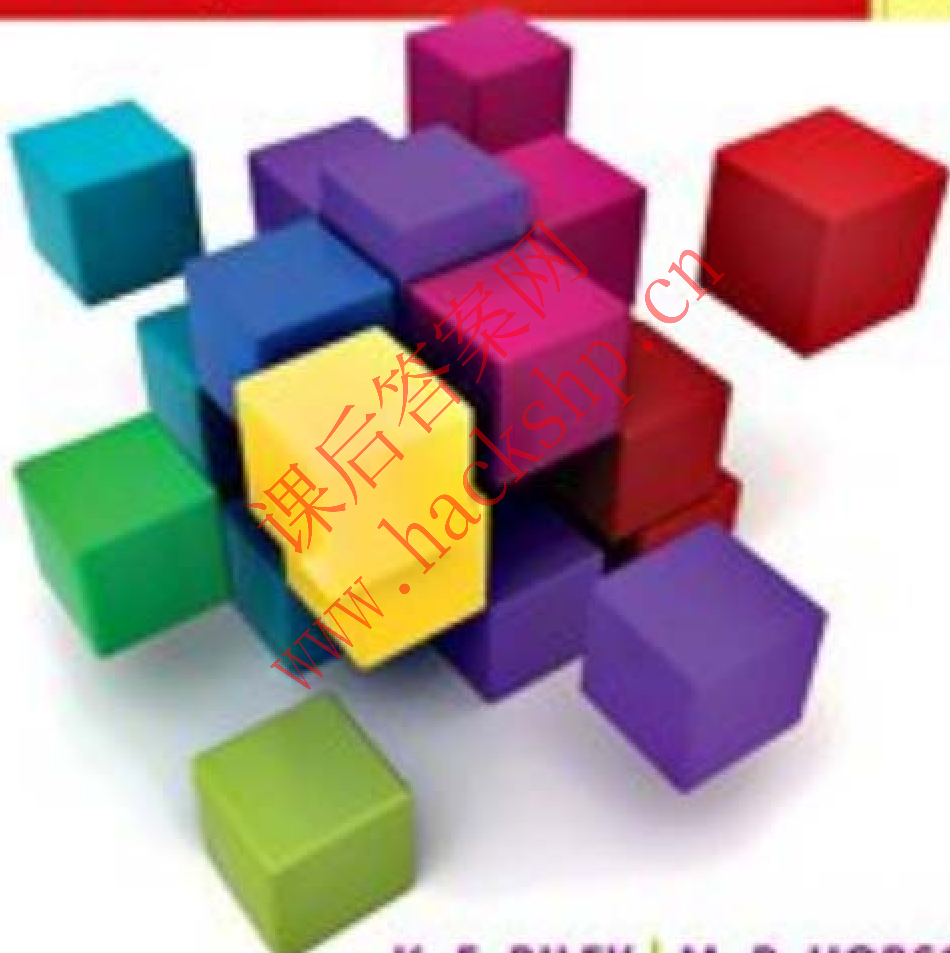
ESSENTIAL MATHEMATICAL METHODS

for the Physical Sciences

STUDENT SOLUTION MANUAL



K. F. RILEY | M. P. HOBSON

CAMBRIDGE www.cambridge.org/9780521141024

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

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.

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 $\{a_0, a_1, a_2, \ldots, a_N\}$ 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:

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

(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

$$\begin{vmatrix} x+2 & 2 & -5 \\ x+3 & -3 & 2 \\ x-2 & 1 & 3 \end{vmatrix} = \begin{vmatrix} x+2 & 2 & -5 \\ 1 & -5 & 7 \\ -4 & -1 & 8 \end{vmatrix}$$

$$= (x+2)(-40+7) + 2(-28-8) - 5(-1-20)$$

$$= -33(x+2) - 72 + 105$$

$$= -33x - 33.$$

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

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

(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 $\mathsf{AB} = \mathsf{BA}$, then

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

i.e. AB is Hermitian.

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

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

i.e. A and B commute.

Thus, AB is Hermitian $\iff$ 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,

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

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:

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

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

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

Continuing in this way, we find:

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

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

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

For $\lambda = 5$, an eigenvector $(x \quad y)^T$ must satisfy $x - 2y = 0$, whilst, for $\lambda = 10$, $4x + 2y = 0$. Thus a pair of independent eigenvectors of A are $(2 \quad 1)^T$ and $(1 \quad -2)^T$. 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 $\mathsf{A}\mathbf{x} = \mathbf{b}$, 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 $\mathsf{Ax} = (21 \quad 7 \quad 22)^{\mathrm{T}}$ 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 $\mathsf{LL}^{\mathrm{T}} = \mathsf{A}$. 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

$$b^2 + c^2 = 3 \quad \Rightarrow \quad c = \sqrt{\tfrac{5}{2}},$$
$$bd + ce = -1 \quad \Rightarrow \quad e = \sqrt{\tfrac{2}{5}}(-1 - \tfrac{3}{2}) = -\sqrt{\tfrac{5}{2}}.$$

And, from the final row,

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

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

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

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

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 $\mathbf{e}^i$, obtain the coefficients α_i involved when column matrix $\mathbf{x}$, which is the solution of

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

is expanded as $\mathbf{x} = \sum_i \alpha_i \mathbf{e}^i$. Here μ is a given constant and $\mathbf{v}$ 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 $\mathbf{v} = (1 \quad 2 \quad 3)^{\mathrm{T}}$.

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

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

$$\mathsf{A}\mathbf{x} - \mu\mathbf{x} = \mathbf{v},$$

$$\sum_i \mathsf{A}\alpha_i \mathbf{e}^i - \sum_i \mu\alpha_i \mathbf{e}^i = \mathbf{v},$$

$$\sum_i (\lambda_i \alpha_i \mathbf{e}^i - \mu\alpha_i \mathbf{e}^i) = \mathbf{v},$$

$$\alpha_j = \frac{(\mathbf{e}^j)^\dagger \mathbf{v}}{\lambda_j - \mu}.$$

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 $(\mathbf{e}^j)^\dagger \mathbf{v} = 0$.

(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,

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

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 $\mathbf{v} = (1 \quad 2 \quad 3)^{\mathrm{T}}$,

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

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 $(\mathbf{e}^3)^\dagger \mathsf{v} = 0$.

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

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

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

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

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

Again as expected, $(\mathbf{x}^1)^\dagger \mathbf{x}^2 = 0$, 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 $J(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathsf{A}\mathbf{x}/(\mathbf{x}^{\mathrm{T}}\mathbf{x})$ 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 $\mathbf{x}^{\mathrm{T}}\mathsf{A}\mathbf{x}$ can be written explicitly as

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

The scalar product $\mathbf{x}^{\mathrm{T}}\mathbf{x}$ 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 $(x_1 \;\; x_2)^{\mathrm{T}}$ 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,

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

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

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

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

$$(7-27)x + 10y - 10z = 0,$$
$$10x + (7-27)y - 10z = 0.$$

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 $S^\dagger$ to obtain

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

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

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

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

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

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

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

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

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

$$= \frac{1}{6}\begin{pmatrix} 1 & \sqrt{2} & \sqrt{3} \\ 2 & -\sqrt{2} & 0 \\ 1 & \sqrt{2} & -\sqrt{3} \end{pmatrix}\begin{pmatrix} \lambda & 2\lambda & \lambda \\ \sqrt{2}\mu & -\sqrt{2}\mu & \sqrt{2}\mu \\ \sqrt{3}\nu & 0 & -\sqrt{3}\nu \end{pmatrix}$$

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

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

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

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

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

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

$$0 = \begin{vmatrix} (c-\omega^2)m+d & -d & 0 \\ -d & (c-\omega^2)M+2d & -d \\ 0 & -d & (c-\omega^2)m+d \end{vmatrix}$$

$$= \begin{vmatrix} (c-\omega^2)m+d & 0 & -(c-\omega^2)m-d \\ -d & (c-\omega^2)M+2d & -d \\ 0 & -d & (c-\omega^2)m+d \end{vmatrix}$$

$$= [(c-\omega^2)m+d]\,\{[(c-\omega^2)M+2d]\,[(c-\omega^2)m+d]-d^2-d^2\}$$

$$= (cm-m\omega^2+d)(c-\omega^2)[Mm(c-\omega^2)+2dm+dM].$$

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

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

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

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

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

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

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

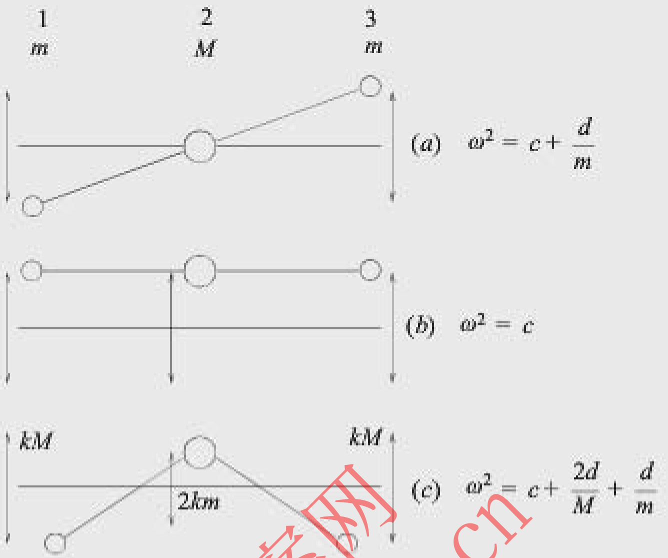


Figure 1.1 The normal modes, as viewed from above, of the coupled pendulums in Problem 1.39.

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

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

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

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

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.

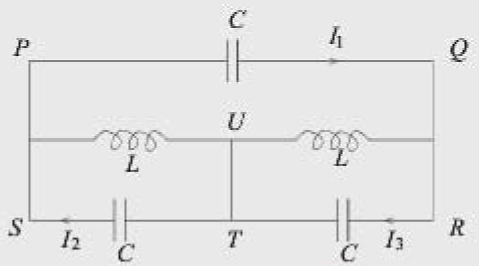


Figure 1.2 The circuit and notation for Problem 1.43.

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

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

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

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

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)

$$T \approx Ml^2\left(\tfrac{8}{3}\dot\theta_1^2 + 2\dot\theta_1\dot\theta_2 + \tfrac{2}{3}\dot\theta_2^2\right),$$
$$V \approx Mgl\left(\tfrac{3}{2}\theta_1^2 + \tfrac{1}{2}\theta_2^2\right).$$

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

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

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 $\begin{pmatrix} 1 & \sqrt{2} & 1 \end{pmatrix}^{\mathrm{T}}$ and $\begin{pmatrix} 1 & -\sqrt{2} & 1 \end{pmatrix}^{\mathrm{T}}$ 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 $f = \frac{y_{i-1} - y_i}{a}T$. 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,

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

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

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

This set of equations can be written as $\mathsf{B}\mathsf{x} = \frac{ma\omega^2}{T}\mathsf{x}$, 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 $\frac{\mathsf{x}^{\mathrm{T}}\mathsf{B}\mathsf{x}}{\mathsf{x}^{\mathrm{T}}\mathsf{x}}$ 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

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

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

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

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:

$$m\ddot{x} = qE - qB\left(\frac{qB}{m}x + v_0\right),$$

$$\Rightarrow \quad \ddot{x} + \left(\frac{qB}{m}\right)^2 x = \frac{q}{m}(E - v_0B).$$

(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

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

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

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

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

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

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

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

Thus, the full Laplacian in spherical polar coordinates reads

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

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 $\rho \leq a$ 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

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

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 $0 \leq \rho \leq a$.

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

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

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

For $0 < \rho \leq a$,

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

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

For $\rho > a$,

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

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

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

Now consider the scalar products:

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

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 $\mathbf{F}$:

$$2xz = F_x = \frac{\partial\phi}{\partial x} \quad\Rightarrow\quad \phi(x, y, z) = x^2z + g(y, z),$$

$$2yz^2 = F_y = \frac{\partial}{\partial y}[x^2z + g(y, z)] \quad\Rightarrow\quad g(y, z) = y^2z^2 + h(z),$$

$$x^2 + 2y^2z - 1 = F_z \quad = \quad \frac{\partial}{\partial z}[x^2z + y^2z^2 + h(z)]$$

$$\Rightarrow\quad h(z) = -z + k.$$

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:

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

(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:

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

(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:

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

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

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

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:

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

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,

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

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:

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

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 **F** in terms of **i**, **j** and **k**,

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

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

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

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, $\phi_1(x,y,z) = \frac{F_0}{2a}(x^2+y^2)\exp\left(-\frac{r^2}{a^2}\right)$, as it stands, is a suitable potential function for **F**(**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 **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 **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.