathbf{F})_y &= 3x^2y + y^3 - 3x^2y - y^3 - y^2z^2 = -y^2z^2,\\
(\nabla \times \mathbf{F})_z &= 3y^2z + 3x^2z + ye^x - 3x^2z - 3y^2z = ye^x.
\end{aligned}$$

Now, S_1 has its normal in the positive z-direction and so only the z-component of $\nabla \times \mathbf{F}$ is needed in the first surface integral of Stokes' theorem. Likewise only the x-component of $\nabla \times \mathbf{F}$ is needed in the second integral, but its value must be subtracted because of the sense in which its contour is traversed:

$$\begin{aligned}
\int_{OABCDEO} (\nabla \times \mathbf{F}) \cdot d\mathbf{r} &= \int_{S_1} (\nabla \times \mathbf{F})_z\,dx\,dy - \int_{S_2} (\nabla \times \mathbf{F})_x\,dy\,dz\\
&= \int_0^1 \int_0^1 ye^x\,dx\,dy - \int_0^1 \int_0^1 2y \times 1 \times z^2\,dy\,dz\\
&= \frac{1}{2}(e-1) - 2\,\frac{1}{2}\,\frac{1}{3} = \frac{e}{2} - \frac{5}{6}.
\end{aligned}$$

As they must, the two methods give the same value.

4 Fourier series

4.1 Prove the orthogonality relations that form the basis of the Fourier series representation of functions.

All of the results are based on the values of the integrals

$$S(n) = \int_{x_0}^{x_0+L} \sin\left(\frac{2\pi nx}{L}\right) dx \quad \text{and} \quad C(n) = \int_{x_0}^{x_0+L} \cos\left(\frac{2\pi nx}{L}\right) dx$$

for integer values of n. Since in all cases with $n \geq 1$ the integrand goes through a whole number of complete cycles, the "area under the curve" is zero. For the case $n = 0$, the integrand in $S(n)$ is zero and so therefore is $S(0)$; for $C(0)$ the integrand is unity and the value of $C(0)$ is L.

We now apply these observations to integrals whose integrands are the products of two sinusoidal functions with arguments that are multiples of a fundamental frequency. The integration interval is equal to the period of that fundamental frequency. To express the integrands in suitable forms, repeated use will be made of the expressions for the sums and differences of sinusoidal functions.

We consider first the product of a sine function and a cosine function:

$$\begin{aligned} I_1 &= \int_{x_0}^{x_0+L} \sin\left(\frac{2\pi rx}{L}\right) \cos\left(\frac{2\pi px}{L}\right) \\ &= \int_{x_0}^{x_0+L} \frac{1}{2}\left[\sin\left(\frac{2\pi(r+p)x}{L}\right) + \sin\left(\frac{2\pi(r-p)x}{L}\right)\right] dx \\ &= \frac{1}{2}[S(r+p) + S(r-p)] = 0, \text{ for all } r \text{ and } p. \end{aligned}$$

Next, we consider the product of two cosines:

$$\begin{aligned} I_2 &= \int_{x_0}^{x_0+L} \cos\left(\frac{2\pi rx}{L}\right) \cos\left(\frac{2\pi px}{L}\right) \\ &= \int_{x_0}^{x_0+L} \frac{1}{2}\left[\cos\left(\frac{2\pi(r+p)x}{L}\right) + \cos\left(\frac{2\pi(r-p)x}{L}\right)\right] dx \\ &= \frac{1}{2}[C(r+p) + C(r-p)] = 0, \end{aligned}$$

unless $r = p > 0$ when $I_2 = \frac{1}{2}L$. If r and p are both zero, then the integrand is unity and $I_2 = L$.

Finally, for the product of two sine functions:

$$\begin{aligned} I_3 &= \int_{x_0}^{x_0+L} \sin\left(\frac{2\pi r x}{L}\right)\sin\left(\frac{2\pi p x}{L}\right) \\ &= \int_{x_0}^{x_0+L} \frac{1}{2}\left[\cos\left(\frac{2\pi(r-p)x}{L}\right) - \cos\left(\frac{2\pi(r+p)x}{L}\right)\right] dx \\ &= \frac{1}{2}[C(r-p) - C(r+p)] = 0, \end{aligned}$$

unless $r = p > 0$ when $I_3 = \frac{1}{2}L$. If either of r and p is zero, then the integrand is zero and $I_3 = 0$.

In summary, all of the integrals have zero value except for those in which the integrand is the square of a single sinusoid. In these cases the integral has value $\frac{1}{2}L$ for all integers r $(= p)$ that are > 0. For r $(= p)$ equal to zero, the $\sin^2$ integral has value zero and the $\cos^2$ integral has value L.

4.3 Which of the following functions of x could be represented by a Fourier series over the range indicated?

(a) $\tanh^{-1}(x)$,	$-\infty < x < \infty$;
(b) $\tan x$,	$-\infty < x < \infty$;
(c) $\lvert \sin x \rvert^{-1/2}$,	$-\infty < x < \infty$;
(d) $\cos^{-1}(\sin 2x)$,	$-\infty < x < \infty$;
(e) $x \sin(1/x)$,	$-\pi^{-1} < x \le \pi^{-1}$, cyclically repeated.

The Dirichlet conditions that a function must satisfy before it can be represented by a Fourier series are:

(i) the function must be periodic;
(ii) it must be single-valued and continuous, except possibly at a finite number of finite discontinuities;
(iii) it must have only a finite number of maxima and minima within one period;
(iv) the integral over one period of $|f(x)|$ must converge.

We now test the given functions against these:

(a) $\tanh^{-1}(x)$ is not a periodic function, since it is only defined for $-1 \le x \le 1$ and changes (monotonically) from $-\infty$ to $+\infty$ as x varies over this restricted range. This function therefore fails condition (i) and *cannot* be represented as a Fourier series.

(b) $\tan x$ is a periodic function but its discontinuities are not finite, nor is its absolute modulus integrable. It therefore fails tests (ii) and (iv) and *cannot* be represented as a Fourier series.

(c) $|\sin x|^{-1/2}$ is a periodic function of period π and, although it becomes infinite at $x = n\pi$, there are no infinite discontinuities. Near $x = 0$, say, it behaves as $|x|^{-1/2}$ and its absolute modulus is therefore integrable. There is only one minimum in any one period. The function therefore satisfies all four Dirichlet conditions and *can* be represented as a Fourier series.

(d) $\cos^{-1}(\sin 2x)$ is clearly a multi-valued function and fails condition (ii); it *cannot* be represented as a Fourier series.

(e) $x\sin(1/x)$, for $-\pi^{-1} < x \le \pi^{-1}$ (cyclically repeated) is clearly cyclic (by definition), continuous, bounded, single-valued and integrable. However, since $\sin(1/x)$ oscillates with unlimited frequency near $x = 0$, there are an infinite number of maxima and minima in any region enclosing $x = 0$. Condition (iii) is therefore not satisfied and the function *cannot* be represented as a Fourier series.

4.5 Find the Fourier series of the function $f(x) = x$ in the range $-\pi < x \le \pi$. Hence show that

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots = \frac{\pi}{4}.$$

This is an odd function in x and so a sine series with period 2π is appropriate. The coefficient of $\sin nx$ will be given by

$$\begin{aligned} b_n &= \frac{2}{2\pi}\int_{-\pi}^{\pi} x\sin nx\,dx \\ &= \frac{1}{\pi}\left\{\left[-\frac{x\cos nx}{n}\right]_{-\pi}^{\pi} + \int_{-\pi}^{\pi}\frac{\cos nx}{n}\,dx\right\} \\ &= \frac{1}{\pi}\left[-\frac{\pi(-1)^n - (-\pi)(-1)^n}{n} + 0\right] = \frac{2(-1)^{n+1}}{n}. \end{aligned}$$

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

We note in passing that although this series is convergent, as it must be, it has poor (i.e. n^{-1}) convergence; this can be put down to the periodic version of the function having a discontinuity (of 2π) at the end of each basic period.

To obtain the sum of a series from such a Fourier representation, we must make a judicious choice for the value of x – making such a choice is rather more of an art than a science! Here, setting $x = \pi/2$ gives

$$\begin{aligned} \frac{\pi}{2} &= 2\sum_{n=1}^{\infty}\frac{(-1)^{n+1}\sin(n\pi/2)}{n} \\ &= 2\sum_{n\text{ odd}}\frac{(-1)^{n+1}(-1)^{(n-1)/2}}{n}, \\ \Rightarrow\quad \frac{\pi}{4} &= \frac{1}{1} - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots. \end{aligned}$$

4.7 For the function

$$f(x) = 1 - x, \qquad 0 \le x \le 1,$$

a Fourier sine series can be found by continuing it in the range $-1 < x \le 0$ as $f(x) = -1 - x$. The function thus has a discontinuity of 2 at $x = 0$. The series is

$$1 - x = f(x) = \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{\sin n\pi x}{n}. \qquad (*)$$

In order to obtain a cosine series, the continuation has to be $f(x) = 1 + x$ in the range $-1 < x \le 0$. The function then has no discontinuity at $x = 0$ and the corresponding series is

$$1 - x = f(x) = \frac{1}{2} + \frac{4}{\pi^2} \sum_{n \text{ odd}} \frac{\cos n\pi x}{n^2}. \qquad (**)$$

For these continued functions and series, consider (i) their derivatives and (ii) their integrals. Do they give meaningful equations? You will probably find it helpful to sketch all the functions involved.

(i) *Derivatives*

(a) The sine series. With the continuation given, the derivative df/dx has the value -1 everywhere, except at the origin where the function is not defined (though $f(0) = 0$ seems the only possible choice), continuous or differentiable. Differentiating the given series $(*)$ for $f(x)$ yields

$$\frac{df}{dx} = 2 \sum_{n=1}^{\infty} \cos n\pi x.$$

This series does not converge and the equation is not meaningful.

(b) The cosine series. With the stated continuation for $f(x)$ the derivative is $+1$ for $-1 < x \le 0$ and is -1 for $0 \le x \le 1$. It is thus the negative of an odd (about $x = 0$) unit square-wave, whose Fourier series is

$$-\frac{4}{\pi} \sum_{n \text{ odd}} \frac{\sin n\pi x}{n}.$$

This is confirmed by differentiating $(**)$ term by term to obtain the same result:

$$\frac{df}{dx} = \frac{4}{\pi^2} \sum_{n \text{ odd}} \frac{-n\pi \sin n\pi x}{n^2} = -\frac{4}{\pi} \sum_{n \text{ odd}} \frac{\sin n\pi x}{n}.$$

(ii) *Integrals*

Since integrals contain an arbitrary constant of integration, we will define $F(-1) = 0$, where $F(x)$ is the indefinite integral of $f(x)$.

(a) The sine series. For $-1 \le x \le 0$,

$$F_a(x) = F(-1) + \int_{-1}^{x} (-1 - x)\, dx = -x - \tfrac{1}{2}x^2 - \tfrac{1}{2}.$$

For $0 \le x \le 1$,

$$F_a(x) = F(0) + \int_0^x (1 - x)\, dx = -\tfrac{1}{2} + \left[x - \tfrac{1}{2}x^2\right]_0^x = x - \tfrac{1}{2}x^2 - \tfrac{1}{2}.$$

This is a continuous function and, like all indefinite integrals, is "smoother" than the function from which it is derived; this latter property will be reflected in the improved convergence of the derived series. Integrating term by term we find that its Fourier series is given by

$$\begin{aligned}
F_a(x) &= \frac{2}{\pi}\int_{-1}^{x}\sum_{n=1}^{\infty}\frac{\sin n\pi x'}{n}\,dx' \\
&= \frac{2}{\pi}\sum_{n=1}^{\infty}\left[-\frac{\cos n\pi x'}{\pi n^2}\right]_{-1}^{x} \\
&= \frac{2}{\pi^2}\sum_{n=1}^{\infty}\frac{(-1)^n - \cos n\pi x}{n^2} \\
&= -\frac{1}{6} - \frac{2}{\pi^2}\sum_{n=1}^{\infty}\frac{\cos n\pi x}{n^2},
\end{aligned}$$

a series that has n^{-2} convergence. Here we have used the result that $\sum_{n=1}^{\infty}(-1)^n n^{-2} = -\pi^2/12$.

(b) The cosine series. The corresponding indefinite integral in this case is

$$\begin{aligned}
F_b(x) &= x + \tfrac{1}{2}x^2 + \tfrac{1}{2} \quad \text{for} \quad -1 \le x \le 0, \\
F_b(x) &= x - \tfrac{1}{2}x^2 + \tfrac{1}{2} \quad \text{for} \quad 0 \le x \le 1,
\end{aligned}$$

and the corresponding integrated series, which has even better convergence (n^{-3}), is given by

$$\frac{1}{2}(x+1) + \frac{4}{\pi^3}\sum_{n \text{ odd}}\frac{\sin n\pi x}{n^3}.$$

However, to have a true Fourier series expression, we must substitute a Fourier series for the $x/2$ term that arises from integrating the constant ($\frac{1}{2}$) in ($**$). This series must be that for $x/2$ across the complete range $-1 \le x \le 1$, and so neither ($*$) nor ($**$) can be rearranged for the purpose. A straightforward calculation (see Problem 4.25 part (b), if necessary) yields the poorly convergent sine series

$$x = 2\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n\pi}\sin n\pi x,$$

and makes the final expression for $F_b(x)$

$$\frac{1}{2} + \sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n\pi}\sin n\pi x + \frac{4}{\pi^3}\sum_{n \text{ odd}}\frac{\sin n\pi x}{n^3}.$$

As will be apparent from a simple sketch, the first series in the above expression dominates; all of its terms are present and it has only n^{-1} convergence. The second series has alternate terms missing and its convergence $\sim n^{-3}$.

4.9 Find the Fourier coefficients in the expansion of $f(x) = \exp x$ over the range $-1 < x < 1$. What value will the expansion have when $x = 2$?

Since the Fourier series will have period 2, we can say immediately that at $x = 2$ the series will converge to the value it has at $x = 0$, namely 1.

As the function $f(x) = \exp x$ is neither even nor odd, its Fourier series will contain both sine and cosine terms. The cosine coefficients are given by

$$\begin{aligned}
a_n &= \frac{2}{2}\int_{-1}^{1} e^x \cos(n\pi x)\,dx \\
&= \left[\cos(n\pi x)\,e^x\right]_{-1}^{1} + \int_{-1}^{1} n\pi\ \sin(n\pi x)\,e^x\,dx \\
&= (-1)^n(e^1 - e^{-1}) + \left[n\pi\ \sin(n\pi x)\,e^x\right]_{-1}^{1} \\
&\quad - \int_{-1}^{1} n^2\pi^2\ \cos(n\pi x)\,e^x\,dx \\
&= 2(-1)^n \sinh 1 - n^2\pi^2 a_n, \\
\Rightarrow \quad a_n &= \frac{2(-1)^n \sinh 1}{1 + n^2\pi^2}.
\end{aligned}$$

Similarly, the sine coefficients are given by

$$\begin{aligned}
b_n &= \frac{2}{2}\int_{-1}^{1} e^x \sin(n\pi x)\,dx \\
&= \left[\sin(n\pi x)\,e^x\right]_{-1}^{1} - \int_{-1}^{1} n\pi\ \cos(n\pi x)\,e^x\,dx \\
&= 0 + \left[-n\pi\ \cos(n\pi x)\,e^x\right]_{-1}^{1} - \int_{-1}^{1} n^2\pi^2\ \sin(n\pi x)\,e^x\,dx \\
&= 2(-1)^{n+1} n\pi \sinh 1 - n^2\pi^2 b_n, \\
\Rightarrow \quad b_n &= \frac{2(-1)^{n+1} n\pi \sinh 1}{1 + n^2\pi^2}.
\end{aligned}$$

4.11 Consider the function $f(x) = \exp(-x^2)$ in the range $0 \le x \le 1$. Show how it should be continued to give as its Fourier series a series (the actual form is not wanted) (a) with only cosine terms, (b) with only sine terms, (c) with period 1 and (d) with period 2.

Would there be any difference between the values of the last two series at (i) $x = 0$, (ii) $x = 1$?

The function and its four continuations are shown as (a)–(d) in Figure 4.1. Note that in the range $0 \le x \le 1$, all four graphs are identical.

Where a continued function has a discontinuity at the ends of its basic period, the series will yield a value at those end-points that is the average of the function's values on the

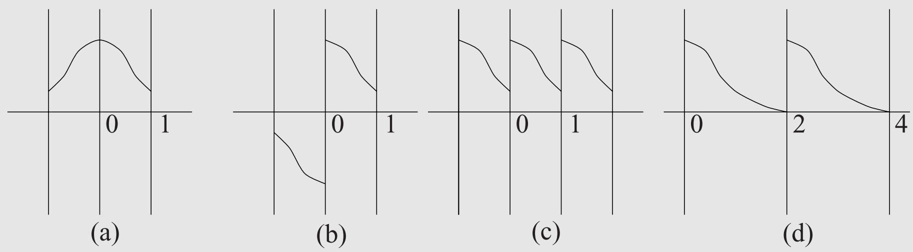

Figure 4.1 The continuations of $\exp(-x^2)$ in $0 \le x \le 1$ to give: (a) cosine terms only; (b) sine terms only; (c) period 1; (d) period 2.

two sides of the discontinuity. Thus for continuation (c) both (i) $x = 0$ and (ii) $x = 1$ are end-points, and the value of the series there will be $(1 + e^{-1})/2$. For continuation (d), $x = 0$ is an end-point, and the series will have value $\frac{1}{2}(1 + e^{-4})$. However, $x = 1$ is not a point of discontinuity, and the series will have the expected value of e^{-1}.

4.13 Consider the representation as a Fourier series of the displacement of a string lying in the interval $0 \le x \le L$ and fixed at its ends, when it is pulled aside by y_0 at the point $x = L/4$. Sketch the continuations for the region outside the interval that will

(a) produce a series of period L,
(b) produce a series that is antisymmetric about $x = 0$, and
(c) produce a series that will contain only cosine terms.
(d) What are (i) the periods of the series in (b) and (c) and (ii) the value of the "a_0-term" in (c)?
(e) Show that a typical term of the series obtained in (b) is

$$\frac{32y_0}{3n^2\pi^2} \sin \frac{n\pi}{4} \sin \frac{n\pi x}{L}.$$

Parts (a), (b) and (c) of Figure 4.2 show the three required continuations. Condition (b) will result in a series containing only sine terms, whilst condition (c) requires the continued function to be symmetric about $x = 0$.

(d) (i) The period in both cases, (b) and (c), is clearly $2L$.

(ii) The average value of the displacement is found from "the area under the triangular curve" to be $(\frac{1}{2}Ly_0)/L = \frac{1}{2}y_0$, and this is the value of the "a_0-term".

(e) For the antisymmetric continuation there will be no cosine terms. The sine term coefficients (for a period of $2L$) are given by

$$b_n = 2\,\frac{2}{2L}\int_0^L f(x)\sin(nkx)\,dx, \text{ where } k = 2\pi/2L = \pi/L.$$

Here use has been made of the antisymmetry about $x = 0$ of both $f(x)$ and the sine functions. However, because of the change in analytic form of $f(x)$ between $x < L/4$ and

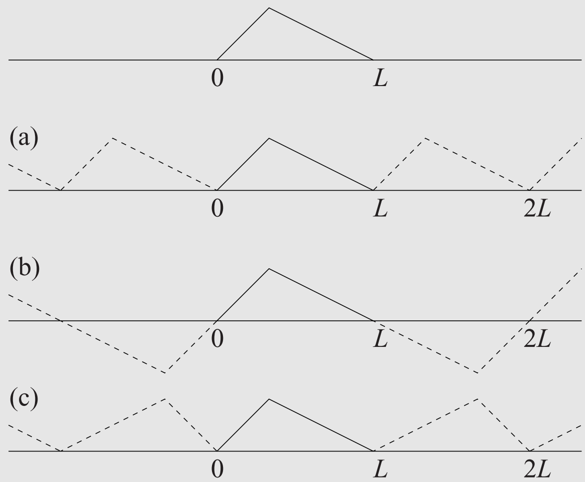

Figure 4.2 Plucked string with fixed ends: (a)–(c) show possible mathematical continuations; (b) is antisymmetric about 0 and (c) is symmetric.

$x > L/4$, the integral will have to be split into two parts. Thus

$$\begin{aligned}
b_n &= \frac{2y_0}{L}\left[\int_0^{L/4} \frac{4x}{L}\sin(nkx)\,dx + \int_{L/4}^{L}\left(\frac{4}{3} - \frac{4x}{3L}\right)\sin(nkx)\,dx\right] \\
&= \frac{8y_0}{3L^2}\left[\int_0^{L/4} 3x\sin(nkx)\,dx + \int_{L/4}^{L}(L-x)\sin(nkx)\,dx\right] \\
&= \frac{8y_0}{3L^2}\left\{\left[-\frac{3x\cos(nkx)}{nk}\right]_0^{L/4} + \int_0^{L/4}\frac{3\cos(nkx)}{nk}\,dx\right. \\
&\quad \left. + \left[-\frac{L\cos(nkx)}{nk}\right]_{L/4}^{L} + \left[\frac{x\cos(nkx)}{nk}\right]_{L/4}^{L} - \int_{L/4}^{L}\frac{\cos(nkx)}{nk}\,dx\right\}.
\end{aligned}$$

Evaluation of the remaining integrals then yields

$$\begin{aligned}
b_n &= \frac{8y_0}{3L^2}\left\{-\frac{3L\cos(n\pi/4)}{4n(\pi/L)} - 0 + \left[\frac{3\sin(nkx)}{n^2k^2}\right]_0^{L/4} - \frac{L\cos(n\pi)}{n(\pi/L)}\right. \\
&\quad \left. + \frac{L\cos(n\pi/4)}{n(\pi/L)} + \frac{L\cos(n\pi)}{n(\pi/L)} - \frac{L\cos(n\pi/4)}{4n(\pi/L)} - \left[\frac{\sin(nkx)}{n^2k^2}\right]_{L/4}^{L}\right\} \\
&= \frac{8y_0}{3L^2}\left[\frac{3L^2\sin(n\pi/4)}{n^2\pi^2} - \frac{L^2\sin(n\pi)}{n^2\pi^2} + \frac{L^2\sin(n\pi/4)}{n^2\pi^2}\right] = \frac{32y_0}{3n^2\pi^2}\sin\left(\frac{n\pi}{4}\right).
\end{aligned}$$

A typical term is therefore

$$\frac{32y_0}{3n^2\pi^2}\sin\left(\frac{n\pi}{4}\right)\sin\left(\frac{n\pi x}{L}\right).$$

We note that every fourth term ($n = 4m$ with m an integer) will be missing.

4.15 The Fourier series for the function $y(x) = |x|$ in the range $-\pi \le x < \pi$ is

$$y(x) = \frac{\pi}{2} - \frac{4}{\pi} \sum_{m=0}^{\infty} \frac{\cos(2m+1)x}{(2m+1)^2}.$$

By integrating this equation term by term from 0 to x, find the function $g(x)$ whose Fourier series is

$$\frac{4}{\pi} \sum_{m=0}^{\infty} \frac{\sin(2m+1)x}{(2m+1)^3}.$$

Using these results, determine, as far as possible by inspection, the form of the functions of which the following are the Fourier series:

(a)

$$\cos\theta + \frac{1}{9}\cos 3\theta + \frac{1}{25}\cos 5\theta + \cdots;$$

(b)

$$\sin\theta + \frac{1}{27}\sin 3\theta + \frac{1}{125}\sin 5\theta + \cdots;$$

(c)

$$\frac{L^2}{3} - \frac{4L^2}{\pi^2}\left[\cos\frac{\pi x}{L} - \frac{1}{4}\cos\frac{2\pi x}{L} + \frac{1}{9}\cos\frac{3\pi x}{L} - \cdots\right].$$

[You may find it helpful to first set $x = 0$ in the quoted result and so obtain values for $S_o = \sum(2m+1)^{-2}$ and other sums derivable from it.]

First, define

$$S = \sum_{\text{all } n \neq 0} n^{-2}, \qquad S_o = \sum_{\text{odd } n} n^{-2}, \qquad S_e = \sum_{\text{even } n \neq 0} n^{-2}.$$

Clearly, $S_e = \frac{1}{4}S$.

Now set $x = 0$ in the quoted result to obtain

$$0 = \frac{\pi}{2} - \frac{4}{\pi} \sum_{m=0}^{\infty} \frac{1}{(2m+1)^2} = \frac{\pi}{2} - \frac{4}{\pi} S_o.$$

Thus, $S_o = \pi^2/8$. Further, $S = S_o + S_e = S_o + \frac{1}{4}S$; it follows that $S = \pi^2/6$ and, by subtraction, that $S_e = \pi^2/24$.

We now consider the integral of $y(x) = |x|$ from 0 to x.

(i) For $x < 0$, $\displaystyle\int_0^x |x|\,dx = \int_0^x (-x)\,dx = -\frac{1}{2}x^2.$

(ii) For $x > 0$, $\displaystyle\int_0^x |x|\,dx = \int_0^x x\,dx = \frac{1}{2}x^2.$

Integrating the series term by term gives

$$\frac{\pi x}{2} - \frac{4}{\pi} \sum_{m=0}^{\infty} \frac{\sin(2m+1)x}{(2m+1)^3}.$$

Equating these two results and isolating the series gives

$$\frac{4}{\pi}\sum_{m=0}^{\infty}\frac{\sin(2m+1)x}{(2m+1)^3} = \tfrac{1}{2}x(\pi - x) \text{ for } x \ge 0,$$
$$= \tfrac{1}{2}x(\pi + x) \text{ for } x \le 0.$$

Questions (a)–(c) are to be solved largely through inspection and so detailed working is not (cannot be) given.

(a) Straightforward substitution of θ for x and rearrangement of the original Fourier series give $g_1(\theta) = \frac{1}{4}\pi(\frac{1}{2}\pi - |\theta|)$.

(b) Straightforward substitution of θ for x and rearrangement of the integrated Fourier series give $g_2(\theta) = \frac{1}{8}\pi\theta(\pi - |\theta|)$.

(c) This contains only cosine terms and is therefore an even function of x. Its average value (given by the a_0 term) is $\frac{1}{3}L^2$. Setting $x = 0$ gives

$$\begin{aligned} f(0) &= \frac{L^2}{3} - \frac{4L^2}{\pi^2}\left(1 - \frac{1}{4} + \frac{1}{9} - \cdots\right) \\ &= \frac{L^2}{3} - \frac{4L^2}{\pi^2}(S_o - S_e) \\ &= \frac{L^2}{3} - \frac{4L^2}{\pi^2}\left(\frac{\pi^2}{8} - \frac{\pi^2}{24}\right) = 0. \end{aligned}$$

Setting $x = L$ gives

$$\begin{aligned} f(L) &= \frac{L^2}{3} - \frac{4L^2}{\pi^2}\left(-1 - \frac{1}{4} - \frac{1}{9} - \cdots\right) \\ &= \frac{L^2}{3} - \frac{4L^2}{\pi^2}(-S) = L^2. \end{aligned}$$

All of this evidence suggests that $f(x) = x^2$ (which it is).

4.17 Find the (real) Fourier series of period 2 for $f(x) = \cosh x$ and $g(x) = x^2$ in the range $-1 \le x \le 1$. By integrating the series for $f(x)$ twice, prove that

$$\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n^2\pi^2(n^2\pi^2+1)} = \frac{1}{2}\left(\frac{1}{\sinh 1} - \frac{5}{6}\right).$$

Since both functions are even, we need consider only constants and cosine terms. The series for x^2 can be calculated directly or, more easily, by using the result of the final part of Problem 4.15 with L set equal to 1:

$$g(x) = x^2 = \frac{1}{3} + \frac{4}{\pi^2}\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\cos\pi n x \text{ for } -1 \le x \le 1.$$

For $f(x) = \cosh x$,

$$a_0 = \frac{2}{2} 2 \int_0^1 \cosh x \, dx = 2\sinh(1),$$

$$\begin{aligned} a_n &= \frac{2}{2} 2 \int_0^1 \cosh x \cos(n\pi x) \, dx \\ &= 2\left[\frac{\cosh x \sin(n\pi x)}{n\pi}\right]_0^1 - 2\int_0^1 \frac{\sinh x \sin(n\pi x)}{n\pi} \, dx \\ &= 0 + 2\left[\frac{\sinh x \cos(n\pi x)}{n^2\pi^2}\right]_0^1 - \frac{a_n}{n^2\pi^2}. \end{aligned}$$

Rearranging this gives

$$a_n = \frac{(-1)^n 2\sinh(1)}{1 + n^2\pi^2}.$$

Thus,

$$\cosh x = \sinh(1)\left(1 + 2\sum_{n=1}^{\infty} \frac{(-1)^n}{1 + n^2\pi^2} \cos n\pi x\right).$$

We now integrate this expansion twice from 0 to x (anticipating that we will recover a hyperbolic cosine function plus some additional terms). Since $\sinh(0) = \sin(m\pi 0) = 0$, the first integration yields

$$\sinh x = \sinh(1)\left(x + 2\sum_{n=1}^{\infty} \frac{(-1)^n}{n\pi(1 + n^2\pi^2)} \sin n\pi x\right).$$

For the second integration we use $\cosh(0) = \cos(m\pi 0) = 1$ to obtain

$$\cosh(x) - 1 = \sinh(1)\left(\frac{1}{2}x^2 + 2\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2\pi^2(1 + n^2\pi^2)}[\cos(n\pi x) - 1]\right).$$

However, this expansion must be the same as the original expansion for $\cosh(x)$ after a Fourier series has been substituted for the $\frac{1}{2}\sinh(1)x^2$ term. The coefficients of $\cos n\pi x$ in the two expressions must be equal; in particular, the equality of the constant terms (formally $\cos n\pi x$ with $n = 0$) requires that

$$\sinh(1) - 1 = \frac{1}{2}\sinh(1)\,\frac{1}{3} + 2\sinh(1)\sum_{n=1}^{\infty} \frac{(-1)^{n+2}}{n^2\pi^2(1 + n^2\pi^2)},$$

i.e.

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2\pi^2(n^2\pi^2 + 1)} = \frac{1}{2}\left(\frac{1}{\sinh 1} - \frac{5}{6}\right),$$

as stated in the question.

4.19 Demonstrate explicitly for the odd (about $x = 0$) square-wave function that Parseval's theorem is valid. You will need to use the relationship

$$\sum_{m=0}^{\infty} \frac{1}{(2m+1)^2} = \frac{\pi^2}{8}.$$

Show that a filter that transmits frequencies only up to $8\pi/T$ will still transmit more than 90% of the power in a square-wave voltage signal of period T.

As stated in the solution to Problem 4.7, and in virtually every textbook, the odd square-wave function has only the odd harmonics present in its Fourier sine series representation. The coefficient of the $\sin(2m+1)\pi x$ term is

$$b_{2m+1} = \frac{4}{(2m+1)\pi}.$$

For a periodic function of period L whose complex Fourier coefficients are c_r, or whose cosine and sine coefficients are a_r and b_r, respectively, Parseval's theorem for one function states that

$$\begin{aligned}\frac{1}{L}\int_{x_0}^{x_0+L} |f(x)|^2 dx &= \sum_{r=-\infty}^{\infty} |c_r|^2 \\ &= \left(\tfrac{1}{2}a_0\right)^2 + \tfrac{1}{2}\sum_{r=1}^{\infty}(a_r^2 + b_r^2),\end{aligned}$$

and therefore requires in this particular case, in which all the a_r are zero and $L = 2$, that

$$\frac{1}{2}\sum_{m=0}^{\infty} \frac{16}{(2m+1)^2\pi^2} = \frac{1}{2}\sum_{n=1}^{\infty} b_n^2 = \frac{1}{2}\int_{-1}^{1} |\pm 1|^2\, dx = 1.$$

Since

$$\sum_{m=0}^{\infty} \frac{1}{(2m+1)^2} = \frac{\pi^2}{8},$$

this reduces to the identity

$$\frac{1}{2}\,\frac{16}{\pi^2}\,\frac{\pi^2}{8} = 1.$$

The power at any particular frequency in an electrical signal is proportional to the square of the amplitude at that frequency, i.e. to $|b_n|^2$ in the present case. If the filter passes only frequencies up to $8\pi/T = 4\omega$, then only the $n = 1$ and the $n = 3$ components will be passed. They contribute a fraction

$$\left(\frac{1}{1} + \frac{1}{9}\right) \div \frac{\pi^2}{8} = 0.901$$

of the total, i.e. more than 90%.

4.21 Find the complex Fourier series for the periodic function of period 2π defined in the range $-\pi \le x \le \pi$ by $y(x) = \cosh x$. By setting $x = 0$ prove that

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2+1} = \frac{1}{2}\left(\frac{\pi}{\sinh \pi} - 1\right).$$

We first note that, although $\cosh x$ is an even function of x, e^{-inx} is neither even nor odd. Consequently it will not be possible to convert the integral into one over the range $0 \le x \le \pi$. The complex Fourier coefficients c_n $(-\infty < n < \infty)$ are therefore calculated as

$$\begin{aligned}
c_n &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \cosh x \, e^{-inx} \, dx \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{2} \left(e^{-inx+x} + e^{-inx-x}\right) dx \\
&= \frac{1}{4\pi} \left[\frac{e^{(1-in)x}}{1-in}\right]_{-\pi}^{\pi} + \frac{1}{4\pi} \left[\frac{e^{(-1-in)x}}{-1-in}\right]_{-\pi}^{\pi} \\
&= \frac{1}{4\pi} \frac{(1+in)(-1)^n(2\sinh \pi) - (1-in)(-1)^n(-2\sinh \pi)}{1+n^2} \\
&= \frac{(-1)^n 4\sinh(\pi)}{4\pi(1+n^2)}.
\end{aligned}$$

Thus,

$$\cosh x = \sum_{n=-\infty}^{\infty} \frac{(-1)^n \sinh \pi}{\pi(1+n^2)} \, e^{inx}.$$

We now set $x = 0$ on both sides of the equation:

$$\begin{aligned}
1 &= \sum_{n=-\infty}^{\infty} \frac{(-1)^n \sinh \pi}{\pi(1+n^2)}, \\
\Rightarrow \quad \sum_{n=-\infty}^{\infty} \frac{(-1)^n}{1+n^2} &= \frac{\pi}{\sinh \pi}.
\end{aligned}$$

Separating out the $n = 0$ term, and noting that $(-1)^n = (-1)^{-n}$, now gives

$$1 + 2\sum_{n=1}^{\infty} \frac{(-1)^n}{1+n^2} = \frac{\pi}{\sinh \pi}$$

and hence the stated result.

4.23 The complex Fourier series for the periodic function generated by $f(t) = \sin t$ for $0 \le t \le \pi/2$, and repeated in every subsequent interval of $\pi/2$, is

$$\sin(t) = \frac{2}{\pi} \sum_{n=-\infty}^{\infty} \frac{4ni - 1}{16n^2 - 1} e^{i4nt}.$$

Apply Parseval's theorem to this series and so derive a value for the sum of the series

$$\frac{17}{(15)^2} + \frac{65}{(63)^2} + \frac{145}{(143)^2} + \cdots + \frac{16n^2 + 1}{(16n^2 - 1)^2} + \cdots.$$

Applying Parseval's theorem (see Solution 4.19) in a straightforward manner to the given equation:

$$\frac{2}{\pi} \int_0^{\pi/2} \sin^2(t)\, dt = \frac{4}{\pi^2} \sum_{n=-\infty}^{\infty} \frac{4ni - 1}{16n^2 - 1} \frac{-4ni - 1}{16n^2 - 1},$$

$$\frac{2}{\pi}\,\frac{1}{2}\,\frac{\pi}{2} = \frac{4}{\pi^2} \sum_{n=-\infty}^{\infty} \frac{16n^2 + 1}{(16n^2 - 1)^2},$$

$$\frac{\pi^2}{8} = 1 + 2\sum_{n=1}^{\infty} \frac{16n^2 + 1}{(16n^2 - 1)^2},$$

$$\Rightarrow \quad \sum_{n=1}^{\infty} \frac{16n^2 + 1}{(16n^2 - 1)^2} = \frac{\pi^2 - 8}{16}.$$

To obtain the second line we have used the standard result that the average value of the square of a sinusoid is $1/2$.

4.25 Show that Parseval's theorem for two real functions whose Fourier expansions have cosine and sine coefficients a_n, b_n and α_n, β_n takes the form

$$\frac{1}{L} \int_0^L f(x)g(x)\, dx = \frac{1}{4} a_0 \alpha_0 + \frac{1}{2} \sum_{n=1}^{\infty} (a_n \alpha_n + b_n \beta_n).$$

(a) Demonstrate that for $g(x) = \sin mx$ or $\cos mx$ this reduces to the definition of the Fourier coefficients.

(b) Explicitly verify the above result for the case in which $f(x) = x$ and $g(x)$ is the square-wave function, both in the interval $-1 \le x \le 1$.

[Note that $g = g^*$, and it is the integral of fg^* that will have to be formally evaluated using the complex Fourier series representations of the two functions.]

If c_n and γ_n are the complex Fourier coefficients for the real functions $f(x)$ and $g(x)$ that have real Fourier coefficients a_n, b_n and α_n, β_n, respectively, then

$$c_n = \tfrac{1}{2}(a_n - ib_n) \quad \text{and} \quad \gamma_n = \alpha_n - i\beta_n,$$
$$c_{-n} = \tfrac{1}{2}(a_n + ib_n) \quad \text{and} \quad \gamma_{-n} = \alpha_n + i\beta_n.$$

The two functions can be written as

$$f(x) = \sum_{n=-\infty}^{\infty} c_n \exp\left(\frac{2\pi inx}{L}\right),$$

$$g(x) = \sum_{n=-\infty}^{\infty} \gamma_n \exp\left(\frac{2\pi inx}{L}\right). \qquad (*)$$

Thus,

$$f(x)g^*(x) = \sum_{n=-\infty}^{\infty} c_n g^*(x) \exp\left(\frac{2\pi inx}{L}\right).$$

Integrating this equation with respect to x over the interval $(0, L)$ and dividing by L, we find

$$\begin{aligned}\frac{1}{L}\int_0^L f(x)g^*(x)\,dx &= \sum_{n=-\infty}^{\infty} c_n \frac{1}{L}\int_0^L g^*(x)\exp\left(\frac{2\pi inx}{L}\right)dx \\ &= \sum_{n=-\infty}^{\infty} c_n \left[\frac{1}{L}\int_0^L g(x)\exp\left(\frac{-2\pi inx}{L}\right)dx\right]^* \\ &= \sum_{n=-\infty}^{\infty} c_n \gamma_n^*.\end{aligned}$$

To obtain the last line we have used the inverse of relationship $(*)$.

Dividing up the sum over all n into a sum over positive n, a sum over negative n and the $n = 0$ term, and then substituting for c_n and γ_n, gives

$$\begin{aligned}\frac{1}{L}\int_0^L f(x)g^*(x)\,dx &= \frac{1}{4}\sum_{n=1}^{\infty}(a_n - ib_n)(\alpha_n + i\beta_n) \\ &\quad + \frac{1}{4}\sum_{n=1}^{\infty}(a_n + ib_n)(\alpha_n - i\beta_n) + \frac{1}{4}a_0\alpha_0 \\ &= \frac{1}{4}\sum_{n=1}^{\infty}(2a_n\alpha_n + 2b_n\beta_n) + \frac{1}{4}a_0\alpha_0 \\ &= \frac{1}{2}\sum_{n=1}^{\infty}(a_n\alpha_n + b_n\beta_n) + \frac{1}{4}a_0\alpha_0,\end{aligned}$$

i.e. the stated result.

(a) For $g(x) = \sin mx$, $\beta_m = 1$ and all other α_n and β_n are zero. The above equation then reduces to

$$\frac{1}{L}\int_0^L f(x)\sin(mx)\,dx = \frac{1}{2}b_n,$$

which is the normal definition of b_n. Similarly, setting $g(x) = \cos mx$ leads to the normal definition of a_n.

(b) For the function $f(x) = x$ in the interval $-1 < x \leq 1$, the sine coefficients are

$$\begin{aligned}
b_n &= \frac{2}{2}\int_{-1}^{1} x \sin n\pi x \, dx \\
&= 2\int_{0}^{1} x \sin n\pi x \, dx \\
&= 2\left\{\left[\frac{-x \cos n\pi x}{n\pi}\right]_0^1 + \int_0^1 \frac{\cos n\pi x}{n\pi}\, dx\right\} \\
&= 2\left\{\frac{(-1)^{n+1}}{n\pi} + \left[\frac{\sin n\pi x}{n^2\pi^2}\right]_0^1\right\} \\
&= \frac{2(-1)^{n+1}}{n\pi}.
\end{aligned}$$

As stated in Problem 4.19, for the (antisymmetric) square-wave function $\beta_n = 4/(n\pi)$ for odd n and $\beta_n = 0$ for even n.

Now the integral

$$\frac{1}{L}\int_0^L f(x)g^*(x)\,dx = \frac{1}{2}\left[\int_{-1}^{0}(-1)x\,dx + \int_0^1 (+1)x\,dx\right] = \frac{1}{2},$$

whilst

$$\frac{1}{2}\sum_{n=1}^{\infty} b_n\beta_n = \frac{1}{2}\sum_{n\text{ odd}} \frac{4}{n\pi}\,\frac{2(-1)^{n+1}}{n\pi} = \frac{4}{\pi^2}\sum_{n\text{ odd}}\frac{1}{n^2} = \frac{4}{\pi^2}\,\frac{\pi^2}{8} = \frac{1}{2}.$$

The value of the sum $\sum n^{-2}$ for odd n is taken from $S_{\rm o}$ in the solution to Problem 4.15. Thus, the two sides of the equation agree, verifying the validity of Parseval's theorem in this case.

5 Integral transforms

5.1 Find the Fourier transform of the function $f(t) = \exp(-|t|)$.

(a) By applying Fourier's inversion theorem prove that

$$\frac{\pi}{2}\exp(-|t|) = \int_0^\infty \frac{\cos\omega t}{1+\omega^2}\,d\omega.$$

(b) By making the substitution $\omega = \tan\theta$, demonstrate the validity of Parseval's theorem for this function.

As the function $|t|$ is not representable by the same integrable function throughout the integration range, we must divide the range into two sections and use different explicit expressions for the integrand in each:

$$\begin{aligned}
\tilde{f}(\omega) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-|t|}\,e^{-i\omega t}\,dt \\
&= \frac{1}{\sqrt{2\pi}}\int_0^\infty e^{-(1+i\omega)t}\,dt + \frac{1}{\sqrt{2\pi}}\int_{-\infty}^0 e^{(1-i\omega)t}\,dt \\
&= \frac{1}{\sqrt{2\pi}}\left(\frac{1}{1+i\omega} + \frac{1}{1-i\omega}\right) \\
&= \frac{1}{\sqrt{2\pi}}\frac{2}{1+\omega^2}.
\end{aligned}$$

(a) Substituting this result into the inversion theorem gives

$$\exp^{-|t|} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{2}{\sqrt{2\pi}(1+\omega^2)}\,e^{i\omega t}\,d\omega.$$

Equating the real parts on the two sides of this equation and noting that the resulting integrand is symmetric in ω, shows that

$$\exp^{-|t|} = \frac{2}{\pi}\int_0^\infty \frac{\cos\omega t}{(1+\omega^2)}\,d\omega,$$

as given in the question.

(b) For Parseval's theorem, which states that

$$\int_{-\infty}^{\infty}|f(t)|^2\,dt = \int_{-\infty}^{\infty}|\tilde{f}(\omega)|^2\,d\omega,$$

we first evaluate

$$\int_{-\infty}^{\infty} |f(t)|^2\,dt = \int_{-\infty}^{0} e^{2t}\,dt + \int_{0}^{\infty} e^{-2t}\,dt$$
$$= 2\int_{0}^{\infty} e^{-2t}\,dt$$
$$= 2\left[\frac{e^{-2t}}{-2}\right]_0^{\infty} = 1.$$

The second integral, over ω, is

$$\int_{-\infty}^{\infty} |\tilde{f}(\omega)|^2\,d\omega = 2\int_{0}^{\infty} \frac{2}{\pi(1+\omega^2)^2}\,d\omega, \quad \text{set } \omega \text{ equal to } \tan\theta,$$
$$= \frac{4}{\pi}\int_{0}^{\pi/2} \frac{1}{\sec^4\theta}\sec^2\theta\,d\theta$$
$$= \frac{4}{\pi}\int_{0}^{\pi/2} \cos^2\theta\,d\theta = \frac{4}{\pi}\,\frac{1}{2}\,\frac{\pi}{2} = 1,$$

i.e. the same as the first one, thus verifying the theorem for this function.

5.3 Find the Fourier transform of $H(x-a)e^{-bx}$, where $H(x)$ is the Heaviside function.

The Heaviside function $H(x)$ has value 0 for $x < 0$ and value 1 for $x \geq 0$. Write $H(x-a)e^{-bx} = h(x)$ with b assumed > 0. Then,

$$\tilde{h}(k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} H(x-a)e^{-bx}\,e^{-ikx}\,dx$$
$$= \frac{1}{\sqrt{2\pi}}\int_{a}^{\infty} e^{-bx-ikx}\,dx$$
$$= \frac{1}{\sqrt{2\pi}}\left[\frac{e^{-bx-ikx}}{-b-ik}\right]_a^{\infty}$$
$$= \frac{1}{\sqrt{2\pi}}\,\frac{e^{-ba}e^{-ika}}{b+ik} = e^{-ika}\,\frac{e^{-ba}}{\sqrt{2\pi}}\,\frac{b-ik}{b^2+k^2}.$$

This same result could be obtained by setting $y = x - a$, finding the transform of $e^{-ba}e^{-by}$, and then using the translation property of Fourier transforms.

5.5 By taking the Fourier transform of the equation

$$\frac{d^2\phi}{dx^2} - K^2\phi = f(x),$$

show that its solution, $\phi(x)$, can be written as

$$\phi(x) = \frac{-1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \frac{e^{ikx}\widetilde{f}(k)}{k^2+K^2}\,dk,$$

where $\widetilde{f}(k)$ is the Fourier transform of $f(x)$.

We take the Fourier transform of each term of

$$\frac{d^2\phi}{dx^2} - K^2\phi = f(x)$$

to give

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \frac{d^2\phi}{dx^2}\, e^{-ikx}\, dx - K^2\tilde{\phi}(k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(x)\, e^{-ikx}\, dx.$$

Since ϕ must vanish at $\pm\infty$, the first term can be integrated twice by parts with no contributions at the end-points. This gives the full equation as

$$-k^2\tilde{\phi}(k) - K^2\tilde{\phi}(k) = \tilde{f}(k).$$

Now, by the Fourier inversion theorem,

$$\begin{aligned}\phi(x) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \tilde{\phi}(k)\, e^{ikx}\, dk \\ &= -\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \frac{\tilde{f}(k)\, e^{ikx}}{k^2 + K^2}\, dk.\end{aligned}$$

Note
The principal advantage of this Fourier approach to a set of one or more linear differential equations is that the differential operators act only on exponential functions whose exponents are linear in x. This means that the derivatives are no more than multiples of the original function and what were originally differential equations are turned into algebraic ones. As the differential equations are linear the algebraic equations can be solved explicitly for the transforms of their solutions, and the solutions themselves may then be found using the inversion theorem. The "price" to be paid for this great simplification is that the inversion integral may not be tractable analytically, but, as a last resort, numerical integration can always be employed.

5.7 Find the Fourier transform of the unit rectangular distribution

$$f(t) = \begin{cases} 1 & |t| < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Determine the convolution of f with itself and, without further integration, determine its transform. Deduce that

$$\int_{-\infty}^{\infty} \frac{\sin^2\omega}{\omega^2}\, d\omega = \pi, \qquad \int_{-\infty}^{\infty} \frac{\sin^4\omega}{\omega^4}\, d\omega = \frac{2\pi}{3}.$$

The function to be transformed is unity in the range $-1 \le t \le 1$ and so

$$\tilde{f}(\omega) = \frac{1}{\sqrt{2\pi}}\int_{-1}^{1} 1\, e^{-i\omega t}\, dt = \frac{1}{\sqrt{2\pi}}\left[\frac{e^{-i\omega} - e^{i\omega}}{-i\omega}\right] = \frac{2\sin\omega}{\sqrt{2\pi}\,\omega}.$$

Denote by $p(t)$ the convolution of f with itself and, in the second line of the calculation below, change the integration variable from s to $u = t - s$:

$$\begin{aligned} p(t) &\equiv \int_{-\infty}^{\infty} f(t-s)f(s)\,ds = \int_{-1}^{1} f(t-s)\,1\,ds \\ &= \int_{t+1}^{t-1} f(u)(-du) = \int_{t-1}^{t+1} f(u)du. \end{aligned}$$

It follows that

$$p(t) = \begin{cases} (t+1)-(-1) & 0 > t > -2 \\ 1-(t-1) & 2 > t > 0 \end{cases} = \begin{cases} 2-|t| & 0 < |t| < 2, \\ 0 & \text{otherwise.} \end{cases}$$

The transform of p is given directly by the convolution theorem [which states that if $h(t)$, given by $h = f * g$, is the convolution of f and g, then $\tilde{h} = \sqrt{2\pi}\,\tilde{f}\,\tilde{g}$] as

$$\tilde{p}(\omega) = \sqrt{2\pi}\,\frac{2\sin\omega}{\sqrt{2\pi}\,\omega}\,\frac{2\sin\omega}{\sqrt{2\pi}\,\omega} = \frac{4}{\sqrt{2\pi}}\,\frac{\sin^2\omega}{\omega^2}.$$

Noting that the two integrals to be evaluated have as integrands the squares of functions that are essentially the known transforms of simple functions, we are led to apply Parseval's theorem to each. Applying the theorem to $f(t)$ and $p(t)$ yields

$$\int_{-\infty}^{\infty} \frac{4\sin^2\omega}{2\pi\omega^2}\,d\omega = \int_{-\infty}^{\infty} |f(t)|^2\,dt = 2 \quad \Rightarrow \quad \int_{-\infty}^{\infty} \frac{\sin^2\omega}{\omega^2} = \pi,$$

and

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{16}{2\pi}\,\frac{\sin^4\omega}{\omega^4}\,d\omega &= \int_{-2}^{0} (2+t)^2\,dt + \int_{0}^{2} (2-t)^2\,dt \\ &= \left[\frac{(2+t)^3}{3}\right]_{-2}^{0} - \left[\frac{(2-t)^3}{3}\right]_{0}^{2} \\ &= \frac{8}{3} + \frac{8}{3}, \\ \Rightarrow \quad \int_{-\infty}^{\infty} \frac{\sin^4\omega}{\omega^4}\,d\omega &= \frac{2\pi}{3}. \end{aligned}$$

5.9 By finding the complex Fourier *series* for its LHS show that either side of the equation

$$\sum_{n=-\infty}^{\infty} \delta(t+nT) = \frac{1}{T}\sum_{n=-\infty}^{\infty} e^{-2\pi nit/T}$$

can represent a periodic train of impulses. By expressing the function $f(t+nX)$, in which X is a constant, in terms of the Fourier *transform* $\tilde{f}(\omega)$ of $f(t)$, show that

$$\sum_{n=-\infty}^{\infty} f(t+nX) = \frac{\sqrt{2\pi}}{X}\sum_{n=-\infty}^{\infty} \tilde{f}\left(\frac{2n\pi}{X}\right) e^{2\pi nit/X}.$$

This result is known as the *Poisson summation formula*.

Denote by $g(t)$ the periodic function $\sum_{n=-\infty}^{\infty} \delta(t+nT)$ with $2\pi/T = \omega$. Its complex Fourier coefficients are given by

$$c_n = \frac{1}{T}\int_0^T g(t)\, e^{-in\omega t}\, dt = \frac{1}{T}\int_0^T \delta(t)\, e^{-in\omega t}\, dt = \frac{1}{T}.$$

Thus, by the inversion theorem, its Fourier *series* representation is

$$g(t) = \sum_{n=-\infty}^{\infty} \frac{1}{T}\, e^{in\omega t} = \sum_{n=-\infty}^{\infty} \frac{1}{T}\, e^{-in\omega t} = \sum_{n=-\infty}^{\infty} \frac{1}{T}\, e^{-i2n\pi t/T},$$

showing that both this sum and the original one are representations of a periodic train of impulses.

In this result,

$$\sum_{n=-\infty}^{\infty} \delta(t+nT) = \frac{1}{T}\sum_{n=-\infty}^{\infty} e^{-2\pi n i t/T},$$

we now make the changes of variable $t \to \omega$, $n \to -n$ and $T \to 2\pi/X$ and obtain

$$\sum_{n=-\infty}^{\infty} \delta\left(\omega - \frac{2\pi n}{X}\right) = \frac{X}{2\pi}\sum_{n=-\infty}^{\infty} e^{inX\omega}. \quad (*)$$

If we denote $f(t+nX)$ by $f_{nX}(t)$ then, by the translation theorem, we have $\tilde{f}_{nX}(\omega) = e^{inX\omega}\tilde{f}(\omega)$ and

$$\begin{aligned}
f(t+nX) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \tilde{f}_{nX}(\omega)\, e^{i\omega t}\, d\omega \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{inX\omega}\, \tilde{f}(\omega)\, e^{i\omega t}\, d\omega, \\
\sum_{n=-\infty}^{\infty} f(t+nX) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \tilde{f}(\omega)\, e^{i\omega t} \sum_{n=-\infty}^{\infty} e^{inX\omega}\, d\omega, \text{ use } (*) \text{ above,} \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \tilde{f}(\omega)\, e^{i\omega t}\, \frac{2\pi}{X} \sum_{n=-\infty}^{\infty} \delta\left(\omega - \frac{2\pi n}{X}\right) d\omega \\
&= \frac{\sqrt{2\pi}}{X}\sum_{n=-\infty}^{\infty} \tilde{f}\left(\frac{2\pi n}{X}\right) e^{i2\pi n t/X}.
\end{aligned}$$

In the final line we have made use of the properties of a δ-function when it appears as a factor in an integrand.

5.11 For a function $f(t)$ that is non-zero only in the range $|t| < T/2$, the full frequency spectrum $\tilde{f}(\omega)$ can be constructed, in principle exactly, from values at discrete sample points $\omega = n(2\pi/T)$. Prove this as follows.

(a) Show that the coefficients of a complex Fourier *series* representation of $f(t)$ with period T can be written as

$$c_n = \frac{\sqrt{2\pi}}{T}\,\tilde{f}\left(\frac{2\pi n}{T}\right).$$

(b) Use this result to represent $f(t)$ as an infinite sum in the defining integral for $\tilde{f}(\omega)$, and hence show that

$$\tilde{f}(\omega) = \sum_{n=-\infty}^{\infty} \tilde{f}\left(\frac{2\pi n}{T}\right)\,\text{sinc}\left(n\pi - \frac{\omega T}{2}\right),$$

where $\text{sinc}\,x$ is defined as $(\sin x)/x$.

(a) The complex coefficients for the Fourier series for $f(t)$ are given by

$$c_n = \frac{1}{T}\int_{-T/2}^{T/2} f(t)\,e^{-i2\pi nt/T}\,dt.$$

But, we also know that the Fourier transform of $f(t)$ is given by

$$\tilde{f}(\omega) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(t)\,e^{-i\omega t}\,dt = \frac{1}{\sqrt{2\pi}}\int_{-T/2}^{T/2} f(t)\,e^{-i\omega t}\,dt.$$

Comparison of these two equations shows that $c_n = \frac{1}{T}\sqrt{2\pi}\,\tilde{f}\left(\frac{2\pi n}{T}\right)$.

(b) Using the Fourier series representation of $f(t)$, the frequency spectrum at a general frequency ω can now be constructed as

$$\begin{aligned}
\tilde{f}(\omega) &= \frac{1}{\sqrt{2\pi}}\int_{-T/2}^{T/2} f(t)\,e^{-i\omega t}\,dt \\
&= \frac{1}{\sqrt{2\pi}}\int_{-T/2}^{T/2}\left[\sum_{n=-\infty}^{\infty}\frac{1}{T}\sqrt{2\pi}\,\tilde{f}\left(\frac{2\pi n}{T}\right)e^{i2\pi nt/T}\right]e^{-i\omega t}\,dt \\
&= \frac{1}{T}\sum_{n=-\infty}^{\infty}\tilde{f}\left(\frac{2\pi n}{T}\right)\frac{2\sin\left(\dfrac{2\pi n}{2}-\dfrac{\omega T}{2}\right)}{\dfrac{2\pi n}{T}-\omega} = \sum_{n=-\infty}^{\infty}\tilde{f}\left(\frac{2\pi n}{T}\right)\text{sinc}\left(n\pi - \frac{\omega T}{2}\right).
\end{aligned}$$

This final formula gives a prescription for calculating the frequency spectrum $\tilde{f}(\omega)$ of $f(t)$ for *any* ω, given the spectrum at the (admittedly infinite number of) discrete values $\omega = 2\pi n/T$. The sinc functions give the weights to be assigned to the known discrete values; of course, the weights vary as ω is varied, with, as expected, the largest weights for the nth contribution occurring when ω is close to $2\pi n/T$.

5.13 Find the Fourier transform specified in part (a) and then use it to answer part (b).

(a) Find the Fourier transform of

$$f(\gamma, p, t) = \begin{cases} e^{-\gamma t} \sin pt & t > 0, \\ 0 & t < 0, \end{cases}$$

where γ (> 0) and p are constant parameters.

(b) The current $I(t)$ flowing through a certain system is related to the applied voltage $V(t)$ by the equation

$$I(t) = \int_{-\infty}^{\infty} K(t-u)V(u)\, du,$$

where

$$K(\tau) = a_1 f(\gamma_1, p_1, \tau) + a_2 f(\gamma_2, p_2, \tau).$$

The function $f(\gamma, p, t)$ is as given in part (a) and all the a_i, γ_i (> 0) and p_i are fixed parameters. By considering the Fourier transform of $I(t)$, find the relationship that must hold between a_1 and a_2 if the total net charge Q passed through the system (over a very long time) is to be zero for an arbitrary applied voltage.

(a) Write the given sine function in terms of exponential functions. Its Fourier transform is then easily calculated as

$$\begin{aligned} \tilde{f}(\omega, \gamma, p) &= \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \frac{e^{(-\gamma - i\omega + ip)t} - e^{(-\gamma - i\omega - ip)t}}{2i}\, dt \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{2i} \left(\frac{-1}{-\gamma - i\omega + ip} + \frac{1}{-\gamma - i\omega - ip} \right) \\ &= \frac{1}{\sqrt{2\pi}} \frac{p}{(\gamma + i\omega)^2 + p^2}. \end{aligned}$$

(b) Since the current is given by the convolution

$$I(t) = \int_{-\infty}^{\infty} K(t-u)V(u)\, du,$$

the convolution theorem implies that the Fourier transforms of I, K and V are related by $\tilde{I}(\omega) = \sqrt{2\pi}\, \tilde{K}(\omega)\, \tilde{V}(\omega)$ with, from part (a),

$$\tilde{K}(\omega) = \frac{1}{\sqrt{2\pi}} \left[\frac{a_1 p_1}{(\gamma_1 + i\omega)^2 + p_1^2} + \frac{a_2 p_2}{(\gamma_2 + i\omega)^2 + p_2^2} \right].$$

Now, by expressing $I(t')$ in its Fourier integral form, we can write

$$Q(\infty) = \int_{-\infty}^{\infty} I(t')\, dt' = \int_{-\infty}^{\infty} dt' \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\, \tilde{I}(\omega)\, e^{i\omega t'}\, d\omega.$$

But $\int_{-\infty}^{\infty} e^{i\omega t'}\,dt' = 2\pi\,\delta(\omega)$ and so

$$\begin{aligned} Q(\infty) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\,\tilde{I}(\omega)\,2\pi\,\delta(\omega)\,d\omega \\ &= \frac{2\pi}{\sqrt{2\pi}}\,\tilde{I}(0) = \sqrt{2\pi}\,\sqrt{2\pi}\,\tilde{K}(0)\,\tilde{V}(0) \\ &= 2\pi\,\frac{1}{\sqrt{2\pi}}\left[\frac{a_1 p_1}{\gamma_1^2 + p_1^2} + \frac{a_2 p_2}{\gamma_2^2 + p_2^2}\right]\tilde{V}(0). \end{aligned}$$

For $Q(\infty)$ to be zero for an arbitrary $V(t)$, we must have

$$\frac{a_1 p_1}{\gamma_1^2 + p_1^2} + \frac{a_2 p_2}{\gamma_2^2 + p_2^2} = 0,$$

and so this is the required relationship.

5.15 A linear amplifier produces an output that is the convolution of its input and its response function. The Fourier transform of the response function for a particular amplifier is

$$\tilde{K}(\omega) = \frac{i\omega}{\sqrt{2\pi}(\alpha + i\omega)^2}.$$

Determine the time variation of its output $g(t)$ when its input is the Heaviside step function.

This result is immediate, since differentiating the definition of a Fourier transform (under the integral sign) gives

$$i\frac{d\tilde{f}(\omega)}{d\omega} = \frac{i}{\sqrt{2\pi}}\frac{\partial}{\partial\omega}\left(\int_{-\infty}^{\infty} f(t)\,e^{-i\omega t}\,dt\right) = \frac{-i^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} tf(t)\,e^{-i\omega t}\,dt,$$

i.e. the transform of $tf(t)$.

Since the amplifier's output is the convolution of its input and response function, we will need the Fourier transforms of both to determine that of its output (using the convolution theorem). We already have that of its response function.

The input Heaviside step function $H(t)$ has a Fourier transform

$$\tilde{H}(\omega) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} H(t)\,e^{-i\omega t}\,dt = \frac{1}{\sqrt{2\pi}}\int_{0}^{\infty} e^{-i\omega t}\,dt = \frac{1}{\sqrt{2\pi}}\,\frac{1}{i\omega}.$$

Thus, using the convolution theorem,

$$\begin{aligned} \tilde{g}(\omega) &= \sqrt{2\pi}\,\frac{i\omega}{\sqrt{2\pi}(\alpha + i\omega)^2}\,\frac{1}{\sqrt{2\pi}}\,\frac{1}{i\omega} \\ &= \frac{1}{\sqrt{2\pi}}\,\frac{1}{(\alpha + i\omega)^2} \\ &= \frac{i}{\sqrt{2\pi}}\,\frac{\partial}{\partial\omega}\left(\frac{1}{\alpha + i\omega}\right) \\ &= i\frac{\partial}{\partial\omega}\left\{\mathcal{F}\left[e^{-\alpha t}H(t)\right]\right\} \\ &= \mathcal{F}\left[te^{-\alpha t}H(t)\right], \end{aligned}$$

where we have used the "library" result to recognize the transform of a decaying exponential in the penultimate line and the result proved above in the final step. The output of the amplifier is therefore of the form $g(t) = te^{-\alpha t}$ for $t > 0$ when its input takes the form of the Heaviside step function.

5.17 [This problem can only be attempted if Problem 5.16 of the main text has been studied.]

For some ion–atom scattering processes, the spherically symmetric potential $V(\mathbf{r})$ of the previous problem may be approximated by $V = |\mathbf{r}_1 - \mathbf{r}_2|^{-1} \exp(-\mu|\mathbf{r}_1 - \mathbf{r}_2|)$. Show, using the result of the worked example in Subsection 5.1.9, that the probability that the ion will scatter from, say, $\mathbf{p}_1$ to $\mathbf{p}'_1$ is proportional to $(\mu^2 + k^2)^{-2}$, where $k = |\mathbf{k}|$ and $\mathbf{k}$ is as given in part (c) of the previous problem.

As shown in Problem 5.16, the Fourier transform of a spherically symmetric potential $V(r)$ is given by

$$\widetilde{V}(\mathbf{k}) = \frac{1}{(2\pi)^{3/2}k} \int_0^\infty 4\pi V(r) r \sin kr \, dr.$$

The ion–atom interaction potential in this particular example is $V(r) = r^{-1}\exp(-\mu r)$. As this is spherically symmetric, we may apply the result to it. Substituting for $V(r)$ gives

$$\begin{aligned} M \propto \widetilde{V}(\mathbf{k}) &\propto \frac{1}{k} \int_0^\infty \frac{e^{-\mu r}}{r} r \sin kr \, dr \\ &= \frac{1}{k} \operatorname{Im} \int_0^\infty e^{-\mu r + ikr} \, dr \\ &= \frac{1}{k} \operatorname{Im} \left[\frac{-1}{-\mu + ik} \right] \\ &= \frac{1}{k} \frac{k}{\mu^2 + k^2}. \end{aligned}$$

Since the probability of the ion scattering from $\mathbf{p}_1$ to $\mathbf{p}'_1$ is proportional to the modulus squared of M, the probability is $\propto |M|^2 \propto (\mu^2 + k^2)^{-2}$.

5.19 Find the Laplace transforms of $t^{-1/2}$ and $t^{1/2}$, by setting $x^2 = ts$ in the result

$$\int_0^\infty \exp(-x^2) \, dx = \tfrac{1}{2}\sqrt{\pi}.$$

Setting $x^2 = st$, and hence $2x\, dx = s\, dt$ and $dx = s\, dt/(2\sqrt{st})$, we obtain

$$\begin{aligned} \int_0^\infty e^{-st} \frac{\sqrt{s}}{2} t^{-1/2} \, dt &= \frac{\sqrt{\pi}}{2}, \\ \Rightarrow \quad \mathcal{L}\left[t^{-1/2}\right] \equiv \int_0^\infty t^{-1/2} e^{-st} \, dt &= \sqrt{\frac{\pi}{s}}. \end{aligned}$$

Integrating the LHS of this result by parts yields

$$\left[e^{-st}\,2t^{1/2}\right]_0^\infty - \int_0^\infty (-s)\,e^{-st}\,2t^{1/2}\,dt = \sqrt{\frac{\pi}{s}}.$$

The first term vanishes at both limits, whilst the second is a multiple of the required Laplace transform of $t^{1/2}$. Hence,

$$\mathcal{L}\left[\,t^{1/2}\,\right] \equiv \int_0^\infty e^{-st}\,t^{1/2}\,dt = \frac{1}{2s}\sqrt{\frac{\pi}{s}}.$$

5.21 Use the properties of Laplace transforms to prove the following without evaluating any Laplace integrals explicitly:

(a) $\mathcal{L}\left[\,t^{5/2}\,\right] = \frac{15}{8}\sqrt{\pi}\,s^{-7/2}$;

(b) $\mathcal{L}\left[\,(\sinh at)/t\,\right] = \frac{1}{2}\ln\left[(s+a)/(s-a)\right], \qquad s > |a|$;

(c) $\mathcal{L}\left[\,\sinh at\cos bt\,\right] = a(s^2 - a^2 + b^2)[(s-a)^2 + b^2]^{-1}[(s+a)^2 + b^2]^{-1}$.

(a) We use the general result for Laplace transforms that

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

If we take $n = 2$, then $f(t)$ becomes $t^{1/2}$, for which we found the Laplace transform in Problem 5.19:

$$\begin{aligned}\mathcal{L}\left[\,t^{5/2}\,\right] = \mathcal{L}\left[\,t^2\,t^{1/2}\,\right] &= (-1)^2 \frac{d^2}{ds^2}\left(\frac{\sqrt{\pi}\,s^{-3/2}}{2}\right) \\ &= \frac{\sqrt{\pi}}{2}\left(-\frac{3}{2}\right)\left(-\frac{5}{2}\right)s^{-7/2} = \frac{15\sqrt{\pi}}{8}s^{-7/2}.\end{aligned}$$

(b) Here we apply a second general result for Laplace transforms which states that

$$\mathcal{L}\left[\frac{f(t)}{t}\right] = \int_s^\infty \bar{f}(u)\,du,$$

provided $\lim_{t\to 0}[f(t)/t]$ exists, which it does in this case.

$$\begin{aligned}\mathcal{L}\left[\frac{\sinh(at)}{t}\right] &= \int_s^\infty \frac{a}{u^2 - a^2}\,du, \qquad u > |a|, \\ &= \frac{1}{2}\int_s^\infty \left(\frac{1}{u-a} - \frac{1}{u+a}\right)du \\ &= \frac{1}{2}\ln\left(\frac{s+a}{s-a}\right), \qquad s > |a|.\end{aligned}$$

(c) The translation property of Laplace transforms can be used here to deal with the $\sinh(at)$ factor, as it can be expressed in terms of exponential functions:

$$\begin{aligned}
\mathcal{L}[\sinh(at)\cos(bt)] &= \mathcal{L}\left[\tfrac{1}{2}e^{at}\cos(bt)\right] - \mathcal{L}\left[\tfrac{1}{2}e^{-at}\cos(bt)\right] \\
&= \frac{1}{2}\frac{s-a}{(s-a)^2+b^2} - \frac{1}{2}\frac{s+a}{(s+a)^2+b^2} \\
&= \frac{1}{2}\frac{(s^2-a^2)2a+2ab^2}{[(s-a)^2+b^2][(s+a)^2+b^2]} \\
&= \frac{a(s^2-a^2+b^2)}{[(s-a)^2+b^2][(s+a)^2+b^2]}.
\end{aligned}$$

The result is valid for $s > |a|$.

5.23 This problem is concerned with the limiting behavior of Laplace transforms.

(a) If $f(t) = A + g(t)$, where A is a constant and the indefinite integral of $g(t)$ is bounded as its upper limit tends to ∞, show that

$$\lim_{s\to 0} s\bar{f}(s) = A.$$

(b) For $t > 0$, the function $y(t)$ obeys the differential equation

$$\frac{d^2y}{dt^2} + a\frac{dy}{dt} + by = c\cos^2\omega t,$$

where a, b and c are positive constants. Find $\bar{y}(s)$ and show that $s\bar{y}(s) \to c/2b$ as $s \to 0$. Interpret the result in the t-domain.

(a) From the definition,

$$\begin{aligned}
\bar{f}(s) &= \int_0^\infty [A + g(t)]\, e^{-st}\, dt \\
&= \left[\frac{A\,e^{-st}}{-s}\right]_0^\infty + \lim_{T\to\infty}\int_0^T g(t)\, e^{-st}\, dt, \\
s\bar{f}(s) &= A + s\lim_{T\to\infty}\int_0^T g(t)\, e^{-st}\, dt.
\end{aligned}$$

Now, for $s \geq 0$,

$$\left|\lim_{T\to\infty}\int_0^T g(t)\, e^{-st}\, dt\right| \leq \left|\lim_{T\to\infty}\int_0^T g(t)\, dt\right| < B, \text{ say.}$$

Thus, taking the limit $s \to 0$,

$$\lim_{s\to 0} s\bar{f}(s) = A \pm \lim_{s\to 0} sB = A.$$

(b) We will need

$$\mathcal{L}\left[\cos^2\omega t\right] = \mathcal{L}\left[\tfrac{1}{2}\cos 2\omega + \tfrac{1}{2}\right] = \frac{s}{2(s^2+4\omega^2)} + \frac{1}{2s}.$$

Taking the transform of the differential equation yields

$$-y'(0) - sy(0) + s^2\bar{y} + a[-y(0) + s\bar{y}] + b\bar{y} = c\left[\frac{s}{2(s^2+4\omega^2)} + \frac{1}{2s}\right].$$

This can be rearranged as

$$s\bar{y} = \frac{c\left(\dfrac{s^2}{2(s^2+4\omega^2)} + \dfrac{1}{2}\right) + sy'(0) + asy(0) + s^2y(0)}{s^2 + as + b}.$$

In the limit $s \to 0$, this tends to $(c/2)/b = c/(2b)$, a value independent of that of a and the initial values of y and y'.

The $s = 0$ component of the transform corresponds to long-term values, when a steady state has been reached and rates of change are negligible. With the first two terms of the differential equation ignored, it reduces to $by = c\cos^2\omega t$, and, as the average value of $\cos^2\omega t$ is $\frac{1}{2}$, the solution is the more or less steady value of $y = \frac{1}{2}c/b$.

5.25 The function $f_a(x)$ is defined as unity for $0 < x < a$ and zero otherwise. Find its Laplace transform $\bar{f}_a(s)$ and deduce that the transform of $xf_a(x)$ is

$$\frac{1}{s^2}\left[1 - (1+as)e^{-sa}\right].$$

Write $f_a(x)$ in terms of Heaviside functions and hence obtain an explicit expression for

$$g_a(x) = \int_0^x f_a(y)f_a(x-y)\,dy.$$

Use the expression to write $\bar{g}_a(s)$ in terms of the functions $\bar{f}_a(s)$ and $\bar{f}_{2a}(s)$, and their derivatives, and hence show that $\bar{g}_a(s)$ is equal to the square of $\bar{f}_a(s)$, in accordance with the convolution theorem.

From their definitions,

$$\bar{f}_a(s) = \int_0^a 1\,e^{-sx}\,dx = \frac{1}{s}(1 - e^{-sa}),$$

$$\int_0^a x\,f_a(x)\,e^{-sx}\,dx = -\frac{d\bar{f}_a}{ds} = \frac{1}{s^2}(1 - e^{-sa}) - \frac{a}{s}e^{-sa}$$

$$= \frac{1}{s^2}\left[1 - (1+as)e^{-sa}\right]. \qquad (*)$$

In terms of Heaviside functions,

$$f(x) = H(x) - H(x-a),$$

and so the expression for $g_a(x) = \int_0^x f_a(y)f_a(x-y)\,dy$ is

$$\int_{-\infty}^{\infty} [H(y) - H(y-a)]\,[H(x-y) - H(x-y-a)]\,dy.$$

This can be expanded as the sum of four integrals, each of which contains the common factors $H(y)$ and $H(x-y)$, implying that, in all cases, unless x is positive and greater than y, the integral has zero value. The other factors in the four integrands are generated analogously to the terms of the expansion $(a-b)(c-d) = ac - ad - bc + bd$:

$$\begin{aligned}&\int_{-\infty}^{\infty} H(y)H(x-y)\,dy\\ &\quad-\int_{-\infty}^{\infty} H(y)H(x-y-a)\,dy\\ &\qquad-\int_{-\infty}^{\infty} H(y-a)H(x-y)\,dy\\ &\qquad\quad+\int_{-\infty}^{\infty} H(y-a)H(x-y-a)\,dy.\end{aligned}$$

In all four integrals the integrand is either 0 or 1 and the value of each integral is equal to the length of the y-interval in which the integrand is non-zero.

- The first integral requires $0 < y < x$ and therefore has value x for $x > 0$.
- The second integral requires $0 < y < x - a$ and therefore has value $x - a$ for $x > a$ and 0 for $x < a$.
- The third integral requires $a < y < x$ and therefore has value $x - a$ for $x > a$ and 0 for $x < a$.
- The final integral requires $a < y < x - a$ and therefore has value $x - 2a$ for $x > 2a$ and 0 for $x < 2a$.

Collecting these together:

$$\begin{array}{ll} x < 0 & g_a(x) = 0 - 0 - 0 + 0 = 0,\\ 0 < x < a & g_a(x) = x - 0 - 0 + 0 = x,\\ a < x < 2a & g_a(x) = x - (x-a) - (x-a) + 0 = 2a - x,\\ 2a < x & g_a(x) = x - (x-a) - (x-a) + (x-2a) = 0.\end{array}$$

Consequently, the transform of $g_a(x)$ is given by

$$\begin{aligned}\bar{g}_a(s) &= \int_0^a xe^{-sx}\,dx + \int_a^{2a} (2a-x)e^{-sx}\,dx\\ &= -\int_0^{2a} xe^{-sx}\,dx + 2\int_0^a xe^{-sx}\,dx + 2a\int_a^{2a} e^{-sx}\,dx\\ &= -\frac{1}{s^2}\left[1-(1+2as)e^{-2sa}\right] + \frac{2}{s^2}\left[1-(1+as)e^{-sa}\right]\\ &\quad+\frac{2a}{s}(e^{-sa} - e^{-2sa})\\ &= \frac{1}{s^2}(1 - 2e^{-sa} + e^{-2sa})\\ &= \frac{1}{s^2}(1-e^{-as})^2 = [\bar{f}_a(s)]^2,\end{aligned}$$

which is as expected. In order to adjust the integral limits in the second line, we both added and subtracted

$$\int_0^a (-x)e^{-sx}\,dx.$$

In the third line we used the result $(*)$ twice, once as it stands and once with a replaced by $2a$.

6 Higher-order ordinary differential equations

6.1 A simple harmonic oscillator, of mass m and natural frequency ω_0, experiences an oscillating driving force $f(t) = ma\cos\omega t$. Therefore, its equation of motion is

$$\frac{d^2x}{dt^2} + \omega_0^2 x = a\cos\omega t,$$

where x is its position. Given that at $t = 0$ we have $x = dx/dt = 0$, find the function $x(t)$. Describe the solution if ω is approximately, but not exactly, equal to ω_0.

To find the full solution given the initial conditions, we need the complete general solution made up of a complementary function (CF) and a particular integral (PI). The CF is clearly of the form $A\cos\omega_0 t + B\sin\omega_0 t$ and, in view of the form of the RHS, we try $x(t) = C\cos\omega t + D\sin\omega t$ as a PI. Substituting this gives

$$-\omega^2 C\cos\omega t - \omega^2 D\sin\omega t + \omega_0^2 C\cos\omega t + \omega_0^2 D\sin\omega t = a\cos\omega t.$$

Equating coefficients of the independent functions $\cos\omega t$ and $\sin\omega t$ requires that

$$-\omega^2 C + \omega_0^2 C = a \quad \Rightarrow \quad C = \frac{a}{\omega_0^2 - \omega^2},$$
$$-\omega^2 D + \omega_0^2 D = 0 \quad \Rightarrow \quad D = 0.$$

Thus, the general solution is

$$x(t) = A\cos\omega_0 t + B\sin\omega_0 t + \frac{a}{\omega_0^2 - \omega^2}\cos\omega t.$$

The initial conditions impose the requirements

$$x(0) = 0 \quad \Rightarrow \quad 0 = A + \frac{a}{\omega_0^2 - \omega^2},$$
$$\text{and } \dot{x}(0) = 0 \quad \Rightarrow \quad 0 = \omega_0 B.$$

Incorporating the implications of these into the general solution gives

$$x(t) = \frac{a}{\omega_0^2 - \omega^2}(\cos\omega t - \cos\omega_0 t)$$
$$= \frac{2a\,\sin[\,\tfrac{1}{2}(\omega + \omega_0)t\,]\,\sin[\,\tfrac{1}{2}(\omega_0 - \omega)t\,]}{(\omega_0 + \omega)(\omega_0 - \omega)}.$$

For $\omega_0 - \omega = \epsilon$ with $|\epsilon|t \ll 1$,

$$x(t) \approx \frac{2a\sin\omega_0 t\,\frac{1}{2}\epsilon t}{2\omega_0\,\epsilon} = \frac{at}{2\omega_0}\sin\omega_0 t.$$

Thus, for moderate t, $x(t)$ is a sine wave of linearly increasing amplitude.

Over a long time, $x(t)$ will vary between $\pm 2a/(\omega_0^2 - \omega^2)$ with sizeable intervals between the two extremes, i.e. it will show beats of amplitude $2a/(\omega_0^2 - \omega^2)$.

6.3 The theory of bent beams shows that at any point in the beam the "bending moment" is given by K/ρ, where K is a constant (that depends upon the beam material and cross-sectional shape) and ρ is the radius of curvature at that point. Consider a light beam of length L whose ends, $x = 0$ and $x = L$, are supported at the same vertical height and which has a weight W suspended from its center. Verify that at any point x ($0 \le x \le L/2$ for definiteness) the net magnitude of the bending moment (bending moment = force $\times$ perpendicular distance) due to the weight and support reactions, evaluated on either side of x, is $Wx/2$.

If the beam is only slightly bent, so that $(dy/dx)^2 \ll 1$, where $y = y(x)$ is the downward displacement of the beam at x, show that the beam profile satisfies the approximate equation

$$\frac{d^2y}{dx^2} = -\frac{Wx}{2K}.$$

By integrating this equation twice and using physically imposed conditions on your solution at $x = 0$ and $x = L/2$, show that the downward displacement at the center of the beam is $WL^3/(48K)$.

The upward reaction of the support at each end of the beam is $\frac{1}{2}W$.

At the position x the moment on the left is due to

(i) the support at $x = 0$ providing a clockwise moment of $\frac{1}{2}Wx$.

The moment on the right is due to

(ii) the support at $x = L$ providing an anticlockwise moment of $\frac{1}{2}W(L - x)$;

(iii) the weight at $x = \frac{1}{2}L$ providing a clockwise moment of $W(\frac{1}{2}L - x)$.

The net clockwise moment on the right is therefore $W(\frac{1}{2}L - x) - \frac{1}{2}W(L - x) = -\frac{1}{2}Wx$, i.e. equal in magnitude, but opposite in sign, to that on the left.

The radius of curvature of the beam is $\rho = [\,1 + (-y')^2\,]^{3/2}/(-y'')$, but if $|y'| \ll 1$ this simplifies to $-1/y''$ and the equation of the beam profile satisfies

$$\frac{Wx}{2} = M = \frac{K}{\rho} = -K\frac{d^2y}{dx^2}.$$

We now need to integrate this, taking into account the boundary conditions $y(0) = 0$ and, on symmetry grounds, $y'(\frac{1}{2}L) = 0$:

$$\begin{aligned}
y' &= -\frac{Wx^2}{4K} + A, \text{ with } y'(\tfrac{1}{2}L) = 0 \quad \Rightarrow \quad A = \frac{WL^2}{16K},\\
y' &= \frac{W}{4K}\left(\frac{L^2}{4} - x^2\right),\\
y &= \frac{W}{4K}\left(\frac{L^2x}{4} - \frac{x^3}{3} + B\right), \text{ with } y(0) = 0 \quad \Rightarrow \quad B = 0.
\end{aligned}$$

The center is lowered by

$$y(\tfrac{1}{2}L) = \frac{W}{4K}\left(\frac{L^2}{4}\frac{L}{2} - \frac{1}{3}\frac{L^3}{8}\right) = \frac{WL^3}{48K}.$$

Note that the derived analytic form for $y(x)$ is not applicable in the range $\frac{1}{2}L \leq x \leq L$; the beam profile is symmetrical about $x = \frac{1}{2}L$, but the expression $\frac{1}{4}L^2 x - \frac{1}{3}x^3$ is not invariant under the substitution $x \to L - x$.

6.5 The function $f(t)$ satisfies the differential equation

$$\frac{d^2 f}{dt^2} + 8\frac{df}{dt} + 12f = 12e^{-4t}.$$

For the following sets of boundary conditions determine whether it has solutions, and, if so, find them:

(a) $f(0) = 0, \quad f'(0) = 0, \quad f(\ln\sqrt{2}) = 0;$
(b) $f(0) = 0, \quad f'(0) = -2, \quad f(\ln\sqrt{2}) = 0.$

Three boundary conditions have been given, and, as this is a second-order linear equation for which only two independent conditions are needed, they may be inconsistent. The plan is to solve it using two of the conditions and then test whether the third one is compatible.

The auxiliary equation for obtaining the CF is

$$\begin{aligned} m^2 + 8m + 12 = 0 \quad &\Rightarrow \quad m = -2 \text{ or } m = -6 \\ &\Rightarrow \quad f(t) = Ae^{-6t} + Be^{-2t}. \end{aligned}$$

Since the form of the RHS, Ce^{-4t}, is not included in the CF, we can try it as the particular integral:

$$16C - 32C + 12C = 12 \quad \Rightarrow \quad C = -3.$$

The general solution is therefore

$$f(t) = Ae^{-6t} + Be^{-2t} - 3e^{-4t}.$$

(a) For boundary conditions $f(0) = 0, \quad f'(0) = 0, \quad f(\ln\sqrt{2}) = 0$:

$$\begin{aligned} f(0) = 0 \quad &\Rightarrow \quad A + B - 3 = 0, \\ f'(0) = 0 \quad &\Rightarrow \quad -6A - 2B + 12 = 0, \\ &\Rightarrow \quad A = \tfrac{3}{2}, \quad B = \tfrac{3}{2}. \\ \text{Hence, } f(t) \quad &= \quad \tfrac{3}{2}e^{-6t} + \tfrac{3}{2}e^{-2t} - 3e^{-4t}. \end{aligned}$$

Recalling that $e^{-(\ln\sqrt{2})} = 1/\sqrt{2}$, we evaluate

$$f(\ln\sqrt{2}) = \frac{3}{2}\frac{1}{8} + \frac{3}{2}\frac{1}{2} - 3\frac{1}{4} = \frac{3}{16} \neq 0.$$

Thus the boundary conditions are inconsistent and there is no solution.

(b) For boundary conditions $f(0) = 0, \quad f'(0) = -2, \quad f(\ln\sqrt{2}) = 0$, we proceed as before:

$$\begin{aligned} f(0) = 0 \quad &\Rightarrow \quad A + B - 3 = 0, \\ f'(0) = 0 \quad &\Rightarrow \quad -6A - 2B + 12 = -2, \\ &\Rightarrow \quad A = 2, \quad B = 1. \\ \text{Hence, } f(t) \quad &= \quad 2e^{-6t} + e^{-2t} - 3e^{-4t}. \end{aligned}$$

We again evaluate

$$f(\ln\sqrt{2}) = 2\,\frac{1}{8} + \frac{1}{2} - 3\,\frac{1}{4} = 0.$$

This time the boundary conditions are consistent and there is a unique solution as given above.

6.7 A solution of the differential equation

$$\frac{d^2y}{dx^2} + 2\frac{dy}{dx} + y = 4e^{-x}$$

takes the value 1 when $x = 0$ and the value e^{-1} when $x = 1$. What is its value when $x = 2$?

The auxiliary equation, $m^2 + 2m + 1 = 0$, has repeated roots $m = -1$, and so the general CF has the special form $y(x) = (A + Bx)e^{-x}$.

Turning to the PI, we note that the form of the RHS of the original equation is contained in the CF, and (to make matters worse) so is x times the RHS. We therefore need to take x^2 times the RHS as a trial PI:

$$y(x) = Cx^2e^{-x}, \quad y' = C(2x - x^2)e^{-x}, \quad y'' = C(2 - 4x + x^2)e^{-x}.$$

Substituting these into the original equation shows that

$$2Ce^{-x} = 4e^{-x} \quad \Rightarrow \quad C = 2$$

and that the full general solution is given by

$$y(x) = (A + Bx)e^{-x} + 2x^2e^{-x}.$$

We now determine the unknown constants using the information given about the solution. Since $y(0) = 1$, $A = 1$. Further, $y(1) = e^{-1}$ requires

$$e^{-1} = (1 + B)e^{-1} + 2e^{-1} \quad \Rightarrow \quad B = -2.$$

Finally, we conclude that $y(x) = (1 - 2x + 2x^2)e^{-x}$ and, therefore, that $y(2) = 5e^{-2}$.

6.9 Find the general solutions of

(a) $\dfrac{d^3y}{dx^3} - 12\dfrac{dy}{dx} + 16y = 32x - 8$,

(b) $\dfrac{d}{dx}\left(\dfrac{1}{y}\dfrac{dy}{dx}\right) + (2a\coth 2ax)\left(\dfrac{1}{y}\dfrac{dy}{dx}\right) = 2a^2$,

where a is a constant.

(a) As this is a third-order equation, we expect three terms in the CF.

Since it is linear with constant coefficients, we can make use of the auxiliary equation, which is

$$m^3 - 12m + 16 = 0.$$

By inspection, $m = 2$ is one root; the other two can be found by factorization:

$$m^3 - 12m + 16 = (m-2)(m^2 + 2m - 8) = (m-2)(m+4)(m-2) = 0.$$

Thus we have one repeated root ($m = 2$) and one other ($m = -4$) leading to a CF of the form

$$y(x) = (A + Bx)e^{2x} + Ce^{-4x}.$$

As the RHS contains no exponentials, we try $y(x) = Dx + E$ for the PI. We then need $16D = 32$ and $-12D + 16E = -8$, giving $D = 2$ and $E = 1$.

The general solution is therefore

$$y(x) = (A + Bx)e^{2x} + Ce^{-4x} + 2x + 1.$$

(b) The equation is already arranged in the form

$$\frac{dg(y)}{dx} + h(x)g(y) = j(x)$$

and so needs only an integrating factor to allow the first integration step to be made. For this equation the IF is

$$\exp\left\{\int 2a \coth 2ax\, dx\right\} = \exp(\ln \sinh 2ax) = \sinh 2ax.$$

After multiplication through by this factor, the equation can be written

$$\begin{aligned}\sinh 2ax \frac{d}{dx}\left(\frac{1}{y}\frac{dy}{dx}\right) + (2a \cosh 2ax)\left(\frac{1}{y}\frac{dy}{dx}\right) &= 2a^2 \sinh 2ax,\\ \frac{d}{dx}\left(\sinh 2ax\, \frac{1}{y}\frac{dy}{dx}\right) &= 2a^2 \sinh 2ax.\end{aligned}$$

Integrating this gives

$$\begin{aligned}\sinh 2ax\, \frac{1}{y}\frac{dy}{dx} &= \frac{2a^2}{2a}\cosh 2ax + A,\\ \Rightarrow \quad \frac{1}{y}\frac{dy}{dx} &= a \coth 2ax + \frac{A}{\sinh 2ax}.\end{aligned}$$

Integrating again,

$$\begin{aligned}\ln y &= \frac{1}{2}\ln(\sinh 2ax) + \int \frac{A}{\sinh 2ax}\, dx + B\\ &= \frac{1}{2}\ln(\sinh 2ax) + \frac{A}{2a}\ln(|\tanh ax|) + B,\\ \Rightarrow \quad y &= C(\sinh 2ax)^{1/2}\,(|\tanh ax|)^D.\end{aligned}$$

The indefinite integral of $(\sinh 2ax)^{-1}$ appearing in the fourth line can be verified by differentiating $y = \ln|\tanh ax|$ in the form $y = \frac{1}{2}\ln(\tanh^2 ax)$ and recalling that

$$\cosh ax \sinh ax = \frac{1}{2}\sinh 2ax.$$

6.11 The quantities $x(t)$, $y(t)$ satisfy the simultaneous equations

$$\begin{aligned}\ddot{x} + 2n\dot{x} + n^2 x &= 0,\\ \ddot{y} + 2n\dot{y} + n^2 y &= \mu\dot{x},\end{aligned}$$

where $x(0) = y(0) = \dot{y}(0) = 0$ and $\dot{x}(0) = \lambda$. Show that

$$y(t) = \tfrac{1}{2}\mu\lambda t^2\left(1 - \tfrac{1}{3}nt\right)\exp(-nt).$$

For these two coupled equations, in which an "output" from the first acts as the "driving input" for the second, we take Laplace transforms and incorporate the boundary conditions:

$$\begin{aligned}(s^2\bar{x} - 0 - \lambda) + 2n(s\bar{x} - 0) + n^2\bar{x} &= 0,\\ (s^2\bar{y} - 0 - 0) + 2n(s\bar{y} - 0) + n^2\bar{y} &= \mu(s\bar{x} - 0).\end{aligned}$$

From the first transformed equation,

$$\bar{x} = \frac{\lambda}{s^2 + 2ns + n^2}.$$

Substituting this into the second transformed equation gives

$$\begin{aligned}\bar{y} &= \frac{\mu s\bar{x}}{(s+n)^2} = \frac{\mu\lambda s}{(s+n)^2(s+n)^2}\\ &= \frac{\mu\lambda}{(s+n)^3} - \frac{\mu\lambda n}{(s+n)^4},\\ \Rightarrow \quad y(t) &= \mu\lambda\left(\frac{t^2}{2!}e^{-nt} - \frac{nt^3}{3!}e^{-nt}\right), \text{ from the look-up table,}\\ &= \frac{1}{2}\mu\lambda t^2\left(1 - \frac{nt}{3}\right)e^{-nt},\end{aligned}$$

i.e. as stated in the question.

6.13 Two unstable isotopes A and B and a stable isotope C have the following decay rates per atom present: $A \to B$, 3 s^{-1}; $A \to C$, 1 s^{-1}; $B \to C$, 2 s^{-1}. Initially a quantity x_0 of A is present but there are no atoms of the other two types. Using Laplace transforms, find the amount of C present at a later time t.

Using the name symbol to represent the corresponding number of atoms and taking Laplace transforms, we have

$$\begin{aligned}
\frac{dA}{dt} = -(3+1)A \quad &\Rightarrow \quad s\bar{A} - x_0 = -4\bar{A} \\
&\Rightarrow \quad \bar{A} = \frac{x_0}{s+4}, \\
\frac{dB}{dt} = 3A - 2B \quad &\Rightarrow \quad s\bar{B} = 3\bar{A} - 2\bar{B} \\
&\Rightarrow \quad \bar{B} = \frac{3x_0}{(s+2)(s+4)}, \\
\frac{dC}{dt} = A + 2B \quad &\Rightarrow \quad s\bar{C} = \bar{A} + 2\bar{B} \\
&\Rightarrow \quad \bar{C} = \frac{x_0(s+2) + 6x_0}{s(s+2)(s+4)}.
\end{aligned}$$

Using the "cover-up" method for finding the coefficients of a partial fraction expansion without repeated factors, e.g. the coefficient of $(s+2)^{-1}$ is $[\,(-2+8)x_0\,]/[\,(-2)(-2+4)\,] = -6x_0/4$, we have

$$\begin{aligned}
\bar{C} &= \frac{x_0(s+8)}{s(s+2)(s+4)} = \frac{x_0}{s} - \frac{6x_0}{4(s+2)} + \frac{4x_0}{8(s+4)} \\
\Rightarrow \quad C(t) &= x_0\left(1 - \tfrac{3}{2}e^{-2t} + \tfrac{1}{2}e^{-4t}\right).
\end{aligned}$$

This is the required expression.

6.15 The "golden mean", which is said to describe the most aesthetically pleasing proportions for the sides of a rectangle (e.g. the ideal picture frame), is given by the limiting value of the ratio of successive terms of the Fibonacci series u_n, which is generated by

$$u_{n+2} = u_{n+1} + u_n,$$

with $u_0 = 0$ and $u_1 = 1$. Find an expression for the general term of the series and verify that the golden mean is equal to the larger root of the recurrence relation's characteristic equation.

The recurrence relation is second order and its characteristic equation, obtained by setting $u_n = A\lambda^n$, is

$$\lambda^2 - \lambda - 1 = 0 \quad \Rightarrow \quad \lambda = \tfrac{1}{2}(1 \pm \sqrt{5}).$$

The general solution is therefore

$$u_n = A\left(\frac{1+\sqrt{5}}{2}\right)^n + B\left(\frac{1-\sqrt{5}}{2}\right)^n.$$

The initial values (boundary conditions) determine A and B:

$$u_0 = 0 \quad \Rightarrow \quad B = -A,$$
$$u_1 = 1 \quad \Rightarrow \quad A\left(\frac{1+\sqrt{5}}{2} - \frac{1-\sqrt{5}}{2}\right) = 1 \quad \Rightarrow \quad A = \frac{1}{\sqrt{5}}.$$
$$\text{Hence, } u_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right].$$

If we write $(1-\sqrt{5})/(1+\sqrt{5}) = r < 1$, the ratio of successive terms in the series is

$$\begin{aligned}\frac{u_{n+1}}{u_n} &= \frac{\frac{1}{2}[\,(1+\sqrt{5})^{n+1} - (1-\sqrt{5})^{n+1}\,]}{(1+\sqrt{5})^n - (1-\sqrt{5})^n} \\ &= \frac{\frac{1}{2}[\,1+\sqrt{5} - (1-\sqrt{5})r^n\,]}{1-r^n} \\ &\to \frac{1+\sqrt{5}}{2} \text{ as } n \to \infty;\end{aligned}$$

i.e. the limiting ratio is the same as the larger value of λ.

This result is a particular example of the more general one that the ratio of successive terms in a series generated by a recurrence relation tends to the largest (in absolute magnitude) of the roots of the characteristic equation. Here there are only two roots, but for an Nth-order relation there will be N roots.

6.17 The first few terms of a series u_n, starting with u_0, are $1, 2, 2, 1, 6, -3$. The series is generated by a recurrence relation of the form

$$u_n = Pu_{n-2} + Qu_{n-4},$$

where P and Q are constants. Find an expression for the general term of the series and show that, in fact, the series consists of two interleaved series given by

$$u_{2m} = \tfrac{2}{3} + \tfrac{1}{3}4^m,$$
$$u_{2m+1} = \tfrac{7}{3} - \tfrac{1}{3}4^m,$$

for $m = 0, 1, 2, \ldots$.

We first find P and Q using

$$\begin{aligned} n &= 4 \qquad 6 = 2P + Q, \\ n &= 5 \quad -3 = P + 2Q, \quad \Rightarrow \quad Q = -4 \text{ and } P = 5.\end{aligned}$$

The recurrence relation is thus

$$u_n = 5u_{n-2} - 4u_{n-4}.$$

To solve this we try $u_n = A + B\lambda^n$ for arbitrary constants A and B and obtain

$$\begin{aligned} A + B\lambda^n &= 5A + 5B\lambda^{n-2} - 4A - 4B\lambda^{n-4}, \\ \Rightarrow \quad 0 &= \lambda^4 - 5\lambda^2 + 4 \\ &= (\lambda^2 - 1)(\lambda^2 - 4) \quad \Rightarrow \quad \lambda = \pm 1, \pm 2. \end{aligned}$$

The general solution is $u_n = A + B(-1)^n + C2^n + D(-2)^n$.

We now need to solve the simultaneous equations for A, B, C and D provided by the values of $u_0, \ldots, u_3$:

$$\begin{aligned} 1 &= A + B + C + D, \\ 2 &= A - B + 2C - 2D, \\ 2 &= A + B + 4C + 4D, \\ 1 &= A - B + 8C - 8D. \end{aligned}$$

These have the straightforward solution

$$A = \frac{3}{2}, \qquad B = -\frac{5}{6}, \qquad C = \frac{1}{12}, \qquad D = \frac{1}{4},$$

and so

$$u_n = \frac{3}{2} - \frac{5}{6}(-1)^n + \frac{1}{12}2^n + \frac{1}{4}(-2)^n.$$

When n is even and equal to $2m$,

$$u_{2m} = \frac{3}{2} - \frac{5}{6} + \frac{4^m}{12} + \frac{4^m}{4} = \frac{2}{3} + \frac{4^m}{3}.$$

When n is odd and equal to $2m + 1$,

$$u_{2m+1} = \frac{3}{2} + \frac{5}{6} + \frac{4^m}{6} - \frac{4^m}{2} = \frac{7}{3} - \frac{4^m}{3}.$$

In passing, we note that the fact that both P and Q, and all of the given values $u_0, \ldots, u_4$, are integers, and hence that all terms in the series are integers, provides an indirect proof that $4^m + 2$ is divisible by 3 (without remainder) for all non-negative integers m. This can be more easily proved by induction, as the reader may like to verify.

6.19 Find the general expression for the u_n satisfying

$$u_{n+1} = 2u_{n-2} - u_n$$

with $u_0 = u_1 = 0$ and $u_2 = 1$, and show that they can be written in the form

$$u_n = \frac{1}{5} - \frac{2^{n/2}}{\sqrt{5}} \cos\left(\frac{3\pi n}{4} - \phi\right),$$

where $\tan\phi = 2$.

The characteristic equation (which will be a cubic since the recurrence relation is third order) and its solution are given by

$$\lambda^{n+1} = 2\lambda^{n-2} - \lambda^n,$$
$$\lambda^3 + \lambda^2 - 2 = 0,$$
$$(\lambda - 1)(\lambda^2 + 2\lambda + 2) = 0 \quad \Rightarrow \quad \lambda = 1 \text{ or } \lambda = -1 \pm i.$$

Thus the general solution of the recurrence relation, which has the generic form $A\lambda_1^n + B\lambda_2^n + C\lambda_3^n$, is

$$\begin{aligned} u_n &= A + B(-1+i)^n + C(-1-i)^n \\ &= A + B\,2^{n/2}e^{i3\pi n/4} + C\,2^{n/2}e^{i5\pi n/4}. \end{aligned}$$

To determine A, B and C we use

$$\begin{aligned} u_0 &= 0, \qquad & 0 &= A + B + C, \\ u_1 &= 0, & 0 &= A + B\,2^{1/2}e^{i3\pi/4} + C\,2^{1/2}e^{i5\pi/4} \\ & & &= A + B(-1+i) + C(-1-i), \\ u_2 &= 1, & 1 &= A + B\,2e^{i6\pi/4} + C\,2e^{i10\pi/4} = A + 2B(-i) + 2C(i). \end{aligned}$$

Adding twice each of the first two equations to the last one gives $5A = 1$. Substituting this into the first and last equations then leads to

$$B + C = -\frac{1}{5} \qquad \text{and} \qquad -B + C = \frac{2}{5i},$$

from which it follows that

$$\begin{aligned} B &= \frac{-1+2i}{10} = \frac{\sqrt{5}}{10}e^{i(\pi-\phi)} \\ \text{and} \quad C &= \frac{-1-2i}{10} = \frac{\sqrt{5}}{10}e^{i(\pi+\phi)}, \end{aligned}$$

where $\tan\phi = 2/1 = 2$.

Thus, collecting these results together, we have

$$\begin{aligned} u_n &= \frac{1}{5} + \frac{2^{n/2}\sqrt{5}}{10}(e^{i3\pi n/4}e^{i(\pi-\phi)} + e^{i5\pi n/4}e^{i(\pi+\phi)}) \\ &= \frac{1}{5} - \frac{2^{n/2}\sqrt{5}}{10}(e^{i3\pi n/4}e^{-i\phi} + e^{-i3\pi n/4}e^{i\phi}) \\ &= \frac{1}{5} - \frac{2^{n/2}\sqrt{5}}{10}\left[2\cos\left(\frac{3\pi n}{4} - \phi\right)\right] \\ &= \frac{1}{5} - \frac{2^{n/2}}{\sqrt{5}}\cos\left(\frac{3\pi n}{4} - \phi\right), \end{aligned}$$

i.e. the form of solution given in the question.

6.21 Find the general solution of

$$x^2\frac{d^2y}{dx^2} - x\frac{dy}{dx} + y = x,$$

given that $y(1) = 1$ and $y(e) = 2e$.

This is Euler's equation and can be solved either by a change of variables, $x = e^t$, or by trying $y = x^\lambda$; we will adopt the second approach. Doing so in the homogeneous equation (RHS set to zero) gives

$$x^2\,\lambda(\lambda-1)x^{\lambda-2} - x\,\lambda x^{\lambda-1} + x^\lambda = 0.$$

The CF is therefore obtained when λ satisfies

$$\lambda(\lambda-1) - \lambda + 1 = 0 \quad\Rightarrow\quad (\lambda-1)^2 = 0 \quad\Rightarrow\quad \lambda = 1 \text{ (repeated)}.$$

Thus, one solution is $y = x$; the other linearly independent solution implied by the repeated root is $x\ln x$ (see a textbook if this is not known).

There is now a further complication as the RHS of the original equation (x) is contained in the CF. We therefore need an extra factor of $\ln x$ in the trial PI, beyond those already in the CF. (This corresponds to the extra power of t needed in the PI if the transformation to a linear equation with constant coefficients is made via the $x = e^t$ change of variable.) As a consequence, the PI to be tried is $y = Cx(\ln x)^2$:

$$x^2\left[2C\frac{\ln x}{x} + \frac{2C}{x}\right] - x\left[Cx\,\frac{2\ln x}{x} + C(\ln x)^2\right] + Cx(\ln x)^2 = x.$$

This implies that $C = \frac{1}{2}$ and gives the general solution as

$$y(x) = Ax + Bx\ln x + \tfrac{1}{2}x(\ln x)^2.$$

It remains only to determine the unknown constants A and B; this is done using the two given values of $y(x)$. The boundary condition $y(1) = 1$ requires that $A = 1$, and $y(e) = 2e$ implies that $B = \frac{1}{2}$; the solution is now completely determined as

$$y(x) = x + \tfrac{1}{2}x\ln x(1 + \ln x).$$

6.23 Prove that the general solution of

$$(x-2)\frac{d^2y}{dx^2} + 3\frac{dy}{dx} + \frac{4y}{x^2} = 0$$

is given by

$$y(x) = \frac{1}{(x-2)^2}\left[k\left(\frac{2}{3x} - \frac{1}{2}\right) + cx^2\right].$$

This equation is not of any plausible standard form, and the only solution method is to try to make it into an exact equation. If this is possible the order of the equation will be reduced by one.

We first multiply through by x^2 and then note that the resulting factor $3x^2$ in the second term can be written as $[\,x^2(x-2)\,]' + 4x$, i.e. as the derivative of the function multiplying y'' together with another simple function. This latter can be combined with the undifferentiated term and allow the whole equation to be written as an exact equation:

$$\begin{aligned}
\frac{d}{dx}\left[x^2(x-2)\frac{dy}{dx}\right] + 4x\frac{dy}{dx} + 4y &= 0,\\
\frac{d}{dx}\left[x^2(x-2)\frac{dy}{dx}\right] + \frac{d(4xy)}{dx} &= 0,\\
\Rightarrow \quad x^2(x-2)\frac{dy}{dx} + 4xy &= k.
\end{aligned}$$

Either by inspection or by use of the standard formula, the IF is $(x-2)/x^4$ and leads to

$$\begin{aligned}
\frac{d}{dx}\left[\frac{(x-2)^2}{x^2}\,y\right] &= \frac{k(x-2)}{x^4},\\
\Rightarrow \quad \frac{(x-2)^2}{x^2}\,y &= k\left(-\frac{1}{2x^2} + \frac{2}{3x^3}\right) + c,\\
\Rightarrow \quad y &= \frac{1}{(x-2)^2}\left(-\frac{k}{2} + \frac{2k}{3x} + cx^2\right).
\end{aligned}$$

6.25 Find the Green's function that satisfies

$$\frac{d^2G(x,\xi)}{dx^2} - G(x,\xi) = \delta(x-\xi) \qquad \text{with} \qquad G(0,\xi) = G(1,\xi) = 0.$$

It is clear from inspection that the CF has solutions of the form $e^{\pm x}$. The other pair of solutions that may suggest themselves are $\sinh x$ and $\cosh x$, but these are merely independent linear combinations of the same two functions.

As both boundary conditions are given at finite values of x (rather than at $x \to \pm\infty$) and both are of the form $y(x) = 0$, it is more convenient to work with those particular linear combinations of e^x and e^{-x} that vanish at the boundary points. The only common linear combination of these two functions that vanishes at a finite value of x is a sinh function. To construct one that vanishes at $x = x_0$ the argument of the sinh function must be made to be $x - x_0$. For the present case the appropriate combinations are

$$\sinh x = \frac{1}{2}(e^x - e^{-x}) \quad \text{and} \quad \sinh(1-x) = \left(\frac{e}{2}\right)e^{-x} - \left(\frac{1}{2e}\right)e^x.$$

Thus, with $0 \le \xi \le 1$, we take

$$G(x,\xi) = \begin{cases} A(\xi)\sinh x & x < \xi,\\ B(\xi)\sinh(1-x) & x > \xi. \end{cases}$$

The continuity requirement on $G(x, \xi)$ at $x = \xi$ and the unit discontinuity requirement on its derivative at the same point give

$$\begin{aligned} A \sinh \xi - B \sinh(1 - \xi) &= 0 \\ \text{and} \qquad - B \cosh(1 - \xi) - A \cosh \xi &= 1, \end{aligned}$$

leading to

$$\begin{aligned} A \sinh \xi \cosh(1 - \xi) + A \cosh \xi \sinh(1 - \xi) &= -\sinh(1 - \xi), \\ A[\sinh(\xi + 1 - \xi)] &= -\sinh(1 - \xi). \end{aligned}$$

Hence,

$$A = -\frac{\sinh(1 - \xi)}{\sinh 1} \qquad \text{and} \qquad B = -\frac{\sinh \xi}{\sinh 1},$$

giving the full Green's function as

$$G(x, \xi) = \begin{cases} -\dfrac{\sinh(1 - \xi)}{\sinh 1} \sinh x & x < \xi, \\ -\dfrac{\sinh \xi}{\sinh 1} \sinh(1 - x) & x > \xi. \end{cases}$$

6.27 Show generally that if $y_1(x)$ and $y_2(x)$ are linearly independent solutions of

$$\frac{d^2 y}{dx^2} + p(x)\frac{dy}{dx} + q(x)y = 0,$$

with $y_1(0) = 0$ and $y_2(1) = 0$, then the Green's function $G(x, \xi)$ for the interval $0 \le x, \xi \le 1$ and with $G(0, \xi) = G(1, \xi) = 0$ can be written in the form

$$G(x, \xi) = \begin{cases} y_1(x)y_2(\xi)/W(\xi) & 0 < x < \xi, \\ y_2(x)y_1(\xi)/W(\xi) & \xi < x < 1, \end{cases}$$

where $W(x) = W[y_1(x), y_2(x)]$ is the Wronskian of $y_1(x)$ and $y_2(x)$.

As usual, we start by writing the general solution as a weighted sum of the linearly independent solutions, whilst leaving the possibility that the weights may be different for different x-ranges:

$$G(x, \xi) = \begin{cases} A(\xi)y_1(x) + B(\xi)y_2(x) & 0 < x < \xi, \\ C(\xi)y_1(x) + D(\xi)y_2(x) & \xi < x < 1. \end{cases}$$

Imposing the boundary conditions and using $y_1(0) = y_2(1) = 0$,

$$\begin{aligned} 0 = G(0, \xi) = A(\xi)y_1(0) + B(\xi)y_2(0) \quad &\Rightarrow \quad B(\xi) = 0, \\ 0 = G(1, \xi) = C(\xi)y_1(1) + D(\xi)y_2(1) \quad &\Rightarrow \quad C(\xi) = 0. \end{aligned}$$

The continuity requirement on $G(x, \xi)$ at $x = \xi$ and the unit discontinuity requirement on its derivative at the same point give

$$A(\xi)y_1(\xi) - D(\xi)y_2(\xi) = 0,$$
$$A(\xi)y_1'(\xi) - D(\xi)y_2'(\xi) = -1,$$

leading to

$$A(\xi)[\,y_1y_2' - y_2y_1'\,] = y_2 \quad \Rightarrow \quad A(\xi) = \frac{y_2(\xi)}{W(\xi)},$$
$$D(\xi) = \frac{y_1(\xi)}{y_2(\xi)}\,A(\xi) = \frac{y_1(\xi)}{W(\xi)}.$$

Thus,

$$G(x, \xi) = \begin{cases} y_1(x)y_2(\xi)/W(\xi) & 0 < x < \xi, \\ y_2(x)y_1(\xi)/W(\xi) & \xi < x < 1. \end{cases}$$

This result is perfectly general for linear second-order equations of the type stated and can be a quick way to find the corresponding Green's function, provided the solutions that vanish at the end-points can be identified easily. Problem 6.25 is a particular example of this general result.

6.29 The equation of motion for a driven damped harmonic oscillator can be written

$$\ddot{x} + 2\dot{x} + (1 + \kappa^2)x = f(t),$$

with $\kappa \neq 0$. If it starts from rest with $x(0) = 0$ and $\dot{x}(0) = 0$, find the corresponding Green's function $G(t, \tau)$ and verify that it can be written as a function of $t - \tau$ only. Find the explicit solution when the driving force is the unit step function, i.e. $f(t) = H(t)$. Confirm your solution by taking the Laplace transforms of both it and the original equation.

The auxiliary equation is

$$m^2 + 2m + (1 + \kappa^2) = 0 \quad \Rightarrow \quad m = -1 \pm i\kappa,$$

and the CF is $x(t) = Ae^{-t}\cos\kappa t + Be^{-t}\sin\kappa t$.

Let

$$G(t, \tau) = \begin{cases} A(\tau)e^{-t}\cos\kappa t + B(\tau)e^{-t}\sin\kappa t & 0 < t < \tau, \\ C(\tau)e^{-t}\cos\kappa t + D(\tau)e^{-t}\sin\kappa t & t > \tau. \end{cases}$$

The boundary condition $x(0) = 0$ implies that $A = 0$, and

$$\dot{x}(0) = 0 \quad \Rightarrow \quad B(-e^{-t}\sin\kappa t + \kappa e^{-t}\cos\kappa t) = 0 \quad \Rightarrow \quad B = 0.$$

Thus $G(t, \tau) = 0$ for $t < \tau$.

The continuity of G at $t = \tau$ gives

$$Ce^{-\tau}\cos\kappa\tau + De^{-\tau}\sin\kappa\tau = 0 \quad \Rightarrow \quad D = -\frac{C\cos\kappa\tau}{\sin\kappa\tau}.$$

The unit discontinuity in the derivative of G at $t = \tau$ requires (using $s = \sin\kappa\tau$ and $c = \cos\kappa\tau$ as shorthand)

$$\begin{aligned} Ce^{-\tau}(-c - \kappa s) + De^{-\tau}(-s + \kappa c) - 0 = 1, \\ C\left[-c - \kappa s - \frac{c}{s}(-s + \kappa c)\right] = e^{\tau}, \\ C(-sc - \kappa s^2 + cs - \kappa c^2) = se^{\tau}, \end{aligned}$$

giving

$$C = -\frac{e^{\tau}\sin\kappa\tau}{\kappa} \quad \text{and} \quad D = \frac{e^{\tau}\cos\kappa\tau}{\kappa}.$$

Thus, for $t > \tau$,

$$\begin{aligned} G(t,\tau) &= \frac{e^{\tau}}{\kappa}(-\sin\kappa\tau\cos\kappa t + \cos\kappa\tau\sin\kappa t)e^{-t} \\ &= \frac{e^{-(t-\tau)}}{\kappa}\sin\kappa(t-\tau). \end{aligned}$$

This form verifies that the Green's function is a function only of the difference $t - \tau$ and not of t and τ separately.

The explicit solution to the given equation when $f(t) = H(t)$ is thus

$$\begin{aligned} x(t) &= \int_0^{\infty} G(t,\tau)f(\tau)\,d\tau \\ &= \int_0^{t} G(t,\tau)H(\tau)\,d\tau, \text{ since } G(t,\tau) = 0 \text{ for } \tau > t, \\ &= \frac{1}{\kappa}\int_0^{t} e^{-(t-\tau)}\sin\kappa(t-\tau)\,d\tau \\ &= \frac{e^{-t}}{\kappa}\text{Im}\int_0^{t} e^{\tau+i\kappa(t-\tau)}\,d\tau \\ &= \frac{e^{-t}}{\kappa}\text{Im}\left[\frac{e^{i\kappa t}e^{\tau-i\kappa\tau}}{1-i\kappa}\right]_{\tau=0}^{\tau=t} \\ &= \frac{e^{-t}}{\kappa}\text{Im}\left[\frac{e^{t}-e^{i\kappa t}}{1-i\kappa}\right]. \end{aligned}$$

Now multiplying both numerator and denominator by $1 + i\kappa$ to make the latter real gives

$$\begin{aligned} x(t) &= \frac{e^{-t}}{\kappa(1+\kappa^2)}\text{Im}\,[\,(e^{t} - e^{i\kappa t})(1+i\kappa)\,] \\ &= \frac{e^{-t}}{\kappa(1+\kappa^2)}[\,\kappa(e^{t} - \cos\kappa t) - \sin\kappa t\,] \\ &= \frac{1}{1+\kappa^2}\left(1 - e^{-t}\cos\kappa t - \frac{1}{\kappa}e^{-t}\sin\kappa t\right). \end{aligned}$$

The Laplace transform of this solution is given by

$$\begin{aligned}\bar{x} &= \frac{1}{1+\kappa^2}\left(\frac{1}{s} - \frac{s+1}{(s+1)^2+\kappa^2} - \frac{1}{\kappa}\frac{\kappa}{(s+1)^2+\kappa^2}\right)\\ &= \frac{(s+1)^2+\kappa^2 - s(s+1) - s}{(1+\kappa^2)s[\,(s+1)^2+\kappa^2\,]}\\ &= \frac{1}{s[\,(s+1)^2+\kappa^2\,]}.\end{aligned}$$

The Laplace transform of the original equation with the given initial conditions reads

$$[\,s^2\bar{x} - 0s - 0\,] + 2[\,s\bar{x} - 0\,] + (1+\kappa^2)\bar{x} = \frac{1}{s},$$

again showing that

$$\bar{x} = \frac{1}{s[\,s^2+2s+1+\kappa^2\,]} = \frac{1}{s[\,(s+1)^2+\kappa^2\,]},$$

and so confirming the solution reached using the Green's function approach.

6.31 Find the Green's function $x = G(t, t_0)$ that solves

$$\frac{d^2x}{dt^2} + \alpha\frac{dx}{dt} = \delta(t - t_0)$$

under the initial conditions $x = dx/dt = 0$ at $t = 0$. Hence solve

$$\frac{d^2x}{dt^2} + \alpha\frac{dx}{dt} = f(t),$$

where $f(t) = 0$ for $t < 0$. Evaluate your answer explicitly for $f(t) = Ae^{-\beta t}$ $(t > 0)$.

It is clear that one solution, $x(t)$, to the homogeneous equation has $\ddot{x} = -\alpha\dot{x}$ and is therefore $x(t) = Ae^{-\alpha t}$. The equation is of second order and therefore has a second solution; this is the trivial (but perfectly valid) x is a constant. The CF is thus $x(t) = Ae^{-\alpha t} + B$.

Let

$$G(t, t_0) = \begin{cases} Ae^{-\alpha t} + B, & 0 \le t \le t_0,\\ Ce^{-\alpha t} + D, & t > t_0.\end{cases}$$

Now, the initial conditions give

$$\begin{aligned} x(0) = 0 \quad &\Rightarrow \quad A + B = 0,\\ \dot{x}(0) = 0 \quad &\Rightarrow \quad -\alpha A = 0 \quad \Rightarrow \quad A = B = 0.\end{aligned}$$

Thus $G(t, t_0) = 0$ for $0 \le t \le t_0$.

The continuity/discontinuity conditions determine C and D through

$$Ce^{-\alpha t_0} + D - 0 = 0,$$
$$-\alpha Ce^{-\alpha t_0} - 0 = 1, \quad \Rightarrow \quad C = -\frac{e^{\alpha t_0}}{\alpha} \text{ and } D = \frac{1}{\alpha}.$$

It follows that $\qquad G(t, t_0) = \dfrac{1}{\alpha}[\,1 - e^{-\alpha(t-t_0)}\,]$ for $t > t_0$.

The general formalism now gives the solution of

$$\frac{d^2x}{dt^2} + \alpha\frac{dx}{dt} = f(t)$$

as

$$x(t) = \int_0^t \frac{1}{\alpha}[\,1 - e^{-\alpha(t-\tau)}\,]f(\tau)\,d\tau.$$

With $f(t) = Ae^{-\beta t}$ this becomes

$$\begin{aligned}
x(t) &= \int_0^t \frac{1}{\alpha}[\,1 - e^{-\alpha(t-\tau)}\,]Ae^{-\beta\tau}\,d\tau \\
&= \frac{A}{\alpha}\int_0^t (e^{-\beta\tau} - e^{-\alpha t}e^{(\alpha-\beta)\tau})\,d\tau \\
&= A\left[\frac{1-e^{-\beta t}}{\alpha\beta} - \frac{e^{-\beta t} - e^{-\alpha t}}{\alpha(\alpha-\beta)}\right] \\
&= A\left[\frac{\alpha - \beta - \alpha e^{-\beta t} + \beta e^{-\alpha t}}{\beta\alpha(\alpha-\beta)}\right] \\
&= A\left[\frac{\alpha(1-e^{-\beta t}) - \beta(1-e^{-\alpha t})}{\beta\alpha(\alpha-\beta)}\right].
\end{aligned}$$

This is the required explicit solution.

6.33 Solve

$$2y\frac{d^3y}{dx^3} + 2\left(y + 3\frac{dy}{dx}\right)\frac{d^2y}{dx^2} + 2\left(\frac{dy}{dx}\right)^2 = \sin x.$$

The only realistic hope for this non-linear equation is to try to arrange it as an exact equation! We note that the second and fourth terms can be written as the derivative of a product, and that adding and subtracting $2y'y''$ will enable the first term to be written in a similar way. We therefore rewrite the equation as

$$\frac{d}{dx}\left(2y\frac{d^2y}{dx^2}\right) + \frac{d}{dx}\left(2y\frac{dy}{dx}\right) + (6-2)\frac{dy}{dx}\frac{d^2y}{dx^2} = \sin x,$$
$$\frac{d}{dx}\left(2y\frac{d^2y}{dx^2}\right) + \frac{d}{dx}\left(2y\frac{dy}{dx}\right) + \frac{d}{dx}\left[2\left(\frac{dy}{dx}\right)^2\right] = \sin x.$$

This second form is obtained by noting that the final term on the LHS of the first equation happens to be an exact differential. Thus the whole of the LHS is an exact differential and one stage of integration can be carried out:

$$2y\frac{d^2y}{dx^2} + 2y\frac{dy}{dx} + 2\left(\frac{dy}{dx}\right)^2 = -\cos x + A.$$

We now note that the first and third terms of this integrated equation can be combined as the derivative of a product, whilst the second term is the derivative of y^2. This allows a further step of integration:

$$\begin{aligned}
\frac{d}{dx}\left(2y\frac{dy}{dx}\right) + 2y\frac{dy}{dx} &= -\cos x + A,\\
\frac{d}{dx}\left(2y\frac{dy}{dx}\right) + \frac{d(y^2)}{dx} &= -\cos x + A,\\
\Rightarrow \quad 2y\frac{dy}{dx} + y^2 &= -\sin x + Ax + B,\\
\frac{d(y^2)}{dx} + y^2 &= -\sin x + Ax + B.
\end{aligned}$$

At this stage an integrating factor is needed. However, as the LHS consists of the sum of the differentiated and undifferentiated forms of the same function, the required IF is simply e^x. After multiplying through by this, we obtain

$$\begin{aligned}
&\frac{d}{dx}\left(e^x y^2\right) = -e^x \sin x + Axe^x + Be^x,\\
\Rightarrow \quad &y^2 = e^{-x}\left[C + \int^x (B + Au - \sin u)e^u \, du\right]\\
&\quad = Ce^{-x} + B + A(x-1) - \tfrac{1}{2}(\sin x - \cos x).
\end{aligned}$$

The last term in this final solution is obtained by considering

$$\begin{aligned}
\int^x e^u \sin u \, du &= \mathrm{Im} \int^x e^{(1+i)u} \, du\\
&= \mathrm{Im}\left[\frac{e^{(1+i)u}}{1+i}\right]^x\\
&= \mathrm{Im}\left[\tfrac{1}{2}(1-i)e^{(1+i)x}\right]\\
&= \tfrac{1}{2}e^x(\sin x - \cos x).
\end{aligned}$$

6.35 Confirm that the equation

$$2x^2 y \frac{d^2y}{dx^2} + y^2 = x^2 \left(\frac{dy}{dx}\right)^2 \qquad (*)$$

is homogeneous in both x and y separately. Make two successive transformations that exploit this fact, starting with a substitution for x, to obtain an equation of the form

$$2\frac{d^2v}{dt^2} + \left(\frac{dv}{dt}\right)^2 - 2\frac{dv}{dt} + 1 = 0.$$

By writing $dv/dt = p$, solve this equation for $v = v(t)$ and hence find the solution to $(*)$.

The "net-power" (weight) of x in each of the terms is $2 - 2 = 0$, 0, $2 - 2 = 0$, i.e. they are all equal (to zero) and the equation is homogeneous in x. Similarly for y, the weights are 2, 2 and 2; and so the equation is also (separately) homogeneous in y.

Using the homogeneity in x, set $x = e^t$ with $\dfrac{d}{dx} = e^{-t}\dfrac{d}{dt}$. The equation then reads

$$2e^{2t} ye^{-t} \frac{d}{dt}\left(e^{-t}\frac{dy}{dt}\right) + y^2 = e^{2t}\left(e^{-t}\frac{dy}{dt}\right)^2,$$

$$2ye^t\left(e^{-t}\frac{d^2y}{dt^2} - e^{-t}\frac{dy}{dt}\right) + y^2 = e^{2t}e^{-2t}\left(\frac{dy}{dt}\right)^2,$$

$$2y\frac{d^2y}{dt^2} - 2y\frac{dy}{dt} + y^2 = \left(\frac{dy}{dt}\right)^2.$$

Now using the homogeneity in y, set $y = e^v$ with

$$\frac{dy}{dt} = \frac{dy}{dv}\frac{dv}{dt} = e^v\frac{dv}{dt} \quad \text{and} \quad \frac{d^2y}{dt^2} = \frac{d}{dt}\left(e^v\frac{dv}{dt}\right) = e^v\frac{dv}{dt}\frac{dv}{dt} + e^v\frac{d^2v}{dt^2}$$

and obtain

$$2e^v\left[e^v\left(\frac{dv}{dt}\right)^2 + e^v\frac{d^2v}{dt^2}\right] - 2e^v e^v\frac{dv}{dt} + e^{2v} = e^{2v}\left(\frac{dv}{dt}\right)^2.$$

Canceling e^{2v} all through, this reduces to

$$2\frac{d^2v}{dt^2} + \left(\frac{dv}{dt}\right)^2 - 2\frac{dv}{dt} + 1 = 0.$$

Since v does not appear in this equation undifferentiated, write $dv/dt = p$, obtaining

$$2\frac{dp}{dt} + p^2 - 2p + 1 = 0 \qquad \Rightarrow \qquad \frac{dp}{dt} = -\tfrac{1}{2}(p-1)^2.$$

This is separable:

$$\frac{dp}{(p-1)^2} = -\tfrac{1}{2}dt \qquad \Rightarrow \qquad (p-1)^{-1} = +\tfrac{1}{2}t + A.$$

Rewriting p as dv/dt, we now have

$$\frac{dv}{dt} - 1 = \frac{1}{\tfrac{1}{2}t + A} \qquad \Rightarrow \quad v - t = 2\ln\left(A + \tfrac{1}{2}t\right) + B.$$

Resubstituting for v and t now gives

$$\begin{aligned} \ln y &= 2\ln(A + \tfrac{1}{2}\ln x) + B + \ln x, \\ \Rightarrow \qquad y &= x(C + D\ln x)^2, \end{aligned}$$

where $C = Ae^{B/2}$ and $D = \frac{1}{2}e^{B/2}$.

7 Series solutions of ordinary differential equations

7.1 Find two power series solutions about $z = 0$ of the differential equation

$$(1 - z^2)y'' - 3zy' + \lambda y = 0.$$

Deduce that the value of λ for which the corresponding power series becomes an Nth-degree polynomial $U_N(z)$ is $N(N + 2)$. Construct $U_2(z)$ and $U_3(z)$.

If the equation is imagined divided through by $(1 - z^2)$ it is straightforward to see that, although $z = \pm 1$ are singular points of the equation, the point $z = 0$ is an *ordinary* point. We therefore expect two (uncomplicated!) series solutions with indicial values $\sigma = 0$ and $\sigma = 1$.

(a) $\sigma = 0$ and $y(z) = \sum_{n=0}^{\infty} a_n z^n$ with $a_0 \neq 0$.
Substituting and equating the coefficients of z^m,

$$(1 - z^2) \sum_{n=0}^{\infty} n(n-1)a_n z^{n-2} - 3 \sum_{n=0}^{\infty} n a_n z^n + \lambda \sum_{n=0}^{\infty} a_n z^n = 0,$$
$$(m+2)(m+1)a_{m+2} - m(m-1)a_m - 3ma_m + \lambda a_m = 0,$$

gives as the recurrence relation

$$a_{m+2} = \frac{m(m-1) + 3m - \lambda}{(m+2)(m+1)} a_m = \frac{m(m+2) - \lambda}{(m+1)(m+2)} a_m.$$

Since this recurrence relation connects alternate coefficients a_m, and $a_0 \neq 0$, only the coefficients with even indices are generated. All such coefficients with index higher than m will become zero, and the series will become an Nth-degree polynomial $U_N(z)$, if $\lambda = m(m+2) = N(N+2)$ for some (even) m appearing in the series; here, this means any positive even integer N.

To construct $U_2(z)$ we need to take $\lambda = 2(2+2) = 8$. The recurrence relation gives a_2 as

$$a_2 = \frac{0 - 8}{(0+1)(0+2)} a_0 = -4a_0 \quad \Rightarrow \quad U_2(z) = a_0(1 - 4z^2).$$

(b) $\sigma = 1$ and $y(z) = z \sum_{n=0}^{\infty} a_n z^n$ with $a_0 \neq 0$.
Substituting and equating the coefficients of z^{m+1},

$$(1 - z^2) \sum_{n=0}^{\infty} (n+1)n a_n z^{n-1} - 3 \sum_{n=0}^{\infty} (n+1) a_n z^{n+1} + \lambda \sum_{n=0}^{\infty} a_n z^{n+1} = 0,$$
$$(m+3)(m+2)a_{m+2} - (m+1)ma_m - 3(m+1)a_m + \lambda a_m = 0,$$

gives as the recurrence relation

$$a_{m+2} = \frac{m(m+1) + 3(m+1) - \lambda}{(m+2)(m+3)} a_m = \frac{(m+1)(m+3) - \lambda}{(m+2)(m+3)} a_m.$$

Again, all coefficients with index higher than m will become zero, and the series will become an Nth-degree polynomial $U_N(z)$, if $\lambda = (m+1)(m+3) = N(N+2)$ for some (even) m appearing in the series; here, this means any positive odd integer N.

To construct $U_3(z)$ we need to take $\lambda = 3(3+2) = 15$. The recurrence relation gives a_2 as

$$a_2 = \frac{3 - 15}{(0+2)(0+3)} a_0 = -2a_0.$$

Thus,

$$U_3(z) = a_0(z - 2z^3).$$

7.3 Find power series solutions in z of the differential equation

$$zy'' - 2y' + 9z^5 y = 0.$$

Identify closed forms for the two series, calculate their Wronskian, and verify that they are linearly independent. Compare the Wronskian with that calculated from the differential equation.

Putting the equation in its standard form shows that $z = 0$ is a singular point of the equation but, as $-2z/z$ and $9z^7/z$ are finite as $z \to 0$, it is a regular singular point. We therefore substitute a Frobenius-type solution,

$$y(z) = z^\sigma \sum_{n=0}^{\infty} a_n z^n \text{ with } a_0 \neq 0,$$

and obtain

$$\sum_{n=0}^{\infty} (n+\sigma)(n+\sigma-1) a_n z^{n+\sigma-1} - 2\sum_{n=0}^{\infty} (n+\sigma) a_n z^{n+\sigma-1} + 9\sum_{n=0}^{\infty} a_n z^{n+\sigma+5} = 0.$$

Equating the coefficient of $z^{\sigma-1}$ to zero gives the indicial equation as

$$\sigma(\sigma-1)a_0 - 2\sigma a_0 = 0 \quad \Rightarrow \quad \sigma = 0,\ 3.$$

These differ by an integer and may or may not yield two independent solutions. The larger root, $\sigma = 3$, will give a solution; the smaller one, $\sigma = 0$, may not.

(a) $\sigma = 3$.

Equating the general coefficient of z^{m+2} to zero (with $\sigma = 3$) gives

$$(m+3)(m+2)a_m - 2(m+3)a_m + 9a_{m-6} = 0.$$

Hence the recurrence relation is

$$a_m = -\frac{9a_{m-6}}{m(m+3)},$$
$$\Rightarrow \quad a_{6p} = -\frac{9}{6p\,(6p+3)}\,a_{6p-6} = -\frac{a_{6p-6}}{2p\,(2p+1)} = \frac{(-1)^p a_0}{(2p+1)!}.$$

The first solution is therefore given by

$$y_1(x) = a_0 z^3 \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{6n} = a_0 \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{3(2n+1)} = a_0 \sin z^3.$$

(b) $\sigma = 0$.
Equating the general coefficient of z^{m-1} to zero (with $\sigma = 0$) gives

$$m(m-1)a_m - 2ma_m + 9a_{m-6} = 0.$$

Hence the recurrence relation is

$$a_m = -\frac{9a_{m-6}}{m(m-3)},$$
$$\Rightarrow \quad a_{6p} = -\frac{9}{6p\,(6p-3)}\,a_{6p-6} = -\frac{a_{6p-6}}{2p\,(2p-1)} = \frac{(-1)^p a_0}{(2p)!}.$$

A second solution is thus

$$y_2(x) = a_0 \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{6n} = a_0 \cos z^3.$$

We see that $\sigma = 0$ does, in fact, produce a (different) series solution. This is because the recurrence relation relates a_n to a_{n+6} and does not involve a_{n+3}; the relevance here of considering the subscripted index "$m + 3$" is that "3" is the difference between the two indicial values.

We now calculate the Wronskian of the two solutions, $y_1 = a_0 \sin z^3$ and $y_2 = b_0 \cos z^3$:

$$\begin{aligned} W(y_1, y_2) &= y_1 y_2' - y_2 y_1' \\ &= a_0 \sin z^3(-3b_0 z^2 \sin z^3) - b_0 \cos z^3(3a_0 z^2 \cos z^3) \\ &= -3a_0 b_0 z^2 \neq 0. \end{aligned}$$

The fact that the Wronskian is non-zero shows that the two solutions are linearly independent.

We can also calculate the Wronskian from the original equation in its standard form,

$$y'' - \frac{2}{z}y' + 9z^4 y = 0,$$

as

$$W = C \exp\left\{-\int^z \frac{-2}{u}\,du\right\} = C \exp(2 \ln z) = Cz^2.$$

This is in agreement with the Wronskian calculated from the solutions, as it must be.

7.5 Investigate solutions of Legendre's equation at one of its singular points as follows.

(a) Verify that $z = 1$ is a regular singular point of Legendre's equation and that the indicial equation for a series solution in powers of $(z - 1)$ has a double root $\sigma = 0$.
(b) Obtain the corresponding recurrence relation and show that a polynomial solution is obtained if ℓ is a positive integer.
(c) Determine the radius of convergence R of the $\sigma = 0$ series and relate it to the positions of the singularities of Legendre's equation.

(a) In standard form, Legendre's equation reads

$$y'' - \frac{2z}{1 - z^2} y' + \frac{\ell(\ell + 1)}{1 - z^2} y = 0.$$

This has a singularity at $z = 1$, but, since

$$\frac{-2z(z - 1)}{1 - z^2} \to 1 \text{ and } \frac{\ell(\ell + 1)(z - 1)^2}{1 - z^2} \to 0 \text{ as } z \to 1,$$

i.e. both limits are finite, the point is a regular singular point.

We next change the origin to the point $z = 1$ by writing $u = z - 1$ and $y(z) = f(u)$. The transformed equation is

$$f'' - \frac{2(u + 1)}{-u(u + 2)} f' + \frac{\ell(\ell + 1)}{-u(u + 2)} y = 0$$

or
$$-u(u + 2) f'' - 2(u + 1) f' + \ell(\ell + 1) f = 0.$$

The point $u = 0$ is a regular singular point of this equation and so we set $f(u) = u^\sigma \sum_{n=0}^{\infty} a_n u^n$ and obtain

$$-u(u + 2) \sum_{n=0}^{\infty} (\sigma + n)(\sigma + n - 1) a_n u^{\sigma + n - 2}$$
$$- 2(u + 1) \sum_{n=0}^{\infty} (\sigma + n) a_n u^{\sigma + n - 1} + \ell(\ell + 1) \sum_{n=0}^{\infty} a_n u^{\sigma + n} = 0.$$

Equating to zero the coefficient of $u^{\sigma - 1}$ gives

$$-2\sigma(\sigma - 1) a_0 - 2\sigma a_0 = 0 \quad \Rightarrow \quad \sigma^2 = 0;$$

i.e. the indicial equation has a double root $\sigma = 0$.

(b) To obtain the recurrence relation we set the coefficient of u^m equal to zero for general m:

$$-m(m - 1) a_m - 2(m + 1) m a_{m+1} - 2m a_m - 2(m + 1) a_{m+1} + \ell(\ell + 1) a_m = 0.$$

Tidying this up gives

$$2(m + 1)(m + 1) a_{m+1} = [\ell(\ell + 1) - m^2 + m - 2m] a_m,$$
$$\Rightarrow \quad a_{m+1} = \frac{\ell(\ell + 1) - m(m + 1)}{2(m + 1)^2} a_m.$$

From this it is clear that, if ℓ is a positive integer, then $a_{\ell+1}$ and all further a_n are zero and that the solution is a polynomial (of degree ℓ).

(c) The limit of the ratio of successive terms in the series is given by

$$\left|\frac{a_{n+1}u^{n+1}}{a_n u^n}\right| = \left|\frac{u[\ell(\ell+1) - m(m+1)]}{2(m+1)^2}\right| \to \frac{|u|}{2} \text{ as } m \to \infty.$$

For convergence this limit needs to be < 1, i.e. $|u| < 2$. Thus the series converges in a circle of radius 2 centered on $u = 0$, i.e. on $z = 1$. The value 2 is to be expected, as it is the distance from $z = 1$ of the next nearest (actually the only other) singularity of the equation (at $z = -1$), excluding $z = 1$ itself.

7.7 The first solution of Bessel's equation for $\nu = 0$ is

$$J_0(z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!\Gamma(n+1)} \left(\frac{z}{2}\right)^{2n}.$$

Use the derivative method to show that

$$J_0(z)\ln z - \sum_{n=1}^{\infty} \frac{(-1)^n}{(n!)^2} \left(\sum_{r=1}^{n} \frac{1}{r}\right) \left(\frac{z}{2}\right)^{2n}$$

is a second (independent) solution.

Bessel's equation with $\nu = 0$ reads

$$zy'' + y' + zy = 0.$$

The recurrence relations that gave rise to the first solution, $J_0(z)$, were $(\sigma + 1)^2 a_1 = 0$ and $(\sigma + n)^2 a_n + a_{n-2} = 0$ for $n \geq 2$. Thus, in a general form as a function of σ, the solution is given by

$$y(\sigma, z) = a_0 z^\sigma \left\{1 - \frac{z^2}{(\sigma+2)^2} + \frac{z^4}{(\sigma+2)^2(\sigma+4)^2} - \cdots \right.$$
$$\left. + \frac{(-1)^n z^{2n}}{[(\sigma+2)(\sigma+4)\dots(\sigma+2n)]^2} + \cdots \right\}.$$

Setting $\sigma = 0$ reproduces the first solution given above.

To obtain a second independent solution, we must differentiate the above expression with respect to σ, before setting σ equal to 0:

$$\frac{\partial y}{\partial \sigma} = \ln z\, J_0(z) + \sum_{n=1}^{\infty} \frac{da_{2n}(\sigma)}{d\sigma} z^{\sigma+2n} \quad \text{at } \sigma = 0.$$

Now

$$\left.\frac{da_{2n}(\sigma)}{d\sigma}\right|_{\sigma=0} = \frac{d}{d\sigma}\left\{\frac{(-1)^n}{[(\sigma+2)(\sigma+4)\ldots(\sigma+2n)]^2}\right\}_{\sigma=0}$$

$$= \frac{(-1)^n(-2)}{[\ldots]^3}\left(\frac{[\ldots]}{\sigma+2}+\frac{[\ldots]}{\sigma+4}+\cdots+\frac{[\ldots]}{\sigma+2n}\right)$$

$$= \frac{(-2)(-1)^n}{[\ldots]^2}\sum_{r=1}^{n}\frac{1}{\sigma+2r}$$

$$= \frac{-2(-1)^n}{2^{2n}(n!)^2}\sum_{r=1}^{n}\frac{1}{2r}, \qquad \text{at } \sigma = 0.$$

Substituting this result, we obtain the second series as

$$J_0(z)\ln z - \sum_{n=1}^{\infty}\frac{(-1)^n}{(n!)^2}\left(\sum_{r=1}^{n}\frac{1}{r}\right)\left(\frac{z}{2}\right)^{2n}.$$

This is the form given in the question.

7.9 Find series solutions of the equation $y'' - 2zy' - 2y = 0$. Identify one of the series as $y_1(z) = \exp z^2$ and verify this by direct substitution. By setting $y_2(z) = u(z)y_1(z)$ and solving the resulting equation for $u(z)$, find an explicit form for $y_2(z)$ and deduce that

$$\int_0^x e^{-v^2}\,dv = e^{-x^2}\sum_{n=0}^{\infty}\frac{n!}{2(2n+1)!}(2x)^{2n+1}.$$

(a) The origin is an ordinary point of the equation and so power series solutions will be possible. Substituting $y(z) = \sum_{n=0}^{\infty} a_n z^n$ gives

$$\sum_{n=0}^{\infty} n(n-1)a_n z^{n-2} - 2\sum_{n=0}^{\infty} n a_n z^n - 2\sum_{n=0}^{\infty} a_n z^n = 0.$$

Equating to zero the coefficient of z^{m-2} yields the recurrence relation

$$a_m = \frac{2m-2}{m(m-1)}\,a_{m-2} = \frac{2}{m}\,a_{m-2}.$$

The solution with $a_0 = 1$ and $a_1 = 0$ is therefore

$$y_1(z) = 1 + \frac{2z^2}{2} + \frac{2^2 z^4}{(2)(4)} + \cdots + \frac{2^n z^{2n}}{2^n\, n!} + \cdots$$

$$= \sum_{n=0}^{\infty}\frac{z^{2n}}{n!} = \exp z^2.$$

Putting this result into the original equation,

$$(4z^2+2)\exp z^2 - 2z\,2z\exp z^2 - 2\exp z^2 = 0,$$

shows directly that it is a valid solution.

The solution with $a_0 = 0$ and $a_1 = 1$ takes the form

$$y_2(z) = z + \frac{2z^3}{3} + \frac{2^2 z^5}{(3)(5)} + \cdots + \frac{2^n\, 2^n\, n!\, z^{2n+1}}{(2n+1)!} + \cdots$$
$$= \sum_{n=0}^{\infty} \frac{n!\,(2z)^{2n+1}}{2(2n+1)!}.$$

(b) We now set $y_2(z) = u(z)y_1(z)$ and substitute it into the original equation. As they must, the terms in which u is undifferentiated cancel and leave

$$u'' \exp z^2 + 2u'(2z \exp z^2) - 2zu' \exp z^2 = 0.$$

It follows that

$$\frac{u''}{u'} = -2z \quad \Rightarrow \quad u' = Ae^{-z^2} \quad \Rightarrow \quad u(x) = A\int^x e^{-v^2}\,dv.$$

Hence, setting the two derived forms for a second solution equal to each other, we have

$$\sum_{n=0}^{\infty} \frac{n!\,(2x)^{2n+1}}{2(2n+1)!} = y_2(x) = y_1(x)u(x) = e^{x^2} A \int^x e^{-v^2}\,dv.$$

For arbitrary small x, only the $n = 0$ term in the series is significant and takes the value $2x/2 = x$, whilst the integral is $A\int^x 1\,dv = Ax$. Thus $A = 1$ and the equality

$$\int_0^x e^{-v^2}\,dv = e^{-x^2} \sum_{n=0}^{\infty} \frac{n!\,(2x)^{2n+1}}{2(2n+1)!}$$

holds for all x.

7.11 For the equation $y'' + z^{-3}y = 0$, show that the origin becomes a regular singular point if the independent variable is changed from z to $x = 1/z$. Hence find a series solution of the form $y_1(z) = \sum_0^\infty a_n z^{-n}$. By setting $y_2(z) = u(z)y_1(z)$ and expanding the resulting expression for du/dz in powers of z^{-1}, show that $y_2(z)$ is a second solution with asymptotic form

$$y_2(z) = c\left[z + \ln z - \tfrac{1}{2} + \mathrm{O}\left(\frac{\ln z}{z}\right)\right],$$

where c is an arbitrary constant.

With the equation in its original form, it is clear that, since $z^2/z^3 \to \infty$ as $z \to 0$, the origin is an irregular singular point. However, if we set $1/z = \xi$ and $y(z) = Y(\xi)$, with

$$\frac{d\xi}{dz} = -\frac{1}{z^2} = -\xi^2 \quad \Rightarrow \quad \frac{d}{dz} = -\xi^2 \frac{d}{d\xi},$$

then

$$-\xi^2 \frac{d}{d\xi}\left(-\xi^2 \frac{dY}{d\xi}\right) + \xi^3 Y = 0,$$

$$\xi^2 \frac{d^2Y}{d\xi^2} + 2\xi \frac{dY}{d\xi} + \xi Y = 0,$$

$$Y'' + \frac{2}{\xi} Y' + \frac{1}{\xi} Y = 0.$$

By inspection, $\xi = 0$ is a regular singular point of this equation, and its indicial equation is

$$\sigma(\sigma - 1) + 2\sigma = 0 \quad \Rightarrow \quad \sigma = 0, \ -1.$$

We start with the larger root, $\sigma = 0$, as this is "guaranteed" to give a valid series solution and assume a solution of the form $Y(\xi) = \sum_{n=0}^{\infty} a_n \xi^n$, leading to

$$\sum_{n=0}^{\infty} n(n-1)a_n \xi^{n-1} + 2\sum_{n=0}^{\infty} n a_n \xi^{n-1} + \sum_{n=0}^{\infty} a_n \xi^n = 0.$$

Equating to zero the coefficient of ξ^{m-1} gives the recurrence relation

$$a_m = \frac{-a_{m-1}}{m(m+1)} \quad \Rightarrow \quad a_m = \frac{(-1)^m}{(m+1)\,(m!)^2}\, a_0$$

and the series solution in inverse powers of z,

$$y_1(z) = a_0 \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)\,(n!)^2\, z^n}.$$

To find the second solution we set $y_2(z) = f(z)y_1(z)$. As usual (and as intended), all terms with f undifferentiated vanish when this is substituted in the original equation. What is left is

$$0 = f''(z)y_1(z) + 2f'(z)y_1'(z),$$

which on rearrangement yields

$$\frac{f''}{f'} = -\frac{2y_1'}{y_1}.$$

This equation, although it contains a second derivative, is in fact only a first-order equation (for f'). It can be integrated directly to give

$$\ln f' = -2 \ln y_1 + c.$$

After exponentiation, this equation can be written as

$$\frac{df}{dz} = \frac{A}{y_1^2(z)} = \frac{A}{a_0^2}\left(1 - \frac{1}{2 \times 1^2\, z} + \frac{1}{3 \times 2^2\, z^2} - \cdots\right)^{-2}$$

$$= \frac{A}{a_0^2}\left[1 + \frac{1}{z} + \mathrm{O}\left(\frac{1}{z^2}\right)\right],$$

where $A = e^c$.

Hence, on integrating a second time, one obtains

$$f(z) = \frac{A}{a_0^2}\left(z + \ln z + \mathrm{O}\left(\frac{1}{z}\right)\right),$$

which in turn implies

$$\begin{aligned} y_2(z) &= \frac{A}{a_0^2}\left[z + \ln z + \mathrm{O}\left(\frac{1}{z}\right)\right] a_0\left(1 - \frac{1}{2z} + \frac{1}{12z^2} - \cdots\right) \\ &= c\left[z + \ln z - \frac{1}{2} + \mathrm{O}\left(\frac{\ln z}{z}\right)\right]. \end{aligned}$$

This establishes the asymptotic form of the second solution.

7.13 The origin is an ordinary point of the Chebyshev equation,

$$(1 - z^2)y'' - zy' + m^2 y = 0,$$

which therefore has series solutions of the form $z^\sigma \sum_0^\infty a_n z^n$ for $\sigma = 0$ and $\sigma = 1$.

(a) Find the recurrence relationships for the a_n in the two cases and show that there exist polynomial solutions $T_m(z)$:
  (i) for $\sigma = 0$, when m is an even integer, the polynomial having $\frac{1}{2}(m + 2)$ terms;
  (ii) for $\sigma = 1$, when m is an odd integer, the polynomial having $\frac{1}{2}(m + 1)$ terms.
(b) $T_m(z)$ is normalized so as to have $T_m(1) = 1$. Find explicit forms for $T_m(z)$ for $m = 0, 1, 2, 3$.
(c) Show that the corresponding non-terminating series solutions $S_m(z)$ have as their first few terms

$$\begin{aligned} S_0(z) &= a_0\left(z + \frac{1}{3!}z^3 + \frac{9}{5!}z^5 + \cdots\right), \\ S_1(z) &= a_0\left(1 - \frac{1}{2!}z^2 - \frac{3}{4!}z^4 - \cdots\right), \\ S_2(z) &= a_0\left(z - \frac{3}{3!}z^3 - \frac{15}{5!}z^5 - \cdots\right), \\ S_3(z) &= a_0\left(1 - \frac{9}{2!}z^2 + \frac{45}{4!}z^4 + \cdots\right). \end{aligned}$$

(a) (i) If, for $\sigma = 0$, $y(z) = \sum_{n=0}^\infty a_n z^n$ with $a_0 \neq 0$, the condition for the coefficient of z^r in

$$(1 - z^2)\sum_{n=0}^\infty n(n-1)a_n z^{n-2} - z\sum_{n=0}^\infty n a_n z^{n-1} + m^2 \sum_{n=0}^\infty a_n z^n$$

to be zero is that

$$\begin{aligned} (r+2)(r+1)a_{r+2} - r(r-1)a_r - ra_r + m^2 a_r &= 0, \\ \Rightarrow \quad a_{r+2} &= \frac{r^2 - m^2}{(r+2)(r+1)}\, a_r. \end{aligned}$$

This relation relates a_{r+2} to a_r and so to a_0 if r is even. For a_{r+2} to vanish, in this case, requires that $r = m$, which must therefore be an even integer. The non-vanishing coefficients will be $a_0, a_2, \ldots, a_m$, i.e. $\frac{1}{2}(m + 2)$ of them in all.

(ii) If, for $\sigma = 1$, $y(z) = \sum_{n=0}^{\infty} a_n z^{n+1}$ with $a_0 \neq 0$, the condition for the coefficient of z^{r+1} in

$$(1 - z^2) \sum_{n=0}^{\infty} (n+1)na_n z^{n-1} - z \sum_{n=0}^{\infty} (n+1)a_n z^n + m^2 \sum_{n=0}^{\infty} a_n z^{n+1}$$

to be zero is that

$$(r+3)(r+2)a_{r+2} - (r+1)ra_r - (r+1)a_r + m^2 a_r = 0,$$
$$\Rightarrow \quad a_{r+2} = \frac{(r+1)^2 - m^2}{(r+3)(r+2)} a_r.$$

This relation relates a_{r+2} to a_r and so to a_0 if r is even. For a_{r+2} to vanish, in this case, requires that $r + 1 = m$, which must therefore be an odd integer. The non-vanishing coefficients will be, as before, $a_0, a_2, \ldots, a_{m-1}$, i.e. $\frac{1}{2}(m+1)$ of them in all.

(b) For $m = 0$, $T_0(z) = a_0$. With the given normalization, $a_0 = 1$ and $T_0(z) = 1$.

For $m = 1$, $T_1(z) = a_0 z$. The required normalization implies that $a_0 = 1$ and so $T_0(z) = z$.

For $m = 2$, we need the recurrence relation in (a)(i). This shows that

$$a_2 = \frac{0^2 - 2^2}{(2)(1)} a_0 = -2a_0 \quad \Rightarrow \quad T_2(z) = a_0(1 - 2z^2).$$

With the given normalization, $a_0 = -1$ and $T_2(z) = 2z^2 - 1$.

For $m = 3$, we use the recurrence relation in (a)(ii) and obtain

$$a_2 = \frac{1^2 - 3^2}{(3)(2)} a_0 = -\frac{4}{3} a_0 \quad \Rightarrow \quad T_3(z) = a_0 \left(z - \frac{4z^3}{3} \right).$$

For the required normalization, we must have $a_0 = -\frac{1}{3}$ and consequently that $T_3(z) = 4z^3 - 3z$.

(c) The non-terminating series solutions $S_m(z)$ arise when $\sigma = 0$ but m is an odd integer and when $\sigma = 1$ with m an even integer. We take each in turn and apply the appropriate recurrence relation to generate the coefficients.

(i) $\sigma = 0$, $m = 1$, using the (a)(i) recurrence relation:

$$a_2 = \frac{0 - 1}{(2)(1)} a_0 = -\frac{1}{2!} a_0, \quad a_4 = \frac{4 - 1}{(4)(3)} a_2 = -\frac{3}{4!} a_0.$$

Hence,

$$S_1(z) = a_0 \left(1 - \frac{1}{2!} z^2 - \frac{3}{4!} z^4 - \cdots \right).$$

(ii) $\sigma = 0$, $m = 3$, using the (a)(i) recurrence relation:

$$a_2 = \frac{0 - 9}{(2)(1)} a_0 = -\frac{9}{2!} a_0, \quad a_4 = \frac{4 - 9}{(4)(3)} a_2 = \frac{45}{4!} a_0.$$

Hence,

$$S_3(z) = a_0 \left(1 - \frac{9}{2!} z^2 + \frac{45}{4!} z^4 + \cdots \right).$$

(iii) $\sigma = 1, m = 0$, using the (a)(ii) recurrence relation:

$$a_2 = \frac{1-0}{(3)(2)}\,a_0 = \frac{1}{3!}\,a_0, \qquad a_4 = \frac{9-0}{(5)(4)}\,a_2 = \frac{9}{5!}\,a_0.$$

Hence,

$$S_0(z) = a_0\left(z + \frac{1}{3!}z^3 + \frac{9}{5!}z^5 + \cdots\right).$$

(iv) $\sigma = 1, m = 2$, using the (a)(ii) recurrence relation:

$$a_2 = \frac{1-4}{(3)(2)}\,a_0 = -\frac{3}{3!}\,a_0, \qquad a_4 = \frac{9-4}{(5)(4)}\,a_2 = -\frac{15}{5!}\,a_0.$$

Hence,

$$S_2(z) = a_0\left(z - \frac{3}{3!}z^3 - \frac{15}{5!}z^5 - \cdots\right).$$

8 Eigenfunction methods for differential equations

8.1 By considering $\langle h|h\rangle$, where $h = f + \lambda g$ with λ real, prove that, for two functions f and g,

$$\langle f|f\rangle\langle g|g\rangle \geq \tfrac{1}{4}[\langle f|g\rangle + \langle g|f\rangle]^2.$$

The function $y(x)$ is real and positive for all x. Its Fourier cosine transform $\tilde{y}_c(k)$ is defined by

$$\tilde{y}_c(k) = \int_{-\infty}^{\infty} y(x)\cos(kx)\,dx,$$

and it is given that $\tilde{y}_c(0) = 1$. Prove that

$$\tilde{y}_c(2k) \geq 2[\tilde{y}_c(k)]^2 - 1.$$

For any $|h\rangle$ we have that $\langle h|h\rangle \geq 0$, with equality only if $|h\rangle = |0\rangle$. Hence, noting that λ is real, we have

$$0 \leq \langle h|h\rangle = \langle f + \lambda g|f + \lambda g\rangle = \langle f|f\rangle + \lambda\langle g|f\rangle + \lambda\langle f|g\rangle + \lambda^2\langle g|g\rangle.$$

This equation, considered as a quadratic inequality in λ, states that the corresponding quadratic equation has no real roots. The condition for this ("$b^2 < 4ac$") is given by

$$[\langle g|f\rangle + \langle f|g\rangle]^2 \leq 4\langle f|f\rangle\langle g|g\rangle, \qquad (*)$$

from which the stated result follows immediately. Note that $\langle g|f\rangle + \langle f|g\rangle$ is real and its square is therefore non-negative.

The given datum is equivalent to

$$1 = \tilde{y}_c(0) = \int_{-\infty}^{\infty} y(x)\cos(0x)\,dx = \int_{-\infty}^{\infty} y(x)\,dx.$$

Now consider

$$\begin{aligned}
\tilde{y}_c(2k) &= \int_{-\infty}^{\infty} y(x)\cos(2kx)\,dx \\
&= 2\int_{-\infty}^{\infty} y(x)\cos^2 kx - \int_{-\infty}^{\infty} y(x)\,dx, \\
\Rightarrow \quad \tilde{y}_c(2k) + 1 &= 2\int_{-\infty}^{\infty} y(x)\cos^2 kx.
\end{aligned}$$

In order to use $(*)$, we need to choose for $f(x)$ and $g(x)$ functions whose product will form the integrand defining $\tilde{y}_c(k)$. With this in mind, we take $f(x) = y^{1/2}(x)\cos kx$ and

$g(x) = y^{1/2}(x)$; we may do this since $y(x) > 0$ for all x. Making these choices gives

$$\left(\int_{-\infty}^{\infty} y\cos kx\,dx + \int_{-\infty}^{\infty} y\cos kx\,dx\right)^2 \le 4\int_{-\infty}^{\infty} y\cos^2 kx\,dx \int_{-\infty}^{\infty} y\,dx,$$

$$\left(\int_{-\infty}^{\infty} 2y\cos kx\,dx\right)^2 \le 4\int_{-\infty}^{\infty} y\cos^2 kx\,dx \times 1,$$

$$4\tilde{y}_c^2(k) \le 4\int_{-\infty}^{\infty} y\cos^2 kx\,dx.$$

Thus,

$$\tilde{y}_c(2k) + 1 = 2\int_{-\infty}^{\infty} y(x)\cos^2 kx \ge 2[\tilde{y}_c(k)]^2$$

and hence the stated result.

8.3 Consider the real eigenfunctions $y_n(x)$ of a Sturm–Liouville equation

$$(py')' + qy + \lambda\rho y = 0, \qquad a \le x \le b,$$

in which $p(x)$, $q(x)$ and $\rho(x)$ are continuously differentiable real functions and $p(x)$ does not change sign in $a \le x \le b$. Take $p(x)$ as positive throughout the interval, if necessary by changing the signs of all eigenvalues. For $a \le x_1 \le x_2 \le b$, establish the identity

$$(\lambda_n - \lambda_m)\int_{x_1}^{x_2} \rho y_n y_m\,dx = \left[y_n\,p\,y_m' - y_m\,p\,y_n'\right]_{x_1}^{x_2}.$$

Deduce that if $\lambda_n > \lambda_m$ then $y_n(x)$ must change sign between two successive zeros of $y_m(x)$.

[The reader may find it helpful to illustrate this result by sketching the first few eigenfunctions of the system $y'' + \lambda y = 0$, with $y(0) = y(\pi) = 0$, and the Legendre polynomials $P_n(z)$ for $n = 2, 3, 4, 5$.]

The function $p(x)$ does not change sign in the interval $a \le x \le b$; we take it as positive, multiplying the equation all through by -1 if necessary. This means that the weight function ρ can still be taken as positive, but that we must consider all possible functions for $q(x)$ and eigenvalues λ of either sign.

We start with the eigenvalue equation for $y_n(x)$, multiply it through by $y_m(x)$ and then integrate from x_1 to x_2. From this result we subtract the same equation with the roles of n and m reversed, as follows. The integration limits are omitted until the explicit integration by parts is carried through:

$$\int y_m(p\,y_n')'\,dx + \int y_m q\,y_n\,dx + \int y_m\lambda_n\rho y_n\,dx = 0,$$

$$\int y_n(p\,y_m')'\,dx + \int y_n q\,y_m\,dx + \int y_n\lambda_m\rho y_m\,dx = 0,$$

$$\int \left[y_m(p\,y_n')' - y_n(p\,y_m')'\right]dx + (\lambda_n - \lambda_m)\int y_m\rho y_n\,dx = 0,$$

$$\left[y_m\,p\,y_n'\right]_{x_1}^{x_2} - \int y_m'\,p\,y_n'\,dx - \left[y_n\,p\,y_m'\right]_{x_1}^{x_2}$$
$$+ \int y_n'\,p\,y_m'\,dx + (\lambda_n - \lambda_m)\int y_m\rho y_n\,dx = 0.$$

Hence

$$(\lambda_n - \lambda_m) \int y_m \rho y_n \, dx = \left[y_n \, p \, y_m' - y_m \, p \, y_n' \right]_{x_1}^{x_2}. \qquad (*)$$

Now, in this general result, take x_1 and x_2 as successive zeros of $y_m(x)$, where m is determined by $\lambda_n > \lambda_m$ (after the signs have been changed, if that was necessary). Clearly the sign of $y_m(x)$ does not change in this interval; let it be α. It follows that the sign of $y_m'(x_1)$ is also α, whilst that of $y_m'(x_2)$ is $-\alpha$. In addition, the second term on the RHS of $(*)$ vanishes at both limits, as $y_m(x_1) = y_m(x_2) = 0$.

Let us now *suppose* that the sign of $y_n(x)$ does not change in this same interval and is always β. Then the sign of the expression on the LHS of $(*)$ is $(+1)(\alpha)(+1)\beta = \alpha\beta$. The first $(+1)$ appears because $\lambda_n > \lambda_m$.

The signs of the upper- and lower-limit contributions of the remaining term on the RHS of $(*)$ are $\beta(+1)(-\alpha)$ and $(-1)\beta(+1)\alpha$, respectively, the additional factor of (-1) in the second product arising from the fact that the contribution comes from a lower limit. The contributions at both limits have the same sign, $-\alpha\beta$, and so the sign of the total RHS must also be $-\alpha\beta$.

This contradicts, however, the sign of $+\alpha\beta$ found for the LHS. It follows that it was wrong to suppose that the sign of $y_n(x)$ does not change in the interval; in other words, a zero of $y_n(x)$ does appear between every pair of zeros of $y_m(x)$.

8.5 Use the properties of Legendre polynomials to solve the following problems.

(a) Find the solution of $(1 - x^2)y'' - 2xy' + by = f(x)$ that is valid in the range $-1 \le x \le 1$ and finite at $x = 0$, in terms of Legendre polynomials.

(b) Find the explicit solution if $b = 14$ and $f(x) = 5x^3$. Verify it by direct substitution.

[Explicit forms for the Legendre polynomials can be found in any textbook. In *Mathematical Methods for Physics and Engineering*, 3rd edition, they are given in Subsection 18.1.1.]

(a) The LHS of the given equation is the same as that of Legendre's equation and so we substitute $y(x) = \sum_{n=0}^{\infty} a_n P_n(x)$ and use the fact that $(1 - x^2)P_n'' - 2xP_n' = -n(n+1)P_n$. This results in

$$\sum_{n=0}^{\infty} a_n [b - n(n+1)] P_n = f(x).$$

Now, using the mutual orthogonality and normalization of the $P_n(x)$, we multiply both sides by $P_m(x)$ and integrate over x:

$$\sum_{n=0}^{\infty} a_n [b - n(n+1)] \, \delta_{mn} \frac{2}{2m+1} = \int_{-1}^{1} f(z) P_m(z) \, dz,$$

$$\Rightarrow \quad a_m = \frac{2m+1}{2[b - m(m+1)]} \int_{-1}^{1} f(z) P_m(z) \, dz.$$

This gives the coefficients in the solution $y(x)$.

(b) We now express $f(x)$ in terms of Legendre polynomials,

$$f(x) = 5x^3 = 2\left[\tfrac{1}{2}(5x^3 - 3x)\right] + 3[x] = 2P_3(x) + 3P_1(x),$$

and conclude that, because of the mutual orthogonality of the Legendre polynomials, only a_3 and a_1 in the series solution will be non-zero. To find them we need to evaluate

$$\int_{-1}^{1} f(z)P_3(z)\,dz = 2\,\frac{2}{2(3)+1} = \frac{4}{7};$$

similarly, $\int_{-1}^{1} f(z)P_1(z)\,dz = 3 \times (2/3) = 2.$

Inserting these values gives

$$a_3 = \frac{7}{2(14-12)}\,\frac{4}{7} = 1 \quad \text{and} \quad a_1 = \frac{3}{2(14-2)}\,2 = \frac{1}{4}.$$

Thus the solution is

$$y(x) = \frac{1}{4}P_1(x) + P_3(x) = \frac{1}{4}x + \frac{1}{2}(5x^3 - 3x) = \frac{5(2x^3 - x)}{4}.$$

Check:

$$(1-x^2)\frac{60x}{4} - 2x\,\frac{30x^2-5}{4} + \frac{140x^3 - 70x}{4} = 5x^3,$$
$$\Rightarrow \quad 60x - 60x^3 - 60x^3 + 10x + 140x^3 - 70x = 20x^3,$$

which is satisfied.

8.7 Consider the set of functions, $\{f(x)\}$, of the real variable x defined in the interval $-\infty < x < \infty$, that $\to 0$ at least as quickly as x^{-1}, as $x \to \pm\infty$. For unit weight function, determine whether each of the following linear operators is Hermitian when acting upon $\{f(x)\}$:

$$\text{(a)}\ \frac{d}{dx} + x; \quad \text{(b)}\ -i\frac{d}{dx} + x^2; \quad \text{(c)}\ ix\frac{d}{dx}; \quad \text{(d)}\ i\frac{d^3}{dx^3}.$$

For an operator $\mathcal{L}$ to be Hermitian over the given range with respect to a unit weight function, the equation

$$\int_{-\infty}^{\infty} f^*(x)[\mathcal{L}g(x)]\,dx = \left\{\int_{-\infty}^{\infty} g^*(x)[\mathcal{L}f(x)]\,dx\right\}^* \qquad (*)$$

must be satisfied for general functions f and g.

(a) For $\mathcal{L} = \dfrac{d}{dx} + x$, the LHS of $(*)$ is

$$\int_{-\infty}^{\infty} f^*(x)\left(\frac{dg}{dx} + xg\right)dx = [f^*g]_{-\infty}^{\infty} - \int_{-\infty}^{\infty}\frac{df^*}{dx}g\,dx + \int_{-\infty}^{\infty} f^*xg\,dx$$
$$= 0 - \int_{-\infty}^{\infty}\frac{df^*}{dx}g\,dx + \int_{-\infty}^{\infty} f^*xg\,dx.$$

The RHS of $(*)$ is

$$\int_{-\infty}^{\infty}\left\{g^*(x)\left(\frac{df}{dx}+xf\right)dx\right\}^* = \left\{\int_{-\infty}^{\infty} g^*\frac{df}{dx}\,dx\right\}^* + \left\{\int_{-\infty}^{\infty} g^* x f\,dx\right\}^*$$
$$= \int_{-\infty}^{\infty} g\frac{df^*}{dx}\,dx + \int_{-\infty}^{\infty} g x f^*\,dx.$$

Since the sign of the first term differs in the two expressions, the LHS $\neq$ RHS and $\mathcal{L}$ is *not* Hermitian. It will also be apparent that purely multiplicative terms in the operator, such as x or x^2, will always be Hermitian; thus we can ignore the x^2 term in part (b).

(b) As explained above, we need only consider

$$\int_{-\infty}^{\infty} f^*(x)\left(-i\frac{dg}{dx}\right)dx = \left[-if^*g\right]_{-\infty}^{\infty} + i\int_{-\infty}^{\infty}\frac{df^*}{dx}g\,dx$$
$$= 0 + i\int_{-\infty}^{\infty}\frac{df^*}{dx}g\,dx$$

and

$$\int_{-\infty}^{\infty}\left\{g^*(x)\left(-i\frac{df}{dx}\right)dx\right\}^* = i\int_{-\infty}^{\infty} g\frac{df^*}{dx}\,dx.$$

These are equal, and so $\mathcal{L} = -i\dfrac{d}{dx}$ is Hermitian, as is $\mathcal{L} = -i\dfrac{d}{dx} + x^2$.

(c) For $\mathcal{L} = ix\dfrac{d}{dx}$, the LHS of $(*)$ is

$$\int_{-\infty}^{\infty} f^*(x)\left(ix\frac{dg}{dx}\right)dx = \left[ixf^*g\right]_{-\infty}^{\infty} - i\int_{-\infty}^{\infty} x\frac{df^*}{dx}g\,dx - i\int_{-\infty}^{\infty} f^*g\,dx$$
$$= 0 - i\int_{-\infty}^{\infty} x\frac{df^*}{dx}g\,dx - i\int_{-\infty}^{\infty} f^*g\,dx.$$

The RHS of $(*)$ is given by

$$\int_{-\infty}^{\infty}\left\{g^*(x)ix\left(\frac{df}{dx}\right)dx\right\}^* = -i\int_{-\infty}^{\infty} gx\frac{df^*}{dx}\,dx.$$

Since, in general, $-i\int_{-\infty}^{\infty} fg^*\,dx \neq 0$, the two sides are not equal; therefore $\mathcal{L}$ is not Hermitian.

(d) Since $\mathcal{L} = i\dfrac{d^3}{dx^3}$ is the cube of the operator $-i\dfrac{d}{dx}$, which was shown in part (b) to be Hermitian, it is expected that $\mathcal{L}$ is Hermitian. This can be verified directly as follows.

The LHS of $(*)$ is given by

$$i\int_{-\infty}^{\infty} f^* \frac{d^3 g}{dx^3}\,dx = \left[i f^* \frac{d^2 g}{dx^2}\right]_{-\infty}^{\infty} - i\int_{-\infty}^{\infty} \frac{df^*}{dx}\frac{d^2 g}{dx^2}\,dx$$

$$= 0 - i\left[\frac{df^*}{dx}\frac{dg}{dx}\right]_{-\infty}^{\infty} + i\int_{-\infty}^{\infty} \frac{d^2 f^*}{dx^2}\frac{dg}{dx}\,dx$$

$$= 0 + i\left[\frac{d^2 f^*}{dx^2}\,g\right]_{-\infty}^{\infty} - i\int_{-\infty}^{\infty} \frac{d^3 f^*}{dx^3}\,g\,dx$$

$$= 0 + \left\{\int_{-\infty}^{\infty} i g^* \frac{d^3 f}{dx^3}\,dx\right\}^* = \text{RHS of } (*).$$

Thus $\mathcal{L}$ is confirmed as Hermitian.

8.9 Find an eigenfunction expansion for the solution with boundary conditions $y(0) = y(\pi) = 0$ of the inhomogeneous equation

$$\frac{d^2 y}{dx^2} + \kappa y = f(x),$$

where κ is a constant and

$$f(x) = \begin{cases} x & 0 \le x \le \pi/2, \\ \pi - x & \pi/2 < x \le \pi. \end{cases}$$

The eigenfunctions of the operator $\mathcal{L} = \dfrac{d^2}{dx^2} + \kappa$ are obviously

$$y_n(x) = A_n \sin nx + B_n \cos nx,$$

with corresponding eigenvalues $\lambda_n = n^2 - \kappa$.

The boundary conditions, $y(0) = y(\pi) = 0$, require that n is a positive integer and that $B_n = 0$, i.e.

$$y_n(x) = A_n \sin nx = \sqrt{\frac{2}{\pi}} \sin nx,$$

where A_n (for $n \ge 1$) has been chosen so that the eigenfunctions are normalized over the interval $x = 0$ to $x = \pi$. Since $\mathcal{L}$ is Hermitian on the range $0 \le x \le \pi$, the eigenfunctions are also mutually orthogonal, and so the $y_n(x)$ form an orthonormal set.

If the required solution is $y(x) = \sum_n a_n y_n(x)$, then direct substitution yields the result

$$\sum_{n=1}^{\infty} (\kappa - n^2) a_n y_n(x) = f(x).$$

Following the usual procedure for analysis using sets of orthonormal functions, this implies that

$$a_m = \frac{1}{\kappa - m^2}\int_0^{\pi} f(z) y_m(z)\,dz$$

and, consequently, that

$$y(x) = \sum_{n=1}^{\infty} \sqrt{\frac{2}{\pi}} \frac{\sin nx}{\kappa - n^2} \sqrt{\frac{2}{\pi}} \int_0^{\pi} f(z) \sin(nz)\, dz.$$

It only remains to evaluate

$$\begin{aligned}
I_n &= \int_0^{\pi} \sin(nx) f(x)\, dx \\
&= \int_0^{\pi/2} x \sin nx\, dx + \int_{\pi/2}^{\pi} (\pi - x) \sin nx\, dx \\
&= \left[\frac{-x \cos nx}{n}\right]_0^{\pi/2} + \int_0^{\pi/2} \frac{\cos nx}{n}\, dx \\
&\quad + \left[\frac{-(\pi - x) \cos nx}{n}\right]_{\pi/2}^{\pi} + \int_{\pi/2}^{\pi} \frac{(-1) \cos nx}{n}\, dx \\
&= -\frac{\pi}{2} \frac{\cos(n\pi/2)}{n} (1 - 1) + \left[\frac{\sin nx}{n^2}\right]_0^{\pi/2} - \left[\frac{\sin nx}{n^2}\right]_{\pi/2}^{\pi} \\
&= 0 + \frac{(-1)^{(n-1)/2}}{n^2} (1 + 1) \text{ for odd } n \text{ and } = 0 \text{ for even } n.
\end{aligned}$$

Thus,

$$y(x) = \frac{4}{\pi} \sum_{n \text{ odd}} \frac{(-1)^{(n-1)/2}}{n^2(\kappa - n^2)} \sin nx$$

is the required solution.

8.11 The differential operator $\mathcal{L}$ is defined by

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

Determine the eigenvalues λ_n of the problem

$$\mathcal{L}y_n = \lambda_n e^x y_n \qquad 0 < x < 1,$$

with boundary conditions

$$y(0) = 0, \qquad \frac{dy}{dx} + \tfrac{1}{2} y = 0 \quad \text{at} \quad x = 1.$$

(a) Find the corresponding unnormalized y_n, and also a weight function $\rho(x)$ with respect to which the y_n are orthogonal. Hence, select a suitable normalization for the y_n.
(b) By making an eigenfunction expansion, solve the equation

$$\mathcal{L}y = -e^{x/2}, \qquad 0 < x < 1,$$

subject to the same boundary conditions as previously.

When written out explicitly, the eigenvalue equation is

$$-\frac{d}{dx}\left(e^x \frac{dy}{dx}\right) - \tfrac{1}{4} e^x y = \lambda e^x y, \qquad (*)$$

or, on differentiating out the product,

$$e^x y'' + e^x y' + \left(\lambda + \tfrac{1}{4}\right)e^x y = 0.$$

The auxiliary equation is

$$m^2 + m + \left(\lambda + \tfrac{1}{4}\right) = 0 \quad \Rightarrow \quad m = -\tfrac{1}{2} \pm i\sqrt{\lambda}.$$

The general solution is thus given by

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

with the condition $y(0) = 0$ implying that $A = 0$. The other boundary condition requires that, at $x = 1$,

$$-\tfrac{1}{2}Be^{-x/2}\sin\sqrt{\lambda}x + \sqrt{\lambda}Be^{-x/2}\cos\sqrt{\lambda}x + \tfrac{1}{2}Be^{-x/2}\sin\sqrt{\lambda}x = 0,$$

i.e. that $\cos\sqrt{\lambda} = 0$ and hence that $\lambda = (n + \frac{1}{2})^2\pi^2$ for non-negative integral n.

(a) The unnormalized eigenfunctions are

$$y_n(x) = B_n e^{-x/2}\sin\left(n + \tfrac{1}{2}\right)\pi x$$

and $(*)$ is in Sturm–Liouville form. However, although $y_n(0) = 0$, the values at the upper limit in $\left[y_m' p\, y_n\right]_0^1$ are $y_n(1) = B_n e^{-1/2}(-1)^n$, $p(1) = e^1$ and $y_m'(1) = -\frac{1}{2}B_m e^{-1/2}(-1)^m$. Consequently, $\left[y_m' p\, y_n\right]_0^1 \neq 0$ and SL theory cannot be applied. We therefore have to find a suitable weight function $\rho(x)$ by inspection. Given the general form of the eigenfunctions, ρ has to be taken as e^x, with the orthonormality integral taking the form

$$\begin{aligned}
I_{nm} &= \int_0^1 \rho(x) y_n(x) y_m^*(x)\,dx \\
&= B_n B_m \int_0^1 e^x e^{-x/2}\sin\left[\left(n + \tfrac{1}{2}\right)\pi x\right] e^{-x/2}\sin\left[\left(m + \tfrac{1}{2}\right)\pi x\right] dx \\
&= \begin{cases} 0 & \text{for } m \neq n, \\ \tfrac{1}{2}B_n B_m & \text{for } m = n. \end{cases}
\end{aligned}$$

It is clear that a suitable normalization is $B_n = \sqrt{2}$ for all n.

(b) We write the solution as $y(x) = \sum_{n=0}^{\infty} a_n y_n(x)$, giving as the equation to be solved

$$\begin{aligned}
-e^{x/2} = \mathcal{L}y &= \mathcal{L}\sum_{n=0}^{\infty} a_n y_n(x) \\
&= \sum_{n=0}^{\infty} a_n[\lambda_n \rho(x) y_n(x)] \\
&= \sum_{n=0}^{\infty} a_n\left(n + \tfrac{1}{2}\right)^2\pi^2 e^x \sqrt{2}e^{-x/2}\sin\left[\left(n + \tfrac{1}{2}\right)\pi x\right] \\
\Rightarrow \quad -1 &= \sum_{n=0}^{\infty} a_n\left(n + \tfrac{1}{2}\right)^2\pi^2\sqrt{2}\sin\left[\left(n + \tfrac{1}{2}\right)\pi x\right].
\end{aligned}$$

After multiplying both sides of this equation by $\sin\left(m+\frac{1}{2}\right)\pi x$ and integrating from 0 to 1, we obtain

$$a_m \int_0^1 \sin^2\left(m+\tfrac{1}{2}\right)\pi x\,dx = \frac{-1}{\left(m+\frac{1}{2}\right)^2\pi^2\sqrt{2}} \int_0^1 \sin\left(m+\tfrac{1}{2}\right)\pi x\,dx,$$

$$\frac{a_m}{2} = \frac{-1}{\left(m+\frac{1}{2}\right)^2\pi^2\sqrt{2}} \int_0^1 \sin\left(m+\tfrac{1}{2}\right)\pi x\,dx$$

$$= \frac{1}{\left(m+\frac{1}{2}\right)^2\pi^2\sqrt{2}} \left[\frac{\cos\left(m+\frac{1}{2}\right)\pi x}{\left(m+\frac{1}{2}\right)\pi}\right]_0^1,$$

$$a_m = -\frac{\sqrt{2}}{\left(m+\frac{1}{2}\right)^3\pi^3}.$$

Substituting this result into the assumed expansion, and recalling that $B_n = \sqrt{2}$, gives as the solution

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

8.13 By substituting $x = \exp t$, find the normalized eigenfunctions $y_n(x)$ and the eigenvalues λ_n of the operator $\mathcal{L}$ defined by

$$\mathcal{L}y = x^2y'' + 2xy' + \tfrac{1}{4}y, \qquad 1 \le x \le e,$$

with $y(1) = y(e) = 0$. Find, as a series $\sum a_n y_n(x)$, the solution of $\mathcal{L}y = x^{-1/2}$.

Putting $x = e^t$ and $y(x) = u(t)$ with $u(0) = u(1) = 0$,

$$\frac{dx}{dt} = e^t \quad \Rightarrow \quad \frac{d}{dx} = e^{-t}\frac{d}{dt}$$

and the eigenvalue equation becomes

$$e^{2t}e^{-t}\frac{d}{dt}\left(e^{-t}\frac{du}{dt}\right) + 2e^te^{-t}\frac{du}{dt} + \frac{1}{4}u = \lambda u,$$

$$\frac{d^2u}{dt^2} - \frac{du}{dt} + 2\frac{du}{dt} + \left(\frac{1}{4} - \lambda\right) = 0.$$

The auxiliary equation to this constant-coefficient linear equation for u is

$$m^2 + m + (\tfrac{1}{4} - \lambda) = 0 \quad \Rightarrow \quad m = -\tfrac{1}{2} \pm \sqrt{\lambda},$$

leading to

$$u(t) = e^{-t/2}\left(Ae^{\sqrt{\lambda}\,t} + Be^{-\sqrt{\lambda}\,t}\right).$$

In view of the requirement that u vanishes at two different values of t (one of which is $t = 0$), we need $\lambda < 0$ and $u(t)$ to take the form

$$u(t) = Ae^{-t/2}\sin\sqrt{-\lambda}\,t \text{ with } \sqrt{-\lambda}\,1 = n\pi, \text{ i.e. } \lambda = -n^2\pi^2,$$

where n is an integer. Thus

$$u_n(t) = A_n e^{-t/2} \sin n\pi t \quad \text{or, in terms of } x, \quad y_n(x) = \frac{A_n}{\sqrt{x}} \sin(n\pi \ln x).$$

Normalization requires that

$$1 = \int_1^e \frac{A_n^2}{x} \sin^2(n\pi \ln x)\, dx = \int_0^1 A_n^2 \sin^2(n\pi t)\, dt = \tfrac{1}{2} A_n^2 \quad \Rightarrow \quad A_n = \sqrt{2}.$$

To solve

$$\mathcal{L}y = x^2 y'' + 2xy' + \tfrac{1}{4} y = \frac{1}{\sqrt{x}},$$

we set $y(x) = \sum_{n=0}^{\infty} a_n y_n(x)$. Then the equation becomes

$$\mathcal{L}y = \sum_{n=0}^{\infty} a_n(-n^2\pi^2) y_n(x) = \sum_{n=0}^{\infty} -n^2\pi^2 a_n \frac{\sqrt{2}}{\sqrt{x}} \sin(n\pi \ln x) = \frac{1}{\sqrt{x}}.$$

Multiplying through by $y_m(x)$ and integrating, as with ordinary Fourier series,

$$\int_1^e \frac{2a_n}{x} \sin(n\pi \ln x)\, \sin(m\pi \ln x)\, dx = -\frac{1}{n^2\pi^2} \int_1^e \frac{\sqrt{2}\sin(m\pi \ln x)}{x}\, dx.$$

The LHS of this equation is the normalization integral just considered and has the value $a_m \delta_{mn}$. Thus

$$\begin{aligned} a_m &= -\frac{\sqrt{2}}{m^2\pi^2} \int_1^e \frac{\sin(m\pi \ln x)}{x}\, dx \\ &= -\frac{\sqrt{2}}{m^2\pi^2} \left[\frac{-\cos(m\pi \ln x)}{m\pi} \right]_1^e \\ &= -\frac{\sqrt{2}}{m^3\pi^3} [1 - (-1)^m] \\ &= \begin{cases} -\dfrac{2\sqrt{2}}{m^3\pi^3} & \text{for } m \text{ odd,} \\ 0 & \text{for } m \text{ even.} \end{cases} \end{aligned}$$

The explicit solution is therefore

$$y(x) = -\frac{4}{\pi^3} \sum_{p=0}^{\infty} \frac{\sin[(2p+1)\pi \ln x]}{(2p+1)^3 \sqrt{x}}.$$

8.15 In the quantum mechanical study of the scattering of a particle by a potential, a Born-approximation solution can be obtained in terms of a function $y(\mathbf{r})$ that satisfies an equation of the form

$$(-\nabla^2 - K^2)y(\mathbf{r}) = F(\mathbf{r}).$$

Assuming that $y_{\mathbf{k}}(\mathbf{r}) = (2\pi)^{-3/2}\exp(i\mathbf{k}\cdot\mathbf{r})$ is a suitably normalized eigenfunction of $-\nabla^2$ corresponding to eigenvalue k^2, find a suitable Green's function $G_K(\mathbf{r},\mathbf{r}')$. By taking the direction of the vector $\mathbf{r}-\mathbf{r}'$ as the polar axis for a $\mathbf{k}$-space integration, show that $G_K(\mathbf{r},\mathbf{r}')$ can be reduced to

$$\frac{1}{4\pi^2|\mathbf{r}-\mathbf{r}'|}\int_{-\infty}^{\infty}\frac{w\sin w}{w^2-w_0^2}\,dw,$$

where $w_0 = K|\mathbf{r}-\mathbf{r}'|$.

[This integral can be evaluated using contour integration and gives the Green's function explicitly as $(4\pi|\mathbf{r}-\mathbf{r}'|)^{-1}\exp(iK|\mathbf{r}-\mathbf{r}'|)$.]

Given that $y_{\mathbf{k}}(\mathbf{r}) = (2\pi)^{-3/2}\exp(i\mathbf{k}\cdot\mathbf{r})$ satisfies

$$-\nabla^2 y_{\mathbf{k}}(\mathbf{r}) = k^2 y_{\mathbf{k}}(\mathbf{r}),$$

it follows that

$$(-\nabla^2 - K^2)y_{\mathbf{k}}(\mathbf{r}) = (k^2 - K^2)y_{\mathbf{k}}(\mathbf{r}).$$

Thus the same functions are suitable eigenfunctions for the extended operator, but with different eigenvalues.

Its Green's function is therefore (from the general expression for Green's functions in terms of eigenfunctions)

$$\begin{aligned}
G_K(\mathbf{r},\mathbf{r}') &= \int \frac{1}{\lambda} y_{\mathbf{k}}(\mathbf{r})y_{\mathbf{k}}^*(\mathbf{r}')\,d\mathbf{k} \\
&= \frac{1}{(2\pi)^3}\int \frac{\exp(i\mathbf{k}\cdot\mathbf{r})\exp(-i\mathbf{k}\cdot\mathbf{r}')}{k^2-K^2}\,d\mathbf{k}.
\end{aligned}$$

We carry out the three-dimensional integration in $\mathbf{k}$-space using the direction $\mathbf{r}-\mathbf{r}'$ as the polar axis (and denote $\mathbf{r}-\mathbf{r}'$ by $\mathbf{R}$). The azimuthal integral is immediate. The remaining two-dimensional integration is as follows:

$$\begin{aligned}
G_K(\mathbf{r},\mathbf{r}') &= \frac{1}{(2\pi)^3}\int_0^{\infty}\int_0^{\pi}\frac{\exp(i\mathbf{k}\cdot\mathbf{R})}{k^2-K^2}\,2\pi k^2\sin\theta_k\,d\theta_k\,dk \\
&= \frac{1}{(2\pi)^2}\int_0^{\infty}\int_0^{\pi}\frac{\exp(ikR\cos\theta_k)}{k^2-K^2}\,k^2\sin\theta_k\,d\theta_k\,dk \\
&= \frac{1}{(2\pi)^2}\int_0^{\infty}\frac{\exp(ikR)-\exp(-ikR)}{ikR(k^2-K^2)}\,k^2\,dk \\
&= \frac{1}{2\pi^2 R}\int_0^{\infty}\frac{k\sin kR}{k^2-K^2}\,dk
\end{aligned}$$

$$= \frac{1}{2\pi^2 R} \int_0^\infty \frac{w \sin w}{w^2 - w_0^2}\, dw, \qquad \text{where } w = kR \text{ and } w_0 = kR,$$

$$= \frac{1}{4\pi^2 R} \int_{-\infty}^\infty \frac{w \sin w}{w^2 - w_0^2}\, dw.$$

Here, the final line is justified by noting that the integrand is an even function of the integration variable w.

9 Special functions

9.1 Use the explicit expressions

$$Y_0^0 = \sqrt{\tfrac{1}{4\pi}}, \qquad Y_1^0 = \sqrt{\tfrac{3}{4\pi}}\cos\theta,$$

$$Y_1^{\pm 1} = \mp\sqrt{\tfrac{3}{8\pi}}\sin\theta\exp(\pm i\phi), \qquad Y_2^0 = \sqrt{\tfrac{5}{16\pi}}(3\cos^2\theta - 1),$$

$$Y_2^{\pm 1} = \mp\sqrt{\tfrac{15}{8\pi}}\sin\theta\cos\theta\exp(\pm i\phi), \qquad Y_2^{\pm 2} = \sqrt{\tfrac{15}{32\pi}}\sin^2\theta\exp(\pm 2i\phi),$$

to verify for $\ell = 0, 1, 2$ that

$$\sum_{m=-\ell}^{\ell} \left|Y_\ell^m(\theta,\phi)\right|^2 = \frac{2\ell+1}{4\pi}$$

and so is independent of the values of θ and ϕ. This is true for any ℓ, but a general proof is more involved. This result helps to reconcile intuition with the apparently arbitrary choice of polar axis in a general quantum mechanical system.

We first note that, since every term is the square of a modulus, factors of the form $\exp(\pm mi\phi)$ never appear in the sums. For each value of ℓ, let us denote the sum by S_ℓ. For $\ell = 0$ and $\ell = 1$, we have

$$S_0 = \sum_{m=0}^{0} \left|Y_0^m(\theta,\phi)\right|^2 = \frac{1}{4\pi},$$

$$S_1 = \sum_{m=-1}^{1} \left|Y_1^m(\theta,\phi)\right|^2 = \frac{3}{4\pi}\cos^2\theta + 2\,\frac{3}{8\pi}\sin^2\theta = \frac{3}{4\pi}.$$

For $\ell = 2$, the summation is more complicated but reads

$$\begin{aligned} S_2 &= \sum_{m=-2}^{2} \left|Y_2^m(\theta,\phi)\right|^2 \\ &= \frac{5}{16\pi}(3\cos^2\theta - 1)^2 + 2\,\frac{15}{8\pi}\sin^2\theta\cos^2\theta + 2\,\frac{15}{32\pi}\sin^4\theta \\ &= \frac{5}{16\pi}(9\cos^4\theta - 6\cos^2\theta + 1 + 12\sin^2\theta\cos^2\theta + 3\sin^4\theta) \end{aligned}$$

$$= \frac{5}{16\pi}[6\cos^4\theta - 6\cos^2\theta + 1 + 6\sin^2\theta\cos^2\theta + 3(\cos^2\theta + \sin^2\theta)^2]$$
$$= \frac{5}{16\pi}[6\cos^2\theta(-\sin^2\theta) + 1 + 6\sin^2\theta\cos^2\theta + 3] = \frac{5}{4\pi}.$$

All three sums are independent of θ and ϕ, and are given by the general formula $(2\ell+1)/4\pi$. It will, no doubt, be noted that $2\ell+1$ is the number of terms in S_ℓ, i.e. the number of m values, and that 4π is the total solid angle subtended at the origin by all space.

9.3 Use the generating function for the Legendre polynomials $P_n(x)$ to show that

$$\int_0^1 P_{2n+1}(x)\,dx = (-1)^n \frac{(2n)!}{2^{2n+1}n!(n+1)!}$$

and that, except for the case $n = 0$,

$$\int_0^1 P_{2n}(x)\,dx = 0.$$

Denote $\int_0^1 P_n(x)\,dx$ by a_n. From the generating function for the Legendre polynomials, we have

$$\frac{1}{(1-2xh+h^2)^{1/2}} = \sum_{n=0}^{\infty} P_n(x)h^n.$$

Integrating this definition with respect to x gives

$$\int_0^1 \frac{dx}{(1-2xh+h^2)^{1/2}} = \sum_{n=0}^{\infty}\left(\int_0^1 P_n(x)\,dx\right)h^n,$$
$$\left[\frac{-(1-2xh+h^2)^{1/2}}{h}\right]_0^1 = \sum_{n=0}^{\infty} a_n h^n,$$
$$\frac{1}{h}[(1+h^2)^{1/2} - 1 + h] = \sum_{n=0}^{\infty} a_n h^n.$$

Now expanding $(1+h^2)^{1/2}$ using the binomial theorem yields

$$\sum_{n=0}^{\infty} a_n h^n = \frac{1}{h}\left[1 + \sum_{m=1}^{\infty} {}^{1/2}C_m h^{2m} - 1 + h\right] = 1 + \sum_{m=1}^{\infty} {}^{1/2}C_m h^{2m-1}.$$

Comparison of the coefficients of h^n on the two sides of the equation shows that all a_{2r} are zero except for $a_0 = 1$. For $n = 2r+1$ we need $2m-1 = n = 2r+1$, i.e. $m = r+1$, and the value of a_{2r+1} is ${}^{1/2}C_{r+1}$.

Now, the binomial coefficient ${}^{1/2}C_m$ can be written as

$$
\begin{aligned}
{}^{1/2}C_m &= \frac{\frac{1}{2}\left(\frac{1}{2}-1\right)\left(\frac{1}{2}-2\right)\cdots\left(\frac{1}{2}-m+1\right)}{m!}, \\
&= \frac{1(1-2)(1-4)\cdots(1-2m+2)}{2^m\, m!} \\
&= (-1)^{m-1}\frac{(1)(1)(3)\cdots(2m-3)}{2^m\, m!} \\
&= (-1)^{m-1}\frac{(2m-2)!}{2^m\, m!\, 2^{m-1}\,(m-1)!} \\
&= (-1)^{m-1}\frac{(2m-2)!}{2^{2m-1}\, m!\,(m-1)!}.
\end{aligned}
$$

Thus, setting $m = r + 1$ gives the value of the integral a_{2r+1} as

$$
a_{2r+1} = {}^{1/2}C_{r+1} = (-1)^r \frac{(2r)!}{2^{2r+1}\,(r+1)!\,r!},
$$

as stated in the question.

9.5 The Hermite polynomials $H_n(x)$ may be defined by

$$
\Phi(x, h) = \exp(2xh - h^2) = \sum_{n=0}^{\infty} \frac{1}{n!} H_n(x) h^n.
$$

Show that

$$
\frac{\partial^2 \Phi}{\partial x^2} - 2x\frac{\partial \Phi}{\partial x} + 2h\frac{\partial \Phi}{\partial h} = 0,
$$

and hence that the $H_n(x)$ satisfy the Hermite equation,

$$
y'' - 2xy' + 2ny = 0,
$$

where n is an integer ≥ 0.

Use Φ to prove that

(a) $H_n'(x) = 2nH_{n-1}(x)$,
(b) $H_{n+1}(x) - 2xH_n(x) + 2nH_{n-1}(x) = 0$.

With

$$
\Phi(x, h) = \exp(2xh - h^2) = \sum_{n=0}^{\infty} \frac{1}{n!} H_n(x) h^n,
$$

we have

$$
\frac{\partial \Phi}{\partial x} = 2h\Phi, \qquad \frac{\partial \Phi}{\partial h} = (2x - 2h)\Phi, \qquad \frac{\partial^2 \Phi}{\partial x^2} = 4h^2\Phi.
$$

It then follows that

$$
\frac{\partial^2 \Phi}{\partial x^2} - 2x\frac{\partial \Phi}{\partial x} + 2h\frac{\partial \Phi}{\partial h} = (4h^2 - 4hx + 4hx - 4h^2)\Phi = 0.
$$

Substituting the series form into this result gives

$$\sum_{n=0}^{\infty}\left(\frac{1}{n!}H_n'' - \frac{2x}{n!}H_n' + \frac{2n}{n!}\right)h^n = 0,$$

$$\Rightarrow \quad H_n'' - 2xH_n' + 2nH_n = 0.$$

This is the equation satisfied by $H_n(x)$, as stated in the question.

(a) From the first relationship derived above, we have that

$$\frac{\partial \Phi}{\partial x} = 2h\Phi,$$

$$\sum_{n=0}^{\infty}\frac{1}{n!}H_n'(x)h^n = 2h\sum_{n=0}^{\infty}\frac{1}{n!}H_n(x)h^n,$$

$$\Rightarrow \quad \frac{1}{m!}H_m' = \frac{2}{(m-1)!}H_{m-1}, \text{ from the coefficients of } h^m.$$

Hence, $$H_n'(x) = 2nH_{n-1}(x).$$

(b) Differentiating result (a) and then applying it again yields

$$H_n'' = 2nH_{n-1}' = 2n\,2(n-1)H_{n-2}.$$

Using this in the differential equation satisfied by the H_n, we obtain

$$4n(n-1)H_{n-2} - 2x\,2nH_{n-1} + 2nH_n = 0.$$

This gives

$$H_{n+1}(x) - 2xH_n(x) + 2nH_{n-1}(x) = 0$$

after dividing through by $2n$ and changing $n \to n+1$.

9.7 For the associated Laguerre polynomials, carry out the following:

(a) Prove the Rodrigues' formula

$$L_n^m(x) = \frac{e^x x^{-m}}{n!}\frac{d^n}{dx^n}(x^{n+m}e^{-x}),$$

taking the polynomials to be defined by

$$L_n^m(x) = \sum_{k=0}^{n}(-1)^k\frac{(n+m)!}{k!(n-k)!(k+m)!}x^k.$$

(b) Prove the recurrence relations

$$(n+1)L_{n+1}^m(x) = (2n+m+1-x)L_n^m(x) - (n+m)L_{n-1}^m(x),$$

$$x(L_n^m)'(x) = nL_n^m(x) - (n+m)L_{n-1}^m(x),$$

but this time taking the polynomial as defined by

$$L_n^m(x) = (-1)^m\frac{d^m}{dx^m}L_{n+m}(x)$$

or the generating function.

(a) It is most convenient to evaluate the nth derivative directly, using Leibnitz' theorem. This gives

$$L_n^m(x) = \frac{e^x x^{-m}}{n!} \sum_{r=0}^{n} \frac{n!}{r!(n-r)!} \frac{d^r}{dx^r}(x^{n+m}) \frac{d^{n-r}}{dx^{n-r}}(e^{-x})$$

$$= e^x x^{-m} \sum_{r=0}^{n} \frac{1}{r!(n-r)!} \frac{(n+m)!}{(n+m-r)!} x^{n+m-r} (-1)^{n-r} e^{-x}$$

$$= \sum_{r=0}^{n} \frac{(-1)^{n-r}}{r!(n-r)!} \frac{(n+m)!}{(n+m-r)!} x^{n-r}.$$

Relabeling the summation using the new index $k = n - r$, we immediately obtain

$$L_n^m(x) = \sum_{k=0}^{n} (-1)^k \frac{(n+m)!}{k!(n-k)!(k+m)!} x^k,$$

which is as given in the question.

(b) The first recurrence relation can be proved using the generating function for the associated Laguerre functions:

$$G(x, h) = \frac{e^{-xh/(1-h)}}{(1-h)^{m+1}} = \sum_{n=0}^{\infty} L_n^m(x) h^n.$$

Differentiating the second equality with respect to h, we obtain

$$\frac{(m+1)(1-h) - x}{(1-h)^{m+3}} e^{-xh/(1-h)} = \sum n L_n^m h^{n-1}.$$

Using the generating function for a second time, we may rewrite this as

$$[(m+1)(1-h) - x] \sum L_n^m h^n = (1-h)^2 \sum n L_n^m h^{n-1}.$$

Equating the coefficients of h^n now yields

$$(m+1)L_n^m - (m+1)L_{n-1}^m - xL_n^m = (n+1)L_{n+1}^m - 2nL_n^m + (n-1)L_{n-1}^m,$$

which can be rearranged and simplified to give the first recurrence relation.

The second result is most easily proved by differentiating one of the standard recurrence relations satisfied by the ordinary Laguerre polynomials, but with n replaced by $n + m$. This standard equality reads

$$xL'_{n+m}(x) = (n+m)L_{n+m}(x) - (n+m)L_{n-1+m}(x).$$

We convert this into an equation for the associated polynomials,

$$L_n^m(x) = (-1)^m \frac{d^m}{dx^m} L_{n+m}(x),$$

by differentiating it m times with respect to x and multiplying through by $(-1)^m$. The result is

$$x(L_n^m)' + mL_n^m = (n+m)L_n^m - (n+m)L_{n-1}^m,$$

which immediately simplifies to give the second recurrence relation satisfied by the associated Laguerre polynomials.

9.9 By initially writing $y(x)$ as $x^{1/2}f(x)$ and then making subsequent changes of variable, reduce Stokes' equation,

$$\frac{d^2y}{dx^2} + \lambda xy = 0,$$

to Bessel's equation. Hence show that a solution that is finite at $x = 0$ is a multiple of $x^{1/2}J_{1/3}(\frac{2}{3}\sqrt{\lambda x^3})$.

With $y(x) = x^{1/2}f(x)$,

$$y' = \frac{f}{2x^{1/2}} + x^{1/2}f' \quad \text{and} \quad y'' = -\frac{f}{4x^{3/2}} + \frac{f'}{x^{1/2}} + x^{1/2}f''$$

and the equation becomes

$$-\frac{f}{4x^{3/2}} + \frac{f'}{x^{1/2}} + x^{1/2}f'' + \lambda x^{3/2}f = 0,$$
$$x^2f'' + xf' + (\lambda x^3 - \tfrac{1}{4})f = 0.$$

Now, guided by the known form of Bessel's equation, change the independent variable to $u = x^{3/2}$ with $f(x) = g(u)$ and

$$\frac{du}{dx} = \frac{3}{2}x^{1/2} \quad \Rightarrow \quad \frac{d}{dx} = \frac{3}{2}u^{1/3}\frac{d}{du}.$$

This gives

$$u^{4/3}\frac{3}{2}u^{1/3}\frac{d}{du}\left(\frac{3}{2}u^{1/3}\frac{dg}{du}\right) + u^{2/3}\frac{3}{2}u^{1/3}\frac{dg}{du} + \left(\lambda u^2 - \frac{1}{4}\right)g = 0,$$
$$\frac{3}{2}u^{5/3}\left(\frac{3}{2}u^{1/3}\frac{d^2g}{du^2} + \frac{1}{2}u^{-2/3}\frac{dg}{du}\right) + \frac{3}{2}u\frac{dg}{du} + \left(\lambda u^2 - \frac{1}{4}\right)g = 0,$$
$$\frac{9}{4}u^2\frac{d^2g}{du^2} + \frac{9}{4}u\frac{dg}{du} + \left(\lambda u^2 - \frac{1}{4}\right)g = 0,$$
$$u^2\frac{d^2g}{du^2} + u\frac{dg}{du} + \left(\frac{4}{9}\lambda u^2 - \frac{1}{9}\right)g = 0.$$

This is close to Bessel's equation but still needs a scaling of the variables. So, set $\frac{2}{3}\sqrt{\lambda}u \equiv \mu u = v$ and $g(u) = h(v)$, obtaining

$$\frac{v^2}{\mu^2}\mu^2\frac{d^2h}{dv^2} + \frac{v}{\mu}\mu\frac{dh}{dv} + \left(v^2 - \frac{1}{9}\right)h = 0.$$

This *is* Bessel's equation and has a general solution

$$h(v) = c_1J_{1/3}(v) + c_2J_{-1/3}(v),$$
$$\Rightarrow \quad g(u) = c_1J_{1/3}\left(\tfrac{2\sqrt{\lambda}}{3}u\right) + c_2J_{-1/3}\left(\tfrac{2\sqrt{\lambda}}{3}u\right),$$
$$\Rightarrow \quad f(x) = c_1J_{1/3}\left(\tfrac{2\sqrt{\lambda}}{3}x^{3/2}\right) + c_2J_{-1/3}\left(\tfrac{2\sqrt{\lambda}}{3}x^{3/2}\right).$$

For a solution that is finite at $x = 0$, only the Bessel function with a positive subscript can be accepted. Therefore the required solution is

$$y(x) = c_1 x^{1/2} J_{1/3}(\tfrac{2\sqrt{\lambda}}{3} x^{3/2}).$$

9.11 The complex function $z!$ is defined by

$$z! = \int_0^\infty u^z e^{-u}\, du \qquad \text{for Re } z > -1.$$

For Re $z \le -1$ it is defined by

$$z! = \frac{(z+n)!}{(z+n)(z+n-1)\cdots(z+1)},$$

where n is any (positive) integer $> -$Re z. Being the ratio of two polynomials, $z!$ is analytic everywhere in the finite complex plane except at the poles that occur when z is a negative integer.

(a) Show that the definition of $z!$ for Re $z \le -1$ is independent of the value of n chosen.
(b) Prove that the residue of $z!$ at the pole $z = -m$, where m is an integer > 0, is $(-1)^{m-1}/(m-1)!$.

(a) Let m and n be two choices of integer with $m > n > -$Re z. Denote the corresponding definitions of $z!$ by $(z!)_m$ and $(z!)_n$ and consider the ratio of these two functions:

$$\begin{aligned}
\frac{(z!)_m}{(z!)_n} &= \frac{(z+m)!}{(z+m)(z+m-1)\cdots(z+1)} \, \frac{(z+n)(z+n-1)\cdots(z+1)}{(z+n)!} \\
&= \frac{(z+m)!}{(z+m)(z+m-1)\cdots(z+n+1)\times(z+n)!} \\
&= \frac{(z+m)!}{(z+m)!} = 1.
\end{aligned}$$

Thus the two functions are identical for all z, i.e the definition of $z!$ is independent of the choice of n, provided that $n > -$Re z.

(b) From the given definition of $z!$ it is clear that its pole at $z = -m$ is a simple one. The residue R at the pole is therefore given by

$$\begin{aligned}
R &= \lim_{z\to -m} (z+m) z! \\
&= \lim_{z\to -m} \frac{(z+m)\,(z+n)!}{(z+n)(z+n-1)\cdots(z+1)} \qquad (\text{integer } n \text{ is chosen} > m) \\
&= \lim_{z\to -m} \frac{(z+n)!}{(z+n)(z+n-1)\cdots(z+m+1)(z+m-1)\cdots(z+1)} \\
&= \frac{(-m+n)!}{(-m+n)\cdots(-m+m+1)(-m+m-1)\cdots(-m+1)} \\
&= \frac{1}{[-1]\,[-2]\cdots[-(m-1)]} \\
&= (-1)^{m-1}\frac{1}{(m-1)!},
\end{aligned}$$

as stated in the question.

9.13 The integral

$$I = \int_{-\infty}^{\infty} \frac{e^{-k^2}}{k^2 + a^2}\, dk, \qquad (*)$$

in which $a > 0$, occurs in some statistical mechanics problems. By first considering the integral

$$J = \int_0^{\infty} e^{iu(k+ia)}\, du$$

and a suitable variation of it, show that $I = (\pi/a)\exp(a^2)\,\mathrm{erfc}(a)$, where $\mathrm{erfc}(x)$ is the complementary error function.

The fact that $a > 0$ will ensure that the improper integral J is well defined. It is

$$J = \int_0^{\infty} e^{iu(k+ia)}\, du = \left[\frac{e^{iu(k+ia)}}{i(k+ia)}\right]_0^{\infty} = \frac{i}{k+ia}.$$

We note that this result contains one of the factors that would appear as a denominator in one term of a partial fraction expansion of the integrand in $(*)$. Another term would contain a factor $(k - ia)^{-1}$, and this can be generated by

$$J' = \int_0^{\infty} e^{-iu(k-ia)}\, du = \left[\frac{e^{-iu(k-ia)}}{-i(k-ia)}\right]_0^{\infty} = \frac{-i}{k-ia}.$$

Now, actually expressing the integrand in partial fractions, using the integral expressions J and J' for the factors, and then reversing the order of integration gives

$$\begin{aligned}
I &= \frac{1}{2a}\int_{-\infty}^{\infty}\left(\frac{ie^{-k^2}}{k+ia} - \frac{ie^{-k^2}}{k-ia}\right) dk \\
&= \frac{1}{2a}\int_{-\infty}^{\infty} e^{-k^2}\, dk \int_0^{\infty} e^{iu(k+ia)}\, du + \frac{1}{2a}\int_{-\infty}^{\infty} e^{-k^2}\, dk \int_0^{\infty} e^{-iu(k-ia)}\, du, \\
\Rightarrow \quad 2aI &= \int_0^{\infty} du \int_{-\infty}^{\infty} e^{-k^2+iuk-ua}\, dk + \int_0^{\infty} du \int_{-\infty}^{\infty} e^{-k^2-iuk-ua}\, dk \\
&= \int_0^{\infty} du \int_{-\infty}^{\infty} e^{-(k-iu/2)^2-u^2/4-ua}\, dk \\
&\quad + \int_0^{\infty} du \int_{-\infty}^{\infty} e^{-(k+iu/2)^2-u^2/4-ua}\, dk \\
&= 2\sqrt{\pi}\int_0^{\infty} e^{-u^2/4-ua}\, du,
\end{aligned}$$

using the standard Gaussian result. We now complete the square in the exponent and set $2v = u + 2$, obtaining

$$\begin{aligned}
2aI &= 2\sqrt{\pi}\int_0^{\infty} e^{-(u+2a)^2/4+a^2}\, du, \\
&= 2\sqrt{\pi}\int_a^{\infty} e^{-v^2}\, e^{a^2}\, 2dv.
\end{aligned}$$

From this it follows that

$$I = \frac{\sqrt{\pi}}{a}\, 2e^{a^2}\, \frac{\sqrt{\pi}}{2}\, \mathrm{erfc}(a) = \frac{\pi}{a}\, e^{a^2}\, \mathrm{erfc}(a),$$

as stated in the question.

9.15 Prove two of the properties of the incomplete gamma function $P(a, x^2)$ as follows.

(a) By considering its form for a suitable value of a, show that the error function can be expressed as a particular case of the incomplete gamma function.

(b) The Fresnel integrals, of importance in the study of the diffraction of light, are given by

$$C(x) = \int_0^x \cos\left(\frac{\pi}{2}t^2\right) dt, \qquad S(x) = \int_0^x \sin\left(\frac{\pi}{2}t^2\right) dt.$$

Show that they can be expressed in terms of the error function by

$$C(x) + iS(x) = A\,\mathrm{erf}\left[\frac{\sqrt{\pi}}{2}(1-i)x\right],$$

where A is a (complex) constant, which you should determine. Hence express $C(x) + iS(x)$ in terms of the incomplete gamma function.

(a) From the definition of the incomplete gamma function, we have

$$P(a, x^2) = \frac{1}{\Gamma(a)} \int_0^{x^2} e^{-t}\, t^{a-1}\, dt.$$

Guided by the x^2 in the upper limit, we now change the integration variable to $y = +\sqrt{t}$, with $2y\,dy = dt$, and obtain

$$P(a, x^2) = \frac{1}{\Gamma(a)} \int_0^{x} e^{-y^2}\, y^{2(a-1)}\, 2y\, dy.$$

To make the RHS into an error function we need to remove the y-term; to do this we choose a such that $2(a-1)+1 = 0$, i.e. $a = \frac{1}{2}$. With this choice, $\Gamma(a) = \sqrt{\pi}$ and

$$P\left(\tfrac{1}{2}, x^2\right) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-y^2}\, dy,$$

i.e. a correctly normalized error function.

(b) Consider the given expression:

$$z = A\,\mathrm{erf}\left[\frac{\sqrt{\pi}}{2}(1-i)x\right] = \frac{2A}{\sqrt{\pi}} \int_0^{\sqrt{\pi}(1-i)x/2} e^{-u^2}\, du.$$

Changing the variable of integration to s, given by $u = \frac{1}{2}\sqrt{\pi}(1-i)s$, and recalling that $(1-i)^2 = -2i$, we obtain

$$\begin{aligned} z &= \frac{2A}{\sqrt{\pi}} \int_0^x e^{-s^2\pi(-2i)/4} \frac{\sqrt{\pi}}{2}(1-i)\,ds \\ &= A(1-i)\int_0^x e^{i\pi s^2/2}\,ds \\ &= A(1-i)\int_0^x \left[\cos\left(\frac{\pi s^2}{2}\right) + i\sin\left(\frac{\pi s^2}{2}\right)\right] ds \\ &= A(1-i)\,[C(x) + iS(x)]. \end{aligned}$$

For the correct normalization we need $A(1-i) = 1$, implying that $A = (1+i)/2$.

Now, from part (a), the error function can be expressed in terms of the incomplete gamma function $P(a, x)$ by

$$\mathrm{erf}(x) = P\left(\tfrac{1}{2}, x^2\right).$$

Here the argument of the error function is $\frac{1}{2}\sqrt{\pi}(1-i)x$, whose square is $-\frac{1}{2}\pi i x^2$, and so

$$C(x) + iS(x) = \frac{1+i}{2}\,P\left(\frac{1}{2}, -\frac{i\pi}{2}x^2\right).$$

10 Partial differential equations

10.1 Determine whether the following can be written as functions of $p = x^2 + 2y$ only, and hence whether they are solutions of

$$\frac{\partial u}{\partial x} = x\frac{\partial u}{\partial y}.$$

(a) $x^2(x^2 - 4) + 4y(x^2 - 2) + 4(y^2 - 1)$;
(b) $x^4 + 2x^2y + y^2$;
(c) $[x^4 + 4x^2y + 4y^2 + 4]/[2x^4 + x^2(8y + 1) + 8y^2 + 2y]$.

As a first step, we verify that any function of $p = x^2 + 2y$ will satisfy the given equation. Using the chain rule, we have

$$\frac{\partial u}{\partial p}\frac{\partial p}{\partial x} = x\frac{\partial u}{\partial p}\frac{\partial p}{\partial y},$$
$$\Rightarrow \quad \frac{\partial u}{\partial p}2x = x\frac{\partial u}{\partial p}2.$$

This is satisfied for *any* function $u(p)$, thus completing the verification.

To test the given functions we substitute for $y = \frac{1}{2}(p - x^2)$ or for $x^2 = p - 2y$ in each of the $f(x, y)$ and then examine whether the resulting forms are independent of x or y, respectively.

$$\begin{aligned}
\text{(a) } f(x, y) &= x^2(x^2 - 4) + 4y(x^2 - 2) + 4(y^2 - 1)\\
&= x^2(x^2 - 4) + 2(p - x^2)(x^2 - 2) + p^2 - 2px^2 + x^4 - 4\\
&= x^4(1 - 2 + 1) + x^2(-4 + 2p + 4 - 2p) - 4p + p^2 - 4\\
&= p^2 - 4p - 4 = g(p).
\end{aligned}$$

This is a function of p only, and therefore the original $f(x, y)$ is a solution of the PDE.

Though not necessary for answering the question, we will repeat the verification, but this time by substituting for x rather than for y:

$$\begin{aligned}
f(x, y) &= x^2(x^2 - 4) + 4y(x^2 - 2) + 4(y^2 - 1)\\
&= (p - 2y)(p - 2y - 4) + 4y(p - 2y - 2) + 4(y^2 - 1)\\
&= p^2 - 4py + 4y^2 - 4p + 8y + 4yp - 8y^2 - 8y + 4y^2 - 4\\
&= p^2 - 4p - 4 = g(p);
\end{aligned}$$

i.e. it is the same as before, as it must be, and again this shows that $f(x, y)$ is a solution of the PDE.

$$\begin{aligned}\text{(b) } f(x, y) &= x^4 + 2x^2 y + y^2 \\ &= (p - 2y)^2 + 2y(p - 2y) + y^2 \\ &= p^2 - 4p\,y + 4y^2 + 2p\,y - 4y^2 + y^2 \\ &= (p - y)^2 \neq g(p).\end{aligned}$$

As this is a function of both p and y, it is not a solution of the PDE.

$$\begin{aligned}\text{(c) } f(x, y) &= \frac{x^4 + 4x^2 y + 4y^2 + 4}{2x^4 + x^2(8y + 1) + 8y^2 + 2y} \\ &= \frac{p^2 - 4p\,y + 4y^2 + 4yp - 8y^2 + 4y^2 + 4}{2p^2 - 8p\,y + 8y^2 + 8yp + p - 16y^2 - 2y + 8y^2 + 2y} \\ &= \frac{p^2 + 4}{2p^2 + p} = g(p).\end{aligned}$$

This is a function of p only and therefore $f(x, y)$ is a solution of the PDE.

10.3 Solve the following partial differential equations for $u(x, y)$ with the boundary conditions given:

(a) $x\dfrac{\partial u}{\partial x} + xy = u, \qquad u = 2y$ on the line $x = 1$;

(b) $1 + x\dfrac{\partial u}{\partial y} = xu, \qquad u(x, 0) = x.$

(a) This can be solved as an ODE for u as a function of x, though the "constant of integration" will be a function of y. In standard form, the equation reads

$$\frac{\partial u}{\partial x} - \frac{u}{x} = -y.$$

By inspection (or formal calculation) the IF for this is x^{-1} and the equation can be rearranged as

$$\begin{aligned}\frac{\partial}{\partial x}\left(\frac{u}{x}\right) &= -\frac{y}{x}, \\ \Rightarrow \quad \frac{u}{x} &= -y \ln x + f(y), \\ u = 2y \text{ on } x = 1 \ \Rightarrow f(y) &= 2y, \\ \text{and so } u(x, y) &= xy(2 - \ln x).\end{aligned}$$

(b) This equation can be written in standard form, with u as a function of y:

$$\frac{\partial u}{\partial y} - u = -\frac{1}{x},$$

for which the IF is clearly e^{-y}, leading to

$$\frac{\partial}{\partial y}\left(e^{-y}u\right) = -\frac{e^{-y}}{x},$$

$$\Rightarrow \quad e^{-y}u = \frac{e^{-y}}{x} + f(x),$$

$$u(x,0) = x \Rightarrow f(x) = x - \frac{1}{x}.$$

Substituting this result, and multiplying through by e^{y}, gives $u(x, y)$ as

$$u(x, y) = \frac{1}{x} + \left(x - \frac{1}{x}\right)e^{y}.$$

10.5 Find solutions of

$$\frac{1}{x}\frac{\partial u}{\partial x} + \frac{1}{y}\frac{\partial u}{\partial y} = 0$$

for which (a) $u(0, y) = y$ and (b) $u(1, 1) = 1$.

As usual, we find $p(x, y)$ from

$$\frac{dx}{x^{-1}} = \frac{dy}{y^{-1}} \quad \Rightarrow \quad x^2 - y^2 = p.$$

(a) On $x = 0$, $p = -y^2$ and

$$u(0, y) = y = (-p)^{1/2} \quad \Rightarrow \quad u(x, y) = [-(x^2 - y^2)]^{1/2} = (y^2 - x^2)^{1/2}.$$

(b) At $(1, 1)$, $p = 0$ and

$$u(1, 1) = 1 \quad \Rightarrow \quad u(x, y) = 1 + g(x^2 - y^2),$$

where g is any function that has $g(0) = 0$.

We note that in part (a) the solution is uniquely determined because the boundary values are given along a line, whereas in part (b), where the value is fixed at only an isolated point, the solution is indeterminate to the extent of a loosely determined function. This is the normal situation, though it is modified if the boundary line in (a) happens to be one along which p has a constant value.

10.7 Solve

$$\sin x\frac{\partial u}{\partial x} + \cos x\frac{\partial u}{\partial y} = \cos x \qquad (*)$$

subject to (a) $u(\pi/2, y) = 0$ and (b) $u(\pi/2, y) = y(y + 1)$.

As usual, the CF is found from

$$\frac{dx}{\sin x} = \frac{dy}{\cos x} \quad \Rightarrow \quad y - \ln\sin x = p.$$

Since the RHS of $(*)$ is a factor in one of the terms on the LHS, a trivial PI is any function of y only whose derivative (with respect to y) is unity, of which the simplest is $u(x, y) = y$. The general solution is therefore

$$u(x, y) = f(y - \ln \sin x) + y.$$

The actual form of the arbitrary function $f(p)$ is determined by the form that $u(x, y)$ takes on the boundary, here the line $x = \pi/2$.

(a) With $u(\pi/2, y) = 0$:

$$\begin{aligned} 0 &= f(y - 0) + y \quad \Rightarrow \quad f(p) = -p \\ \Rightarrow \quad u(x, y) &= \ln \sin x - y + y = \ln \sin x. \end{aligned}$$

(b) With $u(\pi/2, y) = y(y + 1)$:

$$\begin{aligned} y(y + 1) &= f(y - 0) + y \quad \Rightarrow \quad f(p) = p^2 \\ \Rightarrow \quad u(x, y) &= (y - \ln \sin x)^2 + y. \end{aligned}$$

10.9 If $u(x, y)$ satisfies

$$\frac{\partial^2 u}{\partial x^2} - 3\frac{\partial^2 u}{\partial x \partial y} + 2\frac{\partial^2 u}{\partial y^2} = 0$$

and $u = -x^2$ and $\partial u/\partial y = 0$ for $y = 0$ and all x, find the value of $u(0, 1)$.

If we are to find solutions to this homogeneous second-order PDE of the form $u(x, y) = f(x + \lambda y)$, then λ must satisfy

$$1 - 3\lambda + 2\lambda^2 = 0 \quad \Rightarrow \quad \lambda = \tfrac{1}{2},\ 1.$$

Thus $u(x, y) = g(x + \frac{1}{2}y) + f(x + y) \equiv g(p_1) + f(p_2)$.

On $y = 0$, $p_1 = p_2 = x$ and

$$\begin{aligned} -x^2 &= u(x, 0) = g(x) + f(x), \qquad (*) \\ 0 &= \frac{\partial u}{\partial y}(x, 0) = \tfrac{1}{2}g'(x) + f'(x). \end{aligned}$$

$$\begin{aligned} &\text{From } (*), & -2x &= g'(x) + f'(x). \\ &\text{Subtracting,} & 2x &= -\tfrac{1}{2}g'(x). \\ &\text{Integrating,} & g(x) &= -2x^2 + k \quad \Rightarrow \quad f(x) = x^2 - k, \quad \text{from } (*). \\ &\text{Hence,} & u(x, y) &= -2\left(x + \tfrac{1}{2}y\right)^2 + k + (x + y)^2 - k \\ && &= -x^2 + \tfrac{1}{2}y^2. \end{aligned}$$

At the particular point $(0, 1)$ we have $u(0, 1) = -0^2 + \frac{1}{2}(1)^2 = \frac{1}{2}$.

10.11 In those cases in which it is possible to do so, evaluate $u(2, 2)$, where $u(x, y)$ is the solution of

$$2y\frac{\partial u}{\partial x} - x\frac{\partial u}{\partial y} = xy(2y^2 - x^2)$$

that satisfies the (separate) boundary conditions given below.

(a) $u(x, 1) = x^2$.
(b) $u(1, \sqrt{10}) = 5$.
(c) $u(\sqrt{10}, 1) = 5$.

To find the CF, $u(x, y) = f(p)$, we set

$$\frac{dx}{2y} = -\frac{dy}{x} \quad \Rightarrow \quad x^2 + 2y^2 = p.$$

The point $(2, 2)$ corresponds to $p = 2^2 + 2(2^2) = 12$.

For a PI we try $u(x, y) = Ax^n y^m$:

$$2Anx^{n-1}y^{m+1} - Amx^{n+1}y^{m-1} = 2xy^3 - x^3y,$$

which has a solution, $n = m = 2$ with $A = \frac{1}{2}$. Thus the general solution is

$$u(x, y) = f(x^2 + 2y^2) + \tfrac{1}{2}x^2y^2.$$

(a) We must find the function f that makes $u(x, 1) = x^2$. This requires f to satisfy

$$\begin{aligned} x^2 = u(x, 1) &= f(x^2 + 2) + \tfrac{1}{2}x^2 \\ \Rightarrow \qquad f(p) &= \tfrac{1}{2}(p - 2) \\ \Rightarrow \quad u(x, y) &= \tfrac{1}{2}(x^2 + 2y^2 - 2) + \tfrac{1}{2}x^2y^2 \\ &= \tfrac{1}{2}(x^2 + x^2y^2 + 2y^2 - 2). \end{aligned}$$

From which it follows that $u(2, 2) = \frac{1}{2}(4 + 16 + 8 - 2) = 13$.

(b) With $u(1, \sqrt{10}) = 5$: At the point $(1, \sqrt{10})$ the value of p is $1 + 2(10) = 21$. As the "boundary" consists of just this one point, it is only at the points that have $p = 21$ that the value of $u(x, y)$ can be known. Since for the point $(2, 2)$ the value of p is 12, the value of $u(2, 2)$ cannot be determined.

(c) With $u(\sqrt{10}, 1) = 5$: At the point $(\sqrt{10}, 1)$ the value of p is $10 + 2(1) = 12$. Since for $(2, 2)$ it is also 12, the value of $u(2, 2)$ can be determined and is given by $f(12) + \frac{1}{2}(4)(4) = 5 + 8 = 13$.

10.13 Find the most general solution of $\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} = x^2y^2$.

The complementary function for this equation is the solution to the two-dimensional Laplace equation and [either as a general known result or from substituting the trial form $h(x + \lambda y)$ which leads to $\lambda^2 = -1$ and hence to $\lambda = \pm i$] has the form $f(x + iy) + g(x - iy)$ for arbitrary functions f and g.

It therefore remains only to find a suitable PI. As f and g are not specified, there are infinitely many possibilities and which one we finish up with will depend upon the details of the approach adopted. When a solution has been obtained it should be checked by substitution.

As no PI is obvious by inspection, we make a change of variables with the object of obtaining one by means of an explicit integration. To do this, we use as new variables the arguments of the arbitrary functions appearing in the CF.

Setting $\xi = x + iy$ and $\eta = x - iy$, with $u(x, y) = v(\xi, \eta)$, gives

$$\begin{aligned}
&\left(\frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta}\right)\left(\frac{\partial v}{\partial \xi} + \frac{\partial v}{\partial \eta}\right) \\
&\quad + \left(i\frac{\partial}{\partial \xi} - i\frac{\partial}{\partial \eta}\right)\left(i\frac{\partial v}{\partial \xi} - i\frac{\partial v}{\partial \eta}\right) = \left(\frac{\xi + \eta}{2}\right)^2\left(\frac{\xi - \eta}{2i}\right)^2, \\
&(1-1)\frac{\partial^2 v}{\partial \xi^2} + (2+2)\frac{\partial^2 v}{\partial \xi \partial \eta} + (1-1)\frac{\partial^2 v}{\partial \eta^2} = -\frac{1}{16}(\xi^2 - \eta^2)^2, \\
&\frac{\partial^2 v}{\partial \xi \partial \eta} = -\frac{1}{64}(\xi^2 - \eta^2)^2.
\end{aligned}$$

When we integrate this we can set all constants of integration and all arbitrary functions equal to zero as *any* solution will suffice:

$$\begin{aligned}
\frac{\partial^2 v}{\partial \xi \partial \eta} &= -\frac{1}{64}(\xi^4 - 2\xi^2\eta^2 + \eta^4), \\
\frac{\partial v}{\partial \eta} &= -\frac{1}{64}\left(\frac{\xi^5}{5} - \frac{2\xi^3\eta^2}{3} + \xi\eta^4\right), \\
v &= -\frac{1}{64}\left(\frac{\xi^5\eta}{5} - \frac{2\xi^3\eta^3}{9} + \frac{\xi\eta^5}{5}\right).
\end{aligned}$$

Re-expressing this solution as a function of x and y (noting that $\xi\eta = x^2 + y^2$) gives

$$\begin{aligned}
u(x, y) &= \frac{1}{(64)(45)}[10\xi^3\eta^3 - 9\xi\eta(\xi^4 + \eta^4)] \\
&= \frac{1}{(64)(45)}[10(x^2 + y^2)^3 - 18(x^2 + y^2)(x^4 - 6x^2y^2 + y^4)] \\
&= \frac{x^2 + y^2}{(64)(45)}(10x^4 + 20x^2y^2 + 10y^4 - 18x^4 + 108x^2y^2 - 18y^4) \\
&= \frac{x^2 + y^2}{(64)(45)}(128x^2y^2 - 8x^4 - 8y^4) \\
&= \frac{1}{360}(15x^4y^2 - x^6 + 15x^2y^4 - y^6).
\end{aligned}$$

Check

Applying $\dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2}$ to the final expression yields

$$\frac{1}{360}[15(12)x^2y^2 - 30x^4 + 30y^4 + 30x^4 + 15(12)x^2y^2 - 30y^4] = x^2y^2,$$

as it should.

10.15 The non-relativistic Schrödinger equation,

$$-\frac{\hbar^2}{2m}\nabla^2 u + V(\mathbf{r})u = i\hbar\frac{\partial u}{\partial t},$$

is similar to the diffusion equation in having different orders of derivatives in its various terms; this precludes solutions that are arbitrary functions of particular linear combinations of variables. However, since exponential functions do not change their forms under differentiation, solutions in the form of exponential functions of combinations of the variables may still be possible.

Consider the Schrödinger equation for the case of a constant potential, i.e. for a free particle, and show that it has solutions of the form $A\exp(lx + my + nz + \lambda t)$, where the only requirement is that

$$-\frac{\hbar^2}{2m}\left(l^2 + m^2 + n^2\right) = i\hbar\lambda.$$

In particular, identify the equation and wavefunction obtained by taking λ as $-iE/\hbar$, and l, m and n as $ip_x/\hbar$, $ip_y/\hbar$ and $ip_z/\hbar$, respectively, where E is the energy and $\mathbf{p}$ the momentum of the particle; these identifications are essentially the content of the de Broglie and Einstein relationships.

For a free particle we may omit the potential term $V(\mathbf{r})$ from the Schrödinger equation, which then reads (in Cartesian coordinates)

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right) = i\hbar\frac{\partial u}{\partial t}.$$

We try $u(x, y, z, t) = A\exp(lx + my + nz + \lambda t)$, i.e. the product of four exponential functions, and obtain

$$-\frac{\hbar^2}{2m}(l^2 + m^2 + n^2)u = i\hbar\lambda u.$$

This equation is clearly satisfied provided

$$-\frac{\hbar^2}{2m}(l^2 + m^2 + n^2) = i\hbar\lambda.$$

With λ as $-iE/\hbar$, and l, m and n as $ip_x/\hbar$, $ip_y/\hbar$ and $ip_z/\hbar$, respectively, where E is the energy and $\mathbf{p}$ is the momentum of the particle, we have

$$-\frac{\hbar^2}{2m}\left(-\frac{p_x^2}{\hbar^2} - \frac{p_y^2}{\hbar^2} - \frac{p_z^2}{\hbar^2}\right) = E,$$

which can be written more compactly as $E = p^2/2m$, the classical non-relativistic relationship between the (kinetic) energy and momentum of a free particle.

The wavefunction obtained is

$$\begin{aligned}u(\mathbf{r}, t) &= A\exp\left[\frac{i}{\hbar}(p_x x + p_y y + p_z z - Et)\right]\\ &= A\exp\left[\frac{i}{\hbar}(\mathbf{p}\cdot\mathbf{r} - Et)\right],\end{aligned}$$

i.e. a classical plane wave of wave number $\mathbf{k} = \mathbf{p}/\hbar$ and angular frequency $\omega = E/\hbar$ traveling in the direction $\mathbf{p}/p$.

10.17 An incompressible fluid of density ρ and negligible viscosity flows with velocity v along a thin, straight, perfectly light and flexible tube, of cross-section A which is held under tension T. Assume that small transverse displacements u of the tube are governed by

$$\frac{\partial^2 u}{\partial t^2} + 2v\frac{\partial^2 u}{\partial x \partial t} + \left(v^2 - \frac{T}{\rho A}\right)\frac{\partial^2 u}{\partial x^2} = 0.$$

(a) Show that the general solution consists of a superposition of two waveforms traveling with different speeds.

(b) The tube initially has a small transverse displacement $u = a\cos kx$ and is suddenly released from rest. Find its subsequent motion.

(a) This is a second-order equation and will (in general) have two solutions of the form $u(x,t) = f(x + \lambda t)$, where both λ satisfy

$$\lambda^2 + 2v\lambda + \left(v^2 - \frac{T}{\rho A}\right) = 0 \quad \Rightarrow \quad \lambda = -v \pm \sqrt{v^2 - v^2 + \frac{T}{\rho A}} \equiv -v \pm \alpha,$$

and gives (minus) the speed of the corresponding profile. Thus the general displacement consists of a superposition of waveforms traveling with speeds $v \mp \alpha$.

(b) Now $u(x,0) = a\cos kx$ and $\dot{u}(x,0) = 0$, where the dot denotes differentiation with respect to time t. Let the general solution be given by

$$u(x,t) = f[x - (v+\alpha)t] + g[x - (v-\alpha)t],$$

with $$a\cos kx = f(x) + g(x)$$

and $$0 = -(v+\alpha)f'(x) - (v-\alpha)g'(x).$$

We differentiate the first of these with respect to x and then eliminate the function $f'(x)$:

$$\begin{aligned}
-ka\sin kx &= f'(x) + g'(x),\\
-ka(v+\alpha)\sin kx &= (v + \alpha - v + \alpha)g'(x),\\
g'(x) &= -\frac{ka(v+\alpha)}{2\alpha}\sin kx,\\
\Rightarrow \quad g(x) &= \frac{v+\alpha}{2\alpha}a\cos kx + c,\\
\Rightarrow \quad f(x) &= \frac{\alpha - v}{2\alpha}a\cos kx - c.
\end{aligned}$$

Now that the forms of the initially arbitrary functions $f(x)$ and $g(x)$ have been determined, it follows that, for a *general* time t,

$$\begin{aligned}
u(x,t) &= \frac{\alpha - v}{2\alpha}a\cos[kx - k(v+\alpha)t] + \frac{\alpha + v}{2\alpha}a\cos[kx - k(v-\alpha)t]\\
&= \frac{a}{2}2\cos(kx - kvt)\cos k\alpha t + \frac{va}{2\alpha}2\sin(kx - kvt)\sin(-k\alpha t)\\
&= a\cos[k(x - vt)]\cos k\alpha t - \frac{va}{\alpha}\sin[k(x - vt)]\sin k\alpha t.
\end{aligned}$$

10.19 In an electrical cable of resistance R and capacitance C, each per unit length, voltage signals obey the equation $\partial^2 V/\partial x^2 = RC\partial V/\partial t$. This (diffusion-type) equation has solutions of the form

$$f(\zeta) = \frac{2}{\sqrt{\pi}} \int_0^{\zeta} \exp(-\nu^2)\, d\nu, \text{ where } \zeta = \frac{x(RC)^{1/2}}{2t^{1/2}}.$$

It also has solutions of the form $V = Ax + D$.

(a) Find a combination of these that represents the situation after a steady voltage V_0 is applied at $x = 0$ at time $t = 0$.
(b) Obtain a solution describing the propagation of the voltage signal resulting from the application of the signal $V = V_0$ for $0 < t < T$, $V = 0$ otherwise, to the end $x = 0$ of an infinite cable.
(c) Show that for $t \gg T$ the maximum signal occurs at a value of x proportional to $t^{1/2}$ and has a magnitude proportional to t^{-1}.

(a) Consider the given function

$$f(\zeta) = \frac{2}{\sqrt{\pi}} \int_0^{\zeta} \exp(-\nu^2)\, d\nu, \text{ where } \zeta = \frac{x(RC)^{1/2}}{2t^{1/2}}.$$

The requirements to be satisfied by the correct combination of this function and $V(x,t) = Ax + D$ are (i) that, at $t = 0$, V is zero for all x, except $x = 0$ where it is V_0, and (ii) that, as $t \to \infty$, V is V_0 for all x.

(i) At $t = 0$, $\zeta = \infty$ and $f(\zeta) = 1$ for all $x \neq 0$.
(ii) As $t \to \infty$, $\zeta \to 0$ and $f(\zeta) \to 0$ for all finite x.
The required combination is therefore $D = V_0$ and $-V_0 f(\zeta)$, i.e.

$$V(x,t) = V_0 \left[1 - \frac{2}{\sqrt{\pi}} \int_0^{\frac{1}{2}x(CR/t)^{1/2}} \exp(-\nu^2)\, d\nu \right].$$

(b) The equation is linear and so we may superpose solutions. The response to the input $V = V_0$ for $0 < t < T$ can be considered as that to V_0 applied at $t = 0$ and continued, together with $-V_0$ applied at $t = T$ and continued. The solution is therefore the difference between two solutions of the form found in part (a):

$$V(x,t) = \frac{2V_0}{\sqrt{\pi}} \int_{\frac{1}{2}x(CR/t)^{1/2}}^{\frac{1}{2}x[CR/(t-T)]^{1/2}} \exp\left(-\nu^2\right)\, d\nu.$$

(c) To find the maximum signal we set $\partial V/\partial x$ equal to zero. Remembering that we are differentiating with respect to the limits of an integral (whose integrand does not contain x explicitly), we obtain

$$\frac{1}{2}\left(\frac{CR}{t-T}\right)^{1/2} \exp\left[-\frac{x^2CR}{4(t-T)}\right] - \frac{1}{2}\left(\frac{CR}{t}\right)^{1/2} \exp\left[-\frac{x^2CR}{4t}\right] = 0.$$

This requires

$$\begin{aligned}\left(\frac{t-T}{t}\right)^{1/2} &= \exp\left[-\frac{x^2CR}{4(t-T)} + \frac{x^2CR}{4t}\right] \\ &= \exp\left[\frac{x^2CR(-t+t-T)}{4t(t-T)}\right].\end{aligned}$$

For $t \gg T$, we expand both sides:

$$1 - \frac{1}{2}\frac{T}{t} + \cdots = 1 - \frac{Tx^2CR}{4t^2} + \cdots,$$

$$\Rightarrow \quad x^2 \approx \frac{2t}{CR} \quad \Rightarrow \quad \nu = \frac{1}{2}\sqrt{\frac{2t}{CR}}\left(\frac{CR}{t}\right)^{1/2} = \frac{1}{\sqrt{2}}.$$

The corresponding value of V is approximately equal to the value of the integrand, evaluated at this value of ν, multiplied by the difference between the two limits of the integral. Thus

$$\begin{aligned} V_{\max} &\approx \frac{2V_0}{\sqrt{\pi}} \exp(-\nu^2) \frac{x\sqrt{CR}}{2} \left[\frac{1}{(t-T)^{1/2}} - \frac{1}{t^{1/2}} \right] \\ &\approx \frac{2V_0}{\sqrt{\pi}} e^{-1/2} \frac{x\sqrt{CR}}{2} \frac{1}{2} \frac{T}{t^{3/2}} \\ &= \frac{V_0 T e^{-1/2}}{\sqrt{2\pi}\, t}. \end{aligned}$$

In summary, for $t \gg T$ the maximum signal occurs at a value of x proportional to $t^{1/2}$ and has a magnitude proportional to t^{-1}.

10.21 Consider each of the following situations in a qualitative way and determine the equation type, the nature of the boundary curve and the type of boundary conditions involved:

(a) a conducting bar given an initial temperature distribution and then thermally isolated;
(b) two long conducting concentric cylinders, on each of which the voltage distribution is specified;
(c) two long conducting concentric cylinders, on each of which the charge distribution is specified;
(d) a semi-infinite string, the end of which is made to move in a prescribed way.

We use the notation

$$A\frac{\partial^2 u}{\partial x^2} + B\frac{\partial^2 u}{\partial x \partial y} + C\frac{\partial^2 u}{\partial y^2} + D\frac{\partial u}{\partial x} + E\frac{\partial u}{\partial y} + Fu = R(x, y)$$

to express the most general type of PDE, and the following table

Equation type	Boundary	Conditions
hyperbolic	open	Cauchy
parabolic	open	Dirichlet or Neumann
elliptic	closed	Dirichlet or Neumann

to determine the appropriate boundary type and hence conditions.

(a) The diffusion equation $\kappa \dfrac{\partial^2 T}{\partial x^2} = \dfrac{\partial T}{\partial t}$ has $A = \kappa$, $B = 0$ and $C = 0$; thus $B^2 = 4AC$ and the equation is parabolic. This needs an open boundary. In the present case, the initial heat distribution (at the $t = 0$ boundary) is a Dirichlet condition and the insulation (no temperature gradient at the external surfaces) is a Neumann condition.

(b) The governing equation in two-dimensional Cartesians (not the natural choice for this situation, but this does not matter for the present purpose) is the Laplace equation, $\dfrac{\partial^2\phi}{\partial x^2}+\dfrac{\partial^2\phi}{\partial y^2}=0$, which has $A=1$, $B=0$ and $C=1$ and therefore $B^2<4AC$. The equation is therefore elliptic and requires a closed boundary. Since ϕ is specified on the cylinders, the boundary conditions are Dirichlet in this particular situation.

(c) This is the same as part (b) except that the specified charge distribution σ determines $\partial\phi/\partial n$, through $\partial\phi/\partial n=\sigma/\epsilon_0$, and imposes Neumann boundary conditions.

(d) For the wave equation $\dfrac{\partial^2 u}{\partial x^2}-\dfrac{1}{c^2}\dfrac{\partial^2 u}{\partial t^2}=0$, we have $A=1$, $B=0$ and $C=-c^{-2}$, thus making $B^2>4AC$ and the equation hyperbolic. We thus require an open boundary and Cauchy conditions, with the displacement of the end of the string having to be specified at all times – this is equivalent to the displacement and the velocity of the end of the string being specified at all times.

10.23 The Klein–Gordon equation (which is satisfied by the quantum-mechanical wavefunction $\Phi(\mathbf{r})$ of a relativistic spinless particle of non-zero mass m) is

$$\nabla^2\Phi-m^2\Phi=0.$$

Show that the solution for the scalar field $\Phi(\mathbf{r})$ in any volume V bounded by a surface S is unique if either Dirichlet or Neumann boundary conditions are specified on S.

Suppose that, for a given set of boundary conditions ($\Phi=f$ or $\partial\Phi/\partial n=g$ on S), there are two solutions to the Klein–Gordon equation, Φ_1 and Φ_2. Then consider $\Phi_3=\Phi_1-\Phi_2$, which satisfies

$$\nabla^2\Phi_3=\nabla^2\Phi_1-\nabla^2\Phi_2=m^2\Phi_1-m^2\Phi_2=m^2\Phi_3$$

and

$$\text{either } \Phi_3=f-f=0,\ \text{ or } \frac{\partial\Phi_3}{\partial n}=g-g=0 \text{ on } S.$$

Now apply Green's first theorem with the scalar functions equal to Φ_3 and Φ_3^*:

$$\int^S \Phi_3^*\frac{\partial\Phi_3}{\partial n}\,dS=\int_V[\Phi_3^*\nabla^2\Phi_3+(\nabla\Phi_3^*)\cdot(\nabla\Phi_3)]\,dV,$$

$$\Rightarrow\quad 0=\int_V(m^2|\Phi_3|^2+|\nabla\Phi_3|^2)\,dV,$$

whichever set of boundary conditions applies. Since both terms in the integrand on the RHS are non-negative, each must be equal to zero. In particular, $|\Phi_3|=0$ implies that $\Phi_3=0$ everywhere, i.e. $\Phi_1=\Phi_2$ everywhere; the solution is unique.

11 Solution methods for PDEs

11.1 Solve the following first-order partial differential equations by separating the variables:

$$\text{(a)}\ \frac{\partial u}{\partial x} - x\frac{\partial u}{\partial y} = 0; \qquad \text{(b)}\ x\frac{\partial u}{\partial x} - 2y\frac{\partial u}{\partial y} = 0.$$

In each case we write $u(x, y) = X(x)Y(y)$, separate the variables into groups that each depend on only one variable, and then assert that each must be equal to a constant, with the several constants satisfying an arithmetic identity.

(a)
$$\begin{aligned}
\frac{\partial u}{\partial x} - x\frac{\partial u}{\partial y} &= 0,\\
X'Y - xXY' &= 0,\\
\frac{X'}{xX} = \frac{Y'}{Y} = k \quad \Rightarrow &\quad \ln X = \tfrac{1}{2}kx^2 + c_1, \quad \ln Y = ky + c_2,\\
\Rightarrow &\quad X = Ae^{kx^2/2}, \quad Y = Be^{ky},\\
\Rightarrow \quad u(x, y) &= Ce^{\lambda(x^2+2y)}, \text{ where } k = 2\lambda.
\end{aligned}$$

(b)
$$\begin{aligned}
x\frac{\partial u}{\partial x} - 2y\frac{\partial u}{\partial y} &= 0,\\
xX'Y - 2yXY' &= 0,\\
\frac{xX'}{X} = \frac{2yY'}{Y} = k \quad \Rightarrow &\quad \ln X = k\ln x + c_1,\\
&\quad \ln Y = \tfrac{1}{2}k\ln y + c_2,\\
\Rightarrow &\quad X = Ax^k, \quad Y = By^{k/2},\\
\Rightarrow \quad u(x, y) &= C(x^2y)^{\lambda}, \text{ where } k = 2\lambda.
\end{aligned}$$

11.3 The wave equation describing the transverse vibrations of a stretched membrane under tension T and having a uniform surface density ρ is

$$T\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) = \rho\frac{\partial^2 u}{\partial t^2}.$$

Find a separable solution appropriate to a membrane stretched on a frame of length a and width b, showing that the natural angular frequencies of such a membrane are given by

$$\omega^2 = \frac{\pi^2 T}{\rho}\left(\frac{n^2}{a^2} + \frac{m^2}{b^2}\right),$$

where n and m are any positive integers.

We seek solutions $u(x, y, t)$ that are periodic in time and have $u(0, y, t) = u(a, y, t) = u(x, 0, t) = u(x, b, t) = 0$. Write $u(x, y, t) = X(x)Y(y)S(t)$ and substitute, obtaining

$$T(X''YS + XY''S) = \rho XYS'',$$

which, when divided through by XYS, gives

$$\frac{X''}{X} + \frac{Y''}{Y} = \frac{\rho}{T}\frac{S''}{S} = -\frac{\omega^2\rho}{T}.$$

The second equality, obtained by applying the separation of variables principle with separation constant $-\omega^2\rho/T$, gives $S(t)$ as a sinusoidal function of t of frequency ω, i.e. $A\cos(\omega t) + B\sin(\omega t)$.

We then have, on applying the separation of variables principle a second time, that

$$\frac{X''}{X} = \lambda \text{ and } \frac{Y''}{Y} = \mu, \text{ where } \lambda + \mu = -\frac{\omega^2\rho}{T}. \qquad (*)$$

These equations must also have sinusoidal solutions. This is because, since $u(0, y, t) = u(a, y, t) = u(x, 0, t) = u(x, b, t) = 0$, each solution has to have zeros at two different values of its argument. We are thus led to

$$X = A\sin(p\,x) \text{ and } Y = B\sin(q\,x), \text{ where } p^2 = -\lambda \text{ and } q^2 = -\mu.$$

Further, since $u(a, y, t) = u(x, b, t) = 0$, we must have $p = n\pi/a$ and $q = m\pi/b$, where n and m are integers. Putting these values back into $(*)$ gives

$$-p^2 - q^2 = -\frac{\omega^2\rho}{T} \quad\Rightarrow\quad \pi^2\left(\frac{n^2}{a^2} + \frac{m^2}{b^2}\right) = \frac{\omega^2\rho}{T}.$$

Hence the quoted result.

11.5 Denoting the three terms of ∇^2 in spherical polars by ∇_r^2, ∇_θ^2, ∇_ϕ^2 in an obvious way, evaluate $\nabla_r^2 u$, etc. for the two functions given below and verify that, in each case, although the individual terms are not necessarily zero their sum $\nabla^2 u$ is zero. Identify the corresponding values of ℓ and m.

(a) $u(r, \theta, \phi) = \left(Ar^2 + \dfrac{B}{r^3}\right)\dfrac{3\cos^2\theta - 1}{2}.$

(b) $u(r, \theta, \phi) = \left(Ar + \dfrac{B}{r^2}\right)\sin\theta\exp i\phi.$

In both cases we write $u(r,\theta,\phi)$ as $R(r)\Theta(\theta)\Phi(\phi)$ with

$$\nabla_r^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right),\quad \nabla_\theta^2 = \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right),\quad \nabla_\phi^2 = \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}.$$

(a) $u(r,\theta,\phi) = \left(Ar^2 + \dfrac{B}{r^3}\right)\dfrac{3\cos^2\theta - 1}{2}.$

$$\begin{aligned}
\nabla_r^2 u &= \frac{1}{r^2}\frac{\partial}{\partial r}\left(2Ar^3 - \frac{3B}{r^2}\right)\Theta = \left(6A + \frac{6B}{r^5}\right)\Theta = \frac{6u}{r^2},\\
\nabla_\theta^2 u &= \frac{R}{r^2}\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}(-3\sin^2\theta\cos\theta) = \frac{R}{r^2}\left(\frac{-6\sin\theta\cos^2\theta + 3\sin^3\theta}{\sin\theta}\right)\\
&= \frac{R}{r^2}(-9\cos^2\theta + 3) = -\frac{6u}{r^2},\\
\nabla_\phi^2 u &= 0.
\end{aligned}$$

Thus, although $\nabla_r^2 u$ and $\nabla_\theta^2 u$ are not individually zero, their sum is. From $\nabla_r^2 u = \ell(\ell+1)u = 6u$, we deduce that $\ell = 2$ (or -3) and from $\nabla_\phi^2 u = 0$ that $m = 0$.

(b) $u(r,\theta,\phi) = \left(Ar + \dfrac{B}{r^2}\right)\sin\theta\, e^{i\phi}.$

$$\begin{aligned}
\nabla_r^2 u &= \frac{1}{r^2}\frac{\partial}{\partial r}\left(Ar^2 - \frac{2B}{r}\right)\Theta\Phi = \left(\frac{2A}{r} + \frac{2B}{r^4}\right)\Theta\Phi = \frac{2u}{r^2},\\
\nabla_\theta^2 u &= \frac{R\Phi}{r^2}\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta\cos\theta) = \frac{R\Phi}{r^2}\left(\frac{-\sin^2\theta + \cos^2\theta}{\sin\theta}\right)\\
&= -\frac{u}{r^2} + \frac{\cos^2\theta}{\sin^2\theta}\frac{u}{r^2},\\
\nabla_\phi^2 u &= \frac{R\Theta}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}(e^{i\phi}) = -\frac{u}{r^2\sin^2\theta}.
\end{aligned}$$

Hence,

$$\nabla^2 u = \frac{2u}{r^2} - \frac{u}{r^2} + \frac{\cos^2\theta}{\sin^2\theta}\frac{u}{r^2} - \frac{u}{r^2\sin^2\theta} = \frac{u}{r^2}\left(1 + \frac{\cos^2\theta - 1}{\sin^2\theta}\right) = 0.$$

Here each individual term is non-zero, but their sum *is* zero. Further, $\ell(\ell+1) = 2$ and so $\ell = 1$ (or -2), and from $\nabla_\phi^2 u = -u/(r^2\sin\theta)$ it follows that $m^2 = 1$. In fact, from the normal definition of spherical harmonics, $m = +1$.

11.7 If the stream function $\psi(r, \theta)$ for the flow of a very viscous fluid past a sphere is written as $\psi(r, \theta) = f(r)\sin^2\theta$, then $f(r)$ satisfies the equation

$$f^{(4)} - \frac{4f''}{r^2} + \frac{8f'}{r^3} - \frac{8f}{r^4} = 0.$$

At the surface of the sphere $r = a$ the velocity field $\mathbf{u} = \mathbf{0}$, whilst far from the sphere $\psi \simeq (Ur^2\sin^2\theta)/2$.

Show that $f(r)$ can be expressed as a superposition of powers of r, and determine which powers give acceptable solutions. Hence show that

$$\psi(r, \theta) = \frac{U}{4}\left(2r^2 - 3ar + \frac{a^3}{r}\right)\sin^2\theta.$$

For solutions of

$$f^{(4)} - \frac{4f''}{r^2} + \frac{8f'}{r^3} - \frac{8f}{r^4} = 0$$

that are powers of r, i.e. have the form Ar^n, n must satisfy the quartic equation

$$\begin{aligned}
n(n-1)(n-2)(n-3) - 4n(n-1) + 8n - 8 &= 0,\\
(n-1)[n(n-2)(n-3) - 4n + 8] &= 0,\\
(n-1)(n-2)[n(n-3) - 4] &= 0,\\
(n-1)(n-2)(n-4)(n+1) &= 0.
\end{aligned}$$

Thus the possible powers are 1, 2, 4 and -1.

Since $\psi \to \frac{1}{2}Ur^2\sin^2\theta$ as $r \to \infty$, the solution can contain no higher (positive) power of r than the second. Thus there is no $n = 4$ term and the solution has the form

$$\psi(r, \theta) = \left(\frac{Ur^2}{2} + Ar + \frac{B}{r}\right)\sin^2\theta.$$

On the surface of the sphere $r = a$ both velocity components, u_r and u_θ, are zero. These components are given in terms of the stream functions, as shown below; note that u_r is found by differentiating with respect to θ and u_θ by differentiating with respect to r.

$$\begin{aligned}
u_r = 0 \quad &\Rightarrow \quad \frac{1}{a^2\sin\theta}\frac{\partial\psi}{\partial\theta} = 0 \quad \Rightarrow \quad \frac{Ua^2}{2} + Aa + \frac{B}{a} = 0,\\
u_\theta = 0 \quad &\Rightarrow \quad \frac{-1}{a\sin\theta}\frac{\partial\psi}{\partial r} = 0 \quad \Rightarrow \quad Ua + A - \frac{B}{a^2} = 0,\\
&\Rightarrow \quad A = -\tfrac{3}{4}Ua \text{ and } B = \tfrac{1}{4}Ua^3.
\end{aligned}$$

The full solution is thus

$$\psi(r, \theta) = \frac{U}{4}\left(2r^2 - 3ar + \frac{a^3}{r}\right)\sin^2\theta.$$

11.9 A circular disc of radius a is heated in such a way that its perimeter $\rho = a$ has a steady temperature distribution $A + B\cos^2\phi$, where ρ and ϕ are plane polar coordinates and A and B are constants. Find the temperature $T(\rho, \phi)$ everywhere in the region $\rho < a$.

This is a steady state problem, for which the (heat) diffusion equation becomes the Laplace equation. The most general single-valued solution to the Laplace equation in plane polar coordinates is given by

$$T(\rho, \phi) = C \ln \rho + D + \sum_{n=1}^{\infty} (A_n \cos n\phi + B_n \sin n\phi)(C_n \rho^n + D_n \rho^{-n}).$$

The region $\rho < a$ contains the point $\rho = 0$; since $\ln \rho$ and all ρ^{-n} become infinite at that point, $C = D_n = 0$ for all n.

On $\rho = a$

$$T(a, \phi) = A + B\cos^2\phi = A + \tfrac{1}{2}B(\cos 2\phi + 1).$$

Equating the coefficients of $\cos n\phi$, including $n = 0$, gives $A + \frac{1}{2}B = D$, $A_2 C_2 a^2 = \frac{1}{2}B$ and $A_n C_n a^n = 0$ for all $n \neq 2$; further, all $B_n = 0$. The solution everywhere (not just on the perimeter) is therefore

$$T(\rho, \phi) = A + \frac{B}{2} + \frac{B\rho^2}{2a^2}\cos 2\phi.$$

It should be noted that "equating coefficients" to determine unknown constants is justified by the fact that the sinusoidal functions in the sum are mutually orthogonal over the range $0 \leq \phi < 2\pi$.

11.11 The free transverse vibrations of a thick rod satisfy the equation

$$a^4 \frac{\partial^4 u}{\partial x^4} + \frac{\partial^2 u}{\partial t^2} = 0.$$

Obtain a solution in separated-variable form and, for a rod clamped at one end, $x = 0$, and free at the other, $x = L$, show that the angular frequency of vibration ω satisfies

$$\cosh\left(\frac{\omega^{1/2}L}{a}\right) = -\sec\left(\frac{\omega^{1/2}L}{a}\right).$$

[At a clamped end both u and $\partial u/\partial x$ vanish, whilst at a free end, where there is no bending moment, $\partial^2 u/\partial x^2$ and $\partial^3 u/\partial x^3$ are both zero.]

The general solution is written as the product $u(x, t) = X(x)T(t)$, which, on substitution, produces the separated equation

$$a^4 \frac{X^{(4)}}{X} = -\frac{T''}{T} = \omega^2.$$

Here the separation constant has been chosen so as to give oscillatory behavior (in the time variable). The spatial equation then becomes

$$X^{(4)} - \mu^4 X = 0, \quad \text{where } \mu = \omega^{1/2}/a.$$

The required auxiliary equation is $\lambda^4 - \mu^4 = 0$, leading to the general solution

$$X(x) = A \sin \mu x + B \cos \mu x + C \sinh \mu x + D \cosh \mu x.$$

The constants A, B, C and D are to be determined by requiring $X(0) = X'(0) = 0$ and $X''(L) = X'''(L) = 0$.

At the clamped end,

$$\begin{aligned} X(0) = 0 \quad &\Rightarrow \quad D = -B, \\ X' \quad &= \quad \mu(A \cos \mu x - B \sin \mu x + C \cosh \mu x - B \sinh \mu x), \\ X'(0) = 0 \quad &\Rightarrow \quad C = -A. \end{aligned}$$

At the free end,

$$\begin{aligned} X'' &= \mu^2(-A \sin \mu x - B \cos \mu x - A \sinh \mu x - B \cosh \mu x), \\ X''' &= \mu^3(-A \cos \mu x + B \sin \mu x - A \cosh \mu x - B \sinh \mu x), \\ X''(L) = 0 \quad &\Rightarrow \quad A(\sin \mu L + \sinh \mu L) + B(\cos \mu L + \cosh \mu L) = 0, \\ X'''(L) = 0 \quad &\Rightarrow \quad A(-\cos \mu L - \cosh \mu L) + B(\sin \mu L - \sinh \mu L) = 0. \end{aligned}$$

Cross-multiplying then gives

$$\begin{aligned} -\sin^2 \mu L + \sinh^2 \mu L &= \cos^2 \mu L + 2 \cos \mu L \cosh \mu L + \cosh^2 \mu L, \\ 0 &= 1 + 2 \cos \mu L \cosh \mu L + 1, \\ -1 &= \cos \mu L \cosh \mu L, \\ \cosh\left(\frac{\omega^{1/2} L}{a}\right) &= -\sec\left(\frac{\omega^{1/2} L}{a}\right). \end{aligned}$$

Because sinusoidal and hyperbolic functions can all be written in terms of exponential functions, this problem could also be approached by assuming solutions that are (exponential) functions of linear combinations of x and t (as in Chapter 10). However, in practice, eliminating the t-dependent terms leads to involved algebra.

11.13 A string of length L, fixed at its two ends, is plucked at its mid-point by an amount A and then released. Prove that the subsequent displacement is given by

$$u(x, t) = \sum_{n=0}^{\infty} \frac{8A}{\pi^2(2n+1)^2} \sin\left[\frac{(2n+1)\pi x}{L}\right] \cos\left[\frac{(2n+1)\pi ct}{L}\right],$$

where, in the usual notation, $c^2 = T/\rho$.

Find the total kinetic energy of the string when it passes through its unplucked position, by calculating it in each mode (each n) and summing, using the result

$$\sum_{0}^{\infty} \frac{1}{(2n+1)^2} = \frac{\pi^2}{8}.$$

Confirm that the total energy is equal to the work done in plucking the string initially.

We start with the wave equation:

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = 0$$

and assume a separated-variable solution $u(x, t) = X(x)S(t)$. This leads to

$$\frac{X''}{X} = \frac{1}{c^2}\frac{S''}{S} = -k^2.$$

The solution to the spatial equation is given by

$$X(x) = B\cos kx + C\sin kx.$$

Taking the string as anchored at $x = 0$ and $x = L$, we must have $B = 0$ and k constrained by $\sin kL = 0 \Rightarrow k = n\pi/L$ with n an integer.

The solution to the corresponding temporal equation is

$$S(t) = D\cos kct + E\sin kct.$$

Since there is no initial motion, i.e. $\dot{S}(0) = 0$, it follows that $E = 0$.

For any particular value of k, the constants C and D can be amalgamated. The general solution is given by a superposition of the allowed functions, i.e.

$$u(x, t) = \sum_{n=1}^{\infty} C_n \sin\frac{n\pi x}{L}\cos\frac{n\pi ct}{L}.$$

We now have to determine the C_n by making $u(x, 0)$ match the given initial configuration, which is

$$u(x, 0) = \begin{cases} \dfrac{2Ax}{L} & \text{for } 0 \le x \le \dfrac{L}{2}, \\ \dfrac{2A(L-x)}{L} & \dfrac{L}{2} < x \le L. \end{cases}$$

This is now a Fourier series calculation yielding

$$\begin{aligned} \frac{C_n L}{2} &= \int_0^{L/2} \frac{2Ax}{L}\sin\frac{n\pi x}{L}\,dx + \int_{L/2}^{L} \frac{2A(L-x)}{L}\sin\frac{n\pi x}{L}\,dx \\ &= \frac{2A}{L}J_1 + 2AJ_2 - \frac{2A}{L}J_3, \end{aligned}$$

with

$$
\begin{aligned}
J_1 &= \left[-\frac{xL}{n\pi}\cos\frac{n\pi x}{L}\right]_0^{L/2} + \int_0^{L/2}\frac{L}{n\pi}\cos\frac{n\pi x}{L}\,dx \\
&= -\frac{L^2}{2\pi n}\cos\frac{n\pi}{2} + \frac{L^2}{n^2\pi^2}\sin\frac{n\pi}{2}, \\
J_2 &= \int_{L/2}^{L}\sin\frac{n\pi x}{L}\,dx = -\frac{L}{n\pi}\left[\cos\frac{n\pi x}{L}\right]_{L/2}^{L} \\
&= -\frac{L}{n\pi}\left[(-1)^n - \cos\frac{n\pi}{2}\right], \\
J_3 &= \left[-\frac{xL}{n\pi}\cos\frac{n\pi x}{L}\right]_{L/2}^{L} + \int_{L/2}^{L}\frac{L}{n\pi}\cos\frac{n\pi x}{L}\,dx \\
&= \frac{L^2}{2\pi n}\cos\frac{n\pi}{2} - \frac{L^2}{n\pi}(-1)^n - \frac{L^2}{n^2\pi^2}\sin\frac{n\pi}{2}.
\end{aligned}
$$

Thus

$$
\begin{aligned}
J_1 - J_3 &= -\frac{2L^2}{2\pi n}\cos\frac{n\pi}{2} + \frac{L^2}{n\pi}(-1)^n + \frac{2L^2}{n^2\pi^2}\sin\frac{n\pi}{2} \\
&= -LJ_2 + \frac{2L^2}{n^2\pi^2}\sin\frac{n\pi}{2},
\end{aligned}
$$

and so it follows that

$$
\frac{C_n L}{2} = \frac{2A}{L}(J_1 - J_3 + LJ_2) = 2A\frac{2L}{n^2\pi^2}\sin\frac{n\pi}{2}.
$$

This is zero if n is even and $C_n = 8A(-1)^{(n-1)/2}/(n^2\pi^2)$ if n is odd. Write $n = 2m+1$, $m = 0, 1, 2, \ldots$, with $C_{2m+1} = \dfrac{8A(-1)^m}{(2m+1)^2\pi^2}$.

The final solution (in which m is replaced by n, to match the question) is thus

$$
u(x,t) = \sum_{n=0}^{\infty}\frac{8A(-1)^n}{\pi^2(2n+1)^2}\sin\left[\frac{(2n+1)\pi x}{L}\right]\cos\left[\frac{(2n+1)\pi ct}{L}\right].
$$

The velocity profile derived from this is given by

$$
\begin{aligned}
\dot{u}(x,t) = \sum_{n=0}^{\infty}&\frac{8A(-1)^n}{\pi^2(2n+1)^2}\left(\frac{-(2n+1)\pi c}{L}\right) \\
&\times \sin\left[\frac{(2n+1)\pi x}{L}\right]\sin\left[\frac{(2n+1)\pi ct}{L}\right],
\end{aligned}
$$

giving the energy in the $(2n+1)$th mode (evaluated when the time-dependent sine function is maximal) as

$$\begin{aligned} E_{2n+1} &= \int_0^L \tfrac{1}{2}\rho \dot{u}_n^2\, dx \\ &= \int_0^L \frac{\rho}{2}\,\frac{(8A)^2c^2}{L^2(2n+1)^2\pi^2}\sin^2\frac{(2n+1)\pi x}{L} \\ &= \frac{32A^2\rho c^2}{L^2(2n+1)^2\pi^2}\,\frac{L}{2}. \end{aligned}$$

Therefore

$$E = \sum_{n=0}^{\infty} E_{2n+1} = \frac{16A^2\rho c^2}{\pi^2 L}\sum_{n=0}^{\infty}\frac{1}{(2n+1)^2} = \frac{2A^2\rho c^2}{L}.$$

When the mid-point of the string has been displaced sideways by y ($\ll L$), the net (resolved) restoring force is $2T[y/(L/2)] = 4Ty/L$. Thus the total work done to produce a displacement of A is

$$W = \int_0^A \frac{4Ty}{L}\,dy = \frac{2TA^2}{L} = \frac{2\rho c^2 A^2}{L},$$

i.e. the same as the total energy of the subsequent motion.

11.15 Prove that the potential for $\rho < a$ associated with a vertical split cylinder of radius a, the two halves of which ($\cos\phi > 0$ and $\cos\phi < 0$) are maintained at equal and opposite potentials $\pm V$, is given by

$$u(\rho,\phi) = \frac{4V}{\pi}\sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}\left(\frac{\rho}{a}\right)^{2n+1}\cos(2n+1)\phi.$$

The most general solution of the Laplace equation in cylindrical polar coordinates that is independent of z is

$$T(\rho,\phi) = C\ln\rho + D + \sum_{n=1}^{\infty}(A_n\cos n\phi + B_n\sin n\phi)(C_n\rho^n + D_n\rho^{-n}).$$

The required potential must be single-valued and finite in the space inside the cylinder (which includes $\rho = 0$), and on the cylinder it must take the boundary values $u = V$ for $\cos\phi > 0$ and $u = -V$ for $\cos\phi < 0$, i.e the boundary-value function is a square-wave function with average value zero. Although the function is antisymmetric in $\cos\phi$, it is symmetric in ϕ and so the solution will contain only cosine terms (and no sine terms).

These considerations already determine that $C = D = B_n = D_n = 0$, and so have reduced the solution to the form

$$u(\rho,\phi) = \sum_{n=1}^{\infty} A_n\rho^n\cos n\phi.$$

On $\rho = a$ this must match the stated boundary conditions, and so we are faced with a Fourier cosine series calculation. Multiplying through by $\cos m\phi$ and integrating yields

$$\begin{aligned} A_m a^m \frac{1}{2} 2\pi &= 2\int_0^{\pi/2} V \cos m\phi \, d\phi + 2\int_{\pi/2}^{\pi} (-V) \cos m\phi \, d\phi \\ &= 2V \left[\frac{\sin m\phi}{m} \right]_0^{\pi/2} - 2V \left[\frac{\sin m\phi}{m} \right]_{\pi/2}^{\pi} \\ &= \frac{2V}{m} \left(\sin \frac{m\pi}{2} + \sin \frac{m\pi}{2} \right) \\ &= (-1)^{(m-1)/2} \frac{4V}{m} \text{ for } m \text{ odd}, = 0 \text{ for } m \text{ even}. \end{aligned}$$

Writing $m = 2n + 1$ gives the solution as

$$u(\rho, \phi) = \frac{4V}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \left(\frac{\rho}{a} \right)^{2n+1} \cos(2n+1)\phi.$$

11.17 Two identical copper bars are each of length a. Initially, one is at 0 °C and the other at 100 °C; they are then joined together end to end and thermally isolated. Obtain in the form of a Fourier series an expression $u(x, t)$ for the temperature at any point a distance x from the join at a later time t. Bear in mind the heat flow conditions at the free ends of the bars.

Taking $a = 0.5$ m estimate the time it takes for one of the free ends to attain a temperature of 55 °C. The thermal conductivity of copper is 3.8×10^2 J m^{-1} K^{-1} s^{-1}, and its specific heat capacity is 3.4×10^6 J m^{-3} K^{-1}.

The equation governing the heat flow is

$$k \frac{\partial^2 u}{\partial x^2} = s \frac{\partial u}{\partial t},$$

which is the diffusion equation with diffusion constant $\kappa = k/s = 3.8 \times 10^2 / 3.4 \times 10^6 = 1.12 \times 10^{-4}$ m^2 s^{-1}.

Making the usual separation of variables substitution shows that the time variation is of the form $T(t) = T(0)e^{-\kappa\lambda^2 t}$ when the spatial solution is a sinusoidal function of λx. The final common temperature is 50 °C and we make this explicit by writing the general solution as

$$u(x, t) = 50 + \sum_{\lambda} (A_\lambda \sin \lambda x + B_\lambda \cos \lambda x) e^{-\kappa\lambda^2 t}.$$

This term having been taken out, the summation must be antisymmetric about $x = 0$ and therefore contain no cosine terms, i.e. $B_\lambda = 0$.

The boundary condition is that there is no heat flow at $x = \pm a$; this means that $\partial u / \partial x = 0$ at these points and requires

$$\lambda A_\lambda \cos \lambda x|_{x=\pm a} = 0 \quad \Rightarrow \quad \lambda a = (n + \tfrac{1}{2})\pi \quad \Rightarrow \quad \lambda = \frac{(2n+1)\pi}{2a},$$

where n is an integer. This corresponds to a fundamental Fourier period of $4a$. The solution thus takes the form

$$u(x,t) = 50 + \sum_{n=0}^{\infty} A_n \sin \frac{(2n+1)\pi x}{2a} \exp\left(-\frac{(2n+1)^2\pi^2\kappa t}{4a^2}\right).$$

At $t = 0$, the sum must take the values $+50$ for $0 < x < 2a$ and -50 for $-2a < x < 0$. This is (yet) another square-wave function – one that is antisymmetric about $x = 0$ and has amplitude 50. The calculation will not be repeated here but gives $A_n = 200/[(2n+1)\pi]$, making the complete solution

$$u(x,t) = 50 + \frac{200}{\pi}\sum_{n=0}^{\infty} \frac{1}{2n+1} \sin \frac{(2n+1)\pi x}{2a} \exp\left(-\frac{(2n+1)^2\pi^2\kappa t}{4a^2}\right).$$

For a free end, where $x = a$ and $\sin[(2n+1)\pi x/2a] = (-1)^n$, to attain 55 °C needs

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \exp\left(-\frac{(2n+1)^2\,\pi^2\,1.12\times 10^{-4}}{4\times 0.25}\,t\right) = \frac{5\pi}{200} = 0.0785.$$

In principle this is an insoluble equation but, because the RHS $\ll 1$, the $n = 0$ term alone will give a good approximation to t:

$$\exp(-1.105\times 10^{-3}t) \approx 0.0785 \quad \Rightarrow \quad t \approx 2300 \text{ s}.$$

11.19 For an infinite metal bar that has an initial ($t = 0$) temperature distribution $f(x)$ along its length, the temperature distribution at a general time t can be shown to be given by

$$u(x,t) = \frac{1}{\sqrt{4\pi\kappa t}} \int_{-\infty}^{\infty} \exp\left[-\frac{(x-\xi)^2}{4\kappa t}\right] f(\xi)\,d\xi.$$

Find an explicit expression for $u(x,t)$ given that $f(x) = \exp(-x^2/a^2)$.

The given initial distribution is $f(\xi) = \exp(-\xi^2/a^2)$ and so

$$u(x,t) = \frac{1}{\sqrt{4\pi\kappa t}} \int_{-\infty}^{\infty} \exp\left[-\frac{(x-\xi)^2}{4\kappa t}\right] \exp\left(-\frac{\xi^2}{a^2}\right) d\xi.$$

Now consider the exponent in the integrand, writing $1 + \dfrac{4\kappa t}{a^2}$ as τ^2 for compactness:

$$\begin{aligned}
\text{exponent} &= -\frac{\xi^2\tau^2 - 2\xi x + x^2}{4\kappa t} \\
&= -\frac{(\xi\tau - x\tau^{-1})^2 - x^2\tau^{-2} + x^2}{4\kappa t} \\
&\equiv -\eta^2 + \frac{x^2\tau^{-2} - x^2}{4\kappa t}, \quad \text{defining } \eta, \\
\text{with } d\eta &= \frac{\tau\, d\xi}{\sqrt{4\kappa t}}.
\end{aligned}$$

With a change of variable from ξ to η, the integral becomes

$$\begin{aligned} u(x,t) &= \frac{1}{\sqrt{4\pi\kappa t}} \exp\left(\frac{x^2\tau^{-2} - x^2}{4\kappa t}\right) \int_{-\infty}^{\infty} \exp(-\eta^2)\, \frac{\sqrt{4\kappa t}}{\tau}\, d\eta \\ &= \frac{1}{\sqrt{\pi}}\frac{1}{\tau} \exp\left(x^2 \frac{1-\tau^2}{4\kappa t\, \tau^2}\right) \sqrt{\pi} \\ &= \frac{a}{\sqrt{a^2+4\kappa t}} \exp\left(-\frac{x^2}{a^2+4\kappa t}\right). \end{aligned}$$

In words, although it retains a Gaussian shape, the initial distribution spreads symmetrically about the origin, its variance increasing linearly with time ($a^2 \to a^2 + 4\kappa t$). As is typical with diffusion processes, for large enough times the width varies as $\sqrt{t}$.

11.21 In the region $-\infty < x, y < \infty$ and $-t \le z \le t$, a charge-density wave $\rho(\mathbf{r}) = A\cos qx$, in the x-direction, is represented by

$$\rho(\mathbf{r}) = \frac{e^{iqx}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{\rho}(\alpha) e^{i\alpha z}\, d\alpha.$$

The resulting potential is represented by

$$V(\mathbf{r}) = \frac{e^{iqx}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{V}(\alpha) e^{i\alpha z}\, d\alpha.$$

Determine the relationship between $\tilde{V}(\alpha)$ and $\tilde{\rho}(\alpha)$, and hence show that the potential at the point $(0,0,0)$ is given by

$$\frac{A}{\pi\epsilon_0} \int_{-\infty}^{\infty} \frac{\sin kt}{k(k^2+q^2)}\, dk.$$

Poisson's equation,

$$\nabla^2 V(\mathbf{r}) = -\frac{\rho(\mathbf{r})}{\epsilon_0},$$

provides the link between a charge density and the potential it produces.

Taking $V(\mathbf{r})$ in the form of its Fourier representation gives $\nabla^2 V$ as

$$\frac{\partial^2 V(\mathbf{r})}{\partial x^2} + \frac{\partial^2 V(\mathbf{r})}{\partial y^2} + \frac{\partial^2 V(\mathbf{r})}{\partial z^2} = \frac{e^{iqx}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (-q^2 - \alpha^2)\tilde{V}(\alpha) e^{i\alpha z}\, d\alpha,$$

with the $-q^2$ arising from the x-differentiation and the $-\alpha^2$ from the z-differentiation; the $\partial^2 V/\partial y^2$ term contributes nothing.

Comparing this with the integral expression for $-\rho(\mathbf{r})/\epsilon_0$ shows that

$$-\tilde{\rho}(\alpha) = \epsilon_0(-q^2 - \alpha^2)\tilde{V}(\alpha).$$

With the charge-density wave confined in the z-direction to $-t \le z \le t$, the expression for $\rho(\mathbf{r})$ in Cartesian coordinates is (in terms of Heaviside functions)

$$\rho(\mathbf{r}) = Ae^{iqx}[H(z+t) - H(z-t)].$$

The Fourier transform $\tilde{\rho}(\alpha)$ is therefore given by

$$\begin{aligned}
\tilde{\rho}(\alpha) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} A[H(z+t) - H(z-t)]\, e^{-i\alpha z}\, dz \\
&= \frac{A}{\sqrt{2\pi}} \int_{-t}^{t} e^{-i\alpha z}\, dz \\
&= \frac{A}{\sqrt{2\pi}} \frac{e^{-i\alpha t} - e^{i\alpha t}}{-i\alpha} \\
&= \frac{A}{\sqrt{2\pi}} \frac{2\sin\alpha t}{\alpha}.
\end{aligned}$$

Now,

$$\begin{aligned}
V(x,0,z) &= \frac{e^{iqx}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{\tilde{\rho}(\alpha)}{\epsilon_0(q^2+\alpha^2)}\, e^{i\alpha z}\, d\alpha \\
&= \frac{e^{iqx}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{e^{i\alpha z}}{\epsilon_0(q^2+\alpha^2)} \frac{A}{\sqrt{2\pi}} \frac{2\sin\alpha t}{\alpha}\, d\alpha, \\
\Rightarrow \quad V(0,0,0) &= \frac{A}{\pi\epsilon_0} \int_{-\infty}^{\infty} \frac{\sin\alpha t}{\alpha(\alpha^2+q^2)}\, d\alpha,
\end{aligned}$$

as stated in the question.

11.23 Find the Green's function $G(\mathbf{r}, \mathbf{r}_0)$ in the half-space $z > 0$ for the solution of $\nabla^2\Phi = 0$ with Φ specified in cylindrical polar coordinates (ρ, ϕ, z) on the plane $z = 0$ by

$$\Phi(\rho, \phi, z) = \begin{cases} 1 & \text{for } \rho \le 1, \\ 1/\rho & \text{for } \rho > 1. \end{cases}$$

Determine the variation of $\Phi(0, 0, z)$ along the z-axis.

For the half-space $z > 0$ the bounding surface consists of the plane $z = 0$ and the (hemi-spherical) surface at infinity; the Green's function must take zero value on these surfaces. In order to ensure this when a unit point source is introduced at $\mathbf{r} = \mathbf{y}$, we must place a compensating negative unit source at $\mathbf{y}$'s reflection point in the plane. If, in cylindrical polar coordinates, $\mathbf{y} = (\rho, \phi, z_0)$, then the image charge has to be at $\mathbf{y}' = (\rho, \phi, -z_0)$. The resulting Green's function $G(\mathbf{x}, \mathbf{y})$ is given by

$$G(\mathbf{x}, \mathbf{y}) = -\frac{1}{4\pi|\mathbf{x} - \mathbf{y}|} + \frac{1}{4\pi|\mathbf{x} - \mathbf{y}'|}.$$

The solution to the problem with a given potential distribution $f(\rho, \phi)$ on the $z = 0$ part of the bounding surface S is given by

$$\Phi(\mathbf{y}) = \int_S f(\rho, \phi) \left(-\frac{\partial G}{\partial z} \right) \rho\, d\phi\, d\rho,$$

the minus sign arising because the outward normal to the region is in the negative z-direction. Calculating these functions explicitly gives

$$\begin{aligned}
G(\mathbf{x},\mathbf{y}) &= -\frac{1}{4\pi[\rho^2+(z-z_0)^2]^{1/2}} + \frac{1}{4\pi[\rho^2+(z+z_0)^2]^{1/2}},\\
\frac{\partial G}{\partial z} &= \frac{z-z_0}{4\pi[\rho^2+(z-z_0)^2]^{3/2}} - \frac{(z+z_0)}{4\pi[\rho^2+(z+z_0)^2]^{3/2}},\\
-\left.\frac{\partial G}{\partial z}\right|_{z=0} &= -\frac{-2z_0}{4\pi[\rho^2+z_0^2]^{3/2}}.
\end{aligned}$$

Substituting the various factors into the general integral gives

$$\begin{aligned}
\Phi(0,0,z_0) &= \int_0^\infty f(\rho)\frac{2z_0}{4\pi[\rho^2+z_0^2]^{3/2}}\,2\pi\,\rho\,d\rho\\
&= \int_0^1 \frac{z_0\,\rho}{(\rho^2+z_0^2)^{3/2}}\,d\rho + \int_1^\infty \frac{z_0}{(\rho^2+z_0^2)^{3/2}}\,d\rho\\
&= -z_0\left[(\rho^2+z_0^2)^{-1/2}\right]_0^1 + \int_\theta^{\pi/2}\frac{z_0^2\sec^2 u}{z_0^3\sec^3 u}\,du,
\end{aligned}$$

where, in the second integral, we have set $\rho = z_0\tan u$ with $d\rho = z_0\sec^2 u\,du$ and $\theta = \tan^{-1}(1/z_0)$. The integral can now be obtained in closed form as

$$\begin{aligned}
\Phi(0,0,z_0) &= -\frac{z_0}{(1+z_0^2)^{1/2}} + 1 + \frac{1}{z_0}\left[\sin u\right]_\theta^{\pi/2}\\
&= 1 - \frac{z_0}{(1+z_0^2)^{1/2}} + \frac{1}{z_0} - \frac{1}{z_0(1+z_0^2)^{1/2}}.
\end{aligned}$$

Thus the variation of Φ along the z-axis is given by

$$\Phi(0,0,z) = \frac{z(1+z^2)^{1/2} - z^2 + (1+z^2)^{1/2} - 1}{z(1+z^2)^{1/2}}.$$

11.25 Find, in the form of an infinite series, the Green's function of the ∇^2 operator for the Dirichlet problem in the region $-\infty < x < \infty,\ -\infty < y < \infty,\ -c \le z \le c$.

The fundamental solution in three dimensions of $\nabla^2\psi = \delta(\mathbf{r})$ is $\psi(\mathbf{r}) = -1/(4\pi r)$.

For the given problem, $G(\mathbf{r},\mathbf{r}_0)$ has to take the value zero on $z = \pm c$ and $\to 0$ for $|x| \to \infty$ and $|y| \to \infty$. Image charges have to be added in the regions $z > c$ and $z < -c$ to bring this about after a charge q has been placed at $\mathbf{r}_0 = (x_0, y_0, z_0)$ with $-c < z_0 < c$. Clearly all images will be on the line $x = x_0$, $y = y_0$.

Each image placed at $z = \xi$ in the region $z > c$ will require a further image of the same strength but opposite sign at $z = -c - \xi$ (in the region $z < -c$) so as to maintain the plane $z = -c$ as an equipotential. Likewise, each image placed at $z = -\chi$ in the region $z < -c$ will require a further image of the same strength but opposite sign at $z = c + \chi$ (in the region $z > c$) so as to maintain the plane $z = c$ as an equipotential. Thus successive image charges appear as follows:

$$\begin{array}{ccc} -q & 2c - z_0 & -2c - z_0 \\ +q & -3c + z_0 & 3c + z_0 \\ -q & 4c - z_0 & -4c - z_0 \\ +q & \text{etc.} & \text{etc.} \end{array}$$

The terms in the Green's function that are additional to the fundamental solution,

$$-\frac{1}{4\pi}[(x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2]^{-1/2},$$

are therefore

$$-\frac{(-1)}{4\pi}\sum_{n=2}^{\infty}\left\{\frac{(-1)^n}{[(x-x_0)^2 + (y-y_0)^2 + (z+(-1)^n z_0 - nc)^2]^{1/2}} \right.$$
$$\left. + \frac{(-1)^n}{[(x-x_0)^2 + (y-y_0)^2 + (z+(-1)^n z_0 + nc)^2]^{1/2}}\right\}.$$

11.27 Determine the Green's function for the Klein–Gordon equation in a half-space as follows.

(a) By applying the divergence theorem to the volume integral

$$\int_V \left[\phi(\nabla^2 - m^2)\psi - \psi(\nabla^2 - m^2)\phi\right] dV,$$

obtain a Green's function expression, as the sum of a volume integral and a surface integral, for the function $\phi(\mathbf{r}')$ that satisfies

$$\nabla^2\phi - m^2\phi = \rho$$

in V and takes the specified form $\phi = f$ on S, the boundary of V. The Green's function, $G(\mathbf{r}, \mathbf{r}')$, to be used satisfies

$$\nabla^2 G - m^2 G = \delta(\mathbf{r} - \mathbf{r}')$$

and vanishes when $\mathbf{r}$ is on S.

(b) When V is all space, $G(\mathbf{r}, \mathbf{r}')$ can be written as $G(t) = g(t)/t$, where $t = |\mathbf{r} - \mathbf{r}'|$ and $g(t)$ is bounded as $t \to \infty$. Find the form of $G(t)$.

(c) Find $\phi(\mathbf{r})$ in the half-space $x > 0$ if $\rho(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}_1)$ and $\phi = 0$ both on $x = 0$ and as $r \to \infty$.

(a) For general ϕ and ψ we have

$$\begin{aligned} \int_V \left[\phi(\nabla^2 - m^2)\psi - \psi(\nabla^2 - m^2)\phi\right] dV &= \int_V \left[\phi\nabla^2\psi - \psi\nabla^2\phi\right] dV \\ &= \int_V \nabla\cdot(\phi\nabla\psi - \psi\nabla\phi)\, dV \\ &= \int_S (\phi\nabla\psi - \psi\nabla\phi)\cdot\mathbf{n}\, dS. \end{aligned}$$

Now take ϕ as ϕ, with $\nabla^2\phi - m^2\phi = \rho$ and $\phi = f$ on the surface S, and ψ as $G(\mathbf{r}, \mathbf{r}')$ with $\nabla^2 G - m^2 G = \delta(\mathbf{r} - \mathbf{r}')$ and $G(\mathbf{r}, \mathbf{r}') = 0$ on S:

$$\int_V [\phi(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}') - G(\mathbf{r}, \mathbf{r}')\rho(\mathbf{r})]\, dV = \int_S [f(\mathbf{r})\nabla G(\mathbf{r}, \mathbf{r}') - \mathbf{0}] \cdot \mathbf{n}\, dS,$$

which, on rearrangement, gives

$$\phi(\mathbf{r}') = \int_V G(\mathbf{r}, \mathbf{r}')\rho(\mathbf{r})\, dV + \int_S f(\mathbf{r})\nabla G(\mathbf{r}, \mathbf{r}') \cdot \mathbf{n}\, dS.$$

(b) In the following calculation we start by formally integrating the defining Green's equation,

$$\nabla^2 G - m^2 G = \delta(\mathbf{r} - \mathbf{r}'),$$

over a sphere of radius t centered on $\mathbf{r}'$. Having replaced the volume integral of $\nabla^2 G$ with the corresponding surface integral given by the divergence theorem, we move the origin to $\mathbf{r}'$, denote $|\mathbf{r} - \mathbf{r}'|$ by t' and integrate both sides of the equation from $t' = 0$ to $t' = t$:

$$\begin{aligned}
\int_V \nabla^2 G\, dV - \int_V m^2 G\, dV &= \int_V \delta(\mathbf{r} - \mathbf{r}')\, dV, \\
\int_S \nabla G \cdot \mathbf{n}\, dS - m^2 \int_V G\, dV &= 1, \\
4\pi t^2 \frac{dG}{dt} - m^2 \int_0^t G(t')4\pi t'^2\, dt' &= 1, \qquad (*) \\
4\pi t^2 G'' + 8\pi t G' - 4\pi m^2 t^2 G &= 0, \text{ from differentiating w.r.t. } t, \\
t G'' + 2G' - m^2 t G &= 0.
\end{aligned}$$

With $G(t) = g(t)/t$,

$$G' = -\frac{g}{t^2} + \frac{g'}{t} \quad \text{and} \quad G'' = \frac{2g}{t^3} - \frac{2g'}{t^2} + \frac{g''}{t},$$

and the equation becomes

$$\begin{aligned}
0 &= \frac{2g}{t^2} - \frac{2g'}{t} + g'' - \frac{2g}{t^2} + \frac{2g'}{t} - m^2 g, \\
0 &= g'' - m^2 g, \\
\Rightarrow \quad g(t) &= Ae^{-mt}, \text{ since } g \text{ is bounded as } t \to \infty.
\end{aligned}$$

The value of A is determined by resubstituting into $(*)$, which then reads

$$\begin{aligned}
4\pi t^2 \left(-\frac{Ae^{-mt}}{t^2} - \frac{mAe^{-mt}}{t} \right) - m^2 \int_0^t \frac{Ae^{-mt'}}{t'} 4\pi t'^2\, dt' &= 1, \\
-4\pi Ae^{-mt}(1 + mt) - 4\pi Am^2 \left(-\frac{te^{-mt}}{m} + \frac{1 - e^{-mt}}{m^2} \right) &= 1, \\
-4\pi A &= 1,
\end{aligned}$$

making the solution

$$G(\mathbf{r}, \mathbf{r}') = -\frac{e^{-mt}}{4\pi t}, \text{ where } t = |\mathbf{r} - \mathbf{r}'|.$$

(c) For the situation in which $\rho(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}_1)$, i.e. a unit positive charge at $\mathbf{r}_1 = (x_1, y_1, z_1)$, and $\phi = 0$ on the plane $x = 0$, we must have a unit negative image charge at $\mathbf{r}_2 = (-x_1, y_1, z_1)$. The solution in the region $x > 0$ is then

$$\phi(\mathbf{r}) = -\frac{1}{4\pi}\left(\frac{e^{-m|\mathbf{r}-\mathbf{r}_1|}}{|\mathbf{r} - \mathbf{r}_1|} - \frac{e^{-m|\mathbf{r}-\mathbf{r}_2|}}{|\mathbf{r} - \mathbf{r}_2|}\right).$$

12 Calculus of variations

12.1 A surface of revolution, whose equation in cylindrical polar coordinates is $\rho = \rho(z)$, is bounded by the circles $\rho = a$, $z = \pm c$ $(a > c)$. Show that the function that makes the surface integral $I = \int \rho^{-1/2}\, dS$ stationary with respect to small variations is given by $\rho(z) = k + z^2/(4k)$, where $k = [a \pm (a^2 - c^2)^{1/2}]/2$.

The surface element lying between z and $z + dz$ is given by

$$dS = 2\pi\rho\, [(d\rho)^2 + (dz)^2]^{1/2} = 2\pi\rho\, (1 + \rho'^2)^{1/2}\, dz$$

and the integral to be made stationary is

$$I = \int \rho^{-1/2}\, dS = 2\pi \int_{-c}^{c} \rho^{-1/2} \rho\, (1 + \rho'^2)^{1/2}\, dz.$$

The integrand $F(\rho', \rho, z)$ does not in fact contain z explicitly, and so a first integral of the EL equation, symbolically given by $F - \rho' \partial F/\partial \rho' = k$, is

$$\rho^{1/2}(1 + \rho'^2)^{1/2} - \rho' \left[\frac{\rho^{1/2}\rho'}{(1 + \rho'^2)^{1/2}} \right] = A,$$

$$\frac{\rho^{1/2}}{(1 + \rho'^2)^{1/2}} = A.$$

On rearrangement and subsequent integration this gives

$$\frac{d\rho}{dz} = \left(\frac{\rho - A^2}{A^2} \right)^{1/2},$$

$$\int \frac{d\rho}{\sqrt{\rho - A^2}} = \int \frac{dz}{A},$$

$$2\sqrt{\rho - A^2} = \frac{z}{A} + C.$$

Now, $\rho(\pm c) = a$ implies both that $C = 0$ and that $a - A^2 = \dfrac{c^2}{4A^2}$. Thus, writing A^2 as k,

$$4k^2 - 4ka + c^2 = 0 \quad \Rightarrow \quad k = \tfrac{1}{2}[a \pm (a^2 - c^2)^{1/2}].$$

The two stationary functions are therefore

$$\rho = \frac{z^2}{4k} + k,$$

with k as given above. A simple sketch shows that the positive sign in k corresponds to a smaller value of the integral.

12.3 The refractive index n of a medium is a function only of the distance r from a fixed point O. Prove that the equation of a light ray, assumed to lie in a plane through O, traveling in the medium satisfies (in plane polar coordinates)

$$\frac{1}{r^2}\left(\frac{dr}{d\phi}\right)^2 = \frac{r^2}{a^2}\frac{n^2(r)}{n^2(a)} - 1,$$

where a is the distance of the ray from O at the point at which $dr/d\phi = 0$.

If $n = [1 + (\alpha^2/r^2)]^{1/2}$ and the ray starts and ends far from O, find its deviation (the angle through which the ray is turned), if its minimum distance from O is a.

An element of path length is $ds = [(dr)^2 + (r\,d\phi)^2]^{1/2}$ and the time taken for the light to traverse it is $n(r)\,ds/c$, where c is the speed of light *in vacuo*. Fermat's principle then implies that the light follows the curve that minimizes

$$T = \int \frac{n(r)\,ds}{c} = \int \frac{n(r'^2 + r^2)^{1/2}}{c}\,d\phi,$$

where $r' = dr/d\phi$. Since the integrand does not contain ϕ explicitly, the EL equation integrates to (see Problem 12.1)

$$\begin{aligned} n(r'^2 + r^2)^{1/2} - r'\frac{nr'}{(r'^2 + r^2)^{1/2}} &= A, \\ \frac{nr^2}{(r'^2 + r^2)^{1/2}} &= A. \end{aligned}$$

Since $r' = 0$ when $r = a$, $A = n(a)a^2/a$, and the equation is as follows:

$$\begin{aligned} a^2n^2(a)(r'^2 + r^2) &= n^2(r)r^4, \\ r'^2 &= \frac{n^2(r)r^4}{n^2(a)a^2} - r^2, \\ \Rightarrow \quad \frac{1}{r^2}\left(\frac{dr}{d\phi}\right)^2 &= \frac{n^2(r)r^2}{n^2(a)a^2} - 1. \end{aligned}$$

If $n(r) = [1 + (\alpha/r)^2]^{1/2}$, the minimizing curve satisfies

$$\begin{aligned} \left(\frac{dr}{d\phi}\right)^2 &= \frac{r^2(r^2 + \alpha^2)}{a^2 + \alpha^2} - r^2 \\ &= \frac{r^2(r^2 - a^2)}{a^2 + \alpha^2}, \\ \Rightarrow \quad \frac{d\phi}{(a^2 + \alpha^2)^{1/2}} &= \pm\frac{dr}{r\sqrt{r^2 - a^2}}. \end{aligned}$$

By symmetry,

$$\begin{aligned}
\frac{\Delta\phi}{(a^2+\alpha^2)^{1/2}} &\equiv \frac{\phi_{\text{final}}-\phi_{\text{initial}}}{(a^2+\alpha^2)^{1/2}}\\
&= 2\int_a^\infty \frac{dr}{r\sqrt{r^2-a^2}}, \qquad \text{set } r = a\cosh\psi\\
&= 2\int_0^\infty \frac{a\sinh\psi}{a^2\cosh\psi\sinh\psi}\,d\psi\\
&= \frac{2}{a}\int_0^\infty \operatorname{sech}\psi\,d\psi, \qquad \text{set } e^\psi = z\\
&= \frac{2}{a}\int_1^\infty \frac{z^{-1}\,dz}{\frac{1}{2}(z+z^{-1})}\\
&= \frac{2}{a}\int_1^\infty \frac{2\,dz}{z^2+1}\\
&= \frac{4}{a}\left[\tan^{-1} z\right]_1^\infty\\
&= \frac{4}{a}\left(\frac{\pi}{2}-\frac{\pi}{4}\right) = \frac{\pi}{a}.
\end{aligned}$$

If the refractive index were everywhere unity ($\alpha = 0$), $\Delta\phi$ would be π (no deviation). Thus the deviation is given by

$$\frac{\pi}{a}(a^2+\alpha^2)^{1/2} - \pi.$$

12.5 Prove the following results about general systems.

(a) For a system described in terms of coordinates q_i and t, show that if t does not appear explicitly in the expressions for x, y and z ($x = x(q_i, t)$, etc.) then the kinetic energy T is a homogeneous quadratic function of the $\dot{q}_i$ (it may also involve the q_i). Deduce that $\sum_i \dot{q}_i(\partial T/\partial \dot{q}_i) = 2T$.

(b) Assuming that the forces acting on the system are derivable from a potential V, show, by expressing dT/dt in terms of q_i and $\dot{q}_i$, that $d(T+V)/dt = 0$.

To save space we will use the summation convention for summing over the index of the q_i.

(a) The space variables x, y and z are not explicit functions of t and the kinetic energy, T, is given by

$$\begin{aligned}
T &= \tfrac{1}{2}(\alpha_x\dot{x}^2 + \alpha_y\dot{y}^2 + \alpha_z\dot{z}^2)\\
&= \frac{1}{2}\left[\alpha_x\left(\frac{\partial x}{\partial q_i}\dot{q}_i\right)^2 + \alpha_y\left(\frac{\partial y}{\partial q_j}\dot{q}_j\right)^2 + \alpha_z\left(\frac{\partial z}{\partial q_k}\dot{q}_k\right)^2\right]\\
&= A_{mn}\dot{q}_m\dot{q}_n,
\end{aligned}$$

with

$$A_{mn} = \frac{1}{2}\left(\alpha_x\frac{\partial x}{\partial q_m}\frac{\partial x}{\partial q_n} + \alpha_y\frac{\partial y}{\partial q_m}\frac{\partial y}{\partial q_n} + \alpha_z\frac{\partial z}{\partial q_m}\frac{\partial z}{\partial q_n}\right) = A_{nm}.$$

Hence T is a homogeneous quadratic function of the $\dot{q}_i$ (though the A_{mn} may involve the q_i). Further,

$$\frac{\partial T}{\partial \dot{q}_i} = A_{in}\dot{q}_n + A_{mi}\dot{q}_m = 2A_{mi}\dot{q}_m$$

$$\text{and} \quad \dot{q}_i \frac{\partial T}{\partial \dot{q}_i} = 2\dot{q}_i A_{mi}\dot{q}_m = 2T.$$

(b) The Lagrangian is $L = T - V$, with $T = T(q_i, \dot{q}_i)$ and $V = V(q_i)$. Thus

$$\frac{dT}{dt} = \frac{\partial T}{\partial q_i}\dot{q}_i + \frac{dT}{d\dot{q}_i}\ddot{q}_i \quad \text{and} \quad \frac{dV}{dt} = \frac{\partial V}{\partial q_i}\dot{q}_i. \qquad (*)$$

Hamilton's principle requires that

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\partial L}{\partial q_i},$$

$$\Rightarrow \quad \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_i}\right) = \frac{\partial T}{\partial q_i} - \frac{\partial V}{\partial q_i}. \qquad (**)$$

But, from part (a),

$$2T = \dot{q}_i \frac{\partial T}{\partial \dot{q}_i},$$

$$\begin{aligned}\frac{d}{dt}(2T) &= \ddot{q}_i \frac{\partial T}{\partial \dot{q}_i} + \dot{q}_i \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_i}\right)\\ &= \ddot{q}_i \frac{\partial T}{\partial \dot{q}_i} + \dot{q}_i \frac{\partial T}{\partial q_i} - \dot{q}_i \frac{\partial V}{\partial q_i}, \text{ using } (**),\\ &= \frac{dT}{dt} - \frac{dV}{dt}, \qquad \text{using } (*).\end{aligned}$$

This can be rearranged as

$$\frac{d}{dt}(T + V) = 0.$$

12.7 In cylindrical polar coordinates, the curve $(\rho(\theta), \theta, \alpha\rho(\theta))$ lies on the surface of the cone $z = \alpha\rho$. Show that geodesics (curves of minimum length joining two points) on the cone satisfy

$$\rho^4 = c^2[\beta^2\rho'^2 + \rho^2],$$

where c is an arbitrary constant, but β has to have a particular value. Determine the form of $\rho(\theta)$ and hence find the equation of the shortest path on the cone between the points $(R, -\theta_0, \alpha R)$ and $(R, \theta_0, \alpha R)$.

[You will find it useful to determine the form of the derivative of $\cos^{-1}(u^{-1})$.]

In cylindrical polar coordinates the element of length is given by

$$(ds)^2 = (d\rho)^2 + (\rho\, d\theta)^2 + (dz)^2,$$

and the total length of a curve between two points parameterized by θ_0 and θ_1 is

$$\begin{aligned} s &= \int_{\theta_0}^{\theta_1} \sqrt{\left(\frac{d\rho}{d\theta}\right)^2 + \rho^2 + \left(\frac{dz}{d\theta}\right)^2}\, d\theta \\ &= \int_{\theta_0}^{\theta_1} \sqrt{\rho^2 + (1+\alpha^2)\left(\frac{d\rho}{d\theta}\right)^2}\, d\theta, \text{ since } z = \alpha\rho. \end{aligned}$$

Since the independent variable θ does not occur explicitly in the integrand, a first integral of the EL equation is

$$\sqrt{\rho^2 + (1+\alpha^2)\rho'^2} - \rho' \frac{(1+\alpha^2)\rho'}{\sqrt{\rho^2 + (1+\alpha^2)\rho'^2}} = c.$$

After being multiplied through by the square root, this can be arranged as follows:

$$\begin{aligned} \rho^2 + (1+\alpha^2)\rho'^2 - (1+\alpha^2)\rho'^2 &= c\sqrt{\rho^2 + (1+\alpha^2)\rho'^2}, \\ \rho^4 &= c^2[\rho^2 + (1+\alpha^2)\rho'^2]. \end{aligned}$$

This is the given equation of the geodesic, in which c is arbitrary but β^2 must have the value $1+\alpha^2$.

Guided by the hint, we first determine the derivative of $y(u) = \cos^{-1}(u^{-1})$:

$$\frac{dy}{du} = \frac{-1}{\sqrt{1-u^{-2}}}\,\frac{-1}{u^2} = \frac{1}{u\sqrt{u^2-1}}.$$

Now, returning to the geodesic, rewrite it as

$$\begin{aligned} \rho^4 - c^2\rho^2 &= c^2\beta^2\rho'^2, \\ \rho(\rho^2 - c^2)^{1/2} &= c\beta\frac{d\rho}{d\theta}. \end{aligned}$$

Setting $\rho = cu$,

$$\begin{aligned} uc^2(u^2-1)^{1/2} &= c^2\beta\frac{du}{d\theta}, \\ d\theta &= \frac{\beta\, du}{u(u^2-1)^{1/2}}, \end{aligned}$$

which integrates to

$$\theta = \beta\cos^{-1}\left(\frac{1}{u}\right) + k,$$

using the result from the hint.

Since the geodesic must pass through both $(R, -\theta_0, \alpha R)$ and $(R, \theta_0, \alpha R)$, we must have $k = 0$ and

$$\cos\frac{\theta_0}{\beta} = \frac{c}{R}.$$

Further, at a general point on the geodesic,

$$\cos\frac{\theta}{\beta} = \frac{c}{\rho}.$$

Eliminating c then shows that the geodesic on the cone that joins the two given points is

$$\rho(\theta) = \frac{R\cos(\theta_0/\beta)}{\cos(\theta/\beta)}.$$

12.9 You are provided with a line of length $\pi a/2$ and negligible mass and some lead shot of total mass M. Use a variational method to determine how the lead shot must be distributed along the line if the loaded line is to hang in a circular arc of radius a when its ends are attached to two points at the same height. Measure the distance s along the line from its center.

We first note that the total mass of shot available is merely a scaling factor and not a constraint on the minimization process.

The length of string is sufficient to form one-quarter of a complete circle of radius a, and so the ends of the string must be fixed to points that are $2a\sin(\pi/4) = \sqrt{2}a$ apart.

We take the distribution of shot as $\rho = \rho(s)$ and have to minimize the integral $\int gy(s)\rho(s)\,ds$, but subject to the requirement $\int dx = a/\sqrt{2}$. Expressed as an integral over s, this requirement can be written

$$\frac{a}{\sqrt{2}} = \int_{s=0}^{s=\pi a/4} dx = \int_0^{\pi a/4} (1 - y'^2)^{1/2}\,ds,$$

where the derivative y' of y is with respect to s (not x).

We therefore consider the minimization of $\int F(y, y', s)\,ds$, where

$$F(y, y', s) = gy\rho + \lambda\sqrt{1 - y'^2}.$$

The EL equation takes the form

$$\frac{d}{ds}\left(\frac{\partial F}{\partial y'}\right) = \frac{\partial F}{\partial y},$$

$$\lambda\frac{d}{ds}\left(\frac{-y'}{\sqrt{1 - y'^2}}\right) = g\rho(s),$$

$$\frac{-\lambda y'}{\sqrt{1 - y'^2}} = \int_0^s g\rho(s')\,ds' \equiv gP(s),$$

since $y'(0) = 0$ by symmetry.

Now we require $P(s)$ to be such that the solution to this equation takes the form of an arc of a circle, $y(s) = y_0 - a\cos(s/a)$. If this is so, then $y'(s) = \sin(s/a)$ and

$$\frac{-\lambda\sin(s/a)}{\cos(s/a)} = gP(s).$$

When $s = \pi a/4$, $P(s)$ must have the value $M/2$, implying that $\lambda = -Mg/2$ and that, consequently,

$$P(s) = \frac{M}{2} \tan\left(\frac{s}{a}\right).$$

The required distribution $\rho(s)$ is recovered by differentiating this to obtain

$$\rho(s) = \frac{dP}{ds} = \frac{M}{2a} \sec^2\left(\frac{s}{a}\right).$$

12.11 A general result is that light travels through a variable medium by a path that minimizes the travel time (this is an alternative formulation of Fermat's principle). With respect to a particular cylindrical polar coordinate system (ρ, ϕ, z), the speed of light $v(\rho, \phi)$ is independent of z. If the path of the light is parameterized as $\rho = \rho(z)$, $\phi = \phi(z)$, show that

$$v^2(\rho'^2 + \rho^2\phi'^2 + 1)$$

is constant along the path.

For the particular case when $v = v(\rho) = b(a^2 + \rho^2)^{1/2}$, show that the two Euler–Lagrange equations have a common solution in which the light travels along a helical path given by $\phi = Az + B$, $\rho = C$, provided that A has a particular value.

In cylindrical polar coordinates with $\rho = \rho(z)$ and $\phi = \phi(z)$,

$$ds = \left[1 + \left(\frac{d\rho}{dz}\right)^2 + \rho^2 \left(\frac{d\phi}{dz}\right)^2\right]^{1/2} dz.$$

The total travel time of the light is therefore given by

$$\tau = \int \frac{(1 + \rho'^2 + \rho^2\phi'^2)^{1/2}}{v(\rho, \phi)}\, dz.$$

Since z does not appear explicitly in the integrand, we have from the general first integral of the EL equations for more than one dependent variable that

$$\frac{(1 + \rho'^2 + \rho^2\phi'^2)^{1/2}}{v(\rho, \phi)} - \frac{1}{v}\frac{\rho'^2}{(1 + \rho'^2 + \rho^2\phi'^2)^{1/2}} - \frac{1}{v}\frac{\rho^2\phi'^2}{(1 + \rho'^2 + \rho^2\phi'^2)^{1/2}} = k.$$

Rearranging this gives

$$\begin{aligned} 1 + \rho'^2 + \rho^2\phi'^2 - \rho'^2 - \rho^2\phi'^2 &= kv(1 + \rho'^2 + \rho^2\phi'^2)^{1/2}, \\ 1 &= kv(1 + \rho'^2 + \rho^2\phi'^2)^{1/2}, \\ \Rightarrow \quad v^2(1 + \rho'^2 + \rho^2\phi'^2) &= c, \text{ along the path.} \end{aligned}$$

Denoting $(1 + \rho'^2 + \rho^2\phi'^2)$ by $(**)$ for brevity, the EL equations for ρ and ϕ are, respectively,

$$\frac{\rho\phi'^2}{v(**)^{1/2}} - \frac{(**)^{1/2}}{v^2}\frac{\partial v}{\partial \rho} = \frac{d}{dz}\left[\frac{\rho'}{v(**)^{1/2}}\right], \quad (1)$$

and

$$-\frac{(**)^{1/2}}{v^2}\frac{\partial v}{\partial \phi} = \frac{d}{dz}\left[\frac{\rho^2\phi'}{v(**)^{1/2}}\right]. \quad (2)$$

Now, if $v = b(a^2 + \rho^2)^{1/2}$, the only dependence on z in a possible solution $\phi = Az + B$ with $\rho = C$ is through the first of these equations. To see this we note that the square brackets on the RHSs of the two EL equations do not contain any undifferentiated ϕ-terms and so the derivatives (with respect to z) of both are zero. Since $\partial v/\partial\phi$ is also zero, equation (2) is identically satisfied as $0 = 0$. This leaves only (1), which reads

$$\frac{CA^2}{b(a^2 + C^2)^{1/2}(1 + 0 + C^2A^2)^{1/2}} - \frac{(1 + 0 + C^2A^2)^{1/2}bC}{b^2(a^2 + C^2)(a^2 + C^2)^{1/2}} = 0.$$

This is satisfied provided

$$A^2(a^2 + C^2) = 1 + C^2A^2,$$
$$\text{i.e.} \quad A = a^{-1}.$$

Thus, a solution in the form of a helix is possible provided that the helix has a particular pitch, $2\pi a$.

12.13 The Schwarzchild metric for the static field of a non-rotating spherically symmetric black hole of mass M is given by

$$(ds)^2 = c^2\left(1 - \frac{2GM}{c^2 r}\right)(dt)^2 - \frac{(dr)^2}{1 - 2GM/(c^2 r)} - r^2\,(d\theta)^2 - r^2\sin^2\theta\,(d\phi)^2.$$

Considering only motion confined to the plane $\theta = \pi/2$, and assuming that the path of a small test particle is such as to make $\int ds$ stationary, find two first integrals of the equations of motion. From their Newtonian limits, in which GM/r, $\dot{r}^2$ and $r^2\dot{\phi}^2$ are all $\ll c^2$, identify the constants of integration.

For motion confined to the plane $\theta = \pi/2$, $d\theta = 0$ and the corresponding term in the metric can be ignored. With this simplification, we can write

$$ds = \left\{c^2\left(1 - \frac{2GM}{c^2 r}\right) - \frac{\dot{r}^2}{1 - (2GM)/(c^2 r)} - r^2\dot{\phi}^2\right\}^{1/2} dt.$$

Writing the terms in braces as $\{**\}$, the EL equation for ϕ reads

$$\frac{d}{dt}\left(\frac{-r^2\dot{\phi}}{\{**\}^{1/2}}\right) - 0 = 0,$$
$$\Rightarrow \quad \frac{r^2\dot{\phi}}{\{**\}^{1/2}} = A.$$

In the Newtonian limit $\{**\} \to c^2$ and the equation becomes $r^2\dot{\phi} = Ac$. Thus, Ac is a measure of the angular momentum of the particle about the origin.

The EL equation for r is more complicated but, because ds does not contain t explicitly, we can use the general result for the first integral of the EL equations when there is more than one dependent variable: $F - \sum_i \dot{q}_i \dfrac{\partial F}{\partial \dot{q}_i} = k$. This gives us a second equation as follows:

$$F - \dot{r}\frac{\partial F}{\partial \dot{r}} - \dot{\phi}\frac{\partial F}{\partial \dot{\phi}} = B,$$

$$\{**\}^{1/2} + \frac{\dot{r}}{\{**\}^{1/2}}\frac{\dot{r}}{[1-(2GM)/(c^2 r)]} + \frac{\dot{\phi}}{\{**\}^{1/2}}r^2\dot{\phi} = B.$$

Multiplying through by $\{**\}^{1/2}$ and canceling the terms in $\dot{r}^2$ and $\dot{\phi}^2$ now gives

$$c^2 - \frac{2GM}{r} = B\left\{c^2 - \frac{2GM}{r} - \frac{\dot{r}^2}{[1-(2GM)/(c^2 r)]} - r^2\dot{\phi}^2\right\}^{1/2}.$$

In the Newtonian limits, in which GM/r, $\dot{r}^2$ and $r^2\dot{\phi}^2$ are all $\ll c^2$, the equation can be rearranged and the braces expanded to first order in small quantities to give

$$\begin{aligned}
B &= \left(c^2 - \frac{2GM}{r}\right)\left\{c^2 - \frac{2GM}{r} - \frac{\dot{r}^2}{[1-(2GM)/(c^2 r)]} - r^2\dot{\phi}^2\right\}^{-1/2}, \\
cB &= c^2 - \frac{2GM}{r} + \frac{c^2 GM}{c^2 r} + \frac{c^2\dot{r}^2}{2c^2} + \frac{c^2 r^2\dot{\phi}^2}{2c^2} + \cdots, \\
&= c^2 - \frac{GM}{r} + \frac{1}{2}(\dot{r}^2 + r^2\dot{\phi}^2) + \cdots,
\end{aligned}$$

which can be read as "total energy = rest mass energy + gravitational energy + radial and azimuthal kinetic energy". Thus Bc is a measure of the total energy of the test particle.

12.15 Determine the minimum value that the integral

$$J = \int_0^1 [x^4(y'')^2 + 4x^2(y')^2]\,dx$$

can have, given that y is not singular at $x = 0$ and that $y(1) = y'(1) = 1$. Assume that the Euler–Lagrange equation gives the lower limit.

We first set $y'(x) = u(x)$ with $u(1) = y'(1) = 1$. The integral then becomes

$$J = \int_0^1 [x^4(u')^2 + 4x^2u^2]\,dx. \qquad (*)$$

This will be stationary if (using the EL equation)

$$\begin{aligned}
\frac{d}{dx}(2x^4u') - 8x^2u &= 0, \\
8x^3u' + 2x^4u'' - 8x^2u &= 0, \\
x^2u'' + 4xu' - 4u &= 0.
\end{aligned}$$

As this is a homogeneous equation, we try $u(x) = Ax^n$, obtaining

$$n(n-1) + 4n - 4 = 0 \quad \Rightarrow \quad n = -4, \text{ or } n = 1.$$

The form of $y'(x)$ is thus

$$y'(x) = u(x) = \frac{A}{x^4} + Bx \quad \text{with} \quad A + B = 1.$$

Further,

$$y(x) = -\frac{A}{3x^3} + \frac{Bx^2}{2} + C.$$

Since y is not singular at $x = 0$ and $y(1) = 1$, we have that $A = 0$, $B = 1$ and $C = \frac{1}{2}$, yielding $y(x) = \frac{1}{2}(1 + x^2)$. The minimal value of J is thus

$$J_{\text{min}} = \int_0^1 [x^4(1)^2 + 4x^2(x)^2]\,dx = \int_0^1 5x^4\,dx = \left[x^5\right]_0^1 = 1.$$

12.17 Find an appropriate but simple trial function and use it to estimate the lowest eigenvalue λ_0 of Stokes' equation

$$\frac{d^2y}{dx^2} + \lambda xy = 0, \qquad y(0) = y(\pi) = 0.$$

Explain why your estimate must be strictly greater than λ_0.

Stokes' equation is an SL equation with $p = 1$, $q = 0$ and $\rho = x$. For the given boundary conditions the obvious trial function is $y(x) = \sin x$. The lowest eigenvalue $\lambda_0 \leq I/J$, where

$$I = \int_0^\pi py'^2\,dx = \int_0^\pi \cos^2 x\,dx = \frac{\pi}{2}$$

and

$$\begin{aligned}
J &= \int_0^\pi \rho y^2\,dx = \int_0^\pi x\sin^2 x\,dx \\
&= \int_0^\pi \tfrac{1}{2}x(1 - \cos 2x)\,dx \\
&= \left[\frac{x^2}{4}\right]_0^\pi - \left[\frac{x}{2}\,\frac{\sin 2x}{2}\right]_0^\pi + \frac{1}{2}\int_0^\pi \frac{\sin 2x}{2}\,dx \\
&= \frac{\pi^2}{4}.
\end{aligned}$$

Thus $\lambda_0 \leq (\frac{1}{2}\pi)/(\frac{1}{4}\pi^2) = 2/\pi$.

However, if we substitute the trial function directly into the equation we obtain

$$-\sin x + \frac{2}{\pi}x\sin x = 0,$$

which is clearly not satisfied. Thus the trial function is not an eigenfunction, and the actual lowest eigenvalue must be strictly less than the estimate of $2/\pi$.

12.19 A drum skin is stretched across a fixed circular rim of radius a. Small transverse vibrations of the skin have an amplitude $z(\rho, \phi, t)$ that satisfies

$$\nabla^2 z = \frac{1}{c^2}\frac{\partial^2 z}{\partial t^2}$$

in plane polar coordinates. For a normal mode independent of azimuth, in which case $z = Z(\rho)\cos\omega t$, find the differential equation satisfied by $Z(\rho)$. By using a trial function of the form $a^\nu - \rho^\nu$, with adjustable parameter ν, obtain an estimate for the lowest normal mode frequency.

[The exact answer is $(5.78)^{1/2}c/a$.]

In cylindrical polar coordinates, (ρ, ϕ), the wave equation

$$\nabla^2 z = \frac{1}{c^2}\frac{\partial^2 z}{\partial t^2}$$

has azimuth-independent solutions (i.e. independent of ϕ) of the form $z(\rho, t) = Z(\rho)\cos\omega t$, and reduces to

$$\frac{1}{\rho}\frac{d}{d\rho}\left(\rho\frac{dZ}{d\rho}\right)\cos\omega t = -\frac{Z\omega^2}{c^2}\cos\omega t,$$

$$\frac{d}{d\rho}\left(\rho\frac{dZ}{d\rho}\right) + \frac{\omega^2}{c^2}\rho Z = 0.$$

The boundary conditions require that $Z(a) = 0$ and, so that there is no physical discontinuity in the slope of the drum skin at the origin, $Z'(0) = 0$.

This is an SL equation with $p = \rho$, $q = 0$ and weight function $w = \rho$. A suitable trial function is $Z(\rho) = a^\nu - \rho^\nu$, which automatically satisfies $Z(a) = 0$ and, provided $\nu > 1$, has $Z'(0) = -\nu\rho^{\nu-1}|_{\rho=0} = 0$.

We recall that the lowest eigenfrequency satisfies the general formula

$$\frac{\omega^2}{c^2} \le \frac{\displaystyle\int_0^a [(pZ')^2 - qZ^2]\,d\rho}{\displaystyle\int_0^a wZ^2\,d\rho}.$$

In this case

$$\frac{\omega^2}{c^2} \le \frac{\displaystyle\int_0^a \rho\,\nu^2\rho^{2\nu-2}\,d\rho}{\displaystyle\int_0^a \rho(a^\nu - \rho^\nu)^2\,d\rho}$$

$$= \frac{\displaystyle\int_0^a \nu^2\rho^{2\nu-1}\,d\rho}{\displaystyle\int_0^a (\rho a^{2\nu} - 2\rho^{\nu+1}a^\nu + \rho^{2\nu+1})\,d\rho}$$

$$= \frac{(\nu^2 a^{2\nu})/2\nu}{\dfrac{a^{2\nu+2}}{2} - \dfrac{2a^{2\nu+2}}{\nu+2} + \dfrac{a^{2\nu+2}}{2\nu+2}}$$

$$= \frac{1}{a^2}\frac{\nu(\nu+2)(2\nu+2)}{(\nu+2)(2\nu+2) - 4(2\nu+2) + 2(\nu+2)}$$

$$= \frac{(\nu+2)(\nu+1)}{\nu a^2}.$$

Since ν is an adjustable parameter and we know that, however we choose it, the resulting estimate can *never* be less than the lowest true eigenvalue, we choose the value that minimizes the above estimate. Differentiating the estimate with respect to ν gives

$$\nu(2\nu+3) - (\nu^2 + 3\nu + 2) = 0 \quad \Rightarrow \quad \nu^2 - 2 = 0 \quad \Rightarrow \quad \nu = \sqrt{2}.$$

Thus the least upper bound to be found with this parameterization is

$$\omega^2 \le \frac{c^2}{a^2}\frac{(\sqrt{2}+2)(\sqrt{2}+1)}{\sqrt{2}} = \frac{c^2}{2a^2}(\sqrt{2}+2)^2 \quad \Rightarrow \quad \omega = (5.83)^{1/2}\frac{c}{a}.$$

As noted, the actual lowest eigenfrequency is very little below this.

12.21 For the boundary conditions given below, obtain a functional $\Lambda(y)$ whose stationary values give the eigenvalues of the equation

$$(1+x)\frac{d^2y}{dx^2} + (2+x)\frac{dy}{dx} + \lambda y = 0, \qquad y(0) = 0, \; y'(2) = 0.$$

Derive an approximation to the lowest eigenvalue λ_0 using the trial function $y(x) = xe^{-x/2}$. For what value(s) of γ would

$$y(x) = xe^{-x/2} + \beta \sin \gamma x$$

be a suitable trial function for attempting to obtain an improved estimate of λ_0?

Since the derivative of $1 + x$ is not equal to $2 + x$, the given equation is not in self-adjoint form and an integrating factor for the standard form equation,

$$\frac{d^2y}{dx^2} + \frac{2+x}{1+x}\frac{dy}{dx} + \frac{\lambda y}{1+x} = 0,$$

is needed. This will be

$$\exp\left\{\int^x \frac{2+u}{1+u}\,du\right\} = \exp\left\{\int^x \left(1 + \frac{1}{1+u}\right)du\right\} = e^x(1+x).$$

Thus, after multiplying through by this IF, the equation takes the SL form

$$[(1+x)e^x y']' + \lambda e^x y = 0,$$

with $p(x) = (1+x)e^x$, $q(x) = 0$ and $\rho(x) = e^x$.

The required functional is therefore

$$\Lambda(y) = \frac{\int_0^2 [(1+x)e^x y'^2 + 0]\,dx}{\int_0^2 y^2 e^x\,dx},$$

provided that, for the eigenfunctions y_i of the equation, $\left[y_i\,p(x)y_j'(x)\right]_0^2 = 0$; this condition is automatically satisfied with the given boundary conditions.

For the trial function $y(x) = xe^{-x/2}$, clearly $y(0) = 0$ and, less obviously, $y'(x) = (1 - \frac{1}{2}x)e^{-x/2}$, making $y'(2) = 0$. The functional takes the following form:

$$\begin{aligned}
\Lambda &= \frac{\int_0^2 (1+x)e^x \left(1 - \frac{1}{2}x\right)^2 e^{-x}\,dx}{\int_0^2 x^2 e^{-x} e^x\,dx} \\
&= \frac{\int_0^2 (1+x)\left(1 - \frac{1}{2}x\right)^2\,dx}{\int_0^2 x^2\,dx} \\
&= \frac{\int_0^2 \left(1 - x^2 + \frac{1}{4}x^2 + \frac{1}{4}x^3\right)dx}{8/3} \\
&= \frac{3}{8}\left(2 - \frac{3}{4}\frac{8}{3} + \frac{16}{16}\right) = \frac{3}{8}.
\end{aligned}$$

Thus the lowest eigenvalue is $\leq \frac{3}{8}$.

We already know that $xe^{-x/2}$ is a suitable trial function and thus $y_2(x) = \sin\gamma x$ can be considered on its own. It satisfies $y_2(0) = 0$, but must also satisfy $y_2'(2) = \gamma\cos(2\gamma) = 0$. This requires that $\gamma = \frac{1}{2}(n + \frac{1}{2})\pi$ for some integer n; trial functions with γ of this form can be used to try to obtain a better bound on λ_0 by choosing the best value for n and adjusting the parameter β.

12.23 The unnormalized ground-state (i.e. the lowest-energy) wavefunction of the simple harmonic oscillator of classical frequency ω is $\exp(-\alpha x^2)$, where $\alpha = m\omega/2\hbar$. Take as a trial function the orthogonal wavefunction $x^{2n+1}\exp(-\alpha x^2)$, using the integer n as a variable parameter, and apply either Sturm–Liouville theory or the Rayleigh–Ritz principle to show that the energy of the second lowest state of a quantum harmonic oscillator is $\leq 3\hbar\omega/2$.

We first note that, for n a non-negative integer,

$$\int_{-\infty}^{\infty} x^{2n+1} e^{-\alpha x^2} e^{-\alpha x^2}\,dx = 0$$

on symmetry grounds and so confirm that the ground-state wavefunction, $\exp(-\alpha x^2)$, and the trial function, $\psi_{2n+1} = x^{2n+1}\exp(-\alpha x^2)$, are orthogonal with respect to a unit weight function.

The Hamiltonian for the quantum harmonic oscillator in one dimension is given by

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{k}{2}x^2.$$

This means that to prepare the elements required for a Rayleigh–Ritz analysis we will need to find the second derivative of the trial function and evaluate integrals with integrands of the form $x^n \exp(-2\alpha x^2)$. To this end, define

$$I_n = \int_{-\infty}^{\infty} x^n e^{-2\alpha x^2}\, dx, \text{ with recurrence relation } I_n = \frac{n-1}{4\alpha} I_{n-2}.$$

Using Leibnitz' formula shows that

$$\begin{aligned}\frac{d^2\psi_{2n+1}}{dx^2} &= \big[2n(2n+1)x^{2n-1} + 2(2n+1)(-2\alpha)x^{2n+1} \\ &\quad + (4\alpha^2x^2 - 2\alpha)x^{2n+1}\big]\, e^{-\alpha x^2} \\ &= \big[2n(2n+1)x^{2n-1} - 2(4n+3)\alpha x^{2n+1} + 4\alpha^2 x^{2n+3}\big]\, e^{-\alpha x^2}.\end{aligned}$$

Hence, we find that $\langle H\rangle$ is given by

$$\begin{aligned}&-\frac{\hbar^2}{2m}\int_{-\infty}^{\infty} x^{2n+1}e^{-\alpha x^2}\frac{d^2\psi_{2n+1}}{dx^2}\, dx + \frac{k}{2}\int_{-\infty}^{\infty} x^2 x^{4n+2} e^{-2\alpha x^2}\, dx \\ &= -\frac{\hbar^2}{2m}\big[2n(2n+1)I_{4n} - 2(4n+3)\alpha I_{4n+2} + 4\alpha^2 I_{4n+4}\big] + \frac{k}{2} I_{4n+4} \\ &= I_{4n+2}\left\{-\frac{\hbar^2}{2m}\left[\frac{2n(2n+1)4\alpha}{4n+1} - 2(4n+3)\alpha + \frac{4\alpha^2(4n+3)}{4\alpha}\right] + \frac{k(4n+3)}{8\alpha}\right\},\end{aligned}$$

where we have used the recurrence relation to express all integrals in terms of I_{4n+2}. This has been done because the denominator of the Rayleigh–Ritz quotient is this (same) normalization integral, namely

$$\int_{-\infty}^{\infty} \psi_{2n+1}^* \psi_{2n+1}\, dx = I_{4n+2}.$$

Thus, the estimate $E_{2n+1} = \langle H\rangle / I_{4n+2}$ is given by

$$\begin{aligned}E_{2n+1} &= -\frac{\hbar^2\alpha}{2m}\left(\frac{16n^2 + 8n - 16n^2 - 16n - 3}{4n+1}\right) + \frac{k(4n+3)}{8\alpha} \\ &= \frac{\hbar^2\alpha}{2m}\frac{8n+3}{4n+1} + \frac{k(4n+3)}{8\alpha}.\end{aligned}$$

Using $\omega^2 = \dfrac{k}{m}$ and $\alpha = \dfrac{m\omega}{2\hbar}$ then yields

$$E_{2n+1} = \frac{\hbar\omega}{4}\left(\frac{8n+3}{4n+1} + 4n + 3\right) = \frac{\hbar\omega}{2}\frac{8n^2 + 12n + 3}{4n+1}.$$

For non-negative integers n (the orthogonality requirement is not satisfied for non-integer values), this has a minimum value of $\frac{3}{2}\hbar\omega$ when $n = 0$. Thus the second lowest energy level is less than or equal to this value. In fact, it is equal to this value, as can be shown by substituting ψ_1 into $H\psi = E\psi$.

12.25 The upper and lower surfaces of a film of liquid, which has surface energy per unit area (surface tension) γ and density ρ, have equations $z = p(x)$ and $z = q(x)$, respectively. The film has a given volume V (per unit depth in the y-direction) and lies in the region $-L < x < L$, with $p(0) = q(0) = p(L) = q(L) = 0$. The total energy (per unit depth) of the film consists of its surface energy and its gravitational energy, and is expressed by

$$E = \tfrac{1}{2}\rho g \int_{-L}^{L} (p^2 - q^2)\,dx + \gamma \int_{-L}^{L} \left[(1 + p'^2)^{1/2} + (1 + q'^2)^{1/2}\right] dx.$$

(a) Express V in terms of p and q.
(b) Show that, if the total energy is minimized, p and q must satisfy

$$\frac{p'^2}{(1 + p'^2)^{1/2}} - \frac{q'^2}{(1 + q'^2)^{1/2}} = \text{constant}.$$

(c) As an approximate solution, consider the equations

$$p = a(L - |x|), \qquad q = b(L - |x|),$$

where a and b are sufficiently small that a^3 and b^3 can be neglected compared with unity. Find the values of a and b that minimize E.

(a) The total volume constraint is given simply by

$$V = \int_{-L}^{L} [p(x) - q(x)]\,dx.$$

(b) To take account of the constraint, consider the minimization of $E - \lambda V$, where λ is an undetermined Lagrange multiplier. The integrand does not contain x explicitly and so we have two first integrals of the EL equations, one for $p(x)$ and the other for $q(x)$. They are

$$\frac{1}{2}\rho g(p^2 - q^2) + \gamma(1 + p'^2)^{1/2} + \gamma(1 + q'^2)^{1/2} - \lambda(p - q) - p'\frac{\gamma p'}{(1 + p'^2)^{1/2}} = k_1$$

and

$$\frac{1}{2}\rho g(p^2 - q^2) + \gamma(1 + p'^2)^{1/2} + \gamma(1 + q'^2)^{1/2} - \lambda(p - q) - q'\frac{\gamma q'}{(1 + q'^2)^{1/2}} = k_2.$$

Subtracting these two equations gives

$$\frac{p'^2}{(1 + p'^2)^{1/2}} - \frac{q'^2}{(1 + q'^2)^{1/2}} = \text{constant}.$$

(c) If

$$p = a(L - |x|), \qquad q = b(L - |x|),$$

the derivatives of p and q only take the values $\pm a$ and $\pm b$, respectively, and the volume constraint becomes

$$V = \int_{-L}^{L} (a - b)(L - |x|)\,dx = (a - b)L^2 \quad \Rightarrow \quad b = a - \frac{V}{L^2}.$$

The total energy can now be expressed entirely in terms of a and the given parameters, as follows:

$$\begin{aligned}
E &= \frac{1}{2}\rho g \int_{-L}^{L} (a^2 - b^2)(L - |x|)^2\,dx + 2\gamma L(1 + a^2)^{1/2} + 2\gamma L(1 + b^2)^{1/2} \\
&= \frac{1}{2}\rho g(a^2 - b^2)\frac{2L^3}{3} + 2\gamma L(1 + \tfrac{1}{2}a^2 + 1 + \tfrac{1}{2}b^2) + \mathrm{O}(a^4) + \mathrm{O}(b^4) \\
&\approx \frac{\rho g L^3}{3}\left[a^2 - \left(a - \tfrac{V}{L^2}\right)^2\right] + 2\gamma L\left[2 + \tfrac{1}{2}a^2 + \tfrac{1}{2}\left(a - \tfrac{V}{L^2}\right)^2\right] \\
&= \frac{\rho g L^3}{3}\left(\frac{2aV}{L^2} - \frac{V^2}{L^4}\right) + 2\gamma L\left(2 + a^2 - \frac{aV}{L^2} + \frac{V^2}{2L^4}\right).
\end{aligned}$$

This is minimized with respect to a when

$$\begin{aligned}
\frac{2\rho g L^3 V}{3L^2} + 4\gamma L a - \frac{2\gamma L V}{L^2} &= 0, \\
\Rightarrow \quad a &= \frac{V}{2L^2} - \frac{\rho g V}{6\gamma}, \\
\Rightarrow \quad b &= -\frac{V}{2L^2} - \frac{\rho g V}{6\gamma}.
\end{aligned}$$

As might be expected, $|\,b\,| > |\,a\,|$ and there is more of the liquid below the $z = 0$ plane than there is above it.

13 Integral equations

13.1 Solve the integral equation

$$\int_0^\infty \cos(xv)y(v)\,dv = \exp(-x^2/2)$$

for the function $y = y(x)$ for $x > 0$. Note that for $x < 0$, $y(x)$ can be chosen as is most convenient.

Since $\cos uv$ is an even function of v, we will make $y(-v) = y(v)$ so that the complete integrand is also an even function of v. The integral I on the LHS can then be written as

$$I = \frac{1}{2}\int_{-\infty}^{\infty} \cos(xv)y(v)\,dv = \frac{1}{2}\text{Re}\int_{-\infty}^{\infty} e^{ixv}y(v)\,dv = \frac{1}{2}\int_{-\infty}^{\infty} e^{ixv}y(v)\,dv,$$

the last step following because $y(v)$ is symmetric in v. The integral is now $\sqrt{2\pi}\,\times$ a Fourier transform, and it follows from the inversion theorem for Fourier transforms applied to

$$\frac{1}{2}\int_{-\infty}^{\infty} e^{ixv}y(v)\,dv = \exp(-x^2/2)$$

that

$$\begin{aligned}
y(x) &= \frac{2}{2\pi}\int_{-\infty}^{\infty} e^{-u^2/2}e^{-iux}\,du \\
&= \frac{1}{\pi}\int_{-\infty}^{\infty} e^{-(u+ix)^2/2}e^{-x^2/2}\,dx \\
&= \frac{1}{\pi}\sqrt{2\pi}\,e^{-x^2/2} \\
&= \sqrt{\frac{2}{\pi}}\,e^{-x^2/2}.
\end{aligned}$$

Although, as noted in the question, $y(x)$ is arbitrary for $x < 0$, because its form in this range does not affect the value of the integral, for $x > 0$ it *must* have the form given. This is tricky to prove formally, but any second solution $w(x)$ has to satisfy

$$\int_0^\infty \cos(xv)[y(v) - w(v)]\,dv = 0$$

for all $x > 0$. Intuitively, this implies that $y(x)$ and $w(x)$ are identical functions.

13.3 Convert

$$f(x) = \exp x + \int_0^x (x-y) f(y)\, dy$$

into a differential equation, and hence show that its solution is

$$(\alpha + \beta x)\exp x + \gamma\ \exp(-x),$$

where $\alpha,\ \beta,\ \gamma$ are constants that should be determined.

We differentiate the integral equation twice and obtain

$$f'(x) = e^x + (x-x)f(x) + \int_0^x f(y)\, dy,$$
$$f''(x) = e^x + f(x).$$

Expressed in the usual differential equation form, this last equation is

$$f''(x) - f(x) = e^x, \text{ for which the CF is } f(x) = Ae^x + Be^{-x}.$$

Since the complementary function contains the RHS of the equation, we try as a PI $f(x) = Cxe^x$:

$$Cxe^x + 2Ce^x - Cxe^x = e^x \quad \Rightarrow \quad \beta = C = \tfrac{1}{2}.$$

The general solution is therefore $f(x) = Ae^x + Be^{-x} + \frac{1}{2}xe^x$.

The boundary conditions needed to evaluate A and B are constructed by considering the integral equation and its derivative(s) at $x = 0$, because with $x = 0$ the integral on the RHS contributes nothing. We have

$$f(0) = e^0 + 0 = 1 \quad \Rightarrow \quad A + B = 1$$

and

$$f'(0) = e^0 + 0 = 1 \quad \Rightarrow \quad A - B + \tfrac{1}{2} = 1.$$

Solving these yields $\alpha = A = \frac{3}{4}$ and $\gamma = B = \frac{1}{4}$ and makes the complete solution

$$f(x) = \tfrac{3}{4}e^x + \tfrac{1}{4}e^{-x} + \tfrac{1}{2}xe^x.$$

13.5 Solve for $\phi(x)$ the integral equation

$$\phi(x) = f(x) + \lambda \int_0^1 \left[\left(\frac{x}{y}\right)^n + \left(\frac{y}{x}\right)^n\right] \phi(y)\, dy,$$

where $f(x)$ is bounded for $0 < x < 1$ and $-\frac{1}{2} < n < \frac{1}{2}$, expressing your answer in terms of the quantities $F_m = \int_0^1 f(y)y^m\, dy$.

(a) Give the explicit solution when $\lambda = 1$.
(b) For what values of λ are there no solutions unless $F_{\pm n}$ are in a particular ratio? What is this ratio?

This equation has a symmetric degenerate kernel, and so we set

$$\phi(x) = f(x) + a_1 x^n + a_2 x^{-n},$$

giving

$$\begin{aligned}\frac{\phi(x)-f(x)}{\lambda} &= \int_0^1 \left(\frac{x^n}{y^n}+\frac{y^n}{x^n}\right)[\,f(y)+a_1y^n+a_2y^{-n}\,]\,dy\\ &= x^n\int_0^1 \frac{f(y)}{y^n}\,dy + x^{-n}\int_0^1 y^n f(y)\,dy + a_1x^n\\ &\quad + a_2x^{-n} + a_1x^{-n}\int_0^1 y^{2n}\,dy + a_2x^n\int_0^1 y^{-2n}\,dy\\ &= x^n\left(F_{-n}+a_1+\frac{a_2}{1-2n}\right)+x^{-n}\left(F_n+a_2+\frac{a_1}{2n+1}\right).\end{aligned}$$

This is consistent with the assumed form of $\phi(x)$, provided

$$a_1 = \lambda\left(F_{-n}+a_1+\frac{a_2}{1-2n}\right) \quad\text{and}\quad a_2=\lambda\left(F_n+a_2+\frac{a_1}{2n+1}\right).$$

These two simultaneous linear equations can now be solved for a_1 and a_2.
(a) For $\lambda = 1$, the equations simplify and decouple to yield

$$a_2 = -(1-2n)F_{-n} \quad\text{and}\quad a_1 = -(1+2n)F_n,$$

respectively, giving as the explicit solution

$$\phi(x) = f(x) - (1+2n)F_n x^n - (1-2n)F_{-n}x^{-n}.$$

(b) For a general value of λ,

$$\begin{aligned}(1-\lambda)a_1 - \frac{\lambda}{1-2n}\,a_2 &= \lambda F_{-n},\\ -\frac{\lambda}{1+2n}\,a_1 + (1-\lambda)a_2 &= \lambda F_n.\end{aligned}$$

The case $\lambda = 0$ is trivial, with $\phi(x) = f(x)$, and so suppose that $\lambda \neq 0$. Then, after being divided through by λ, the equations can be written in the matrix and vector form $\mathsf{Aa} = \mathsf{F}$:

$$\begin{pmatrix} \dfrac{1}{\lambda}-1 & -\dfrac{1}{1-2n} \\ -\dfrac{1}{1+2n} & \dfrac{1}{\lambda}-1 \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} F_{-n} \\ F_n \end{pmatrix}.$$

In general, this matrix equation will have no solution if $|\mathsf{A}| = 0$. This will be the case if

$$\left(\frac{1}{\lambda}-1\right)^2 - \frac{1}{1-4n^2} = 0,$$

which, on rearrangement, shows that λ would have to be given by

$$\frac{1}{\lambda} = 1 \pm \frac{1}{\sqrt{1-4n^2}}.$$

We note that this value for λ is real because n lies in the range $-\frac{1}{2} < n < \frac{1}{2}$. In fact $-\infty < \lambda < \frac{1}{2}$. Even for these two values of λ, however, if *either* $F_n = F_{-n} = 0$ *or* the matrix equation

$$\begin{pmatrix} \pm\dfrac{1}{\sqrt{1-4n^2}} & -\dfrac{1}{1-2n} \\ -\dfrac{1}{1+2n} & \pm\dfrac{1}{\sqrt{1-4n^2}} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} F_{-n} \\ F_n \end{pmatrix}$$

is equivalent to two linear equations that are multiples of each other, there will still be a solution. In this latter case, we must have

$$\frac{F_n}{F_{-n}} = \mp\sqrt{\frac{1-2n}{1+2n}}.$$

Again we note that, because of the range in which n lies, this ratio is real; this condition can, however, require any value in the range $-\infty$ to ∞ for F_n/F_{-n}.

13.7 The kernel of the integral equation

$$\psi(x) = \lambda \int_a^b K(x, y)\psi(y)\, dy$$

has the form

$$K(x, y) = \sum_{n=0}^{\infty} h_n(x)g_n(y),$$

where the $h_n(x)$ form a complete orthonormal set of functions over the interval $[a, b]$.

(a) Show that the eigenvalues λ_i are given by

$$|\mathsf{M} - \lambda^{-1}\mathsf{I}| = 0,$$

where M is the matrix with elements

$$M_{kj} = \int_a^b g_k(u)h_j(u)\, du.$$

If the corresponding solutions are $\psi^{(i)}(x) = \sum_{n=0}^{\infty} a_n^{(i)} h_n(x)$, find an expression for $a_n^{(i)}$.

(b) Obtain the eigenvalues and eigenfunctions over the interval $[0, 2\pi]$ if

$$K(x, y) = \sum_{n=1}^{\infty} \frac{1}{n} \cos nx \cos ny.$$

(a) We write the ith eigenfunction as

$$\psi^{(i)}(x) = \sum_{n=0}^{\infty} a_n^{(i)} h_n(x).$$

From the orthonormality of the $h_n(x)$, it follows immediately that

$$a_m^{(i)} = \int_a^b h_m(x)\psi^{(i)}(x)\, dx.$$

However, the coefficients $a_m^{(i)}$ have to be found as the components of the eigenvectors $\mathsf{a}^{(i)}$ defined below, since the $\psi^{(i)}$ are not initially known.

Substituting this assumed form of solution, we obtain

$$\sum_{m=0}^{\infty} a_m^{(i)} h_m(x) = \lambda_i \int_a^b \sum_{n=0}^{\infty} h_n(x) g_n(y) \sum_{l=0}^{\infty} a_l^{(i)} h_l(y)\, dy$$
$$= \lambda_i \sum_{n,l} a_l^{(i)} M_{nl} h_n(x).$$

Since the $\{h_n\}$ are an orthonormal set, it follows that

$$a_m^{(i)} = \lambda_i \sum_{n,l} a_l^{(i)} M_{nl} \delta_{mn} = \lambda_i \sum_{l=0}^{\infty} M_{ml} a_l^{(i)},$$

$$\text{i.e.} \quad (\mathsf{M} - \lambda_i^{-1}\mathsf{I})\mathsf{a}^{(i)} = 0.$$

Thus, the allowed values of λ_i are given by $|\mathsf{M} - \lambda^{-1}\mathsf{I}| = 0$, and the expansion coefficients $a_m^{(i)}$ by the components of the corresponding eigenvectors.
(b) To make the set $\{h_n(x) = \cos nx\}$ into a complete orthonormal set we need to add the set of functions $\{\eta_\nu(x) = \sin \nu x\}$ and then normalize all the functions by multiplying them by $1/\sqrt{\pi}$. For this particular kernel the general functions $g_n(x)$ are given by $g_n(x) = n^{-1}\sqrt{\pi} \cos nx$.

The matrix elements are then

$$M_{kj} = \int_0^{2\pi} \frac{1}{\sqrt{\pi}} \cos ju \frac{\sqrt{\pi}}{k} \cos ku\, du = \frac{\pi}{k}\delta_{kj},$$
$$M_{k\nu} = \int_0^{2\pi} \frac{1}{\sqrt{\pi}} \sin \nu u \frac{\sqrt{\pi}}{k} \cos ku\, du = 0.$$

Thus the matrix M is diagonal and particularly simple. The eigenvalue equation reads

$$\sum_{j=0}^{\infty} \left(\frac{\pi}{k}\delta_{kj} - \lambda_i^{-1}\delta_{kj} \right) a_j^{(i)} = 0,$$

giving the immediate result that $\lambda_k = k/\pi$ with $a_k^{(k)} = 1$ and all other $a_j^{(k)} = a_\nu^{(k)} = 0$. The eigenfunction corresponding to eigenvalue k/π is therefore

$$\psi^{(k)}(x) = h_k(x) = \frac{1}{\sqrt{\pi}} \cos kx.$$

13.9 For $f(t) = \exp(-t^2/2)$, use the relationships of the Fourier transforms of $f'(t)$ and $tf(t)$ to that of $f(t)$ itself to find a simple differential equation satisfied by $\tilde{f}(\omega)$, the Fourier transform of $f(t)$, and hence determine $\tilde{f}(\omega)$ to within a constant. Use this result to solve for $h(t)$ the integral equation

$$\int_{-\infty}^{\infty} e^{-t(t-2x)/2} h(t)\, dt = e^{3x^2/8}.$$

As a standard result,

$$\mathcal{F}\left[\, f'(t) \,\right] = i\omega \tilde{f}(\omega),$$

though we will not need this relationship in the following solution.

From its definition,

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

Now, for the particular given function,

$$\tilde{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-t^2/2}\, e^{-i\omega t}\, dt$$
$$= \frac{1}{\sqrt{2\pi}} \left[\frac{e^{-t^2/2}\, e^{-i\omega t}}{-i\omega} \right]_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{te^{-t^2/2}\, e^{-i\omega t}}{-i\omega}\, dt$$
$$= 0 - \frac{1}{i\omega}\, i \frac{d\tilde{f}}{d\omega}.$$

Hence,

$$\frac{d\tilde{f}}{d\omega} = -\omega \tilde{f} \quad \Rightarrow \quad \ln \tilde{f} = -\tfrac{1}{2}\omega^2 + k \quad \Rightarrow \quad \tilde{f} = Ae^{-\omega^2/2}, \qquad (*)$$

giving $\tilde{f}(\omega)$ to within a multiplicative constant. Now, we are also given

$$\int_{-\infty}^{\infty} e^{-t(t-2x)/2}\, h(t)\, dt = e^{3x^2/8},$$
$$\Rightarrow \quad \int_{-\infty}^{\infty} e^{-(t-x)^2/2}\, e^{x^2/2}\, h(t)\, dt = e^{3x^2/8},$$
$$\Rightarrow \quad \int_{-\infty}^{\infty} e^{-(x-t)^2/2}\, h(t)\, dt = e^{-x^2/8}. \qquad (**)$$

The LHS of $(**)$ is a convolution integral, and so applying the convolution theorem for Fourier transforms and result $(*)$, used twice, yields

$$\sqrt{2\pi}\, Ae^{-\omega^2/2} \tilde{h}(\omega) = \mathcal{F}\left[e^{-(x/2)^2/2} \right] = Ae^{-(2\omega)^2/2},$$
$$\Rightarrow \quad \sqrt{2\pi}\, \tilde{h}(\omega) = e^{-3\omega^2/2} = e^{-(\sqrt{3}\omega)^2/2},$$
$$\Rightarrow \quad h(t) = \frac{1}{\sqrt{2\pi}\, A}\, e^{-(t/\sqrt{3})^2/2} = \frac{1}{\sqrt{2\pi}\, A}\, e^{-t^2/6}.$$

We now substitute in $(**)$ to determine A:

$$\int_{-\infty}^{\infty} e^{-(x-t)^2/2} \frac{1}{\sqrt{2\pi}\, A}\, e^{-t^2/6}\, dt = e^{-x^2/8},$$
$$\frac{1}{\sqrt{2\pi}\, A} \int_{-\infty}^{\infty} e^{-2t^2/3}\, e^{xt}\, e^{-x^2/2}\, e^{x^2/8}\, dt = 1,$$
$$\frac{1}{\sqrt{2\pi}\, A} \int_{-\infty}^{\infty} \exp\left[-\frac{2}{3}\left(t - \frac{3x}{4} \right)^2 \right] dt = 1.$$

From the normalization of the Gaussian integral, this implies that

$$\frac{1}{\sqrt{2\pi}\,A} = \frac{2}{\sqrt{2\pi}\,\sqrt{3}},$$

which in turn means $A = \sqrt{3}/2$, giving finally that

$$h(t) = \sqrt{\frac{2}{3\pi}}\,e^{-t^2/6}.$$

13.11 At an international "peace" conference a large number of delegates are seated around a circular table with each delegation sitting near its allies and diametrically opposite the delegation most bitterly opposed to it. The position of a delegate is denoted by θ, with $0 \le \theta \le 2\pi$. The fury $f(\theta)$ felt by the delegate at θ is the sum of his own natural hostility $h(\theta)$ and the influences on him of each of the other delegates; a delegate at position ϕ contributes an amount $K(\theta - \phi) f(\phi)$. Thus

$$f(\theta) = h(\theta) + \int_0^{2\pi} K(\theta - \phi) f(\phi)\,d\phi.$$

Show that if $K(\psi)$ takes the form $K(\psi) = k_0 + k_1 \cos\psi$ then

$$f(\theta) = h(\theta) + p + q\cos\theta + r\sin\theta$$

and evaluate p, q and r. A positive value for k_1 implies that delegates tend to placate their opponents but upset their allies, whilst negative values imply that they calm their allies but infuriate their opponents. A walkout will occur if $f(\theta)$ exceeds a certain threshold value for some θ. Is this more likely to happen for positive or for negative values of k_1?

Given that $K(\psi) = k_0 + k_1\cos\psi$, we try a solution $f(\theta) = h(\theta) + p + q\cos\theta + r\sin\theta$, reducing the equation to

$$\begin{aligned}
p &+ q\cos\theta + r\sin\theta \\
&= \int_0^{2\pi} [k_0 + k_1(\cos\theta\cos\phi + \sin\theta\sin\phi)] \\
&\quad \times [h(\phi) + p + q\cos\phi + r\sin\phi]\,d\phi \\
&= k_0(H + 2\pi p) + k_1(H_c\cos\theta + H_s\sin\theta + \pi q\cos\theta + \pi r\sin\theta),
\end{aligned}$$

where $H = \int_0^{2\pi} h(z)\,dz$, $H_c = \int_0^{2\pi} h(z)\cos z\,dz$ and $H_s = \int_0^{2\pi} h(z)\sin z\,dz$.

Thus, on equating the constant terms and the coefficients of $\cos\theta$ and $\sin\theta$, we have

$$\begin{aligned}
p &= k_0 H + 2\pi k_0 p \quad &\Rightarrow \quad p &= \frac{k_0 H}{1 - 2\pi k_0},\\
q &= k_1 H_c + k_1 \pi q \quad &\Rightarrow \quad q &= \frac{k_1 H_c}{1 - k_1\pi},\\
r &= k_1 H_s + k_1 \pi r \quad &\Rightarrow \quad r &= \frac{k_1 H_s}{1 - k_1\pi}.
\end{aligned}$$

And so the full solution for $f(\theta)$ is given by

$$\begin{aligned} f(\theta) &= h(\theta) + \frac{k_0 H}{1 - 2\pi k_0} + \frac{k_1 H_c}{1 - k_1 \pi}\cos\theta + \frac{k_1 H_s}{1 - k_1\pi}\sin\theta \\ &= h(\theta) + \frac{k_0 H}{1 - 2\pi k_0} + \frac{k_1}{1 - k_1\pi}\left(H_c^2 + H_s^2\right)^{1/2}\cos(\theta - \alpha), \end{aligned}$$

where $\tan\alpha = H_s/H_c$.

Clearly, the maximum value of $f(\theta)$ will depend upon $h(\theta)$ and its various integrals, but it is most likely to exceed any particular value if k_1 is positive and $\approx \pi^{-1}$. Stick with your friends!

13.13 The operator $\mathcal{M}$ is defined by

$$\mathcal{M}f(x) \equiv \int_{-\infty}^{\infty} K(x, y) f(y)\, dy,$$

where $K(x, y) = 1$ inside the square $|x| < a$, $|y| < a$ and $K(x, y) = 0$ elsewhere. Consider the possible eigenvalues of $\mathcal{M}$ and the eigenfunctions that correspond to them; show that the only possible eigenvalues are 0 and $2a$ and determine the corresponding eigenfunctions. Hence find the general solution of

$$f(x) = g(x) + \lambda \int_{-\infty}^{\infty} K(x, y) f(y)\, dy.$$

From the given properties of $K(x, y)$ we can assert the following.

(i) No matter what the form of $f(x)$, $\mathcal{M}f(x) = 0$ if $|x| > a$.

(ii) All functions for which both $\int_{-a}^{a} f(y)\, dy = 0$ and $f(x) = 0$ for $|x| > a$ are eigenfunctions corresponding to eigenvalue 0.

(iii) For any function $f(x)$, the integral $\int_{-a}^{a} f(y)\, dy$ is equal to a constant whose value is independent of x; thus $f(x)$ can only be an eigenfunction if it is equal to a constant, μ, for $-a \le x \le a$ and is zero otherwise. For this case $\int_{-a}^{a} f(y)\, dy = 2a\mu$ and the eigenvalue is $2a$.

Point (iii) gives the only possible non-zero eigenvalue, whilst point (ii) shows that eigenfunctions corresponding to zero eigenvalues do exist.

Denote by $S(x, a)$ the function that has unit value for $|x| \le a$ and zero value otherwise; $K(x, y)$ could be expressed as $K(x, y) = S(x, a)S(y, a)$. Substitute the trial solution $f(x) = g(x) + kS(x, a)$ into

$$f(x) = g(x) + \lambda \int_{-\infty}^{\infty} K(x, y) f(y)\, dy.$$

This gives

$$\begin{aligned} g(x) + kS(x, a) &= g(x) + \lambda \int_{-\infty}^{\infty} K(x, y)[\, g(y) + kS(y, a)\,]\, dy, \\ kS(x, a) &= \lambda S(x, a) \int_{-a}^{a} g(y)\, dy + \lambda k\, 2a S(x, a). \end{aligned}$$

Here, having replaced $K(x, y)$ by $S(x, a)S(y, a)$, we use the factor $S(y, a)$ to reduce the limits of the y-integration from $\pm\infty$ to $\pm a$. As this result is to hold for all x we must have

$$k = \frac{\lambda G}{1 - 2a\lambda}, \quad \text{where } G = \int_{-a}^{a} g(y)\,dy.$$

The general solution is thus

$$f(x) = \begin{cases} g(x) + \dfrac{\lambda G}{1 - 2a\lambda} & \text{for } |x| \le a, \\ g(x) & \text{for } |x| > a. \end{cases}$$

13.15 Use Fredholm theory to show that, for the kernel

$$K(x, z) = (x + z)\exp(x - z)$$

over the interval [0, 1], the resolvent kernel is

$$R(x, z; \lambda) = \frac{\exp(x - z)\left[(x + z) - \lambda\left(\frac{1}{2}x + \frac{1}{2}z - xz - \frac{1}{3}\right)\right]}{1 - \lambda - \frac{1}{12}\lambda^2},$$

and hence solve

$$y(x) = x^2 + 2\int_0^1 (x + z)\exp(x - z)\,y(z)\,dz,$$

expressing your answer in terms of I_n, where $I_n = \int_0^1 u^n \exp(-u)\,du$.

We calculate successive values of d_n and $D_n(x, z)$ using the Fredholm recurrence relations:

$$d_n = \int_a^b D_{n-1}(x, x)\,dx,$$

$$D_n(x, z) = K(x, z)d_n - n\int_a^b K(x, z_1)D_{n-1}(z_1, z)\,dz_1,$$

starting from $d_0 = 1$ and $D_0(x, z) = (x + z)e^{x-z}$. In the first iteration we obtain

$$d_1 = \int_0^1 (u + u)e^{u-u}\,du = 1,$$

$$\begin{aligned} D_1(x, z) &= (x + z)e^{x-z}(1) - 1\int_0^1 (x + u)e^{x-u}(u + z)e^{u-z}\,du \\ &= (x + z)e^{x-z} - e^{x-z}\int_0^1 [xz + (x + z)u + u^2]\,du \\ &= e^{x-z}\left[\tfrac{1}{2}(x + z) - xz - \tfrac{1}{3}\right]. \end{aligned}$$

Performing the second iteration gives

$$d_2 = \int_0^1 e^{u-u}\left(u - u^2 - \tfrac{1}{3}\right) du = \tfrac{1}{2} - \tfrac{1}{3} - \tfrac{1}{3} = -\tfrac{1}{6},$$

$$\begin{aligned} D_2(x,z) &= (x+z)e^{x-z}\left(-\tfrac{1}{6}\right) \\ &\quad -2\int_0^1 (x+u)e^{x-u}e^{u-z}\left[\tfrac{1}{2}(u+z) - uz - \tfrac{1}{3}\right] du \\ &= e^{x-z}\left\{-\tfrac{1}{6}(x+z) - 2\left[x\left(\tfrac{1}{4} + \tfrac{z}{2} - \tfrac{z}{2} - \tfrac{1}{3}\right) + \left(\tfrac{1}{6} + \tfrac{z}{4} - \tfrac{z}{3} - \tfrac{1}{6}\right)\right]\right\} \\ &= e^{x-z}\left\{-\tfrac{1}{6}(x+z) - 2\left[-\tfrac{x}{12} - \tfrac{z}{12}\right]\right\} = 0. \end{aligned}$$

Since $D_2(x,z) = 0$, $d_3 = 0$, $D_3(x,z) = 0$, etc. Consequently both $D(x,z;\lambda)$ and $d(\lambda)$ are finite, rather than infinite, series:

$$D(x,z;\lambda) = (x+z)e^{x-z} - \lambda\left[\tfrac{1}{2}(x+z) - xz - \tfrac{1}{3}\right]e^{x-z},$$

$$d(\lambda) = 1 - \lambda + \left(-\tfrac{1}{6}\right)\frac{\lambda^2}{2!} = 1 - \lambda - \tfrac{1}{12}\lambda^2.$$

The resolvent kernel $R(x,z;\lambda)$, given by the ratio $D(x,z;\lambda)/d(\lambda)$, is therefore as stated in the question.

For the particular integral equation, $\lambda = 2$ and $f(x) = x^2$. It follows that

$$d(\lambda) = 1 - 2 - \tfrac{4}{12} = -\tfrac{4}{3} \quad \text{and} \quad D(x,z:\lambda) = \left(2xz + \tfrac{2}{3}\right)e^{x-z}.$$

The solution is therefore given by

$$\begin{aligned} y(x) &= f(x) + \lambda\int_0^1 R(x,z;\lambda) f(z)\, dz \\ &= x^2 + 2\int_0^1 \frac{\left(2xz + \frac{2}{3}\right)z^2 e^{x-z}}{-\frac{4}{3}}\, dz \\ &= x^2 - \int_0^1 (3xz^3 + z^2)e^{x-z}\, dz \\ &= x^2 - (3xI_3 + I_2)e^x. \end{aligned}$$

14 Complex variables

14.1 Find an analytic function of $z = x + iy$ whose imaginary part is

$$(y \cos y + x \sin y) \exp x.$$

If the required function is $f(z) = u + iv$, with $v = (y \cos y + x \sin y) \exp x$, then, from the Cauchy–Riemann equations,

$$\frac{\partial v}{\partial x} = e^x (y \cos y + x \sin y + \sin y) = -\frac{\partial u}{\partial y}.$$

Integrating with respect to y gives

$$\begin{aligned} u &= -e^x \int (y \cos y + x \sin y + \sin y)\, dy + f(x) \\ &= -e^x \left(y \sin y - \int \sin y\, dy - x \cos y - \cos y \right) + f(x) \\ &= -e^x (y \sin y + \cos y - x \cos y - \cos y) + f(x) \\ &= e^x (x \cos y - y \sin y) + f(x). \end{aligned}$$

We determine $f(x)$ by applying the second Cauchy–Riemann equation, which equates $\partial u/\partial x$ with $\partial v/\partial y$:

$$\begin{aligned} \frac{\partial u}{\partial x} &= e^x (x \cos y - y \sin y + \cos y) + f'(x), \\ \frac{\partial v}{\partial y} &= e^x (\cos y - y \sin y + x \cos y). \end{aligned}$$

By comparison, $\qquad f'(x) = 0 \quad \Rightarrow \quad f(x) = k,$

where k is a real constant that can be taken as zero. Hence, the analytic function is given by

$$\begin{aligned} f(z) = u + iv &= e^x (x \cos y - y \sin y + iy \cos y + ix \sin y) \\ &= e^x [\,(\cos y + i \sin y)(x + iy)\,] \\ &= e^x\, e^{iy} (x + iy) \\ &= z e^z. \end{aligned}$$

The final line confirms explicitly that this is a function of z alone (as opposed to a function of both z and z^*).

14.3 Find the radii of convergence of the following Taylor series:

$$\text{(a)}\ \sum_{n=2}^{\infty} \frac{z^n}{\ln n}, \qquad \text{(b)}\ \sum_{n=1}^{\infty} \frac{n!z^n}{n^n},$$
$$\text{(c)}\ \sum_{n=1}^{\infty} z^n n^{\ln n}, \quad \text{(d)}\ \sum_{n=1}^{\infty} \left(\frac{n+p}{n}\right)^{n^2} z^n, \ \text{ with } p \text{ real.}$$

In each case we consider the series as $\sum_n a_n z^n$ and apply the formula

$$\frac{1}{R} = \lim_{n\to\infty} |a_n|^{1/n}$$

derived from considering the Cauchy root test for absolute convergence.

(a) $$\frac{1}{R} = \lim_{n\to\infty} \left(\frac{1}{\ln n}\right)^{1/n} = 1, \text{ since } -n^{-1} \ln \ln n \to 0 \text{ as } n \to \infty.$$

Thus $R = 1$. For interest, we also note that at the point $z = 1$ the series is

$$\sum_{n=2}^{\infty} \frac{1}{\ln n} > \sum_{n=2}^{\infty} \frac{1}{n},$$

which diverges. This shows that the given series diverges at this point on its circle of convergence.

(b) $$\frac{1}{R} = \lim_{n\to\infty} \left(\frac{n!}{n^n}\right)^{1/n}.$$

Since the nth root of $n!$ tends to n as $n \to \infty$, the limit of this ratio is that of n/n, namely unity. Thus $R = 1$ and the series converges inside the unit circle.

(c) $$\frac{1}{R} = \lim_{n\to\infty} \left(n^{\ln n}\right)^{1/n} = \lim_{n\to\infty} n^{(\ln n)/n}$$
$$= \lim_{n\to\infty} \exp\left[\frac{\ln n}{n} \ln n\right] = \exp(0) = 1.$$

Thus $R = 1$ and the series converges inside the unit circle. It is obvious that the series diverges at the point $z = 1$.

(d) $$\frac{1}{R} = \lim_{n\to\infty} \left[\left(\frac{n+p}{n}\right)^{n^2}\right]^{1/n} = \lim_{n\to\infty} \left(\frac{n+p}{n}\right)^n$$
$$= \lim_{n\to\infty} \left(1 + \frac{p}{n}\right)^n = e^p.$$

Thus $R = e^{-p}$ and the series converges inside a circle of this radius centered on the origin $z = 0$.

14.5 Determine the types of singularities (if any) possessed by the following functions at $z = 0$ and $z = \infty$:

(a) $(z-2)^{-1}$, (b) $(1+z^3)/z^2$, (c) $\sinh(1/z)$,
(d) e^z/z^3, (e) $z^{1/2}/(1+z^2)^{1/2}$.

(a) Although $(z-2)^{-1}$ has a simple pole at $z = 2$, at both $z = 0$ and $z = \infty$ it is well behaved and analytic.

b) Near $z = 0$, $f(z) = (1+z^3)/z^2$ behaves like $1/z^2$ and so has a double pole there. It is clear that as $z \to \infty$ $f(z)$ behaves as z and so has a simple pole there; this can be made more formal by setting $z = 1/\xi$ to obtain $g(\xi) = \xi^2 + \xi^{-1}$ and considering $\xi \to 0$. This leads to the same conclusion.

(c) As $z \to \infty$, $f(z) = \sinh(1/z)$ behaves like $\sinh\xi$ as $\xi \to 0$, i.e. analytically. However, the definition of the sinh function involves an infinite series – in this case an infinite series of inverse powers of z. Thus, no finite n for which

$$\lim_{z\to 0}[\, z^n f(z)\,] \text{ is finite}$$

can be found, and $f(z)$ has an essential singularity at $z = 0$.

(d) Near $z = 0$, $f(z) = e^z/z^3$ behaves as $1/z^3$ and has a pole of order 3 at the origin. At $z = \infty$ it has an obvious essential singularity; formally, the series expansion of $e^{1/\xi}$ about $\xi = 0$ contains arbitrarily high inverse powers of ξ.

(e) Near $z = 0$, $f(z) = z^{1/2}/(1+z^2)^{1/2}$ behaves as $z^{1/2}$ and therefore has a branch point there. To investigate its behavior as $z \to \infty$, we set $z = 1/\xi$ and obtain

$$f(z) = g(\xi) = \left(\frac{\xi^{-1}}{1+\xi^{-2}}\right)^{1/2} = \left(\frac{\xi}{\xi^2+1}\right)^{1/2} \sim \xi^{1/2} \text{ as } \xi \to 0.$$

Hence $f(z)$ also has a branch point at $z = \infty$.

14.7 Find the real and imaginary parts of the functions (i) z^2, (ii) e^z, and (iii) $\cosh \pi z$. By considering the values taken by these parts on the boundaries of the region $x \geq 0$, $y \leq 1$, determine the solution of Laplace's equation in that region that satisfies the boundary conditions

$$\phi(x,0) = 0, \qquad \phi(0,y) = 0,$$
$$\phi(x,1) = x, \qquad \phi(1,y) = y + \sin \pi y.$$

Writing $f_k(z) = u_k(x,y) + i v_k(x,y)$, we have

$$\begin{aligned}
\text{(i)} \quad f_1(z) &= z^2 = (x+iy)^2 \\
&\Rightarrow u_1 = x^2 - y^2 \text{ and } v_1 = 2xy, \\
\text{(ii)} \quad f_2(z) &= e^z = e^{x+iy} = e^x(\cos y + i \sin y) \\
&\Rightarrow u_2 = e^x \cos y \text{ and } v_2 = e^x \sin y, \\
\text{(iii)} \quad f_3(z) &= \cosh \pi z = \cosh \pi x \cos \pi y + i \sinh \pi x \sin \pi y \\
&\Rightarrow u_3 = \cosh \pi x \cos \pi y \text{ and } v_3 = \sinh \pi x \sin \pi y.
\end{aligned}$$

All of these u and v are necessarily solutions of Laplace's equation (this follows from the Cauchy–Riemann equations), and, since Laplace's equation is linear, we can form any linear combination of them and it will also be a solution. We need to choose the combination that matches the given boundary conditions.

Since the third and fourth conditions involve x and $\sin \pi y$, and these appear only in v_1 and v_3, respectively, let us try a linear combination of them:

$$\phi(x, y) = A(2xy) + B(\sinh \pi x \sin \pi y).$$

The requirement $\phi(x, 0) = 0$ is clearly satisfied, as is $\phi(0, y) = 0$. The condition $\phi(x, 1) = x$ becomes $2Ax + 0 = x$, requiring $A = \frac{1}{2}$, and the remaining condition, $\phi(1, y) = y + \sin \pi y$, takes the form $y + B \sinh \pi \sin \pi y = y + \sin \pi y$, thus determining B as $1/\sinh \pi$.

With ϕ a solution of Laplace's equation and all of the boundary conditions satisfied, the uniqueness theorem guarantees that

$$\phi(x, y) = xy + \frac{\sinh \pi x \sin \pi y}{\sinh \pi}$$

is the correct solution.

14.9 The *fundamental theorem of algebra* states that, for a complex polynomial $p_n(z)$ of degree n, the equation $p_n(z) = 0$ has precisely n complex roots. By applying Liouville's theorem, which reads

If $f(z)$ is analytic and bounded for all z then f is a constant,

to $f(z) = 1/p_n(z)$, prove that $p_n(z) = 0$ has at least one complex root. Factor out that root to obtain $p_{n-1}(z)$ and, by repeating the process, prove the fundamental theorem.

We prove this result by the method of contradiction. Suppose $p_n(z) = 0$ has no roots in the complex plane, then $f_n(z) = 1/p_n(z)$ is bounded for all z and, by Liouville's theorem, is therefore a constant. It follows that $p_n(z)$ is also a constant and that $n = 0$. However, if $n > 0$ we have a contradiction and it was wrong to suppose that $p_n(z) = 0$ has no roots; it must have at least one. Let one of them be $z = z_1$; i.e. $p_n(z)$, being a polynomial, can be written $p_n(z) = (z - z_1)p_{n-1}(z)$.

Now, by considering $f_{n-1}(z) = 1/p_{n-1}(z)$ in just the same way, we can conclude that either $n - 1 = 0$ or a further reduction is possible. It is clear that n such reductions are needed to make f_0 a constant, thus establishing that $p_n(z) = 0$ has precisely n (complex) roots.

14.11 The function

$$f(z) = (1 - z^2)^{1/2}$$

of the complex variable z is defined to be real and positive on the real axis for $-1 < x < 1$. Using cuts running along the real axis for $1 < x < +\infty$ and $-\infty < x < -1$, show how $f(z)$ is made single-valued and evaluate it on the upper and lower sides of both cuts.

Use these results and a suitable contour in the complex z-plane to evaluate the integral

$$I = \int_1^\infty \frac{dx}{x(x^2 - 1)^{1/2}}.$$

Confirm your answer by making the substitution $x = \sec \theta$.

As usual when dealing with branch cuts aimed at making a multi-valued function into a single-valued one, we introduce polar coordinates centered on the branch points. For $f(z)$ the branch points are at $z = \pm 1$, and so we define r_1 as the distance of z from the point 1 and θ_1 as the angle the line joining 1 to z makes with the part of the x-axis for which $1 < x < +\infty$, with $0 \le \theta_1 \le 2\pi$. Similarly, r_2 and θ_2 are centered on the point -1, but θ_2 lies in the range $-\pi \le \theta_2 \le \pi$.

With these definitions,

$$\begin{aligned} f(z) = (1-z^2)^{1/2} &= (1-z)^{1/2}(1+z)^{1/2} \\ &= \left[(-r_1 e^{i\theta_1})(r_2 e^{i\theta_2})\right]^{1/2} \\ &= (r_1 r_2)^{1/2} e^{i(\theta_1+\theta_2-\pi)/2}. \end{aligned}$$

In the final line the choice between $\exp(+i\pi)$ and $\exp(-i\pi)$ for dealing with the minus sign appearing before r_1 in the second line was resolved by the requirement that $f(z)$ is real and positive when $-1 < x < 1$ with $y = 0$. For these values of z, $r_1 = 1 - x$, $r_2 = 1 + x$, $\theta_1 = \pi$ and $\theta_2 = 0$. Thus,

$$f(z) = [(1-x)(1+x)]^{1/2}\, e^{(\pi+0-\pi)/2} = (1-x^2)^{1/2} e^{i0} = +(1-x^2)^{1/2},$$

as required.

Now applying the same prescription to points lying just above and just below each of the cuts, we have

$$\begin{aligned}
x > 1,\ y = 0_+ \qquad & r_1 = x-1 \quad r_2 = x+1 \quad \theta_1 = 0 \quad \theta_2 = 0 \\
\Rightarrow \quad & f(z) = (x^2-1)^{1/2} e^{i(0+0-\pi)/2} = -i(x^2-1)^{1/2}, \\
x > 1,\ y = 0_- \qquad & r_1 = x-1 \quad r_2 = x+1 \quad \theta_1 = 2\pi \quad \theta_2 = 0 \\
\Rightarrow \quad & f(z) = (x^2-1)^{1/2} e^{i(2\pi+0-\pi)/2} = i(x^2-1)^{1/2}, \\
x < -1,\ y = 0_+ \qquad & r_1 = 1-x \quad r_2 = -x-1 \quad \theta_1 = \pi \quad \theta_2 = \pi \\
\Rightarrow \quad & f(z) = (x^2-1)^{1/2} e^{i(\pi+\pi-\pi)/2} = i(x^2-1)^{1/2}, \\
x < -1,\ y = 0_- \qquad & r_1 = 1-x \quad r_2 = -x-1 \quad \theta_1 = \pi \quad \theta_2 = -\pi \\
\Rightarrow \quad & f(z) = (x^2-1)^{1/2} e^{i(\pi-\pi-\pi)/2} = -i(x^2-1)^{1/2}.
\end{aligned}$$

To use these results to evaluate the given integral I, consider the contour integral

$$J = \int_C \frac{dz}{z(1-z^2)^{1/2}} = \int_c \frac{dz}{zf(z)}.$$

Here C is a large circle (consisting of arcs Γ_1 and Γ_2 in the upper and lower half-planes, respectively) of radius R centered on the origin but indented along the positive and negative x-axes by the cuts considered earlier. At the ends of the cuts are two small circles γ_1 and γ_2 that enclose the branch points $z = 1$ and $z = -1$, respectively. Thus the complete closed contour, starting from γ_1 and moving along the positive real axis, consists of, in order, circle γ_1, cut C_1, arc Γ_1, cut C_2, circle γ_2, cut C_3, arc Γ_2 and cut C_4, leading back to γ_1.

On the arcs Γ_1 and Γ_2 the integrand is $O(R^{-2})$ and the contributions to the contour integral $\to 0$ as $R \to \infty$. For the small circle γ_1, where we can set $z = 1 + \rho e^{i\phi}$ with $dz = i\rho e^{i\phi}\, d\phi$, we have

$$\int_{\gamma_1} \frac{dz}{z(1+z)^{1/2}(1-z)^{1/2}} = \int_0^{2\pi} \frac{i\rho e^{i\phi}}{(1+\rho e^{i\phi})(2+\rho e^{i\phi})^{1/2}(-\rho e^{i\phi})^{1/2}}\, d\phi,$$

and this $\to 0$ as $\rho \to 0$. Similarly, the small circle γ_2 contributes nothing to the contour integral. This leaves only the contributions from the four arms of the branch cuts. To relate these to I we use our previous results about the value of $f(z)$ on the various arms:

$$\text{on } C_1,\ z = x \text{ and } \int_{C_1} = \int_1^{\infty} \frac{dx}{x[-i(x^2-1)^{1/2}]} = iI;$$

$$\text{on } C_2,\ z = -x \text{ and } \int_{C_2} = \int_{\infty}^{1} \frac{-dx}{-x[\,i(x^2-1)^{1/2}]} = iI;$$

$$\text{on } C_3,\ z = -x \text{ and } \int_{C_3} = \int_1^{\infty} \frac{-dx}{-x[-i(x^2-1)^{1/2}]} = iI;$$

$$\text{on } C_4,\ z = x \text{ and } \int_{C_1} = \int_{\infty}^{1} \frac{dx}{x[\,i(x^2-1)^{1/2}]} = iI.$$

So the full contour integral around C has the value $4iI$. But, this must be the same as $2\pi i$ times the residue of $z^{-1}(1-z^2)^{-1/2}$ at $z = 0$, which is the only pole of the integrand inside the contour. The residue is clearly unity, and so we deduce that $I = \pi/2$.

This particular integral can be evaluated much more simply using elementary methods. Setting $x = \sec\theta$ with $dx = \sec\theta\tan\theta\, d\theta$ gives

$$\begin{aligned} I &= \int_1^{\infty} \frac{dx}{x(x^2-1)^{1/2}} \\ &= \int_0^{\pi/2} \frac{\sec\theta\tan\theta\, d\theta}{\sec\theta\,(\sec^2\theta - 1)^{1/2}} = \int_0^{\pi/2} d\theta = \frac{\pi}{2}, \end{aligned}$$

and so verifies the result obtained by contour integration.

14.13 The following is an alternative (and roundabout!) way of evaluating the Gaussian integral.

(a) Prove that the integral of $[\exp(i\pi z^2)]\operatorname{cosec}\pi z$ around the parallelogram with corners $\pm 1/2 \pm R\exp(i\pi/4)$ has the value $2i$.
(b) Show that the parts of the contour parallel to the real axis give no contribution when $R \to \infty$.
(c) Evaluate the integrals along the other two sides by putting $z' = r\exp(i\pi/4)$ and working in terms of $z' + \frac{1}{2}$ and $z' - \frac{1}{2}$. Hence by letting $R \to \infty$ show that

$$\int_{-\infty}^{\infty} e^{-\pi r^2}\, dr = 1.$$

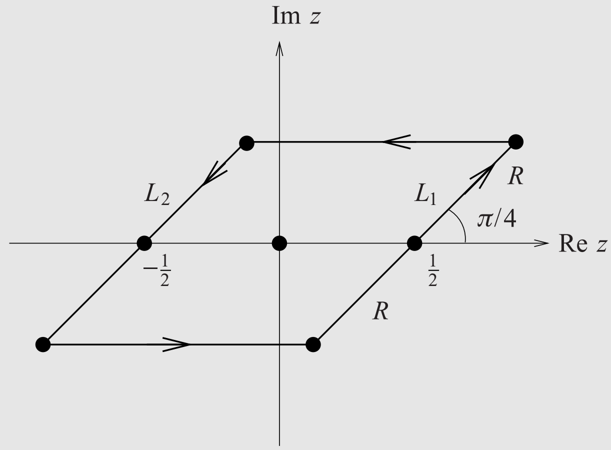

Figure 14.1 The parallelogram contour used in Problem 14.13.

The integral is

$$\int_C e^{i\pi z^2} \operatorname{cosec} \pi z \, dz = \int_C \frac{e^{i\pi z^2}}{\sin \pi z} \, dz$$

and the suggested contour C is shown in Figure 14.1.

(a) The integrand has (simple) poles only on the real axis at $z = n$, where n is an integer. The only such pole enclosed by C is at $z = 0$. The residue there is

$$a_{-1} = \lim_{z \to 0} \frac{z e^{i\pi z^2}}{\sin \pi z} = \frac{1}{\pi}.$$

The value of the integral around C is therefore $2\pi i \times (\pi^{-1}) = 2i$.

(b) On the parts of C parallel to the real axis, $z = \pm R e^{i\pi/4} + x'$, where $-\frac{1}{2} \le x' \le \frac{1}{2}$. The integrand is thus given by

$$\begin{aligned} f(z) &= \frac{1}{\sin \pi z} \exp\left[i\pi \left(\pm R e^{i\pi/4} + x' \right)^2 \right] \\ &= \frac{1}{\sin \pi z} \exp\left[i\pi \left(R^2 e^{i\pi/2} \pm 2Rx' e^{i\pi/4} + x'^2 \right) \right] \\ &= \frac{1}{\sin \pi z} \exp\left[-\pi R^2 \pm \frac{2\pi i R x'}{\sqrt{2}} (1+i) + i\pi x'^2 \right] \\ &= \mathrm{O}\left(\exp[-\pi R^2 \mp \sqrt{2}\pi R x'] \right) \\ &\to 0 \text{ as } R \to \infty. \end{aligned}$$

Since the integration range is finite ($-\frac{1}{2} \le x' \le \frac{1}{2}$), the integrals $\to 0$ as $R \to \infty$.

(c) On the first of the other two sides, let us set $z = \frac{1}{2} + re^{i\pi/4}$ with $-R \le r \le R$. The corresponding integral I_1 is

$$\begin{aligned}
I_1 &= \int_{L_1} e^{i\pi z^2} \operatorname{cosec} \pi z \, dz \\
&= \int_{-R}^{R} \frac{\exp\left[\, i\pi\left(\frac{1}{2} + re^{i\pi/4}\right)^2 \right]}{\sin\left[\pi\left(\frac{1}{2} + re^{i\pi/4}\right)\right]} e^{i\pi/4} \, dr \\
&= \int_{-R}^{R} \frac{e^{i\pi/4} \exp(i\pi re^{i\pi/4}) \exp(i\pi r^2 i) e^{i\pi/4}}{\cos(\pi re^{i\pi/4})} \, dr \\
&= \int_{-R}^{R} \frac{i \exp(i\pi re^{i\pi/4}) e^{-\pi r^2}}{\cos(\pi re^{i\pi/4})} \, dr.
\end{aligned}$$

Similarly (remembering the sense of integration), the remaining side contributes

$$I_2 = -\int_{-R}^{R} \frac{i \exp(-i\pi re^{i\pi/4}) e^{-\pi r^2}}{-\cos(\pi re^{i\pi/4})} \, dr.$$

Adding together all four contributions gives

$$0 + 0 + \int_{-R}^{R} \frac{i[\exp(i\pi re^{i\pi/4}) + \exp(-i\pi re^{i\pi/4})] e^{-\pi r^2}}{\cos(\pi re^{i\pi/4})} \, dr,$$

which simplifies to

$$\int_{-R}^{R} 2ie^{-\pi r^2} \, dr.$$

From part (a), this must be equal to $2i$ as $R \to \infty$, and so $\displaystyle\int_{-\infty}^{\infty} e^{-\pi r^2} \, dr = 1$.

15 Applications of complex variables

Many of the problems in this chapter involve contour integration and the choice of a suitable contour. In order to save the space taken by drawing several broadly similar contours that differ only in notation, the positions of poles, the values of lengths or angles, or other minor details, we make reference to Figure 15.1 which shows a number of typical contour types.

15.1 In the method of complex impedances for a.c. circuits, an inductance L is represented by a complex impedance $Z_L = i\omega L$ and a capacitance C by $Z_C = 1/(i\omega C)$. Kirchhoff's circuit laws,

$$\sum_i I_i = 0 \text{ at a node and } \sum_i Z_i I_i = \sum_j V_j \text{ around any closed loop,}$$

are then applied as if the circuit were a d.c. one.

Apply this method to the a.c. bridge connected as in Figure 15.2 to show that if the resistance R is chosen as $R = (L/C)^{1/2}$ then the amplitude of the current I_R through it is independent of the angular frequency ω of the applied a.c. voltage $V_0\, e^{i\omega t}$.

Determine how the phase of I_R, relative to that of the voltage source, varies with the angular frequency ω.

Omitting the common factor $e^{i\omega t}$ from all currents and voltages, let the current drawn from the voltage source be (the complex quantity) I and the current flowing from A to D be I_1. Then the currents in the remaining branches are $AE:\ I - I_1$, $DB:\ I_1 - I_R$ and $EB:\ I - I_1 + I_R$.

Applying $\sum_i Z_i I_i = \sum_j V_j$ to three separate loops yields

$$\text{loop } ADBA \qquad i\omega L\, I_1 + \frac{1}{i\omega C}(I_1 - I_R) = V_0,$$

$$\text{loop } ADEA \qquad i\omega L\, I_1 + R\, I_R - \frac{1}{i\omega C}(I - I_1) = 0,$$

$$\text{loop } DBED \qquad \frac{1}{i\omega C}(I_1 - I_R) - i\omega L\,(I - I_1 + I_R) - R\, I_R = 0.$$

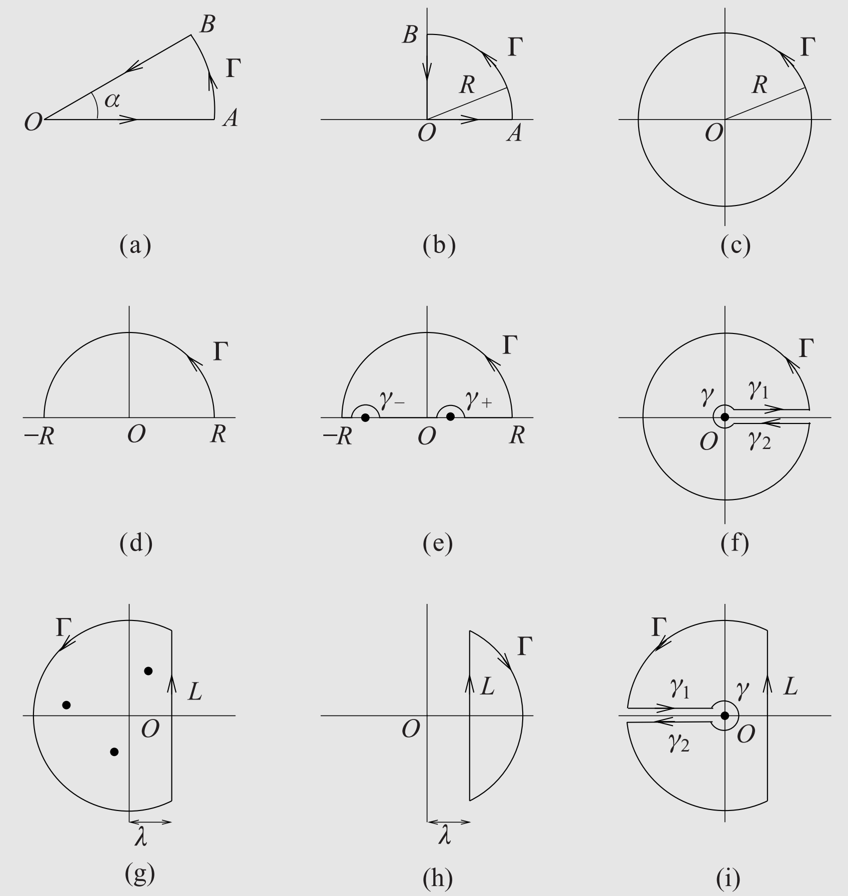

Figure 15.1 Typical contours for use in contour integration.

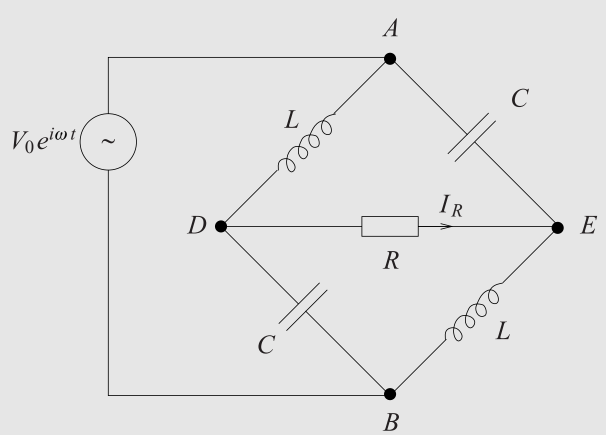

Figure 15.2 The inductor–capacitor–resistor network for Problem 15.1.

Now, denoting $(LC)^{-1}$ by ω_0^2 and choosing R as $(L/C)^{1/2} = (\omega_0 C)^{-1}$, we can write these equations as follows:

$$\left(1 - \frac{\omega^2}{\omega_0^2}\right) I_1 - I_R = i\omega C V_0,$$

$$-I + \left(1 - \frac{\omega^2}{\omega_0^2}\right) I_1 + i\frac{\omega}{\omega_0} I_R = 0,$$

$$\frac{\omega^2}{\omega_0^2} I + \left(1 - \frac{\omega^2}{\omega_0^2}\right) I_1 + \left(-1 + \frac{\omega^2}{\omega_0^2} - i\frac{\omega}{\omega_0}\right) I_R = 0.$$

Eliminating I from the last two of these yields

$$\left(1 + \frac{\omega^2}{\omega_0^2}\right)\left(1 - \frac{\omega^2}{\omega_0^2}\right) I_1 - \left(\frac{i\omega}{\omega_0} + 1\right)\left(1 - \frac{\omega^2}{\omega_0^2}\right) I_R = 0.$$

Thus,

$$I_R = \frac{1 + \dfrac{\omega^2}{\omega_0^2}}{1 + i\dfrac{\omega}{\omega_0}} I_1 = \frac{\omega_0^2 + \omega^2}{\omega_0(\omega_0 + i\omega)} \frac{\omega_0^2(i\omega C V_0 + I_R)}{\omega_0^2 - \omega^2}.$$

After some cancellation and rearrangement,

$$\left(\omega_0^2 - \omega^2\right) I_R = \omega_0(\omega_0 - i\omega)(i\omega C V_0 + I_R),$$
$$(i\omega\omega_0 - \omega^2) I_R = \omega_0\omega(i\omega_0 + \omega) C V_0,$$

and so

$$I_R = \omega_0 C V_0 \frac{i\omega_0 + \omega}{i\omega_0 - \omega} = \omega_0 C V_0 \frac{(i\omega_0 + \omega)(-i\omega_0 - \omega)}{(i\omega_0 - \omega)(-i\omega_0 - \omega)}$$
$$= \omega_0 C V_0 \frac{\omega_0^2 - \omega^2 - 2i\omega\omega_0}{\omega_0^2 + \omega^2}.$$

From this we can read off

$$|I_R| = \omega_0 C V_0 \frac{\left[\left(\omega^2 - \omega_0^2\right)^2 + 4\omega^2\omega_0^2\right]^{1/2}}{\omega_0^2 + \omega^2} = \omega_0 C V_0, \text{ i.e. independent of } \omega,$$

and

$$\phi = \text{phase of } I_R = \tan^{-1} \frac{-2\omega\omega_0}{\omega_0^2 - \omega^2}.$$

Thus I_R (which was arbitrarily and notionally defined as flowing from D to E in the equivalent d.c. circuit) has an imaginary part that is always negative but a real part that changes sign as ω passes through ω_0. Its phase ϕ, relative to that of the voltage source, therefore varies from 0 when ω is small to $-\pi$ when ω is large.

15.3 For the function

$$f(z) = \ln\left(\frac{z+c}{z-c}\right),$$

where c is real, show that the real part u of f is constant on a circle of radius $c \operatorname{cosech} u$ centered on the point $z = c \coth u$. Use this result to show that the electrical capacitance per unit length of two parallel cylinders of radii a, placed with their axes $2d$ apart, is proportional to $[\cosh^{-1}(d/a)]^{-1}$.

From

$$f(z) = \ln\left(\frac{z+c}{z-c}\right) = \ln\left|\frac{z+c}{z-c}\right| + i \arg\left(\frac{z+c}{z-c}\right),$$

we have that

$$u = \ln\left|\frac{z+c}{z-c}\right| = \frac{1}{2}\ln\frac{(x+c)^2+y^2}{(x-c)^2+y^2} \quad \Rightarrow \quad e^{2u} = \frac{(x+c)^2+y^2}{(x-c)^2+y^2}.$$

The curve upon which $u(x, y)$ is constant is therefore given by

$$(x^2 - 2cx + c^2 + y^2)e^{2u} = x^2 + 2xc + c^2 + y^2.$$

This can be rewritten as

$$\begin{aligned} x^2(e^{2u}-1) - 2xc(e^{2u}+1) + y^2(e^{2u}-1) + c^2(e^{2u}-1) &= 0, \\ x^2 - 2xc\frac{e^{2u}+1}{e^{2u}-1} + y^2 + c^2 &= 0, \\ x^2 - 2xc\coth u + y^2 + c^2 &= 0, \end{aligned}$$

which, in conic-section form, becomes

$$(x - c\coth u)^2 + y^2 = c^2\coth^2 u - c^2 = c^2 \operatorname{cosech}^2 u.$$

This is a circle with center $(c \coth u, 0)$ and radius $|c \operatorname{cosech} u|$.

Now consider two such circles with the same value of $|c \operatorname{cosech} u|$, equal to a, but different values of u satisfying $c \coth u_1 = -d$ and $c \coth u_2 = +d$. These two equations imply that $u_1 = -u_2$, corresponding physically to equal but opposite charges $-Q$ and $+Q$ placed on identical cylindrical conductors that coincide with the circles; the conductors are raised to potentials u_1 and u_2.

We have already established that we need $c \coth u_2 = d$ and $c \operatorname{cosech} u_2 = a$. Dividing these two equations gives $\cosh u_2 = d/a$.

The capacitance (per unit length) of the arrangement is given by the magnitude of the charge on one conductor divided by the potential difference between the conductors that results from the presence of that charge, i.e.

$$C = \frac{Q}{u_2 - u_1} \propto \frac{1}{2u_2} = \frac{1}{2\cosh^{-1}(d/a)},$$

as stated in the question.

15.5 By considering in turn the transformations

$$z = \tfrac{1}{2}c(w + w^{-1}) \quad \text{and} \quad w = \exp \zeta,$$

where $z = x + iy$, $w = r \exp i\theta$, $\zeta = \xi + i\eta$ and c is a real positive constant, show that $z = c \cosh \zeta$ maps the strip $\xi \geq 0$, $0 \leq \eta \leq 2\pi$, onto the whole z-plane. Which curves in the z-plane correspond to the lines $\xi =$ constant and $\eta =$ constant? Identify those corresponding to $\xi = 0$, $\eta = 0$ and $\eta = 2\pi$.

The electric potential ϕ of a charged conducting strip $-c \leq x \leq c$, $y = 0$, satisfies

$$\phi \sim -k \ln(x^2 + y^2)^{1/2} \text{ for large values of } (x^2 + y^2)^{1/2},$$

with ϕ constant on the strip. Show that $\phi = \text{Re}\,[-k \cosh^{-1}(z/c)]$ and that the magnitude of the electric field near the strip is $k(c^2 - x^2)^{-1/2}$.

We first note that the combined transformation is given by

$$z = \frac{c}{2}(e^{\zeta} + e^{-\zeta}) = c \cosh \zeta \quad \Rightarrow \quad \zeta = \cosh^{-1} \frac{z}{c}.$$

The successive connections linking the strip in the ζ-plane and its image in the z-plane are

$$\begin{aligned} z &= c \cosh \zeta = c \cosh(\xi + i\eta) \\ &= c \cosh \xi \cos \eta + ic \sinh \xi \sin \eta, \text{ with } \xi > 0, 0 \leq \eta \leq 2\pi, \end{aligned}$$

$$re^{i\theta} = w = e^{\zeta} = e^{\xi} e^{i\eta}, \text{ with the strip as } 1 < r < \infty, 0 \leq \theta \leq 2\pi,$$

$$\begin{aligned} x + iy = z &= \frac{c}{2}(w + w^{-1}) \\ &= \frac{c}{2}[\,r(\cos\theta + i \sin\theta) + r^{-1}(\cos\theta - i \sin\theta)\,] \\ &= \frac{c}{2}\left(r + \frac{1}{r}\right)\cos\theta + i\frac{c}{2}\left(r - \frac{1}{r}\right)\sin\theta. \end{aligned}$$

This last expression for z and the previous specification of the strip in terms of r and θ show that both x and y can take all values, i.e. that the original strip in the ζ-plane is mapped onto the whole of the z-plane. From the two expressions for z we also see that $x = c \cosh \xi \cos \eta$ and $y = c \sinh \xi \sin \eta$.

For ξ constant, the contour in the xy-plane, obtained by eliminating η, is

$$\frac{x^2}{c^2 \cosh^2 \xi} + \frac{y^2}{c^2 \sinh^2 \xi} = 1, \quad \text{i.e. an ellipse.}$$

The eccentricity of the ellipse is given by

$$e = \left(\frac{c^2 \cosh^2 \xi - c^2 \sinh^2 \xi}{c^2 \cosh^2 \xi}\right)^{1/2} = \frac{1}{\cosh \xi}.$$

The foci of the ellipse are at $\pm\, e \times$ the major semi-axis, i.e. $\pm 1/\cosh \xi \times c \cosh \xi = \pm c$. This is independent of ξ and so all the ellipses are confocal.

Similarly, for η constant, the contour is

$$\frac{x^2}{c^2\cos^2\eta} - \frac{y^2}{c^2\sin^2\eta} = 1.$$

This is one of a set of confocal hyperbolae.

(i) $\xi = 0 \quad \Rightarrow \quad y = 0,\ x = c\cos\eta.$

This is the finite line (degenerate ellipse) on the x-axis, $-c \le x \le c$.

(ii) $\eta = 0 \quad \Rightarrow \quad y = 0,\ x = c\cosh\xi.$

This is a part of the x-axis not covered in (i), $c < x < \infty$. The other part, $-\infty < x < -c$, corresponds to $\eta = \pi$.

(iii) This is the same as (the first case) in (ii).

Now, in the ζ-plane, consider the real part of the function $F(\zeta) = -k\zeta$, with k real. On $\xi = 0$ [case (i) above] it reduces to Re $\{-ik\eta\}$, which is zero for all η, i.e. a constant. This implies that the real part of the transformed function will be a constant (actually zero) on $-c \le x \le c$ in the z-plane.

Further,

$$\begin{aligned}(x^2+y^2)^{1/2} &= (c^2\cosh^2\xi\cos^2\eta + c^2\sinh^2\xi\sin^2\eta)^{1/2}\\ &\approx \tfrac{1}{2}ce^{\xi} \text{ for large } \xi,\\ \Rightarrow \quad \xi &\approx \ln(x^2+y^2)^{1/2} + \text{fixed constant.}\end{aligned}$$

Hence,

$$\text{Re}\,\{-k\zeta\} = -k\xi \approx -k\ln(x^2+y^2)^{1/2} \text{ for large } (x^2+y^2)^{1/2}.$$

Thus, the transformation

$$F(\zeta) = -k\zeta \quad \rightarrow \quad f(z) = -k\cosh^{-1}\frac{z}{c}$$

produces a function in the z-plane that satisfies the stated boundary conditions (as well as satisfying Laplace's equation). It is therefore the required solution.

The electric field near the conducting strip, where $y = 0$ and $z^2 = x^2$, can have no component in the x-direction (except at the points $x = \pm c$), but its magnitude is still given by

$$E = |\,f'(z)\,| = \left|-\frac{k}{\sqrt{z^2-c^2}}\right| = \frac{k}{(c^2-x^2)^{1/2}}.$$

15.7 Prove that if $f(z)$ has a simple zero at z_0 then $1/f(z)$ has residue $1/f'(z_0)$ there. Hence evaluate

$$\int_{-\pi}^{\pi} \frac{\sin\theta}{a-\sin\theta}\,d\theta,$$

where a is real and > 1.

If $f(z)$ is analytic and has a simple zero at $z = z_0$ then it can be written as

$$f(z) = \sum_{n=1}^{\infty} a_n(z-z_0)^n, \qquad \text{with } a_1 \neq 0.$$

Using a binomial expansion,

$$\begin{aligned}\frac{1}{f(z)} &= \frac{1}{a_1(z-z_0)\left(1+\sum_{n=2}^{\infty}\frac{a_n}{a_1}(z-z_0)^{n-1}\right)}\\ &= \frac{1}{a_1(z-z_0)}(1+b_1(z-z_0)+b_2(z-z_0)^2+\cdots),\end{aligned}$$

for some coefficients, b_i. The residue at $z = z_0$ is clearly a_1^{-1}.

But, from differentiating the Taylor expansion,

$$\begin{aligned} f'(z) &= \sum_{n=1}^{\infty} na_n(z-z_0)^{n-1},\\ \Rightarrow \quad f'(z_0) &= a_1+0+0+\cdots = a_1,\end{aligned}$$

i.e. the residue $= \dfrac{1}{a_1}$ can also be expressed as $\dfrac{1}{f'(z_0)}$.

Denote the required integral by I and consider the contour integral

$$J = \int_C \frac{dz}{a-\dfrac{1}{2i}(z-z^{-1})} = \int_C \frac{2iz\,dz}{2aiz-z^2+1},$$

where C is the unit circle, i.e. contour (c) of Figure 15.1 with $R = 1$. The denominator has simple zeros at $z = ai \pm \sqrt{-a^2+1} = i(a \pm \sqrt{a^2-1})$. Since a is strictly greater than 1, $\alpha = i(a-\sqrt{a^2-1})$ lies strictly inside the unit circle, whilst $\beta = i(a+\sqrt{a^2-1})$ lies strictly outside it (and need not be considered further).

Extending the previous result to the case of $h(z) = g(z)/f(z)$, where $g(z)$ is analytic at z_0, the residue of $h(z)$ at $z = z_0$ can be seen to be $g(z_0)/f'(z_0)$. Applying this, we find that the residue of the integrand at $z = \alpha$ is given by

$$\left|\frac{2iz}{2ai-2z}\right|_{z=\alpha} = \frac{i\alpha}{ai-ai+i\sqrt{a^2-1}} = \frac{\alpha}{\sqrt{a^2-1}}.$$

Now on the unit circle, $z = e^{i\theta}$ with $dz = i\,e^{i\theta}\,d\theta$, and J can be written as

$$J = \int_{-\pi}^{\pi} \frac{i\,e^{i\theta}\,d\theta}{a-\dfrac{1}{2i}(e^{i\theta}-e^{-i\theta})} = \int_{-\pi}^{\pi} \frac{i(\cos\theta + i\sin\theta)\,d\theta}{a-\sin\theta}.$$

Hence,

$$\begin{aligned} I = -\mathrm{Re}\;J &= -\mathrm{Re}\;2\pi i\frac{i(a-\sqrt{a^2-1})}{\sqrt{a^2-1}}\\ &= 2\pi\left(\frac{a}{\sqrt{a^2-1}}-1\right).\end{aligned}$$

Although it is not asked for, we can also deduce from the fact that the residue at $z = \alpha$ is purely imaginary that

$$\int_{-\pi}^{\pi} \frac{\cos\theta}{a - \sin\theta}\, d\theta = 0,$$

a result that can also be obtained by more elementary means, when it is noted that the numerator of the integrand is the derivative of the denominator.

15.9 Prove that

$$\int_0^{\infty} \frac{\cos mx}{4x^4 + 5x^2 + 1}\, dx = \frac{\pi}{6}\left(4e^{-m/2} - e^{-m}\right) \qquad \text{for } m > 0.$$

Since, when z is on the real axis, the integrand is equal to

$$\text{Re}\, \frac{e^{imz}}{(z^2+1)(4z^2+1)} = \text{Re}\, \frac{e^{imz}}{(z+i)(z-i)(2z+i)(2z-i)},$$

we consider the integral of $f(z) = \dfrac{e^{imz}}{(z+i)(z-i)(2z+i)(2z-i)}$ around contour (d) in Figure 15.1.

As $|f(z)| \sim |z|^{-4}$ as $z \to \infty$ and $m > 0$, all the conditions for Jordan's lemma to hold are satisfied and the integral around the large semi-circle contributes nothing. For this integrand there are two poles inside the contour, at $z = i$ and at $z = \frac{1}{2}i$. The respective residues are

$$\frac{e^{-m}}{2i\; 3i\; i} = \frac{ie^{-m}}{6} \quad \text{and} \quad \frac{e^{-m/2}}{\frac{3i}{2}\left(-\frac{i}{2}\right) 2i} = \frac{-2ie^{-m/2}}{3}.$$

The residue theorem therefore reads

$$\int_{-\infty}^{\infty} \frac{e^{imx}}{4x^4 + 5x^2 + 1}\, dx + 0 = 2\pi i \left(\frac{ie^{-m}}{6} - \frac{2ie^{-m/2}}{3} \right),$$

and the stated result follows from equating real parts and changing the lower integration limit, recognizing that the integrand is symmetric about $x = 0$ and so the integral from 0 to ∞ is equal to half of that from $-\infty$ to ∞.

15.11 Using a suitable cut plane, prove that if α is real and $0 < \alpha < 1$ then

$$\int_0^{\infty} \frac{x^{-\alpha}}{1+x}\, dx$$

has the value $\pi \operatorname{cosec} \pi\alpha$.

As α is not an integer, the complex form of the integrand $f(z) = \dfrac{z^{-\alpha}}{1+z}$ is not single-valued. We therefore need to perform the contour integration in a cut plane; contour (f) of Figure 15.1 is a suitable contour. We will be making use of the fact that, because the integrand takes different values on γ_1 and γ_2, the contributions coming from these two parts of the complete contour, although related, do *not* cancel.

The contributions from γ and Γ are both zero because:

(i) around γ, $|zf(z)| \sim \dfrac{z\,z^{-\alpha}}{1} = z^{1-\alpha} \to 0$ as $|z| \to 0$, since $\alpha < 1$;

(ii) around Γ, $|zf(z)| \sim \dfrac{z\,z^{-\alpha}}{z} = z^{-\alpha} \to 0$ as $|z| \to \infty$, since $\alpha > 0$.

Therefore, the only contributions come from the cut; on γ_1, $z = xe^{0i}$, whilst on γ_2, $z = xe^{2\pi i}$. The only pole inside the contour is a simple one at $z = -1 = e^{i\pi}$, where the residue is $e^{-i\pi\alpha}$. The residue theorem now reads

$$0 + \int_0^\infty \frac{x^{-\alpha}}{1+x}\,dx + 0 - \int_0^\infty \frac{x^{-\alpha}e^{-2\pi i\alpha}}{1+xe^{2\pi i}}\,dx = 2\pi i\,e^{-i\pi\alpha},$$

$$\Rightarrow \quad (1 - e^{-2\pi i\alpha})\int_0^\infty \frac{x^{-\alpha}}{1+x}\,dx = 2\pi i\,e^{-i\pi\alpha}.$$

This can be rearranged to read

$$\int_0^\infty \frac{x^{-\alpha}}{1+x}\,dx = \frac{2\pi i\,e^{-i\pi\alpha}}{(1-e^{-2\pi i\alpha})} = \frac{2\pi i}{e^{i\pi\alpha} - e^{-i\pi\alpha}} = \frac{\pi}{\sin\pi\alpha},$$

thus establishing the stated result.

15.13 By integrating a suitable function around a large semi-circle in the upper half plane and a small semi-circle centered on the origin, determine the value of

$$I = \int_0^\infty \frac{(\ln x)^2}{1+x^2}\,dx$$

and deduce, as a by-product of your calculation, that

$$\int_0^\infty \frac{\ln x}{1+x^2}\,dx = 0.$$

The suggested contour is that shown in Figure 15.1(e), but with only one indentation γ on the real axis (at $z = 0$) and with $R = \infty$. The appropriate complex function is

$$f(z) = \frac{(\ln z)^2}{1+z^2}.$$

The only pole inside the contour is at $z = i$, and the residue there is given by

$$\frac{(\ln i)^2}{i+i} = \frac{(\ln 1 + i(\pi/2))^2}{2i} = -\frac{\pi^2}{8i}.$$

To evaluate the integral around γ, we set $z = \rho\,e^{i\theta}$ with $\ln z = \ln\rho + i\theta$ and $dz = i\rho\,e^{i\theta}\,d\theta$; the integral becomes

$$\int_\pi^0 \frac{\ln^2\rho + 2i\theta\ln\rho - \theta^2}{1+\rho^2e^{2i\theta}}\,i\rho\,e^{i\theta}\,d\theta, \text{ which} \to 0 \text{ as } \rho \to 0.$$

Thus γ contributes nothing. Even more obviously, on Γ, $|zf(z)| \sim z^{-1}$ and tends to zero as $|z| \to \infty$, showing that Γ also contributes nothing.

On γ_+, $z = xe^{i0}$ and the contribution is equal to I.

On γ_-, $z = xe^{i\pi}$ and the contribution is (remembering that the contour actually runs from $x = \infty$ to $x = 0$) given by

$$\begin{aligned} I_- &= -\int_0^\infty \frac{(\ln x + i\pi)^2}{1+x^2}\, e^{i\pi}\, dx \\ &= I + 2i\pi \int_0^\infty \frac{\ln x}{1+x^2}\, dx - \pi^2 \int_0^\infty \frac{1}{1+x^2}\, dx. \end{aligned}$$

The residue theorem for the complete closed contour thus reads

$$0 + I + 0 + I + 2i\pi \int_0^\infty \frac{\ln x}{1+x^2}\, dx - \pi^2 \left[\tan^{-1} x\right]_0^\infty = 2\pi i \left(\frac{-\pi^2}{8i}\right).$$

Equating the real parts $\Rightarrow$ $2I - \frac{1}{2}\pi^3 = -\frac{1}{4}\pi^3$ $\Rightarrow$ $I = \frac{1}{8}\pi^3$.

Equating the imaginary parts gives the stated by-product.

15.15 Prove that

$$\sum_{-\infty}^{\infty} \frac{1}{n^2 + \frac{3}{4}n + \frac{1}{8}} = 4\pi.$$

Carry out the summation numerically, say between -4 and 4, and note how much of the sum comes from values near the poles of the contour integration.

In order to evaluate this sum, we must first find a function of z that takes the value of the corresponding term in the sum whenever z is an integer. Clearly this is

$$\frac{1}{z^2 + \frac{3}{4}z + \frac{1}{8}}.$$

Further, to make use of the properties of contour integrals, we need to multiply this function by one that has simple poles at the same points, each with unit residue. An appropriate choice of integrand is therefore

$$f(z) = \frac{\pi \cot \pi z}{z^2 + \frac{3}{4}z + \frac{1}{8}} = \frac{\pi \cot \pi z}{\left(z + \frac{1}{2}\right)\left(z + \frac{1}{4}\right)}.$$

The contour to be used must enclose all integer values of z, both positive and negative and, in practical terms, must give zero contribution for $|z| \to \infty$, except possibly on the real axis. A large circle C, centered on the origin (see contour (c) in Figure 15.1) suggests itself.

As $|zf(z)| \to 0$ on C, the contour integral has value zero. This implies that the residues at the enclosed poles add up to zero. The residues are

$$\frac{\pi \cot\left(-\frac{1}{2}\pi\right)}{-\frac{1}{2}+\frac{1}{4}} = 0 \qquad \text{at } z = -\tfrac{1}{2},$$
$$\frac{\pi \cot\left(-\frac{1}{4}\pi\right)}{-\frac{1}{4}+\frac{1}{2}} = -4\pi \;\; \text{at } z = -\tfrac{1}{4},$$
$$\sum_{n=-\infty}^{\infty} \frac{1}{\left(n+\frac{1}{2}\right)\left(n+\frac{1}{4}\right)} \qquad \text{at } z = n,\ -\infty < n < \infty.$$

The quoted result follows immediately.

For the rough numerical summation we tabulate n, $D(n) = n^2 + \frac{3}{4}n + \frac{1}{8}$ and the nth term of the series, $1/D(n)$:

n	$D(n)$	$1/D(n)$
-4	13.125	0.076
-3	6.875	0.146
-2	2.625	0.381
-1	0.375	2.667
0	0.125	8.000
1	1.875	0.533
2	5.625	0.178
3	11.375	0.088
4	19.125	0.052

The total of these nine terms is 12.121; this is to be compared with the total for the entire infinite series (of positive terms), which is $4\pi = 12.566$. It will be seen that the sum is dominated by the terms for $n = 0$ and $n = -1$. These two values bracket the positions on the real axis of the poles at $z = -\frac{1}{2}$ and $z = -\frac{1}{4}$.

15.17 By considering the integral of

$$\left(\frac{\sin \alpha z}{\alpha z}\right)^2 \frac{\pi}{\sin \pi z}, \qquad \alpha < \frac{\pi}{2},$$

around a circle of large radius, prove that

$$\sum_{m=1}^{\infty} (-1)^{m-1} \frac{\sin^2 m\alpha}{(m\alpha)^2} = \frac{1}{2}.$$

Denote the given function by $f(z)$ and consider its integral around contour (c) in Figure 15.1.

As $|z| \to \infty$, $\sin \alpha z \sim e^{\alpha|z|}$, and so $f(z) \sim |z|^{-2} e^{2\alpha|z|} e^{-\pi|z|} = z^{-2} e^{(2\alpha-\pi)|z|}$, and, since $\alpha < \frac{1}{2}\pi$, $|zf(z)\,dz| \to 0$ as $|z| \to \infty$ and the integral around the contour has value zero for $R = \infty$.

The function $f(z)$ has simple poles at $z = n$, where n is an integer, $-\infty < n < \infty$. The pole at $z = 0$ is only a first-order pole as the term in parentheses $\to 1$ as $z \to 0$ and has no singularity there. It follows that the sum of the residues of $f(z)$ at all of its poles is zero. For $n \neq 0$, that residue is

$$\begin{aligned}\pi \left(\frac{\sin n\alpha}{n\alpha}\right)^2 \left(\left.\frac{d(\sin \pi z)}{dz}\right|_{z=n}\right)^{-1} &= \left(\frac{\sin n\alpha}{n\alpha}\right)^2 \frac{1}{\cos \pi n} \\ &= (-1)^n \left(\frac{\sin n\alpha}{n\alpha}\right)^2.\end{aligned}$$

For $n = 0$ the residue is 1.

Since the general residue is an even function of n, the sum for $-\infty < n \leq -1$ is equal to that for $1 \leq n < \infty$, and the zero sum of the residues can be written

$$1 + 2\sum_{n=1}^{\infty} (-1)^n \left(\frac{\sin n\alpha}{n\alpha}\right)^2 = 0,$$

leading immediately to the stated result.

15.19 Find the function $f(t)$ whose Laplace transform is

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

Consider first the Taylor series expansion of $\bar{f}(s)$ about $s = 0$:

$$\bar{f}(s) = \frac{e^{-s} - 1 + s}{s^2} = \frac{\left(1 - s + \frac{1}{2}s^2 + \cdots\right) - 1 + s}{s^2} \sim \frac{1}{2} + O(s).$$

Thus $\bar{f}$ has no pole at $s = 0$, and λ in the Bromwich integral can be as small as we wish (but > 0). When the integration line is made part of a closed contour C, the inversion integral becomes

$$f(t) = \int_C \frac{e^{-s}e^{st} - e^{st} + se^{st}}{s^2}\, ds.$$

For $t < 0$, all the terms $\to 0$ as $\text{Re}\, s \to \infty$, and so we close the contour in the right half-plane, as in contour (h) of Figure 15.1. On Γ, s times the integrand $\to 0$, and, as the contour encloses no poles, it follows that the integral along L is zero. Thus $f(t) = 0$ for $t < 0$.

For $t > 1$, all terms $\to 0$ as $\text{Re}\, s \to -\infty$, and so we close the contour in the left half-plane, as in contour (g) of Figure 15.1. On Γ, s times the integrand again $\to 0$, and, as this contour also encloses no poles, it again follows that the integral along L is zero. Thus $f(t) = 0$ for $t > 1$, as well as for $t < 0$.

For $0 < t < 1$, we need to separate the Bromwich integral into two parts (guided by the different ways in which the parts behave as $|s| \to \infty$):

$$f(t) = \int_L \frac{e^{-s}e^{st}}{s^2}\, ds + \int_L \frac{(s-1)e^{st}}{s^2}\, ds \equiv I_1 + I_2.$$

For I_1 the exponent is $s(t-1)$; $t-1$ is negative and so, as in the case $t < 0$, we close the contour in the right half-plane [contour (h)]. No poles are included in this contour, and we conclude that $I_1 = 0$.

For I_2 the exponent is st, indicating that (g) is the appropriate contour. However, $(s-1)/s^2$ does have a pole at $s = 0$ and that is inside the contour. The integral around Γ contributes nothing (that is why it was chosen), and the integral along L must be equal to the residue of $(s-1)e^{st}/s^2$ at $s = 0$. Now,

$$\frac{(s-1)e^{st}}{s^2} = \left(\frac{1}{s} - \frac{1}{s^2}\right)\left(1 + st + \frac{s^2t^2}{2!} + \cdots\right) = -\frac{1}{s^2} + \frac{1}{s}(1-t) + \cdots .$$

The residue, and hence the value of I_2, is therefore $1-t$. Since I_1 has been shown to have value 0, $1-t$ is also the expression for $f(t)$ for $0 < t < 1$.

15.21 Use contour (i) in Figure 15.1 to show that the function with Laplace transform $s^{-1/2}$ is $(\pi x)^{-1/2}$. [For an integrand of the form $r^{-1/2}\exp(-rx)$, change variable to $t = r^{1/2}$.]

With the suggested contour no poles of $s^{-1/2}e^{sx}$ are enclosed and so the integral of $(2\pi i)^{-1}s^{-1/2}e^{sx}$ around the closed curve must have the value zero. It is also clear that the integral along Γ will be zero since $\text{Re}\, s < 0$ on Γ.

For the small circle γ enclosing the origin, set $s = \rho\, e^{i\theta}$, with $ds = i\rho\, e^{i\theta}\, d\theta$, and consider

$$\lim_{\rho\to 0} \int_0^{2\pi} \rho^{-1/2}e^{-i\theta/2} \exp(x\rho\, e^{i\theta}) i\rho\, e^{i\theta}\, d\theta.$$

This $\to 0$ as $\rho \to 0$ (as $\rho^{1/2}$).

On the upper cut, γ_1, $s = re^{i\pi}$ and the contribution to the integral is

$$\frac{1}{2\pi i}\int_\infty^0 \frac{e^{-i\pi/2}}{r^{1/2}} \exp(rxe^{i\pi})e^{i\pi}\, dr,$$

whilst, on the lower cut, γ_2, $s = re^{-i\pi}$, and its contribution to the integral is

$$\frac{1}{2\pi i}\int_0^\infty \frac{e^{i\pi/2}}{r^{1/2}} \exp(rxe^{-i\pi})e^{-i\pi}\, dr.$$

Combining the two (and making both integrals run over the same range) gives

$$-\frac{1}{2\pi i}\int_0^\infty \frac{2i}{r^{1/2}}\,e^{-rx}\,dr = -\frac{1}{\pi}\int_0^\infty \frac{1}{t}\,e^{-t^2x}2t\,dt, \text{ after setting } r = t^2,$$
$$= -\frac{2}{\pi}\frac{\sqrt{\pi}}{2\sqrt{x}}.$$

Since this must add to the Bromwich integral along L to make zero, it follows that the function with Laplace transform $s^{-1/2}$ is $(\pi x)^{-1/2}$.

16 Probability

16.1 By shading or numbering Venn diagrams, determine which of the following are valid relationships between events. For those that are, prove the relationship using de Morgan's laws.

(a) $\overline{(\bar{X} \cup Y)} = X \cap \bar{Y}$.
(b) $\bar{X} \cup \bar{Y} = \overline{(X \cup Y)}$.
(c) $(X \cup Y) \cap Z = (X \cup Z) \cap Y$.
(d) $X \cup \overline{(Y \cap Z)} = (X \cup \bar{Y}) \cap \bar{Z}$.
(e) $X \cup \overline{(Y \cap Z)} = (X \cup \bar{Y}) \cup \bar{Z}$.

For each part of this question we refer to the corresponding part of Figure 16.1.

(a) This relationship is correct as both expressions define the shaded region that is both inside X and outside Y.

(b) This relationship is *not* valid. The LHS specifies the whole sample space *apart from* the region marked with the heavy shading. The RHS defines the region that is lightly shaded. The unmarked regions of X and Y are included in the former but not in the latter.

(c) This relationship is *not* valid. The LHS specifies the sum of the regions marked 2, 3 and 4 in the figure, whilst the RHS defines the sum of the regions marked 1, 3 and 4.

(d) This relationship is *not* valid. On the LHS, $\overline{Y \cap Z}$ is the whole sample space apart from regions 3 and 4. So $X \cup \overline{(Y \cap Z)}$ consists of all regions except for region 3. On the RHS, $X \cup \bar{Y}$ contains all regions except 3 and 7. The events $\bar{Z}$ contain regions 1, 6, 7 and 8 and so $(X \cup \bar{Y}) \cap \bar{Z}$ consists of regions 1, 6 and 8. Thus regions 2, 4, 5 and 7 are in one specification but not in the other.

(e) This relationship is valid. The LHS is as found in (d), namely all regions except for region 3. The RHS consists of the union (as opposed to the intersection) of the two subregions found in (d) and thus contains those regions found in either or both of $X \cup \bar{Y}$ (1, 2, 4, 5, 6 and 8) and $\bar{Z}$ (1, 6, 7 and 8). This covers all regions except region 3 – in agreement with those found for the LHS.

For the two valid relationships, their proofs using de Morgan's laws are:

(a) $$\overline{(\bar{X} \cup Y)} = \bar{\bar{X}} \cap \bar{Y} = X \cap \bar{Y},$$

(e) $$X \cup \overline{(Y \cap Z)} = X \cup (\bar{Y} \cup \bar{Z}) = (X \cup \bar{Y}) \cup \bar{Z}.$$

Here we have also used the result that the complement of the complement of a set is the set itself.

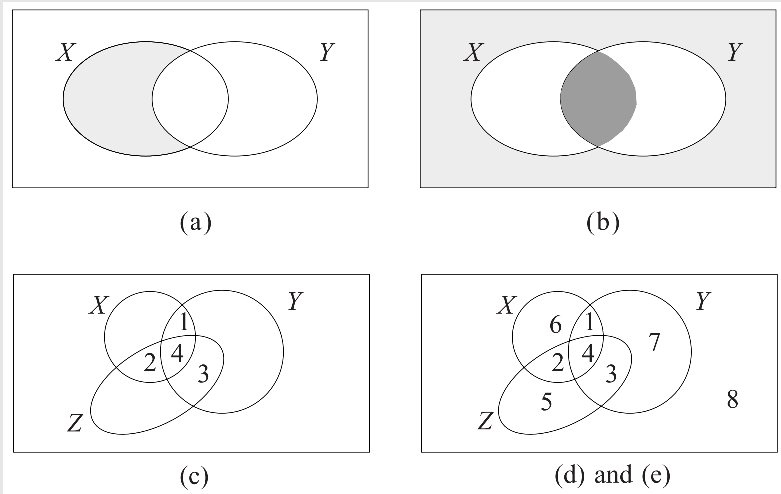

Figure 16.1 The Venn diagrams used in Problem 16.1.

16.3 A and B each have two unbiased four-faced dice, the four faces being numbered 1, 2, 3 and 4. Without looking, B tries to guess the sum x of the numbers on the bottom faces of A's two dice after they have been thrown onto a table. If the guess is correct B receives x^2 euros, but if not he loses x euros.

Determine B's expected gain per throw of A's dice when he adopts each of the following strategies:

(a) he selects x at random in the range $2 \le x \le 8$;
(b) he throws his own two dice and guesses x to be whatever they indicate;
(c) he takes your advice and always chooses the same value for x. Which number would you advise?

We first calculate the probabilities $p(x)$ and the corresponding gains $g(x) = p(x)x^2 - [1 - p(x)]x$ for each value of the total x. Expressing both in units of $1/16$, they are as follows:

x	2	3	4	5	6	7	8
$p(x)$	1	2	3	4	3	2	1
$g(x)$	−26	−24	−4	40	30	0	−56

(a) If B's guess is random in the range $2 \le x \le 8$ then his expected return is

$$\frac{1}{16}\frac{1}{7}(-26 - 24 - 4 + 40 + 30 + 0 - 56) = -\frac{40}{112} = -0.36 \text{ euros.}$$

(b) If he picks by throwing his own dice then his distribution of guesses is the same as that of $p(x)$ and his expected return is

$$\frac{1}{16}\frac{1}{16}[1(-26)+2(-24)+3(-4)+4(40)+3(30)+2(0)+1(-56)]$$
$$=\frac{108}{256}=0.42 \text{ euros.}$$

(c) If B chooses y, then his expected return is

$$h(y)=p(y)y^2-\sum_{x\neq y}p(x)x.$$

An additional line in the table (in the same units) would read $h(x)$, -74, -56, -20, 40, 46, 32, -8. You should not advise B, but take his place, guess "6" each time, and expect an average profit of $46/16 = 2.87$ euros per throw.

16.5 Two duelists, A and B, take alternate shots at each other, and the duel is over when a shot (fatal or otherwise!) hits its target. Each shot fired by A has a probability α of hitting B, and each shot fired by B has a probability β of hitting A. Calculate the probabilities P_1 and P_2, defined as follows, that A will win such a duel: P_1, A fires the first shot; P_2, B fires the first shot.

If they agree to fire simultaneously, rather than alternately, what is the probability P_3 that A will win, i.e. hit B without being hit himself?

Each shot has only two possible outcomes, a hit or a miss. P_1 is the probability that A will win when it is his turn to fire the next shot, and he is still able to do so (event W). There are three possible outcomes of the first two shots: C_1, A hits with his shot; C_2, A misses but B hits; C_3, both miss. Thus

$$\begin{aligned}
P_1 &= \sum_i \Pr(C_i)\Pr(W|C_i)\\
&= [\alpha\times 1]+[(1-\alpha)\beta\times 0]+[(1-\alpha)(1-\beta)\times P_1]\\
\Rightarrow\quad P_1 &= \frac{\alpha}{\alpha+\beta-\alpha\beta}.
\end{aligned}$$

When B fires first but misses, the situation is the one just considered. But if B hits with his first shot then clearly A's chances of winning are zero. Since these are the only two possible outcomes of B's first shot, we can write

$$P_2=[\beta\times 0]+[(1-\beta)\times P_1]\quad\Rightarrow\quad P_2=\frac{(1-\beta)\alpha}{\alpha+\beta-\alpha\beta}.$$

When both fire at the same time there are four possible outcomes D_i to the first round: D_1, A hits and B misses; D_2, B hits but A misses; D_3, they both hit; D_4, they both miss. If getting hit, even if you manage to hit your opponent, does not count as a win, then

$$\begin{aligned}
P_3 &= \sum_i \Pr(D_i)\Pr(W|D_i)\\
&= [\alpha(1-\beta)\times 1]+[(1-\alpha)\beta\times 0]+[\alpha\beta\times 0]+[(1-\alpha)(1-\beta)\times P_3].
\end{aligned}$$

This can be rearranged as

$$P_3 = \frac{\alpha(1-\beta)}{\alpha+\beta-\alpha\beta} = P_2.$$

Thus the result is the same as if B had fired first. However, we also note that if all that matters to A is that B is hit, whether or not he is hit himself, then the third bracket takes the value $\alpha\beta \times 1$ and P_3 takes the same value as P_1.

16.7 An electronics assembly firm buys its microchips from three different suppliers; half of them are bought from firm X, whilst firms Y and Z supply 30% and 20%, respectively. The suppliers use different quality-control procedures and the percentages of defective chips are 2%, 4% and 4% for X, Y and Z, respectively. The probabilities that a defective chip will fail two or more assembly-line tests are 40%, 60% and 80%, respectively, whilst all defective chips have a 10% chance of escaping detection. An assembler finds a chip that fails only one test. What is the probability that it came from supplier X?

Since the number of tests failed by a defective chip are mutually exclusive outcomes (0, 1 or ≥ 2), a chip supplied by X has a probability of failing just one test given by $0.02(1-0.1-0.4) = 0.010$. The corresponding probabilities for chips supplied by Y and Z are $0.04(1-0.1-0.6) = 0.012$ and $0.04(1-0.1-0.8) = 0.004$, respectively.

Using "1" to denote failing a single test, Bayes' theorem gives the probability that the chip was supplied by X as

$$\begin{aligned}\Pr(X|1) &= \frac{\Pr(1|X)\Pr(X)}{\Pr(1|X)\Pr(X)+\Pr(1|Y)\Pr(Y)+\Pr(1|Z)\Pr(Z)} \\ &= \frac{0.010\times 0.5}{0.010\times 0.5+0.012\times 0.3+0.004\times 0.2} = \frac{50}{94}.\end{aligned}$$

16.9 A boy is selected at random from amongst the children belonging to families with n children. It is known that he has at least two sisters. Show that the probability that he has $k-1$ brothers is

$$\frac{(n-1)!}{(2^{n-1}-n)(k-1)!(n-k)!},$$

for $1 \leq k \leq n-2$ and zero for other values of k. Assume that boys and girls are equally likely.

The boy has $n-1$ siblings. Let A_j be the event that $j-1$ of them are brothers, i.e. his family contains j boys and $n-j$ girls. The probability of event A_j is

$$\Pr(A_j) = \frac{{}^{n-1}C_{j-1}\left(\frac{1}{2}\right)^{n-1}}{\sum_{j=1}^{n} {}^{n-1}C_{j-1}\left(\frac{1}{2}\right)^{n-1}} = \frac{(n-1)!}{2^{n-1}(j-1)!(n-j)!}.$$

If B is the event that the boy has at least two sisters, then

$$\Pr(B|A_j) = \begin{cases} 1 & 1 \leq j \leq n-2, \\ 0 & n-1 \leq j \leq n. \end{cases}$$

Now we apply Bayes' theorem to give the probability that he has $k-1$ brothers:

$$\Pr(A_k|B) = \frac{1\ \Pr(A_k)}{\sum_{j=1}^{n-2} 1\ \Pr(A_j)},$$

for $1 \le k \le n-2$. The denominator of this expression is the sum $1 = (\frac{1}{2} + \frac{1}{2})^{n-1} = \sum_{j=1}^{n} {}^{n-1}C_{j-1} \left(\frac{1}{2}\right)^{n-1}$, but omitting the $j = n-1$ and the $j = n$ terms, and so is equal to

$$1 - \frac{(n-1)!}{2^{n-1}\,(n-2)!\,1!} - \frac{(n-1)!}{2^{n-1}\,(n-1)!\,0!} = \frac{1}{2^{n-1}}\left[2^{n-1} - (n-1) - 1\right].$$

Thus,

$$\Pr(A_k|B) = \frac{(n-1)!}{2^{n-1}(k-1)!(n-k)!}\,\frac{2^{n-1}}{2^{n-1} - n} = \frac{(n-1)!}{(2^{n-1} - n)(k-1)!(n-k)!},$$

as given in the question.

16.11 A set of $2N+1$ rods consists of one of each integer length $1, 2, \ldots, 2N, 2N+1$. Three, of lengths a, b and c, are selected, of which a is the longest. By considering the possible values of b and c, determine the number of ways in which a non-degenerate triangle (i.e. one of non-zero area) can be formed (i) if a is even, and (ii) if a is odd. Combine these results appropriately to determine the total number of non-degenerate triangles that can be formed using three of the $2N+1$ rods, and hence show that the probability that such a triangle can be formed from a random selection (without replacement) of three rods is

$$\frac{(N-1)(4N+1)}{2(4N^2-1)}.$$

Rod a is the longest of the three rods. As no two are the same length, let $a > b > c$. To form a non-degenerate triangle we require that $b + c > a$, and, in consequence, $4 \le a \le 2N+1$.

(i) With a even. Consider each b $(< a)$ in turn and determine how many values of c allow a triangle to be made:

b	Values of c	Number of c values
$a-1$	$2,\ 3,\ \cdots,\ a-2$	$a-3$
$a-2$	$3,\ 4,\ \cdots,\ a-3$	$a-5$
$\cdots$	$\cdots$	$\cdots$
$\frac{1}{2}a+1$	$\frac{1}{2}a$	1

Thus, there are $1 + 3 + 5 + \cdots + (a-3)$ possible triangles when a is even.
(ii) A table for odd a is similar, except that the last line will read $b = \frac{1}{2}(a+3)$, $c = \frac{1}{2}(a-1)$ or $\frac{1}{2}(a+1)$, and the number of c values $= 2$. Thus there are $2 + 4 + 6 + \cdots + (a-3)$ possible triangles when a is odd.

To find the total number $n(N)$ of possible triangles, we group together the cases $a = 2m$ and $a = 2m+1$, where $m = 1, 2, \ldots, N$. Then,

$$\begin{aligned}
n(N) &= \sum_{m=2}^{N}[1+3+\cdots+(2m-3)] + [2+4+\cdots+(2m+1-3)] \\
&= \sum_{m=2}^{N}\sum_{k=1}^{2m-2} k = \sum_{m=2}^{N}\tfrac{1}{2}(2m-2)(2m-1) = \textstyle\sum_{m=2}^{N} 2m^2 - 3m + 1 \\
&= 2\left[\tfrac{1}{6}N(N+1)(2N+1) - 1\right] - 3\left[\tfrac{1}{2}N(N+1) - 1\right] + N - 1 \\
&= \frac{N}{6}[2(N+1)(2N+1) - 9(N+1) + 6] \\
&= \frac{N}{6}(4N^2 - 3N - 1) = \frac{N}{6}(4N+1)(N-1).
\end{aligned}$$

The number of ways that three rods can be drawn at random (without replacement) is $(2N+1)(2N)(2N-1)/3!$ and so the probability that they can form a triangle is

$$\frac{N(4N+1)(N-1)}{6}\;\frac{3!}{(2N+1)(2N)(2N-1)} = \frac{(N-1)(4N+1)}{2(4N^2-1)},$$

as stated in the question.

16.13 The duration (in minutes) of a telephone call made from a public call-box is a random variable T. The probability density function of T is

$$f(t) = \begin{cases} 0 & t < 0, \\ \frac{1}{2} & 0 \le t < 1, \\ ke^{-2t} & t \ge 1, \end{cases}$$

where k is a constant. To pay for the call, 20 pence has to be inserted at the beginning, and a further 20 pence after each subsequent half-minute. Determine by how much the average cost of a call exceeds the cost of a call of average length charged at 40 pence per minute.

From the normalization of the PDF, we must have

$$1 = \int_0^\infty f(t)\,dt = \frac{1}{2} + \int_1^\infty ke^{-2t}\,dt = \frac{1}{2} + \frac{ke^{-2}}{2} \quad \Rightarrow \quad k = e^2.$$

The average length of a call is given by

$$\begin{aligned}
\bar{t} &= \int_0^1 t\,\frac{1}{2}\,dt + \int_1^\infty t\,e^2 e^{-2t}\,dt \\
&= \frac{1}{2}\,\frac{1}{2} + \left[\frac{te^2e^{-2t}}{-2}\right]_1^\infty + \int_1^\infty \frac{e^2\,e^{-2t}}{2}\,dt = \frac{1}{4} + \frac{1}{2} + \frac{e^2}{2}\left[\frac{e^{-2t}}{-2}\right]_1^\infty = \frac{3}{4} + \frac{1}{4} = 1.
\end{aligned}$$

Let $p_n = \Pr\{\frac{1}{2}(n-1) < t < \frac{1}{2}n\}$. The corresponding cost is $c_n = 20n$.
Clearly, $p_1 = p_2 = \frac{1}{4}$ and, for $n > 2$,

$$p_n = e^2 \int_{(n-1)/2}^{n/2} e^{-2t}\,dt = e^2\left[\frac{e^{-2t}}{-2}\right]_{(n-1)/2}^{n/2} = \frac{1}{2}e^2(e-1)e^{-n}.$$

The average cost of a call is therefore

$$\bar{c} = 20\left[\frac{1}{4} + 2\frac{1}{4} + \sum_{n=3}^{\infty} n\frac{1}{2}e^2(e-1)e^{-n}\right] = 15 + 10e^2(e-1)\sum_{n=3}^{\infty} ne^{-n}.$$

Now, the final summation might be recognized as part of an arithmetico-geometric series whose sum can be found from the standard formula

$$S = \frac{a}{1-r} + \frac{rd}{(1-r)^2},$$

with $a = 0, d = 1$ and $r = e^{-1}$, or could be evaluated directly by noting that as a geometric series,

$$\sum_{n=0}^{\infty} e^{-nx} = \frac{1}{1-e^{-x}}.$$

Differentiating this with respect to x and then setting $x = 1$ gives

$$-\sum_{n=0}^{\infty} ne^{-nx} = -\frac{e^{-x}}{(1-e^{-x})^2} \quad \Rightarrow \quad \sum_{n=0}^{\infty} ne^{-n} = \frac{e^{-1}}{(1-e^{-1})^2}.$$

From either method it follows that

$$\begin{aligned}\sum_{n=3}^{\infty} ne^{-n} &= \frac{e}{(e-1)^2} - e^{-1} - 2e^{-2}\\ &= \frac{e - e + 2 - e^{-1} - 2 + 4e^{-1} - 2e^{-2}}{(e-1)^2} = \frac{3e^{-1} - 2e^{-2}}{(e-1)^2}.\end{aligned}$$

The total charge therefore exceeds that of a call of average length (1 minute) charged at 40 pence per minute by the amount (in pence)

$$15 + 10e^2(e-1)\frac{3e^{-1} - 2e^{-2}}{(e-1)^2} - 40 = \frac{10(3e-2) - 25e + 25}{e-1} = \frac{5e+5}{e-1} = 10.82.$$

16.15 A tennis tournament is arranged on a straight knockout basis for 2^n players, and for each round, except the final, opponents for those still in the competition are drawn at random. The quality of the field is so even that in any match it is equally likely that either player will win. Two of the players have surnames that begin with "Q". Find the probabilities that they play each other

(a) in the final,
(b) at some stage in the tournament.

Let p_r be the probability that *before* the rth round the two players are both still in the tournament (and, by implication, have not met each other). Clearly, $p_1 = 1$.

Before the rth round there are 2^{n+1-r} players left in. For both "Q" players to still be in before the $(r+1)$th round, Q_1 must avoid Q_2 in the draw and both must win their matches. Thus

$$p_{r+1} = \frac{2^{n+1-r} - 2}{2^{n+1-r} - 1}\left(\frac{1}{2}\right)^2 p_r.$$

(a) The probability that they meet in the final is p_n, given by

$$\begin{aligned}
p_n &= 1\,\frac{2^n-2}{2^n-1}\,\frac{1}{4}\,\frac{2^{n-1}-2}{2^{n-1}-1}\,\frac{1}{4}\cdots\frac{2^2-2}{2^2-1}\,\frac{1}{4}\\
&= \left(\frac{1}{4}\right)^{n-1} 2^{n-1}\left[\frac{(2^{n-1}-1)(2^{n-2}-1)\cdots(2^1-1)}{(2^n-1)(2^{n-1}-1)\cdots(2^2-1)}\right]\\
&= \left(\frac{1}{4}\right)^{n-1} 2^{n-1}\,\frac{1}{2^n-1}\\
&= \frac{1}{2^{n-1}(2^n-1)}.
\end{aligned}$$

(b) The more general solution to the recurrence relation derived above is

$$\begin{aligned}
p_r &= 1\,\frac{2^n-2}{2^n-1}\,\frac{1}{4}\,\frac{2^{n-1}-2}{2^{n-1}-1}\,\frac{1}{4}\cdots\frac{2^{n+2-r}-2}{2^{n+2-r}-1}\,\frac{1}{4}\\
&= \left(\frac{1}{4}\right)^{r-1} 2^{r-1}\left[\frac{(2^{n-1}-1)(2^{n-2}-1)\cdots(2^{n+1-r}-1)}{(2^n-1)(2^{n-1}-1)\cdots(2^{n+2-r}-1)}\right]\\
&= \left(\frac{1}{2}\right)^{r-1}\frac{2^{n+1-r}-1}{2^n-1}.
\end{aligned}$$

Before the rth round, if they are both still in the tournament, the probability that they will be drawn against each other is $(2^{n-r+1}-1)^{-1}$. Consequently, the chance that they will meet at *some* stage is

$$\begin{aligned}
\sum_{r=1}^{n} p_r\,\frac{1}{2^{n-r+1}-1} &= \sum_{r=1}^{n}\left(\frac{1}{2}\right)^{r-1}\frac{2^{n+1-r}-1}{2^n-1}\,\frac{1}{2^{n-r+1}-1}\\
&= \frac{1}{2^n-1}\sum_{r=1}^{n}\left(\frac{1}{2}\right)^{r-1}\\
&= \frac{1}{2^n-1}\,\frac{1-(\frac{1}{2})^n}{1-\frac{1}{2}} = \frac{1}{2^{n-1}}.
\end{aligned}$$

This same conclusion can also be reached in the following way.

The probability that Q_1 *is not* put out of (i.e. wins) the tournament is $(\frac{1}{2})^n$. It follows that the probability that Q_1 *is* put out is $1-(\frac{1}{2})^n$ and that the player responsible is Q_2 with probability $[1-(\frac{1}{2})^n]/(2^n-1) = 2^{-n}$. Similarly, the probability that Q_2 is put out and that the player responsible is Q_1 is also 2^{-n}. These are exclusive events but cover all cases in which Q_1 and Q_2 meet during the tournament, the probability of which is therefore $2 \times 2^{-n} = 2^{1-n}$.

16.17 This problem is about interrelated binomial trials.

(a) In two sets of binomial trials T and t, the probabilities that a trial has a successful outcome are P and p, respectively, with corresponding probabilities of failure of $Q = 1 - P$ and $q = 1 - p$. One "game" consists of a trial T, followed, if T is successful, by a trial t and then a further trial T. The two trials continue to alternate until one of the T-trials fails, at which point the game ends. The score S for the game is the total number of successes in the t-trials. Find the PGF for S and use it to show that

$$E[S] = \frac{Pp}{Q}, \qquad V[S] = \frac{Pp(1 - Pq)}{Q^2}.$$

(b) Two normal unbiased six-faced dice A and B are rolled alternately starting with A; if A shows a 6 the experiment ends. If B shows an odd number no points are scored, one point is scored for a 2 or a 4, and two points are awarded for a 6. Find the average and standard deviation of the score for the experiment and show that the latter is the greater.

(a) This is a situation in which the score for the game is a variable-length sum, the length N being determined by the outcome of the T-trials. The probability that $N = n$ is given by $h_n = P^n Q$, since n T-trials must succeed and then be followed by a failing T-trial. Thus the PGF for the length of each "game" is given by

$$\chi_N(t) \equiv \sum_{n=0}^{\infty} h_n t^n = \sum_{n=0}^{\infty} P^n Q t^n = \frac{Q}{1 - Pt}.$$

For each permitted Bernoulli t-trial, $X_i = 1$ with probability p and $X_i = 0$ with probability q; its PGF is thus $\Phi_X(t) = q + pt$. The score for the game is $S = \sum_{i=1}^{N} X_i$ and its PGF is given by the compound function

$$\begin{aligned}\Xi_S(t) &= \chi_N(\Phi_X(t)) \\ &= \frac{Q}{1 - P(q + pt)},\end{aligned}$$

in which the PGF for a single t-trial forms the argument of the PGF for the length of each "game".

It follows that the mean of S is found from

$$\Xi_S'(t) = \frac{QPp}{(1 - Pq - Ppt)^2} \quad \Rightarrow \quad E[S] = \Xi_S'(1) = \frac{QPp}{(1 - P)^2} = \frac{Pp}{Q}.$$

To calculate the variance of S we need to find $\Xi_S''(1)$. This second derivative is

$$\Xi_S''(t) = \frac{2QP^2p^2}{(1 - Pq - Ppt)^3} \quad \Rightarrow \quad \Xi_S''(1) = \frac{2P^2p^2}{Q^2}.$$

The variance is therefore

$$\begin{aligned}V[S] &= \Xi_S''(1) + \Xi_S'(1) - [\Xi_S'(1)]^2 \\ &= \frac{2P^2p^2}{Q^2} + \frac{Pp}{Q} - \frac{P^2p^2}{Q^2} \\ &= \frac{Pp(Pp + Q)}{Q^2} = \frac{Pp(P - Pq + Q)}{Q^2} = \frac{Pp(1 - Pq)}{Q^2}.\end{aligned}$$

(b) For die A: $P = \frac{5}{6}$ and $Q = \frac{1}{6}$ giving $\chi_N(t) = 1/(6-5t)$.

For die B: $\Pr(X = 0) = \frac{3}{6}$, $\Pr(X = 1) = \frac{2}{6}$ and $\Pr(X = 2) = \frac{1}{6}$ giving $\Phi_X(t) = (3 + 2t + t^2)/6$.

The PGF for the game score S is thus

$$\Xi_S(t) = \frac{1}{6 - \frac{5}{6}(3 + 2t + t^2)} = \frac{6}{21 - 10t - 5t^2}.$$

We need to evaluate the first two derivatives of $\Xi_S(t)$ at $t = 1$, as follows:

$$\begin{aligned}
\Xi_S'(t) &= \frac{-6(-10-10t)}{(21-10t-5t^2)^2} = \frac{60+60t}{(21-10t-5t^2)^2} \\
\Rightarrow \quad E[S] = \Xi_S'(1) &= \frac{120}{6^2} = \frac{10}{3} = 3.33, \\
\Xi_S''(t) &= \frac{60}{(21-10t-5t^2)^2} - \frac{2(60+60t)(-10-10t)}{(21-10t-5t^2)^3} \\
\Rightarrow \quad \Xi_S''(1) &= \frac{60}{36} - \frac{2(120)(-20)}{(6)^3} = \frac{215}{9}.
\end{aligned}$$

Substituting the calculated values gives $V[S]$ as

$$V[S] = \frac{215}{9} + \frac{10}{3} - \left(\frac{10}{3}\right)^2 = \frac{145}{9},$$

from which it follows that

$$\sigma_S = \sqrt{V[S]} = 4.01, \text{ i.e. greater than the mean.}$$

16.19 A point P is chosen at random on the circle $x^2 + y^2 = 1$. The random variable X denotes the distance of P from $(1, 0)$. Find the mean and variance of X and the probability that X is greater than its mean.

With O as the center of the unit circle and Q as the point $(1, 0)$, let OP make an angle θ with the x-axis OQ. The random variable X then has the value $2\sin(\theta/2)$ with θ uniformly distributed on $(0, 2\pi)$, i.e.

$$f(x)\,dx = \frac{1}{2\pi}\,d\theta.$$

The mean of X is given straightforwardly by

$$\langle X\rangle = \int_0^2 Xf(x)\,dx = \int_0^{2\pi} 2\sin\left(\frac{\theta}{2}\right)\frac{1}{2\pi}\,d\theta = \frac{1}{\pi}\left[-2\cos\frac{\theta}{2}\right]_0^{2\pi} = \frac{4}{\pi}.$$

For the variance we have

$$\sigma_X^2 = \langle X^2\rangle - \langle X\rangle^2 = \int_0^{2\pi} 4\sin^2\left(\frac{\theta}{2}\right)\frac{1}{2\pi}\,d\theta - \frac{16}{\pi^2} = \frac{4}{2\pi}\,\frac{1}{2}\,2\pi - \frac{16}{\pi^2} = 2 - \frac{16}{\pi^2}.$$

When $X = \langle X \rangle = 4/\pi$, the angle $\theta = 2\sin^{-1}(2/\pi)$ and so

$$\Pr(X > \langle X \rangle) = \frac{2\pi - 4\sin^{-1}\dfrac{2}{\pi}}{2\pi} = 0.561.$$

16.21 The number of errors needing correction on each page of a set of proofs follows a Poisson distribution of mean μ. The cost of the first correction on any page is α and that of each subsequent correction on the same page is β. Prove that the average cost of correcting a page is

$$\alpha + \beta(\mu - 1) - (\alpha - \beta)e^{-\mu}.$$

Since the number of errors on a page is Poisson distributed, the probability of n errors on any particular page is

$$\Pr(n \text{ errors}) = p_n = e^{-\mu}\frac{\mu^n}{n!}.$$

The average cost per page, found by averaging the corresponding cost over all values of n, is

$$\begin{aligned} c &= 0\,p_0 + \alpha p_1 + \sum_{n=2}^{\infty}[\alpha + (n-1)\beta]p_n \\ &= \alpha\mu e^{-\mu} + (\alpha - \beta)\sum_{n=2}^{\infty} p_n + \beta\sum_{n=2}^{\infty} np_n. \end{aligned}$$

Now, $\sum_{n=0}^{\infty} p_n = 1$ and, for a Poisson distribution, $\sum_{n=0}^{\infty} np_n = \mu$. These can be used to evaluate the above, once the $n = 0$ and $n = 1$ terms have been removed. Thus

$$\begin{aligned} c &= \alpha\mu e^{-\mu} + (\alpha - \beta)(1 - e^{-\mu} - \mu e^{-\mu}) + \beta(\mu - 0 - \mu e^{-\mu}) \\ &= \alpha + \beta(\mu - 1) + e^{-\mu}(\alpha\mu - \alpha + \beta - \mu\alpha + \mu\beta - \mu\beta) \\ &= \alpha + \beta(\mu - 1) + e^{-\mu}(\beta - \alpha), \end{aligned}$$

as given in the question.

16.23 The probability distribution for the number of eggs in a clutch is Po(λ), and the probability that each egg will hatch is p (independently of the size of the clutch). Show by direct calculation that the probability distribution for the number of chicks that hatch is Po(λp).

Clearly, to determine the probability that a clutch produces k chicks, we must consider clutches of size n, for all $n \geq k$, and for each such clutch find the probability that exactly k of the n chicks do hatch. We then average over all n, weighting the results according to the distribution of n.

The probability that k chicks hatch from a clutch of size n is ${}^nC_k p^k q^{n-k}$, where $q = 1 - p$. The probability that the clutch is of size n is $e^{-\lambda}\lambda^n/n!$. Consequently, the overall probability of k chicks hatching from a clutch is

$$\begin{aligned}
\Pr(k \text{ chicks}) &= \sum_{n=k}^{\infty} e^{-\lambda} \frac{\lambda^n}{n!} \, {}^nC_k \, p^k \, q^{n-k} \\
&= e^{-\lambda} p^k \lambda^k \sum_{n=k}^{\infty} \frac{(\lambda q)^{n-k}}{n!} \, \frac{n!}{k!\,(n-k)!}, \qquad \text{set } n-k=m, \\
&= e^{-\lambda} \frac{(\lambda p)^k}{k!} \sum_{m=0}^{\infty} \frac{(\lambda q)^m}{m!} \\
&= e^{-\lambda} \frac{(\lambda p)^k}{k!} \, e^{\lambda q} \\
&= \frac{e^{-\lambda p} (\lambda p)^k}{k!},
\end{aligned}$$

since $q = 1 - p$. Thus $\Pr(k$ chicks) is distributed as a Poisson distribution with parameter $\mu = \lambda p$.

16.25 Under EU legislation on harmonization, all kippers are to weigh 0.2000 kg and vendors who sell underweight kippers must be fined by their government. The weight of a kipper is normally distributed with a mean of 0.2000 kg and a standard deviation of 0.0100 kg. They are packed in cartons of 100 and large quantities of them are sold.

Every day a carton is to be selected at random from each vendor and tested according to one of the following schemes, which have been approved for the purpose.

(a) The entire carton is weighed and the vendor is fined 2500 euros if the average weight of a kipper is less than 0.1975 kg.
(b) Twenty five kippers are selected at random from the carton; the vendor is fined 100 euros if the average weight of a kipper is less than 0.1980 kg.
(c) Kippers are removed one at a time, at random, until one has been found that weighs *more* than 0.2000 kg; the vendor is fined $4n(n-1)$ euros, where n is the number of kippers removed.

Which scheme should the Chancellor of the Exchequer be urging his government to adopt?

For these calculations we measure weights in grammes.

(a) For this scheme we have a normal distribution with mean $\mu = 200$ and s.d. $\sigma = 10$. The s.d. for a carton is $\sqrt{100}\,\sigma = 100$ and the mean weight is 20 000. There is a penalty if the weight of a carton is less than 19 750. This critical value represents a standard variable of

$$Z = \frac{19\,750 - 20\,000}{100} = -2.5.$$

The probability that $Z < -2.5 = 1 - \Phi(2.5) = 1 - 0.9938 = 0.0062$. Thus the average fine per carton tested on this scheme is $0.0062 \times 2500 = 15.5$ euros.

(b) For this scheme the general parameters are the same but the mean weight of the sample measured is 5000 and its s.d. is $\sqrt{25}\,(10) = 50$. The Z-value at which a fine is imposed is

$$Z = \frac{(198 \times 25) - 5000}{50} = -1.$$

The probability that $Z < -1.0 = 1 - \Phi(1.0) = 1 - 0.8413 = 0.1587$. Thus the average fine per carton tested on this scheme is $0.1587 \times 100 = 15.9$ euros.

(c) This scheme is a series of Bernoulli trials in which the probability of success is $\frac{1}{2}$ (since half of all kippers weigh more than 200 and the distribution is normal). The probability that it will take n kippers to find one that passes the test is $q^{n-1}p = (\frac{1}{2})^n$. The expected fine is therefore

$$f = \sum_{n=2}^{\infty} 4n(n-1)\left(\frac{1}{2}\right)^n = 4\,\frac{2(\frac{1}{4})}{(\frac{1}{2})^3} = 16 \text{ euros.}$$

The expression for the sum was found by twice differentiating the sum of the geometric series $\sum r^n$ with respect to r, as follows:

$$\begin{aligned}
\sum_{n=0}^{\infty} r^n = \frac{1}{1-r} \quad &\Rightarrow \quad \sum_{n=1}^{\infty} nr^{n-1} = \frac{1}{(1-r)^2} \\
&\Rightarrow \quad \sum_{n=2}^{\infty} n(n-1)r^{n-2} = \frac{2}{(1-r)^3} \\
&\Rightarrow \quad \sum_{n=2}^{\infty} n(n-1)r^{n} = \frac{2r^2}{(1-r)^3}.
\end{aligned}$$

There is, in fact, little to choose between the schemes on monetary grounds; no doubt political considerations, such as the current unemployment rate, will decide!

16.27 A practical-class demonstrator sends his 12 students to the storeroom to collect apparatus for an experiment, but forgets to tell each which type of component to bring. There are three types, A, B and C, held in the stores (in large numbers) in the proportions 20%, 30% and 50%, respectively, and each student picks a component at random. In order to set up one experiment, one unit each of A and B and two units of C are needed. Let $\Pr(N)$ be the probability that at least N experiments can be set up.

(a) Evaluate $\Pr(3)$.

(b) Find an expression for $\Pr(N)$ in terms of k_1 and k_2, the numbers of components of types A and B, respectively, selected by the students. Show that $\Pr(2)$ can be written in the form

$$\Pr(2) = (0.5)^{12} \sum_{i=2}^{6} {}^{12}C_i\,(0.4)^i \sum_{j=2}^{8-i} {}^{12-i}C_j\,(0.6)^j.$$

(c) By considering the conditions under which no experiments can be set up, show that $\Pr(1) = 0.9145$.

(a) To make three experiments possible the 12 components picked must be three each of A and B and six of C. The probability of this is given by the multinomial distribution as

$$\Pr(3) = \frac{(12)!}{3!\,3!\,6!}(0.2)^3(0.3)^3(0.5)^6 = 0.06237.$$

(b) Let the numbers of A, B and C selected be k_1, k_2 and k_3, respectively, and consider when *at least* N experiments can be set up. We have the obvious inequalities $k_1 \geq N$, $k_2 \geq N$ and $k_3 \geq 2N$. In addition $k_3 = 12 - k_1 - k_2$, implying that $k_2 \leq 12 - 2N - k_1$. Further, k_1 cannot be greater than $12 - 3N$ if at least N experiments are to be set up, as each requires three other components that are not of type A. These inequalities set the limits on the acceptable values of k_1 and k_2 (k_3 is not a third independent variable). Thus $\Pr(N)$ is given by

$$\sum_{k_1 \geq N}^{12-3N} \sum_{k_2 \geq N}^{12-2N-k_1} \frac{(12)!}{k_1!\,k_2!\,(12-k_1-k_2)!}(0.2)^{k_1}(0.3)^{k_2}(0.5)^{12-k_1-k_2}.$$

The answer to part (a) is a particular case of this with $N = 3$, when each summation reduces to a single term.

For $N = 2$ the expression becomes

$$\begin{aligned}
\Pr(2) &= \sum_{k_1 \geq 2}^{6} \sum_{k_2 \geq 2}^{8-k_1} \frac{(12)!}{k_1!\,k_2!\,(12-k_1-k_2)!}(0.2)^{k_1}(0.3)^{k_2}(0.5)^{12-k_1-k_2} \\
&= (0.5)^{12} \sum_{i=2}^{6} \sum_{j=2}^{8-i} \frac{(12)!\,(0.2/0.5)^i}{i!\,(12-i)!} \frac{(12-i)!\,(0.3/0.5)^j}{j!\,(12-i-j)!} \\
&= (0.5)^{12} \sum_{i=2}^{6} {}^{12}C_i\,(0.4)^i \sum_{j=2}^{8-i} {}^{12-i}C_j\,(0.6)^j.
\end{aligned}$$

(c) No experiment can be set up if any one of the following four events occurs: $A_1 = (k_1 = 0)$, $A_2 = (k_2 = 0)$, $A_3 = (k_3 = 0)$ and $A_4 = (k_3 = 1)$. The probability for the union of these four events is given by

$$\Pr(A_1 \cup A_2 \cup A_3 \cup A_4) = \sum_{i=1}^{4} \Pr(A_i) - \sum_{i,j} \Pr(A_i \cap A_j) + \cdots.$$

The probabilities $\Pr(A_i)$ are straightforward to calculate as follows:

$$\begin{aligned}
\Pr(A_1) &= (1-0.2)^{12}, & \Pr(A_2) &= (1-0.3)^{12}, \\
\Pr(A_3) &= (1-0.5)^{12}, & \Pr(A_4) &= {}^{12}C_1(1-0.5)^{12}(0.5).
\end{aligned}$$

The calculation of the probability for the intersection of two events is typified by

$$\begin{aligned}
\Pr(A_1 \cap A_2) &= [1-(0.2+0.3)]^{12} \\
\text{and } \Pr(A_1 \cap A_4) &= {}^{12}C_1[1-(0.2+0.5)]^{11}(0.5)^1.
\end{aligned}$$

A few trial evaluations show that these are of order 10^{-4} and can be ignored by comparison with the larger terms in the first sum, which are (after rounding)

$$\sum_{i=1}^{4} \Pr(A_i) = (0.8)^{12} + (0.7)^{12} + (0.5)^{12} + 12(0.5)^{11}(0.5)$$
$$= 0.0687 + 0.0138 + 0.0002 + 0.0029 = 0.0856.$$

Since the probability of no experiments being possible is 0.0856, it follows that $\Pr(1) = 0.9144$.

16.29 The continuous random variables X and Y have a joint PDF proportional to $xy(x-y)^2$ with $0 \le x \le 1$ and $0 \le y \le 1$. Find the marginal distributions for X and Y and show that they are negatively correlated with correlation coefficient $-\frac{2}{3}$.

This PDF is clearly symmetric between x and y. We start by finding its normalization constant c:

$$\int_0^1 \int_0^1 c(x^3y - 2x^2y^2 + xy^3)\,dx\,dy = c\left(\frac{1}{4}\frac{1}{2} - 2\frac{1}{3}\frac{1}{3} + \frac{1}{2}\frac{1}{4}\right) = \frac{c}{36}.$$

Thus, we must have that $c = 36$.

The marginal distribution for x is given by

$$f(x) = 36\int_0^1 (x^3y - 2x^2y^2 + xy^3)\,dy$$
$$= 36\left(\tfrac{1}{2}x^3 - \tfrac{2}{3}x^2 + \tfrac{1}{4}x\right)$$
$$= 18x^3 - 24x^2 + 9x,$$

and the mean of x by

$$\mu_X = \bar{x} = \int_0^1 (18x^4 - 24x^3 + 9x^2)\,dx = \frac{18}{5} - \frac{24}{4} + \frac{9}{3} = \frac{3}{5}.$$

By symmetry, the marginal distribution and the mean for y are $18y^3 - 24y^2 + 9y$ and $\frac{3}{5}$, respectively.

To calculate the correlation coefficient we also need the variances of x and y and their covariance. The variances, obviously equal, are given by

$$\sigma_X^2 = \int_0^1 x^2(18x^3 - 24x^2 + 9x)\,dx - \left(\tfrac{3}{5}\right)^2$$
$$= \frac{18}{6} - \frac{24}{5} + \frac{9}{4} - \frac{9}{25}$$
$$= \frac{900 - 1440 + 675 - 108}{300} = \frac{9}{100}.$$

The standard deviations σ_X and σ_Y are therefore both equal to $3/10$.

The covariance is calculated next; it is given by

$$\begin{aligned}
\mathrm{Cov}\,[X, Y] &= \langle XY \rangle - \mu_X \mu_Y \\
&= 36 \int_0^1 \int_0^1 (x^4 y^2 - 2x^3 y^3 + x^2 y^4)\, dx\, dy - \frac{3}{5}\frac{3}{5} \\
&= \frac{36}{5 \times 3} - \frac{72}{4 \times 4} + \frac{36}{3 \times 5} - \frac{9}{25} \\
&= \frac{12}{5} - \frac{9}{2} + \frac{12}{5} - \frac{9}{25} \\
&= \frac{120 - 225 + 120 - 18}{50} = -\frac{3}{50}.
\end{aligned}$$

Finally,

$$\mathrm{Corr}\,[X, Y] = \frac{\mathrm{Cov}\,[X, Y]}{\sigma_X\, \sigma_Y} = \frac{-\frac{3}{50}}{\frac{3}{10}\frac{3}{10}} = -\frac{2}{3}.$$

16.31 Two continuous random variables X and Y have a joint probability distribution

$$f(x, y) = A(x^2 + y^2),$$

where A is a constant and $0 \le x \le a$, $0 \le y \le a$. Show that X and Y are negatively correlated with correlation coefficient $-15/73$. By sketching a rough contour map of $f(x, y)$ and marking off the regions of positive and negative correlation, convince yourself that this (perhaps counter-intuitive) result is plausible.

The calculations of the various parameters of the distribution are straightforward (see Problem 16.29). The parameter A is determined by the normalization condition:

$$1 = \int_0^a \int_0^a A(x^2 + y^2)\, dx\, dy = A\left(\frac{a^4}{3} + \frac{a^4}{3}\right) \quad \Rightarrow \quad A = \frac{3}{2a^4}.$$

The two expectation values required are given by

$$\begin{aligned}
E[X] &= \int_0^a \int_0^a Ax(x^2 + y^2)\, dx\, dy \\
&= \frac{3}{2a^4}\left(\frac{a^5}{4 \times 1} + \frac{a^5}{2 \times 3}\right) = \frac{5a}{8}, \qquad (E[Y] = E[X]), \\
E[X^2] &= \int_0^a \int_0^a Ax^2(x^2 + y^2)\, dx\, dy \\
&= \frac{3}{2a^4}\left(\frac{a^6}{5 \times 1} + \frac{a^6}{3 \times 3}\right) = \frac{7a^2}{15}.
\end{aligned}$$

Hence the variance, calculated from the general result $V[X] = E[X^2] - (E[X])^2$, is

$$V[X] = \frac{7a^2}{15} - \left(\frac{5a}{8}\right)^2 = \frac{73}{960}a^2,$$

and the standard deviations are given by

$$\sigma_X = \sigma_Y = \sqrt{\frac{73}{960}}\, a.$$

To obtain the correlation coefficient we need also to calculate the following:

$$\begin{aligned} E[XY] &= \int_0^a \int_0^a Axy(x^2+y^2)\, dx\, dy \\ &= \frac{3}{2a^4}\left(\frac{a^6}{4\times 2} + \frac{a^6}{2\times 4}\right) = \frac{3a^2}{8}. \end{aligned}$$

Then the covariance, given by $\text{Cov}\,[X, Y] = E[XY] - E[X]E[Y]$, is evaluated as

$$\text{Cov}\,[X, Y] = \frac{3}{8}a^2 - \frac{5a}{8}\frac{5a}{8} = -\frac{a^2}{64}.$$

Combining this last result with the standard deviations calculated above, we then obtain

$$\text{Corr}\,[X, Y] = \frac{-(a^2/64)}{\sqrt{\frac{73}{960}}\, a\sqrt{\frac{73}{960}}\, a} = -\frac{15}{73}.$$

As the means of both X and Y are $\frac{5}{8}a = 0.62a$, the areas of the square of side a for which $X - \mu_X$ and $Y - \mu_Y$ have the same sign (i.e. regions of positive correlation) are about $(0.62)^2 \approx 39\%$ and $(0.38)^2 \approx 14\%$ of the total area of the square. The regions of negative correlation occupy some 47% of the square.

However, $f(x, y) = A(x^2 + y^2)$ favors the regions where one or both of x and y are large and close to unity. Broadly speaking, this gives little weight to the region in which both X and Y are less than their means, and so, although it is the largest region in area, it contributes relatively little to the overall correlation. The two (equal area) regions of negative correlation together outweigh the smaller high probability region of positive correlation in the top right-hand corner of the square; the overall result is a net negative correlation coefficient.

17 Statistics

17.1 A group of students uses a pendulum experiment to measure g, the acceleration of free fall, and obtains the following values (in m s^{-2}): 9.80, 9.84, 9.72, 9.74, 9.87, 9.77, 9.28, 9.86, 9.81, 9.79, 9.82. What would you give as the best value and standard error for g as measured by the group?

We first note that the reading of $9.28\ \text{m s}^{-2}$ is so far from the others that it is almost certainly in error and should not be used in the calculation. The mean of the ten remaining values is 9.802 and the standard deviation of the sample about its mean is 0.04643. After including Bessel's correction factor, the estimate of the population s.d. is $\sigma = 0.0489$, leading to a s.d. in the measured value of the mean of $0.0489/\sqrt{10} = 0.0155$. We therefore give the best value and standard error for g as $9.80 \pm 0.02\ \text{m s}^{-2}$.

17.3 The following are the values obtained by a class of 14 students when measuring a physical quantity x: 53.8, 53.1, 56.9, 54.7, 58.2, 54.1, 56.4, 54.8, 57.3, 51.0, 55.1, 55.0, 54.2, 56.6.

(a) Display these results as a histogram and state what you would give as the best value for x.
(b) Without calculation, estimate how much reliance could be placed upon your answer to (a).
(c) Data books give the value of x as 53.6 with negligible error. Are the data obtained by the students in conflict with this?

(a) The histogram in Figure 17.1 shows no reading that is an obvious mistake and there is no reason to suppose other than a Gaussian distribution. The best value for x is the arithmetic mean of the 14 values given, i.e. 55.1.

(b) We note that 11 values, i.e. approximately two-thirds of the 14 readings, lie within ± 2 bins of the mean. This estimates the s.d. for the population as 2.0 and gives a standard error in the mean of $\approx 2.0/\sqrt{14} \approx 0.6$.

(c) Within the accuracy we are likely to achieve by estimating σ for the sample by eye, the value of Student's t is $(55.1 - 53.6)/0.6$, i.e. about 2.5. With 14 readings there are 13 degrees of freedom. From standard tables for the Student's t-test, $C_{13}(2.5) \approx 0.985$. It is therefore likely at the $2 \times 0.015 = 3\%$ significance level that the data are in conflict with the accepted value.

[Numerical analysis of the data, rather than a visual estimate, gives the lower value 0.51 for the standard error in the mean and implies that there is a conflict between the data and the accepted value at the 1.0% significance level.]

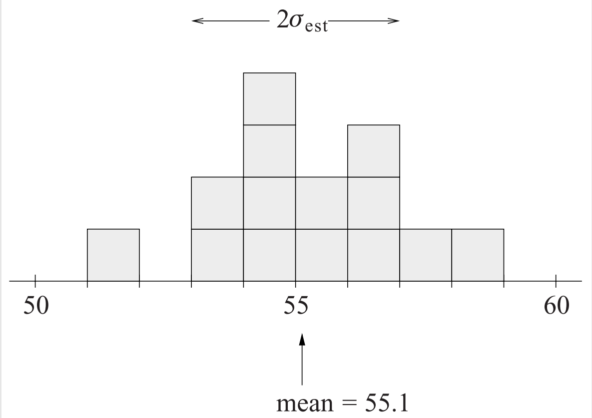

Figure 17.1 Histogram of the data in Problem 17.3.

17.5 A population contains individuals of k types in equal proportions. A quantity X has mean μ_i amongst individuals of type i and variance σ^2, which has the same value for all types. In order to estimate the mean of X over the whole population, two schemes are considered; each involves a total sample size of nk. In the first the sample is drawn randomly from the whole population, whilst in the second (*stratified sampling*) n individuals are randomly selected from each of the k types.

Show that in both cases the estimate has expectation

$$\mu = \frac{1}{k}\sum_{i=1}^{k}\mu_i,$$

but that the variance of the first scheme exceeds that of the second by an amount

$$\frac{1}{k^2 n}\sum_{i=1}^{k}(\mu_i-\mu)^2.$$

(i) For the first scheme the estimator $\hat{\mu}$ has expectation

$$\langle\hat{\mu}\rangle = \frac{1}{nk}\sum_{j=1}^{nk}\langle x_j\rangle,$$

where

$$\langle x_j\rangle = \frac{1}{k}\sum_{i=1}^{k}\mu_i \text{ for all } j,$$

since the k types are in equal proportions in the population. Thus,

$$\langle\hat{\mu}\rangle = \frac{1}{nk}\sum_{j=1}^{nk}\frac{1}{k}\sum_{i=1}^{k}\mu_i = \frac{1}{k}\sum_{i=1}^{k}\mu_i = \mu.$$

The variance of $\hat{\mu}$ is given by

$$\begin{aligned} V[\hat{\mu}] &= \frac{1}{n^2k^2}\,nk\,V[x] \\ &= \frac{1}{nk}\left(\langle x^2\rangle - \mu^2\right) \\ &= \frac{1}{nk}\left(\frac{1}{k}\sum_{i=1}^{k}\langle x_i^2\rangle - \mu^2\right), \end{aligned}$$

again since the k types are in equal proportions in the population.

Now we use the relationship $\sigma^2 = \langle x_i^2\rangle - \mu_i^2$ to replace $\langle x_i^2\rangle$ for each type, noting that σ^2 has the same value in each case. The expression for the variance becomes

$$\begin{aligned} V[\,\hat{\mu}\,] &= \frac{1}{nk}\left[\frac{1}{k}\sum_{i=1}^{k}\left(\mu_i^2+\sigma^2\right) - \mu^2\right] \\ &= \frac{\sigma^2-\mu^2}{nk} + \frac{1}{nk^2}\sum_{i=1}^{k}(\mu_i-\mu+\mu)^2 \\ &= \frac{\sigma^2-\mu^2}{nk} + \frac{1}{nk^2}\sum_{i=1}^{k}\left[\,(\mu_i-\mu)^2 + 2\mu(\mu_i-\mu) + \mu^2\,\right] \\ &= \frac{\sigma^2-\mu^2}{nk} + \frac{1}{nk^2}\sum_{i=1}^{k}(\mu_i-\mu)^2 + 0 + \frac{k\mu^2}{nk^2} \\ &= \frac{\sigma^2}{nk} + \frac{1}{nk^2}\sum_{i=1}^{k}(\mu_i-\mu)^2. \end{aligned}$$

(ii) For the second scheme the calculations are more straightforward. The expectation value of the estimator $\hat{\mu} = (nk)^{-1}\sum_{i=1}^{k}\langle x_i\rangle$ is

$$\langle\hat{\mu}\rangle = \frac{1}{nk}\sum_{i=1}^{k}n\mu_i = \frac{1}{k}\sum_{i=1}^{k}\mu_i = \mu,$$

whilst the variance is given by

$$V[\,\hat{\mu}\,] = \frac{1}{n^2k^2}\sum_{i=1}^{k}V[\,\langle x_i\rangle\,] = \frac{1}{n^2k^2}\sum_{i=1}^{k}n\sigma_i^2 = \frac{1}{k^2}\frac{k\sigma^2}{n} = \frac{\sigma^2}{kn},$$

since $\sigma_i^2 = \sigma^2$ for all i.

Comparing the results from (i) and (ii), we see that the variance of the estimator in the first scheme is larger by

$$\frac{1}{nk^2}\sum_{i=1}^{k}(\mu_i-\mu)^2.$$

17.7 According to a particular theory, two dimensionless quantities X and Y have equal values. Nine measurements of X gave values of 22, 11, 19, 19, 14, 27, 8, 24 and 18, whilst seven measured values of Y were 11, 14, 17, 14, 19, 16 and 14. Assuming that the measurements of both quantities are Gaussian distributed with a common variance, are they consistent with the theory? An alternative theory predicts that $Y^2=\pi^2X$; are the data consistent with this proposal?

On the hypothesis that $X=Y$ and both quantities have Gaussian distributions with a common variance, we need to calculate the value of t given by

$$t=\frac{\bar{w}-\omega}{\hat{\sigma}}\left(\frac{N_1N_2}{N_1+N_2}\right)^{1/2},$$

where $\bar{w}=\bar{x}_1-\bar{x}_2$, $\omega=\mu_1-\mu_2=0$ and

$$\hat{\sigma}=\left[\frac{N_1s_1^2+N_2s_2^2}{N_1+N_2-2}\right]^{1/2}.$$

The nine measurements of X have a mean of 18.0 and a value for s^2 of 33.33. The corresponding values for the seven measurements of Y are 15.0 and 5.71. Substituting these values gives

$$\hat{\sigma}=\left[\frac{9\times 33.33+7\times 5.71}{9+7-2}\right]^{1/2}=4.93,$$

$$t=\frac{18.0-15.0-0}{4.93}\left(\frac{9\times 7}{9+7}\right)^{1/2}=1.21.$$

This variable follows a Student's t-distribution for $9+7-2=14$ degrees of freedom. Interpolation in standard tables gives $C_{14}(1.21)\approx 0.874$, showing that a larger value of t could be expected in about $2\times(1-0.874)=25\%$ of cases. Thus no inconsistency between the data and the first theory has been established.

For the second theory we are testing Y^2 against π^2X; the former will not be Gaussian distributed and the two distributions will not have a common variance. Thus the best we can do is to compare the difference between the two expressions, evaluated with the mean values of X and Y, against the estimated error in that difference.

The difference in the expressions is $(15.0)^2-18.0\pi^2=47.3$. The error in the difference between the functions of Y and X is given approximately by

$$\begin{aligned}V(Y^2-\pi^2X)&=(2Y)^2\,V[\,Y\,]+(\pi^2)^2\,V[\,X\,]\\&=(30.0)^2\frac{5.71}{7-1}+(\pi^2)^2\frac{33.33}{9-1}\\&=1262\quad\Rightarrow\quad\sigma\approx 35.5.\end{aligned}$$

The difference is thus about $47.3/35.5 = 1.33$ standard deviations away from the theoretical value of 0. The distribution will not be truly Gaussian but, if it were, this figure would have a probability of being exceeded in magnitude some $2 \times (1 - 0.908) = 18\%$ of the time. Again no inconsistency between the data and theory has been established.

17.9 During an investigation into possible links between mathematics and classical music, pupils at a school were asked whether they had preferences (a) between mathematics and English, and (b) between classical and pop music. The results are given below.

	Classical	None	Pop
Mathematics	23	13	14
None	17	17	36
English	30	10	40

Determine whether there is any evidence for

(a) a link between academic and musical tastes, and
(b) a claim that pupils either had preferences in both areas or had no preference.

You will need to consider the appropriate value for the number of degrees of freedom to use when applying the χ^2 test.

We first note that there were 200 pupils taking part in the survey. Denoting no academic preference between mathematics and English by NA and no musical preference by NM, we draw up an enhanced table of the actual numbers m_{XY} of preferences for the various combinations that also shows the overall probabilities p_X and p_Y of the three choices in each selection.

	C	NM	P	Total	p_X
M	23	13	14	50	0.25
NA	17	17	36	70	0.35
E	30	10	40	80	0.40
Total	70	40	90	200	
p_Y	0.35	0.20	0.45		

(a) If we now assume the (null) hypothesis that there are no correlations in the data and that any apparent correlations are the result of statistical fluctuations, then the expected number of pupils opting for the combination X and Y is $n_{XY} = 200 \times p_X \times p_Y$. A table of n_{XY} is as follows:

	C	NM	P	Total
M	17.5	10	22.5	50
NA	24.5	14	31.5	70
E	28	16	36	80
Total	70	40	90	200

Taking the standard deviation as the square root of the expected number of votes for each particular combination, the value of χ^2 is given by

$$\chi^2 = \sum_{\text{all XY combinations}} \left(\frac{n_i - m_i}{\sqrt{n_i}} \right)^2 = 12.3.$$

For an $n \times n$ correlation table (here $n = 3$), the $(n-1) \times (n-1)$ block of entries in the upper left can be filled in arbitrarily. But, as the totals for each row and column are predetermined, the remaining $2n - 1$ entries are not arbitrary. Thus the number of degrees of freedom (d.o.f.) for such a table is $(n-1)^2$, here 4 d.o.f. From tables, a χ^2 of 12.3 for 4 d.o.f. makes the assumed hypothesis less than 2% likely, and so it is almost certain that a correlation between academic and musical tastes does exist.

(b) To investigate a claim that pupils either had preferences in both areas or had no preference, we must combine expressed preferences for classical or pop into one set labeled PM meaning "expressed a musical preference"; similarly for academic subjects. The correlation table is now a 2×2 one and will have only one degree of freedom. The actual data table is

	PM	NM	Total	p_X
PA	107	23	130	0.65
NA	53	17	70	0.35
Total	160	40	200	
p_Y	0.80	0.20		

and the expected ($n_{XY} = 200 p_X p_Y$) one is

	PM	NM	Total
PA	104	26	130
NA	56	14	70
Total	160	40	200

The value of χ^2 is

$$\chi^2 = \frac{(-3)^2}{104} + \frac{(3)^2}{26} + \frac{(3)^2}{56} + \frac{(-3)^2}{14} = 1.24.$$

This is close to the expected value (1) of χ^2 for 1 d.o.f. and is neither too big nor too small. Thus there is no evidence for the claim (or for any tampering with the data!).

17.11 A particle detector consisting of a shielded scintillator is being tested by placing it near a particle source whose intensity can be controlled by the use of absorbers. It might register counts even in the absence of particles from the source because of the cosmic ray background.

The number of counts n registered in a fixed time interval as a function of the source strength s is given as:

source strength s:	0	1	2	3	4	5	6
counts n:	6	11	20	42	44	62	61

At any given source strength, the number of counts is expected to be Poisson distributed with mean

$$n = a + bs,$$

where a and b are constants. Analyze the data for a fit to this relationship and obtain the best values for a and b together with their standard errors.

(a) How well is the cosmic ray background determined?
(b) What is the value of the correlation coefficient between a and b? Is this consistent with what would happen if the cosmic ray background were imagined to be negligible?
(c) Do the data fit the expected relationship well? Is there any evidence that the reported data "are too good a fit"?

Because in this problem the independent variable s takes only consecutive integer values, we will use it as a label i and denote the number of counts corresponding to $s = i$ by n_i. As the data are expected to be Poisson distributed, the best estimate of the variance of each reading is equal to the best estimate of the reading itself, namely the actual measured value. Thus each reading n_i has an error of $\sqrt{n_i}$, and the covariance matrix N takes the form $N = \text{diag}(n_0, n_1, \dots, n_6)$, i.e. it is diagonal, but not a multiple of the unit matrix.

The expression for χ^2 is

$$\chi^2(a, b) = \sum_{i=0}^{6} \left(\frac{n_i - a - bi}{\sqrt{n_i}} \right)^2 \qquad (*).$$

Minimization with respect to a and b gives the simultaneous equations

$$0 = \frac{\partial \chi^2}{\partial a} = -2 \sum_{i=0}^{6} \frac{n_i - a - bi}{n_i},$$

$$0 = \frac{\partial \chi^2}{\partial b} = -2 \sum_{i=0}^{6} \frac{i(n_i - a - bi)}{n_i}.$$

As is shown more generally in textbooks on numerical computing (e.g. William H. Press *et al.*, *Numerical Recipes in C*, 2nd edn (Cambridge: Cambridge University Press, 1996), Sect. 15.2), these equations are most conveniently solved by defining the quantities

$$S \equiv \sum_{i=0}^{6} \frac{1}{n_i}, \quad S_x \equiv \sum_{i=0}^{6} \frac{i}{n_i}, \quad S_y \equiv \sum_{i=0}^{6} \frac{n_i}{n_i},$$

$$S_{xx} \equiv \sum_{i=0}^{6} \frac{i^2}{n_i}, \quad S_{xy} \equiv \sum_{i=0}^{6} \frac{in_i}{n_i}, \quad \Delta \equiv SS_{xx} - (S_x)^2.$$

With these definitions (which correspond to the quantities calculated and accessibly stored in most calculators programmed to perform least-squares fitting), the solutions for the best estimators of a and b are

$$\hat{a} = \frac{S_{xx}S_y - S_x S_{xy}}{\Delta},$$
$$\hat{b} = \frac{S_{xy}S - S_x S_y}{\Delta},$$

with variances and covariance given by

$$\sigma_a^2 = \frac{S_{xx}}{\Delta}, \quad \sigma_b^2 = \frac{S}{\Delta}, \quad \text{Cov}\,(a, b) = -\frac{S_x}{\Delta}.$$

The computed values of these quantities are: $S = 0.38664$; $S_x = 0.53225$; $S_y = 7$; $S_{xx} = 1.86221$; $S_{xy} = 21$; $\Delta = 0.43671$.

From these values, the best estimates of $\hat{a}$, $\hat{b}$ and the variances σ_a^2 and σ_b^2 are

$$\hat{a} = 4.2552, \quad \hat{b} = 10.061, \quad \sigma_a^2 = 4.264, \quad \sigma_b^2 = 0.8853.$$

The covariance is $\text{Cov}\,(a, b) = -1.2187$, giving estimates for a and b of

$$a = 4.3 \pm 2.1 \qquad \text{and} \qquad b = 10.06 \pm 0.94,$$

with a correlation coefficient $r_{ab} = -0.63$.

(a) The cosmic ray background must be present, since $n(0) \neq 0$, but its value of about 4 is uncertain to within a factor of 2.

(b) The correlation between a and b is negative and quite strong. This is as expected since, if the cosmic ray background represented by a were reduced towards zero, then b would have to be increased to compensate when fitting to the measured data for non-zero source strengths.

(c) A measure of the goodness-of-fit is the value of χ^2 achieved using the best-fit values for a and b. Direct resubstitution of the values found into $(*)$ gives $\chi^2 = 4.9$. If the weight of a particular reading is taken as the square root of the predicted (rather than the measured) value, then χ^2 rises slightly to 5.1. In either case the result is almost exactly that "expected" for 5 d.o.f. – neither too good nor too bad. There are five degrees of freedom because there are seven data points and two parameters have been chosen to give a best fit.

17.13 The following are the values and standard errors of a physical quantity $f(\theta)$ measured at various values of θ (in which there is negligible error):

θ	0	$\pi/6$	$\pi/4$	$\pi/3$
$f(\theta)$	3.72 ± 0.2	1.98 ± 0.1	-0.06 ± 0.1	-2.05 ± 0.1

θ	$\pi/2$	$2\pi/3$	$3\pi/4$	π
$f(\theta)$	-2.83 ± 0.2	1.15 ± 0.1	3.99 ± 0.2	9.71 ± 0.4

Theory suggests that f should be of the form $a_1 + a_2 \cos\theta + a_3 \cos 2\theta$. Show that the normal equations for the coefficients a_i are

$$\begin{aligned}
481.3a_1 + 158.4a_2 - 43.8a_3 &= 284.7, \\
158.4a_1 + 218.8a_2 + 62.1a_3 &= -31.1, \\
-43.8a_1 + 62.1a_2 + 131.3a_3 &= 368.4.
\end{aligned}$$

(a) If you have matrix inversion routines available on a computer, determine the best values and variances for the coefficients a_i and the correlation between the coefficients a_1 and a_2.
(b) If you have only a calculator available, solve for the values using a Gauss–Seidel iteration and start from the approximate solution $a_1 = 2$, $a_2 = -2$, $a_3 = 4$.

Assume that the measured data have uncorrelated errors. The quoted errors are not all equal and so the covariance matrix N, whilst being diagonal, will not be a multiple of the unit matrix; it will be

$$\mathsf{N} = \text{diag}(0.04,\ 0.01,\ 0.01,\ 0.01,\ 0.04,\ 0.01,\ 0.04,\ 0.16).$$

Using as base functions the three functions $h_1(\theta) = 1$, $h_2(\theta) = \cos\theta$ and $h_3(\theta) = \cos 2\theta$, we calculate the elements of the 8×3 response matrix $\mathsf{R}_{ij} = h_j(\theta_i)$. To save space we display its 3×8 transpose:

$$\mathsf{R}^{\mathrm{T}} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0.866 & 0.707 & 0.500 & 0 & -0.500 & -0.707 & -1 \\ 1 & 0.500 & 0 & -0.500 & -1 & -0.500 & 0 & 1 \end{pmatrix}.$$

Then

$$\mathsf{R}^{\mathrm{T}}\mathsf{N}^{-1} = \begin{pmatrix} 25 & 100 & 100 & 100 & 25 & 100 & 25 & 6.25 \\ 25 & 86.6 & 70.7 & 50 & 0 & -50 & -17.7 & -6.25 \\ 25 & 50.0 & 0 & -50.0 & -25 & -50 & 0 & 6.25 \end{pmatrix}$$

and

$$\mathsf{R}^{\mathrm{T}}\mathsf{N}^{-1}\mathsf{R} = \mathsf{R}^{\mathrm{T}}\mathsf{N}^{-1}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0.866 & 0.500 \\ 1 & 0.707 & 0 \\ 1 & 0.500 & -0.500 \\ 1 & 0 & -1 \\ 1 & -0.500 & -0.500 \\ 1 & -0.707 & 0 \\ 1 & -1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 481.25 & 158.35 & -43.75 \\ 158.35 & 218.76 & 62.05 \\ -43.75 & 62.05 & 131.25 \end{pmatrix}.$$

From the measured values,

$$\mathsf{f} = (3.72,\ 1.98,\ -0.06,\ -2.05,\ -2.83,\ 1.15,\ 3.99,\ 9.71)^{\mathrm{T}},$$

we need to calculate $\mathsf{R}^{\mathrm{T}}\mathsf{N}^{-1}\mathsf{f}$, which is given by

$$\begin{pmatrix} 25 & 100 & 100 & 100 & 25 & 100 & 25 & 6.25 \\ 25 & 86.6 & 70.7 & 50 & 0 & -50 & -17.7 & -6.25 \\ 25 & 50.0 & 0 & -50 & -25 & -50 & 0 & 6.25 \end{pmatrix}\begin{pmatrix} 3.72 \\ 1.98 \\ -0.06 \\ -2.05 \\ -2.83 \\ 1.15 \\ 3.99 \\ 9.71 \end{pmatrix},$$

i.e. $(284.7,\ -31.08,\ 368.44)^{\mathrm{T}}$.

The vector of LS estimators of a_i satisfies $\mathsf{R}^{\mathrm{T}}\mathsf{N}^{-1}\mathsf{R}\hat{\mathsf{a}} = \mathsf{R}^{\mathrm{T}}\mathsf{N}^{-1}\mathsf{f}$. Substituting the forms calculated above into the two sides of the equality gives the set of equations stated in the question.

(a) Machine (or manual!) inversion gives

$$(\mathsf{R}^{\mathrm{T}}\mathsf{N}^{-1}\mathsf{R})^{-1} = 10^{-3}\begin{pmatrix} 3.362 & -3.177 & 2.623 \\ -3.177 & 8.282 & -4.975 \\ 2.623 & -4.975 & 10.845 \end{pmatrix}.$$

From this (covariance matrix) we can calculate the standard errors on the a_i from the square roots of the terms on the leading diagonal as ± 0.058, ± 0.091 and ± 0.104. We can further calculate the correlation coefficient r_{12} between a_1 and a_2 as

$$r_{12} = \frac{-3.177 \times 10^{-3}}{0.058 \times 0.091} = -0.60.$$

The best values for the a_i are given by the result of multiplying the column matrix $(284.7,\ -31.08,\ 368.44)^{\mathrm{T}}$ by the above inverted matrix. This yields $(2.022,\ -2.944,\ 4.897)^{\mathrm{T}}$ to give the best estimates of the a_i as

$$a_1 = 2.02 \pm 0.06, \quad a_2 = -2.99 \pm 0.09, \quad a_3 = 4.90 \pm 0.10.$$

(b) Denote the given set of equations by $\mathsf{Aa} = \mathsf{b}$ and start by dividing each equation by the quantity needed to make the diagonal elements of A each equal to unity; this produces $\mathsf{Ca} = \mathsf{d}$. Then, writing $\mathsf{C} = \mathsf{I} - \mathsf{F}$ yields the basis of the iteration scheme,

$$\mathsf{a}_{n+1} = \mathsf{F}\mathsf{a}_n + \mathsf{d}.$$

We use only the simplest form of Gauss–Seidel iteration (with no separation into upper and lower diagonal matrices).

The explicit form of $\mathsf{Ca} = \mathsf{d}$ is

$$\begin{pmatrix} 1 & 0.3290 & -0.0909 \\ 0.7239 & 1 & 0.2836 \\ -0.3333 & 0.4728 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 0.5916 \\ -0.1421 \\ 2.8072 \end{pmatrix}$$

and

$$\mathsf{F} = \begin{pmatrix} 0 & -0.3290 & 0.0909 \\ -0.7239 & 0 & -0.2836 \\ 0.3333 & -0.4728 & 0 \end{pmatrix}.$$

Starting with the approximate solution $a_1 = 2$, $a_2 = -2$, $a_3 = 4$ gives as the result of the first ten iterations

a_1	a_2	a_3
2.000	−2.000	4.000
1.613	−2.724	4.419
1.890	−2.563	4.633
1.856	−2.824	4.649
1.943	−2.804	4.761
1.947	−2.899	4.781
1.980	−2.907	4.827
1.987	−2.944	4.842
2.000	−2.953	4.861
2.005	−2.969	4.870
2.011	−2.975	4.879

This final set of values is in close agreement with that obtained by direct inversion; in fact, after 18 iterations the values agree exactly to three significant figures. Of course, using this method makes it difficult to estimate the errors in the derived values.