16

Series solutions of ordinary differential equations

16.1 *Find two power series solutions about $z = 0$ of the differential equation*

$$(1 - z^2)y'' - 3zy' + \lambda y = 0.$$

Deduce that the value of λ for which the corresponding power series becomes an Nth-degree polynomial $U_N(z)$ is $N(N+2)$. Construct $U_2(z)$ and $U_3(z)$.

If the equation is imagined divided through by $(1 - z^2)$ it is straightforward to see that, although $z = \pm 1$ are singular points of the equation, the point $z = 0$ is an *ordinary* point. We therefore expect two (uncomplicated!) series solutions with indicial values $\sigma = 0$ and $\sigma = 1$.

(a) $\sigma = 0$ and $y(z) = \sum_{n=0}^{\infty} a_n z^n$ with $a_0 \neq 0$.

Substituting and equating the coefficients of z^m,

$$(1 - z^2)\sum_{n=0}^{\infty} n(n-1)a_n z^{n-2} - 3\sum_{n=0}^{\infty} n a_n z^n + \lambda \sum_{n=0}^{\infty} a_n z^n = 0,$$
$$(m+2)(m+1)a_{m+2} - m(m-1)a_m - 3m a_m + \lambda a_m = 0,$$

gives as the recurrence relation

$$a_{m+2} = \frac{m(m-1) + 3m - \lambda}{(m+2)(m+1)}\, a_m = \frac{m(m+2) - \lambda}{(m+1)(m+2)}\, a_m.$$

Since this recurrence relation connects alternate coefficients a_m, and $a_0 \neq 0$, only the coefficients with even indices are generated. All such coefficients with index higher than m will become zero, and the series will become an Nth-degree polynomial $U_N(z)$, if $\lambda = m(m+2) = N(N+2)$ for some (even) m appearing in the series; here, this means any positive even integer N.

To construct $U_2(z)$ we need to take $\lambda = 2(2+2) = 8$. The recurrence relation gives a_2 as

$$a_2 = \frac{0-8}{(0+1)(0+2)}a_0 = -4a_0 \quad \Rightarrow \quad U_2(z) = a_0(1-4z^2).$$

(b) $\sigma = 1$ and $y(z) = z\sum_{n=0}^{\infty} a_n z^n$ with $a_0 \neq 0$.

Substituting and equating the coefficients of z^{m+1},

$$(1-z^2)\sum_{n=0}^{\infty}(n+1)na_n z^{n-1} - 3\sum_{n=0}^{\infty}(n+1)a_n z^{n+1} + \lambda\sum_{n=0}^{\infty} a_n z^{n+1} = 0,$$
$$(m+3)(m+2)a_{m+2} - (m+1)ma_m - 3(m+1)a_m + \lambda a_m = 0,$$

gives as the recurrence relation

$$a_{m+2} = \frac{m(m+1)+3(m+1)-\lambda}{(m+2)(m+3)}a_m = \frac{(m+1)(m+3)-\lambda}{(m+2)(m+3)}a_m.$$

Again, all coefficients with index higher than m will become zero, and the series will become an Nth-degree polynomial $U_N(z)$, if $\lambda = (m+1)(m+3) = N(N+2)$ for some (even) m appearing in the series; here, this means any positive odd integer N.

To construct $U_3(z)$ we need to take $\lambda = 3(3+2) = 15$. The recurrence relation gives a_2 as

$$a_2 = \frac{3-15}{(0+2)(0+3)}a_0 = -2a_0.$$

Thus,

$$U_3(z) = a_0(z-2z^3).$$

> ***16.3*** *Find power series solutions in z of the differential equation*
>
> $$zy'' - 2y' + 9z^5 y = 0.$$
>
> *Identify closed forms for the two series, calculate their Wronskian, and verify that they are linearly independent. Compare the Wronskian with that calculated from the differential equation.*

Putting the equation in its standard form shows that $z = 0$ is a singular point of the equation but, as $-2z/z$ and $9z^7/z$ are finite as $z \to 0$, it is a regular singular point. We therefore substitute a Frobenius type solution,

$$y(z) = z^{\sigma}\sum_{n=0}^{\infty} a_n z^n \text{ with } a_0 \neq 0,$$

and obtain

$$\sum_{n=0}^{\infty}(n+\sigma)(n+\sigma-1)a_n z^{n+\sigma-1} \\ -2\sum_{n=0}^{\infty}(n+\sigma)a_n z^{n+\sigma-1}+9\sum_{n=0}^{\infty}a_n z^{n+\sigma+5}=0.$$

Equating the coefficient of $z^{\sigma-1}$ to zero gives the indicial equation as

$$\sigma(\sigma-1)a_0-2\sigma a_0=0 \quad\Rightarrow\quad \sigma=0,\ 3.$$

These differ by an integer and may or may not yield two independent solutions. The larger root, $\sigma=3$, will give a solution; the smaller one, $\sigma=0$, may not.

(a) $\sigma=3$.

Equating the general coefficient of z^{m+2} to zero (with $\sigma=3$) gives

$$(m+3)(m+2)a_m-2(m+3)a_m+9a_{m-6}=0.$$

Hence the recurrence relation is

$$a_m=-\frac{9a_{m-6}}{m(m+3)},$$
$$\Rightarrow\quad a_{6p}=-\frac{9}{6p\,(6p+3)}\,a_{6p-6}=-\frac{a_{6p-6}}{2p\,(2p+1)}=\frac{(-1)^p a_0}{(2p+1)!}.$$

The first solution is therefore given by

$$y_1(x)=a_0z^3\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}z^{6n}=a_0\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}z^{3(2n+1)}=a_0\sin z^3.$$

(b) $\sigma=0$.

Equating the general coefficient of z^{m-1} to zero (with $\sigma=0$) gives

$$m(m-1)a_m-2ma_m+9a_{m-6}=0.$$

Hence the recurrence relation is

$$a_m=-\frac{9a_{m-6}}{m(m-3)},$$
$$\Rightarrow\quad a_{6p}=-\frac{9}{6p\,(6p-3)}\,a_{6p-6}=-\frac{a_{6p-6}}{2p\,(2p-1)}=\frac{(-1)^p a_0}{(2p)!}.$$

A second solution is thus

$$y_2(x)=a_0\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n)!}z^{6n}=a_0\cos z^3.$$

We see that $\sigma=0$ does, in fact, produce a (different) series solution. This is because the recurrence relation relates a_n to a_{n+6} and does not involve a_{n+3};

the relevance here of considering the subscripted index '$m+3$' is that '3' is the difference between the two indicial values.

We now calculate the Wronskian of the two solutions, $y_1 = a_0 \sin z^3$ and $y_2 = b_0 \cos z^3$:

$$\begin{aligned} W(y_1, y_2) &= y_1 y_2' - y_2 y_1' \\ &= a_0 \sin z^3(-3b_0 z^2 \sin z^3) - b_0 \cos z^3(3a_0 z^2 \cos z^3) \\ &= -3a_0 b_0 z^2 \neq 0. \end{aligned}$$

The fact that the Wronskian is non-zero shows that the two solutions are linearly independent.

We can also calculate the Wronskian from the original equation in its standard form,

$$y'' - \frac{2}{z} y' + 9z^4 y = 0,$$

as

$$W = C \exp\left\{ -\int^z \frac{-2}{u}\, du \right\} = C \exp(2 \ln z) = Cz^2.$$

This is in agreement with the Wronskian calculated from the solutions, as it must be.

16.5 *Investigate solutions of Legendre's equation at one of its singular points as follows.*

(a) *Verify that $z = 1$ is a regular singular point of Legendre's equation and that the indicial equation for a series solution in powers of $(z-1)$ has a double root $\sigma = 0$.*
(b) *Obtain the corresponding recurrence relation and show that a polynomial solution is obtained if ℓ is a positive integer.*
(c) *Determine the radius of convergence R of the $\sigma = 0$ series and relate it to the positions of the singularities of Legendre's equation.*

(a) In standard form, Legendre's equation reads

$$y'' - \frac{2z}{1-z^2} y' + \frac{\ell(\ell+1)}{1-z^2} y = 0.$$

This has a singularity at $z = 1$, but, since

$$\frac{-2z(z-1)}{1-z^2} \to 1 \text{ and } \frac{\ell(\ell+1)(z-1)^2}{1-z^2} \to 0 \text{ as } z \to 1,$$

i.e. both limits are finite, the point is a regular singular point.

We next change the origin to the point $z = 1$ by writing $u = z - 1$ and $y(z) = f(u)$. The transformed equation is

$$f'' - \frac{2(u+1)}{-u(u+2)} f' + \frac{\ell(\ell+1)}{-u(u+2)} y = 0$$

or $$-u(u+2)f'' - 2(u+1)f' + \ell(\ell+1)f = 0.$$

The point $u = 0$ is a regular singular point of this equation and so we set $f(u) = u^\sigma \sum_{n=0}^{\infty} a_n u^n$ and obtain

$$-u(u+2)\sum_{n=0}^{\infty}(\sigma+n)(\sigma+n-1)a_n u^{\sigma+n-2}$$
$$- 2(u+1)\sum_{n=0}^{\infty}(\sigma+n)a_n u^{\sigma+n-1} + \ell(\ell+1)\sum_{n=0}^{\infty} a_n u^{\sigma+n} = 0.$$

Equating to zero the coefficient of $u^{\sigma-1}$ gives

$$-2\sigma(\sigma-1)a_0 - 2\sigma a_0 = 0 \quad \Rightarrow \quad \sigma^2 = 0;$$

i.e. the indicial equation has a double root $\sigma = 0$.

(b) To obtain the recurrence relation we set the coefficient of u^m equal to zero for general m:

$$-m(m-1)a_m - 2(m+1)ma_{m+1} - 2ma_m - 2(m+1)a_{m+1} + \ell(\ell+1)a_m = 0.$$

Tidying this up gives

$$2(m+1)(m+1)a_{m+1} = [\,\ell(\ell+1) - m^2 + m - 2m\,]a_m,$$
$$\Rightarrow \quad a_{m+1} = \frac{\ell(\ell+1) - m(m+1)}{2(m+1)^2}\, a_m.$$

From this it is clear that, if ℓ is a positive integer, then $a_{\ell+1}$ and all further a_n are zero and that the solution is a polynomial (of degree ℓ).

(c) The limit of the ratio of successive terms in the series is given by

$$\left|\frac{a_{n+1}u^{n+1}}{a_n u^n}\right| = \left|\frac{u[\,\ell(\ell+1) - m(m+1)\,]}{2(m+1)^2}\right| \to \frac{|u|}{2} \text{ as } m \to \infty.$$

For convergence this limit needs to be < 1, i.e. $|u| < 2$. Thus the series converges in a circle of radius 2 centred on $u = 0$, i.e. on $z = 1$. The value 2 is to be expected, as it is the distance from $z = 1$ of the next nearest (actually the only other) singularity of the equation (at $z = -1$), excluding $z = 1$ itself.

16.7 *The first solution of Bessel's equation for $\nu = 0$ is*

$$J_0(z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!\Gamma(n+1)} \left(\frac{z}{2}\right)^{2n}.$$

Use the derivative method to show that

$$J_0(z)\ln z - \sum_{n=1}^{\infty} \frac{(-1)^n}{(n!)^2} \left(\sum_{r=1}^{n} \frac{1}{r}\right) \left(\frac{z}{2}\right)^{2n}$$

is a second solution.

Bessel's equation with $\nu = 0$ reads

$$zy'' + y' + zy = 0.$$

The recurrence relations that gave rise to the first solution, $J_0(z)$, were $(\sigma+1)^2 a_1 = 0$ and $(\sigma+n)^2 a_n + a_{n-2} = 0$ for $n \geq 2$. Thus, in a general form as a function of σ, the solution is given by

$$y(\sigma, z) = a_0 z^{\sigma} \left\{ 1 - \frac{z^2}{(\sigma+2)^2} + \frac{z^4}{(\sigma+2)^2(\sigma+4)^2} - \cdots + \frac{(-1)^n z^{2n}}{[\,(\sigma+2)(\sigma+4)\ldots(\sigma+2n)\,]^2} + \cdots \right\}.$$

Setting $\sigma = 0$ reproduces the first solution given above.

To obtain a second independent solution, we must differentiate the above expression with respect to σ, before setting σ equal to 0:

$$\frac{\partial y}{\partial \sigma} = \ln z\, J_0(z) + \sum_{n=1}^{\infty} \frac{da_{2n}(\sigma)}{d\sigma}\, z^{\sigma+2n} \text{ at } \sigma = 0.$$

Now

$$\begin{aligned}
\left.\frac{da_{2n}(\sigma)}{d\sigma}\right|_{\sigma=0} &= \frac{d}{d\sigma}\left\{\frac{(-1)^n}{[\,(\sigma+2)(\sigma+4)\ldots(\sigma+2n)\,]^2}\right\}_{\sigma=0} \\
&= \frac{(-1)^n(-2)}{[\ldots]^3}\left(\frac{[\ldots]}{\sigma+2} + \frac{[\ldots]}{\sigma+4} + \cdots + \frac{[\ldots]}{\sigma+2n}\right) \\
&= \frac{(-2)(-1)^n}{[\ldots]^2} \sum_{r=1}^{n} \frac{1}{\sigma+2r} \\
&= \frac{-2(-1)^n}{2^{2n}(n!)^2} \sum_{r=1}^{n} \frac{1}{2r}, \qquad \text{at } \sigma = 0.
\end{aligned}$$

Substituting this result, we obtain the second series as

$$J_0(z)\ln z - \sum_{n=1}^{\infty} \frac{(-1)^n}{(n!)^2}\left(\sum_{r=1}^{n}\frac{1}{r}\right)\left(\frac{z}{2}\right)^{2n}.$$

This is the form given in the question.

16.9 *Find series solutions of the equation $y'' - 2zy' - 2y = 0$. Identify one of the series as $y_1(z) = \exp z^2$ and verify this by direct substitution. By setting $y_2(z) = u(z)y_1(z)$ and solving the resulting equation for $u(z)$, find an explicit form for $y_2(z)$ and deduce that*

$$\int_0^x e^{-v^2}\,dv = e^{-x^2}\sum_{n=0}^{\infty}\frac{n!}{2(2n+1)!}(2x)^{2n+1}.$$

(a) The origin is an ordinary point of the equation and so power series solutions will be possible. Substituting $y(z) = \sum_{n=0}^{\infty} a_n z^n$ gives

$$\sum_{n=0}^{\infty} n(n-1)a_n z^{n-2} - 2\sum_{n=0}^{\infty} n a_n z^n - 2\sum_{n=0}^{\infty} a_n z^n = 0.$$

Equating to zero the coefficient of z^{m-2} yields the recurrence relation

$$a_m = \frac{2m-2}{m(m-1)}\,a_{m-2} = \frac{2}{m}\,a_{m-2}.$$

The solution with $a_0 = 1$ and $a_1 = 0$ is therefore

$$\begin{aligned} y_1(z) &= 1 + \frac{2z^2}{2} + \frac{2^2 z^4}{(2)(4)} + \cdots + \frac{2^n z^{2n}}{2^n\, n!} + \cdots \\ &= \sum_{n=0}^{\infty}\frac{z^{2n}}{n!} = \exp z^2. \end{aligned}$$

Putting this result into the original equation,

$$(4z^2+2)\exp z^2 - 2z\,2z\exp z^2 - 2\exp z^2 = 0,$$

shows directly that it is a valid solution.

The solution with $a_0 = 0$ and $a_1 = 1$ takes the form

$$\begin{aligned} y_2(z) &= z + \frac{2z^3}{3} + \frac{2^2 z^5}{(3)(5)} + \cdots + \frac{2^n\, 2^n\, n!\, z^{2n+1}}{(2n+1)!} + \cdots \\ &= \sum_{n=0}^{\infty}\frac{n!\,(2z)^{2n+1}}{2(2n+1)!}. \end{aligned}$$

We now set $y_2(z) = u(z)y_1(z)$ and substitute it into the original equation. As they must, the terms in which u is undifferentiated cancel and leave

$$u'' \exp z^2 + 2u'(2z \exp z^2) - 2zu' \exp z^2 = 0.$$

It follows that

$$\frac{u''}{u'} = -2z \quad \Rightarrow \quad u' = Ae^{-z^2} \quad \Rightarrow \quad u(x) = A\int^x e^{-v^2}\,dv.$$

Hence, setting the two derived forms for a second solution equal to each other, we have

$$\sum_{n=0}^{\infty} \frac{n!\,(2x)^{2n+1}}{2(2n+1)!} = y_2(x) = y_1(x)u(x) = e^{x^2} A\int^x e^{-v^2}\,dv.$$

For arbitrary small x, only the $n = 0$ term in the series is significant and takes the value $2x/2 = x$, whilst the integral is $A\int^x 1\,dv = Ax$. Thus $A = 1$ and the equality

$$\int_0^x e^{-v^2}\,dv = e^{-x^2}\sum_{n=0}^{\infty} \frac{n!\,(2x)^{2n+1}}{2(2n+1)!}$$

holds for all x.

16.11 *Find the general power series solution about $z = 0$ of the equation*

$$z\frac{d^2y}{dz^2} + (2z-3)\frac{dy}{dz} + \frac{4}{z}y = 0.$$

The origin is clearly a singular point of this equation but, since $z(2z-3)/z$ and $4z^2/z^2$ are finite as $z \to 0$, it is a regular singular point. The equation will therefore have at least one Frobenius-type solution of the form $y(z) = z^\sigma \sum_{n=0}^{\infty} a_n z^n$.

The indicial equation for the solution can be read off directly from $z^2y'' + z(2z-3)y' + 4y = 0$ as

$$\sigma(\sigma-1) - 3\sigma + 4 = (\sigma-2)^2 = 0 \quad \Rightarrow \quad \sigma = 2 \text{ (repeated root)}.$$

The recurrence relation in terms of a general σ is needed and is provided by setting the coefficient of $z^{m+\sigma}$ equal to 0:

$$(m+\sigma)(m-1+\sigma)a_m + 2(m-1+\sigma)a_{m-1} - 3(m+\sigma)a_m + 4a_m = 0.$$

This relation can be simplified and then applied repeatedly to give a_m in terms of

a_0 and hence an explicit expression for $y(\sigma, z)$:

$$\begin{aligned} a_m &= \frac{-2(m-1+\sigma)}{(m+\sigma)^2 - (m+\sigma) - 3(m+\sigma) + 4}\, a_{m-1} \\ &= \frac{-2(m-1+\sigma)}{(m+\sigma-2)^2}\, a_{m-1} \qquad\qquad \text{for } m \geq 1 \\ &= (-2)^m \frac{(m-1+\sigma)(m-2+\sigma)\ldots\sigma}{(m-2+\sigma)^2(m-3+\sigma)^2\ldots(\sigma-1)^2}\, a_0 \\ &= (-2)^m \frac{(m-1+\sigma)}{(m-2+\sigma)(m-3+\sigma)\ldots\sigma(\sigma-1)^2}\, a_0. \end{aligned}$$

Because of the form of the recurrence relation, we write the $n = 0$ and $n = 1$ terms explicitly:

$$\begin{aligned} y(\sigma, z) = a_0 z^\sigma &- \frac{2\sigma}{(\sigma-1)^2}\, a_0 z^{\sigma+1} \\ &+ z^\sigma \sum_{n=2}^{\infty} \frac{(n-1+\sigma)(-2z)^n}{(n-2+\sigma)(n-3+\sigma)\ldots\sigma(\sigma-1)^2}. \end{aligned}$$

We also need the derivative of this with respect to σ. As always, the derivative consists of two terms, the first of which is $y(\sigma, z)\ln z$. The second, in this case, is

$$\begin{aligned} &\frac{2(\sigma+1)}{(\sigma-1)^3}\, a_0 z^{\sigma+1} + a_0 z^\sigma \sum_{n=2}^{\infty} \frac{(n-1+\sigma)(-2z)^n}{(n-2+\sigma)(n-3+\sigma)\ldots\sigma(\sigma-1)^2} \\ &\qquad \times \left[\frac{1}{n-1+\sigma} - \frac{1}{n-2+\sigma} - \frac{1}{n-3+\sigma} - \cdots - \frac{1}{\sigma} - \frac{2}{\sigma-1}\right]. \end{aligned}$$

The factor in square brackets is obtained by considering $a_n(\sigma)$ as the product of factors of the form $(\sigma + \alpha)^\beta$; differentiation of the product with respect to σ produces a sum of terms, each of which is the original product divided by $(\sigma + \alpha)$, for some α, and multiplied by the corresponding β. In the actual expression, β takes the values $+1$ (once), -1 (on $n - 1$ occasions) and -2 (once).

To obtain two independent solutions, we finally set $\sigma = 2$ and $a_0 = 1$ obtaining

$$\begin{aligned} y_1(z) &= \sum_{n=0}^{\infty} \frac{(n+1)\,(-2)^n\, z^{n+2}}{n!}, \\ y_2(z) &= y_1(z)\ln z + 6a_0 z^3 \\ &\quad + \sum_{n=2}^{\infty} \frac{(n+1)\,(-2)^n\, z^{n+2}}{n!} \left[\frac{1}{n+1} - \frac{1}{n} - \frac{1}{n-1} - \cdots - \frac{1}{2} - 2\right]. \end{aligned}$$

The general solution is any linear combination of $y_1(z)$ and $y_2(z)$.

16.13 *For the equation $y'' + z^{-3}y = 0$, show that the origin becomes a regular singular point if the independent variable is changed from z to $x = 1/z$. Hence find a series solution of the form $y_1(z) = \sum_0^\infty a_n z^{-n}$. By setting $y_2(z) = u(z)y_1(z)$ and expanding the resulting expression for du/dz in powers of z^{-1}, show that $y_2(z)$ is a second solution with asymptotic form*

$$y_2(z) = c\left[z + \ln z - \tfrac{1}{2} + O\left(\frac{\ln z}{z}\right)\right],$$

where c is an arbitrary constant.

With the equation in its original form, it is clear that, since $z^2/z^3 \to \infty$ as $z \to 0$, the origin is an irregular singular point. However, if we set $1/z = \xi$ and $y(z) = Y(\xi)$, with

$$\frac{d\xi}{dz} = -\frac{1}{z^2} = -\xi^2 \quad \Rightarrow \quad \frac{d}{dz} = -\xi^2\frac{d}{d\xi},$$

then

$$\begin{aligned}
-\xi^2\frac{d}{d\xi}\left(-\xi^2\frac{dY}{d\xi}\right) + \xi^3 Y &= 0,\\
\xi^2\frac{d^2Y}{d\xi^2} + 2\xi\frac{dY}{d\xi} + \xi Y &= 0,\\
Y'' + \frac{2}{\xi}Y' + \frac{1}{\xi}Y &= 0.
\end{aligned}$$

By inspection, $\xi = 0$ is a regular singular point of this equation, and its indicial equation is

$$\sigma(\sigma - 1) + 2\sigma = 0 \quad \Rightarrow \quad \sigma = 0,\ -1.$$

We start with the larger root, $\sigma = 0$, as this is 'guaranteed' to give a valid series solution and assume a solution of the form $Y(\xi) = \sum_{n=0}^\infty a_n\xi^n$, leading to

$$\sum_{n=0}^\infty n(n-1)a_n\xi^{n-1} + 2\sum_{n=0}^\infty na_n\xi^{n-1} + \sum_{n=0}^\infty a_n\xi^n = 0.$$

Equating to zero the coefficient of ξ^{m-1} gives the recurrence relation

$$a_m = \frac{-a_{m-1}}{m(m+1)} \quad \Rightarrow \quad a_m = \frac{(-1)^m}{(m+1)\,(m!)^2}\,a_0$$

and the series solution in inverse powers of z,

$$y_1(z) = a_0\sum_{n=0}^\infty \frac{(-1)^n}{(n+1)\,(n!)^2\,z^n}.$$

To find the second solution we set $y_2(z) = f(z)y_1(z)$. As usual (and as intended), all terms with f undifferentiated vanish when this is substituted in the original equation. What is left is

$$0 = f''(z)y_1(z) + 2f'(z)y_1'(z),$$

which on rearrangement yields

$$\frac{f''}{f'} = -\frac{2y_1'}{y_1}.$$

This equation, although it contains a second derivative, is in fact only a first-order equation (for f'). It can be integrated directly to give

$$\ln f' = -2\ln y_1 + c.$$

After exponentiation, this equation can be written as

$$\begin{aligned}\frac{df}{dz} = \frac{A}{y_1^2(z)} &= \frac{A}{a_0^2}\left(1 - \frac{1}{2 \times 1^2\, z} + \frac{1}{3 \times 2^2\, z^2} - \cdots\right)^{-2} \\ &= \frac{A}{a_0^2}\left[1 + \frac{1}{z} + \mathrm{O}\left(\frac{1}{z^2}\right)\right],\end{aligned}$$

where $A = e^c$.

Hence, on integrating a second time, one obtains

$$f(z) = \frac{A}{a_0^2}\left(z + \ln z + \mathrm{O}\left(\frac{1}{z}\right)\right),$$

which in turn implies

$$\begin{aligned}y_2(z) &= \frac{A}{a_0^2}\left[z + \ln z + \mathrm{O}\left(\frac{1}{z}\right)\right] a_0\left(1 - \frac{1}{2z} + \frac{1}{12z^2} - \cdots\right) \\ &= c\left[z + \ln z - \frac{1}{2} + \mathrm{O}\left(\frac{\ln z}{z}\right)\right].\end{aligned}$$

This establishes the asymptotic form of the second solution.

16.15 *The origin is an ordinary point of the Chebyshev equation,*

$$(1-z^2)y'' - zy' + m^2 y = 0,$$

which therefore has series solutions of the form $z^\sigma \sum_0^\infty a_n z^n$ *for* $\sigma = 0$ *and* $\sigma = 1$.

(a) *Find the recurrence relationships for the* a_n *in the two cases and show that there exist polynomial solutions* $T_m(z)$:

(i) *for* $\sigma = 0$, *when* m *is an even integer, the polynomial having* $\frac{1}{2}(m+2)$ *terms;*

(ii) *for* $\sigma = 1$, *when* m *is an odd integer, the polynomial having* $\frac{1}{2}(m+1)$ *terms.*

(b) $T_m(z)$ *is normalised so as to have* $T_m(1) = 1$. *Find explicit forms for* $T_m(z)$ *for* $m = 0, 1, 2, 3$.

(c) *Show that the corresponding non-terminating series solutions* $S_m(z)$ *have as their first few terms*

$$S_0(z) = a_0\left(z + \frac{1}{3!}z^3 + \frac{9}{5!}z^5 + \cdots\right),$$

$$S_1(z) = a_0\left(1 - \frac{1}{2!}z^2 - \frac{3}{4!}z^4 - \cdots\right),$$

$$S_2(z) = a_0\left(z - \frac{3}{3!}z^3 - \frac{15}{5!}z^5 - \cdots\right),$$

$$S_3(z) = a_0\left(1 - \frac{9}{2!}z^2 + \frac{45}{4!}z^4 + \cdots\right).$$

(a)(i) If, for $\sigma = 0$, $y(z) = \sum_{n=0}^\infty a_n z^n$ with $a_0 \neq 0$, the condition for the coefficient of z^r in

$$(1-z^2)\sum_{n=0}^\infty n(n-1)a_n z^{n-2} - z\sum_{n=0}^\infty n a_n z^{n-1} + m^2 \sum_{n=0}^\infty a_n z^n$$

to be zero is that

$$(r+2)(r+1)a_{r+2} - r(r-1)a_r - ra_r + m^2 a_r = 0,$$

$$\Rightarrow \quad a_{r+2} = \frac{r^2 - m^2}{(r+2)(r+1)}\, a_r.$$

This relation relates a_{r+2} to a_r and so to a_0 if r is even. For a_{r+2} to vanish, in this case, requires that $r = m$, which must therefore be an even integer. The non-vanishing coefficients will be $a_0, a_2, \ldots, a_m$, i.e. $\frac{1}{2}(m+2)$ of them in all.

(ii) If, for $\sigma = 1$, $y(z) = \sum_{n=0}^\infty a_n z^{n+1}$ with $a_0 \neq 0$, the condition for the coefficient of z^{r+1} in

$$(1-z^2)\sum_{n=0}^{\infty}(n+1)na_n z^{n-1} - z\sum_{n=0}^{\infty}(n+1)a_n z^n + m^2\sum_{n=0}^{\infty}a_n z^{n+1}$$

to be zero is that

$$(r+3)(r+2)a_{r+2} - (r+1)ra_r - (r+1)a_r + m^2 a_r = 0,$$
$$\Rightarrow \quad a_{r+2} = \frac{(r+1)^2 - m^2}{(r+3)(r+2)}\, a_r.$$

This relation relates a_{r+2} to a_r and so to a_0 if r is even. For a_{r+2} to vanish, in this case, requires that $r+1=m$, which must therefore be an odd integer. The non-vanishing coefficients will be, as before, $a_0, a_2, \dots, a_{m-1}$, i.e. $\frac{1}{2}(m+1)$ of them in all.

(b) For $m=0$, $T_0(z)=a_0$. With the given normalisation, $a_0=1$ and $T_0(z)=1$.

For $m=1$, $T_1(z)=a_0 z$. The required normalisation implies that $a_0=1$ and so $T_0(z)=z$.

For $m=2$, we need the recurrence relation in (a)(i). This shows that

$$a_2 = \frac{0^2-2^2}{(2)(1)}\, a_0 = -2a_0 \quad \Rightarrow \quad T_2(z) = a_0(1-2z^2).$$

With the given normalisation, $a_0=-1$ and $T_2(z)=2z^2-1$.

For $m=3$, we use the recurrence relation in (a)(ii) and obtain

$$a_2 = \frac{1^2-3^2}{(3)(2)}\, a_0 = -\frac{4}{3}a_0 \quad \Rightarrow \quad T_3(z) = a_0(z - \frac{4z^3}{3}).$$

For the required normalisation, we must have $a_0 = -\frac{1}{3}$ and consequently that $T_3(z) = 4z^3 - 3z$.

(c) The non-terminating series solutions $S_m(z)$ arise when $\sigma=0$ but m is an odd integer and when $\sigma=1$ with m an even integer. We take each in turn and apply the appropriate recurrence relation to generate the coefficients.

(i) $\sigma=0$, $m=1$, using the (a)(i) recurrence relation:

$$a_2 = \frac{0-1}{(2)(1)}\, a_0 = -\frac{1}{2!}\, a_0, \quad a_4 = \frac{4-1}{(4)(3)}\, a_2 = -\frac{3}{4!}\, a_0.$$

Hence,

$$S_1(z) = a_0\left(1 - \frac{1}{2!}z^2 - \frac{3}{4!}z^4 - \cdots\right).$$

(ii) $\sigma=0$, $m=3$, using the (a)(i) recurrence relation:

$$a_2 = \frac{0-9}{(2)(1)}\, a_0 = -\frac{9}{2!}\, a_0, \quad a_4 = \frac{4-9}{(4)(3)}\, a_2 = \frac{45}{4!}\, a_0.$$

Hence,

$$S_3(z) = a_0 \left(1 - \frac{9}{2!} z^2 + \frac{45}{4!} z^4 + \cdots \right).$$

(iii) $\sigma = 1$, $m = 0$, using the (a)(ii) recurrence relation:

$$a_2 = \frac{1 - 0}{(3)(2)} a_0 = \frac{1}{3!} a_0, \quad a_4 = \frac{9 - 0}{(5)(4)} a_2 = \frac{9}{5!} a_0.$$

Hence,

$$S_0(z) = a_0 \left(z + \frac{1}{3!} z^3 + \frac{9}{5!} z^5 + \cdots \right).$$

(iv) $\sigma = 1$, $m = 2$, using the (a)(ii) recurrence relation:

$$a_2 = \frac{1 - 4}{(3)(2)} a_0 = -\frac{3}{3!} a_0, \quad a_4 = \frac{9 - 4}{(5)(4)} a_2 = -\frac{15}{5!} a_0.$$

Hence,

$$S_2(z) = a_0 \left(z - \frac{3}{3!} z^3 - \frac{15}{5!} z^5 - \cdots \right).$$

17

Eigenfunction methods for differential equations

17.1 *By considering* $\langle h|h\rangle$, *where* $h = f + \lambda g$ *with* λ *real, prove that, for two functions* f *and* g,

$$\langle f|f\rangle\langle g|g\rangle \geq \tfrac{1}{4}[\langle f|g\rangle + \langle g|f\rangle]^2.$$

The function $y(x)$ *is real and positive for all* x. *Its Fourier cosine transform* $\tilde{y}_c(k)$ *is defined by*

$$\tilde{y}_c(k) = \int_{-\infty}^{\infty} y(x)\cos(kx)\,dx,$$

and it is given that $\tilde{y}_c(0) = 1$. *Prove that*

$$\tilde{y}_c(2k) \geq 2[\tilde{y}_c(k)]^2 - 1.$$

For any $|h\rangle$ we have that $\langle h|h\rangle \geq 0$, with equality only if $|h\rangle = |0\rangle$. Hence, noting that λ is real, we have

$$0 \leq \langle h|h\rangle = \langle f + \lambda g|f + \lambda g\rangle = \langle f|f\rangle + \lambda\langle g|f\rangle + \lambda\langle f|g\rangle + \lambda^2\langle g|g\rangle.$$

This equation, considered as a quadratic inequality in λ, states that the corresponding quadratic equation has no real roots. The condition for this ('$b^2 < 4ac$') is given by

$$[\,\langle g|f\rangle + \langle f|g\rangle\,]^2 \leq 4\langle f|f\rangle\langle g|g\rangle, \qquad (*)$$

from which the stated result follows immediately. Note that $\langle g|f\rangle + \langle f|g\rangle$ is real and its square is therefore non-negative.

The given datum is equivalent to

$$1 = \tilde{y}_c(0) = \int_{-\infty}^{\infty} y(x)\cos(0x)\,dx = \int_{-\infty}^{\infty} y(x)\,dx.$$

Now consider

$$\begin{aligned}
\tilde{y}_c(2k) &= \int_{-\infty}^{\infty} y(x)\cos(2kx)\,dx \\
&= 2\int_{-\infty}^{\infty} y(x)\cos^2 kx - \int_{-\infty}^{\infty} y(x)\,dx, \\
\Rightarrow \quad \tilde{y}_c(2k) + 1 &= 2\int_{-\infty}^{\infty} y(x)\cos^2 kx.
\end{aligned}$$

In order to use (∗), we need to choose for $f(x)$ and $g(x)$ functions whose product will form the integrand defining $\tilde{y}_c(k)$. With this in mind, we take $f(x) = y^{1/2}(x)\cos kx$ and $g(x) = y^{1/2}(x)$; we may do this since $y(x) > 0$ for all x. Making these choices gives

$$\begin{aligned}
\left(\int_{-\infty}^{\infty} y\cos kx\,dx + \int_{-\infty}^{\infty} y\cos kx\,dx\right)^2 &\leq 4\int_{-\infty}^{\infty} y\cos^2 kx\,dx \int_{-\infty}^{\infty} y\,dx, \\
\left(\int_{-\infty}^{\infty} 2y\cos kx\,dx\right)^2 &\leq 4\int_{-\infty}^{\infty} y\cos^2 kx\,dx \times 1, \\
4\tilde{y}_c^2(k) &\leq 4\int_{-\infty}^{\infty} y\cos^2 kx\,dx.
\end{aligned}$$

Thus,

$$\tilde{y}_c(2k) + 1 = 2\int_{-\infty}^{\infty} y(x)\cos^2 kx \geq 2[\tilde{y}_c(k)]^2$$

and hence the stated result.

17.3 *Consider the real eigenfunctions $y_n(x)$ of a Sturm–Liouville equation*

$$(py')' + qy + \lambda\rho y = 0, \qquad a \leq x \leq b,$$

in which $p(x)$, $q(x)$ and $\rho(x)$ are continuously differentiable real functions and $p(x)$ does not change sign in $a \leq x \leq b$. Take $p(x)$ as positive throughout the interval, if necessary by changing the signs of all eigenvalues. For $a \leq x_1 \leq x_2 \leq b$, establish the identity

$$(\lambda_n - \lambda_m)\int_{x_1}^{x_2} \rho y_n y_m\,dx = \left[\, y_n\, p\, y_m' - y_m\, p\, y_n' \,\right]_{x_1}^{x_2}.$$

Deduce that if $\lambda_n > \lambda_m$ then $y_n(x)$ must change sign between two successive zeros of $y_m(x)$.

[The reader may find it helpful to illustrate this result by sketching the first few eigenfunctions of the system $y'' + \lambda y = 0$, with $y(0) = y(\pi) = 0$, and the Legendre polynomials $P_n(z)$ for $n = 2, 3, 4, 5$.]

The function $p(x)$ does not change sign in the interval $a \leq x \leq b$; we take it as positive, multiplying the equation all through by -1 if necessary. This means that the weight function ρ can still be taken as positive, but that we must consider all possible functions for $q(x)$ and eigenvalues λ of either sign.

We start with the eigenvalue equation for $y_n(x)$, multiply it through by $y_m(x)$ and then integrate from x_1 to x_2. From this result we subtract the same equation with the roles of n and m reversed, as follows. The integration limits are omitted until the explicit integration by parts is carried through:

$$\int y_m(p\,y_n')'\,dx + \int y_m q\,y_n\,dx + \int y_m\lambda_n\rho y_n\,dx = 0,$$

$$\int y_n(p\,y_m')'\,dx + \int y_n q\,y_m\,dx + \int y_n\lambda_m\rho y_m\,dx = 0,$$

$$\int \left[\,y_m(p\,y_n')' - y_n(p\,y_m')'\,\right]\,dx + (\lambda_n - \lambda_m)\int y_m\rho y_n\,dx = 0,$$

$$\left[\,y_m p\,y_n'\,\right]_{x_1}^{x_2} - \int y_m' p\,y_n'\,dx - \left[\,y_n p\,y_m'\,\right]_{x_1}^{x_2} + \int y_n' p\,y_m'\,dx + (\lambda_n - \lambda_m)\int y_m\rho y_n\,dx = 0.$$

Hence

$$(\lambda_n - \lambda_m)\int y_m\rho y_n\,dx = \left[\,y_n p\,y_m' - y_m p\,y_n'\,\right]_{x_1}^{x_2}. \qquad (*)$$

Now, in this general result, take x_1 and x_2 as successive zeros of $y_m(x)$, where m is determined by $\lambda_n > \lambda_m$ (after the signs have been changed, if that was necessary). Clearly the sign of $y_m(x)$ does not change in this interval; let it be α. It follows that the sign of $y_m'(x_1)$ is also α, whilst that of $y_m'(x_2)$ is $-\alpha$. In addition, the second term on the RHS of $(*)$ vanishes at both limits, as $y_m(x_1) = y_m(x_2) = 0$.

Let us now *suppose* that the sign of $y_n(x)$ does not change in this same interval and is always β. Then the sign of the expression on the LHS of $(*)$ is $(+1)(\alpha)(+1)\beta = \alpha\beta$. The first $(+1)$ appears because $\lambda_n > \lambda_m$.

The signs of the upper- and lower-limit contributions of the remaining term on the RHS of $(*)$ are $\beta(+1)(-\alpha)$ and $(-1)\beta(+1)\alpha$, respectively, the additional factor of (-1) in the second product arising from the fact that the contribution comes from a lower limit. The contributions at both limits have the same sign, $-\alpha\beta$, and so the sign of the total RHS must also be $-\alpha\beta$.

This contradicts, however, the sign of $+\alpha\beta$ found for the LHS. It follows that it was wrong to suppose that the sign of $y_n(x)$ does not change in the interval; in other words, a zero of $y_n(x)$ does appear between every pair of zeros of $y_m(x)$.

17.5 *Use the properties of Legendre polynomials to carry out the following exercises.*

(a) *Find the solution of* $(1-x^2)y'' - 2xy' + by = f(x)$ *that is valid in the range* $-1 \le x \le 1$ *and finite at* $x = 0$*, in terms of Legendre polynomials.*
(b) *Find the explicit solution if* $b = 14$ *and* $f(x) = 5x^3$*. Verify it by direct substitution.*

[*Explicit forms for the Legendre polynomials can be found in any textbook. In* Mathematical Methods for Physics and Engineering, 3rd edition*, they are given in Subsection 18.1.1.*]

(a) The LHS of the given equation is the same as that of Legendre's equation and so we substitute $y(x) = \sum_{n=0}^{\infty} a_n P_n(x)$ and use the fact that $(1-x^2)P_n'' - 2xP_n' = -n(n+1)P_n$. This results in

$$\sum_{n=0}^{\infty} a_n [\, b - n(n+1)\,] P_n = f(x).$$

Now, using the mutual orthogonality and normalisation of the $P_n(x)$, we multiply both sides by $P_m(x)$ and integrate over x:

$$\sum_{n=0}^{\infty} a_n [\, b - n(n+1)\,]\, \delta_{mn} \frac{2}{2m+1} = \int_{-1}^{1} f(z) P_m(z)\, dz,$$

$$\Rightarrow \quad a_m = \frac{2m+1}{2[\, b - m(m+1)\,]} \int_{-1}^{1} f(z) P_m(z)\, dz.$$

This gives the coefficients in the solution $y(x)$.

(b) We now express $f(x)$ in terms of Legendre polynomials,

$$f(x) = 5x^3 = 2[\, \tfrac{1}{2}(5x^3 - 3x)\,] + 3[\, x\,] = 2P_3(x) + 3P_1(x),$$

and conclude that, because of the mutual orthogonality of the Legendre polynomials, only a_3 and a_1 in the series solution will be non-zero. To find them we need to evaluate

$$\int_{-1}^{1} f(z) P_3(z)\, dz = 2\, \frac{2}{2(3)+1} = \frac{4}{7};$$

similarly, $\int_{-1}^{1} f(z) P_1(z)\, dz = 3 \times (2/3) = 2$.

Inserting these values gives

$$a_3 = \frac{7}{2(14-12)}\, \frac{4}{7} = 1 \text{ and } a_1 = \frac{3}{2(14-2)}\, 2 = \frac{1}{4}.$$

Thus the solution is

$$y(x) = \frac{1}{4}P_1(x) + P_3(x) = \frac{1}{4}x + \frac{1}{2}(5x^3 - 3x) = \frac{5(2x^3 - x)}{4}.$$

Check:

$$(1 - x^2)\frac{60x}{4} - 2x\frac{30x^2 - 5}{4} + \frac{140x^3 - 70x}{4} = 5x^3,$$
$$\Rightarrow \quad 60x - 60x^3 - 60x^3 + 10x + 140x^3 - 70x = 20x^3,$$

which is satisfied.

17.7 *Consider the set of functions, $\{f(x)\}$, of the real variable x defined in the interval $-\infty < x < \infty$, that $\to 0$ at least as quickly as x^{-1}, as $x \to \pm\infty$. For unit weight function, determine whether each of the following linear operators is Hermitian when acting upon $\{f(x)\}$:*

$$\text{(a)}\ \frac{d}{dx} + x; \quad \text{(b)}\ -i\frac{d}{dx} + x^2; \quad \text{(c)}\ ix\frac{d}{dx}; \quad \text{(d)}\ i\frac{d^3}{dx^3}.$$

For an operator $\mathcal{L}$ to be Hermitian over the given range with respect to a unit weight function, the equation

$$\int_{-\infty}^{\infty} f^*(x)[\mathcal{L}g(x)]\,dx = \left\{\int_{-\infty}^{\infty} g^*(x)[\mathcal{L}f(x)]\,dx\right\}^* \qquad (*)$$

must be satisfied for general functions f and g.

(a) For $\mathcal{L} = \dfrac{d}{dx} + x$, the LHS of $(*)$ is

$$\begin{aligned}\int_{-\infty}^{\infty} f^*(x)\left(\frac{dg}{dx} + xg\right)\,dx &= \left[\,f^*g\,\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{df^*}{dx}g\,dx + \int_{-\infty}^{\infty} f^*xg\,dx \\ &= 0 - \int_{-\infty}^{\infty} \frac{df^*}{dx}g\,dx + \int_{-\infty}^{\infty} f^*xg\,dx.\end{aligned}$$

The RHS of $(*)$ is

$$\begin{aligned}\int_{-\infty}^{\infty} \left\{g^*(x)\left(\frac{df}{dx} + xf\right)\,dx\right\}^* &= \left\{\int_{-\infty}^{\infty} g^*\frac{df}{dx}\,dx\right\}^* + \left\{\int_{-\infty}^{\infty} g^*xf\,dx\right\}^* \\ &= \int_{-\infty}^{\infty} g\frac{df^*}{dx}\,dx + \int_{-\infty}^{\infty} gxf^*\,dx.\end{aligned}$$

Since the sign of the first term differs in the two expressions, the LHS $\neq$ RHS and $\mathcal{L}$ is *not* Hermitian. It will also be apparent that purely multiplicative terms in the operator, such as x or x^2, will always be Hermitian; thus we can ignore the x^2 term in part (b).

(b) As explained above, we need only consider

$$\int_{-\infty}^{\infty} f^*(x)\left(-i\frac{dg}{dx}\right)dx = \left[-if^*g\right]_{-\infty}^{\infty} + i\int_{-\infty}^{\infty}\frac{df^*}{dx}g\,dx$$
$$= 0 + i\int_{-\infty}^{\infty}\frac{df^*}{dx}g\,dx$$

and

$$\int_{-\infty}^{\infty}\left\{g^*(x)\left(-i\frac{df}{dx}\right)dx\right\}^* = i\int_{-\infty}^{\infty} g\frac{df^*}{dx}\,dx.$$

These are equal, and so $\mathcal{L} = -i\dfrac{d}{dx}$ is Hermitian, as is $\mathcal{L} = -i\dfrac{d}{dx} + x^2$.

(c) For $\mathcal{L} = ix\dfrac{d}{dx}$, the LHS of (∗) is

$$\int_{-\infty}^{\infty} f^*(x)\left(ix\frac{dg}{dx}\right)dx = \left[ixf^*g\right]_{-\infty}^{\infty} - i\int_{-\infty}^{\infty} x\frac{df^*}{dx}g\,dx - i\int_{-\infty}^{\infty} f^*g\,dx$$
$$= 0 - i\int_{-\infty}^{\infty} x\frac{df^*}{dx}g\,dx - i\int_{-\infty}^{\infty} f^*g\,dx.$$

The RHS of (∗) is given by

$$\int_{-\infty}^{\infty}\left\{g^*(x)ix\left(\frac{df}{dx}\right)dx\right\}^* = -i\int_{-\infty}^{\infty} gx\frac{df^*}{dx}\,dx.$$

Since, in general, $-i\int_{-\infty}^{\infty} fg^*\,dx \neq 0$, the two sides are not equal; therefore $\mathcal{L}$ is not Hermitian.

(d) Since $\mathcal{L} = i\dfrac{d^3}{dx^3}$ is the cube of the operator $-i\dfrac{d}{dx}$, which was shown in part (b) to be Hermitian, it is expected that $\mathcal{L}$ is Hermitian. This can be verified directly as follows.

The LHS of (∗) is given by

$$i\int_{-\infty}^{\infty} f^*\frac{d^3g}{dx^3}\,dx = \left[if^*\frac{d^2g}{dx^2}\right]_{-\infty}^{\infty} - i\int_{-\infty}^{\infty}\frac{df^*}{dx}\frac{d^2g}{dx^2}\,dx$$
$$= 0 - i\left[\frac{df^*}{dx}\frac{dg}{dx}\right]_{-\infty}^{\infty} + i\int_{-\infty}^{\infty}\frac{d^2f^*}{dx^2}\frac{dg}{dx}\,dx$$
$$= 0 + i\left[\frac{d^2f^*}{dx^2}g\right]_{-\infty}^{\infty} - i\int_{-\infty}^{\infty}\frac{d^3f^*}{dx^3}g\,dx$$
$$= 0 + \left\{\int_{-\infty}^{\infty} ig^*\frac{d^3f}{dx^3}\,dx\right\}^* = \text{ RHS of } (*).$$

Thus $\mathcal{L}$ is confirmed as Hermitian.

17.9 *Find an eigenfunction expansion for the solution with boundary conditions $y(0) = y(\pi) = 0$ of the inhomogeneous equation*

$$\frac{d^2y}{dx^2} + \kappa y = f(x),$$

where κ is a constant and

$$f(x) = \begin{cases} x & 0 \le x \le \pi/2, \\ \pi - x & \pi/2 < x \le \pi. \end{cases}$$

The eigenfunctions of the operator $\mathcal{L} = \dfrac{d^2}{dx^2} + \kappa$ are obviously

$$y_n(x) = A_n \sin nx + B_n \cos nx,$$

with corresponding eigenvalues $\lambda_n = n^2 - \kappa$.

The boundary conditions, $y(0) = y(\pi) = 0$, require that n is a positive integer and that $B_n = 0$, i.e.

$$y_n(x) = A_n \sin nx = \sqrt{\frac{2}{\pi}} \sin nx,$$

where A_n (for $n \ge 1$) has been chosen so that the eigenfunctions are normalised over the interval $x = 0$ to $x = \pi$. Since $\mathcal{L}$ is Hermitian on the range $0 \le x \le \pi$, the eigenfunctions are also mutually orthogonal, and so the $y_n(x)$ form an orthonormal set.

If the required solution is $y(x) = \sum_n a_n y_n(x)$, then direct substitution yields the result

$$\sum_{n=1}^{\infty} (\kappa - n^2) a_n y_n(x) = f(x).$$

Following the usual procedure for analysis using sets of orthonormal functions, this implies that

$$a_m = \frac{1}{\kappa - m^2} \int_0^{\pi} f(z) y_m(z)\, dz$$

and, consequently, that

$$y(x) = \sum_{n=1}^{\infty} \sqrt{\frac{2}{\pi}} \frac{\sin nx}{\kappa - n^2} \sqrt{\frac{2}{\pi}} \int_0^{\pi} f(z) \sin(nz)\, dz.$$

It only remains to evaluate

$$
\begin{aligned}
I_n &= \int_0^{\pi} \sin(nx) f(x)\,dx \\
&= \int_0^{\pi/2} x \sin nx\,dx + \int_{\pi/2}^{\pi} (\pi - x)\sin nx\,dx \\
&= \left[\frac{-x\cos nx}{n}\right]_0^{\pi/2} + \int_0^{\pi/2} \frac{\cos nx}{n}\,dx \\
&\quad + \left[\frac{-(\pi - x)\cos nx}{n}\right]_{\pi/2}^{\pi} + \int_{\pi/2}^{\pi} \frac{(-1)\cos nx}{n}\,dx \\
&= -\frac{\pi}{2}\frac{\cos(n\pi/2)}{n}(1-1) + \left[\frac{\sin nx}{n^2}\right]_0^{\pi/2} - \left[\frac{\sin nx}{n^2}\right]_{\pi/2}^{\pi} \\
&= 0 + \frac{(-1)^{(n-1)/2}}{n^2}(1+1) \text{ for odd } n \text{ and } = 0 \text{ for even } n.
\end{aligned}
$$

Thus,

$$
y(x) = \frac{4}{\pi}\sum_{n \text{ odd}} \frac{(-1)^{(n-1)/2}}{n^2(\kappa - n^2)} \sin nx
$$

is the required solution.

17.11 *The differential operator $\mathcal{L}$ is defined by*

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

Determine the eigenvalues λ_n of the problem

$$
\mathcal{L}y_n = \lambda_n e^x y_n \qquad 0 < x < 1,
$$

with boundary conditions

$$
y(0) = 0, \qquad \frac{dy}{dx} + \tfrac{1}{2}y = 0 \quad \textit{at} \quad x = 1.
$$

(a) *Find the corresponding unnormalised y_n, and also a weight function $\rho(x)$ with respect to which the y_n are orthogonal. Hence, select a suitable normalisation for the y_n.*
(b) *By making an eigenfunction expansion, solve the equation*

$$
\mathcal{L}y = -e^{x/2}, \qquad 0 < x < 1,
$$

subject to the same boundary conditions as previously.

When written out explicitly, the eigenvalue equation is

$$-\frac{d}{dx}\left(e^x\frac{dy}{dx}\right) - \tfrac{1}{4}e^x y = \lambda e^x y, \qquad (*)$$

or, on differentiating out the product,

$$e^x y'' + e^x y' + (\lambda + \tfrac{1}{4})e^x y = 0.$$

The auxiliary equation is

$$m^2 + m + (\lambda + \tfrac{1}{4}) = 0 \quad \Rightarrow \quad m = -\tfrac{1}{2} \pm i\sqrt{\lambda}.$$

The general solution is thus given by

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

with the condition $y(0) = 0$ implying that $A = 0$. The other boundary condition requires that, at $x = 1$,

$$-\tfrac{1}{2}Be^{-x/2}\sin\sqrt{\lambda}x + \sqrt{\lambda}Be^{-x/2}\cos\sqrt{\lambda}x + \tfrac{1}{2}Be^{-x/2}\sin\sqrt{\lambda}x = 0,$$

i.e. that $\cos\sqrt{\lambda} = 0$ and hence that $\lambda = (n + \frac{1}{2})^2\pi^2$ for non-negative integral n.

(a) The unnormalised eigenfunctions are

$$y_n(x) = B_n e^{-x/2}\sin(n + \tfrac{1}{2})\pi x$$

and $(*)$ is in Sturm–Liouville form. However, although $y_n(0) = 0$, the values at the upper limit in $\left[\, y_m' p\, y_n \right]_0^1$ are $y_n(1) = B_n e^{-1/2}(-1)^n$, $p(1) = e^1$ and $y_m'(1) = -\frac{1}{2}B_m e^{-1/2}(-1)^m$. Consequently, $\left[\, y_m' p\, y_n \right]_0^1 \neq 0$ and S–L theory cannot be applied. We therefore have to find a suitable weight function $\rho(x)$ by inspection. Given the general form of the eigenfunctions, ρ has to be taken as e^x, with the orthonormality integral taking the form

$$\begin{aligned} I_{nm} &= \int_0^1 \rho(x) y_n(x) y_m^*(x)\, dx \\ &= B_n B_m \int_0^1 e^x e^{-x/2}\sin[\,(n + \tfrac{1}{2})\pi x\,]e^{-x/2}\sin[\,(m + \tfrac{1}{2})\pi x\,]\, dx \\ &= \begin{cases} 0 & \text{for } m \neq n, \\ \tfrac{1}{2}B_n B_m & \text{for } m = n. \end{cases} \end{aligned}$$

It is clear that a suitable normalisation is $B_n = \sqrt{2}$ for all n.

(b) We write the solution as $y(x) = \sum_{n=0}^{\infty} a_n y_n(x)$, giving as the equation to be solved

$$\begin{aligned}
-e^{x/2} = \mathcal{L}y &= \mathcal{L}\sum_{n=0}^{\infty} a_n y_n(x) \\
&= \sum_{n=0}^{\infty} a_n [\,\lambda_n \rho(x) y_n(x)\,] \\
&= \sum_{n=0}^{\infty} a_n (n+\tfrac{1}{2})^2 \pi^2 e^x \sqrt{2} e^{-x/2} \sin[\,(n+\tfrac{1}{2})\pi x\,] \\
\Rightarrow \quad -1 &= \sum_{n=0}^{\infty} a_n (n+\tfrac{1}{2})^2 \pi^2 \sqrt{2} \sin[\,(n+\tfrac{1}{2})\pi x\,].
\end{aligned}$$

After multiplying both sides of this equation by $\sin(m+\frac{1}{2})\pi x$ and integrating from 0 to 1, we obtain

$$\begin{aligned}
a_m \int_0^1 \sin^2(m+\tfrac{1}{2})\pi x\, dx &= \frac{-1}{(m+\frac{1}{2})^2\pi^2\sqrt{2}} \int_0^1 \sin(m+\tfrac{1}{2})\pi x\, dx, \\
\frac{a_m}{2} &= \frac{-1}{(m+\frac{1}{2})^2\pi^2\sqrt{2}} \int_0^1 \sin(m+\tfrac{1}{2})\pi x\, dx \\
&= \frac{1}{(m+\frac{1}{2})^2\pi^2\sqrt{2}} \left[\frac{\cos(m+\frac{1}{2})\pi x}{(m+\frac{1}{2})\pi} \right]_0^1, \\
a_m &= -\frac{\sqrt{2}}{(m+\frac{1}{2})^3\pi^3}.
\end{aligned}$$

Substituting this result into the assumed expansion, and recalling that $B_n = \sqrt{2}$, gives as the solution

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

17.13 *By substituting $x = \exp t$, find the normalised eigenfunctions $y_n(x)$ and the eigenvalues λ_n of the operator $\mathcal{L}$ defined by*

$$\mathcal{L}y = x^2 y'' + 2xy' + \tfrac{1}{4}y, \qquad 1 \le x \le e,$$

with $y(1) = y(e) = 0$. Find, as a series $\sum a_n y_n(x)$, the solution of $\mathcal{L}y = x^{-1/2}$.

Putting $x = e^t$ and $y(x) = u(t)$ with $u(0) = u(1) = 0$,

$$\frac{dx}{dt} = e^t \quad \Rightarrow \quad \frac{d}{dx} = e^{-t}\frac{d}{dt}$$

and the eigenvalue equation becomes

$$e^{2t}e^{-t}\frac{d}{dt}\left(e^{-t}\frac{du}{dt}\right) + 2e^t e^{-t}\frac{du}{dt} + \frac{1}{4}u = \lambda u,$$
$$\frac{d^2u}{dt^2} - \frac{du}{dt} + 2\frac{du}{dt} + \left(\frac{1}{4} - \lambda\right) = 0.$$

The auxiliary equation to this constant-coefficient linear equation for u is

$$m^2 + m + (\tfrac{1}{4} - \lambda) = 0 \quad \Rightarrow \quad m = -\tfrac{1}{2} \pm \sqrt{\lambda},$$

leading to

$$u(t) = e^{-t/2}\left(Ae^{\sqrt{\lambda}\,t} + Be^{-\sqrt{\lambda}\,t}\right).$$

In view of the requirement that u vanishes at two different values of t (one of which is $t = 0$), we need $\lambda < 0$ and $u(t)$ to take the form

$$u(t) = Ae^{-t/2}\sin\sqrt{-\lambda}\,t \text{ with } \sqrt{-\lambda}\,1 = n\pi, \text{ i.e. } \lambda = -n^2\pi^2,$$

where n is an integer. Thus

$$u_n(t) = A_n e^{-t/2}\sin n\pi t \quad \text{or, in terms of } x, \quad y_n(x) = \frac{A_n}{\sqrt{x}}\sin(n\pi\ln x).$$

Normalisation requires that

$$1 = \int_1^e \frac{A_n^2}{x}\sin^2(n\pi\ln x)\,dx = \int_0^1 A_n^2\sin^2(n\pi t)\,dt = \tfrac{1}{2}A_n^2 \quad \Rightarrow \quad A_n = \sqrt{2}.$$

To solve

$$\mathcal{L}y = x^2y'' + 2xy' + \tfrac{1}{4}y = \frac{1}{\sqrt{x}},$$

we set $y(x) = \sum_{n=0}^{\infty} a_n y_n(x)$. Then the equation becomes

$$\mathcal{L}y = \sum_{n=0}^{\infty} a_n(-n^2\pi^2)y_n(x) = \sum_{n=0}^{\infty} -n^2\pi^2 a_n\frac{\sqrt{2}}{\sqrt{x}}\sin(n\pi\ln x) = \frac{1}{\sqrt{x}}.$$

Multiplying through by $y_m(x)$ and integrating, as with ordinary Fourier series,

$$\int_1^e \frac{2a_n}{x}\sin(n\pi\ln x)\,\sin(m\pi\ln x)\,dx = -\frac{1}{n^2\pi^2}\int_1^e \frac{\sqrt{2}\sin(m\pi\ln x)}{x}\,dx.$$

The LHS of this equation is the normalisation integral just considered and has the value $a_m\delta_{mn}$. Thus

$$\begin{aligned}
a_m &= -\frac{\sqrt{2}}{m^2\pi^2}\int_1^e \frac{\sin(m\pi \ln x)}{x}\,dx \\
&= -\frac{\sqrt{2}}{m^2\pi^2}\left[\frac{-\cos(m\pi\ln x)}{m\pi}\right]_1^e \\
&= -\frac{\sqrt{2}}{m^3\pi^3}[\,1-(-1)^m\,] \\
&= \begin{cases} -\dfrac{2\sqrt{2}}{m^3\pi^3} & \text{for } m \text{ odd,} \\ \quad 0 & \text{for } m \text{ even.} \end{cases}
\end{aligned}$$

The explicit solution is therefore

$$y(x) = -\frac{4}{\pi^3}\sum_{p=0}^{\infty}\frac{\sin[(2p+1)\pi\ln x]}{(2p+1)^3\sqrt{x}}.$$

17.15 *In the quantum mechanical study of the scattering of a particle by a potential, a Born-approximation solution can be obtained in terms of a function $y(\mathbf{r})$ that satisfies an equation of the form*

$$(-\nabla^2 - K^2)y(\mathbf{r}) = F(\mathbf{r}).$$

Assuming that $y_{\mathbf{k}}(\mathbf{r}) = (2\pi)^{-3/2}\exp(i\mathbf{k}\cdot\mathbf{r})$ is a suitably normalised eigenfunction of $-\nabla^2$ corresponding to eigenvalue k^2, find a suitable Green's function $G_K(\mathbf{r},\mathbf{r}')$. By taking the direction of the vector $\mathbf{r}-\mathbf{r}'$ as the polar axis for a $\mathbf{k}$-space integration, show that $G_K(\mathbf{r},\mathbf{r}')$ can be reduced to

$$\frac{1}{4\pi^2|\mathbf{r}-\mathbf{r}'|}\int_{-\infty}^{\infty}\frac{w\sin w}{w^2-w_0^2}\,dw,$$

where $w_0 = K|\mathbf{r}-\mathbf{r}'|$.

[This integral can be evaluated using contour integration and gives the Green's function explicitly as $(4\pi|\mathbf{r}-\mathbf{r}'|)^{-1}\exp(iK|\mathbf{r}-\mathbf{r}'|)$.]

Given that $y_{\mathbf{k}}(\mathbf{r}) = (2\pi)^{-3/2}\exp(i\mathbf{k}\cdot\mathbf{r})$ satisfies

$$-\nabla^2 y_{\mathbf{k}}(\mathbf{r}) = k^2 y_{\mathbf{k}}(\mathbf{r}),$$

it follows that

$$(-\nabla^2 - K^2)y_{\mathbf{k}}(\mathbf{r}) = (k^2-K^2)y_{\mathbf{k}}(\mathbf{r}).$$

Thus the same functions are suitable eigenfunctions for the extended operator, but with different eigenvalues.

Its Green's function is therefore (from the general expression for Green's functions in terms of eigenfunctions)

$$\begin{aligned}G_K(\mathbf{r},\mathbf{r}') &= \int \frac{1}{\lambda} y_{\mathbf{k}}(\mathbf{r}) y^*_{\mathbf{k}}(\mathbf{r}')\, d\mathbf{k} \\ &= \frac{1}{(2\pi)^3} \int \frac{\exp(i\mathbf{k}\cdot\mathbf{r})\, \exp(-i\mathbf{k}\cdot\mathbf{r}')}{k^2 - K^2}\, d\mathbf{k}.\end{aligned}$$

We carry out the three-dimensional integration in $\mathbf{k}$-space using the direction $\mathbf{r}-\mathbf{r}'$ as the polar axis (and denote $\mathbf{r}-\mathbf{r}'$ by $\mathbf{R}$). The azimuthal integral is immediate. The remaining two-dimensional integration is as follows:

$$\begin{aligned}G_K(\mathbf{r},\mathbf{r}') &= \frac{1}{(2\pi)^3} \int_0^\infty \int_0^\pi \frac{\exp(i\mathbf{k}\cdot\mathbf{R})}{k^2 - K^2}\, 2\pi k^2 \sin\theta_k\, d\theta_k\, dk \\ &= \frac{1}{(2\pi)^2} \int_0^\infty \int_0^\pi \frac{\exp(ikR\cos\theta_k)}{k^2 - K^2}\, k^2 \sin\theta_k\, d\theta_k\, dk \\ &= \frac{1}{(2\pi)^2} \int_0^\infty \frac{\exp(ikR) - \exp(-ikR)}{ikR(k^2 - K^2)}\, k^2\, dk \\ &= \frac{1}{2\pi^2 R} \int_0^\infty \frac{k \sin kR}{k^2 - K^2}\, dk \\ &= \frac{1}{2\pi^2 R} \int_0^\infty \frac{w \sin w}{w^2 - w_0^2}\, dw, \quad \text{where } w = kR \text{ and } w_0 = kR, \\ &= \frac{1}{4\pi^2 R} \int_{-\infty}^\infty \frac{w \sin w}{w^2 - w_0^2}\, dw.\end{aligned}$$

Here, the final line is justified by noting that the integrand is an even function of the integration variable w.

18

Special functions

18.1 *Use the explicit expressions*

$$
\begin{aligned}
Y_0^0 &= \sqrt{\tfrac{1}{4\pi}}, & Y_1^0 &= \sqrt{\tfrac{3}{4\pi}}\cos\theta, \\
Y_1^{\pm 1} &= \mp\sqrt{\tfrac{3}{8\pi}}\sin\theta\exp(\pm i\phi), & Y_2^0 &= \sqrt{\tfrac{5}{16\pi}}(3\cos^2\theta - 1), \\
Y_2^{\pm 1} &= \mp\sqrt{\tfrac{15}{8\pi}}\sin\theta\cos\theta\exp(\pm i\phi), & Y_2^{\pm 2} &= \sqrt{\tfrac{15}{32\pi}}\sin^2\theta\exp(\pm 2i\phi),
\end{aligned}
$$

to verify for $\ell = 0, 1, 2$ *that*

$$\sum_{m=-\ell}^{\ell} |Y_\ell^m(\theta, \phi)|^2 = \frac{2\ell + 1}{4\pi}$$

and so is independent of the values of θ *and* ϕ. *This is true for any* ℓ, *but a general proof is more involved. This result helps to reconcile intuition with the apparently arbitrary choice of polar axis in a general quantum mechanical system.*

We first note that, since every term is the square of a modulus, factors of the form $\exp(\pm mi\phi)$ never appear in the sums. For each value of ℓ, let us denote the sum by S_ℓ. For $\ell = 0$ and $\ell = 1$, we have

$$
\begin{aligned}
S_0 &= \sum_{m=0}^{0} |Y_0^m(\theta, \phi)|^2 = \frac{1}{4\pi}, \\
S_1 &= \sum_{m=-1}^{1} |Y_1^m(\theta, \phi)|^2 = \frac{3}{4\pi}\cos^2\theta + 2\,\frac{3}{8\pi}\sin^2\theta = \frac{3}{4\pi}.
\end{aligned}
$$

For $\ell = 2$, the summation is more complicated but reads

$$\begin{aligned}
S_2 &= \sum_{m=-2}^{2} |Y_2^m(\theta,\phi)|^2 \\
&= \frac{5}{16\pi}(3\cos^2\theta - 1)^2 + 2\,\frac{15}{8\pi}\sin^2\theta\cos^2\theta + 2\,\frac{15}{32\pi}\sin^4\theta \\
&= \frac{5}{16\pi}(9\cos^4\theta - 6\cos^2\theta + 1 + 12\sin^2\theta\cos^2\theta + 3\sin^4\theta) \\
&= \frac{5}{16\pi}[\,6\cos^4\theta - 6\cos^2\theta + 1 + 6\sin^2\theta\cos^2\theta + 3(\cos^2\theta + \sin^2\theta)^2\,] \\
&= \frac{5}{16\pi}[\,6\cos^2\theta(-\sin^2\theta) + 1 + 6\sin^2\theta\cos^2\theta + 3\,] = \frac{5}{4\pi}.
\end{aligned}$$

All three sums are independent of θ and ϕ, and are given by the general formula $(2\ell + 1)/4\pi$. It will, no doubt, be noted that $2\ell + 1$ is the number of terms in S_ℓ, i.e. the number of m values, and that 4π is the total solid angle subtended at the origin by all space.

18.3 *Use the generating function for the Legendre polynomials $P_n(x)$ to show that*

$$\int_0^1 P_{2n+1}(x)\,dx = (-1)^n \frac{(2n)!}{2^{2n+1}n!(n+1)!}$$

and that, except for the case $n = 0$,

$$\int_0^1 P_{2n}(x)\,dx = 0.$$

Denote $\int_0^1 P_n(x)\,dx$ by a_n. From the generating function for the Legendre polynomials, we have

$$\frac{1}{(1 - 2xh + h^2)^{1/2}} = \sum_{n=0}^{\infty} P_n(x)h^n.$$

Integrating this definition with respect to x gives

$$\begin{aligned}
\int_0^1 \frac{dx}{(1 - 2xh + h^2)^{1/2}} &= \sum_{n=0}^{\infty} \left(\int_0^1 P_n(x)\,dx \right) h^n, \\
\left[\frac{-(1 - 2xh + h^2)^{1/2}}{h} \right]_0^1 &= \sum_{n=0}^{\infty} a_n h^n, \\
\frac{1}{h}[\,(1 + h^2)^{1/2} - 1 + h\,] &= \sum_{n=0}^{\infty} a_n h^n.
\end{aligned}$$

Now expanding $(1+h^2)^{1/2}$ using the binomial theorem yields

$$\sum_{n=0}^{\infty} a_n h^n = \frac{1}{h}\left[1+\sum_{m=1}^{\infty} {}^{1/2}C_m h^{2m} - 1 + h\right] = 1 + \sum_{m=1}^{\infty} {}^{1/2}C_m h^{2m-1}.$$

Comparison of the coefficients of h^n on the two sides of the equation shows that all a_{2r} are zero except for $a_0 = 1$. For $n = 2r+1$ we need $2m-1 = n = 2r+1$, i.e. $m = r+1$, and the value of a_{2r+1} is ${}^{1/2}C_{r+1}$.

Now, the binomial coefficient ${}^{1/2}C_m$ can be written as

$$\begin{aligned} {}^{1/2}C_m &= \frac{\frac{1}{2}(\frac{1}{2}-1)(\frac{1}{2}-2)\cdots(\frac{1}{2}-m+1)}{m!}, \\ &= \frac{1(1-2)(1-4)\cdots(1-2m+2)}{2^m\, m!} \\ &= (-1)^{m-1}\frac{(1)(1)(3)\cdots(2m-3)}{2^m\, m!} \\ &= (-1)^{m-1}\frac{(2m-2)!}{2^m\, m!\, 2^{m-1}\,(m-1)!} \\ &= (-1)^{m-1}\frac{(2m-2)!}{2^{2m-1}\, m!\,(m-1)!}. \end{aligned}$$

Thus, setting $m = r+1$ gives the value of the integral a_{2r+1} as

$$a_{2r+1} = {}^{1/2}C_{r+1} = (-1)^r \frac{(2r)!}{2^{2r+1}\,(r+1)!\,r!},$$

as stated in the question.

18.5 *The Hermite polynomials $H_n(x)$ may be defined by*

$$\Phi(x,h) = \exp(2xh - h^2) = \sum_{n=0}^{\infty} \frac{1}{n!} H_n(x) h^n.$$

Show that

$$\frac{\partial^2 \Phi}{\partial x^2} - 2x\frac{\partial \Phi}{\partial x} + 2h\frac{\partial \Phi}{\partial h} = 0,$$

and hence that the $H_n(x)$ satisfy the Hermite equation,

$$y'' - 2xy' + 2ny = 0,$$

where n is an integer ≥ 0.

Use Φ to prove that

(a) $H_n'(x) = 2nH_{n-1}(x)$,
(b) $H_{n+1}(x) - 2xH_n(x) + 2nH_{n-1}(x) = 0.$

With

$$\Phi(x, h) = \exp(2xh - h^2) = \sum_{n=0}^{\infty} \frac{1}{n!} H_n(x) h^n,$$

we have

$$\frac{\partial \Phi}{\partial x} = 2h\Phi, \quad \frac{\partial \Phi}{\partial h} = (2x - 2h)\Phi, \quad \frac{\partial^2 \Phi}{\partial x^2} = 4h^2\Phi.$$

It then follows that

$$\frac{\partial^2 \Phi}{\partial x^2} - 2x\frac{\partial \Phi}{\partial x} + 2h\frac{\partial \Phi}{\partial h} = (4h^2 - 4hx + 4hx - 4h^2)\Phi = 0.$$

Substituting the series form into this result gives

$$\sum_{n=0}^{\infty} \left(\frac{1}{n!} H_n'' - \frac{2x}{n!} H_n' + \frac{2n}{n!} \right) h^n = 0,$$

$$\Rightarrow \quad H_n'' - 2xH_n' + 2nH_n = 0.$$

This is the equation satisfied by $H_n(x)$, as stated in the question.

(a) From the first relationship derived above, we have that

$$\frac{\partial \Phi}{\partial x} = 2h\Phi,$$

$$\sum_{n=0}^{\infty} \frac{1}{n!} H_n'(x) h^n = 2h \sum_{n=0}^{\infty} \frac{1}{n!} H_n(x) h^n,$$

$$\Rightarrow \quad \frac{1}{m!} H_m' = \frac{2}{(m-1)!} H_{m-1}, \text{ from the coefficients of } h^m.$$

Hence, $$H_n'(x) = 2nH_{n-1}(x).$$

(b) Differentiating result (a) and then applying it again yields

$$H_n'' = 2nH_{n-1}' = 2n\,2(n-1)H_{n-2}.$$

Using this in the differential equation satisfied by the H_n, we obtain

$$4n(n-1)H_{n-2} - 2x\,2nH_{n-1} + 2nH_n = 0.$$

This gives

$$H_{n+1}(x) - 2xH_n(x) + 2nH_{n-1}(x) = 0$$

after dividing through by $2n$ and changing $n \to n+1$.

18.7 *For the associated Laguerre polynomials, carry through the following exercises.*

(a) *Prove the Rodrigues' formula*

$$L_n^m(x) = \frac{e^x x^{-m}}{n!} \frac{d^n}{dx^n}(x^{n+m} e^{-x}),$$

taking the polynomials to be defined by

$$L_n^m(x) = \sum_{k=0}^{n} (-1)^k \frac{(n+m)!}{k!(n-k)!(k+m)!} x^k.$$

(b) *Prove the recurrence relations*

$$(n+1)L_{n+1}^m(x) = (2n+m+1-x)L_n^m(x) - (n+m)L_{n-1}^m(x),$$

$$x(L_n^m)'(x) = nL_n^m(x) - (n+m)L_{n-1}^m(x),$$

but this time taking the polynomial as defined by

$$L_n^m(x) = (-1)^m \frac{d^m}{dx^m} L_{n+m}(x)$$

or the generating function.

(a) It is most convenient to evaluate the nth derivative directly, using Leibnitz' theorem. This gives

$$\begin{aligned}
L_n^m(x) &= \frac{e^x x^{-m}}{n!} \sum_{r=0}^{n} \frac{n!}{r!(n-r)!} \frac{d^r}{dx^r}(x^{n+m}) \frac{d^{n-r}}{dx^{n-r}}(e^{-x}) \\
&= e^x x^{-m} \sum_{r=0}^{n} \frac{1}{r!(n-r)!} \frac{(n+m)!}{(n+m-r)!} x^{n+m-r} (-1)^{n-r} e^{-x} \\
&= \sum_{r=0}^{n} \frac{(-1)^{n-r}}{r!(n-r)!} \frac{(n+m)!}{(n+m-r)!} x^{n-r}.
\end{aligned}$$

Relabelling the summation using the new index $k = n - r$, we immediately obtain

$$L_n^m(x) = \sum_{k=0}^{n} (-1)^k \frac{(n+m)!}{k!(n-k)!(k+m)!} x^k,$$

which is as given in the question.

(b) The first recurrence relation can be proved using the generating function for the associated Laguerre functions:

$$G(x,h) = \frac{e^{-xh/(1-h)}}{(1-h)^{m+1}} = \sum_{n=0}^{\infty} L_n^m(x)h^n.$$

Differentiating the second equality with respect to h, we obtain

$$\frac{(m+1)(1-h)-x}{(1-h)^{m+3}}\, e^{-xh/(1-h)} = \sum nL_n^m h^{n-1}.$$

Using the generating function for a second time, we may rewrite this as

$$[(m+1)(1-h)-x]\sum L_n^m h^n = (1-h)^2 \sum nL_n^m h^{n-1}.$$

Equating the coefficients of h^n now yields

$$(m+1)L_n^m - (m+1)L_{n-1}^m - xL_n^m = (n+1)L_{n+1}^m - 2nL_n^m + (n-1)L_{n-1}^m,$$

which can be rearranged and simplified to give the first recurrence relation.

The second result is most easily proved by differentiating one of the standard recurrence relations satisfied by the ordinary Laguerre polynomials, but with n replaced by $n+m$. This standard equality reads

$$xL'_{n+m}(x) = (n+m)L_{n+m}(x) - (n+m)L_{n-1+m}(x).$$

We convert this into an equation for the associated polynomials,

$$L_n^m(x) = (-1)^m \frac{d^m}{dx^m} L_{n+m}(x),$$

by differentiating it m times with respect to x and multiplying through by $(-1)^m$. The result is

$$x(L_n^m)' + mL_n^m = (n+m)L_n^m - (n+m)L_{n-1}^m,$$

which immediately simplifies to give the second recurrence relation satisfied by the associated Laguerre polynomials.

18.9 *By initially writing $y(x)$ as $x^{1/2}f(x)$ and then making subsequent changes of variable, reduce Stokes' equation,*

$$\frac{d^2y}{dx^2} + \lambda xy = 0,$$

to Bessel's equation. Hence show that a solution that is finite at $x=0$ is a multiple of $x^{1/2}J_{1/3}(\frac{2}{3}\sqrt{\lambda x^3})$.

With $y(x) = x^{1/2}f(x)$,

$$y' = \frac{f}{2x^{1/2}} + x^{1/2}f' \text{ and } y'' = -\frac{f}{4x^{3/2}} + \frac{f'}{x^{1/2}} + x^{1/2}f''$$

and the equation becomes

$$-\frac{f}{4x^{3/2}}+\frac{f'}{x^{1/2}}+x^{1/2}f''+\lambda x^{3/2}f=0,$$
$$x^2f''+xf'+(\lambda x^3-\tfrac{1}{4})f=0.$$

Now, guided by the known form of Bessel's equation, change the independent variable to $u=x^{3/2}$ with $f(x)=g(u)$ and

$$\frac{du}{dx}=\frac{3}{2}x^{1/2}\quad\Rightarrow\quad\frac{d}{dx}=\frac{3}{2}u^{1/3}\frac{d}{du}.$$

This gives

$$u^{4/3}\frac{3}{2}u^{1/3}\frac{d}{du}\left(\frac{3}{2}u^{1/3}\frac{dg}{du}\right)+u^{2/3}\frac{3}{2}u^{1/3}\frac{dg}{du}+\left(\lambda u^2-\frac{1}{4}\right)g=0,$$
$$\frac{3}{2}u^{5/3}\left(\frac{3}{2}u^{1/3}\frac{d^2g}{du^2}+\frac{1}{2}u^{-2/3}\frac{dg}{du}\right)+\frac{3}{2}u\frac{dg}{du}+\left(\lambda u^2-\frac{1}{4}\right)g=0,$$
$$\frac{9}{4}u^2\frac{d^2g}{du^2}+\frac{9}{4}u\frac{dg}{du}+\left(\lambda u^2-\frac{1}{4}\right)g=0,$$
$$u^2\frac{d^2g}{du^2}+u\frac{dg}{du}+\left(\frac{4}{9}\lambda u^2-\frac{1}{9}\right)g=0.$$

This is close to Bessel's equation but still needs a scaling of the variables. So, set $\frac{2}{3}\sqrt{\lambda}u\equiv\mu u=v$ and $g(u)=h(v)$, obtaining

$$\frac{v^2}{\mu^2}\,\mu^2\frac{d^2h}{dv^2}+\frac{v}{\mu}\,\mu\frac{dh}{dv}+\left(v^2-\frac{1}{9}\right)h=0.$$

This *is* Bessel's equation and has a general solution

$$h(v)=c_1J_{1/3}(v)+c_2J_{-1/3}(v),$$
$$\Rightarrow\quad g(u)=c_1J_{1/3}(\tfrac{2\sqrt{\lambda}}{3}u)+c_2J_{-1/3}(\tfrac{2\sqrt{\lambda}}{3}u),$$
$$\Rightarrow\quad f(x)=c_1J_{1/3}(\tfrac{2\sqrt{\lambda}}{3}x^{3/2})+c_2J_{-1/3}(\tfrac{2\sqrt{\lambda}}{3}x^{3/2}).$$

For a solution that is finite at $x=0$, only the Bessel function with a positive subscript can be accepted. Therefore the required solution is

$$y(x)=c_1x^{1/2}J_{1/3}(\tfrac{2\sqrt{\lambda}}{3}x^{3/2}).$$

> ***18.11*** *Identify the series for the following hypergeometric functions, writing them in terms of better-known functions.*
>
> (a) $F(a, b, b; z)$,
> (b) $F(1, 1, 2; -x)$,
> (c) $F(\frac{1}{2}, 1, \frac{3}{2}; -x^2)$,
> (d) $F(\frac{1}{2}, \frac{1}{2}, \frac{3}{2}; x^2)$,
> (e) $F(-a, a, \frac{1}{2}; \sin^2 x)$*; this is a much more difficult exercise.*

The hypergeometric equation is

$$z(1-z)y'' + [\,c - (a+b+1)z\,]\,y' - aby = 0.$$

The $(n+1)$th term of its series solution, the hypergeometric function $F(a, b, c; z)$, is given by

$$\frac{a(a+1)\cdots(a+n-1)\,b(b+1)\cdots(b+n-1)}{c(c+1)\cdots(c+n-1)}\,\frac{z^n}{n!}$$

for $n \geq 1$ and unity for $n = 0$.

(a) $F(a, b, b; z)$. In each term the equal factors arising from the second and third parameters cancel, as one is in the numerator and the other in the denominator. Thus,

$$\begin{aligned} F(a, b, b; z) &= 1 + az + \frac{a(a+1)}{2!}\,z^2 + \frac{a(a+1)(a+2)}{3!}\,z^3 + \cdots \\ &= (1-z)^{-a}. \end{aligned}$$

(b) $F(1, 1, 2; -x)$. The $n+1$th term is

$$\frac{(n!)\,(n!)}{(n+1)!\,(n!)}\,(-x)^n = \frac{(-1)^n x^n}{n+1}$$

making the series

$$\sum_{n=0}^{\infty} \frac{(-1)^n x^n}{n+1} = 1 - \frac{x}{2} + \frac{x^2}{3} - \frac{x^3}{4} + \cdots = \frac{1}{x}\ln(1+x).$$

(c) $F(\frac{1}{2}, 1, \frac{3}{2}; -x^2)$. Directly from the series:

$$\begin{aligned} F(\tfrac{1}{2}, 1, \tfrac{3}{2}; -x^2) &= 1 + \frac{(\frac{1}{2})\,(1)}{1!\,(\frac{3}{2})}\,(-x^2) + \frac{(\frac{1}{2})\,(\frac{3}{2})\,(1)\,(2)}{2!\,(\frac{3}{2})\,(\frac{5}{2})}\,(-x^2)^2 + \cdots \\ &= 1 - \frac{x^2}{3} + \frac{x^4}{5} - \frac{x^6}{7} + \cdots. \end{aligned}$$

The coefficients are those of $\tan^{-1} x$, though the powers of x are all too small by one. Thus $F(\frac{1}{2}, 1, \frac{3}{2}; -x^2) = x^{-1}\tan^{-1} x$.

(d) $F(\frac{1}{2},\frac{1}{2},\frac{3}{2};x^2)$. Again, directly from the series:

$$F(\tfrac{1}{2},\tfrac{1}{2},\tfrac{3}{2};x^2) = 1 + \frac{(\frac{1}{2})^2}{1!\,(\frac{3}{2})}(x^2) + \frac{(\frac{1}{2})^2(\frac{3}{2})^2}{2!\,(\frac{3}{2})(\frac{5}{2})}(x^2)^2 + \cdots$$

$$= 1 + \tfrac{1}{6}x^2 + \tfrac{3}{40}5x^4 + \frac{15}{336}x^6 + \cdots.$$

From the larger standard tables of Maclaurin series it can be seen that, although the successive coefficients are those of $\sin^{-1}x$, the powers of x are all too small by one. Thus $F(\frac{1}{2},\frac{1}{2},\frac{3}{2};-x^2) = x^{-1}\sin^{-1}x$.

(e) $F(-a,a,\frac{1}{2};\sin^2 x)$. Since we will obtain a series involving terms such as $\sin^{2m}x$, the series may be difficult to identify. The series is

$$1 + \frac{(-a)(a)}{(\frac{1}{2})}\sin^2 x + \frac{(-a)(-a+1)(a)(a+1)}{2!\,(\frac{1}{2})(\frac{3}{2})}\sin^4 x + \cdots. \qquad (*)$$

Clearly, this contains only even powers of x, though just the first two terms alone constitute an infinite power series in x. However, a term containing x^{2m} can only arise from the first $m+1$ terms of $(*)$ and a few trials may be helpful.

If $F(-a,a,\frac{1}{2};\sin^2 x) = \sum_{n=0}^{\infty} b_n x^{2n}$, then $b_0 = 1$ and $b_1 = -2a^2$ since the corresponding powers of x can only arise from the first and second terms of $(*)$, respectively. The coefficient of x^4 is determined by the second and third terms of $(*)$ and is given by

$$b_2 = -2a^2\left(-\frac{2}{3!}\right) + \frac{2a^2(a^2-1)}{3}(1) = \frac{2a^4}{3}.$$

The coefficient of x^6, namely b_3, has contributions from the second, third and fourth terms of $(*)$ and is given by

$$-2a^2\left[\left(\frac{1}{3!}\right)^2 + \frac{2}{5!}\right] + \frac{2a^2(a^2-1)}{3}\left(\frac{-4}{3!}\right) + \frac{4a^2(a^2-1)(4-a^2)}{45}(1)$$

$$= -2a^2\left(\frac{20+12}{720}\right) - \frac{8a^4}{18} + \frac{8a^2}{18} + \frac{4}{45}(-4a^2+5a^4-a^6)$$

$$= \left(-\frac{64}{720} + \frac{8}{18} - \frac{16}{45}\right)a^2 + \left(-\frac{8}{18} + \frac{20}{45}\right)a^4 + \left(-\frac{4}{45}\right)a^6$$

$$= -\frac{4}{45}a^6.$$

Thus, in powers up to x^6,

$$F(-a,a,\tfrac{1}{2};\sin^2 x) = 1 - 2a^2x^2 + \tfrac{2}{3}a^4x^4 - \tfrac{4}{45}a^6x^6$$

$$= 1 - \frac{(2ax)^2}{2!} + \frac{(2ax)^4}{4!} - \frac{(2ax)^6}{6!}.$$

Though not totally conclusive, this sequence of coefficients strongly suggests that $F(-a, a, \frac{1}{2}; \sin^2 x) = \cos 2ax$. Note that a does not need to be an integer.

This tentative conclusion can be tested by transforming the original hypergeometric equation as follows. With $z = \sin^2 x$, we have that $dz/dx = 2\sin x \cos x = \sin 2x$, implying that $d/dz = (\sin 2x)^{-1}\, d/dx$. The equation becomes

$$\sin^2 x(1-\sin^2 x)\frac{1}{\sin 2x}\frac{d}{dx}\left(\frac{1}{\sin 2x}\frac{dy}{dx}\right) + \left[\frac{1}{2} - (-a+a+1)\sin^2 x\right]\frac{1}{\sin 2x}\frac{dy}{dx} + a^2 y = 0.$$

This can be simplified as follows:

$$\frac{1}{4}\sin 2x\left(\frac{1}{\sin 2x}\frac{d^2y}{dx^2} - \frac{2\cos 2x}{\sin^2 2x}\frac{dy}{dx}\right) + \frac{1-2\sin^2 x}{2\sin 2x}\frac{dy}{dx} + a^2 y = 0,$$

$$\frac{1}{4}\frac{d^2y}{dx^2} - \frac{\cos 2x}{2\sin 2x}\frac{dy}{dx} + \frac{\cos 2x}{2\sin 2x}\frac{dy}{dx} + a^2 y = 0,$$

$$\frac{d^2y}{dx^2} + 4a^2 y = 0.$$

For a solution with $y(0) = 1$, this implies that $y(x) = \cos 2ax$, thus confirming our provisional conclusion.

18.13 *Find a change of variable that will allow the integral*

$$I = \int_1^\infty \frac{\sqrt{u-1}}{(u+1)^2}\,du$$

to be expressed in terms of the beta function and so evaluate it.

The beta function is normally expressed in terms of an integral, over the range 0 to 1, of an integrand of the form $v^m(1-v)^n$, with $m, n > -1$. We therefore need a change of variable $u = f(x)$ such that $u+1$ is an inverse power of x; this being so, we also need $f(0) = \infty$ and $f(1) = 1$.

Consider

$$u + 1 = \frac{A}{x}, \quad \text{i.e.} \quad f(x) = \frac{A}{x} - 1.$$

This satisfies the first two requirements, and also satisfies the third one if we choose $A = 2$.

So, substitute $u = \dfrac{2}{x} - 1 = \dfrac{2-x}{x}$, with $u - 1 = \dfrac{2(1-x)}{x}$ and $du = -\dfrac{2}{x^2}$. The integral then becomes

$$\begin{aligned}
I &= \int_1^0 \frac{2^{1/2}(1-x)^{1/2}\,x^2\,(-2)}{x^{1/2}\,2^2\,x^2}\,dx \\
&= \frac{1}{\sqrt{2}}\int_0^1 (1-x)^{1/2}x^{-1/2}\,dx \\
&= \frac{1}{\sqrt{2}}\,B(\tfrac{1}{2},\tfrac{3}{2}) = \frac{\Gamma(\frac{1}{2})\,\Gamma(\frac{3}{2})}{\sqrt{2}\,\Gamma(\frac{1}{2}+\frac{3}{2})} \\
&= \frac{\sqrt{\pi}\,\frac{1}{2}\sqrt{\pi}}{\sqrt{2}\,1} = \frac{\pi}{2\sqrt{2}}.
\end{aligned}$$

18.15 *The complex function $z!$ is defined by*

$$z! = \int_0^\infty u^z e^{-u}\,du \qquad \text{for } \operatorname{Re} z > -1.$$

For $\operatorname{Re} z \le -1$ it is defined by

$$z! = \frac{(z+n)!}{(z+n)(z+n-1)\cdots(z+1)},$$

where n is any (positive) integer $> -\operatorname{Re} z$. Being the ratio of two polynomials, $z!$ is analytic everywhere in the finite complex plane except at the poles that occur when z is a negative integer.

(a) *Show that the definition of $z!$ for $\operatorname{Re} z \le -1$ is independent of the value of n chosen.*
(b) *Prove that the residue of $z!$ at the pole $z = -m$, where m is an integer > 0, is $(-1)^{m-1}/(m-1)!$.*

(a) Let m and n be two choices of integer with $m > n > -\operatorname{Re} z$. Denote the corresponding definitions of $z!$ by $(z!)_m$ and $(z!)_n$ and consider the ratio of these two functions:

$$\begin{aligned}
\frac{(z!)_m}{(z!)_n} &= \frac{(z+m)!}{(z+m)(z+m-1)\cdots(z+1)}\,\frac{(z+n)(z+n-1)\cdots(z+1)}{(z+n)!} \\
&= \frac{(z+m)!}{(z+m)(z+m-1)\cdots(z+n+1)\times(z+n)!} \\
&= \frac{(z+m)!}{(z+m)!} = 1.
\end{aligned}$$

Thus the two functions are identical for all z, i.e the definition of $z!$ is independent of the choice of n, provided that $n > -\operatorname{Re} z$.

(b) From the given definition of $z!$ it is clear that its pole at $z = -m$ is a simple one. The residue R at the pole is therefore given by

$$\begin{aligned}
R &= \lim_{z \to -m} (z+m) z! \\
&= \lim_{z \to -m} \frac{(z+m)\,(z+n)!}{(z+n)(z+n-1)\cdots(z+1)} \qquad \text{(integer } n \text{ is chosen} > m) \\
&= \lim_{z \to -m} \frac{(z+n)!}{(z+n)(z+n-1)\cdots(z+m+1)(z+m-1)\cdots(z+1)} \\
&= \frac{(-m+n)!}{(-m+n)\cdots(-m+m+1)(-m+m-1)\cdots(-m+1)} \\
&= \frac{1}{[-1]\,[-2]\cdots[-(m-1)]} \\
&= (-1)^{m-1} \frac{1}{(m-1)!},
\end{aligned}$$

as stated in the question.

18.17 *The integral*

$$I = \int_{-\infty}^{\infty} \frac{e^{-k^2}}{k^2 + a^2}\, dk, \qquad (*)$$

in which $a > 0$, occurs in some statistical mechanics problems. By first considering the integral

$$J = \int_0^{\infty} e^{iu(k+ia)}\, du$$

and a suitable variation of it, show that $I = (\pi/a)\exp(a^2)\,\mathrm{erfc}(a)$, where $\mathrm{erfc}(x)$ is the complementary error function.

The fact that $a > 0$ will ensure that the improper integral J is well defined. It is

$$J = \int_0^{\infty} e^{iu(k+ia)}\, du = \left[\frac{e^{iu(k+ia)}}{i(k+ia)}\right]_0^{\infty} = \frac{i}{k+ia}.$$

We note that this result contains one of the factors that would appear as a denominator in one term of a partial fraction expansion of the integrand in (*). Another term would contain a factor $(k - ia)^{-1}$, and this can be generated by

$$J' = \int_0^{\infty} e^{-iu(k-ia)}\, du = \left[\frac{e^{-iu(k-ia)}}{-i(k-ia)}\right]_0^{\infty} = \frac{-i}{k-ia}.$$

Now, actually expressing the integrand in partial fractions, using the integral expressions J and J' for the factors, and then reversing the order of integration gives

$$\begin{aligned}
I &= \frac{1}{2a}\int_{-\infty}^{\infty}\left(\frac{ie^{-k^2}}{k+ia} - \frac{ie^{-k^2}}{k-ia}\right)dk \\
&= \frac{1}{2a}\int_{-\infty}^{\infty} e^{-k^2}\,dk \int_0^{\infty} e^{iu(k+ia)}\,du + \frac{1}{2a}\int_{-\infty}^{\infty} e^{-k^2}\,dk \int_0^{\infty} e^{-iu(k-ia)}\,du, \\
\Rightarrow \quad 2aI &= \int_0^{\infty} du \int_{-\infty}^{\infty} e^{-k^2+iuk-ua}\,dk + \int_0^{\infty} du \int_{-\infty}^{\infty} e^{-k^2-iuk-ua}\,dk \\
&= \int_0^{\infty} du \int_{-\infty}^{\infty} e^{-(k-iu/2)^2-u^2/4-ua}\,dk \\
&\qquad\qquad + \int_0^{\infty} du \int_{-\infty}^{\infty} e^{-(k+iu/2)^2-u^2/4-ua}\,dk \\
&= 2\sqrt{\pi}\int_0^{\infty} e^{-u^2/4-ua}\,du,
\end{aligned}$$

using the standard Gaussian result. We now complete the square in the exponent and set $2v = u + 2$, obtaining

$$\begin{aligned}
2aI &= 2\sqrt{\pi}\int_0^{\infty} e^{-(u+2a)^2/4+a^2}\,du, \\
&= 2\sqrt{\pi}\int_a^{\infty} e^{-v^2}\,e^{a^2}\,2dv.
\end{aligned}$$

From this it follows that

$$I = \frac{\sqrt{\pi}}{a}\,2e^{a^2}\,\frac{\sqrt{\pi}}{2}\,\mathrm{erfc}(a) = \frac{\pi}{a}\,e^{a^2}\,\mathrm{erfc}(a),$$

as stated in the question.

18.19 *For the functions $M(a, c; z)$ that are the solutions of the confluent hypergeometric equation:*

(a) *use their series representation to prove that*

$$c\,\frac{d}{dz}M(a, c; z) = a\,M(a+1, c+1; z);$$

(b) *use an integral representation to prove that*

$$M(a, c; z) = e^z M(c-a, c; -z).$$

(a) Directly differentiating the explicit series term by term gives

$$\begin{aligned}\frac{d}{dz}M(a,c;z) &= \frac{d}{dz}\left[1+\frac{a}{c}z+\frac{a(a+1)}{2!\,c(c+1)}z^2+\cdots\right]\\ &= \frac{a}{c}+\frac{2a(a+1)}{2!\,c(c+1)}z+\frac{3a(a+1)(a+2)}{3!\,c(c+1)(c+2)}z^2+\cdots\\ &= \frac{a}{c}\left[1+\frac{a+1}{c+1}z+\frac{(a+1)(a+2)}{2!\,(c+1)(c+2)}z^2+\cdots\right]\\ &= \frac{a}{c}\,M(a+1,c+1;z).\end{aligned}$$

The quoted result follows immediately.

(b) This will be achieved most simply if we choose a representation in which the parameters can be rearranged without having to perform any actual integration. We therefore take the representation

$$M(a,c;z)=\frac{\Gamma(c)}{\Gamma(c-a)\,\Gamma(a)}\int_0^1 e^{zt}\,t^{a-1}\,(1-t)^{c-a-1}\,dt$$

and change the variable of integration to $s=1-t$ whilst regrouping the parameters (without changing their values, of course). This gives

$$\begin{aligned}&M(a,c;z)\\ &= \frac{\Gamma(c)}{\Gamma(a)\,\Gamma(c-a)}\int_1^0 e^z\,e^{-zs}\,(1-s)^{a-1}\,s^{c-a-1}\,(-ds)\\ &= \frac{\Gamma(c)\,e^z}{\Gamma[\,c-(c-a)\,]\,\Gamma[\,(c-a)\,]}\int_0^1 e^{-zs}\,(1-s)^{c-(c-a)-1}\,s^{(c-a)-1}\,ds\\ &= e^z\,M(c-a,c,-z),\end{aligned}$$

thus establishing the identity, in which $a\to c-a$ and $z\to -z$ whilst c remains unchanged.

18.21 *Find the differential equation satisfied by the function $y(x)$ defined by*

$$y(x)=Ax^{-n}\int_0^x e^{-t}t^{n-1}\,dt\equiv Ax^{-n}\gamma(n,x),$$

and, by comparing it with the confluent hypergeometric function, express y as a multiple of the solution $M(a,c;z)$ of that equation. Determine the value of A that makes y equal to M.

As the comparison is to be made with the hypergeometric equation, which is a second-order differential equation, we must calculate the first two derivatives of

$y(x)$. Further, as it is a homogeneous equation, we may omit the multiplicative constant A for the time being:

$$\begin{aligned}
y(x) &= x^{-n}\gamma(n,x),\\
y'(x) &= -nx^{-n-1}\gamma(n,x) + x^{-n}e^{-x}x^{n-1}\\
&= -nx^{-1}y + x^{-1}e^{-x},\\
y''(x) &= nx^{-2}y - nx^{-1}y' - x^{-2}e^{-x} - x^{-1}e^{-x}\\
&= nx^{-2}y - nx^{-1}y' - (x^{-1}+1)(y' + nx^{-1}y),\\
x^2y'' &= (-nx - x - x^2)y' + (n - n - nx)y.
\end{aligned}$$

The second line uses the standard result for differentiating an indefinite integral with respect to its upper limit. In the fifth line we substituted for $x^{-1}e^{-x}$ from the expression obtained for $y'(x)$ in the third line. Thus the equation to be compared with the confluent hypergeometric equation is

$$xy'' + (n+1+x)y' + ny = 0.$$

This has to be compared with

$$zw'' + (c-z)w' - aw = 0.$$

Now xy'' and xy' terms have the same signs (both positive), whereas the zw'' and zw' terms have opposite signs. To deal with this, we must set $z = -x$ in the confluent hypergeometric equation; renaming $w(z) = y(x)$ gives $w' = -y'$ and $w'' = y''$. The equation then becomes (after an additional overall sign change)

$$xy'' + (c+x)y' + ay = 0.$$

The obvious assignments, to go with $z \to -x$, are now $a \to n$ and $c \to n+1$. We therefore conclude that $y(x)$ is a multiple of $M(n, n+1; -x)$.

To determine the constant A in the given form of $y(x)$ we expand both its definition and $M(n, n+1; -x)$ in powers of x. Strictly, only the first term is necessary, but the second acts as a check.

From the hypergeometric series,

$$M(n, n+1; -x) = 1 + \frac{n(-x)}{n+1} + \cdots.$$

From the definition of $y(x)$,

$$\begin{aligned}
y(x) &= Ax^{-n}\int_0^x e^{-t}t^{n-1}\,dt \\
&= Ax^{-n}\int_0^x \left[t^{n-1} - t(t^{n-1}) + \frac{t^2}{2!}(t^{n-1}) + \cdots\right] \\
&= Ax^{-n}\left[\frac{t^n}{n} - \frac{t^{n+1}}{n+1} + \cdots\right]_0^x \\
&= Ax^{-n}\left[\frac{x^n}{n} - \frac{x^{n+1}}{n+1} + \cdots\right] \\
&= \frac{A}{n} - \frac{Ax}{n+1} + \cdots .
\end{aligned}$$

This reproduces the first two terms of $M(n, n+1; -x)$ if $A = n$, yielding, finally, that

$$y(x) = nx^{-n}\gamma(n, x) = M(n, n+1; -x).$$

18.23 *Prove two of the properties of the incomplete gamma function $P(a, x^2)$ as follows.*

(a) *By considering its form for a suitable value of a, show that the error function can be expressed as a particular case of the incomplete gamma function.*
(b) *The Fresnel integrals, of importance in the study of the diffraction of light, are given by*

$$C(x) = \int_0^x \cos\left(\frac{\pi}{2}t^2\right)\,dt, \qquad S(x) = \int_0^x \sin\left(\frac{\pi}{2}t^2\right)\,dt.$$

Show that they can be expressed in terms of the error function by

$$C(x) + iS(x) = A\,\mathrm{erf}\left[\frac{\sqrt{\pi}}{2}(1-i)x\right],$$

where A is a (complex) constant, which you should determine. Hence express $C(x) + iS(x)$ in terms of the incomplete gamma function.

(a) From the definition of the incomplete gamma function, we have

$$P(a, x^2) = \frac{1}{\Gamma(a)}\int_0^{x^2} e^{-t}\,t^{a-1}\,dt.$$

Guided by the x^2 in the upper limit, we now change the integration variable to

$y = +\sqrt{t}$, with $2y\,dy = dt$, and obtain

$$P(a, x^2) = \frac{1}{\Gamma(a)} \int_0^x e^{-y^2}\, y^{2(a-1)}\, 2y\, dy.$$

To make the RHS into an error function we need to remove the y-term; to do this we choose a such that $2(a-1)+1 = 0$, i.e. $a = \frac{1}{2}$. With this choice, $\Gamma(a) = \sqrt{\pi}$ and

$$P\left(\tfrac{1}{2}, x^2\right) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-y^2}\, dy,$$

i.e. a correctly normalised error function.

(b) Consider the given expression:

$$z = A \operatorname{erf}\left[\frac{\sqrt{\pi}}{2}(1-i)x\right] = \frac{2A}{\sqrt{\pi}} \int_0^{\sqrt{\pi}(1-i)x/2} e^{-u^2}\, du.$$

Changing the variable of integration to s, given by $u = \frac{1}{2}\sqrt{\pi}(1-i)s$, and recalling that $(1-i)^2 = -2i$, we obtain

$$\begin{aligned}
z &= \frac{2A}{\sqrt{\pi}} \int_0^x e^{-s^2\pi(-2i)/4}\, \frac{\sqrt{\pi}}{2}(1-i)\, ds \\
&= A(1-i) \int_0^x e^{i\pi s^2/2}\, ds \\
&= A(1-i) \int_0^x \left[\cos\left(\frac{\pi s^2}{2}\right) + i \sin\left(\frac{\pi s^2}{2}\right)\right] ds \\
&= A(1-i)\,[\,C(x) + iS(x)\,].
\end{aligned}$$

For the correct normalisation we need $A(1-i) = 1$, implying that $A = (1+i)/2$.

Now, from part (a), the error function can be expressed in terms of the incomplete gamma function $P(a, x)$ by

$$\operatorname{erf}(x) = P(\tfrac{1}{2}, x^2).$$

Here the argument of the error function is $\frac{1}{2}\sqrt{\pi}(1-i)x$, whose square is $-\frac{1}{2}\pi i x^2$, and so

$$C(x) + iS(x) = \frac{1+i}{2}\, P\left(\frac{1}{2}, -\frac{i\pi}{2}\, x^2\right).$$

19

Quantum operators

19.1 *Show that the commutator of two operators that correspond to two physical observables cannot itself correspond to another physical observable.*

Let the two operators be A and B, both of which must be Hermitian since they correspond to physical variables, and consider the Hermitian conjugate of their commutator:

$$[A, B]^\dagger = (AB)^\dagger - (BA)^\dagger = B^\dagger A^\dagger - A^\dagger B^\dagger = BA - AB = -[A, B].$$

Thus, the commutator is anti-Hermitian or zero and therefore cannot represent a non-trivial physical variable (as its eigenvalues are imaginary).

19.3 *In quantum mechanics, the time dependence of the state function $|\psi\rangle$ of a system is given, as a further postulate, by the equation*

$$i\hbar \frac{\partial}{\partial t}|\psi\rangle = H|\psi\rangle,$$

where H is the Hamiltonian of the system. Use this to find the time dependence of the expectation value $\langle A\rangle$ of an operator A that itself has no explicit time dependence. Hence show that operators that commute with the Hamiltonian correspond to the classical 'constants of the motion'.

For a particle of mass m moving in a one-dimensional potential $V(x)$, prove Ehrenfest's theorem:

$$\frac{d\langle p_x\rangle}{dt} = -\left\langle \frac{dV}{dx} \right\rangle \qquad \text{and} \qquad \frac{d\langle x\rangle}{dt} = \frac{\langle p_x\rangle}{m}.$$

The expectation value of A at any time is $\langle\psi(x,t)|A|\psi(x,t)\rangle$, where we have explicitly indicated that the state function varies with time. Now

$$\frac{d}{dt}\langle\psi\,|\,A\,|\,\psi\rangle = \left(\frac{\partial}{\partial t}\langle\psi|\right)\,A\,|\,\psi\rangle + \langle\psi\,|\,\frac{\partial A}{\partial t}\,|\,\psi\rangle + \langle\psi\,|\,A\left(\frac{\partial}{\partial t}|\,\psi\rangle\right).$$

Since A has no explicit time dependence, $\partial A/\partial t = 0$ and the second term drops out.

The given (quantum) equation of motion and its Hermitian conjugate are

$$i\hbar\frac{\partial}{\partial t}|\psi\rangle = H|\psi\rangle \quad \text{and} \quad \frac{\partial}{\partial t}\langle\psi\,| = -\frac{1}{i\hbar}\langle\psi\,|H^{\dagger} = -\frac{1}{i\hbar}\langle\psi\,|H,$$

since H is Hermitian. Thus,

$$\begin{aligned}
\frac{d}{dt}\langle\psi\,|\,A\,|\,\psi\rangle &= -\frac{1}{i\hbar}\langle\psi\,|\,HA\,|\,\psi\rangle + \frac{1}{i\hbar}\langle\psi\,|\,AH\,|\,\psi\rangle \\
&= -\frac{1}{i\hbar}\langle\psi\,|\,[\,H,A\,]\,|\,\psi\rangle \\
&= \frac{i}{\hbar}\langle\psi\,|\,[\,H,A\,]\,|\,\psi\rangle.
\end{aligned}$$

This shows that the rate of change of the expectation value of A is proportional to the expectation value of the commutator of A and the Hamiltonian. If A and H commute, the RHS is zero, the expectation value of A has a zero rate of change, and $\langle\psi\,|\,A\,|\,\psi\rangle$ is a constant of the motion.

For the particle moving in the one-dimensional potential $V(x)$,

$$H = \frac{p_x^2}{2m} + V(x).$$

(i) For $\langle p_x\rangle$,

$$\begin{aligned}
[\,H,p_x\,]\,|\,\psi\rangle &= [\,V,p_x\,]\,|\,\psi\rangle, \text{ since } p_x \text{ clearly commutes with } p_x^2, \\
&= -i\hbar V\frac{\partial}{\partial x}|\,\psi\rangle + i\hbar\frac{\partial}{\partial x}\left(V|\,\psi\rangle\right) \\
&= -i\hbar V\frac{\partial}{\partial x}|\,\psi\rangle + i\hbar V\frac{\partial}{\partial x}|\,\psi\rangle + i\hbar\,\frac{\partial V}{\partial x}\,|\,\psi\rangle \\
&= i\hbar\,\frac{\partial V}{\partial x}\,|\,\psi\rangle,
\end{aligned}$$

implying that

$$\begin{aligned}
\frac{d}{dt}\langle p_x\rangle = \frac{d}{dt}\langle\psi\,|\,p_x\,|\,\psi\rangle &= \frac{i}{\hbar}\langle\psi\,|\,[\,H,p_x\,]\,|\,\psi\rangle \\
&= \frac{i}{\hbar}\langle\psi\,|\,i\hbar\frac{\partial V}{\partial x}\,|\,\psi\rangle = -\left\langle\frac{\partial V}{\partial x}\right\rangle.
\end{aligned}$$

(ii) For $\langle x\rangle$ we will need the general commutator property $[\,AB,C\,] = A\,[\,B,C\,] +$

$[A, C]B$ to evaluate $[p_x^2, x]$:

$$[H, x] | \psi \rangle = \frac{1}{2m} [p_x^2, x] | \psi \rangle, \text{ since } x \text{ clearly commutes with } V(x),$$
$$= \frac{1}{2m} \{ p_x [p_x, x] | \psi \rangle + [p_x, x] p_x | \psi \rangle \}, \quad \text{as above},$$
$$= \frac{1}{2m} \{ -i\hbar p_x | \psi \rangle - i\hbar p_x | \psi \rangle \} = -\frac{i\hbar}{m} p_x | \psi \rangle,$$

implying that

$$\frac{d}{dt} \langle x \rangle = \frac{d}{dt} \langle \psi | x | \psi \rangle = \frac{i}{\hbar} \langle \psi | [H, x] | \psi \rangle$$
$$= \frac{i}{\hbar} \langle \psi | -\frac{i\hbar}{m} p_x | \psi \rangle = \frac{1}{m} \langle \psi | p_x | \psi \rangle = \frac{1}{m} \langle p_x \rangle.$$

Ehrenfest's theorem should be compared to the classical statements 'momentum equals mass times velocity', 'the force is given by minus the gradient of the potential' and 'force is equal to the rate of change of momentum'.

19.5 *Find closed-form expressions for* $\cos \mathsf{C}$ *and* $\sin \mathsf{C}$, *where* C *is the matrix*

$$\mathsf{C} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Demonstrate that the 'expected' relationships

$$\cos^2 \mathsf{C} + \sin^2 \mathsf{C} = \mathsf{I} \quad \textit{and} \quad \sin 2\mathsf{C} = 2 \sin \mathsf{C} \cos \mathsf{C}$$

are valid.

Consider the square of C:

$$\mathsf{C}^2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = 2\mathsf{I}.$$

Now

$$\cos \mathsf{C} = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} \mathsf{C}^{2n} = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} 2^n \, \mathsf{I}^n = (\cos \sqrt{2}) \, \mathsf{I}$$

and

$$\sin \mathsf{C} = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} \mathsf{C}^{2n+1} = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} 2^n \, \mathsf{I}^n \mathsf{C} = \frac{1}{\sqrt{2}} (\sin \sqrt{2}) \, \mathsf{C}.$$

To test the analogue of '$\cos^2 \theta + \sin^2 \theta = 1$':

$$\cos^2 \mathsf{C} + \sin^2 \mathsf{C} = (\cos^2 \sqrt{2}) \mathsf{I} + \tfrac{1}{2} (\sin^2 \sqrt{2}) \mathsf{C}^2$$
$$= (\cos^2 \sqrt{2}) \mathsf{I} + \tfrac{1}{2} (\sin^2 \sqrt{2}) \, 2 \, \mathsf{I} = \mathsf{I},$$

as expected.

To test the analogue of '$\sin 2\theta = 2\sin\theta\cos\theta$', we note that $(2\mathsf{C})^{2n} = 2^{2n}(2\mathsf{I})^n = (2^2\,2\,\mathsf{I})^n = 8^n\mathsf{I}^n$ and obtain

$$\begin{aligned}
\sin 2\mathsf{C} &= \sin\begin{pmatrix} 2 & 2 \\ 2 & -2 \end{pmatrix} \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}\, 8^n\, \mathsf{I}^n \begin{pmatrix} 2 & 2 \\ 2 & -2 \end{pmatrix} \\
&= \frac{1}{\sqrt{8}}\,(\sin\sqrt{8})\begin{pmatrix} 2 & 2 \\ 2 & -2 \end{pmatrix} \\
&= \frac{\sin 2\sqrt{2}}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.
\end{aligned}$$

But we also have that

$$\begin{aligned}
2\sin\mathsf{C}\cos\mathsf{C} &= 2\,\frac{1}{\sqrt{2}}(\sin\sqrt{2})\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}(\cos\sqrt{2})\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
&= \frac{2\,\sin\sqrt{2}\,\cos\sqrt{2}}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \\
&= \frac{\sin 2\sqrt{2}}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},
\end{aligned}$$

thus confirming the relationship (at least in this case).

> ***19.7*** *Expressed in terms of the annihilation and creation operators A and $A^\dagger$, a system has an unperturbed Hamiltonian $H_0 = \hbar\omega A^\dagger A$. The system is disturbed by the addition of a perturbing Hamiltonian $H_1 = g\hbar\omega(A + A^\dagger)$, where g is real. Show that the effect of the perturbation is to move the whole energy spectrum of the system down by $g^2\hbar\omega$.*

The total Hamiltonian H for the system is $H_0 + H_1$, where

$$H_0 = \hbar\omega A^\dagger A \qquad \text{and} \qquad H_1 = g\hbar\omega(A + A^\dagger).$$

We note that both terms are Hermitian ($H_0^\dagger = H_0$, $H_1^\dagger = H_1$) and that the energy spectrum of the system is given by the eigenvalues μ_i for which

$$(H_0 + H_1)\,|\,\psi_i\rangle = \mu_i\,|\,\psi_i\rangle$$

has solutions.

Now,
$$\begin{aligned}
H &= H_0 + H_1 \\
&= \hbar\omega[\,A^\dagger A + g(A + A^\dagger)\,] \\
&= \hbar\omega[\,(A^\dagger + gI)(A + gI) - g^2 I\,].
\end{aligned}$$

We define B by $B = A + gI$, with $B^\dagger = A^\dagger + gI$, and consider

$$\left[\, B, B^\dagger \,\right] = \left[\, A + gI, A^\dagger + gI \,\right] = \left[\, A, A^\dagger \,\right],$$

since $[\, C, I \,] = 0$ for any C and clearly $[\, I, I \,] = 0$.

Thus, H is expressible as

$$H = \hbar\omega B^\dagger B - g^2\hbar\omega I$$

with $\left[\, B, B^\dagger \,\right] = \left[\, A, A^\dagger \,\right]$. That is, H has the same structure with respect to B as H_0 has with respect to A (apart from an additional term proportional to the identity operator) and B and $B^\dagger$ have the same commutation relation as A and $A^\dagger$. This implies that H has the same spectrum of eigenvalues λ_i as H_0, except for a (downward) shift of $-g^2\hbar\omega$, i.e. $\mu_i = \lambda_i - g^2\hbar\omega$ for each value of i. Thus the whole spectrum is lowered by this amount.

19.9 *By considering the function*

$$F(\lambda) = \exp(\lambda A) B \exp(-\lambda A),$$

where A and B are linear operators and λ a parameter, and finding its derivatives with respect to λ, prove that

$$e^A B e^{-A} = B + [\, A, B \,] + \frac{1}{2!}\,[\, A, [\, A, B \,]\,] + \frac{1}{3!}\,[\, A, [\, A, [\, A, B \,]\,]\,] + \cdots.$$

Use this result to express

$$\exp\left(\frac{iL_x\theta}{\hbar}\right) L_y \exp\left(\frac{-iL_x\theta}{\hbar}\right)$$

as a linear combination of the angular momentum operators L_x, L_y and L_z.

Starting from the definition of $F(\lambda)$, we calculate its first few derivatives with respect to λ, remembering that operator A commutes with any function of A but not necessarily with any function of B:

$$\begin{aligned}
F(\lambda) &= \exp(\lambda A) B \exp(-\lambda A),\\
\frac{dF(\lambda)}{d\lambda} &= A \exp(\lambda A) B \exp(-\lambda A) - \exp(\lambda A) B \exp(-\lambda A) A\\
&= AF(\lambda) - F(\lambda)A = [\, A, F(\lambda) \,],\\
\frac{d^2F(\lambda)}{d\lambda^2} &= A\frac{dF}{d\lambda} - \frac{dF}{d\lambda}A = \left[\, A, \frac{dF(\lambda)}{d\lambda} \,\right] = [\, A, [\, A, F(\lambda) \,]\,],\\
\frac{d^3F(\lambda)}{d\lambda^3} &= [\, A, [\, A, [\, A, F(\lambda) \,]\,]\,], \text{ and so on for higher derivatives.}
\end{aligned}$$

Now we use a Taylor series in λ, based on the values of the derivatives at $\lambda = 0$, to evaluate $F(1)$. At $\lambda = 0$, $F(\lambda) = B$, and we obtain

$$\begin{aligned}
e^A B e^{-A} &= F(1) \\
&= F(0) + \frac{dF(0)}{d\lambda} + \frac{1}{2!}\,\frac{d^2F(0)}{d\lambda^2} + \frac{1}{3!}\,\frac{d^3F(0)}{d\lambda^3} + \cdots \\
&= B + [\,A, B\,] + \frac{1}{2!}\,[\,A, [\,A, B\,]\,] + \frac{1}{3!}\,[\,A, [\,A, [\,A, B\,]\,]\,] + \cdots .
\end{aligned}$$

To apply this result to

$$\Theta \equiv \exp\left(\frac{iL_x\theta}{\hbar}\right) L_y \exp\left(\frac{-iL_x\theta}{\hbar}\right),$$

we need to take A as $iL_x\theta/\hbar$ and B as L_y. The corresponding commutator is given by

$$\left[\frac{iL_x\theta}{\hbar}, L_y\right] = \frac{i\theta}{\hbar}\,\left[\,L_x, L_y\,\right] = -\theta L_z.$$

Because multiple commutators are involved, we will also need

$$\left[\frac{iL_x\theta}{\hbar}, L_z\right] = \frac{i\theta}{\hbar}\,[\,L_x, L_z\,] = \theta L_y.$$

Substituting in the derived series, we obtain

$$\begin{aligned}
\Theta &= L_y + (-\theta L_z) + \frac{1}{2!}\left[\frac{i\theta L_x}{\hbar}, -\theta L_z\right] + \cdots \\
&= L_y - \theta L_z + \frac{1}{2!}(-\theta^2 L_y) + \frac{1}{3!}\left[\frac{i\theta L_x}{\hbar}, -\theta^2 L_y\right] + \cdots \\
&= L_y - \theta L_z - \frac{1}{2!}\,\theta^2 L_y + \frac{1}{3!}(\theta^3 L_z) + \frac{1}{4!}\left[\frac{i\theta L_x}{\hbar}, \theta^3 L_z\right] + \cdots \\
&= L_y - \theta L_z - \frac{1}{2!}\,\theta^2 L_y + \frac{1}{3!}\,\theta^3 L_z + \frac{1}{4!}(\theta^4 L_y) + \frac{1}{5!}\left[\frac{i\theta L_x}{\hbar}, \theta^4 L_y\right] + \\
&= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \cdots\right) L_y - \left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \cdots\right) L_z \\
&= \cos\theta\, L_y - \sin\theta\, L_z.
\end{aligned}$$

At each stage in the above calculation, the value of the commutator in the nth term of the series has been used to reduce the $(n+1)$th term from a multiple commutator, with n levels of nesting, to a single commutator.

20

Partial differential equations: general and particular solutions

20.1 *Determine whether the following can be written as functions of $p = x^2 + 2y$ only, and hence whether they are solutions of*

$$\frac{\partial u}{\partial x} = x\frac{\partial u}{\partial y}.$$

(a) $x^2(x^2 - 4) + 4y(x^2 - 2) + 4(y^2 - 1)$;
(b) $x^4 + 2x^2y + y^2$;
(c) $[x^4 + 4x^2y + 4y^2 + 4]/[2x^4 + x^2(8y + 1) + 8y^2 + 2y]$.

As a first step, we verify that any function of $p = x^2 + 2y$ will satisfy the given equation. Using the chain rule, we have

$$\frac{\partial u}{\partial p}\frac{\partial p}{\partial x} = x\frac{\partial u}{\partial p}\frac{\partial p}{\partial y},$$
$$\Rightarrow \quad \frac{\partial u}{\partial p}2x = x\frac{\partial u}{\partial p}2.$$

This is satisfied for *any* function $u(p)$, thus completing the verification.

To test the given functions we substitute for $y = \frac{1}{2}(p - x^2)$ or for $x^2 = p - 2y$ in each of the $f(x, y)$ and then examine whether the resulting forms are independent of x or y, respectively.

$$\begin{aligned}
\text{(a) } f(x, y) &= x^2(x^2 - 4) + 4y(x^2 - 2) + 4(y^2 - 1)\\
&= x^2(x^2 - 4) + 2(p - x^2)(x^2 - 2) + p^2 - 2p\,x^2 + x^4 - 4\\
&= x^4(1 - 2 + 1) + x^2(-4 + 2p + 4 - 2p) - 4p + p^2 - 4\\
&= p^2 - 4p - 4 = g(p).
\end{aligned}$$

This is a function of p only, and therefore the original $f(x,y)$ is a solution of the PDE.

Though not necessary for answering the question, we will repeat the verification, but this time by substituting for x rather than for y:

$$\begin{aligned}
f(x,y) &= x^2(x^2-4) + 4y(x^2-2) + 4(y^2-1) \\
&= (p-2y)(p-2y-4) + 4y(p-2y-2) + 4(y^2-1) \\
&= p^2 - 4py + 4y^2 - 4p + 8y + 4yp - 8y^2 - 8y + 4y^2 - 4 \\
&= p^2 - 4p - 4 = g(p);
\end{aligned}$$

i.e. it is the same as before, as it must be, and again this shows that $f(x,y)$ is a solution of the PDE.

(b)
$$\begin{aligned}
f(x,y) &= x^4 + 2x^2y + y^2 \\
&= (p-2y)^2 + 2y(p-2y) + y^2 \\
&= p^2 - 4p\,y + 4y^2 + 2p\,y - 4y^2 + y^2 \\
&= (p-y)^2 \neq g(p).
\end{aligned}$$

As this is a function of both p and y, it is not a solution of the PDE.

(c)
$$\begin{aligned}
f(x,y) &= \frac{x^4 + 4x^2y + 4y^2 + 4}{2x^4 + x^2(8y+1) + 8y^2 + 2y} \\
&= \frac{p^2 - 4p\,y + 4y^2 + 4yp - 8y^2 + 4y^2 + 4}{2p^2 - 8p\,y + 8y^2 + 8yp + p - 16y^2 - 2y + 8y^2 + 2y} \\
&= \frac{p^2+4}{2p^2+p} = g(p).
\end{aligned}$$

This is a function of p only and therefore $f(x,y)$ is a solution of the PDE.

20.3 *Solve the following partial differential equations for $u(x,y)$ with the boundary conditions given:*

(a) $x\dfrac{\partial u}{\partial x} + xy = u,\quad u = 2y$ *on the line* $x = 1$;

(b) $1 + x\dfrac{\partial u}{\partial y} = xu,\quad u(x,0) = x.$

(a) This can be solved as an ODE for u as a function of x, though the 'constant of integration' will be a function of y. In standard form, the equation reads

$$\frac{\partial u}{\partial x} - \frac{u}{x} = -y.$$

By inspection (or formal calculation) the IF for this is x^{-1} and the equation can be rearranged as

$$\frac{\partial}{\partial x}\left(\frac{u}{x}\right) = -\frac{y}{x},$$

$$\Rightarrow \quad \frac{u}{x} = -y\ln x + f(y),$$

$$u = 2y \text{ on } x = 1 \ \Rightarrow f(y) = 2y,$$

$$\text{and so } u(x,y) = xy(2 - \ln x).$$

(b) This equation can be written in standard form, with u as a function of y:

$$\frac{\partial u}{\partial y} - u = -\frac{1}{x},$$

for which the IF is clearly e^{-y}, leading to

$$\frac{\partial}{\partial y}\left(e^{-y}u\right) = -\frac{e^{-y}}{x},$$

$$\Rightarrow \quad e^{-y}u = \frac{e^{-y}}{x} + f(x),$$

$$u(x,0) = x \Rightarrow f(x) = x - \frac{1}{x}.$$

Substituting this result, and multiplying through by e^{y}, gives $u(x,y)$ as

$$u(x,y) = \frac{1}{x} + \left(x - \frac{1}{x}\right)e^{y}.$$

20.5 *Find solutions of*

$$\frac{1}{x}\frac{\partial u}{\partial x} + \frac{1}{y}\frac{\partial u}{\partial y} = 0$$

for which (a) $u(0,y) = y$ *and* (b) $u(1,1) = 1$.

As usual, we find $p(x,y)$ from

$$\frac{dx}{x^{-1}} = \frac{dy}{y^{-1}} \quad \Rightarrow \quad x^2 - y^2 = p.$$

(a) On $x = 0$, $p = -y^2$ and

$$u(0,y) = y = (-p)^{1/2} \quad \Rightarrow \quad u(x,y) = [-(x^2 - y^2)]^{1/2} = (y^2 - x^2)^{1/2}.$$

(b) At $(1,1)$, $p = 0$ and

$$u(1,1) = 1 \quad \Rightarrow \quad u(x,y) = 1 + g(x^2 - y^2),$$

where g is any function that has $g(0) = 0$.

We note that in part (a) the solution is uniquely determined because the boundary values are given along a line, whereas in part (b), where the value is fixed at only an isolated point, the solution is indeterminate to the extent of a loosely determined function. This is the normal situation, though it is modified if the line in (a) happens to be a characteristic of the PDE.

20.7 *Solve*

$$\sin x \frac{\partial u}{\partial x} + \cos x \frac{\partial u}{\partial y} = \cos x \qquad (*)$$

subject to (a) $u(\pi/2, y) = 0$ *and* (b) $u(\pi/2, y) = y(y+1)$.

As usual, the CF is found from

$$\frac{dx}{\sin x} = \frac{dy}{\cos x} \quad \Rightarrow \quad y - \ln \sin x = p.$$

Since the RHS of $(*)$ is a factor in one of the terms on the LHS, a trivial PI is any function of y only whose derivative (with respect to y) is unity, of which the simplest is $u(x, y) = y$. The general solution is therefore

$$u(x, y) = f(y - \ln \sin x) + y.$$

The actual form of the arbitrary function $f(p)$ is determined by the form that $u(x, y)$ takes on the boundary, here the line $x = \pi/2$.

(a) With $u(\pi/2, y) = 0$:

$$0 = f(y - 0) + y \quad \Rightarrow \quad f(p) = -p$$
$$\Rightarrow \quad u(x, y) = \ln \sin x - y + y = \ln \sin x.$$

(b) With $u(\pi/2, y) = y(y+1)$:

$$y(y+1) = f(y - 0) + y \quad \Rightarrow \quad f(p) = p^2$$
$$\Rightarrow \quad u(x, y) = (y - \ln \sin x)^2 + y.$$

20.9 *If* $u(x, y)$ *satisfies*

$$\frac{\partial^2 u}{\partial x^2} - 3\frac{\partial^2 u}{\partial x \partial y} + 2\frac{\partial^2 u}{\partial y^2} = 0$$

and $u = -x^2$ *and* $\partial u/\partial y = 0$ *for* $y = 0$ *and all* x, *find the value of* $u(0, 1)$.

If we are to find solutions to this homogeneous second-order PDE of the form $u(x,y) = f(x + \lambda y)$, then λ must satisfy

$$1 - 3\lambda + 2\lambda^2 = 0 \quad \Rightarrow \quad \lambda = \tfrac{1}{2},\ 1.$$

Thus $u(x,y) = g(x + \frac{1}{2}y) + f(x+y) \equiv g(p_1) + f(p_2)$.

On $y = 0$, $p_1 = p_2 = x$ and

$$-x^2 = u(x,0) = g(x) + f(x), \qquad (*)$$

$$0 = \frac{\partial u}{\partial y}(x,0) = \tfrac{1}{2}g'(x) + f'(x).$$

From $(*)$, $\quad -2x = g'(x) + f'(x).$

Subtracting, $\quad 2x = -\frac{1}{2}g'(x).$

Integrating, $\quad g(x) = -2x^2 + k \quad \Rightarrow \quad f(x) = x^2 - k, \quad$ from $(*)$.

Hence,
$$u(x,y) = -2(x + \tfrac{1}{2}y)^2 + k + (x+y)^2 - k$$
$$= -x^2 + \tfrac{1}{2}y^2.$$

At the particular point $(0,1)$ we have $u(0,1) = -0^2 + \frac{1}{2}(1)^2 = \frac{1}{2}$.

20.11 *In those cases in which it is possible to do so, evaluate $u(2,2)$, where $u(x,y)$ is the solution of*

$$2y\frac{\partial u}{\partial x} - x\frac{\partial u}{\partial y} = xy(2y^2 - x^2)$$

that satisfies the (separate) boundary conditions given below.

(a) $u(x,1) = x^2$ *for all* x.
(b) $u(x,1) = x^2$ *for* $x \geq 0$.
(c) $u(x,1) = x^2$ *for* $0 \leq x \leq 3$.
(d) $u(x,0) = x$ *for* $x \geq 0$.
(e) $u(x,0) = x$ *for all* x.
(f) $u(1, \sqrt{10}) = 5$.
(g) $u(\sqrt{10}, 1) = 5$.

To find the CF, $u(x,y) = f(p)$, we set

$$\frac{dx}{2y} = -\frac{dy}{x} \quad \Rightarrow \quad x^2 + 2y^2 = p.$$

This result also defines the characteristic curves, which are right ellipses centred on the origin with semi-axes of lengths $\sqrt{p}$ and $\sqrt{p/2}$. The point $(2,2)$ lies on the characteristic with $p = 2^2 + 2(2^2) = 12$; we will only be able to determine the value of $u(2,2)$ if this curve cuts the boundary on which u is specified.

For a PI we try $u(x, y) = Ax^n y^m$:

$$2Anx^{n-1}y^{m+1} - Amx^{n+1}y^{m-1} = 2xy^3 - x^3 y,$$

which has a solution, $n = m = 2$ with $A = \frac{1}{2}$. Thus the general solution is

$$u(x, y) = f(x^2 + 2y^2) + \tfrac{1}{2}x^2 y^2.$$

(a) With $u(x, 1) = x^2$ for all x:

$$\begin{aligned}
x^2 = u(x,1) &= f(x^2 + 2) + \tfrac{1}{2}x^2 \\
\Rightarrow \qquad f(p) &= \tfrac{1}{2}(p - 2) \\
\Rightarrow \quad u(x, y) &= \tfrac{1}{2}(x^2 + 2y^2 - 2) + \tfrac{1}{2}x^2 y^2 \\
&= \tfrac{1}{2}(x^2 + x^2 y^2 + 2y^2 - 2), \\
u(2, 2) &= \tfrac{1}{2}(4 + 16 + 8 - 2) = 13.
\end{aligned}$$

The line $y = 1$ cuts each characteristic in zero (for $p < 2$) or two (for $p > 2$) distinct points. Here $p = 12$ (> 2) and the characteristic (ellipse) that passes through $(2, 2)$ cuts the boundary (the line $y = 1$) in two places. In general, a double intersection can lead to inconsistencies and hence to no solution. However, it causes no difficulty with the given boundary conditions since the required values of x^2 at $x = \pm\sqrt{12 - 2(1^2)}$ are equal and u is a even function of x.

(b) With $u(x, 1) = x^2$ for $x \geq 0$.
Since every characteristic ellipse (with $p > 2$) cuts the line $y = 1$ once (and only once in $x > 0$), this gives the same result as in part (a).

(c) With $u(x, 1) = x^2$ for $0 \leq x \leq 3$.
The ellipses that cut the line $y = 1$ with $0 \leq x \leq 3$ have p-values lying between $0^2 + 2(1)^2 = 2$ and $3^2 + 2(1)^2 = 11$. Thus the $p = 12$ curve does not do so and $u(2, 2)$ is undetermined.

(d) With $u(x, 0) = x$ for $x \geq 0$:

$$\begin{aligned}
x = u(x, 0) &= f(x^2 + 0) + 0 \\
\Rightarrow \qquad f(p) &= p^{1/2} \\
\Rightarrow \quad u(x, y) &= (x^2 + 2y^2)^{1/2} + \tfrac{1}{2}x^2 y^2 \\
\Rightarrow \quad u(2, 2) &= \sqrt{12} + 8.
\end{aligned}$$

The characteristic (ellipse) $p = 12$ cuts the positive x-axis (i.e. $y = 0$) in one and only one place ($x = +\sqrt{12}$) and so the solution is well defined and the above evaluation valid.

(e) With $u(x, 0) = x$ for all x.
This is as in part (d) except that now a characteristic ellipse cuts the defining boundary in two places, $x = \pm\sqrt{p}$, and requires both $u(\sqrt{p}, 0) = \sqrt{p}$ and

$u(-\sqrt{p}, 0) = -\sqrt{p}$. Since $\sqrt{z}$ is not differentiable at $z = 0$, this is not possible and no solution exists.

(f) With $u(1, \sqrt{10}) = 5$.
At the point $(1, \sqrt{10})$ the value of p is $1 + 2(10) = 21$. As the 'boundary' consists of just this one point, it is only at the points that lie on the characteristic $p = 21$ that the value of $u(x, y)$ can be known. Since for the point $(2, 2)$ the value of p is 12, the value of $u(2, 2)$ cannot be determined.

(g) With $u(\sqrt{10}, 1) = 5$.
At the point $(\sqrt{10}, 1)$ the value of p is $10 + 2(1) = 12$. Since for $(2, 2)$ it is also 12 the value of $u(2, 2)$ is determined and is given by $f(12) + \frac{1}{2}(4)(4) = 5 + 8 = 13$.

20.13 *The solution to the equation*

$$6\frac{\partial^2 u}{\partial x^2} - 5\frac{\partial^2 u}{\partial x \partial y} + \frac{\partial^2 u}{\partial y^2} = 14$$

that satisfies $u = 2x + 1$ and $\partial u/\partial y = 4 - 6x$, both on the line $y = 0$, is

$$u(x, y) = -8y^2 - 6xy + 2x + 4y + 1.$$

By changing the independent variables in the equation to

$$\xi = x + 2y \qquad \text{and} \qquad \eta = x + 3y,$$

show that it must be possible to write $14(x^2 + 5xy + 6y^2)$ in the form

$$f_1(x + 2y) + f_2(x + 3y) - (x^2 + y^2),$$

and determine the forms of $f_1(z)$ and $f_2(z)$.

Let $u(x, y) = v(\xi, \eta)$, with $\xi = x + 2y$ and $\eta = x + 3y$. We must first express the differential operators $\partial/\partial x$ and $\partial/\partial y$ in terms of differentiation with respect to ξ and η; to do this we use the chain rule:

$$\frac{\partial}{\partial x} = \frac{\partial \xi}{\partial x}\frac{\partial}{\partial \xi} + \frac{\partial \eta}{\partial x}\frac{\partial}{\partial \eta} = \frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta}; \text{ similarly } \frac{\partial}{\partial y} = 2\frac{\partial}{\partial \xi} + 3\frac{\partial}{\partial \eta}.$$

The equation becomes

$$6\left(\frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta}\right)\left(\frac{\partial v}{\partial \xi} + \frac{\partial v}{\partial \eta}\right) - 5\left(\frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta}\right)\left(2\frac{\partial v}{\partial \xi} + 3\frac{\partial v}{\partial \eta}\right)$$
$$+ \left(2\frac{\partial}{\partial \xi} + 3\frac{\partial}{\partial \eta}\right)\left(2\frac{\partial v}{\partial \xi} + 3\frac{\partial v}{\partial \eta}\right) = 14,$$

$$(6 - 10 + 4)\frac{\partial^2 v}{\partial \xi^2} + (12 - 25 + 12)\frac{\partial^2 v}{\partial \xi \partial \eta} + (6 - 15 + 9)\frac{\partial^2 v}{\partial \eta^2} = 14.$$

Collecting similar terms together, we find

$$\frac{\partial^2 v}{\partial \xi \partial \eta} = -14.$$

This equation has a CF of the form $f(\xi) + g(\eta)$ and a PI of $-14\xi\eta$. Thus its general solution is

$$v(\xi, \eta) = f(\xi) + g(\eta) - 14\xi\eta.$$

This must be the same as the given answer, i.e.

$$-8y^2 - 6xy + 2x + 4y + 1 = f(x + 2y) + g(x + 3y) - 14(x + 2y)(x + 3y)$$

for some functions f and g. Thus

$$\begin{aligned}
w(x, y) &= 14(x^2 + 5xy + 6y^2) \\
&= 14(x + 2y)(x + 3y) \\
&= f(x + 2y) + g(x + 3y) + 8y^2 + 6xy - 2x - 4y - 1 \\
&= f(x + 2y) + g(x + 3y) - (x^2 + y^2) + h(x, y), \\
\text{where} \quad h(x, y) &= x^2 + 9y^2 + 6xy - 2x - 4y - 1 \\
&= (x + 3y)^2 - 2(x + 2y) - 1 \\
&= F(x + 2y) + G(x + 3y).
\end{aligned}$$

It follows that

$$w(x, y) = f_1(x + 2y) + f_2(x + 3y) - (x^2 + y^2),$$

where $f_1(z) = f(z) - 2z - 1$ and $f_2(z) = g(z) + z^2$.

After rearrangement this reads

$$15x^2 + 70xy + 85y^2 = f_1(x + 2y) + f_2(x + 3y). \qquad (**)$$

Taking second derivatives with respect to x and y separately,

$$\begin{aligned}
30 &= f_1'' + f_2'', \\
170 &= 4f_1'' + 9f_2'', \\
\Rightarrow \quad 50 &= 5f_2'' \quad \Rightarrow \quad f_2(z) = 5z^2 + \alpha z + \beta \\
\text{and } 100 &= 5f_1'' \quad \Rightarrow \quad f_1(z) = 10z^2 + \gamma z + \delta.
\end{aligned}$$

Equating the coefficients of xy, x and y and the constants in $(**)$ gives $70 = 40+30$, $0 = \alpha + \gamma$, $0 = 3\alpha + 2\gamma$, $0 = \beta + \delta$. These equations have the solution $\alpha = \gamma = 0$ and $\beta = k = -\delta$. Thus

$$f_1(z) = 10z^2 - k \qquad \text{and} \qquad f_2(z) = 5z^2 + k.$$

Clearly, k can take any value without affecting the final form given in the question.

***20.15** Find the most general solution of $\partial^2 u/\partial x^2 + \partial^2 u/\partial y^2 = x^2y^2$.*

The complementary function for this equation is the solution to the two-dimensional Laplace equation and [either as a general known result or from substituting the trial form $h(x+\lambda y)$ which leads to $\lambda^2 = -1$ and hence to $\lambda = \pm i$] has the form $f(x+iy) + g(x-iy)$ for arbitrary functions f and g.

It therefore remains only to find a suitable PI. As f and g are not specified, there are infinitely many possibilities and which one we finish up with will depend upon the details of the approach adopted. When a solution has been obtained it should be checked by substitution.

As no PI is obvious by inspection, we make a change of variables with the object of obtaining one by means of an explicit integration. To do this, we use as new variables the arguments of the arbitrary functions appearing in the CF.

Setting $\xi = x+iy$ and $\eta = x-iy$, with $u(x,y) = v(\xi,\eta)$, gives

$$\begin{aligned}
&\left(\frac{\partial}{\partial \xi}+\frac{\partial}{\partial \eta}\right)\left(\frac{\partial v}{\partial \xi}+\frac{\partial v}{\partial \eta}\right) \\
&\qquad + \left(i\frac{\partial}{\partial \xi}-i\frac{\partial}{\partial \eta}\right)\left(i\frac{\partial v}{\partial \xi}-i\frac{\partial v}{\partial \eta}\right) = \left(\frac{\xi+\eta}{2}\right)^2\left(\frac{\xi-\eta}{2i}\right)^2, \\
&(1-1)\frac{\partial^2 v}{\partial \xi^2}+(2+2)\frac{\partial^2 v}{\partial \xi\,\partial \eta}+(1-1)\frac{\partial^2 v}{\partial \eta^2} = -\frac{1}{16}(\xi^2-\eta^2)^2, \\
&\frac{\partial^2 v}{\partial \xi\,\partial \eta} = -\frac{1}{64}(\xi^2-\eta^2)^2.
\end{aligned}$$

When we integrate this we can set all constants of integration and all arbitrary functions equal to zero as *any* solution will suffice:

$$\begin{aligned}
\frac{\partial^2 v}{\partial \xi\,\partial \eta} &= -\frac{1}{64}(\xi^4 - 2\xi^2\eta^2 + \eta^4), \\
\frac{\partial v}{\partial \eta} &= -\frac{1}{64}\left(\frac{\xi^5}{5} - \frac{2\xi^3\eta^2}{3} + \xi\eta^4\right), \\
v &= -\frac{1}{64}\left(\frac{\xi^5\eta}{5} - \frac{2\xi^3\eta^3}{9} + \frac{\xi\eta^5}{5}\right).
\end{aligned}$$

Re-expressing this solution as a function of x and y (noting that $\xi\eta = x^2+y^2$)

gives

$$\begin{aligned}
u(x,y) &= \frac{1}{(64)(45)}[\,10\xi^3\eta^3 - 9\xi\eta(\xi^4+\eta^4)\,] \\
&= \frac{1}{(64)(45)}[\,10(x^2+y^2)^3 - 18(x^2+y^2)(x^4-6x^2y^2+y^4)\,] \\
&= \frac{x^2+y^2}{(64)(45)}(10x^4+20x^2y^2+10y^4-18x^4+108x^2y^2-18y^4) \\
&= \frac{x^2+y^2}{(64)(45)}(128x^2y^2-8x^4-8y^4) \\
&= \frac{1}{360}(15x^4y^2-x^6+15x^2y^4-y^6).
\end{aligned}$$

Check

Applying $\dfrac{\partial^2}{\partial x^2}+\dfrac{\partial^2}{\partial y^2}$ to the final expression yields

$$\frac{1}{360}[\,15(12)x^2y^2-30x^4+30y^4+30x^4+15(12)x^2y^2-30y^4\,] = x^2y^2,$$

as it should.

20.17 *The non-relativistic Schrödinger equation,*

$$-\frac{\hbar^2}{2m}\nabla^2 u + V(\mathbf{r})u = i\hbar\frac{\partial u}{\partial t},$$

is similar to the diffusion equation in having different orders of derivatives in its various terms; this precludes solutions that are arbitrary functions of particular linear combinations of variables. However, since exponential functions do not change their forms under differentiation, solutions in the form of exponential functions of combinations of the variables may still be possible.

Consider the Schrödinger equation for the case of a constant potential, i.e. for a free particle, and show that it has solutions of the form $A\exp(lx+my+nz+\lambda t)$, where the only requirement is that

$$-\frac{\hbar^2}{2m}\left(l^2+m^2+n^2\right) = i\hbar\lambda.$$

In particular, identify the equation and wavefunction obtained by taking λ as $-iE/\hbar$, and l, m and n as $ip_x/\hbar, ip_y/\hbar$ and $ip_z/\hbar$, respectively, where E is the energy and $\mathbf{p}$ the momentum of the particle; these identifications are essentially the content of the de Broglie and Einstein relationships.

For a free particle we may omit the potential term $V(\mathbf{r})$ from the Schrödinger equation, which then reads (in Cartesian coordinates)

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}\right)=i\hbar\frac{du}{dt}.$$

We try $u(x,y,z,t)=A\exp(lx+my+nz+\lambda t)$, i.e. the product of four exponential functions, and obtain

$$-\frac{\hbar^2}{2m}(l^2+m^2+n^2)u=i\hbar\lambda u.$$

This equation is clearly satisfied provided

$$-\frac{\hbar^2}{2m}(l^2+m^2+n^2)=i\hbar\lambda.$$

With λ as $-iE/\hbar$, and l,m and n as $ip_x/\hbar, ip_y/\hbar$ and $ip_z/\hbar$, respectively, where E is the energy and $\mathbf{p}$ is the momentum of the particle, we have

$$-\frac{\hbar^2}{2m}\left(-\frac{p_x^2}{\hbar^2}-\frac{p_y^2}{\hbar^2}-\frac{p_z^2}{\hbar^2}\right)=E,$$

which can be written more compactly as $E=p^2/2m$, the classical non-relativistic relationship between the (kinetic) energy and momentum of a free particle.

The wavefunction obtained is

$$\begin{aligned}u(\mathbf{r},t)&=A\exp\left[\frac{i}{\hbar}(p_x x+p_y y+p_z z-Et)\right]\\&=A\exp\left[\frac{i}{\hbar}(\mathbf{p}\cdot\mathbf{r}-Et)\right],\end{aligned}$$

i.e. a classical plane wave of wave number $\mathbf{k}=\mathbf{p}/\hbar$ and angular frequency $\omega=E/\hbar$ travelling in the direction $\mathbf{p}/p$.

20.19 *An incompressible fluid of density ρ and negligible viscosity flows with velocity v along a thin, straight, perfectly light and flexible tube, of cross-section A which is held under tension T. Assume that small transverse displacements u of the tube are governed by*

$$\frac{\partial^2 u}{\partial t^2}+2v\frac{\partial^2 u}{\partial x\partial t}+\left(v^2-\frac{T}{\rho A}\right)\frac{\partial^2 u}{\partial x^2}=0.$$

(a) *Show that the general solution consists of a superposition of two waveforms travelling with different speeds.*

(b) *The tube initially has a small transverse displacement $u=a\cos kx$ and is suddenly released from rest. Find its subsequent motion.*

(a) This is a second-order equation and will (in general) have two solutions of the form $u(x,t) = f(x + \lambda t)$, where both λ satisfy

$$\lambda^2 + 2v\lambda + \left(v^2 - \frac{T}{\rho A}\right) = 0 \quad \Rightarrow \quad \lambda = -v \pm \sqrt{v^2 - v^2 + \frac{T}{\rho A}} \equiv -v \pm \alpha,$$

and gives (minus) the speed of the corresponding profile. Thus the general displacement consists of a superposition of waveforms travelling with speeds $v \mp \alpha$.

(b) Now $u(x,0) = a\cos kx$ and $\dot{u}(x,0) = 0$, where the dot denotes differentiation with respect to time t. Let the general solution be given by

$$u(x,t) = f[\,x - (v+\alpha)t\,] + g[\,x - (v-\alpha)t\,],$$

with $\quad a\cos kx = f(x) + g(x)$

and $\quad 0 = -(v+\alpha)f'(x) - (v-\alpha)g'(x).$

We differentiate the first of these with respect to x and then eliminate the function $f'(x)$:

$$\begin{aligned}
-ka\sin kx &= f'(x) + g'(x),\\
-ka(v+\alpha)\sin kx &= (v+\alpha-v+\alpha)g'(x),\\
g'(x) &= -\frac{ka(v+\alpha)}{2\alpha}\,\sin kx,\\
\Rightarrow \quad g(x) &= \frac{v+\alpha}{2\alpha}\,a\cos kx + c,\\
\Rightarrow \quad f(x) &= \frac{\alpha - v}{2\alpha}\,a\cos kx - c.
\end{aligned}$$

Now that the forms of the initially arbitrary functions $f(x)$ and $g(x)$ have been determined, it follows that, for a *general* time t,

$$\begin{aligned}
u(x,t) &= \frac{\alpha - v}{2\alpha}\,a\cos[\,kx - k(v+\alpha)t\,] + \frac{\alpha+v}{2\alpha}\,a\cos[\,kx - k(v-\alpha)t\,]\\
&= \frac{a}{2}2\cos(kx - kvt)\cos k\alpha t + \frac{va}{2\alpha}\,2\sin(kx-kvt)\sin(-k\alpha t)\\
&= a\cos[\,k(x-vt)\,]\cos k\alpha t - \frac{va}{\alpha}\,\sin[\,k(x-vt)\,]\sin k\alpha t.
\end{aligned}$$

20.21 *In an electrical cable of resistance R and capacitance C, each per unit length, voltage signals obey the equation $\partial^2 V/\partial x^2 = RC\partial V/\partial t$. This (diffusion-type) equation has solutions of the form*

$$f(\zeta) = \frac{2}{\sqrt{\pi}} \int_0^{\zeta} \exp(-v^2)\, dv, \ \textit{where } \zeta = \frac{x(RC)^{1/2}}{2t^{1/2}}.$$

It also has solutions of the form $V = Ax + D$.

(a) *Find a combination of these that represents the situation after a steady voltage V_0 is applied at $x = 0$ at time $t = 0$.*
(b) *Obtain a solution describing the propagation of the voltage signal resulting from the application of the signal $V = V_0$ for $0 < t < T$, $V = 0$ otherwise, to the end $x = 0$ of an infinite cable.*
(c) *Show that for $t \gg T$ the maximum signal occurs at a value of x proportional to $t^{1/2}$ and has a magnitude proportional to t^{-1}.*

(a) Consider the given function

$$f(\zeta) = \frac{2}{\sqrt{\pi}} \int_0^{\zeta} \exp(-v^2)\, dv, \ \text{where } \zeta = \frac{x(RC)^{1/2}}{2t^{1/2}}.$$

The requirements to be satisfied by the correct combination of this function and $V(x,t) = Ax + D$ are (i) that, at $t = 0$, V is zero for all x, except $x = 0$ where it is V_0, and (ii) that, as $t \to \infty$, V is V_0 for all x.

(i) At $t = 0$, $\zeta = \infty$ and $f(\zeta) = 1$ for all $x \neq 0$.
(ii) As $t \to \infty$, $\zeta \to 0$ and $f(\zeta) \to 0$ for all finite x.

The required combination is therefore $D = V_0$ and $-V_0 f(\zeta)$, i.e.

$$V(x,t) = V_0 \left[1 - \frac{2}{\sqrt{\pi}} \int_0^{\frac{1}{2}x(CR/t)^{1/2}} \exp(-v^2)\, dv \right].$$

(b) The equation is linear and so we may superpose solutions. The response to the input $V = V_0$ for $0 < t < T$ can be considered as that to V_0 applied at $t = 0$ and continued, together with $-V_0$ applied at $t = T$ and continued. The solution is therefore the difference between two solutions of the form found in part (a):

$$V(x,t) = \frac{2V_0}{\sqrt{\pi}} \int_{\frac{1}{2}x(CR/t)^{1/2}}^{\frac{1}{2}x[CR/(t-T)]^{1/2}} \exp\left(-v^2\right)\, dv.$$

(c) To find the maximum signal we set $\partial V/\partial x$ equal to zero. Remembering that we are differentiating with respect to the limits of an integral (whose integrand does not contain x explicitly), we obtain

$$\frac{1}{2}\left(\frac{CR}{t-T}\right)^{1/2}\exp\left[-\frac{x^2CR}{4(t-T)}\right]-\frac{1}{2}\left(\frac{CR}{t}\right)^{1/2}\exp\left[-\frac{x^2CR}{4t}\right]=0.$$

This requires

$$\begin{aligned}\left(\frac{t-T}{t}\right)^{1/2}&=\exp\left[-\frac{x^2CR}{4(t-T)}+\frac{x^2CR}{4t}\right]\\&=\exp\left[\frac{x^2CR(-t+t-T)}{4t(t-T)}\right].\end{aligned}$$

For $t \gg T$, we expand both sides:

$$\begin{aligned}&1-\frac{1}{2}\frac{T}{t}+\cdots=1-\frac{Tx^2CR}{4t^2}+\cdots,\\&\Rightarrow \quad x^2\approx\frac{2t}{CR}\quad\Rightarrow\quad v=\frac{1}{2}\sqrt{\frac{2t}{CR}}\left(\frac{CR}{t}\right)^{1/2}=\frac{1}{\sqrt{2}}.\end{aligned}$$

The corresponding value of V is approximately equal to the value of the integrand, evaluated at this value of v, multiplied by the difference between the two limits of the integral. Thus

$$\begin{aligned}V_{\max}&\approx\frac{2V_0}{\sqrt{\pi}}\exp(-v^2)\frac{x\sqrt{CR}}{2}\left[\frac{1}{(t-T)^{1/2}}-\frac{1}{t^{1/2}}\right]\\&\approx\frac{2V_0}{\sqrt{\pi}}e^{-1/2}\frac{x\sqrt{CR}}{2}\frac{1}{2}\frac{T}{t^{3/2}}\\&=\frac{V_0Te^{-1/2}}{\sqrt{2\pi}\,t}.\end{aligned}$$

In summary, for $t \gg T$ the maximum signal occurs at a value of x proportional to $t^{1/2}$ and has a magnitude proportional to t^{-1}.

20.23 *Consider each of the following situations in a qualitative way and determine the equation type, the nature of the boundary curve and the type of boundary conditions involved:*

(a) *a conducting bar given an initial temperature distribution and then thermally isolated;*
(b) *two long conducting concentric cylinders, on each of which the voltage distribution is specified;*
(c) *two long conducting concentric cylinders, on each of which the charge distribution is specified;*
(d) *a semi-infinite string, the end of which is made to move in a prescribed way.*

We use the notation

$$A\frac{\partial^2 u}{\partial x^2} + B\frac{\partial^2 u}{\partial x \partial y} + C\frac{\partial^2 u}{\partial y^2} + D\frac{\partial u}{\partial x} + E\frac{\partial u}{\partial y} + Fu = R(x, y)$$

to express the most general type of PDE, and the following table

Equation type	Boundary	Conditions
hyperbolic	open	Cauchy
parabolic	open	Dirichlet or Neumann
elliptic	closed	Dirichlet or Neumann

to determine the appropriate boundary type and hence conditions.

(a) The diffusion equation $\kappa\frac{\partial^2 T}{\partial x^2} = \frac{\partial T}{\partial t}$ has $A = \kappa$, $B = 0$ and $C = 0$; thus $B^2 = 4AC$ and the equation is parabolic. This needs an open boundary. In the present case, the initial heat distribution (at the $t = 0$ boundary) is a Dirichlet condition and the insulation (no temperature gradient at the external surfaces) is a Neumann condition.

(b) The governing equation in two-dimensional Cartesians (not the natural choice for this situation, but this does not matter for the present purpose) is the Laplace equation, $\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = 0$, which has $A = 1$, $B = 0$ and $C = 1$ and therefore $B^2 < 4AC$. The equation is therefore elliptic and requires a closed boundary. Since ϕ is specified on the cylinders, the boundary conditions are Dirichlet in this particular situation.

(c) This is the same as part (b) except that the specified charge distribution σ determines $\partial\phi/\partial n$, through $\partial\phi/\partial n = \sigma/\epsilon_0$, and imposes Neumann boundary conditions.

(d) For the wave equation $\frac{\partial^2 u}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = 0$, we have $A = 1$, $B = 0$ and $C = -c^{-2}$, thus making $B^2 > 4AC$ and the equation hyperbolic. We thus require an open boundary and Cauchy conditions, with the displacement of the end of the string having to be specified at all times — this is equivalent to the displacement and the velocity of the end of the string being specified at all times.

20.25 *The Klein–Gordon equation (which is satisfied by the quantum-mechanical wavefunction $\Phi(\mathbf{r})$ of a relativistic spinless particle of non-zero mass m) is*

$$\nabla^2\Phi - m^2\Phi = 0.$$

Show that the solution for the scalar field $\Phi(\mathbf{r})$ in any volume V bounded by a surface S is unique if either Dirichlet or Neumann boundary conditions are specified on S.

Suppose that, for a given set of boundary conditions ($\Phi = f$ or $\partial\Phi/\partial n = g$ on S), there are two solutions to the Klein–Gordon equation, Φ_1 and Φ_2. Then consider $\Phi_3 = \Phi_1 - \Phi_2$, which satisfies

$$\nabla^2\Phi_3 = \nabla^2\Phi_1 - \nabla^2\Phi_2 = m^2\Phi_1 - m^2\Phi_2 = m^2\Phi_3$$

and

$$\text{either } \Phi_3 = f - f = 0, \text{ or } \frac{\partial\Phi_3}{\partial n} = g - g = 0 \text{ on } S.$$

Now apply Green's first theorem with the scalar functions equal to Φ_3 and Φ_3^*:

$$\int^S \Phi_3^* \frac{\partial\Phi_3}{\partial n}\, dS = \int_V [\,\Phi_3^*\nabla^2\Phi_3 + (\nabla\Phi_3^*)\cdot(\nabla\Phi_3)\,]\, dV,$$

$$\Rightarrow \quad 0 = \int_V (m^2|\Phi_3|^2 + |\nabla\Phi_3|^2)\, dV,$$

whichever set of boundary conditions applies. Since both terms in the integrand on the RHS are non-negative, each must be equal to zero. In particular, $|\Phi_3| = 0$ implies that $\Phi_3 = 0$ everywhere, i.e $\Phi_1 = \Phi_2$ everywhere; the solution is unique.

21

Partial differential equations: separation of variables and other methods

21.1 *Solve the following first-order partial differential equations by separating the variables:*

$$\text{(a)}\ \frac{\partial u}{\partial x} - x\frac{\partial u}{\partial y} = 0; \qquad \text{(b)}\ x\frac{\partial u}{\partial x} - 2y\frac{\partial u}{\partial y} = 0.$$

In each case we write $u(x, y) = X(x)Y(y)$, separate the variables into groups that each depend on only one variable, and then assert that each must be equal to a constant, with the several constants satisfying an arithmetic identity.

(a)

$$\begin{aligned}
\frac{\partial u}{\partial x} - x\frac{\partial u}{\partial y} &= 0, \\
X'Y - xXY' &= 0, \\
\frac{X'}{xX} = \frac{Y'}{Y} = k \quad &\Rightarrow \quad \ln X = \tfrac{1}{2}kx^2 + c_1, \quad \ln Y = ky + c_2, \\
&\Rightarrow \quad X = Ae^{kx^2/2}, \quad Y = Be^{ky}, \\
\Rightarrow \quad u(x, y) &= Ce^{\lambda(x^2+2y)}, \text{ where } k = 2\lambda.
\end{aligned}$$

(b)

$$\begin{aligned}
x\frac{\partial u}{\partial x} - 2y\frac{\partial u}{\partial y} &= 0, \\
xX'Y - 2yXY' &= 0, \\
\frac{xX'}{X} = \frac{2yY'}{Y} = k \quad &\Rightarrow \quad \ln X = k\ln x + c_1, \\
&\qquad \ln Y = \tfrac{1}{2}k\ln y + c_2, \\
&\Rightarrow \quad X = Ax^k, \quad Y = By^{k/2}, \\
\Rightarrow \quad u(x, y) &= C(x^2y)^{\lambda}, \text{ where } k = 2\lambda.
\end{aligned}$$

21.3 *The wave equation describing the transverse vibrations of a stretched membrane under tension T and having a uniform surface density ρ is*

$$T\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right)=\rho\frac{\partial^2 u}{\partial t^2}.$$

Find a separable solution appropriate to a membrane stretched on a frame of length a and width b, showing that the natural angular frequencies of such a membrane are given by

$$\omega^2=\frac{\pi^2 T}{\rho}\left(\frac{n^2}{a^2}+\frac{m^2}{b^2}\right),$$

where n and m are any positive integers.

We seek solutions $u(x,y,t)$ that are periodic in time and have $u(0,y,t)=u(a,y,t)=u(x,0,t)=u(x,b,t)=0$. Write $u(x,y,t)=X(x)Y(y)S(t)$ and substitute, obtaining

$$T(X''YS+XY''S)=\rho XYS'',$$

which, when divided through by XYS, gives

$$\frac{X''}{X}+\frac{Y''}{Y}=\frac{\rho}{T}\frac{S''}{S}=-\frac{\omega^2\rho}{T}.$$

The second equality, obtained by applying the separation of variables principle with separation constant $-\omega^2\rho/T$, gives $S(t)$ as a sinusoidal function of t of frequency ω, i.e. $A\cos(\omega t)+B\sin(\omega t)$.

We then have, on applying the separation of variables principle a second time, that

$$\frac{X''}{X}=\lambda \text{ and } \frac{Y''}{Y}=\mu, \text{ where } \lambda+\mu=-\frac{\omega^2\rho}{T}. \qquad (*)$$

These equations must also have sinusoidal solutions. This is because, since $u(0,y,t)=u(a,y,t)=u(x,0,t)=u(x,b,t)=0$, each solution has to have zeros at two different values of its argument. We are thus led to

$$X=A\sin(px) \text{ and } Y=B\sin(qx), \text{ where } p^2=-\lambda \text{ and } q^2=-\mu.$$

Further, since $u(a,y,t)=u(x,b,t)=0$, we must have $p=n\pi/a$ and $q=m\pi/b$, where n and m are integers. Putting these values back into $(*)$ gives

$$-p^2-q^2=-\frac{\omega^2\rho}{T} \quad\Rightarrow\quad \pi^2\left(\frac{n^2}{a^2}+\frac{m^2}{b^2}\right)=\frac{\omega^2\rho}{T}.$$

Hence the quoted result.

21.5 *Denoting the three terms of ∇^2 in spherical polars by ∇_r^2, ∇_θ^2, ∇_ϕ^2 in an obvious way, evaluate $\nabla_r^2 u$, etc. for the two functions given below and verify that, in each case, although the individual terms are not necessarily zero their sum $\nabla^2 u$ is zero. Identify the corresponding values of ℓ and m.*

$$\text{(a)}\ \ u(r,\theta,\phi) = \left(Ar^2 + \frac{B}{r^3}\right)\frac{3\cos^2\theta - 1}{2}.$$

$$\text{(b)}\ \ u(r,\theta,\phi) = \left(Ar + \frac{B}{r^2}\right)\sin\theta \exp i\phi.$$

In both cases we write $u(r,\theta,\phi)$ as $R(r)\Theta(\theta)\Phi(\phi)$ with

$$\nabla_r^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right),\ \nabla_\theta^2 = \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right),\ \nabla_\phi^2 = \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}.$$

(a) $u(r,\theta,\phi) = \left(Ar^2 + \dfrac{B}{r^3}\right)\dfrac{3\cos^2\theta - 1}{2}.$

$$\nabla_r^2 u = \frac{1}{r^2}\frac{\partial}{\partial r}\left(2Ar^3 - \frac{3B}{r^2}\right)\Theta = \left(6A + \frac{6B}{r^5}\right)\Theta = \frac{6u}{r^2},$$

$$\nabla_\theta^2 u = \frac{R}{r^2}\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}(-3\sin^2\theta\cos\theta) = \frac{R}{r^2}\left(\frac{-6\sin\theta\cos^2\theta + 3\sin^3\theta}{\sin\theta}\right)$$

$$= \frac{R}{r^2}(-9\cos^2\theta + 3) = -\frac{6u}{r^2},$$

$$\nabla_\phi^2 u = 0.$$

Thus, although $\nabla_r^2 u$ and $\nabla_\theta^2 u$ are not individually zero, their sum is. From $\nabla_r^2 u = \ell(\ell+1)u = 6u$, we deduce that $\ell = 2$ (or -3) and from $\nabla_\phi^2 u = 0$ that $m = 0$.

(b) $u(r,\theta,\phi) = \left(Ar + \dfrac{B}{r^2}\right)\sin\theta\, e^{i\phi}.$

$$\nabla_r^2 u = \frac{1}{r^2}\frac{\partial}{\partial r}\left(Ar^2 - \frac{2B}{r}\right)\Theta\Phi = \left(\frac{2A}{r} + \frac{2B}{r^4}\right)\Theta\Phi = \frac{2u}{r^2},$$

$$\nabla_\theta^2 u = \frac{R\Phi}{r^2}\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta\cos\theta) = \frac{R\Phi}{r^2}\left(\frac{-\sin^2\theta + \cos^2\theta}{\sin\theta}\right)$$

$$= -\frac{u}{r^2} + \frac{\cos^2\theta}{\sin^2\theta}\frac{u}{r^2},$$

$$\nabla_\phi^2 u = \frac{R\Theta}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}(\,e^{i\phi}) = -\frac{u}{r^2\sin^2\theta}.$$

Hence,

$$\nabla^2 u = \frac{2u}{r^2} - \frac{u}{r^2} + \frac{\cos^2\theta}{\sin^2\theta}\,\frac{u}{r^2} - \frac{u}{r^2\sin^2\theta} = \frac{u}{r^2}\left(1 + \frac{\cos^2\theta - 1}{\sin^2\theta}\right) = 0.$$

Here each individual term is non-zero, but their sum *is* zero. Further, $\ell(\ell+1) = 2$ and so $\ell = 1$ (or -2), and from $\nabla^2_\phi u = -u/(r^2\sin\theta)$ it follows that $m^2 = 1$. In fact, from the normal definition of spherical harmonics, $m = +1$.

21.7 *If the stream function $\psi(r,\theta)$ for the flow of a very viscous fluid past a sphere is written as $\psi(r,\theta) = f(r)\sin^2\theta$, then $f(r)$ satisfies the equation*

$$f^{(4)} - \frac{4f''}{r^2} + \frac{8f'}{r^3} - \frac{8f}{r^4} = 0.$$

At the surface of the sphere $r = a$ the velocity field $\mathbf{u} = \mathbf{0}$, whilst far from the sphere $\psi \simeq (Ur^2\sin^2\theta)/2$.

Show that $f(r)$ can be expressed as a superposition of powers of r, and determine which powers give acceptable solutions. Hence show that

$$\psi(r,\theta) = \frac{U}{4}\left(2r^2 - 3ar + \frac{a^3}{r}\right)\sin^2\theta.$$

For solutions of

$$f^{(4)} - \frac{4f''}{r^2} + \frac{8f'}{r^3} - \frac{8f}{r^4} = 0$$

that are powers of r, i.e. have the form Ar^n, n must satisfy the quartic equation

$$\begin{aligned} n(n-1)(n-2)(n-3) - 4n(n-1) + 8n - 8 &= 0,\\ (n-1)[\,n(n-2)(n-3) - 4n + 8\,] &= 0,\\ (n-1)(n-2)[\,n(n-3) - 4\,] &= 0,\\ (n-1)(n-2)(n-4)(n+1) &= 0. \end{aligned}$$

Thus the possible powers are 1, 2, 4 and -1.

Since $\psi \to \frac{1}{2}Ur^2\sin^2\theta$ as $r \to \infty$, the solution can contain no higher (positive) power of r than the second. Thus there is no $n = 4$ term and the solution has the form

$$\psi(r,\theta) = \left(\frac{Ur^2}{2} + Ar + \frac{B}{r}\right)\sin^2\theta.$$

On the surface of the sphere $r = a$ both velocity components, u_r and u_θ, are zero. These components are given in terms of the stream functions, as shown below;

note that u_r is found by differentiating with respect to θ and u_θ by differentiating with respect to r.

$$u_r = 0 \quad \Rightarrow \quad \frac{1}{a^2 \sin\theta}\frac{\partial\psi}{\partial\theta} = 0 \quad \Rightarrow \quad \frac{Ua^2}{2} + Aa + \frac{B}{a} = 0,$$
$$u_\theta = 0 \quad \Rightarrow \quad \frac{-1}{a\sin\theta}\frac{\partial\psi}{\partial r} = 0 \quad \Rightarrow \quad Ua + A - \frac{B}{a^2} = 0,$$
$$\Rightarrow \quad A = -\tfrac{3}{4}Ua \text{ and } B = \tfrac{1}{4}Ua^3.$$

The full solution is thus

$$\psi(r,\theta) = \frac{U}{4}\left(2r^2 - 3ar + \frac{a^3}{r}\right)\sin^2\theta.$$

21.9 *A circular disc of radius a is heated in such a way that its perimeter $\rho = a$ has a steady temperature distribution $A + B\cos^2\phi$, where ρ and ϕ are plane polar coordinates and A and B are constants. Find the temperature $T(\rho,\phi)$ everywhere in the region $\rho < a$.*

This is a steady state problem, for which the (heat) diffusion equation becomes the Laplace equation. The most general single-valued solution to the Lapace equation in plane polar coordinates is given by

$$T(\rho,\phi) = C\ln\rho + D + \sum_{n=1}^{\infty}(A_n\cos n\phi + B_n\sin n\phi)(C_n\rho^n + D_n\rho^{-n}).$$

The region $\rho < a$ contains the point $\rho = 0$; since $\ln\rho$ and all ρ^{-n} become infinite at that point, $C = D_n = 0$ for all n.

On $\rho = a$

$$T(a,\phi) = A + B\cos^2\phi = A + \tfrac{1}{2}B(\cos 2\phi + 1).$$

Equating the coefficients of $\cos n\phi$, including $n = 0$, gives $A + \frac{1}{2}B = D$, $A_2C_2a^2 = \frac{1}{2}B$ and $A_nC_na^n = 0$ for all $n \neq 2$; further, all $B_n = 0$. The solution everywhere (not just on the perimeter) is therefore

$$T(\rho,\phi) = A + \frac{B}{2} + \frac{B\rho^2}{2a^2}\cos 2\phi.$$

It should be noted that ‘equating coefficients’ to determine unknown constants is justified by the fact that the sinusoidal functions in the sum are mutually orthogonal over the range $0 \le \phi < 2\pi$.

21.11 *The free transverse vibrations of a thick rod satisfy the equation*

$$a^4 \frac{\partial^4 u}{\partial x^4} + \frac{\partial^2 u}{\partial t^2} = 0.$$

Obtain a solution in separated-variable form and, for a rod clamped at one end, $x = 0$, and free at the other, $x = L$, show that the angular frequency of vibration ω satisfies

$$\cosh\left(\frac{\omega^{1/2} L}{a}\right) = -\sec\left(\frac{\omega^{1/2} L}{a}\right).$$

[*At a clamped end both u and $\partial u/\partial x$ vanish, whilst at a free end, where there is no bending moment, $\partial^2 u/\partial x^2$ and $\partial^3 u/\partial x^3$ are both zero.*]

The general solution is written as the product $u(x,t) = X(x)T(t)$, which, on substitution, produces the separated equation

$$a^4 \frac{X^{(4)}}{X} = -\frac{T''}{T} = \omega^2.$$

Here the separation constant has been chosen so as to give oscillatory behaviour (in the time variable). The spatial equation then becomes

$$X^{(4)} - \mu^4 X = 0, \text{ where } \mu = \omega^{1/2}/a.$$

The required auxiliary equation is $\lambda^4 - \mu^4 = 0$, leading to the general solution

$$X(x) = A \sin \mu x + B \cos \mu x + C \sinh \mu x + D \cosh \mu x.$$

The constants A, B, C and D are to be determined by requiring $X(0) = X'(0) = 0$ and $X''(L) = X'''(L) = 0$.

At the clamped end,

$$\begin{aligned}
X(0) = 0 \quad &\Rightarrow \quad D = -B, \\
X' \quad &= \quad \mu(A \cos \mu x - B \sin \mu x + C \cosh \mu x - B \sinh \mu x), \\
X'(0) = 0 \quad &\Rightarrow \quad C = -A.
\end{aligned}$$

At the free end,

$$\begin{aligned}
X'' &= \mu^2(-A \sin \mu x - B \cos \mu x - A \sinh \mu x - B \cosh \mu x), \\
X''' &= \mu^3(-A \cos \mu x + B \sin \mu x - A \cosh \mu x - B \sinh \mu x), \\
X''(L) = 0 \quad &\Rightarrow \quad A(\sin \mu L + \sinh \mu L) + B(\cos \mu L + \cosh \mu L) = 0, \\
X'''(L) = 0 \quad &\Rightarrow \quad A(-\cos \mu L - \cosh \mu L) + B(\sin \mu L - \sinh \mu L) = 0.
\end{aligned}$$

Cross-multiplying then gives

$$-\sin^2\mu L + \sinh^2\mu L = \cos^2\mu L + 2\cos\mu L\cosh\mu L + \cosh^2\mu L,$$
$$0 = 1 + 2\cos\mu L\cosh\mu L + 1,$$
$$-1 = \cos\mu L\cosh\mu L,$$
$$\cosh\left(\frac{\omega^{1/2}L}{a}\right) = -\sec\left(\frac{\omega^{1/2}L}{a}\right).$$

Because sinusoidal and hyperbolic functions can all be written in terms of exponential functions, this problem could also be approached by assuming solutions that are (exponential) functions of linear combinations of x and t (as in Chapter 20). However, in practice, eliminating the t-dependent terms leads to involved algebra.

21.13 *A string of length L, fixed at its two ends, is plucked at its mid-point by an amount A and then released. Prove that the subsequent displacement is given by*

$$u(x,t) = \sum_{n=0}^{\infty} \frac{8A}{\pi^2(2n+1)^2} \sin\left[\frac{(2n+1)\pi x}{L}\right] \cos\left[\frac{(2n+1)\pi ct}{L}\right],$$

where, in the usual notation, $c^2 = T/\rho$.

Find the total kinetic energy of the string when it passes through its unplucked position, by calculating it in each mode (each n) and summing, using the result

$$\sum_{0}^{\infty} \frac{1}{(2n+1)^2} = \frac{\pi^2}{8}.$$

Confirm that the total energy is equal to the work done in plucking the string initially.

We start with the wave equation:

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = 0$$

and assume a separated-variable solution $u(x,t) = X(x)S(t)$. This leads to

$$\frac{X''}{X} = \frac{1}{c^2}\frac{S''}{S} = -k^2.$$

The solution to the spatial equation is given by

$$X(x) = B\cos kx + C\sin kx.$$

Taking the string as anchored at $x = 0$ and $x = L$, we must have $B = 0$ and k constrained by $\sin kL = 0 \Rightarrow k = n\pi/L$ with n an integer.

The solution to the corresponding temporal equation is

$$S(t) = D\cos kct + E\sin kct.$$

Since there is no initial motion, i.e. $\dot{S}(0) = 0$, it follows that $E = 0$.
For any particular value of k, the constants C and D can be amalgamated. The general solution is given by a superposition of the allowed functions, i.e.

$$u(x,t) = \sum_{n=1}^{\infty} C_n \sin\frac{n\pi x}{L}\cos\frac{n\pi ct}{L}.$$

We now have to determine the C_n by making $u(x,0)$ match the given initial configuration, which is

$$u(x,0) = \begin{cases} \dfrac{2Ax}{L} & \text{for } 0 \le x \le \dfrac{L}{2}, \\ \dfrac{2A(L-x)}{L} & \dfrac{L}{2} < x \le L. \end{cases}$$

This is now a Fourier series calculation yielding

$$\begin{aligned} \frac{C_n L}{2} &= \int_0^{L/2} \frac{2Ax}{L}\sin\frac{n\pi x}{L}\,dx + \int_{L/2}^{L} \frac{2A(L-x)}{L}\sin\frac{n\pi x}{L}\,dx \\ &= \frac{2A}{L}J_1 + 2AJ_2 - \frac{2A}{L}J_3, \end{aligned}$$

with

$$\begin{aligned} J_1 &= \left[-\frac{xL}{n\pi}\cos\frac{n\pi x}{L}\right]_0^{L/2} + \int_0^{L/2}\frac{L}{n\pi}\cos\frac{n\pi x}{L}\,dx \\ &= -\frac{L^2}{2\pi n}\cos\frac{n\pi}{2} + \frac{L^2}{n^2\pi^2}\sin\frac{n\pi}{2}, \\ J_2 &= \int_{L/2}^{L}\sin\frac{n\pi x}{L}\,dx = -\frac{L}{n\pi}\left[\cos\frac{n\pi x}{L}\right]_{L/2}^{L} \\ &= -\frac{L}{n\pi}\left[(-1)^n - \cos\frac{n\pi}{2}\right], \\ J_3 &= \left[-\frac{xL}{n\pi}\cos\frac{n\pi x}{L}\right]_{L/2}^{L} + \int_{L/2}^{L}\frac{L}{n\pi}\cos\frac{n\pi x}{L}\,dx \\ &= \frac{L^2}{2\pi n}\cos\frac{n\pi}{2} - \frac{L^2}{n\pi}(-1)^n - \frac{L^2}{n^2\pi^2}\sin\frac{n\pi}{2}. \end{aligned}$$

Thus

$$\begin{aligned} J_1 - J_3 &= -\frac{2L^2}{2\pi n}\cos\frac{n\pi}{2} + \frac{L^2}{n\pi}(-1)^n + \frac{2L^2}{n^2\pi^2}\sin\frac{n\pi}{2} \\ &= -LJ_2 + \frac{2L^2}{n^2\pi^2}\sin\frac{n\pi}{2}, \end{aligned}$$

and so it follows that

$$\frac{C_n L}{2} = \frac{2A}{L}(J_1 - J_3 + LJ_2) = 2A\frac{2L}{n^2\pi^2}\sin\frac{n\pi}{2}.$$

This is zero if n is even and $C_n = 8A(-1)^{(n-1)/2}/(n^2\pi^2)$ if n is odd. Write $n = 2m+1$, $m = 0, 1, 2, \ldots$, with $C_{2m+1} = \dfrac{8A(-1)^m}{(2m+1)^2\pi^2}$.

The final solution (in which m is replaced by n, to match the question) is thus

$$u(x,t) = \sum_{n=0}^{\infty} \frac{8A(-1)^n}{\pi^2(2n+1)^2} \sin\left[\frac{(2n+1)\pi x}{L}\right] \cos\left[\frac{(2n+1)\pi ct}{L}\right].$$

The velocity profile derived from this is given by

$$\begin{aligned}\dot{u}(x,t) = \sum_{n=0}^{\infty} &\frac{8A(-1)^n}{\pi^2(2n+1)^2}\left(\frac{-(2n+1)\pi c}{L}\right)\\ &\times\ \sin\left[\frac{(2n+1)\pi x}{L}\right] \sin\left[\frac{(2n+1)\pi ct}{L}\right],\end{aligned}$$

giving the energy in the $(2n+1)$th mode (evaluated when the time-dependent sine function is maximal) as

$$\begin{aligned}E_{2n+1} &= \int_0^L \tfrac{1}{2}\rho\dot{u}_n^2\,dx\\ &= \int_0^L \frac{\rho}{2}\,\frac{(8A)^2c^2}{L^2(2n+1)^2\pi^2}\sin^2\frac{(2n+1)\pi x}{L}\\ &= \frac{32A^2\rho c^2}{L^2(2n+1)^2\pi^2}\,\frac{L}{2}.\end{aligned}$$

Therefore

$$E = \sum_{n=0}^{\infty} E_{2n+1} = \frac{16A^2\rho c^2}{\pi^2 L}\sum_{n=0}^{\infty}\frac{1}{(2n+1)^2} = \frac{2A^2\rho c^2}{L}.$$

When the mid-point of the string has been displaced sideways by y ($\ll L$), the net (resolved) restoring force is $2T[\,y/(L/2)\,] = 4Ty/L$. Thus the total work done to produce a displacement of A is

$$W = \int_0^A \frac{4Ty}{L}\,dy = \frac{2TA^2}{L} = \frac{2\rho c^2 A^2}{L},$$

i.e. the same as the total energy of the subsequent motion.

21.15 *Prove that the potential for $\rho < a$ associated with a vertical split cylinder of radius a, the two halves of which ($\cos\phi > 0$ and $\cos\phi < 0$) are maintained at equal and opposite potentials $\pm V$, is given by*

$$u(\rho, \phi) = \frac{4V}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \left(\frac{\rho}{a}\right)^{2n+1} \cos(2n+1)\phi.$$

The most general solution of the Laplace equation in cylindrical polar coordinates that is independent of z is

$$T(\rho, \phi) = C \ln \rho + D + \sum_{n=1}^{\infty} (A_n \cos n\phi + B_n \sin n\phi)(C_n \rho^n + D_n \rho^{-n}).$$

The required potential must be single-valued and finite in the space inside the cylinder (which includes $\rho = 0$), and on the cylinder it must take the boundary values $u = V$ for $\cos\phi > 0$ and $u = -V$ for $\cos\phi < 0$, i.e the boundary-value function is a square-wave function with average value zero. Although the function is antisymmetric in $\cos\phi$, it is symmetric in ϕ and so the solution will contain only cosine terms (and no sine terms).

These considerations already determine that $C = D = B_n = D_n = 0$, and so have reduced the solution to the form

$$u(\rho, \phi) = \sum_{n=1}^{\infty} A_n \rho^n \cos n\phi.$$

On $\rho = a$ this must match the stated boundary conditions, and so we are faced with a Fourier cosine series calculation. Multiplying through by $\cos m\phi$ and integrating yields

$$\begin{aligned}
A_m a^m \frac{1}{2} 2\pi &= 2 \int_0^{\pi/2} V \cos m\phi \, d\phi + 2 \int_{\pi/2}^{\pi} (-V) \cos m\phi \, d\phi \\
&= 2V \left[\frac{\sin m\phi}{m} \right]_0^{\pi/2} - 2V \left[\frac{\sin m\phi}{m} \right]_{\pi/2}^{\pi} \\
&= \frac{2V}{m} \left(\sin \frac{m\pi}{2} + \sin \frac{m\pi}{2} \right) \\
&= (-1)^{(m-1)/2} \frac{4V}{m} \text{ for } m \text{ odd}, = 0 \text{ for } m \text{ even.}
\end{aligned}$$

Writing $m = 2n + 1$ gives the solution as

$$u(\rho, \phi) = \frac{4V}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \left(\frac{\rho}{a}\right)^{2n+1} \cos(2n+1)\phi.$$

21.17 *Two identical copper bars are each of length a. Initially, one is at* 0 °C *and the other at* 100 °C*; they are then joined together end to end and thermally isolated. Obtain in the form of a Fourier series an expression $u(x,t)$ for the temperature at any point a distance x from the join at a later time t. Bear in mind the heat flow conditions at the free ends of the bars.*

Taking $a = 0.5$ m estimate the time it takes for one of the free endsto attain a temperature of 55 °C. *The thermal conductivity of copper is* 3.8×10^2 J m^{-1} K^{-1} s^{-1}, *and its specific heat capacity is* 3.4×10^6 J m^{-3} K^{-1}.

The equation governing the heat flow is

$$k\frac{\partial^2 u}{\partial x^2} = s\frac{\partial u}{\partial t},$$

which is the diffusion equation with diffusion constant $\kappa = k/s = 3.8 \times 10^2/3.4 \times 10^6 = 1.12 \times 10^{-4}$ m^2 s^{-1}.

Making the usual separation of variables substitution shows that the time variation is of the form $T(t) = T(0)e^{-\kappa\lambda^2 t}$ when the spatial solution is a sinusoidal function of λx. The final common temperature is 50 °C and we make this explicit by writing the general solution as

$$u(x,t) = 50 + \sum_{\lambda}(A_\lambda \sin \lambda x + B_\lambda \cos \lambda x)e^{-\kappa\lambda^2 t}.$$

This term having been taken out, the summation must be antisymmetric about $x = 0$ and therefore contain no cosine terms, i.e. $B_\lambda = 0$.

The boundary condition is that there is no heat flow at $x = \pm a$; this means that $\partial u/\partial x = 0$ at these points and requires

$$\lambda A_\lambda \cos \lambda x|_{x=\pm a} = 0 \quad \Rightarrow \quad \lambda a = (n + \tfrac{1}{2})\pi \quad \Rightarrow \quad \lambda = \frac{(2n+1)\pi}{2a},$$

where n is an integer. This corresponds to a fundamental Fourier period of $4a$. The solution thus takes the form

$$u(x,t) = 50 + \sum_{n=0}^{\infty} A_n \sin \frac{(2n+1)\pi x}{2a} \exp\left(-\frac{(2n+1)^2\pi^2\kappa t}{4a^2}\right).$$

At $t = 0$, the sum must take the values +50 for $0 < x < 2a$ and −50 for $-2a < x < 0$. This is (yet) another square-wave function — one that is antisymmetric about $x = 0$ and has amplitude 50. The calculation will not be repeated here but gives $A_n = 200/[(2n+1)\pi]$, making the complete solution

$$u(x,t) = 50 + \frac{200}{\pi}\sum_{n=0}^{\infty} \frac{1}{2n+1} \sin \frac{(2n+1)\pi x}{2a} \exp\left(-\frac{(2n+1)^2\pi^2\kappa t}{4a^2}\right).$$

For a free end, where $x = a$ and $\sin[(2n+1)\pi x/2a] = (-1)^n$, to attain 55 °C needs

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \exp\left(-\frac{(2n+1)^2\,\pi^2\,1.12\times 10^{-4}}{4\times 0.25}\,t\right) = \frac{5\pi}{200} = 0.0785.$$

In principle this is an insoluble equation but, because the RHS $\ll 1$, the $n = 0$ term alone will give a good approximation to t:

$$\exp(-1.105\times 10^{-3}t) \approx 0.0785 \quad \Rightarrow \quad t \approx 2300 \text{ s}.$$

21.19 *For an infinite metal bar that has an initial ($t = 0$) temperature distribution $f(x)$ along its length, the temperature distribution at a general time t can be shown to be given by*

$$u(x,t) = \frac{1}{\sqrt{4\pi\kappa t}} \int_{-\infty}^{\infty} \exp\left[-\frac{(x-\xi)^2}{4\kappa t}\right] f(\xi)\, d\xi.$$

Find an explicit expression for $u(x,t)$ given that $f(x) = \exp(-x^2/a^2)$.

The given initial distribution is $f(\xi) = \exp(-\xi^2/a^2)$ and so

$$u(x,t) = \frac{1}{\sqrt{4\pi\kappa t}} \int_{-\infty}^{\infty} \exp\left[-\frac{(x-\xi)^2}{4\kappa t}\right] \exp\left(-\frac{\xi^2}{a^2}\right) d\xi.$$

Now consider the exponent in the integrand, writing $1 + \dfrac{4\kappa t}{a^2}$ as τ^2 for compactness:

$$\begin{aligned}
\text{exponent} &= -\frac{\xi^2\tau^2 - 2\xi x + x^2}{4\kappa t} \\
&= -\frac{(\xi\tau - x\tau^{-1})^2 - x^2\tau^{-2} + x^2}{4\kappa t} \\
&\equiv -\eta^2 + \frac{x^2\tau^{-2} - x^2}{4\kappa t}, \quad \text{defining } \eta, \\
\text{with } d\eta &= \frac{\tau\, d\xi}{\sqrt{4\kappa t}}.
\end{aligned}$$

With a change of variable from ξ to η, the integral becomes

$$\begin{aligned}
u(x,t) &= \frac{1}{\sqrt{4\pi\kappa t}} \exp\left(\frac{x^2\tau^{-2} - x^2}{4\kappa t}\right) \int_{-\infty}^{\infty} \exp(-\eta^2)\, \frac{\sqrt{4\kappa t}}{\tau}\, d\eta \\
&= \frac{1}{\sqrt{\pi}} \frac{1}{\tau} \exp\left(x^2 \frac{1-\tau^2}{4\kappa t\,\tau^2}\right) \sqrt{\pi} \\
&= \frac{a}{\sqrt{a^2 + 4\kappa t}} \exp\left(-\frac{x^2}{a^2 + 4\kappa t}\right).
\end{aligned}$$

In words, although it retains a Gaussian shape, the initial distribution spreads symmetrically about the origin, its variance increasing linearly with time ($a^2 \to a^2 + 4\kappa t$). As is typical with diffusion processes, for large enough times the width varies as $\sqrt{t}$.

21.21 *In the region $-\infty < x,\ y < \infty$ and $-t \le z \le t$, a charge-density wave $\rho(\mathbf{r}) = A\cos qx$, in the x-direction, is represented by*

$$\rho(\mathbf{r}) = \frac{e^{iqx}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{\rho}(\alpha) e^{i\alpha z}\, d\alpha.$$

The resulting potential is represented by

$$V(\mathbf{r}) = \frac{e^{iqx}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{V}(\alpha) e^{i\alpha z}\, d\alpha.$$

Determine the relationship between $\tilde{V}(\alpha)$ and $\tilde{\rho}(\alpha)$, and hence show that the potential at the point $(0,0,0)$ is given by

$$\frac{A}{\pi\epsilon_0} \int_{-\infty}^{\infty} \frac{\sin kt}{k(k^2+q^2)}\, dk.$$

Poisson's equation,

$$\nabla^2 V(\mathbf{r}) = -\frac{\rho(\mathbf{r})}{\epsilon_0},$$

provides the link between a charge density and the potential it produces.

Taking $V(\mathbf{r})$ in the form of its Fourier representation gives $\nabla^2 V$ as

$$\frac{\partial^2 V(\mathbf{r})}{\partial x^2} + \frac{\partial^2 V(\mathbf{r})}{\partial y^2} + \frac{\partial^2 V(\mathbf{r})}{\partial z^2} = \frac{e^{iqx}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (-q^2 - \alpha^2) \tilde{V}(\alpha) e^{i\alpha z}\, d\alpha,$$

with the $-q^2$ arising from the x-differentiation and the $-\alpha^2$ from the z-differentiation; the $\partial^2 V/\partial y^2$ term contributes nothing.

Comparing this with the integral expression for $-\rho(\mathbf{r})/\epsilon_0$ shows that

$$-\tilde{\rho}(\alpha) = \epsilon_0(-q^2 - \alpha^2)\tilde{V}(\alpha).$$

With the charge-density wave confined in the z-direction to $-t \le z \le t$, the expression for $\rho(\mathbf{r})$ in Cartesian coordinates is (in terms of Heaviside functions)

$$\rho(\mathbf{r}) = Ae^{iqx}[\,H(z+t) - H(z-t)\,].$$

The Fourier transform $\tilde{\rho}(\alpha)$ is therefore given by

$$\begin{aligned}\tilde{\rho}(\alpha) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} A[\,H(z+t) - H(z-t)\,]\, e^{-i\alpha z}\, dz \\ &= \frac{A}{\sqrt{2\pi}} \int_{-t}^{t} e^{-i\alpha z}\, dz \\ &= \frac{A}{\sqrt{2\pi}} \frac{e^{-i\alpha t} - e^{i\alpha t}}{-i\alpha} \\ &= \frac{A}{\sqrt{2\pi}} \frac{2\sin\alpha t}{\alpha}.\end{aligned}$$

Now,

$$\begin{aligned}V(x,0,z) &= \frac{e^{iqx}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{\tilde{\rho}(\alpha)}{\epsilon_0(q^2+\alpha^2)}\, e^{i\alpha z}\, d\alpha \\ &= \frac{e^{iqx}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{e^{i\alpha z}}{\epsilon_0(q^2+\alpha^2)} \frac{A}{\sqrt{2\pi}} \frac{2\sin\alpha t}{\alpha}\, d\alpha, \\ \Rightarrow \quad V(0,0,0) &= \frac{A}{\pi\epsilon_0} \int_{-\infty}^{\infty} \frac{\sin\alpha t}{\alpha(\alpha^2+q^2)}\, d\alpha,\end{aligned}$$

as stated in the question.

21.23 *Find the Green's function $G(\mathbf{r},\mathbf{r}_0)$ in the half-space $z>0$ for the solution of $\nabla^2\Phi = 0$ with Φ specified in cylindrical polar coordinates (ρ,ϕ,z) on the plane $z=0$ by*

$$\Phi(\rho,\phi,z) = \begin{cases} 1 & \text{for } \rho \le 1, \\ 1/\rho & \text{for } \rho > 1. \end{cases}$$

Determine the variation of $\Phi(0,0,z)$ along the z-axis.

For the half-space $z>0$ the bounding surface consists of the plane $z=0$ and the (hemi-spherical) surface at infinity; the Green's function must take zero value on these surfaces. In order to ensure this when a unit point source is introduced at $\mathbf{r} = \mathbf{y}$, we must place a compensating negative unit source at $\mathbf{y}$'s reflection point in the plane. If, in cylindrical polar coordinates, $\mathbf{y} = (\rho,\phi,z_0)$, then the image charge has to be at $\mathbf{y}' = (\rho,\phi,-z_0)$. The resulting Green's function $G(\mathbf{x},\mathbf{y})$ is given by

$$G(\mathbf{x},\mathbf{y}) = -\frac{1}{4\pi|\mathbf{x}-\mathbf{y}|} + \frac{1}{4\pi|\mathbf{x}-\mathbf{y}'|}.$$

The solution to the problem with a given potential distribution $f(\rho,\phi)$ on the $z=0$ part of the bounding surface S is given by

$$\Phi(\mathbf{y}) = \int_S f(\rho,\phi)\left(-\frac{\partial G}{\partial z}\right) \rho\, d\phi\, d\rho,$$

the minus sign arising because the outward normal to the region is in the negative z-direction. Calculating these functions explicitly gives

$$\begin{aligned}
G(\mathbf{x},\mathbf{y}) &= -\frac{1}{4\pi[\,\rho^2+(z-z_0)^2\,]^{1/2}}+\frac{1}{4\pi[\,\rho^2+(z+z_0)^2\,]^{1/2}},\\
\frac{\partial G}{\partial z} &= \frac{z-z_0}{4\pi[\,\rho^2+(z-z_0)^2\,]^{3/2}}-\frac{(z+z_0)}{4\pi[\,\rho^2+(z+z_0)^2\,]^{3/2}},\\
-\left.\frac{\partial G}{\partial z}\right|_{z=0} &= -\frac{-2z_0}{4\pi[\,\rho^2+z_0^2\,]^{3/2}}.
\end{aligned}$$

Substituting the various factors into the general integral gives

$$\begin{aligned}
\Phi(0,0,z_0) &= \int_0^\infty f(\rho)\frac{2z_0}{4\pi[\,\rho^2+z_0^2\,]^{3/2}}\,2\pi\,\rho\,d\rho\\
&= \int_0^1 \frac{z_0\,\rho}{(\rho^2+z_0^2)^{3/2}}\,d\rho+\int_1^\infty \frac{z_0}{(\rho^2+z_0^2)^{3/2}}\,d\rho\\
&= -z_0\left[\,(\rho^2+z_0^2)^{-1/2}\,\right]_0^1+\int_\theta^{\pi/2}\frac{z_0^2\sec^2 u}{z_0^3\sec^3 u}\,du,
\end{aligned}$$

where, in the second integral, we have set $\rho = z_0\tan u$ with $d\rho = z_0\sec^2 u\,du$ and $\theta = \tan^{-1}(1/z_0)$. The integral can now be obtained in closed form as

$$\begin{aligned}
\Phi(0,0,z_0) &= -\frac{z_0}{(1+z_0^2)^{1/2}}+1+\frac{1}{z_0}\,[\,\sin u\,]_\theta^{\pi/2}\\
&= 1-\frac{z_0}{(1+z_0^2)^{1/2}}+\frac{1}{z_0}-\frac{1}{z_0(1+z_0^2)^{1/2}}.
\end{aligned}$$

Thus the variation of Φ along the z-axis is given by

$$\Phi(0,0,z) = \frac{z(1+z^2)^{1/2}-z^2+(1+z^2)^{1/2}-1}{z(1+z^2)^{1/2}}.$$

21.25 *Find, in the form of an infinite series, the Green's function of the ∇^2 operator for the Dirichlet problem in the region $-\infty < x < \infty$, $-\infty < y < \infty$, $-c \le z \le c$.*

The fundamental solution in three dimensions of $\nabla^2\psi = \delta(\mathbf{r})$ is $\psi(\mathbf{r}) = -1/(4\pi r)$.

For the given problem, $G(\mathbf{r},\mathbf{r}_0)$ has to take the value zero on $z=\pm c$ and $\to 0$ for $|x|\to\infty$ and $|y|\to\infty$. Image charges have to be added in the regions $z>c$ and $z<-c$ to bring this about after a charge q has been placed at $\mathbf{r}_0 = (x_0,y_0,z_0)$ with $-c<z_0<c$. Clearly all images will be on the line $x=x_0$, $y=y_0$.

Each image placed at $z=\xi$ in the region $z>c$ will require a further image of the same strength but opposite sign at $z=-c-\xi$ (in the region $z<-c$) so as to maintain the plane $z = -c$ as an equipotential. Likewise, each image placed at $z = -\chi$ in the region $z < -c$ will require a further image of the same strength but opposite sign at $z = c + \chi$ (in the region $z > c$) so as to maintain the plane $z = c$ as an equipotential. Thus succesive image charges appear as follows:

$$\begin{array}{ccc} -q & 2c - z_0 & -2c - z_0 \\ +q & -3c + z_0 & 3c + z_0 \\ -q & 4c - z_0 & -4c - z_0 \\ +q & \text{etc.} & \text{etc.} \end{array}$$

The terms in the Green's function that are additional to the fundamental solution,

$$-\frac{1}{4\pi}[(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2]^{-1/2},$$

are therefore

$$-\frac{(-1)}{4\pi}\sum_{n=2}^{\infty}\left\{\frac{(-1)^n}{[\,(x - x_0)^2 + (y - y_0)^2 + (z + (-1)^n z_0 - nc)^2\,]^{1/2}}\right.$$
$$\left. + \quad \frac{(-1)^n}{[\,(x - x_0)^2 + (y - y_0)^2 + (z + (-1)^n z_0 + nc)^2\,]^{1/2}}\right\}.$$

21.27 *Determine the Green's function for the Klein–Gordon equation in a half-space as follows.*

(a) *By applying the divergence theorem to the volume integral*

$$\int_V \left[\phi(\nabla^2 - m^2)\psi - \psi(\nabla^2 - m^2)\phi\right]\, dV,$$

obtain a Green's function expression, as the sum of a volume integral and a surface integral, for the function $\phi(\mathbf{r}')$ that satisfies

$$\nabla^2\phi - m^2\phi = \rho$$

in V and takes the specified form $\phi = f$ on S, the boundary of V. The Green's function, $G(\mathbf{r}, \mathbf{r}')$, to be used satisfies

$$\nabla^2 G - m^2 G = \delta(\mathbf{r} - \mathbf{r}')$$

and vanishes when $\mathbf{r}$ is on S.

(b) *When V is all space, $G(\mathbf{r}, \mathbf{r}')$ can be written as $G(t) = g(t)/t$, where $t = |\mathbf{r} - \mathbf{r}'|$ and $g(t)$ is bounded as $t \to \infty$. Find the form of $G(t)$.*

(c) *Find $\phi(\mathbf{r})$ in the half-space $x > 0$ if $\rho(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}_1)$ and $\phi = 0$ both on $x = 0$ and as $r \to \infty$.*

(a) For general ϕ and ψ we have

$$\begin{aligned}\int_V \left[\phi(\nabla^2 - m^2)\psi - \psi(\nabla^2 - m^2)\phi\right] \, dV &= \int_V \left[\phi\nabla^2\psi - \psi\nabla^2\phi\right] \, dV \\ &= \int_V \nabla \cdot (\phi\nabla\psi - \psi\nabla\phi) \, dV \\ &= \int_S (\phi\nabla\psi - \psi\nabla\phi) \cdot \mathbf{n} \, dS.\end{aligned}$$

Now take ϕ as ϕ, with $\nabla^2\phi - m^2\phi = \rho$ and $\phi = f$ on the surface S, and ψ as $G(\mathbf{r}, \mathbf{r}')$ with $\nabla^2 G - m^2 G = \delta(\mathbf{r} - \mathbf{r}')$ and $G(\mathbf{r}, \mathbf{r}') = 0$ on S:

$$\int_V [\, \phi(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}') - G(\mathbf{r}, \mathbf{r}')\rho(\mathbf{r}) \,] \, dV = \int_S [\, f(\mathbf{r})\nabla G(\mathbf{r}, \mathbf{r}') - \mathbf{0} \,] \cdot \mathbf{n} \, dS,$$

which, on rearrangement, gives

$$\phi(\mathbf{r}') = \int_V G(\mathbf{r}, \mathbf{r}')\rho(\mathbf{r}) \, dV + \int_S f(\mathbf{r})\nabla G(\mathbf{r}, \mathbf{r}') \cdot \mathbf{n} \, dS.$$

(b) In the following calculation we start by formally integrating the defining Green's equation,

$$\nabla^2 G - m^2 G = \delta(\mathbf{r} - \mathbf{r}'),$$

over a sphere of radius t centred on $\mathbf{r}'$. Having replaced the volume integral of $\nabla^2 G$ with the corresponding surface integral given by the divergence theorem, we move the origin to $\mathbf{r}'$, denote $|\mathbf{r} - \mathbf{r}'|$ by t' and integrate both sides of the equation from $t' = 0$ to $t' = t$:

$$\begin{aligned}\int_V \nabla^2 G \, dV - \int_V m^2 G \, dV &= \int_V \delta(\mathbf{r} - \mathbf{r}') \, dV, \\ \int_S \nabla G \cdot \mathbf{n} \, dS - m^2 \int_V G \, dV &= 1, \\ 4\pi t^2 \frac{dG}{dt} - m^2 \int_0^t G(t') 4\pi t'^2 \, dt' &= 1, \qquad (*) \\ 4\pi t^2 G'' + 8\pi t G' - 4\pi m^2 t^2 G &= 0, \text{ from differentiating w.r.t. } t, \\ tG'' + 2G' - m^2 tG &= 0.\end{aligned}$$

With $G(t) = g(t)/t$,

$$G' = -\frac{g}{t^2} + \frac{g'}{t} \quad \text{and} \quad G'' = \frac{2g}{t^3} - \frac{2g'}{t^2} + \frac{g''}{t},$$

and the equation becomes

$$\begin{aligned}0 &= \frac{2g}{t^2} - \frac{2g'}{t} + g'' - \frac{2g}{t^2} + \frac{2g'}{t} - m^2 g, \\ 0 &= g'' - m^2 g, \\ \Rightarrow \quad g(t) &= Ae^{-mt}, \text{ since } g \text{ is bounded as } t \to \infty.\end{aligned}$$

The value of A is determined by resubstituting into $(*)$, which then reads

$$\begin{aligned}4\pi t^2\left(-\frac{Ae^{-mt}}{t^2}-\frac{mAe^{-mt}}{t}\right)-m^2\int_0^t\frac{Ae^{-mt'}}{t'}\,4\pi t'^2\,dt'&=1,\\ -4\pi Ae^{-mt}(1+mt)-4\pi Am^2\left(-\frac{te^{-mt}}{m}+\frac{1-e^{-mt}}{m^2}\right)&=1,\\ -4\pi A&=1,\end{aligned}$$

making the solution

$$G(\mathbf{r},\mathbf{r}')=-\frac{e^{-mt}}{4\pi t},\ \text{where } t=|\mathbf{r}-\mathbf{r}'|.$$

(c) For the situation in which $\rho(\mathbf{r})=\delta(\mathbf{r}-\mathbf{r}_1)$, i.e. a unit positive charge at $\mathbf{r}_1=(x_1,y_1,z_1)$, and $\phi=0$ on the plane $x=0$, we must have a unit negative image charge at $\mathbf{r}_2=(-x_1,y_1,z_1)$. The solution in the region $x>0$ is then

$$\phi(\mathbf{r})=-\frac{1}{4\pi}\left(\frac{e^{-m|\mathbf{r}-\mathbf{r}_1|}}{|\mathbf{r}-\mathbf{r}_1|}-\frac{e^{-m|\mathbf{r}-\mathbf{r}_2|}}{|\mathbf{r}-\mathbf{r}_2|}\right).$$

22

Calculus of variations

22.1 *A surface of revolution, whose equation in cylindrical polar coordinates is $\rho = \rho(z)$, is bounded by the circles $\rho = a$, $z = \pm c$ $(a > c)$. Show that the function that makes the surface integral $I = \int \rho^{-1/2}\, dS$ stationary with respect to small variations is given by $\rho(z) = k + z^2/(4k)$, where $k = [a \pm (a^2 - c^2)^{1/2}]/2$.*

The surface element lying between z and $z + dz$ is given by

$$dS = 2\pi\rho\,[\,(d\rho)^2 + (dz)^2\,]^{1/2} = 2\pi\rho\,(1 + \rho'^2)^{1/2}\,dz$$

and the integral to be made stationary is

$$I = \int \rho^{-1/2}\, dS = 2\pi \int_{-c}^{c} \rho^{-1/2}\rho\,(1 + \rho'^2)^{1/2}\, dz.$$

The integrand $F(\rho', \rho, z)$ does not in fact contain z explicitly, and so a first integral of the E–L equation, symbolically given by $F - \rho'\partial F/\partial\rho' = k$, is

$$\rho^{1/2}(1 + \rho'^2)^{1/2} - \rho'\left[\frac{\rho^{1/2}\rho'}{(1 + \rho'^2)^{1/2}}\right] = A,$$

$$\frac{\rho^{1/2}}{(1 + \rho'^2)^{1/2}} = A.$$

On rearrangement and subsequent integration this gives

$$\frac{d\rho}{dz} = \left(\frac{\rho - A^2}{A^2}\right)^{1/2},$$

$$\int \frac{d\rho}{\sqrt{\rho - A^2}} = \int \frac{dz}{A},$$

$$2\sqrt{\rho - A^2} = \frac{z}{A} + C.$$

Now, $\rho(\pm c) = a$ implies both that $C = 0$ and that $a - A^2 = \dfrac{c^2}{4A^2}$. Thus, writing A^2 as k,

$$4k^2 - 4ka + c^2 = 0 \quad \Rightarrow \quad k = \tfrac{1}{2}[\,a \pm (a^2 - c^2)^{1/2}\,].$$

The two stationary functions are therefore

$$\rho = \frac{z^2}{4k} + k,$$

with k as given above. A simple sketch shows that the positive sign in k corresponds to a smaller value of the integral.

22.3 *The refractive index n of a medium is a function only of the distance r from a fixed point O. Prove that the equation of a light ray, assumed to lie in a plane through O, travelling in the medium satisfies (in plane polar coordinates)*

$$\frac{1}{r^2}\left(\frac{dr}{d\phi}\right)^2 = \frac{r^2}{a^2}\frac{n^2(r)}{n^2(a)} - 1,$$

where a is the distance of the ray from O at the point at which $dr/d\phi = 0$.

If $n = [1 + (\alpha^2/r^2)]^{1/2}$ and the ray starts and ends far from O, find its deviation (the angle through which the ray is turned), if its minimum distance from O is a.

An element of path length is $ds = [\,(dr)^2 + (r\,d\phi)^2\,]^{1/2}$ and the time taken for the light to traverse it is $n(r)\,ds/c$, where c is the speed of light *in vacuo*. Fermat's principle then implies that the light follows the curve that minimises

$$T = \int \frac{n(r)\,ds}{c} = \int \frac{n(r'^2 + r^2)^{1/2}}{c}\,d\phi,$$

where $r' = dr/d\phi$. Since the integrand does not contain ϕ explicitly, the E–L equation integrates to (see exercise 22.1)

$$n(r'^2 + r^2)^{1/2} - r'\frac{nr'}{(r'^2 + r^2)^{1/2}} = A,$$
$$\frac{nr^2}{(r'^2 + r^2)^{1/2}} = A.$$

Since $r' = 0$ when $r = a$, $A = n(a)a^2/a$, and the equation is as follows:

$$a^2 n^2(a)(r'^2 + r^2) = n^2(r)r^4,$$
$$r'^2 = \frac{n^2(r)r^4}{n^2(a)a^2} - r^2,$$
$$\Rightarrow \quad \frac{1}{r^2}\left(\frac{dr}{d\phi}\right)^2 = \frac{n^2(r)r^2}{n^2(a)a^2} - 1.$$

If $n(r) = [\,1 + (\alpha/r)^2\,]^{1/2}$, the minimising curve satisfies

$$\begin{aligned} \left(\frac{dr}{d\phi}\right)^2 &= \frac{r^2(r^2+\alpha^2)}{a^2+\alpha^2} - r^2 \\ &= \frac{r^2(r^2-a^2)}{a^2+\alpha^2}, \\ \Rightarrow \qquad \frac{d\phi}{(a^2+\alpha^2)^{1/2}} &= \pm\frac{dr}{r\sqrt{r^2-a^2}}. \end{aligned}$$

By symmetry,

$$\begin{aligned} \frac{\Delta\phi}{(a^2+\alpha^2)^{1/2}} &\equiv \frac{\phi_{\,\text{final}} - \phi_{\,\text{initial}}}{(a^2+\alpha^2)^{1/2}} \\ &= 2\int_a^\infty \frac{dr}{r\sqrt{r^2-a^2}}, \qquad \text{set } r = a\cosh\psi, \\ &= 2\int_0^\infty \frac{a\sinh\psi}{a^2\cosh\psi\sinh\psi}\,d\psi \\ &= \frac{2}{a}\int_0^\infty \operatorname{sech}\,\psi\,d\psi, \qquad \text{set } e^\psi = z, \\ &= \frac{2}{a}\int_1^\infty \frac{z^{-1}\,dz}{\frac{1}{2}(z+z^{-1})} \\ &= \frac{2}{a}\int_1^\infty \frac{2\,dz}{z^2+1} \\ &= \frac{4}{a}\left[\,\tan^{-1} z\,\right]_1^\infty \\ &= \frac{4}{a}\left(\frac{\pi}{2} - \frac{\pi}{4}\right) = \frac{\pi}{a}. \end{aligned}$$

If the refractive index were everywhere unity ($\alpha = 0$), $\Delta\phi$ would be π (no deviation). Thus the deviation is given by

$$\frac{\pi}{a}(a^2+\alpha^2)^{1/2} - \pi.$$

22.5 *Prove the following results about general systems.*

(a) *For a system described in terms of coordinates q_i and t, show that if t does not appear explicitly in the expressions for x, y and z ($x = x(q_i, t)$, etc.) then the kinetic energy T is a homogeneous quadratic function of the $\dot{q}_i$ (it may also involve the q_i). Deduce that $\sum_i \dot{q}_i(\partial T/\partial \dot{q}_i) = 2T$.*

(b) *Assuming that the forces acting on the system are derivable from a potential V, show, by expressing dT/dt in terms of q_i and $\dot{q}_i$, that $d(T+V)/dt = 0$.*

To save space we will use the summation convention for summing over the index of the q_i.

(a) The space variables x, y and z are not explicit functions of t and the kinetic energy, T, is given by

$$\begin{aligned}T &= \tfrac{1}{2}(\alpha_x \dot{x}^2 + \alpha_y \dot{y}^2 + \alpha_z \dot{z}^2)\\ &= \frac{1}{2}\left[\alpha_x\left(\frac{\partial x}{\partial q_i}\dot{q}_i\right)^2 + \alpha_y\left(\frac{\partial y}{\partial q_j}\dot{q}_j\right)^2 + \alpha_z\left(\frac{\partial z}{\partial q_k}\dot{q}_k\right)^2\right]\\ &= A_{mn}\dot{q}_m\dot{q}_n,\end{aligned}$$

with

$$A_{mn} = \frac{1}{2}\left(\alpha_x \frac{\partial x}{\partial q_m}\frac{\partial x}{\partial q_n} + \alpha_y \frac{\partial y}{\partial q_m}\frac{\partial y}{\partial q_n} + \alpha_z \frac{\partial z}{\partial q_m}\frac{\partial z}{\partial q_n}\right) = A_{nm}.$$

Hence T is a homogeneous quadratic function of the $\dot{q}_i$ (though the A_{mn} may involve the q_i). Further,

$$\frac{\partial T}{\partial \dot{q}_i} = A_{in}\dot{q}_n + A_{mi}\dot{q}_m = 2A_{mi}\dot{q}_m$$

$$\text{and} \quad \dot{q}_i \frac{\partial T}{\partial \dot{q}_i} = 2\dot{q}_i A_{mi}\dot{q}_m = 2T.$$

(b) The Lagrangian is $L = T - V$, with $T = T(q_i, \dot{q}_i)$ and $V = V(q_i)$. Thus

$$\frac{dT}{dt} = \frac{\partial T}{\partial q_i}\dot{q}_i + \frac{dT}{d\dot{q}_i}\ddot{q}_i \quad \text{and} \quad \frac{dV}{dt} = \frac{\partial V}{\partial q_i}\dot{q}_i. \qquad (*)$$

Hamilton's principle requires that

$$\begin{aligned}\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) &= \frac{\partial L}{\partial q_i},\\ \Rightarrow \quad \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_i}\right) &= \frac{\partial T}{\partial q_i} - \frac{\partial V}{\partial q_i}. \qquad (**)\end{aligned}$$

But, from part (a),

$$\begin{aligned}2T &= \dot{q}_i \frac{\partial T}{\partial \dot{q}_i},\\ \frac{d}{dt}(2T) &= \ddot{q}_i \frac{\partial T}{\partial \dot{q}_i} + \dot{q}_i \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_i}\right)\\ &= \ddot{q}_i \frac{\partial T}{\partial \dot{q}_i} + \dot{q}_i \frac{\partial T}{\partial q_i} - \dot{q}_i \frac{\partial V}{\partial q_i}, \text{ using } (**),\\ &= \frac{dT}{dt} - \frac{dV}{dt}, \qquad \text{using } (*).\end{aligned}$$

This can be rearranged as

$$\frac{d}{dt}(T + V) = 0.$$

22.7 *In cylindrical polar coordinates, the curve* $(\rho(\theta), \theta, \alpha\rho(\theta))$ *lies on the surface of the cone* $z = \alpha\rho$. *Show that geodesics (curves of minimum length joining two points) on the cone satisfy*

$$\rho^4 = c^2[\beta^2\rho'^2 + \rho^2],$$

where c *is an arbitrary constant, but* β *has to have a particular value. Determine the form of* $\rho(\theta)$ *and hence find the equation of the shortest path on the cone between the points* $(R, -\theta_0, \alpha R)$ *and* $(R, \theta_0, \alpha R)$.

[*You will find it useful to determine the form of the derivative of* $\cos^{-1}(u^{-1})$.]

In cylindrical polar coordinates the element of length is given by

$$(ds)^2 = (d\rho)^2 + (\rho\, d\theta)^2 + (dz)^2,$$

and the total length of a curve between two points parameterised by θ_0 and θ_1 is

$$\begin{aligned} s &= \int_{\theta_0}^{\theta_1} \sqrt{\left(\frac{d\rho}{d\theta}\right)^2 + \rho^2 + \left(\frac{dz}{d\theta}\right)^2}\, d\theta \\ &= \int_{\theta_0}^{\theta_1} \sqrt{\rho^2 + (1+\alpha^2)\left(\frac{d\rho}{d\theta}\right)^2}\, d\theta, \ \text{since } z = \alpha\rho. \end{aligned}$$

Since the independent variable θ does not occur explicitly in the integrand, a first integral of the E–L equation is

$$\sqrt{\rho^2 + (1+\alpha^2)\rho'^2} - \rho' \frac{(1+\alpha^2)\rho'}{\sqrt{\rho^2 + (1+\alpha^2)\rho'^2}} = c.$$

After being multiplied through by the square root, this can be arranged as follows:

$$\begin{aligned} \rho^2 + (1+\alpha^2)\rho'^2 - (1+\alpha^2)\rho'^2 &= c\sqrt{\rho^2 + (1+\alpha^2)\rho'^2}, \\ \rho^4 &= c^2[\,\rho^2 + (1+\alpha^2)\rho'^2\,]. \end{aligned}$$

This is the given equation of the geodesic, in which c is arbitrary but β^2 must have the value $1 + \alpha^2$.

Guided by the hint, we first determine the derivative of $y(u) = \cos^{-1}(u^{-1})$:

$$\frac{dy}{du} = \frac{-1}{\sqrt{1-u^{-2}}}\,\frac{-1}{u^2} = \frac{1}{u\sqrt{u^2-1}}.$$

Now, returning to the geodesic, rewrite it as

$$\begin{aligned} \rho^4 - c^2\rho^2 &= c^2\beta^2\rho'^2, \\ \rho(\rho^2 - c^2)^{1/2} &= c\beta\frac{d\rho}{d\theta}. \end{aligned}$$

Setting $\rho = cu$,

$$uc^2(u^2-1)^{1/2} = c^2\beta\frac{du}{d\theta},$$
$$d\theta = \frac{\beta\,du}{u(u^2-1)^{1/2}},$$

which integrates to

$$\theta = \beta\cos^{-1}\left(\frac{1}{u}\right) + k,$$

using the result from the hint.

Since the geodesic must pass through both $(R, -\theta_0, \alpha R)$ and $(R, \theta_0, \alpha R)$, we must have $k = 0$ and

$$\cos\frac{\theta_0}{\beta} = \frac{c}{R}.$$

Further, at a general point on the geodesic,

$$\cos\frac{\theta}{\beta} = \frac{c}{\rho}.$$

Eliminating c then shows that the geodesic on the cone that joins the two given points is

$$\rho(\theta) = \frac{R\cos(\theta_0/\beta)}{\cos(\theta/\beta)}.$$

22.9 *You are provided with a line of length $\pi a/2$ and negligible mass and some lead shot of total mass M. Use a variational method to determine how the lead shot must be distributed along the line if the loaded line is to hang in a circular arc of radius a when its ends are attached to two points at the same height. Measure the distance s along the line from its centre.*

We first note that the total mass of shot available is merely a scaling factor and not a constraint on the minimisation process.

The length of string is sufficient to form one-quarter of a complete circle of radius a, and so the ends of the string must be fixed to points that are $2a\sin(\pi/4) = \sqrt{2}a$ apart.

We take the distribution of shot as $\rho = \rho(s)$ and have to minimise the integral $\int gy(s)\rho(s)\,ds$, but subject to the requirement $\int dx = a/\sqrt{2}$. Expressed as an integral over s, this requirement can be written

$$\frac{a}{\sqrt{2}} = \int_{s=0}^{s=\pi a/4} dx = \int_0^{\pi a/4} (1-y'^2)^{1/2}\,ds,$$

where the derivative y' of y is with respect to s (not x).

We therefore consider the minimisation of $\int F(y, y', s)\, ds$, where

$$F(y, y', s) = gy\rho + \lambda\sqrt{1 - y'^2}.$$

The E–L equation takes the form

$$\frac{d}{ds}\left(\frac{\partial F}{\partial y'}\right) = \frac{\partial F}{\partial y},$$

$$\lambda\frac{d}{ds}\left(\frac{-y'}{\sqrt{1 - y'^2}}\right) = g\rho(s),$$

$$\frac{-\lambda y'}{\sqrt{1 - y'^2}} = \int_0^s g\rho(s')\, ds' \equiv gP(s),$$

since $y'(0) = 0$ by symmetry.

Now we require $P(s)$ to be such that the solution to this equation takes the form of an arc of a circle, $y(s) = y_0 - a\cos(s/a)$. If this is so, then $y'(s) = \sin(s/a)$ and

$$\frac{-\lambda\sin(s/a)}{\cos(s/a)} = gP(s).$$

When $s = \pi a/4$, $P(s)$ must have the value $M/2$, implying that $\lambda = -Mg/2$ and that, consequently,

$$P(s) = \frac{M}{2}\tan\left(\frac{s}{a}\right).$$

The required distribution $\rho(s)$ is recovered by differentiating this to obtain

$$\rho(s) = \frac{dP}{ds} = \frac{M}{2a}\sec^2\left(\frac{s}{a}\right).$$

22.11 *A general result is that light travels through a variable medium by a path that minimises the travel time (this is an alternative formulation of Fermat's principle). With respect to a particular cylindrical polar coordinate system (ρ, ϕ, z), the speed of light $v(\rho, \phi)$ is independent of z. If the path of the light is parameterised as $\rho = \rho(z), \phi = \phi(z)$, show that*

$$v^2(\rho'^2 + \rho^2\phi'^2 + 1)$$

is constant along the path.

For the particular case when $v = v(\rho) = b(a^2 + \rho^2)^{1/2}$, show that the two Euler–Lagrange equations have a common solution in which the light travels along a helical path given by $\phi = Az + B$, $\rho = C$, provided that A has a particular value.

In cylindrical polar coordinates with $\rho = \rho(z)$ and $\phi = \phi(z)$,

$$ds = \left[1 + \left(\frac{d\rho}{dz}\right)^2 + \rho^2\left(\frac{d\phi}{dz}\right)^2\right]^{1/2} dz.$$

The total travel time of the light is therefore given by

$$\tau = \int \frac{(1+\rho'^2+\rho^2\phi'^2)^{1/2}}{v(\rho,\phi)}\,dz.$$

Since z does not appear explicitly in the integrand, we have from the general first integral of the E–L equations for more than one dependent variable that

$$\frac{(1+\rho'^2+\rho^2\phi'^2)^{1/2}}{v(\rho,\phi)} - \frac{1}{v}\,\frac{\rho'^2}{(1+\rho'^2+\rho^2\phi'^2)^{1/2}} - \frac{1}{v}\,\frac{\rho^2\phi'^2}{(1+\rho'^2+\rho^2\phi'^2)^{1/2}} = k.$$

Rearranging this gives

$$\begin{aligned} 1+\rho'^2+\rho^2\phi'^2-\rho'^2-\rho^2\phi'^2 &= kv(1+\rho'^2+\rho^2\phi'^2)^{1/2},\\ 1 &= kv(1+\rho'^2+\rho^2\phi'^2)^{1/2},\\ \Rightarrow \quad v^2(1+\rho'^2+\rho^2\phi'^2) &= c, \text{ along the path.}\end{aligned}$$

Denoting $(1+\rho'^2+\rho^2\phi'^2)$ by $(**)$ for brevity, the E–L equations for ρ and ϕ are, respectively,

$$\frac{\rho\phi'^2}{v(**)^{1/2}} - \frac{(**)^{1/2}}{v^2}\frac{\partial v}{\partial \rho} = \frac{d}{dz}\left[\frac{\rho'}{v(**)^{1/2}}\right], \quad (1)$$

and

$$-\frac{(**)^{1/2}}{v^2}\frac{\partial v}{\partial \phi} = \frac{d}{dz}\left[\frac{\rho^2\phi'}{v(**)^{1/2}}\right]. \quad (2)$$

Now, if $v = b(a^2+\rho^2)^{1/2}$, the only dependence on z in a possible solution $\phi = Az + B$ with $\rho = C$ is through the first of these equations. To see this we note that the square brackets on the RHS's of the two E–L equations do not contain any undifferentiated ϕ-terms and so the derivatives (with respect to z) of both are zero. Since $\partial v/\partial\phi$ is also zero, equation (2) is identically satisfied as $0 = 0$. This leaves only (1), which reads

$$\frac{CA^2}{b(a^2+C^2)^{1/2}(1+0+C^2A^2)^{1/2}} - \frac{(1+0+C^2A^2)^{1/2}bC}{b^2(a^2+C^2)(a^2+C^2)^{1/2}} = 0.$$

This is satisfied provided

$$\begin{aligned} A^2(a^2+C^2) &= 1 + C^2A^2,\\ \text{i.e.}\quad A &= a^{-1}.\end{aligned}$$

Thus, a solution in the form of a helix is possible provided that the helix has a particular pitch, $2\pi a$.

22.13 *A dam of capacity V (less than $\pi b^2 h/2$) is to be constructed on level ground next to a long straight wall which runs from $(-b, 0)$ to $(b, 0)$. This is to be achieved by joining the ends of a new wall, of height h, to those of the existing wall. Show that, in order to minimise the length L of new wall to be built, it should form part of a circle, and that L is then given by*

$$\int_{-b}^{b} \frac{dx}{(1-\lambda^2 x^2)^{1/2}},$$

where λ is found from

$$\frac{V}{hb^2} = \frac{\sin^{-1}\mu}{\mu^2} - \frac{(1-\mu^2)^{1/2}}{\mu}$$

and $\mu = \lambda b$.

The objective is to chose the wall profile, $y = y(x)$, so as to minimise

$$L = \int_{-b}^{b} \sqrt{(dx)^2 + (dy)^2} = \int_{-b}^{b} (1 + y'^2)^{1/2}\, dx$$

subject to the constraint that the capacity of the dam formed is

$$V = h \int_{-b}^{b} y\, dx.$$

For this constrained variation problem we consider the minimisation of

$$K = \int_{-b}^{b} [\,(1 + y'^2)^{1/2} - \lambda y\,]\, dx,$$

where λ is a Lagrange multiplier.

Since x does not appear in the integrand, a first integral of the E–L equation is

$$(1 + y'^2)^{1/2} - \lambda y - y' \frac{y'}{(1 + y'^2)^{1/2}} = k,$$
$$\frac{1}{(1 + y'^2)^{1/2}} = k + \lambda y.$$

Rearranging this and integrating gives

$$\frac{1}{(k + \lambda y)^2} - 1 = y'^2,$$
$$\frac{(k + \lambda y)\, dy}{\sqrt{1 - (k + \lambda y)^2}} = dx,$$
$$\Rightarrow \quad -\frac{\sqrt{1 - (k + \lambda y)^2}}{\lambda} = x + c.$$

This result can be arranged in a more familiar form as

$$\lambda^2(x+c)^2 + (k+\lambda y)^2 = 1.$$

This is the equation of a circle that is centred on $(-c, -k/\lambda)$; from symmetry $c = 0$. Further, since $(\pm b, 0)$ lies on the curve, we must have

$$\lambda^2 b^2 + k^2 = 1, \qquad (*)$$

giving a connection between the Lagrange multiplier and one of the constants of integration. The length of the wall is given by

$$L = \int_{-b}^{b} (1 + y'^2)^{1/2}\, dx = \int_{-b}^{b} \frac{1}{k+\lambda y}\, dx = \int_{-b}^{b} \frac{1}{(1-\lambda^2 x^2)^{1/2}}\, dx.$$

The remaining constraint determines the value of λ and is that

$$\begin{aligned}
\frac{V}{h} &= \int_{-b}^{b} y\, dx = \frac{1}{\lambda}\int_{-b}^{b} \left(\sqrt{1-\lambda^2 x^2} - k\right) dx \\
&= \frac{1}{\lambda}\int_{-b}^{b} \left(\sqrt{1-\lambda^2 x^2} - \sqrt{1-\lambda^2 b^2}\right) dx, \text{ using}(*), \\
\frac{\lambda V}{h} &= \left[x\sqrt{1-\lambda^2 x^2} \right]_{-b}^{b} - \int_{-b}^{b} \frac{-\lambda^2 x\, x}{\sqrt{1-\lambda^2 x^2}}\, dx - \left[x\sqrt{1-\lambda^2 b^2} \right]_{-b}^{b} \\
&= \lambda^2 \int_{-b}^{b} \frac{x^2}{\sqrt{1-\lambda^2 x^2}}\, dx.
\end{aligned}$$

To evaluate this integral we set $\lambda x = \sin\theta$ and $\mu = \lambda b = \sin\phi$, to give

$$\begin{aligned}
\frac{\lambda V}{h} &= \int_{-\phi}^{\phi} \frac{\sin^2\theta\, \cos\theta}{\lambda\cos\theta}\, d\theta, \\
\frac{\lambda^2 V}{h} &= \int_{-\phi}^{\phi} \frac{1}{2}(1-\cos 2\theta)\, d\theta \\
&= \phi - \frac{2}{4}\sin 2\phi, \\
\frac{\mu^2 V}{hb^2} &= \sin^{-1}\mu - \frac{1}{2}\, 2\mu(1-\mu^2)^{1/2}, \\
\frac{V}{hb^2} &= \frac{\sin^{-1}\mu}{\mu^2} - \frac{(1-\mu^2)^{1/2}}{\mu}.
\end{aligned}$$

This equation determines μ and hence λ.

22.15 *The Schwarzchild metric for the static field of a non-rotating spherically symmetric black hole of mass M is given by*

$$(ds)^2 = c^2\left(1 - \frac{2GM}{c^2 r}\right)(dt)^2 - \frac{(dr)^2}{1 - 2GM/(c^2 r)} - r^2\,(d\theta)^2 - r^2\sin^2\theta\,(d\phi)^2.$$

Considering only motion confined to the plane $\theta = \pi/2$, and assuming that the path of a small test particle is such as to make $\int ds$ stationary, find two first integrals of the equations of motion. From their Newtonian limits, in which GM/r, $\dot{r}^2$ and $r^2\dot{\phi}^2$ are all $\ll c^2$, identify the constants of integration.

For motion confined to the plane $\theta = \pi/2$, $d\theta = 0$ and the corresponding term in the metric can be ignored. With this simplification, we can write

$$ds = \left\{ c^2\left(1 - \frac{2GM}{c^2 r}\right) - \frac{\dot{r}^2}{1 - (2GM)/(c^2 r)} - r^2\dot{\phi}^2 \right\}^{1/2} dt.$$

Writing the terms in braces as $\{**\}$, the E–L equation for ϕ reads

$$\frac{d}{dt}\left(\frac{-r^2\dot{\phi}}{\{**\}^{1/2}}\right) - 0 = 0,$$

$$\Rightarrow \quad \frac{r^2\dot{\phi}}{\{**\}^{1/2}} = A.$$

In the Newtonian limit $\{**\} \to c^2$ and the equation becomes $r^2\dot{\phi} = Ac$. Thus, Ac is a measure of the angular momentum of the particle about the origin.

The E–L equation for r is more complicated but, because ds does not contain t explicitly, we can use the general result for the first integral of the E–L equations when there is more than one dependent variable: $F - \sum_i \dot{q}_i \dfrac{\partial F}{\partial \dot{q}_i} = k$. This gives us a second equation as follows:

$$F - \dot{r}\frac{\partial F}{\partial \dot{r}} - \dot{\phi}\frac{\partial F}{\partial \dot{\phi}} = B,$$

$$\{**\}^{1/2} + \frac{\dot{r}}{\{**\}^{1/2}}\,\frac{\dot{r}}{[1 - (2GM)/(c^2 r)]} + \frac{\dot{\phi}}{\{**\}^{1/2}}\,r^2\dot{\phi} = B.$$

Multiplying through by $\{**\}^{1/2}$ and cancelling the terms in $\dot{r}^2$ and $\dot{\phi}^2$ now gives

$$c^2 - \frac{2GM}{r} = B\left\{c^2 - \frac{2GM}{r} - \frac{\dot{r}^2}{[1 - (2GM)/(c^2 r)]} - r^2\dot{\phi}^2\right\}^{1/2}.$$

In the Newtonian limits, in which GM/r, $\dot{r}^2$ and $r^2\dot{\phi}^2$ are all $\ll c^2$, the equation can be rearranged and the braces expanded to first order in small quantities to give

$$B = \left(c^2 - \frac{2GM}{r}\right)\left\{c^2 - \frac{2GM}{r} - \frac{\dot{r}^2}{[1-(2GM)/(c^2 r)]} - r^2\dot{\phi}^2\right\}^{-1/2},$$

$$cB = c^2 - \frac{2GM}{r} + \frac{c^2 GM}{c^2 r} + \frac{c^2\dot{r}^2}{2c^2} + \frac{c^2 r^2\dot{\phi}^2}{2c^2} + \cdots,$$

$$= c^2 - \frac{GM}{r} + \frac{1}{2}(\dot{r}^2 + r^2\dot{\phi}^2) + \cdots,$$

which can be read as 'total energy = rest mass energy + gravitational energy + radial and azimuthal kinetic energy'. Thus Bc is a measure of the total energy of the test particle.

22.17 *Determine the minimum value that the integral*

$$J = \int_0^1 [x^4(y'')^2 + 4x^2(y')^2]\,dx$$

can have, given that y is not singular at $x = 0$ and that $y(1) = y'(1) = 1$. Assume that the Euler–Lagrange equation gives the lower limit and verify retrospectively that your solution satisfies the end-point condition

$$\left[\eta\frac{\partial F}{\partial y'}\right]_a^b = 0,$$

where $F = F(y', y, x)$ and $\eta(x)$ is the variation from the minimising curve.

We first set $y'(x) = u(x)$ with $u(1) = y'(1) = 1$. The integral then becomes

$$J = \int_0^1 [x^4(u')^2 + 4x^2u^2]\,dx. \qquad (*)$$

This will be stationary if (using the E–L equation)

$$\frac{d}{dx}(2x^4u') - 8x^2u = 0,$$

$$8x^3u' + 2x^4u'' - 8x^2u = 0,$$

$$x^2u'' + 4xu' - 4u = 0.$$

As this is a homogeneous equation, we try $u(x) = Ax^n$, obtaining

$$n(n-1) + 4n - 4 = 0 \quad \Rightarrow \quad n = -4, \text{ or } n = 1.$$

The form of $y'(x)$ is thus

$$y'(x) = u(x) = \frac{A}{x^4} + Bx \quad \text{with} \quad A + B = 1.$$

Further,

$$y(x) = -\frac{A}{3x^3} + \frac{Bx^2}{2} + C.$$

Since y is not singular at $x = 0$ and $y(1) = 1$, we have that $A = 0$, $B = 1$ and $C = \frac{1}{2}$, yielding $y(x) = \frac{1}{2}(1 + x^2)$. The minimal value of J is thus

$$J_{\min} = \int_0^1 [\,x^4(1)^2 + 4x^2(x)^2\,]\,dx = \int_0^1 5x^4\,dx = \left[\,x^5\,\right]_0^1 = 1.$$

In (∗) the integrand is $G(u', u, x)$ and so the end-point condition reads

$$\left[\eta \frac{\partial G}{\partial u'}\right]_0^1 = 0.$$

At the upper limit $\eta(1) = 0$, since $u(1) = y'(1) = 1$ is fixed. At the lower limit,

$$\left.\frac{\partial G}{\partial u'}\right|_{x=0} = 2x^4u'\big|_{x=0} = 0.$$

Thus the contributions at the two limits are individually zero and the boundary condition is satisfied in the simplest way.

22.19 *Find an appropriate but simple trial function and use it to estimate the lowest eigenvalue λ_0 of Stokes' equation*

$$\frac{d^2y}{dx^2} + \lambda xy = 0, \qquad y(0) = y(\pi) = 0.$$

Explain why your estimate must be strictly greater than λ_0.

Stokes' equation is an S–L equation with $p = 1$, $q = 0$ and $\rho = x$. For the given boundary conditions the obvious trial function is $y(x) = \sin x$. The lowest eigenvalue $\lambda_0 \leq I/J$, where

$$I = \int_0^\pi py'^2\,dx = \int_0^\pi \cos^2 x\,dx = \frac{\pi}{2}$$

and

$$\begin{aligned} J &= \int_0^\pi \rho y^2\,dx = \int_0^\pi x\sin^2 x\,dx \\ &= \int_0^\pi \tfrac{1}{2}x(1 - \cos 2x)\,dx \\ &= \left[\frac{x^2}{4}\right]_0^\pi - \left[\frac{x}{2}\,\frac{\sin 2x}{2}\right]_0^\pi + \frac{1}{2}\int_0^\pi \frac{\sin 2x}{2}\,dx \\ &= \frac{\pi^2}{4}. \end{aligned}$$

Thus $\lambda_0 \leq (\frac{1}{2}\pi)/(\frac{1}{4}\pi^2) = 2/\pi$.

However, if we substitute the trial function directly into the equation we obtain

$$-\sin x + \frac{2}{\pi} x \sin x = 0,$$

which is clearly not satisfied. Thus the trial function is not an eigenfunction, and the actual lowest eigenvalue must be strictly less than the estimate of $2/\pi$.

22.21 *A drumskin is stretched across a fixed circular rim of radius a. Small transverse vibrations of the skin have an amplitude $z(\rho, \phi, t)$ that satisfies*

$$\nabla^2 z = \frac{1}{c^2} \frac{\partial^2 z}{\partial t^2}$$

in plane polar coordinates. For a normal mode independent of azimuth, in which case $z = Z(\rho) \cos \omega t$, find the differential equation satisfied by $Z(\rho)$. By using a trial function of the form $a^\nu - \rho^\nu$, with adjustable parameter ν, obtain an estimate for the lowest normal mode frequency.

[*The exact answer is* $(5.78)^{1/2} c/a$.]

In cylindrical polar coordinates, (ρ, ϕ), the wave equation,

$$\nabla^2 z = \frac{1}{c^2} \frac{\partial^2 z}{\partial t^2},$$

has azimuth-independent solutions (i.e. independent of ϕ) of the form $z(\rho, t) = Z(\rho) \cos \omega t$, and reduces to

$$\frac{1}{\rho} \frac{d}{d\rho} \left(\rho \frac{dZ}{d\rho} \right) \cos \omega t = -\frac{Z \omega^2}{c^2} \cos \omega t,$$

$$\frac{d}{d\rho} \left(\rho \frac{dZ}{d\rho} \right) + \frac{\omega^2}{c^2} \rho Z = 0.$$

The boundary conditions require that $Z(a) = 0$ and, so that there is no physical discontinuity in the slope of the drumskin at the origin, $Z'(0) = 0$.

This is an S–L equation with $p = \rho$, $q = 0$ and weight function $w = \rho$. A suitable trial function is $Z(\rho) = a^\nu - \rho^\nu$, which automatically satisfies $Z(a) = 0$ and, provided $\nu > 1$, has $Z'(0) = -\nu \rho^{\nu-1}|_{\rho=0} = 0$.

We recall that the lowest eigenfrequency satisfies the general formula

$$\frac{\omega^2}{c^2} \leq \frac{\displaystyle\int_0^a [(pZ')^2 - qZ^2]\, d\rho}{\displaystyle\int_0^a wZ^2\, d\rho}.$$

In this case

$$\frac{\omega^2}{c^2} \le \frac{\int_0^a \rho\, \nu^2 \rho^{2\nu-2}\, d\rho}{\int_0^a \rho(a^\nu - \rho^\nu)^2\, d\rho}$$

$$= \frac{\int_0^a \nu^2 \rho^{2\nu-1}\, d\rho}{\int_0^a (\rho a^{2\nu} - 2\rho^{\nu+1} a^\nu + \rho^{2\nu+1})\, d\rho}$$

$$= \frac{(\nu^2 a^{2\nu})/2\nu}{\dfrac{a^{2\nu+2}}{2} - \dfrac{2a^{2\nu+2}}{\nu+2} + \dfrac{a^{2\nu+2}}{2\nu+2}}$$

$$= \frac{1}{a^2}\, \frac{\nu(\nu+2)(2\nu+2)}{(\nu+2)(2\nu+2) - 4(2\nu+2) + 2(\nu+2)}$$

$$= \frac{(\nu+2)(\nu+1)}{\nu a^2}.$$

Since ν is an adjustable parameter and we know that, however we choose it, the resulting estimate can *never* be less than the lowest true eigenvalue, we choose the value that minimises the above estimate. Differentiating the estimate with respect to ν gives

$$\nu(2\nu+3) - (\nu^2 + 3\nu + 2) = 0 \quad \Rightarrow \quad \nu^2 - 2 = 0 \quad \Rightarrow \quad \nu = \sqrt{2}.$$

Thus the least upper bound to be found with this parameterisation is

$$\omega^2 \le \frac{c^2}{a^2}\, \frac{(\sqrt{2}+2)(\sqrt{2}+1)}{\sqrt{2}} = \frac{c^2}{2a^2}(\sqrt{2}+2)^2 \quad \Rightarrow \quad \omega = (5.83)^{1/2}\frac{c}{a}.$$

As noted, the actual lowest eigenfrequency is very little below this.

22.23 *For the boundary conditions given below, obtain a functional $\Lambda(y)$ whose stationary values give the eigenvalues of the equation*

$$(1+x)\frac{d^2y}{dx^2} + (2+x)\frac{dy}{dx} + \lambda y = 0, \qquad y(0) = 0,\ y'(2) = 0.$$

Derive an approximation to the lowest eigenvalue λ_0 using the trial function $y(x) = xe^{-x/2}$. For what value(s) of γ would

$$y(x) = xe^{-x/2} + \beta \sin \gamma x$$

be a suitable trial function for attempting to obtain an improved estimate of λ_0?

Since the derivative of $1 + x$ is not equal to $2 + x$, the given equation is not in self-adjoint form and an integrating factor for the standard form equation,

$$\frac{d^2y}{dx^2} + \frac{2+x}{1+x}\frac{dy}{dx} + \frac{\lambda y}{1+x} = 0,$$

is needed. This will be

$$\exp\left\{\int^x \frac{2+u}{1+u}\,du\right\} = \exp\left\{\int^x \left(1 + \frac{1}{1+u}\right)\,du\right\} = e^x(1+x).$$

Thus, after multiplying through by this IF, the equation takes the S–L form

$$[\,(1+x)e^x y'\,]' + \lambda e^x y = 0,$$

with $p(x) = (1+x)e^x$, $q(x) = 0$ and $\rho(x) = e^x$.

The required functional is therefore

$$\Lambda(y) = \frac{\int_0^2 [\,(1+x)e^x y'^2 + 0\,]\,dx}{\int_0^2 y^2 e^x\,dx},$$

provided that, for the eigenfunctions y_i of the equation, $\left[\,y_i p(x) y_j'(x)\,\right]_0^2 = 0$; this condition is automatically satisfied with the given boundary conditions.

For the trial function $y(x) = xe^{-x/2}$, clearly $y(0) = 0$ and, less obviously, $y'(x) = (1 - \frac{1}{2}x)e^{-x/2}$, making $y'(2) = 0$. The functional takes the following form:

$$\begin{aligned}
\Lambda &= \frac{\int_0^2 (1+x)e^x\,(1-\frac{1}{2}x)^2 e^{-x}\,dx}{\int_0^2 x^2 e^{-x} e^x\,dx} \\
&= \frac{\int_0^2 (1+x)(1-\frac{1}{2}x)^2\,dx}{\int_0^2 x^2\,dx} \\
&= \frac{\int_0^2 (1 - x^2 + \frac{1}{4}x^2 + \frac{1}{4}x^3)\,dx}{8/3} \\
&= \frac{3}{8}\left(2 - \frac{3}{4}\frac{8}{3} + \frac{16}{16}\right) = \frac{3}{8}.
\end{aligned}$$

Thus the lowest eigenvalue is $\leq \frac{3}{8}$.

We already know that $xe^{-x/2}$ is a suitable trial function and thus $y_2(x) = \sin\gamma x$ can be considered on its own. It satisfies $y_2(0) = 0$, but must also satisfy $y_2'(2) = \gamma\cos(2\gamma) = 0$. This requires that $\gamma = \frac{1}{2}(n + \frac{1}{2})\pi$ for some integer n; trial functions with γ of this form can be used to try to obtain a better bound on λ_0 by choosing the best value for n and adjusting the parameter β.

22.25 *The unnormalised ground-state (i.e. the lowest-energy) wavefunction of the simple harmonic oscillator of classical frequency ω is $\exp(-\alpha x^2)$, where $\alpha = m\omega/2\hbar$. Take as a trial function the orthogonal wavefunction $x^{2n+1}\exp(-\alpha x^2)$, using the integer n as a variable parameter, and apply either Sturm–Liouville theory or the Rayleigh–Ritz principle to show that the energy of the second lowest state of a quantum harmonic oscillator is $\leq 3\hbar\omega/2$.*

We first note that, for n a non-negative integer,

$$\int_{-\infty}^{\infty} x^{2n+1} e^{-\alpha x^2} e^{-\alpha x^2}\, dx = 0$$

on symmetry grounds and so confirm that the ground-state wavefunction, $\exp(-\alpha x^2)$, and the trial function, $\psi_{2n+1} = x^{2n+1}\exp(-\alpha x^2)$, are orthogonal with respect to a unit weight function.

The Hamiltonian for the quantum harmonic oscillator in one-dimension is given by

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{k}{2}x^2.$$

This means that to prepare the elements required for a Rayleigh–Ritz analysis we will need to find the second derivative of the trial function and evaluate integrals with integrands of the form $x^n\exp(-2\alpha x^2)$. To this end, define

$$I_n = \int_{-\infty}^{\infty} x^n e^{-2\alpha x^2}\, dx, \text{ with recurrence relation } I_n = \frac{n-1}{4\alpha} I_{n-2}.$$

Using Leibnitz' formula shows that

$$\begin{aligned}\frac{d^2\psi_{2n+1}}{dx^2} &= \big[\,2n(2n+1)x^{2n-1} + 2(2n+1)(-2\alpha)x^{2n+1} \\ &\qquad + (4\alpha^2x^2 - 2\alpha)x^{2n+1}\,\big]\, e^{-\alpha x^2} \\ &= \big[\,2n(2n+1)x^{2n-1} - 2(4n+3)\alpha x^{2n+1} + 4\alpha^2 x^{2n+3}\,\big]\, e^{-\alpha x^2}.\end{aligned}$$

Hence, we find that $\langle H\rangle$ is given by

$$\begin{aligned}&-\frac{\hbar^2}{2m}\int_{-\infty}^{\infty} x^{2n+1}e^{-\alpha x^2}\frac{d^2\psi_{2n+1}}{dx^2}\, dx + \frac{k}{2}\int_{-\infty}^{\infty} x^2 x^{4n+2} e^{-2\alpha x^2}\, dx \\ &= -\frac{\hbar^2}{2m}\big[\,2n(2n+1)I_{4n} - 2(4n+3)\alpha I_{4n+2} + 4\alpha^2 I_{4n+4}\,\big] + \frac{k}{2}I_{4n+4} \\ &= I_{4n+2}\left\{-\frac{\hbar^2}{2m}\left[\frac{2n(2n+1)4\alpha}{4n+1} - 2(4n+3)\alpha + \frac{4\alpha^2(4n+3)}{4\alpha}\right] + \frac{k(4n+3)}{8\alpha}\right\},\end{aligned}$$

where we have used the recurrence relation to express all integrals in terms of

I_{4n+2}. This has been done because the denominator of the Rayleigh–Ritz quotient is this (same) normalisation integral, namely

$$\int_{-\infty}^{\infty} \psi_{2n+1}^{*}\psi_{2n+1}\,dx = I_{4n+2}.$$

Thus, the estimate $E_{2n+1} = \langle H\rangle/I_{4n+2}$ is given by

$$\begin{aligned} E_{2n+1} &= -\frac{\hbar^2\alpha}{2m}\left(\frac{16n^2+8n-16n^2-16n-3}{4n+1}\right)+\frac{k(4n+3)}{8\alpha} \\ &= \frac{\hbar^2\alpha}{2m}\,\frac{8n+3}{4n+1}+\frac{k(4n+3)}{8\alpha}. \end{aligned}$$

Using $\omega^2 = \dfrac{k}{m}$ and $\alpha = \dfrac{m\omega}{2\hbar}$ then yields

$$E_{2n+1} = \frac{\hbar\omega}{4}\left(\frac{8n+3}{4n+1}+4n+3\right) = \frac{\hbar\omega}{2}\,\frac{8n^2+12n+3}{4n+1}.$$

For non-negative integers n (the orthogonality requirement is not satisfied for non-integer values), this has a minimum value of $\frac{3}{2}\hbar\omega$ when $n = 0$. Thus the second lowest energy level is less than or equal to this value. In fact, it is equal to this value, as can be shown by substituting ψ_1 into $H\psi = E\psi$.

22.27 *The upper and lower surfaces of a film of liquid, which has surface energy per unit area (surface tension) γ and density ρ, have equations $z = p(x)$ and $z = q(x)$, respectively. The film has a given volume V (per unit depth in the y-direction) and lies in the region $-L < x < L$, with $p(0) = q(0) = p(L) = q(L) = 0$. The total energy (per unit depth) of the film consists of its surface energy and its gravitational energy, and is expressed by*

$$E = \tfrac{1}{2}\rho g\int_{-L}^{L}(p^2-q^2)\,dx + \gamma\int_{-L}^{L}\left[(1+p'^2)^{1/2}+(1+q'^2)^{1/2}\right]dx.$$

(a) *Express V in terms of p and q.*

(b) *Show that, if the total energy is minimised, p and q must satisfy*

$$\frac{p'^2}{(1+p'^2)^{1/2}} - \frac{q'^2}{(1+q'^2)^{1/2}} = \textit{constant}.$$

(c) *As an approximate solution, consider the equations*

$$p = a(L-|x|), \qquad q = b(L-|x|),$$

where a and b are sufficiently small that a^3 and b^3 can be neglected compared with unity. Find the values of a and b that minimise E.

(a) The total volume constraint is given simply by

$$V = \int_{-L}^{L} [\, p(x) - q(x)\,]\, dx.$$

(b) To take account of the constraint, consider the minimisation of $E - \lambda V$, where λ is an undetermined Lagrange multiplier. The integrand does not contain x explicitly and so we have two first integrals of the E–L equations, one for $p(x)$ and the other for $q(x)$. They are

$$\frac{1}{2}\rho g(p^2 - q^2) + \gamma(1 + p'^2)^{1/2} + \gamma(1 + q'^2)^{1/2} - \lambda(p - q) - p'\frac{\gamma p'}{(1 + p'^2)^{1/2}} = k_1$$

and

$$\frac{1}{2}\rho g(p^2 - q^2) + \gamma(1 + p'^2)^{1/2} + \gamma(1 + q'^2)^{1/2} - \lambda(p - q) - q'\frac{\gamma q'}{(1 + q'^2)^{1/2}} = k_2.$$

Subtracting these two equations gives

$$\frac{p'^2}{(1 + p'^2)^{1/2}} - \frac{q'^2}{(1 + q'^2)^{1/2}} = \text{constant}.$$

(c) If

$$p = a(L - |x|), \qquad q = b(L - |x|),$$

the derivatives of p and q only take the values $\pm a$ and $\pm b$, respectively, and the volume constraint becomes

$$V = \int_{-L}^{L} (a - b)(L - |x|)\, dx = (a - b)L^2 \quad \Rightarrow \quad b = a - \frac{V}{L^2}.$$

The total energy can now be expressed entirely in terms of a and the given parameters, as follows:

$$\begin{aligned}
E &= \frac{1}{2}\rho g \int_{-L}^{L} (a^2 - b^2)(L - |x|)^2\, dx + 2\gamma L(1 + a^2)^{1/2} + 2\gamma L(1 + b^2)^{1/2} \\
&= \frac{1}{2}\rho g(a^2 - b^2)\frac{2L^3}{3} + 2\gamma L(1 + \tfrac{1}{2}a^2 + 1 + \tfrac{1}{2}b^2) + O(a^4) + O(b^4) \\
&\approx \frac{\rho g L^3}{3}\left[a^2 - \left(a - \tfrac{V}{L^2}\right)^2\right] + 2\gamma L\left[2 + \tfrac{1}{2}a^2 + \tfrac{1}{2}\left(a - \tfrac{V}{L^2}\right)^2\right] \\
&= \frac{\rho g L^3}{3}\left(\frac{2aV}{L^2} - \frac{V^2}{L^4}\right) + 2\gamma L\left(2 + a^2 - \frac{aV}{L^2} + \frac{V^2}{2L^4}\right).
\end{aligned}$$

This is minimised with respect to a when

$$\frac{2\rho g L^3 V}{3L^2} + 4\gamma L a - \frac{2\gamma L V}{L^2} = 0,$$

$$\Rightarrow \quad a = \frac{V}{2L^2} - \frac{\rho g V}{6\gamma},$$

$$\Rightarrow \quad b = -\frac{V}{2L^2} - \frac{\rho g V}{6\gamma}.$$

As might be expected, $|b| > |a|$ and there is more of the liquid below the $z = 0$ plane than there is above it.

22.29 *The 'stationary value of an integral' approach to finding the eigenvalues of a Sturm–Liouville equation can be extended to two independent variables, x and z, with little modification. In the integral to be minimised, y'^2 is replaced by $(\nabla y)^2$ and the integrals of the various functions of $y(x,z)$ become two-dimensional, i.e. the infinitesimal is $dx\,dz$.*

The vibrations of a trampoline 4 units long and 1 unit wide satisfy the equation

$$\nabla^2 y + k^2 y = 0.$$

By taking the simplest possible permissible polynomial as a trial function, show that the lowest mode of vibration has $k^2 \le 10.63$ and, by direct solution, that the actual value is 10.49.

Written explicitly, the equation is

$$\frac{\partial^2 y}{\partial x^2} + \frac{\partial^2 y}{\partial z^2} + k^2 y = 0.$$

This is an extended S–L equation with $p(x,z) = 1$, $q(x,z) = 0$, $\rho(x,z) = 1$ and eigenvalue λ. We therefore consider the stationary values of $\Lambda = I/J$, where

$$I = \int\int \left[\left(\frac{\partial y}{\partial x}\right)^2 + \left(\frac{\partial y}{\partial z}\right)^2 \right] dx\,dz$$

and J is the normalisation integral $\int y^2(x,z)\,dx\,dz$.

For a trampoline 4 units long and 1 unit wide, the simplest trial function that satisfies $y(0,z) = y(4,z) = y(x,0) = y(x,1) = 0$ is

$$y(x,z) = x(4-x)z(1-z).$$

For this function,

$$\frac{\partial y}{\partial x} = (4-2x)z(1-z) \quad \text{and} \quad \frac{\partial y}{\partial z} = (1-2z)x(4-x).$$

Thus, I is given by

$$\begin{aligned}
&\int_0^4 (4-2x)^2\,dx \int_0^1 z^2(1-z)^2\,dz + \int_0^4 x^2(4-x)^2\,dx \int_0^1 (1-2z)^2\,dz \\
&= \left[16(4) - 16\left(\frac{16}{2}\right) + 4\left(\frac{64}{3}\right)\right]\left[\left(\frac{1}{3}\right) - 2\left(\frac{1}{4}\right) + \left(\frac{1}{5}\right)\right] \\
&\quad + \left[16\left(\frac{64}{3}\right) - 8\left(\frac{256}{4}\right) + \left(\frac{1024}{5}\right)\right]\left[1(1) - 4\left(\frac{1}{2}\right) + 4\left(\frac{1}{3}\right)\right] \\
&= 64\left(1 - 2 + \frac{4}{3}\right)\frac{1}{30} + \frac{1024}{30}\left(1 - 2 + \frac{4}{3}\right) = \frac{1088}{90}.
\end{aligned}$$

Similarly, J is given by

$$\begin{aligned}
&\int_0^4 x^2(4-x)^2\,dx \int_0^1 z^2(1-z)^2\,dz \\
&= \left[16\left(\frac{64}{3}\right) - 8\left(\frac{256}{4}\right) + \left(\frac{1024}{5}\right)\right]\left[\left(\frac{1}{3}\right) - 2\left(\frac{1}{4}\right) + \left(\frac{1}{5}\right)\right] \\
&= 1024\left(\frac{1}{3} - \frac{2}{4} + \frac{1}{5}\right)^2 = \frac{1024}{900}.
\end{aligned}$$

Thus the lowest eigenvalue $k^2 \le (1088/90) \div (1024/900) = 10.63$.

The obvious direct solution satisfying the boundary conditions is

$$y(x,z) = A\sin\frac{\pi x}{4}\sin\pi z.$$

Substituting this into the original equation gives

$$-\frac{\pi^2}{16}y(x,z) - \pi^2 y(x,z) + k^2 y(x,z) = 0,$$

which is clearly satisfied if

$$k^2 = \frac{17\pi^2}{16} = 10.49.$$

23

Integral equations

23.1 *Solve the integral equation*

$$\int_0^\infty \cos(xv)y(v)\,dv = \exp(-x^2/2)$$

for the function $y = y(x)$ for $x > 0$. Note that for $x < 0$, $y(x)$ can be chosen as is most convenient.

Since $\cos uv$ is an even function of v, we will make $y(-v) = y(v)$ so that the complete integrand is also an even function of v. The integral I on the LHS can then be written as

$$I = \frac{1}{2}\int_{-\infty}^{\infty} \cos(xv)y(v)\,dv = \frac{1}{2}\text{Re}\int_{-\infty}^{\infty} e^{ixv}y(v)\,dv = \frac{1}{2}\int_{-\infty}^{\infty} e^{ixv}y(v)\,dv,$$

the last step following because $y(v)$ is symmetric in v. The integral is now $\sqrt{2\pi}\,\times$ a Fourier transform, and it follows from the inversion theorem for Fourier transforms applied to

$$\frac{1}{2}\int_{-\infty}^{\infty} e^{ixv}y(v)\,dv = \exp(-x^2/2)$$

that

$$\begin{aligned}
y(x) &= \frac{2}{2\pi}\int_{-\infty}^{\infty} e^{-u^2/2}e^{-iux}\,du \\
&= \frac{1}{\pi}\int_{-\infty}^{\infty} e^{-(u+ix)^2/2}e^{-x^2/2}\,dx \\
&= \frac{1}{\pi}\sqrt{2\pi}\,e^{-x^2/2} \\
&= \sqrt{\frac{2}{\pi}}\,e^{-x^2/2}.
\end{aligned}$$

Although, as noted in the question, $y(x)$ is arbitrary for $x < 0$, because its form in this range does not affect the value of the integral, for $x > 0$ it *must* have the form given. This is tricky to prove formally, but any second solution $w(x)$ has to satisfy

$$\int_0^\infty \cos(xv)[y(v) - w(v)]\, dv = 0$$

for all $x > 0$. Intuitively, this implies that $y(x)$ and $w(x)$ are identical functions.

23.3 *Convert*

$$f(x) = \exp x + \int_0^x (x - y)f(y)\, dy$$

into a differential equation, and hence show that its solution is

$$(\alpha + \beta x)\exp x + \gamma\ \exp(-x),$$

where α, β, γ *are constants that should be determined.*

We differentiate the integral equation twice and obtain

$$\begin{aligned} f'(x) &= e^x + (x - x)f(x) + \int_0^x f(y)\, dy, \\ f''(x) &= e^x + f(x). \end{aligned}$$

Expressed in the usual differential equation form, this last equation is

$$f''(x) - f(x) = e^x, \text{ for which the CF is } f(x) = Ae^x + Be^{-x}.$$

Since the complementary function contains the RHS of the equation, we try as a PI $f(x) = Cxe^x$:

$$Cxe^x + 2Ce^x - Cxe^x = e^x \quad \Rightarrow \quad \beta = C = \tfrac{1}{2}.$$

The general solution is therefore $f(x) = Ae^x + Be^{-x} + \frac{1}{2}xe^x$.

The boundary conditions needed to evaluate A and B are constructed by considering the integral equation and its derivative(s) at $x = 0$, because with $x = 0$ the integral on the RHS contributes nothing. We have

$$f(0) = e^0 + 0 = 1 \quad \Rightarrow \quad A + B = 1$$

and

$$f'(0) = e^0 + 0 = 1 \quad \Rightarrow \quad A - B + \tfrac{1}{2} = 1.$$

Solving these yields $\alpha = A = \frac{3}{4}$ and $\gamma = B = \frac{1}{4}$ and makes the complete solution

$$f(x) = \tfrac{3}{4}e^x + \tfrac{1}{4}e^{-x} + \tfrac{1}{2}xe^x.$$

23.5 *Solve for $\phi(x)$ the integral equation*

$$\phi(x) = f(x) + \lambda \int_0^1 \left[\left(\frac{x}{y}\right)^n + \left(\frac{y}{x}\right)^n \right] \phi(y)\, dy,$$

where $f(x)$ is bounded for $0 < x < 1$ and $-\frac{1}{2} < n < \frac{1}{2}$, expressing your answer in terms of the quantities $F_m = \int_0^1 f(y) y^m\, dy$.

(a) *Give the explicit solution when $\lambda = 1$.*
(b) *For what values of λ are there no solutions unless $F_{\pm n}$ are in a particular ratio? What is this ratio?*

This equation has a symmetric degenerate kernel, and so we set

$$\phi(x) = f(x) + a_1 x^n + a_2 x^{-n},$$

giving

$$\begin{aligned}
\frac{\phi(x) - f(x)}{\lambda} &= \int_0^1 \left(\frac{x^n}{y^n} + \frac{y^n}{x^n}\right) [\, f(y) + a_1 y^n + a_2 y^{-n}\,]\, dy \\
&= x^n \int_0^1 \frac{f(y)}{y^n}\, dy + x^{-n} \int_0^1 y^n f(y)\, dy + a_1 x^n \\
&\qquad + a_2 x^{-n} + a_1 x^{-n} \int_0^1 y^{2n}\, dy + a_2 x^n \int_0^1 y^{-2n}\, dy \\
&= x^n \left(F_{-n} + a_1 + \frac{a_2}{1 - 2n} \right) + x^{-n} \left(F_n + a_2 + \frac{a_1}{2n + 1} \right).
\end{aligned}$$

This is consistent with the assumed form of $\phi(x)$, provided

$$a_1 = \lambda \left(F_{-n} + a_1 + \frac{a_2}{1 - 2n} \right) \quad \text{and} \quad a_2 = \lambda \left(F_n + a_2 + \frac{a_1}{2n + 1} \right).$$

These two simultaneous linear equations can now be solved for a_1 and a_2.

(a) For $\lambda = 1$, the equations simplify and decouple to yield

$$a_2 = -(1 - 2n) F_{-n} \quad \text{and} \quad a_1 = -(1 + 2n) F_n,$$

respectively, giving as the explicit solution

$$\phi(x) = f(x) - (1 + 2n) F_n x^n - (1 - 2n) F_{-n} x^{-n}.$$

(b) For a general value of λ,

$$\begin{aligned}
(1 - \lambda) a_1 - \frac{\lambda}{1 - 2n}\, a_2 &= \lambda F_{-n}, \\
-\frac{\lambda}{1 + 2n}\, a_1 + (1 - \lambda) a_2 &= \lambda F_n.
\end{aligned}$$

The case $\lambda = 0$ is trivial, with $\phi(x) = f(x)$, and so suppose that $\lambda \neq 0$. Then, after being divided through by λ, the equations can be written in the matrix and vector form $\mathsf{Aa} = \mathsf{F}$:

$$\begin{pmatrix} \dfrac{1}{\lambda} - 1 & -\dfrac{1}{1-2n} \\ -\dfrac{1}{1+2n} & \dfrac{1}{\lambda} - 1 \end{pmatrix} \begin{pmatrix} a_1 \\ \\ a_2 \end{pmatrix} = \begin{pmatrix} F_{-n} \\ \\ F_n \end{pmatrix}.$$

In general, this matrix equation will have no solution if $|\mathsf{A}| = 0$. This will be the case if

$$\left(\frac{1}{\lambda} - 1\right)^2 - \frac{1}{1-4n^2} = 0,$$

which, on rearrangement, shows that λ would have to be given by

$$\frac{1}{\lambda} = 1 \pm \frac{1}{\sqrt{1-4n^2}}.$$

We note that this value for λ is real because n lies in the range $-\frac{1}{2} < n < \frac{1}{2}$. In fact $-\infty < \lambda < \frac{1}{2}$. Even for these two values of λ, however, if *either* $F_n = F_{-n} = 0$ *or* the matrix equation

$$\begin{pmatrix} \pm\dfrac{1}{\sqrt{1-4n^2}} & -\dfrac{1}{1-2n} \\ -\dfrac{1}{1+2n} & \pm\dfrac{1}{\sqrt{1-4n^2}} \end{pmatrix} \begin{pmatrix} a_1 \\ \\ a_2 \end{pmatrix} = \begin{pmatrix} F_{-n} \\ \\ F_n \end{pmatrix}$$

is equivalent to two linear equations that are multiples of each other, there will still be a solution. In this latter case, we must have

$$\frac{F_n}{F_{-n}} = \mp\sqrt{\frac{1-2n}{1+2n}}.$$

Again we note that, because of the range in which n lies, this ratio is real; this condition can, however, require any value in the range $-\infty$ to ∞ for F_n/F_{-n}.

23.7 *The kernel of the integral equation*

$$\psi(x) = \lambda \int_a^b K(x,y)\psi(y)\,dy$$

has the form

$$K(x,y) = \sum_{n=0}^{\infty} h_n(x)g_n(y),$$

where the $h_n(x)$ form a complete orthonormal set of functions over the interval $[a,b]$.

(a) *Show that the eigenvalues λ_i are given by*

$$|\mathsf{M} - \lambda^{-1}\mathsf{I}| = 0,$$

where M is the matrix with elements

$$M_{kj} = \int_a^b g_k(u)h_j(u)\,du.$$

If the corresponding solutions are $\psi^{(i)}(x) = \sum_{n=0}^{\infty} a_n^{(i)} h_n(x)$, find an expression for $a_n^{(i)}$.

(b) *Obtain the eigenvalues and eigenfunctions over the interval $[0, 2\pi]$ if*

$$K(x,y) = \sum_{n=1}^{\infty} \frac{1}{n}\cos nx \cos ny.$$

(a) We write the ith eigenfunction as

$$\psi^{(i)}(x) = \sum_{n=0}^{\infty} a_n^{(i)} h_n(x).$$

From the orthonormality of the $h_n(x)$, it follows immediately that

$$a_m^{(i)} = \int_a^b h_m(x)\psi^{(i)}(x)\,dx.$$

However, the coefficients $a_m^{(i)}$ have to be found as the components of the eigenvectors $\mathsf{a}^{(i)}$ defined below, since the $\psi^{(i)}$ are not initially known.

Substituting this assumed form of solution, we obtain

$$\begin{aligned}\sum_{m=0}^{\infty} a_m^{(i)} h_m(x) &= \lambda_i \int_a^b \sum_{n=0}^{\infty} h_n(x)g_n(y) \sum_{l=0}^{\infty} a_l^{(i)} h_l(y)\,dy \\ &= \lambda_i \sum_{n,l} a_l^{(i)} M_{nl} h_n(x).\end{aligned}$$

Since the $\{h_n\}$ are an orthonormal set, it follows that

$$a_m^{(i)} = \lambda_i \sum_{n,l} a_l^{(i)} M_{nl} \delta_{mn} = \lambda_i \sum_{l=0}^{\infty} M_{ml} a_l^{(i)},$$

$$\text{i.e.} \quad (\mathsf{M} - \lambda_i^{-1}\mathsf{I})\mathsf{a}^{(i)} = 0.$$

Thus, the allowed values of λ_i are given by $|\mathsf{M} - \lambda^{-1}\mathsf{I}| = 0$, and the expansion coefficients $a_m^{(i)}$ by the components of the corresponding eigenvectors.

(b) To make the set $\{h_n(x) = \cos nx\}$ into a complete orthonormal set we need to add the set of functions $\{\eta_\nu(x) = \sin \nu x\}$ and then normalise all the functions by multiplying them by $1/\sqrt{\pi}$. For this particular kernel the general functions $g_n(x)$ are given by $g_n(x) = n^{-1}\sqrt{\pi}\cos nx$.

The matrix elements are then

$$M_{kj} = \int_0^{2\pi} \frac{1}{\sqrt{\pi}} \cos ju \, \frac{\sqrt{\pi}}{k} \cos ku \, du = \frac{\pi}{k} \delta_{kj},$$

$$M_{k\nu} = \int_0^{2\pi} \frac{1}{\sqrt{\pi}} \sin \nu u \, \frac{\sqrt{\pi}}{k} \cos ku \, du = 0.$$

Thus the matrix M is diagonal and particularly simple. The eigenvalue equation reads

$$\sum_{j=0}^{\infty} \left(\frac{\pi}{k} \delta_{kj} - \lambda_i^{-1} \delta_{kj} \right) a_j^{(i)} = 0,$$

giving the immediate result that $\lambda_k = k/\pi$ with $a_k^{(k)} = 1$ and all other $a_j^{(k)} = a_\nu^{(k)} = 0$. The eigenfunction corresponding to eigenvalue k/π is therefore

$$\psi^{(k)}(x) = h_k(x) = \frac{1}{\sqrt{\pi}} \cos kx.$$

23.9 *For $f(t) = \exp(-t^2/2)$, use the relationships of the Fourier transforms of $f'(t)$ and $tf(t)$ to that of $f(t)$ itself to find a simple differential equation satisfied by $\tilde{f}(\omega)$, the Fourier transform of $f(t)$, and hence determine $\tilde{f}(\omega)$ to within a constant. Use this result to solve for $h(t)$ the integral equation*

$$\int_{-\infty}^{\infty} e^{-t(t-2x)/2} h(t) \, dt = e^{3x^2/8}.$$

As a standard result,

$$\mathscr{F}\left[f'(t) \right] = i\omega \tilde{f}(\omega),$$

though we will not need this relationship in the following solution.

From its definition,

$$\begin{aligned}\mathscr{F}\left[tf(t)\right] &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} tf(t)\,e^{-i\omega t}\,dt\\ &= \frac{1}{-i}\frac{d}{d\omega}\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(t)\,e^{-i\omega t}\,dt\right) = i\frac{d\tilde{f}}{d\omega}.\end{aligned}$$

Now, for the particular given function,

$$\begin{aligned}\tilde{f}(\omega) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-t^2/2}\,e^{-i\omega t}\,dt\\ &= \frac{1}{\sqrt{2\pi}}\left[\frac{e^{-t^2/2}\,e^{-i\omega t}}{-i\omega}\right]_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{te^{-t^2/2}\,e^{-i\omega t}}{-i\omega}\,dt\\ &= 0 - \frac{1}{i\omega}\,i\frac{d\tilde{f}}{d\omega}.\end{aligned}$$

Hence,

$$\frac{d\tilde{f}}{d\omega} = -\omega\tilde{f} \quad\Rightarrow\quad \ln\tilde{f} = -\tfrac{1}{2}\omega^2 + k \quad\Rightarrow\quad \tilde{f} = Ae^{-\omega^2/2}, \qquad (*)$$

giving $\tilde{f}(\omega)$ to within a multiplicative constant.

Now, we are given

$$\begin{aligned}&\int_{-\infty}^{\infty} e^{-t(t-2x)/2}\,h(t)\,dt = e^{3x^2/8},\\ \Rightarrow\quad &\int_{-\infty}^{\infty} e^{-(t-x)^2/2}\,e^{x^2/2}\,h(t)\,dt = e^{3x^2/8},\\ \Rightarrow\quad &\int_{-\infty}^{\infty} e^{-(x-t)^2/2}\,h(t)\,dt = e^{-x^2/8}. \qquad (**)\end{aligned}$$

The LHS of $(**)$ is a convolution integral, and so applying the convolution theorem for Fourier transforms and result $(*)$, used twice, yields

$$\begin{aligned}&\sqrt{2\pi}Ae^{-\omega^2/2}\tilde{h}(\omega) = \mathscr{F}\left[e^{-(x/2)^2/2}\right] = Ae^{-(2\omega)^2/2},\\ \Rightarrow\quad &\sqrt{2\pi}\,\tilde{h}(\omega) = e^{-3\omega^2/2} = e^{-(\sqrt{3}\omega)^2/2},\\ \Rightarrow\quad &h(t) = \frac{1}{\sqrt{2\pi}A}\,e^{-(t/\sqrt{3})^2/2} = \frac{1}{\sqrt{2\pi}A}\,e^{-t^2/6}.\end{aligned}$$

We now substitute in $(**)$ to determine A:

$$\int_{-\infty}^{\infty} e^{-(x-t)^2/2} \frac{1}{\sqrt{2\pi}A} e^{-t^2/6}\, dt = e^{-x^2/8},$$
$$\frac{1}{\sqrt{2\pi}A} \int_{-\infty}^{\infty} e^{-2t^2/3} e^{xt} e^{-x^2/2} e^{x^2/8}\, dt = 1,$$
$$\frac{1}{\sqrt{2\pi}A} \int_{-\infty}^{\infty} \exp\left[-\frac{2}{3}\left(t - \frac{3x}{4}\right)^2\right] dt = 1.$$

From the normalisation of the Gaussian integral, this implies that

$$\frac{1}{\sqrt{2\pi}A} = \frac{2}{\sqrt{2\pi}\,\sqrt{3}},$$

which in turn means $A = \sqrt{3}/2$, giving finally that

$$h(t) = \sqrt{\frac{2}{3\pi}}\, e^{-t^2/6}.$$

This solution can be checked by resubstitution.

23.11 *At an international 'peace' conference a large number of delegates are seated around a circular table with each delegation sitting near its allies and diametrically opposite the delegation most bitterly opposed to it. The position of a delegate is denoted by θ, with $0 \le \theta \le 2\pi$. The fury $f(\theta)$ felt by the delegate at θ is the sum of his own natural hostility $h(\theta)$ and the influences on him of each of the other delegates; a delegate at position ϕ contributes an amount $K(\theta - \phi)f(\phi)$. Thus*

$$f(\theta) = h(\theta) + \int_0^{2\pi} K(\theta - \phi)f(\phi)\, d\phi.$$

Show that if $K(\psi)$ takes the form $K(\psi) = k_0 + k_1 \cos\psi$ then

$$f(\theta) = h(\theta) + p + q\cos\theta + r\sin\theta$$

and evaluate p, q and r. A positive value for k_1 implies that delegates tend to placate their opponents but upset their allies, whilst negative values imply that they calm their allies but infuriate their opponents. A walkout will occur if $f(\theta)$ exceeds a certain threshold value for some θ. Is this more likely to happen for positive or for negative values of k_1?

Given that $K(\psi) = k_0 + k_1\cos\psi$, we try a solution $f(\theta) = h(\theta) + p + q\cos\theta + r\sin\theta$, reducing the equation to

$$\begin{aligned}
&p + q\cos\theta + r\sin\theta \\
&= \int_0^{2\pi} [k_0 + k_1(\cos\theta\cos\phi + \sin\theta\sin\phi)] \\
&\qquad \times\ [h(\phi) + p + q\cos\phi + r\sin\phi]\ d\phi \\
&= k_0(H + 2\pi p) + k_1(H_c\cos\theta + H_s\sin\theta + \pi q\cos\theta + \pi r\sin\theta),
\end{aligned}$$

where $H = \int_0^{2\pi} h(z)\,dz$, $H_c = \int_0^{2\pi} h(z)\cos z\,dz$ and $H_s = \int_0^{2\pi} h(z)\sin z\,dz$.

Thus, on equating the constant terms and the coefficients of $\cos\theta$ and $\sin\theta$, we have

$$\begin{aligned}
p &= k_0 H + 2\pi k_0 p \quad \Rightarrow \quad p = \frac{k_0 H}{1 - 2\pi k_0}, \\
q &= k_1 H_c + k_1 \pi q \quad \Rightarrow \quad q = \frac{k_1 H_c}{1 - k_1 \pi}, \\
r &= k_1 H_s + k_1 \pi r \quad \Rightarrow \quad r = \frac{k_1 H_s}{1 - k_1 \pi}.
\end{aligned}$$

And so the full solution for $f(\theta)$ is given by

$$\begin{aligned}
f(\theta) &= h(\theta) + \frac{k_0 H}{1 - 2\pi k_0} + \frac{k_1 H_c}{1 - k_1\pi}\cos\theta + \frac{k_1 H_s}{1 - k_1\pi}\sin\theta \\
&= h(\theta) + \frac{k_0 H}{1 - 2\pi k_0} + \frac{k_1}{1 - k_1\pi}(H_c^2 + H_s^2)^{1/2}\cos(\theta - \alpha),
\end{aligned}$$

where $\tan\alpha = H_s/H_c$.

Clearly, the maximum value of $f(\theta)$ will depend upon $h(\theta)$ and its various integrals, but it is most likely to exceed any particular value if k_1 is positive and $\approx \pi^{-1}$. Stick with your friends!

23.13 *The operator $\mathscr{M}$ is defined by*

$$\mathscr{M} f(x) \equiv \int_{-\infty}^{\infty} K(x, y) f(y)\,dy,$$

where $K(x, y) = 1$ inside the square $|x| < a, |y| < a$ and $K(x, y) = 0$ elsewhere. Consider the possible eigenvalues of $\mathscr{M}$ and the eigenfunctions that correspond to them; show that the only possible eigenvalues are 0 and $2a$ and determine the corresponding eigenfunctions. Hence find the general solution of

$$f(x) = g(x) + \lambda \int_{-\infty}^{\infty} K(x, y) f(y)\,dy.$$

From the given properties of $K(x, y)$ we can assert the following.

(i) No matter what the form of $f(x)$, $\mathscr{M}f(x) = 0$ if $|x| > a$.

(ii) All functions for which both $\int_{-a}^{a} f(y)\,dy = 0$ and $f(x) = 0$ for $|x| > a$ are eigenfunctions corresponding to eigenvalue 0.

(iii) For any function $f(x)$, the integral $\int_{-a}^{a} f(y)\,dy$ is equal to a constant whose value is independent of x; thus $f(x)$ can only be an eigenfunction if it is equal to a constant, μ, for $-a \le x \le a$ and is zero otherwise. For this case $\int_{-a}^{a} f(y)\,dy = 2a\mu$ and the eigenvalue is $2a$.

Point (iii) gives the only possible non-zero eigenvalue, whilst point (ii) shows that eigenfunctions corresponding to zero eigenvalues do exist.

Denote by $S(x, a)$ the function that has unit value for $|x| \le a$ and zero value otherwise; $K(x, y)$ could be expressed as $K(x, y) = S(x, a)S(y, a)$. Substitute the trial solution $f(x) = g(x) + kS(x, a)$ into

$$f(x) = g(x) + \lambda \int_{-\infty}^{\infty} K(x, y) f(y)\,dy.$$

This gives

$$g(x) + kS(x, a) = g(x) + \lambda \int_{-\infty}^{\infty} K(x, y)[\,g(y) + kS(y, a)\,]\,dy,$$
$$kS(x, a) = \lambda S(x, a) \int_{-a}^{a} g(y)\,dy + \lambda k\, 2aS(x, a).$$

Here, having replaced $K(x, y)$ by $S(x, a)S(y, a)$, we use the factor $S(y, a)$ to reduce the limits of the y-integration from $\pm\infty$ to $\pm a$. As this result is to hold for all x we must have

$$k = \frac{\lambda G}{1 - 2a\lambda}, \quad \text{where } G = \int_{-a}^{a} g(y)\,dy.$$

The general solution is thus

$$f(x) = \begin{cases} g(x) + \dfrac{\lambda G}{1 - 2a\lambda} & \text{for } |x| \le a, \\ g(x) & \text{for } |x| > a. \end{cases}$$

> ***23.15*** *Use Fredholm theory to show that, for the kernel*
>
> $$K(x,z) = (x+z)\exp(x-z)$$
>
> *over the interval* $[0,1]$*, the resolvent kernel is*
>
> $$R(x,z;\lambda) = \frac{\exp(x-z)[(x+z) - \lambda(\frac{1}{2}x + \frac{1}{2}z - xz - \frac{1}{3})]}{1 - \lambda - \frac{1}{12}\lambda^2},$$
>
> *and hence solve*
>
> $$y(x) = x^2 + 2\int_0^1 (x+z)\exp(x-z)\,y(z)\,dz,$$
>
> *expressing your answer in terms of* I_n*, where* $I_n = \int_0^1 u^n \exp(-u)\,du$*.*

We calculate successive values of d_n and $D_n(x,z)$ using the Fredholm recurrence relations:

$$d_n = \int_a^b D_{n-1}(x,x)\,dx,$$

$$D_n(x,z) = K(x,z)d_n - n\int_a^b K(x,z_1)D_{n-1}(z_1,z)\,dz_1,$$

starting from $d_0 = 1$ and $D_0(x,z) = (x+z)e^{x-z}$. In the first iteration we obtain

$$d_1 = \int_0^1 (u+u)e^{u-u}\,du = 1,$$

$$\begin{aligned}
D_1(x,z) &= (x+z)e^{x-z}(1) - 1\int_0^1 (x+u)e^{x-u}(u+z)e^{u-z}\,du \\
&= (x+z)e^{x-z} - e^{x-z}\int_0^1 [\,xz + (x+z)u + u^2\,]\,du \\
&= e^{x-z}[\,\tfrac{1}{2}(x+z) - xz - \tfrac{1}{3}\,].
\end{aligned}$$

Performing the second iteration gives

$$d_2 = \int_0^1 e^{u-u}(u - u^2 - \tfrac{1}{3})\,du = \tfrac{1}{2} - \tfrac{1}{3} - \tfrac{1}{3} = -\tfrac{1}{6},$$

$$\begin{aligned}
D_2(x,z) &= (x+z)e^{x-z}(-\tfrac{1}{6}) \\
&\qquad - 2\int_0^1 (x+u)e^{x-u}e^{u-z}\left[\,\tfrac{1}{2}(u+z) - uz - \tfrac{1}{3}\,\right]\,du \\
&= e^{x-z}\left\{-\tfrac{1}{6}(x+z) - 2\left[x\left(\frac{1}{4} + \frac{z}{2} - \frac{z}{2} - \frac{1}{3}\right) + \left(\frac{1}{6} + \frac{z}{4} - \frac{z}{3} - \frac{1}{6}\right)\right]\right\} \\
&= e^{x-z}\left\{-\tfrac{1}{6}(x+z) - 2\left[-\frac{x}{12} - \frac{z}{12}\right]\right\} = 0.
\end{aligned}$$

Since $D_2(x,z)=0$, $d_3=0$, $D_3(x,z)=0$, etc. Consequently both $D(x,z;\lambda)$ and $d(\lambda)$ are finite, rather than infinite, series:

$$D(x,z;\lambda)=(x+z)e^{x-z}-\lambda\left[\tfrac{1}{2}(x+z)-xz-\tfrac{1}{3}\right]e^{x-z},$$
$$d(\lambda)=1-\lambda+\left(-\tfrac{1}{6}\right)\frac{\lambda^2}{2!}=1-\lambda-\tfrac{1}{12}\lambda^2.$$

The resolvent kernel $R(x,z;\lambda)$, given by the ratio $D(x,z;\lambda)/d(\lambda)$, is therefore as stated in the question.

For the particular integral equation, $\lambda=2$ and $f(x)=x^2$. It follows that

$$d(\lambda)=1-2-\tfrac{4}{12}=-\tfrac{4}{3}\quad\text{and}\quad D(x,z\,:\lambda)=(2xz+\tfrac{2}{3})e^{x-z}.$$

The solution is therefore given by

$$\begin{aligned}
y(x)&=f(x)+\lambda\int_0^1 R(x,z;\lambda)f(z)\,dz\\
&=x^2+2\int_0^1\frac{(2xz+\frac{2}{3})z^2e^{x-z}}{-\frac{4}{3}}\,dz\\
&=x^2-\int_0^1(3xz^3+z^2)e^{x-z}\,dz\\
&=x^2-(3xI_3+I_2)e^x.
\end{aligned}$$

24

Complex variables

24.1 *Find an analytic function of $z = x + iy$ whose imaginary part is*

$$(y\cos y + x\sin y)\exp x.$$

If the required function is $f(z) = u + iv$, with $v = (y\cos y + x\sin y)\exp x$, then, from the Cauchy–Riemann equations,

$$\frac{\partial v}{\partial x} = e^x(y\cos y + x\sin y + \sin y) = -\frac{\partial u}{\partial y}.$$

Integrating with respect to y gives

$$\begin{aligned}
u &= -e^x \int (y\cos y + x\sin y + \sin y)\,dy + f(x) \\
&= -e^x \left(y\sin y - \int \sin y\,dy - x\cos y - \cos y \right) + f(x) \\
&= -e^x(y\sin y + \cos y - x\cos y - \cos y) + f(x) \\
&= e^x(x\cos y - y\sin y) + f(x).
\end{aligned}$$

We determine $f(x)$ by applying the second Cauchy–Riemann equation, which equates $\partial u/\partial x$ with $\partial v/\partial y$:

$$\begin{aligned}
\frac{\partial u}{\partial x} &= e^x(x\cos y - y\sin y + \cos y) + f'(x), \\
\frac{\partial v}{\partial y} &= e^x(\cos y - y\sin y + x\cos y).
\end{aligned}$$

By comparison, $\qquad f'(x) = 0 \quad \Rightarrow \quad f(x) = k,$

where k is a real constant that can be taken as zero. Hence, the analytic function is given by

$$\begin{aligned} f(z) = u + iv &= e^x(x\cos y - y\sin y + iy\cos y + ix\sin y) \\ &= e^x[(\cos y + i\sin y)(x + iy)] \\ &= e^x\, e^{iy}(x + iy) \\ &= ze^z. \end{aligned}$$

The final line confirms explicitly that this is a function of z alone (as opposed to a function of both z and z^*).

24.3 *Find the radii of convergence of the following Taylor series:*

$$\text{(a)} \sum_{n=2}^{\infty} \frac{z^n}{\ln n}, \qquad \text{(b)} \sum_{n=1}^{\infty} \frac{n!z^n}{n^n},$$
$$\text{(c)} \sum_{n=1}^{\infty} z^n n^{\ln n}, \qquad \text{(d)} \sum_{n=1}^{\infty} \left(\frac{n+p}{n}\right)^{n^2} z^n, \textit{ with } p \textit{ real.}$$

In each case we consider the series as $\sum_n a_n z^n$ and apply the formula

$$\frac{1}{R} = \lim_{n\to\infty} |a_n|^{1/n}$$

derived from considering the Cauchy root test for absolute convergence.

(a) $$\frac{1}{R} = \lim_{n\to\infty} \left(\frac{1}{\ln n}\right)^{1/n} = 1, \text{ since } -n^{-1}\ln\ln n \to 0 \text{ as } n \to \infty.$$

Thus $R = 1$. For interest, we also note that at the point $z = 1$ the series is

$$\sum_{n=2}^{\infty} \frac{1}{\ln n} > \sum_{n=2}^{\infty} \frac{1}{n},$$

which diverges. This shows that the given series diverges at this point on its circle of convergence.

(b) $$\frac{1}{R} = \lim_{n\to\infty} \left(\frac{n!}{n^n}\right)^{1/n}.$$

Since the nth root of $n!$ tends to n as $n \to \infty$, the limit of this ratio is that of n/n, namely unity. Thus $R = 1$ and the series converges inside the unit circle.

(c) $$\begin{aligned} \frac{1}{R} &= \lim_{n\to\infty} \left(n^{\ln n}\right)^{1/n} = \lim_{n\to\infty} n^{(\ln n)/n} \\ &= \lim_{n\to\infty} \exp\left[\frac{\ln n}{n} \ln n\right] = \exp(0) = 1. \end{aligned}$$

Thus $R = 1$ and the series converges inside the unit circle. It is obvious that the series diverges at the point $z = 1$.

(d) $$\frac{1}{R} = \lim_{n\to\infty}\left[\left(\frac{n+p}{n}\right)^{n^2}\right]^{1/n} = \lim_{n\to\infty}\left(\frac{n+p}{n}\right)^{n}$$
$$= \lim_{n\to\infty}\left(1+\frac{p}{n}\right)^{n} = e^{p}.$$

Thus $R = e^{-p}$ and the series converges inside a circle of this radius centred on the origin $z = 0$.

***24.5** Determine the types of singularities (if any) possessed by the following functions at $z = 0$ and $z = \infty$:*

(a) $(z-2)^{-1}$, (b) $(1+z^3)/z^2$, (c) $\sinh(1/z)$,
(d) e^z/z^3, (e) $z^{1/2}/(1+z^2)^{1/2}$.

(a) Although $(z-2)^{-1}$ has a simple pole at $z = 2$, at both $z = 0$ and $z = \infty$ it is well behaved and analytic.

(b) Near $z = 0$, $f(z) = (1+z^3)/z^2$ behaves like $1/z^2$ and so has a double pole there. It is clear that as $z \to \infty$ $f(z)$ behaves as z and so has a simple pole there; this can be made more formal by setting $z = 1/\xi$ to obtain $g(\xi) = \xi^2 + \xi^{-1}$ and considering $\xi \to 0$. This leads to the same conclusion.

(c) As $z \to \infty$, $f(z) = \sinh(1/z)$ behaves like $\sinh\xi$ as $\xi \to 0$, i.e. analytically. However, the definition of the sinh function involves an infinite series — in this case an infinite series of inverse powers of z. Thus, no finite n for which

$$\lim_{z\to 0}[\,z^n f(z)\,] \text{ is finite}$$

can be found, and $f(z)$ has an essential singularity at $z = 0$.

(d) Near $z = 0$, $f(z) = e^z/z^3$ behaves as $1/z^3$ and has a pole of order 3 at the origin. At $z = \infty$ it has an obvious essential singularity; formally, the series expansion of $e^{1/\xi}$ about $\xi = 0$ contains arbitrarily high inverse powers of ξ.

(e) Near $z = 0$, $f(z) = z^{1/2}/(1+z^2)^{1/2}$ behaves as $z^{1/2}$ and therefore has a branch point there. To investigate its behaviour as $z \to \infty$, we set $z = 1/\xi$ and obtain

$$f(z) = g(\xi) = \left(\frac{\xi^{-1}}{1+\xi^{-2}}\right)^{1/2} = \left(\frac{\xi}{\xi^2+1}\right)^{1/2} \sim \xi^{1/2} \text{ as } \xi \to 0.$$

Hence $f(z)$ also has a branch point at $z = \infty$.

24.7 *Find the real and imaginary parts of the functions* (i) z^2, (ii) e^z, *and* (iii) $\cosh \pi z$. *By considering the values taken by these parts on the boundaries of the region* $x \geq 0$, $y \leq 1$, *determine the solution of Laplace's equation in that region that satisfies the boundary conditions*

$$\phi(x,0) = 0, \qquad \phi(0,y) = 0,$$
$$\phi(x,1) = x, \qquad \phi(1,y) = y + \sin \pi y.$$

Writing $f_k(z) = u_k(x,y) + iv_k(x,y)$, we have

$$\begin{aligned}
\text{(i)} \quad f_1(z) &= z^2 = (x+iy)^2 \\
&\Rightarrow u_1 = x^2 - y^2 \text{ and } v_1 = 2xy, \\
\text{(ii)} \quad f_2(z) &= e^z = e^{x+iy} = e^x(\cos y + i \sin y) \\
&\Rightarrow u_2 = e^x \cos y \text{ and } v_2 = e^x \sin y, \\
\text{(iii)} \quad f_3(z) &= \cosh \pi z = \cosh \pi x \cos \pi y + i \sinh \pi x \sin \pi y \\
&\Rightarrow u_3 = \cosh \pi x \cos \pi y \text{ and } v_3 = \sinh \pi x \sin \pi y.
\end{aligned}$$

All of these u and v are necessarily solutions of Laplace's equation (this follows from the Cauchy–Riemann equations), and, since Laplace's equation is linear, we can form any linear combination of them and it will also be a solution. We need to choose the combination that matches the given boundary conditions.

Since the third and fourth conditions involve x and $\sin \pi y$, and these appear only in v_1 and v_3, respectively, let us try a linear combination of them:

$$\phi(x,y) = A(2xy) + B(\sinh \pi x \sin \pi y).$$

The requirement $\phi(x,0) = 0$ is clearly satisfied, as is $\phi(0,y) = 0$. The condition $\phi(x,1) = x$ becomes $2Ax + 0 = x$, requiring $A = \frac{1}{2}$, and the ,remaining condition, $\phi(1,y) = y + \sin \pi y$, takes the form $y + B \sinh \pi \sin \pi y = y + \sin \pi y$, thus determining B as $1/\sinh \pi$.

With ϕ a solution of Laplace's equation and all of the boundary conditions satisfied, the uniqueness theorem guarantees that

$$\phi(x,y) = xy + \frac{\sinh \pi x \sin \pi y}{\sinh \pi}$$

is the correct solution.

24.9 *The* fundamental theorem of algebra *states that, for a complex polynomial $p_n(z)$ of degree n, the equation $p_n(z) = 0$ has precisely n complex roots. By applying Liouville's theorem, which reads*

If $f(z)$ is analytic and bounded for all z then f is a constant,

to $f(z) = 1/p_n(z)$, prove that $p_n(z) = 0$ has at least one complex root. Factor out that root to obtain $p_{n-1}(z)$ and, by repeating the process, prove the fundamental theorem.

We prove this result by the method of contradiction. Suppose $p_n(z) = 0$ has no roots in the complex plane, then $f_n(z) = 1/p_n(z)$ is bounded for all z and, by Liouville's theorem, is therefore a constant. It follows that $p_n(z)$ is also a constant and that $n = 0$. However, if $n > 0$ we have a contradiction and it was wrong to suppose that $p_n(z) = 0$ has no roots; it must have at least one. Let one of them be $z = z_1$; i.e. $p_n(z)$, being a polynomial, can be written $p_n(z) = (z - z_1)p_{n-1}(z)$.

Now, by considering $f_{n-1}(z) = 1/p_{n-1}(z)$ in just the same way, we can conclude that either $n - 1 = 0$ or a further reduction is possible. It is clear that n such reductions are needed to make f_0 a constant, thus establishing that $p_n(z) = 0$ has precisely n (complex) roots.

Many of the remaining exercises in this chapter involve contour integration and the choice of a suitable contour. In order to save the space taken by drawing several broadly similar contours that differ only in notation, the positions of poles, the values of lengths or angles, or other minor details, we show in figure 24.1 a number of typical contour types to which reference can be made.

24.11 *The function*

$$f(z) = (1 - z^2)^{1/2}$$

of the complex variable z is defined to be real and positive on the real axis for $-1 < x < 1$. Using cuts running along the real axis for $1 < x < +\infty$ and $-\infty < x < -1$, show how $f(z)$ is made single-valued and evaluate it on the upper and lower sides of both cuts.

Use these results and a suitable contour in the complex z-plane to evaluate the integral

$$I = \int_1^\infty \frac{dx}{x(x^2 - 1)^{1/2}}.$$

Confirm your answer by making the substitution $x = \sec\theta$.

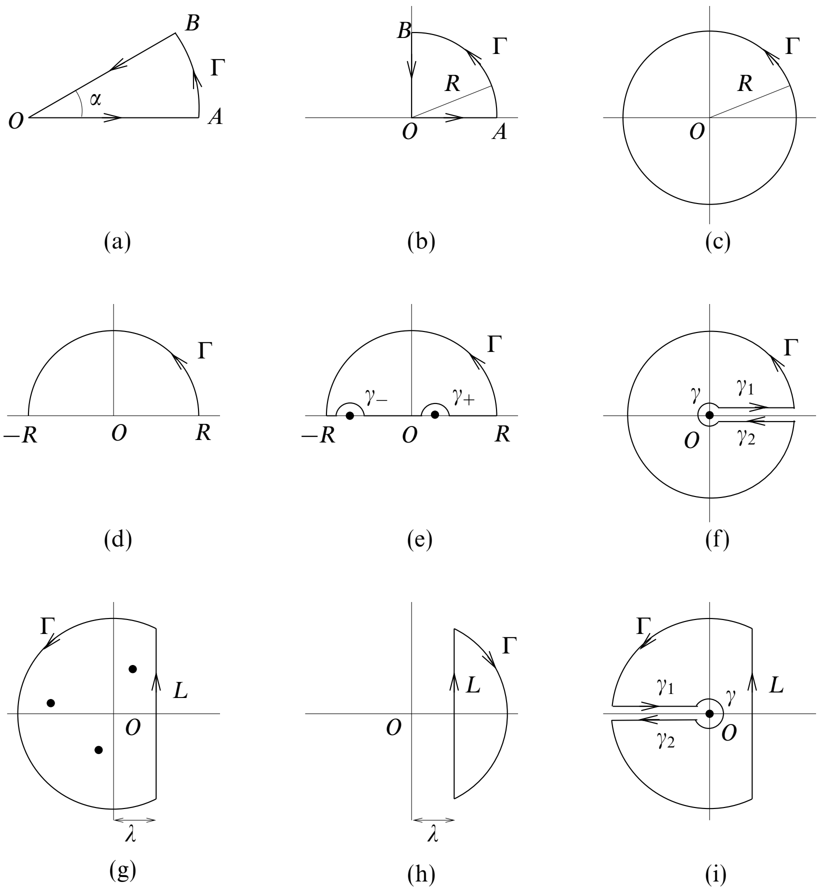

Figure 24.1 Typical contours for use in contour integration.

As usual when dealing with branch cuts aimed at making a multi-valued function into a single-valued one, we introduce polar coordinates centred on the branch points. For $f(z)$ the branch points are at $z = \pm 1$, and so we define r_1 as the distance of z from the point 1 and θ_1 as the angle the line joining 1 to z makes with the part of the x-axis for which $1 < x < +\infty$, with $0 \le \theta_1 \le 2\pi$. Similarly, r_2 and θ_2 are centred on the point -1, but θ_2 lies in the range $-\pi \le \theta_2 \le \pi$.

With these definitions,

$$\begin{aligned} f(z) = (1 - z^2)^{1/2} &= (1 - z)^{1/2}(1 + z)^{1/2} \\ &= \left[(-r_1 e^{i\theta_1})(r_2 e^{i\theta_2}) \right]^{1/2} \\ &= (r_1 r_2)^{1/2} e^{i(\theta_1 + \theta_2 - \pi)/2}. \end{aligned}$$

In the final line the choice between $\exp(+i\pi)$ and $\exp(-i\pi)$ for dealing with the minus sign appearing before r_1 in the second line was resolved by the requirement that $f(z)$ is real and positive when $-1 < x < 1$ with $y = 0$. For these values of z, $r_1 = 1 - x$, $r_2 = 1 + x$, $\theta_1 = \pi$ and $\theta_2 = 0$. Thus,

$$f(z) = [(1-x)(1+x)]^{1/2}\, e^{(\pi+0-\pi)/2} = (1-x^2)^{1/2} e^{i0} = +(1-x^2)^{1/2},$$

as required.

Now applying the same prescription to points lying just above and just below each of the cuts, we have

$$\begin{array}{lll}
x > 1, y = 0_+ & & r_1 = x-1 \quad r_2 = x+1 \quad \theta_1 = 0 \quad \theta_2 = 0 \\
 & \Rightarrow & f(z) = (x^2-1)^{1/2} e^{i(0+0-\pi)/2} = -i(x^2-1)^{1/2}, \\
x > 1, y = 0_- & & r_1 = x-1 \quad r_2 = x+1 \quad \theta_1 = 2\pi \quad \theta_2 = 0 \\
 & \Rightarrow & f(z) = (x^2-1)^{1/2} e^{i(2\pi+0-\pi)/2} = i(x^2-1)^{1/2}, \\
x < -1, y = 0_+ & & r_1 = 1-x \quad r_2 = -x-1 \quad \theta_1 = \pi \quad \theta_2 = \pi \\
 & \Rightarrow & f(z) = (x^2-1)^{1/2} e^{i(\pi+\pi-\pi)/2} = i(x^2-1)^{1/2}, \\
x < -1, y = 0_- & & r_1 = 1-x \quad r_2 = -x-1 \quad \theta_1 = \pi \quad \theta_2 = -\pi \\
 & \Rightarrow & f(z) = (x^2-1)^{1/2} e^{i(\pi-\pi-\pi)/2} = -i(x^2-1)^{1/2}.
\end{array}$$

To use these results to evaluate the given integral I, consider the contour integral

$$J = \int_C \frac{dz}{z(1-z^2)^{1/2}} = \int_c \frac{dz}{zf(z)}.$$

Here C is a large circle (consisting of arcs Γ_1 and Γ_2 in the upper and lower half-planes, respectively) of radius R centred on the origin but indented along the positive and negative x-axes by the cuts considered earlier. At the ends of the cuts are two small circles γ_1 and γ_2 that enclose the branch points $z = 1$ and $z = -1$, respectively. Thus the complete closed contour, starting from γ_1 and moving along the positive real axis, consists of, in order, circle γ_1, cut C_1, arc Γ_1, cut C_2, circle γ_2, cut C_3, arc Γ_2 and cut C_4, leading back to γ_1.

On the arcs Γ_1 and Γ_2 the integrand is $O(R^{-2})$ and the contributions to the contour integral $\to 0$ as $R \to \infty$. For the small circle γ_1, where we can set $z = 1 + \rho e^{i\phi}$ with $dz = i\rho e^{i\phi}\, d\phi$, we have

$$\int_{\gamma_1} \frac{dz}{z(1+z)^{1/2}(1-z)^{1/2}} = \int_0^{2\pi} \frac{i\rho e^{i\phi}}{(1+\rho e^{i\phi})(2+\rho e^{i\phi})^{1/2}(-\rho e^{i\phi})^{1/2}}\, d\phi,$$

and this $\to 0$ as $\rho \to 0$. Similarly, the small circle γ_2 contributes nothing to the contour integral. This leaves only the contributions from the four arms of the branch cuts. To relate these to I we use our previous results about the value of $f(z)$ on the various arms:

$$\text{on } C_1,\ z = x \text{ and } \int_{C_1} = \int_1^\infty \frac{dx}{x[-i(x^2-1)^{1/2}]} = iI;$$

$$\text{on } C_2,\ z = -x \text{ and } \int_{C_2} = \int_\infty^1 \frac{-dx}{-x[i(x^2-1)^{1/2}]} = iI;$$

$$\text{on } C_3,\ z = -x \text{ and } \int_{C_3} = \int_1^\infty \frac{-dx}{-x[-i(x^2-1)^{1/2}]} = iI;$$

$$\text{on } C_4,\ z = x \text{ and } \int_{C_1} = \int_\infty^1 \frac{dx}{x[i(x^2-1)^{1/2}]} = iI.$$

So the full contour integral around C has the value $4iI$. But, this must be the same as $2\pi i$ times the residue of $z^{-1}(1-z^2)^{-1/2}$ at $z = 0$, which is the only pole of the integrand inside the contour. The residue is clearly unity, and so we deduce that $I = \pi/2$.

This particular integral can be evaluated much more simply using elementary methods. Setting $x = \sec\theta$ with $dx = \sec\theta\tan\theta\, d\theta$ gives

$$\begin{aligned} I &= \int_1^\infty \frac{dx}{x(x^2-1)^{1/2}} \\ &= \int_0^{\pi/2} \frac{\sec\theta\tan\theta\, d\theta}{\sec\theta\,(\sec^2\theta-1)^{1/2}} = \int_0^{\pi/2} d\theta = \frac{\pi}{2}, \end{aligned}$$

and so verifies the result obtained by contour integration.

***24.13** Prove that if $f(z)$ has a simple zero at z_0 then $1/f(z)$ has residue $1/f'(z_0)$ there. Hence evaluate*

$$\int_{-\pi}^{\pi} \frac{\sin\theta}{a-\sin\theta}\, d\theta,$$

where a is real and > 1.

If $f(z)$ is analytic and has a simple zero at $z = z_0$ then it can be written as

$$f(z) = \sum_{n=1}^{\infty} a_n(z-z_0)^n, \qquad \text{with } a_1 \neq 0.$$

Using a binomial expansion,

$$\frac{1}{f(z)} = \frac{1}{a_1(z-z_0)\left(1+\sum_{n=2}^{\infty}\frac{a_n}{a_1}(z-z_0)^{n-1}\right)}$$
$$= \frac{1}{a_1(z-z_0)}(1+b_1(z-z_0)+b_2(z-z_0)^2+\cdots),$$

for some coefficients, b_i. The residue at $z = z_0$ is clearly a_1^{-1}.

But, from differentiating the Taylor expansion,

$$f'(z) = \sum_{n=1}^{\infty} na_n(z-z_0)^{n-1},$$
$$\Rightarrow \quad f'(z_0) = a_1 + 0 + 0 + \cdots = a_1,$$

i.e. the residue $= \dfrac{1}{a_1}$ can also be expressed as $\dfrac{1}{f'(z_0)}$.

Denote the required integral by I and consider the contour integral

$$J = \int_C \frac{dz}{a - \dfrac{1}{2i}(z - z^{-1})} = \int_C \frac{2iz\,dz}{2aiz - z^2 + 1},$$

where C is the unit circle, i.e. contour (c) of figure 24.1 with $R = 1$. The denominator has simple zeros at $z = ai \pm \sqrt{-a^2+1} = i(a \pm \sqrt{a^2-1})$. Since a is strictly greater than 1, $\alpha = i(a - \sqrt{a^2-1})$ lies strictly inside the unit circle, whilst $\beta = i(a + \sqrt{a^2-1})$ lies strictly outside it (and need not be considered further).

Extending the previous result to the case of $h(z) = g(z)/f(z)$, where $g(z)$ is analytic at z_0, the residue of $h(z)$ at $z = z_0$ can be seen to be $g(z_0)/f'(z_0)$. Applying this, we find that the residue of the integrand at $z = \alpha$ is given by

$$\left|\frac{2iz}{2ai - 2z}\right|_{z=\alpha} = \frac{i\alpha}{ai - ai + i\sqrt{a^2-1}} = \frac{\alpha}{\sqrt{a^2-1}}.$$

Now on the unit circle, $z = e^{i\theta}$ with $dz = i\,e^{i\theta}\,d\theta$, and J can be written as

$$J = \int_{-\pi}^{\pi} \frac{i\,e^{i\theta}\,d\theta}{a - \dfrac{1}{2i}(\,e^{i\theta} - \,e^{-i\theta})} = \int_{-\pi}^{\pi} \frac{i(\cos\theta + i\sin\theta)\,d\theta}{a - \sin\theta}.$$

Hence,

$$I = -\mathrm{Re}\,J = -\mathrm{Re}\,2\pi i\frac{i(a - \sqrt{a^2-1})}{\sqrt{a^2-1}}$$
$$= 2\pi\left(\frac{a}{\sqrt{a^2-1}} - 1\right).$$

Although it is not asked for, we can also deduce from the fact that the residue at $z = \alpha$ is purely imaginary that

$$\int_{-\pi}^{\pi} \frac{\cos\theta}{a - \sin\theta}\, d\theta = 0,$$

a result that can also be obtained by more elementary means, when it is noted that the numerator of the integrand is the derivative of the denominator.

24.15 *Prove that*

$$\int_0^{\infty} \frac{\cos mx}{4x^4 + 5x^2 + 1}\, dx = \frac{\pi}{6}\left(4e^{-m/2} - e^{-m}\right) \qquad \textit{for } m > 0.$$

Since, when z is on the real axis, the integrand is equal to

$$\text{Re}\, \frac{e^{imz}}{(z^2 + 1)(4z^2 + 1)} = \text{Re}\, \frac{e^{imz}}{(z + i)(z - i)(2z + i)(2z - i)},$$

we consider the integral of $f(z) = \dfrac{e^{imz}}{(z + i)(z - i)(2z + i)(2z - i)}$ around contour (d) in figure 24.1.

As $|f(z)| \sim |z|^{-4}$ as $z \to \infty$ and $m > 0$, all the conditions for Jordan's lemma to hold are satisfied and the integral around the large semicircle contributes nothing. For this integrand there are two poles inside the contour, at $z = i$ and at $z = \frac{1}{2}i$. The respective residues are

$$\frac{e^{-m}}{2i\, 3i\, i} = \frac{ie^{-m}}{6} \quad \text{and} \quad \frac{e^{-m/2}}{\frac{3i}{2}\,(-\frac{i}{2})\, 2i} = \frac{-2ie^{-m/2}}{3}.$$

The residue theorem therefore reads

$$\int_{-\infty}^{\infty} \frac{e^{imx}}{4x^4 + 5x^2 + 1}\, dx + 0 = 2\pi i\left(\frac{ie^{-m}}{6} - \frac{2ie^{-m/2}}{3}\right),$$

and the stated result follows from equating real parts and changing the lower integration limit, recognising that the integrand is symmetric about $x = 0$ and so the integral from 0 to ∞ is equal to half of that from $-\infty$ to ∞.

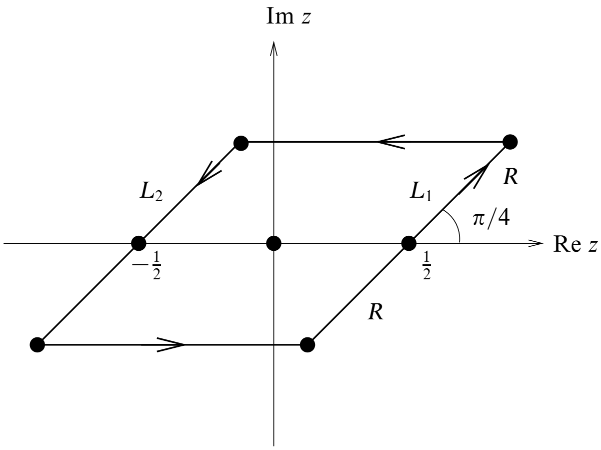

Figure 24.2 The contour used in exercise 24.17.

24.17 *The following is an alternative (and roundabout!) way of evaluating the Gaussian integral.*

(a) *Prove that the integral of* $[\exp(i\pi z^2)]\operatorname{cosec}\pi z$ *around the parallelogram with corners* $\pm 1/2 \pm R\exp(i\pi/4)$ *has the value* $2i$.
(b) *Show that the parts of the contour parallel to the real axis give no contribution when* $R \to \infty$.
(c) *Evaluate the integrals along the other two sides by putting* $z' = r\exp(i\pi/4)$ *and working in terms of* $z' + \frac{1}{2}$ *and* $z' - \frac{1}{2}$. *Hence by letting* $R \to \infty$ *show that*

$$\int_{-\infty}^{\infty} e^{-\pi r^2}\, dr = 1.$$

The integral is

$$\int_C e^{i\pi z^2} \operatorname{cosec}\, \pi z\, dz = \int_C \frac{e^{i\pi z^2}}{\sin \pi z}\, dz$$

and the suggested contour C is shown in figure 24.2.

(a) The integrand has (simple) poles only on the real axis at $z = n$, where n is an integer. The only such pole enclosed by C is at $z = 0$. The residue there is

$$a_{-1} = \lim_{z\to 0} \frac{z e^{i\pi z^2}}{\sin \pi z} = \frac{1}{\pi}.$$

The value of the integral around C is therefore $2\pi i \times (\pi^{-1}) = 2i$.

(b) On the parts of C parallel to the real axis, $z = \pm Re^{i\pi/4} + x'$, where $-\frac{1}{2} \le x' \le \frac{1}{2}$. The integrand is thus given by

$$\begin{aligned}
f(z) &= \frac{1}{\sin \pi z} \exp\left[\, i\pi(\pm Re^{i\pi/4} + x')^2 \,\right] \\
&= \frac{1}{\sin \pi z} \exp\left[\, i\pi \left(R^2 e^{i\pi/2} \pm 2Rx' e^{i\pi/4} + x'^2 \right) \right] \\
&= \frac{1}{\sin \pi z} \exp\left[-\pi R^2 \pm \frac{2\pi i R x'}{\sqrt{2}}(1+i) + i\pi x'^2 \right] \\
&= \mathrm{O}\left(\exp[-\pi R^2 \mp \sqrt{2}\pi R x'\,] \right) \\
&\to 0 \text{ as } R \to \infty.
\end{aligned}$$

Since the integration range is finite ($-\frac{1}{2} \le x' \le \frac{1}{2}$), the integrals $\to 0$ as $R \to \infty$.

(c) On the first of the other two sides, let us set $z = \frac{1}{2} + re^{i\pi/4}$ with $-R \le r \le R$. The corresponding integral I_1 is

$$\begin{aligned}
I_1 &= \int_{L_1} e^{i\pi z^2} \operatorname{cosec} \pi z \, dz \\
&= \int_{-R}^{R} \frac{\exp\left[\, i\pi(\frac{1}{2} + re^{i\pi/4})^2 \,\right]}{\sin\left[\, \pi(\frac{1}{2} + re^{i\pi/4}) \,\right]} e^{i\pi/4} \, dr \\
&= \int_{-R}^{R} \frac{e^{i\pi/4} \exp(i\pi r e^{i\pi/4}) \exp(i\pi r^2 i) e^{i\pi/4}}{\cos(\pi r e^{i\pi/4})} \, dr \\
&= \int_{-R}^{R} \frac{i \exp(i\pi r e^{i\pi/4}) e^{-\pi r^2}}{\cos(\pi r e^{i\pi/4})} \, dr.
\end{aligned}$$

Similarly (remembering the sense of integration), the remaining side contributes

$$I_2 = -\int_{-R}^{R} \frac{i \exp(-i\pi r e^{i\pi/4}) e^{-\pi r^2}}{-\cos(\pi r e^{i\pi/4})} \, dr.$$

Adding together all four contributions gives

$$0 + 0 + \int_{-R}^{R} \frac{i[\exp(i\pi r e^{i\pi/4}) + \exp(-i\pi r e^{i\pi/4})\,] e^{-\pi r^2}}{\cos(\pi r e^{i\pi/4})} \, dr,$$

which simplifies to

$$\int_{-R}^{R} 2ie^{-\pi r^2} \, dr.$$

From part (a), this must be equal to $2i$ as $R \to \infty$, and so $\displaystyle\int_{-\infty}^{\infty} e^{-\pi r^2} \, dr = 1.$

24.19 *Using a suitable cut plane, prove that if α is real and $0 < \alpha < 1$ then*

$$\int_0^\infty \frac{x^{-\alpha}}{1+x}\,dx$$

has the value $\pi \operatorname{cosec} \pi\alpha$.

As α is not an integer, the complex form of the integrand $f(z) = \dfrac{z^{-\alpha}}{1+z}$ is not single-valued. We therefore need to perform the contour integration in a cut plane; contour (f) of figure 24.1 is a suitable contour. We will be making use of the fact that, because the integrand takes different values on γ_1 and γ_2, the contributions coming from these two parts of the complete contour, although related, do *not* cancel.

The contributions from γ and Γ are both zero because:

(i) around γ, $|zf(z)| \sim \dfrac{z\,z^{-\alpha}}{1} = z^{1-\alpha} \to 0$ as $|z| \to 0$, since $\alpha < 1$;

(ii) around Γ, $|zf(z)| \sim \dfrac{z\,z^{-\alpha}}{z} = z^{-\alpha} \to 0$ as $|z| \to \infty$, since $\alpha > 0$.

Therefore, the only contributions come from the cut; on γ_1, $z = xe^{0i}$, whilst on γ_2, $z = xe^{2\pi i}$. The only pole inside the contour is a simple one at $z = -1 = e^{i\pi}$, where the residue is $e^{-i\pi\alpha}$. The residue theorem now reads

$$0 + \int_0^\infty \frac{x^{-\alpha}}{1+x}\,dx + 0 - \int_0^\infty \frac{x^{-\alpha}e^{-2\pi i\alpha}}{1+xe^{2\pi i}}\,dx = 2\pi i\, e^{-i\pi\alpha},$$

$$\Rightarrow \quad (1 - e^{-2\pi i\alpha}) \int_0^\infty \frac{x^{-\alpha}}{1+x}\,dx = 2\pi i\, e^{-i\pi\alpha}.$$

This can be rearranged to read

$$\int_0^\infty \frac{x^{-\alpha}}{1+x}\,dx = \frac{2\pi i\, e^{-i\pi\alpha}}{(1 - e^{-2\pi i\alpha})} = \frac{2\pi i}{e^{i\pi\alpha} - e^{-i\pi\alpha}} = \frac{\pi}{\sin \pi\alpha},$$

thus establishing the stated result.

24.21 *By integrating a suitable function around a large semicircle in the upper half plane and a small semicircle centred on the origin, determine the value of*

$$I = \int_0^\infty \frac{(\ln x)^2}{1+x^2}\,dx$$

and deduce, as a by-product of your calculation, that

$$\int_0^\infty \frac{\ln x}{1+x^2}\,dx = 0.$$

The suggested contour is that shown in figure 24.1(e), but with only one indentation γ on the real axis (at $z = 0$) and with $R = \infty$. The appropriate complex function is

$$f(z) = \frac{(\ln z)^2}{1 + z^2}.$$

The only pole inside the contour is at $z = i$, and the residue there is given by

$$\frac{(\ln i)^2}{i + i} = \frac{(\ln 1 + i(\pi/2))^2}{2i} = -\frac{\pi^2}{8i}.$$

To evaluate the integral around γ, we set $z = \rho\, e^{i\theta}$ with $\ln z = \ln \rho + i\theta$ and $dz = i\rho\, e^{i\theta}\, d\theta$; the integral becomes

$$\int_{\pi}^{0} \frac{\ln^2 \rho + 2i\theta \ln \rho - \theta^2}{1 + \rho^2 e^{2i\theta}}\, i\rho\, e^{i\theta}\, d\theta, \text{ which} \to 0 \text{ as } \rho \to 0.$$

Thus γ contributes nothing. Even more obviously, on Γ, $|zf(z)| \sim z^{-1}$ and tends to zero as $|z| \to \infty$, showing that Γ also contributes nothing.

On γ_+, $z = xe^{i0}$ and the contribution is equal to I.

On γ_-, $z = xe^{i\pi}$ and the contribution is (remembering that the contour actually runs from $x = \infty$ to $x = 0$) given by

$$\begin{aligned} I_- &= -\int_0^{\infty} \frac{(\ln x + i\pi)^2}{1 + x^2}\, e^{i\pi}\, dx \\ &= I + 2i\pi \int_0^{\infty} \frac{\ln x}{1 + x^2}\, dx - \pi^2 \int_0^{\infty} \frac{1}{1 + x^2}\, dx. \end{aligned}$$

The residue theorem for the complete closed contour thus reads

$$0 + I + 0 + I + 2i\pi \int_0^{\infty} \frac{\ln x}{1 + x^2}\, dx - \pi^2 \left[\tan^{-1} x \right]_0^{\infty} = 2\pi i \left(\frac{-\pi^2}{8i} \right).$$

Equating the real parts $\quad \Rightarrow \quad 2I - \frac{1}{2}\pi^3 = -\frac{1}{4}\pi^3 \quad \Rightarrow \quad I = \frac{1}{8}\pi^3.$

Equating the imaginary parts gives the stated by-product.

25

Applications of complex variables

Many of the exercises in this chapter involve contour integration and the choice of a suitable contour. In order to save the space taken by drawing several broadly similar contours that differ only in notation, the positions of poles, the values of lengths or angles, or other minor details, we make reference to figure 24.1 which shows a number of typical contour types.

25.1 *In the method of complex impedances for a.c. circuits, an inductance L is represented by a complex impedance $Z_L = i\omega L$ and a capacitance C by $Z_C = 1/(i\omega C)$. Kirchhoff's circuit laws,*

$$\sum_i I_i = 0 \text{ at a node and } \sum_i Z_i I_i = \sum_j V_j \text{ around any closed loop,}$$

are then applied as if the circuit were a d.c. one.

Apply this method to the a.c. bridge connected as in figure 25.1 to show that if the resistance R is chosen as $R = (L/C)^{1/2}$ then the amplitude of the current I_R through it is independent of the angular frequency ω of the applied a.c. voltage $V_0\, e^{i\omega t}$.

Determine how the phase of I_R, relative to that of the voltage source, varies with the angular frequency ω.

Omitting the common factor $e^{i\omega t}$ from all currents and voltages, let the current drawn from the voltage source be (the complex quantity) I and the current flowing from A to D be I_1. Then the currents in the remaining branches are $AE : I - I_1$, $DB : I_1 - I_R$ and $EB : I - I_1 + I_R$.

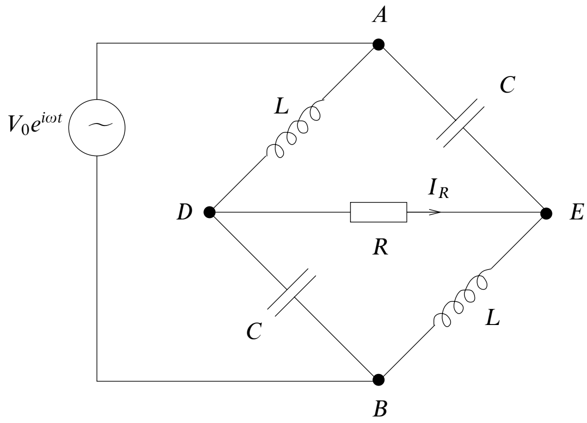

Figure 25.1 The inductor–capacitor–resistor network for exercise 25.1.

Applying $\sum_i Z_i I_i = \sum_j V_j$ to three separate loops yields

$$\text{loop } ADBA \qquad i\omega L I_1 + \frac{1}{i\omega C}(I_1 - I_R) = V_0,$$

$$\text{loop } ADEA \qquad i\omega L I_1 + R I_R - \frac{1}{i\omega C}(I - I_1) = 0,$$

$$\text{loop } DBED \qquad \frac{1}{i\omega C}(I_1 - I_R) - i\omega L(I - I_1 + I_R) - R I_R = 0.$$

Now, denoting $(LC)^{-1}$ by ω_0^2 and choosing R as $(L/C)^{1/2} = (\omega_0 C)^{-1}$, we can write these equations as follows:

$$\left(1 - \frac{\omega^2}{\omega_0^2}\right) I_1 - I_R = i\omega C V_0,$$

$$-I + \left(1 - \frac{\omega^2}{\omega_0^2}\right) I_1 + i\frac{\omega}{\omega_0} I_R = 0,$$

$$\frac{\omega^2}{\omega_0^2} I + \left(1 - \frac{\omega^2}{\omega_0^2}\right) I_1 + \left(-1 + \frac{\omega^2}{\omega_0^2} - i\frac{\omega}{\omega_0}\right) I_R = 0.$$

Eliminating I from the last two of these yields

$$\left(1 + \frac{\omega^2}{\omega_0^2}\right)\left(1 - \frac{\omega^2}{\omega_0^2}\right) I_1 - \left(\frac{i\omega}{\omega_0} + 1\right)\left(1 - \frac{\omega^2}{\omega_0^2}\right) I_R = 0.$$

Thus,

$$I_R = \frac{1 + \dfrac{\omega^2}{\omega_0^2}}{1 + i\dfrac{\omega}{\omega_0}} I_1 = \frac{\omega_0^2 + \omega^2}{\omega_0(\omega_0 + i\omega)} \, \frac{\omega_0^2(i\omega C V_0 + I_R)}{\omega_0^2 - \omega^2}.$$

After some cancellation and rearrangement,

$$(\omega_0^2 - \omega^2) I_R = \omega_0(\omega_0 - i\omega)(i\omega C V_0 + I_R),$$
$$(i\omega\omega_0 - \omega^2) I_R = \omega_0\omega(i\omega_0 + \omega) C V_0,$$

and so

$$I_R = \omega_0 C V_0 \frac{i\omega_0 + \omega}{i\omega_0 - \omega} = \omega_0 C V_0 \frac{(i\omega_0 + \omega)(-i\omega_0 - \omega)}{(i\omega_0 - \omega)(-i\omega_0 - \omega)}$$
$$= \omega_0 C V_0 \frac{\omega_0^2 - \omega^2 - 2i\omega\omega_0}{\omega_0^2 + \omega^2}.$$

From this we can read off

$$|I_R| = \omega_0 C V_0 \frac{\left[(\omega^2 - \omega_0^2)^2 + 4\omega^2\omega_0^2\right]^{1/2}}{\omega_0^2 + \omega^2} = \omega_0 C V_0, \text{ i.e. independent of } \omega,$$

and

$$\phi = \text{phase of } I_R \; = \tan^{-1} \frac{-2\omega\omega_0}{\omega_0^2 - \omega^2}.$$

Thus I_R (which was arbitrarily and notionally defined as flowing from D to E in the equivalent d.c. circuit) has an imaginary part that is always negative but a real part that changes sign as ω passes through ω_0. Its phase ϕ, relative to that of the voltage source, therefore varies from 0 when ω is small to $-\pi$ when ω is large.

25.3 *For the function*

$$f(z) = \ln\left(\frac{z+c}{z-c}\right),$$

where c is real, show that the real part u of f is constant on a circle of radius $c\,\text{cosech}\, u$ centred on the point $z = c\coth u$. Use this result to show that the electrical capacitance per unit length of two parallel cylinders of radii a, placed with their axes $2d$ apart, is proportional to $[\cosh^{-1}(d/a)]^{-1}$.

From

$$f(z) = \ln\left(\frac{z+c}{z-c}\right) = \ln\left|\frac{z+c}{z-c}\right| + i\arg\left(\frac{z+c}{z-c}\right),$$

we have that

$$u = \ln\left|\frac{z+c}{z-c}\right| = \frac{1}{2}\ln\frac{(x+c)^2+y^2}{(x-c)^2+y^2} \quad \Rightarrow \quad e^{2u} = \frac{(x+c)^2+y^2}{(x-c)^2+y^2}.$$

The curve upon which $u(x,y)$ is constant is therefore given by

$$(x^2 - 2cx + c^2 + y^2)e^{2u} = x^2 + 2xc + c^2 + y^2.$$

This can be rewritten as

$$x^2(e^{2u}-1)-2xc(e^{2u}+1)+y^2(e^{2u}-1)+c^2(e^{2u}-1)=0,$$
$$x^2-2xc\frac{e^{2u}+1}{e^{2u}-1}+y^2+c^2=0,$$
$$x^2-2xc\coth u+y^2+c^2=0,$$

which, in conic-section form, becomes

$$(x-c\coth u)^2+y^2=c^2\coth^2 u-c^2=c^2\text{cosech}^2 u.$$

This is a circle with centre $(c\coth u, 0)$ and radius $|c\,\text{cosech}\, u|$.

Now consider two such circles with the same value of $|c\text{cosech}\ u|$, equal to a, but different values of u satisfying $c\coth u_1=-d$ and $c\coth u_2=+d$. These two equations imply that $u_1=-u_2$, corresponding physically to equal but opposite charges $-Q$ and $+Q$ placed on identical cylindrical conductors that coincide with the circles; the conductors are raised to potentials u_1 and u_2.

We have already established that we need $c\coth u_2=d$ and $c\,\text{cosech}\ u_2=a$. Dividing these two equations gives $\cosh u_2=d/a$.

The capacitance (per unit length) of the arrangement is given by the magnitude of the charge on one conductor divided by the potential difference between the conductors that results from the presence of that charge, i.e.

$$C=\frac{Q}{u_2-u_1}\propto\frac{1}{2u_2}=\frac{1}{2\cosh^{-1}(d/a)},$$

as stated in the question.

25.5 *By considering in turn the transformations*

$$z=\tfrac{1}{2}c(w+w^{-1})\quad \textit{and}\quad w=\exp\zeta,$$

where $z=x+iy$, $w=r\exp i\theta$, $\zeta=\xi+i\eta$ and c is a real positive constant, show that $z=c\cosh\zeta$ maps the strip $\xi\geq 0$, $0\leq\eta\leq 2\pi$, onto the whole z-plane. Which curves in the z-plane correspond to the lines $\xi=$ constant and $\eta=$ constant? Identify those corresponding to $\xi=0$, $\eta=0$ and $\eta=2\pi$.

The electric potential ϕ of a charged conducting strip $-c\leq x\leq c$, $y=0$, satisfies

$$\phi\sim -k\ln(x^2+y^2)^{1/2}\textit{for large values of }(x^2+y^2)^{1/2},$$

with ϕ constant on the strip. Show that $\phi=\text{Re}\ [-k\cosh^{-1}(z/c)]$ and that the magnitude of the electric field near the strip is $k(c^2-x^2)^{-1/2}$.

We first note that the combined transformation is given by

$$z = \frac{c}{2}(e^{\zeta} + e^{-\zeta}) = c\cosh\zeta \quad \Rightarrow \quad \zeta = \cosh^{-1}\frac{z}{c}.$$

The successive connections linking the strip in the ζ-plane and its image in the z-plane are

$$\begin{aligned} z &= c\cosh\zeta = c\cosh(\xi + i\eta) \\ &= c\cosh\xi\cos\eta + ic\sinh\xi\sin\eta, \text{ with } \xi > 0,\ 0 \le \eta \le 2\pi, \end{aligned}$$

$$re^{i\theta} = w = e^{\zeta} = e^{\xi}e^{i\eta}, \text{ with the strip as } 1 < r < \infty,\ 0 \le \theta \le 2\pi,$$

$$\begin{aligned} x + iy = z &= \frac{c}{2}(w + w^{-1}) \\ &= \frac{c}{2}[\,r(\cos\theta + i\sin\theta) + r^{-1}(\cos\theta - i\sin\theta)\,] \\ &= \frac{c}{2}\left(r + \frac{1}{r}\right)\cos\theta + i\frac{c}{2}\left(r - \frac{1}{r}\right)\sin\theta. \end{aligned}$$

This last expression for z and the previous specification of the strip in terms of r and θ show that both x and y can take all values, i.e. that the original strip in the ζ-plane is mapped onto the whole of the z-plane. From the two expressions for z we also see that $x = c\cosh\xi\cos\eta$ and $y = c\sinh\xi\sin\eta$.

For ξ constant, the contour in the xy-plane, obtained by eliminating η, is

$$\frac{x^2}{c^2\cosh^2\xi} + \frac{y^2}{c^2\sinh^2\xi} = 1, \quad \text{i.e. an ellipse.}$$

The eccentricity of the ellipse is given by

$$e = \left(\frac{c^2\cosh^2\xi - c^2\sinh^2\xi}{c^2\cosh^2\xi}\right)^{1/2} = \frac{1}{\cosh\xi}.$$

The foci of the ellipse are at $\pm\, e\times$ the major semi-axis, i.e. $\pm 1/\cosh\xi \times c\cosh\xi = \pm c$. This is independent of ξ and so all the ellipses are confocal.

Similarly, for η constant, the contour is

$$\frac{x^2}{c^2\cos^2\eta} - \frac{y^2}{c^2\sin^2\eta} = 1.$$

This is one of a set of confocal hyperbolae.

(i) $\xi = 0 \quad \Rightarrow \quad y = 0,\ x = c\cos\eta$.
This is the finite line (degenerate ellipse) on the x-axis, $-c \le x \le c$.

(ii) $\eta = 0 \quad \Rightarrow \quad y = 0,\ x = c\cosh\xi$.
This is a part of the x-axis not covered in (i), $c < x < \infty$. The other part, $-\infty < x < -c$, corresponds to $\eta = \pi$.

(iii) This is the same as (the first case) in (ii).

Now, in the ζ-plane, consider the real part of the function $F(\zeta) = -k\zeta$, with k real. On $\xi = 0$ [case (i) above] it reduces to $\text{Re}\,\{-ik\eta\}$, which is zero for all η, i.e. a constant. This implies that the real part of the transformed function will be a constant (actually zero) on $-c \le x \le c$ in the z-plane.

Further,

$$\begin{aligned}(x^2 + y^2)^{1/2} &= (c^2 \cosh^2 \xi \cos^2 \eta + c^2 \sinh^2 \xi \sin^2 \eta)^{1/2}\\ &\approx \tfrac{1}{2} c e^{\xi} \text{ for large } \xi,\\ \Rightarrow \quad \xi &\approx \ln(x^2 + y^2)^{1/2} + \text{fixed constant.}\end{aligned}$$

Hence,

$$\text{Re}\,\{-k\zeta\} = -k\xi \approx -k \ln(x^2 + y^2)^{1/2} \text{ for large } (x^2 + y^2)^{1/2}.$$

Thus, the transformation

$$F(\zeta) = -k\zeta \quad \rightarrow \quad f(z) = -k \cosh^{-1} \frac{z}{c}$$

produces a function in the z-plane that satisfies the stated boundary conditions (as well as satisfying Laplace's equation). It is therefore the required solution.

The electric field near the conducting strip, where $y = 0$ and $z^2 = x^2$, can have no component in the x-direction (except at the points $x = \pm c$), but its magnitude is still given by

$$E = |\,f'(z)\,| = \left| -\frac{k}{\sqrt{z^2 - c^2}} \right| = \frac{k}{(c^2 - x^2)^{1/2}}.$$

25.7 *Use contour integration to answer the following questions about the complex zeros of a polynomial equation.*

(a) *Prove that* $z^8 + 3z^3 + 7z + 5$ *has two zeros in the first quadrant.*
(b) *Find in which quadrants the zeros of* $2z^3 + 7z^2 + 10z + 6$ *lie. Try to locate them.*

(a) Consider the principle of the argument applied to the integral of $f(z) = z^8 + 3z^3 + 7z + 5$ around contour (b) in figure 24.1.

On OA $f(z)$ is always real and $\Delta_{AB} \arg(f) = 0$.

On AB the argument of f increases by $8 \times \frac{1}{2}\pi = 4\pi$.

On BO $z = iy$ and $f(z) = h(y) = y^8 - 3iy^3 + 7iy + 5$. The argument of $h(y)$ is therefore

$$\tan^{-1} \frac{-3y^3 + 7y}{y^8 + 5}.$$

The appropriate choice at $y = \infty$ for this multi-valued function is 4π, as we have just shown. As y decreases from ∞ the argument initially decreases, but passes through 4π again when $y = \sqrt{7/3}$. After that it remains greater than 4π until returning to that value at $y = 0$. Further, since $y^8 + 5$ has no zeros for real y, $\arg(h)$ can reach neither $\frac{7}{2}\pi$ nor $\frac{9}{2}\pi$. Consequently, we deduce that $\Delta_{BO} \arg(f) = 0$.

In summary, $\Delta \arg(f)$ around the closed contour is 4π, and it follows from the principle of the argument that the first quadrant must contain 2 zeros of $f(z)$.

(b) For $f(z) = 2z^3 + 7z^2 + 10z + 6$ we initially follow the same procedure as in part (a), although, as it is a cubic with all of its coefficients positive, we know that it must have at least one negative real zero.

It is straightforward to conclude that $\Delta_{OA} \arg(f) = 0$ and that around the curve AB the change of argument is $\Delta_{AB} \arg(f) = \frac{3}{2}\pi$. On BO,

$$\arg(f) = \tan^{-1} \frac{10y - 2y^3}{6 - 7y^2}.$$

At $y = \infty$ this is $\frac{3}{2}\pi$ (as we have just established) and, as y decreases towards 0, it passes through π at $y = \sqrt{5}$ and $\frac{1}{2}\pi$ at $y = \sqrt{6/7}$, and finally becomes zero at $y = 0$. Thus the net change around the whole closed contour is zero, and we conclude that there are no zeros in the first quadrant. Since the zeros of polynomials with real coefficients occur in complex conjugate pairs, it follows that the fourth quadrant also contains no zeros. This shows that the complex conjugate zeros of $f(z)$ are located in the second and third quadrants.

We start our search for the negative real zero by tabulating some easy-to-calculate values of $f(x)$, the choice of successive values of x being guided by previous results:

z	0	-1	-2	-1.5
$f(z)$	6	1	-2	0

By chance, we have hit upon an exact zero, $z = -\frac{3}{2}$. It follows that $(2z + 3)$ is a factor of $f(z)$, which can be written

$$f(z) = (2z + 3)(z^2 + 2z + 2).$$

The other two zeros are therefore

$$z = -1 \pm \sqrt{1 - 2} = -1 \pm i.$$

As expected, these are in the second and third quadrants.

25.9 *Prove that*

$$\sum_{-\infty}^{\infty} \frac{1}{n^2 + \frac{3}{4}n + \frac{1}{8}} = 4\pi.$$

Carry out the summation numerically, say between -4 *and* 4, *and note how much of the sum comes from values near the poles of the contour integration.*

In order to evaluate this sum, we must first find a function of z that takes the value of the corresponding term in the sum whenever z is an integer. Clearly this is

$$\frac{1}{z^2 + \frac{3}{4}z + \frac{1}{8}}.$$

Further, too make use of the properties of contour integrals, we need to multiply this function by one that has simple poles at the same points, each with unit residue. An appropriate choice of integrand is therefore

$$f(z) = \frac{\pi \cot \pi z}{z^2 + \frac{3}{4}z + \frac{1}{8}} = \frac{\pi \cot \pi z}{(z + \frac{1}{2})(z + \frac{1}{4})}.$$

The contour to be used must enclose all integer values of z, both positive and negative and, in practical terms, must give zero contribution for $|z| \to \infty$, except possibly on the real axis. A large circle C, centred on the origin (see contour (c) in figure 24.1) suggests itself.

As $|zf(z)| \to 0$ on C, the contour integral has value zero. This implies that the residues at the enclosed poles add up to zero. The residues are

$$\frac{\pi \cot(-\frac{1}{2}\pi)}{-\frac{1}{2} + \frac{1}{4}} = 0 \qquad \text{at } z = -\tfrac{1}{2},$$

$$\frac{\pi \cot(-\frac{1}{4}\pi)}{-\frac{1}{4} + \frac{1}{2}} = -4\pi \qquad \text{at } z = -\tfrac{1}{4},$$

$$\sum_{n=-\infty}^{\infty} \frac{1}{(n + \frac{1}{2})(n + \frac{1}{4})} \qquad \text{at } z = n, \; -\infty < n < \infty.$$

The quoted result follows immediately.

For the rough numerical summation we tabulate n, $D(n) = n^2 + \frac{3}{4}n + \frac{1}{8}$ and the

nth term of the series, $1/D(n)$:

n	$D(n)$	$1/D(n)$
−4	13.125	0.076
−3	6.875	0.146
−2	2.625	0.381
−1	0.375	2.667
0	0.125	8.000
1	1.875	0.533
2	5.625	0.178
3	11.375	0.088
4	19.125	0.052

The total of these nine terms is 12.121; this is to be compared with the total for the entire infinite series (of positive terms), which is $4\pi = 12.566$. It will be seen that the sum is dominated by the terms for $n = 0$ and $n = -1$. These two values bracket the positions on the real axis of the poles at $z = -\frac{1}{2}$ and $z = -\frac{1}{4}$.

25.11 *By considering the integral of*

$$\left(\frac{\sin \alpha z}{\alpha z}\right)^2 \frac{\pi}{\sin \pi z}, \qquad \alpha < \frac{\pi}{2},$$

around a circle of large radius, prove that

$$\sum_{m=1}^{\infty} (-1)^{m-1} \frac{\sin^2 m\alpha}{(m\alpha)^2} = \frac{1}{2}.$$

Denote the given function by $f(z)$ and consider its integral around contour (c) in figure 24.1.

As $|z| \to \infty$, $\sin \alpha z \sim e^{\alpha|z|}$, and so $f(z) \sim |z|^{-2} e^{2\alpha|z|} e^{-\pi|z|} = z^{-2} e^{(2\alpha-\pi)|z|}$, and, since $\alpha < \frac{1}{2}\pi$, $|zf(z)\,dz| \to 0$ as $|z| \to \infty$ and the integral around the contour has value zero for $R = \infty$.

The function $f(z)$ has simple poles at $z = n$, where n is an integer, $-\infty < n < \infty$. The pole at $z = 0$ is only a first-order pole as the term in parentheses $\to 1$ as $z \to 0$ and has no singularity there. It follows that the sum of the residues of $f(z)$ at all of its poles is zero. For $n \neq 0$, that residue is

$$\pi \left(\frac{\sin n\alpha}{n\alpha}\right)^2 \left(\left.\frac{d(\sin \pi z)}{dz}\right|_{z=n}\right)^{-1} = \left(\frac{\sin n\alpha}{n\alpha}\right)^2 \frac{1}{\cos \pi n}$$
$$= (-1)^n \left(\frac{\sin n\alpha}{n\alpha}\right)^2.$$

For $n = 0$ the residue is 1.

Since the general residue is an even function of n, the sum for $-\infty < n \leq -1$ is equal to that for $1 \leq n < \infty$, and the zero sum of the residues can be written

$$1 + 2\sum_{n=1}^{\infty}(-1)^n \left(\frac{\sin n\alpha}{n\alpha}\right)^2 = 0,$$

leading immediately to the stated result.

25.13 *Find the function $f(t)$ whose Laplace transform is*

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

Consider first the Taylor series expansion of $\bar{f}(s)$ about $s = 0$:

$$\bar{f}(s) = \frac{e^{-s} - 1 + s}{s^2} = \frac{(1 - s + \frac{1}{2}s^2 + \cdots) - 1 + s}{s^2} \sim \frac{1}{2} + \mathrm{O}(s).$$

Thus $\bar{f}$ has no pole at $s = 0$, and λ in the Bromwich integral can be as small as we wish (but > 0). When the integration line is made part of a closed contour C, the inversion integral becomes

$$f(t) = \int_C \frac{e^{-s}e^{st} - e^{st} + se^{st}}{s^2}\, ds.$$

For $t < 0$, all the terms $\to 0$ as Re $s \to \infty$, and so we close the contour in the right half-plane, as in contour (h) of figure 24.1. On Γ, s times the integrand $\to 0$, and, as the contour encloses no poles, it follows that the integral along L is zero. Thus $f(t) = 0$ for $t < 0$.

For $t > 1$, all terms $\to 0$ as Re $s \to -\infty$, and so we close the contour in the left half-plane, as in contour (g) of figure 24.1. On Γ, s times the integrand again $\to 0$, and, as this contour also encloses no poles, it again follows that the integral along L is zero. Thus $f(t) = 0$ for $t > 1$, as well as for $t < 0$.

For $0 < t < 1$, we need to separate the Bromwich integral into two parts (guided by the different ways in which the parts behave as $|s| \to \infty$):

$$f(t) = \int_L \frac{e^{-s}e^{st}}{s^2}\, ds + \int_L \frac{(s-1)e^{st}}{s^2}\, ds \equiv I_1 + I_2.$$

For I_1 the exponent is $s(t-1)$; $t - 1$ is negative and so, as in the case $t < 0$, we close the contour in the right half-plane [contour (h)]. No poles are included in this contour, and we conclude that $I_1 = 0$.

For I_2 the exponent is st, indicating that (g) is the appropriate contour. However,

$(s-1)/s^2$ does have a pole at $s = 0$ and that is inside the contour. The integral around Γ contributes nothing (that is why it was chosen), and the integral along L must be equal to the residue of $(s-1)e^{st}/s^2$ at $s = 0$. Now,

$$\frac{(s-1)e^{st}}{s^2} = \left(\frac{1}{s} - \frac{1}{s^2}\right)\left(1 + st + \frac{s^2t^2}{2!} + \cdots\right) = -\frac{1}{s^2} + \frac{1}{s}(1-t) + \cdots .$$

The residue, and hence the value of I_2, is therefore $1-t$. Since I_1 has been shown to have value 0, $1-t$ is also the expression for $f(t)$ for $0 < t < 1$.

25.15 *Use contour* (i) *in figure 24.1 to show that the function with Laplace transform* $s^{-1/2}$ *is* $(\pi x)^{-1/2}$.

[*For an integrand of the form* $r^{-1/2}\exp(-rx)$, *change variable to* $t = r^{1/2}$.]

With the suggested contour no poles of $s^{-1/2}e^{sx}$ are enclosed and so the integral of $(2\pi i)^{-1}s^{-1/2}e^{sx}$ around the closed curve must have the value zero. It is also clear that the integral along Γ will be zero since Re $s < 0$ on Γ.

For the small circle γ enclosing the origin, set $s = \rho\, e^{i\theta}$, with $ds = i\rho\, e^{i\theta}\, d\theta$, and consider

$$\lim_{\rho \to 0} \int_0^{2\pi} \rho^{-1/2}e^{-i\theta/2}\exp(x\rho\, e^{i\theta}) i\rho\, e^{i\theta}\, d\theta.$$

This $\to 0$ as $\rho \to 0$ (as $\rho^{1/2}$).

On the upper cut, γ_1, $s = re^{i\pi}$ and the contribution to the integral is

$$\frac{1}{2\pi i}\int_\infty^0 \frac{e^{-i\pi/2}}{r^{1/2}}\exp(rxe^{i\pi})e^{i\pi}\, dr,$$

whilst, on the lower cut, γ_2, $s = re^{-i\pi}$, and its contribution to the integral is

$$\frac{1}{2\pi i}\int_0^\infty \frac{e^{i\pi/2}}{r^{1/2}}\exp(rxe^{-i\pi})e^{-i\pi}\, dr.$$

Combining the two (and making both integrals run over the same range) gives

$$\begin{aligned} -\frac{1}{2\pi i}\int_0^\infty \frac{2i}{r^{1/2}}\, e^{-rx}\, dr &= -\frac{1}{\pi}\int_0^\infty \frac{1}{t}\, e^{-t^2x} 2t\, dt, \text{ after setting } r = t^2, \\ &= -\frac{2}{\pi}\frac{\sqrt{\pi}}{2\sqrt{x}}. \end{aligned}$$

Since this must add to the Bromwich integral along L to make zero, it follows that the function with Laplace transform $s^{-1/2}$ is $(\pi x)^{-1/2}$.

25.17 *The equation*

$$\frac{d^2y}{dz^2} + \left(\nu + \frac{1}{2} - \frac{1}{4}z^2\right) y = 0,$$

sometimes called the Weber–Hermite equation, has solutions known as parabolic cylinder functions. Find, to within (possibly complex) multiplicative constants, the two W.K.B. solutions of this equation that are valid for large $|z|$. In each case, determine the leading term and show that the multiplicative correction factor is of the form $1 + O(\nu^2/z^2)$.

Identify the Stokes and anti-Stokes lines for the equation. On which of the Stokes lines is the W.K.B. solution that tends to zero for z large, real and negative, the dominant solution?

If we consider the equation to be of the generic form

$$\frac{d^2y}{dz^2} + f(z)y = 0,$$

then the W.K.B. solutions are, to within a constant multiplier,

$$y_\pm(z) = \frac{1}{[f(z)]^{1/4}} \exp\left\{\pm i \int^z \sqrt{f(u)}\, du\right\}.$$

In this particular case, writing $\nu + \frac{1}{2}$ as μ for the time being, these solutions are

$$y_\pm(z) = \frac{1}{(\mu - \frac{1}{4}z^2)^{1/4}} \exp\left\{\pm i \int^z \sqrt{\mu - \frac{u^2}{4}}\, du\right\}.$$

Now we seek solutions for large z, and, in this spirit, make binomial expansions of both roots in inverse powers of the relevant variable, z or u. This enables us to write, for a succession of multiplicative complex constants and working to $O(z^{-2})$,

$$\begin{aligned}
y_\pm(z) &= \frac{A}{(\frac{1}{4}z^2 - \mu)^{1/4}} \exp\left\{\pm i^2 \int^z \sqrt{\frac{u^2}{4} - \mu}\, du\right\} \\
&= \frac{B}{\sqrt{z}} \left(1 + \frac{\mu}{z^2} + \cdots\right) \exp\left\{\pm i^2 \int^z \frac{u}{2}\left(1 - \frac{4\mu}{u^2}\right)^{1/2} du\right\} \\
&= \frac{B}{\sqrt{z}} \left(1 + \frac{\mu}{z^2}\right) \exp\left\{\mp \int^z \left(\frac{u}{2} - \frac{\mu}{u} - \frac{\mu^2}{u^3} + \cdots\right) du\right\}.
\end{aligned}$$

Performing the indefinite integral in the exponent yields

$$\begin{aligned} y_\pm(z) &= \frac{B}{\sqrt{z}}\left(1+\frac{\mu}{z^2}\right)\exp\left\{\mp\left(\frac{z^2}{4}-\mu\ln z+\frac{\mu^2}{2z^2}+\cdots\right)\right\} \\ &= \frac{B}{\sqrt{z}}\left(1+\frac{\mu}{z^2}\right)e^{\mp z^2/4}z^{\pm\mu}\left(1\mp\frac{\mu^2}{2z^2}+\cdots\right) \\ &= \frac{B}{\sqrt{z}}\,e^{\mp z^2/4}z^{\pm\mu}\left(1+\frac{2\mu\mp\mu^2}{2z^2}+\cdots\right). \end{aligned}$$

Replacing μ by $\nu+\frac{1}{2}$ and writing the two solutions separately, we have

$$y_1(z) = C\,e^{-z^2/4}\,z^\nu\left[1+\mathrm{O}\left(\frac{\nu^2}{z^2}\right)\right], \qquad y_2(z) = D\,e^{z^2/4}\,z^{-(\nu+1)}\left[1+\mathrm{O}\left(\frac{\nu^2}{z^2}\right)\right].$$

The Stokes lines are determined by the argument(s) of z that make the exponent in the solutions purely real, resulting in one solution being very large (dominant) and one very small (subdominant). As the exponent is proportional to z^2, the Stokes lines are given by $\arg z$ equals 0, $\pi/2$, π or $3\pi/2$. For z large, real and negative, the solution that tends to zero is $y_1(z) \propto e^{-z^2/4}$. This is dominant when z^2 is real and negative, i.e. when z lies on either $\arg z = \pi/2$ or $\arg z = 3\pi/2$.

The anti-Stokes lines, on which the exponent is purely imaginary and consequently the two solutions are comparable in magnitude, are clearly given by the four lines $\arg z = (2n+1)\pi/4$ for $n = 0, 1, 2, 3$.

25.19 *The function $h(z)$ of the complex variable z is defined by the integral*

$$h(z) = \int_{-i\infty}^{i\infty} \exp(t^2 - 2zt)\,dt.$$

(a) *Make a change of integration variable, $t = iu$, and evaluate $h(z)$ using a standard integral. Is your answer valid for all finite z?*

(b) *Evaluate the integral using the method of steepest descents, considering in particular the cases* (i) *z is real and positive,* (ii) *z is real and negative and* (iii) *z is purely imaginary and equal to $i\beta$, where β is real. In each case sketch the corresponding contour in the complex t-plane.*

(c) *Evaluate the integral for the same three cases as specified in part* (b) *using the method of stationary phases. To determine an appropriate contour that passes through a saddle point $t = t_0$, write $t = t_0 + (u + iv)$ and apply the criterion for determining a level line. Sketch the relevant contour in each case, indicating what freedom there is to distort it.*

Comment on the accuracy of the results obtained using the approximate methods adopted in (b) *and* (c).

Before we consider the three different methods of evaluating the integral, we note that its limits lie one in each of the $\pi/2$ sectors of the complex t-plane that are centred on the negative and positive parts of the imaginary axis. All contours that we employ must do the same, though it will not matter exactly where in these sectors they formally end, as, within them, the integrand, which behaves like $\exp(-|t|^2)$, goes (rapidly) to zero as $|t| \to \infty$.

(a) Making the change of integration variable $t = iu$ with $dt = i\,du$ gives $h(z)$ as

$$\begin{aligned} h(z) &= \int_{-\infty}^{\infty} \exp(-u^2 - 2izu)\, i\, du \\ &= \int_{-\infty}^{\infty} \exp[-(u + iz)^2]\, \exp(-z^2)\, i\, du \\ &= i\sqrt{\pi}\, e^{-z^2}. \end{aligned}$$

It is the behaviour of the dominant term in the exponent that determines the convergence or otherwise of the integral. In this case, the t^2 term dominates the term containing z, and, since, as discussed above, it produces convergence, the result is valid for all (finite) values of z.

(b) We first identify the saddle point(s) t_0 of the integrand by setting the derivative of the exponent equal to zero:

$$0 = \frac{d}{dt}(t^2 - 2zt) = 2t - 2z \quad \Rightarrow \quad t_0 = z; \text{ only one saddle point.}$$

The second derivative of the exponent is 2 (independent of the value of z in this case), and so, in the standard notation $f''(t_0) = Ae^{i\alpha}$, we have $A = 2$ and $\alpha = 0$. The value of $f_0 \equiv f(t_0)$ is $t_0^2 - 2zt_0 = -z^2$.

The remaining task is to determine the orientation and direction of traversal of the saddle point. With $t - t_0 = se^{i\theta}$, the possible lines of steepest descents (l.s.d.) are given by $2\theta + \alpha = 0$, $\pm\pi$ or 2π. Of these, the need for $\frac{1}{2}As^2\cos(2\theta + \alpha)$ to be negative picks out $\theta = \pm\frac{1}{2}\pi$. Thus the l.s.d. through the saddle point is parallel to the imaginary axis and the direction of traversal is $+\frac{1}{2}\pi$. Since this lies (just) in the range $-\frac{1}{2}\pi < \theta \leq \frac{1}{2}\pi$, we take the positive sign from the general formula

$$\pm\left(\frac{2\pi}{A}\right)^{1/2} \exp(f_0) \exp[\,\tfrac{1}{2}i(\pi - \alpha)\,]$$

and obtain

$$\begin{aligned} h(z) &= +\left(\frac{2\pi}{2}\right)^{1/2} \exp(-z^2) \exp[\,\tfrac{1}{2}i(\pi - 0)\,] \\ &= i\sqrt{\pi}\, e^{-z^2}. \end{aligned}$$

The conclusion about the orientation and sense of traversal of the saddle point did not depend upon the value of z (because $f''(t_0)$ did not). Consequently the

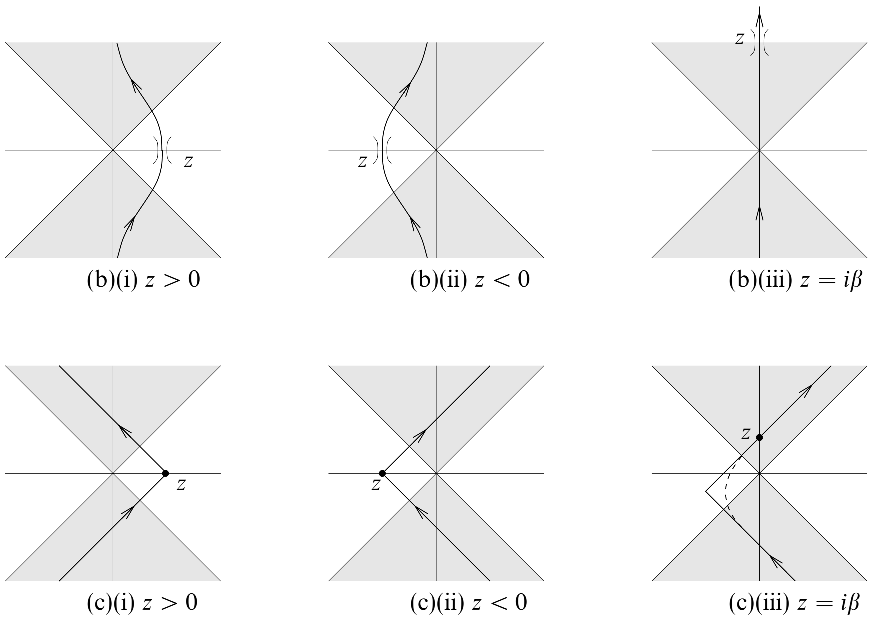

Figure 25.2 The contours following (b) the lines of steepest descents and (c) the lines of stationary phase for the integral in exercise 25.19.

value of the integral is the same for all three cases, though the path in the complex t-plane is determined by z, as is shown in the upper row of sketches in figure 25.2.

(c) We know from general theory that the directions of the level lines at a saddle point make an angle of $\pi/4$ with the l.s.d. through the point. From this and the results of part (b) we can say that the level lines at $t_0 = z$ have directions $\theta = \pm\pi/4$ and $\pm 3\pi/4$.

The same conclusion can be reached, and an indication of suitable contours obtained, by writing $t = t_0 + u + iv$ and requiring that the resulting integrand has a constant magnitude for all u and v. That magnitude must be the same as it is at the saddle point, i.e. when $u = v = 0$.

We consider first cases (i) and (ii) in which $t_0 = z$ is real and $t = (z + u) + iv$. The integrand is then

$$g(u, v) = \exp(t^2 - 2zt) = \exp[\,(z + u)^2 - v^2 + 2iv(z + u) - 2z(z + u + iv)\,],$$

with $g(0, 0) = \exp(-z^2)$. For the integrand to have a constant magnitude, the real part of the exponent must not depend upon u and v. The u- and v-dependent part of the real part of the exponent is $2zu + u^2 - v^2 - 2zu$, and this must therefore have the value 0 for all u and v, i.e. $v = \pm u$. These are the same lines as $\theta = \pm\pi/4$ and $\theta = \pm 3\pi/4$.

Now, although the saddle point at $t_0 = z$ lies outside both of the regions in which the contour must begin and end, the contour must go through it. It is therefore necessary for the contour to turn through a right angle at the saddle point; it transfers from one of the level lines that pass through the saddle to the other one. As will be seen from sketches (c)(i) and (c)(ii), the contour in case (i), $z > 0$, turns to the left by $\pi/2$ as it passes through the saddle; that for $z < 0$ turns to the right through $\pi/2$.

The formula for the total contribution to the integral from integrating through the saddle point along a level line is the same as that for an l.s.d. evaluation, though the former is a Fresnel integral and the latter is an error integral. The stationary phase calculation therefore also yields the value $i\sqrt{\pi}\,\exp(-z^2)$ for $h(z)$.

In both of the present cases, the sharp turn through a right angle at the saddle point means that the vector diagram for the integral consists of one-half from each of two Cornu spirals that are mirror images of each other. Each is broken at its centre point where the phase of the integrand is stationary. The two half spirals join at right angles at the point that is midway between their 'winding points'.

We now turn to case (iii), in which $z = i\beta$ is imaginary. In this case the saddle point lies within one of the two regions that each contain one end of the contour. However, a parallel analysis to that for cases (i) and (ii), setting $t = u + i(\beta + v)$, yields the same conclusion, namely that $v = \pm u$ are appropriate level lines through the saddle.

It is a matter of choice whether the solid line shown in sketch (c)(iii), or its mirror image in the imaginary axis, is chosen; the calculated value for the integral will be the same. The result for $h(z)$ will also be the same as for cases (i) and (ii), i.e. $i\sqrt{\pi}\,\exp(-z^2)$, or, more explicitly in this case, $i\sqrt{\pi}\,\exp(\beta^2)$. Since the contour does not have to go through any particular point other than $t = i\beta$ (and does not need to take a right-angled turn there) and the integrand is analytic, the contour in the end-region not containing z can follow almost any path. One variation from two intersecting straight lines is shown dashed in figure 25.2(c)(iii).

Finally, we note that the fact that all methods give the same answer for $h(z)$, even though the l.s.d. and stationary phase calculations are, in general, approximations, can be put down to the particular form of the integrand. The exponent, $t^2 - 2zt$, is a quadratic function, and so its Taylor series terminates after three terms (of which the second vanishes at the saddle point). Consequently, the l.s.d. and stationary phase approaches which ignore the cubic and higher terms in the Taylor series are *not* approximations. This, together with the fact that there is only one saddle point in the whole t-plane, means that the methods produce exact results for this form of integrand.

25.21 *The stationary phase approximation to an integral of the form*

$$F(\nu) = \int_a^b g(t)e^{i\nu f(t)}\,dt, \qquad |\nu| \gg 1,$$

where $f(t)$ is a real function of t and $g(t)$ is a slowly varying function (when compared with the argument of the exponential), can be written as

$$F(\nu) \sim \left(\frac{2\pi}{|\nu|}\right)^{1/2} \sum_{n=1}^{N} \frac{g(t_n)}{\sqrt{A_n}} \exp\left\{ i \left[\nu f(t_n) + \frac{\pi}{4}\,\mathrm{sgn}\Big(\nu f''(t_n)\Big) \right] \right\},$$

where the t_n are the N stationary points of $f(t)$ that lie in $a < t_1 < t_2 < \cdots < t_N < b$, $A_n = |\,f''(t_n)\,|$, and sgn(x) is the sign of x.

Use this result to find an approximation, valid for large positive values of ν, to the integral

$$F(\nu, z) = \int_{-\infty}^{\infty} \frac{1}{1+t^2}\,\cos[\,(2t^3 - 3zt^2 - 12z^2t)\nu\,]\,dt,$$

where z is a real positive parameter.

Since the argument of the cosine function is everywhere real, we can consider the required integral as the real part of

$$\int_{-\infty}^{\infty} \frac{1}{1+t^2}\,\exp\{\,i\,[\,(2t^3 - 3zt^2 - 12z^2t)\nu\,]\}\,dt,$$

to which we can apply the stated approximation directly. To do so, we need to calculate values for all of the terms appearing in the quoted 'omnibus' formula.

We start by determining the stationary points involved, given by

$$\begin{aligned} 0 = f'(t) = 6t^2 - 6zt - 12z^2 \quad &\Rightarrow \quad (t+z)(t-2z) = 0 \\ &\Rightarrow \quad t_1 = -z \text{ and } t_2 = 2z. \end{aligned}$$

Thus $N = 2$ and the required second derivatives, $f''(t) = 12t - 6z$, and values, $f_n = f(t_n)$, are given by

$$\begin{gathered} f_1 = -2z^3 - 3z^3 + 12z^3 = 7z^3,\ \ f_2 = 16z^3 - 12z^3 - 24z^3 = -20z^3, \\ f''(t_1) = -12z - 6z = -18z,\ \ f''(t_2) = 24z - 6z = 18z. \end{gathered}$$

The two corresponding values of the multiplicative function $g(t) = (1+t^2)^{-1}$ are

$$g(t_1) = (1+z^2)^{-1} \quad \text{and} \quad g(t_2) = (1+4z^2)^{-1}.$$

Substituting all of these gives

$$F(\nu,z) \sim \text{Re}\, \left(\frac{2\pi}{\nu}\right)^{1/2} \left\{ \frac{\exp[\,i(7\nu z^3 - \frac{1}{4}\pi)\,]}{\sqrt{18z}\,(1+z^2)} + \frac{\exp[\,i(-20\nu z^3 + \frac{1}{4}\pi)\,]}{\sqrt{18z}\,(1+4z^2)} \right\}$$

$$= \left(\frac{\pi}{9z\nu}\right)^{1/2} \left[\frac{\cos(7\nu z^3 - \frac{1}{4}\pi)}{1+z^2} + \frac{\cos(20\nu z^3 - \frac{1}{4}\pi)}{1+4z^2} \right]$$

as the stationary phase approximation to the integral.

25.23 *Use the method of steepest descents to find an asymptotic approximation, valid for z large, real and positive, to the function defined by*

$$F_\nu(z) = \int_C \exp(-iz\sin t + i\nu t)\,dt,$$

where ν is real and non-negative and C is a contour that starts at $t = -\pi + i\infty$ and ends at $t = -i\infty$.

Let us denote the integrand by $g(t)$ and the exponent by $f(t)$; thus $g(t) = \exp[\,f(t)\,]$.

We first check that the integrand $\to 0$ at the two end-points; if it did not, the method could not be even approximately correct. As the end-points involve $\pm\infty$, we should formally consider a limiting process, but in practice we need only identify the dominant term in each expression and determine its behaviour as $t \to \infty$.

At $t = -\pi + i\infty$,

$$\sin t = \frac{1}{2i}\,[\,e^{i(-\pi+i\infty)} - e^{-i(-\pi+i\infty)}\,] = \frac{1}{2i}\,(-0 + e^{\infty}) = \frac{1}{2i}\,e^{\infty}.$$

Thus

$$g(-\pi + i\infty) = \exp[\,-iz(e^{\infty}/2i) + i\nu(-\pi + i\infty)\,] = 0,$$

for z real and > 0 and for all ν. Similarly, at $t = -i\infty$,

$$\sin(-i\infty) = \frac{1}{2i}\,(e^{-i^2\infty} - e^{i^2\infty}) = \frac{1}{2i}\,e^{\infty}$$

and

$$g(-i\infty) = \exp[\,-iz(e^{\infty}/2i) + i\nu(-i\infty)\,] = 0.$$

In each case, the behaviour of $f(t)$ is dominated by the exponentiation appearing in the sine term; as this produces a negative exponent for the exponential function determining $g(t)$, the latter $\to 0$ at both end-points.

We next determine the position(s) and properties of the saddle points. These are given by

$$\begin{aligned}
0 = \frac{df}{dt} &= -iz\cos t + iv \\
&\Rightarrow \quad t_0 = \cos^{-1}\frac{v}{z} \text{ (real) with } -\pi < t_0 < 0 \text{ and } z > v, \\
\frac{d^2 f}{dt^2} &= iz\sin t, \\
f''(t_0) &= iz\sin\left(\cos^{-1}\frac{v}{z}\right) = \frac{-iz\sqrt{z^2-v^2}}{z} \\
&= -i\sqrt{z^2-v^2} \\
&\equiv Ae^{i\alpha} \text{ with } A = \sqrt{z^2-v^2} \text{ and } \alpha = \frac{3\pi}{2}, \\
f_0 \equiv f(t_0) &= -iz\sin\left(\cos^{-1}\frac{v}{z}\right) + iv\left(\cos^{-1}\frac{v}{z}\right) \\
&= i\sqrt{z^2-v^2} + iv\left(\cos^{-1}\frac{v}{z}\right).
\end{aligned}$$

Thus the only saddle point is at $t_0 = \cos^{-1}(v/z)$, and the values of $f(t_0)$ and $f''(t_0)$ are given above.

The final step before evaluating the approximate expression for the integral is to determine the direction of the contour through the saddle point. A line of steepest descents (l.s.d.), on which the phase of $f(t)$ is constant, is given by $\sin(2\theta + \alpha) = 0$, where $t - t_0 = se^{i\theta}$ and α is as determined above by $f''(t_0)$. Thus $2\theta + 3\pi/2 = 0,\ \pm\pi,\ 2\pi$ are possible lines, and, of the resulting possible values of $\pm\pi/4$ and $\pm 3\pi/4$ for θ, it is clear that approaching from the direction $\theta = 3\pi/4$ and leaving in the direction $\theta = -\pi/4$ is appropriate. This can be verified by considering the first non-constant, non-vanishing term in the Taylor expansion of $f(t)$, namely

$$\frac{1}{2!}(t-t_0)^2 f''(t_0) = \frac{1}{2!}s^2 e^{-2i\pi/4}(-i\sqrt{z^2-v^2}) = -\frac{s^2\sqrt{z^2-v^2}}{2}.$$

This is real and negative (in both cases, since $e^{-2i\pi/4} = e^{6i\pi/4}$), thus confirming that the standard result for integrating over the saddle point can be used. This is

$$I \approx \pm\left(\frac{2\pi}{A}\right)^{1/2} \exp(f_0)\exp[\tfrac{1}{2}i(\pi - \alpha)],$$

with the $\pm$ choice being resolved by the direction in which the l.s.d. passes through the saddle point; it is positive if $|\theta| < \pi/2$ and negative otherwise. In this particular case, the l.s.d. is traversed in the direction $-\pi/4$ through the saddle point and the plus sign is appropriate.

Finally, inserting all of the specific data for this case into the general formula, we find that

$$\begin{aligned} F_\nu(z) &\approx \left[\frac{2\pi}{\sqrt{z^2-\nu^2}}\right]^{1/2} \exp\left[i\sqrt{z^2-\nu^2} + i\nu\cos^{-1}\frac{\nu}{z}\right] \exp[\tfrac{1}{2}i(\pi - \tfrac{3}{2}\pi)] \\ &= \left[\frac{2\pi}{\sqrt{z^2-\nu^2}}\right]^{1/2} \exp\left[i\left(\sqrt{z^2-\nu^2} + \nu\cos^{-1}\frac{\nu}{z} - \frac{\pi}{4}\right)\right] \\ &\approx \left[\frac{2\pi}{z}\right]^{1/2} \exp\left[i\left(z - \frac{\nu\pi}{2} - \frac{\pi}{4}\right)\right], \text{ for } z \gg \nu. \end{aligned}$$

This last approximation enables us to identify the function $F_\nu(z)$ as probably being a multiple of the Hankel function (Bessel function of the third kind) $H_\nu^{(1)}(z)$, though, as different functions can have the same asymptotic form, this cannot be certain.

26

Tensors

26.1 *Use the basic definition of a Cartesian tensor to show the following.*

(a) *That for any general, but fixed,* ϕ,

$$(u_1, u_2) = (x_1 \cos\phi - x_2 \sin\phi,\ x_1 \sin\phi + x_2 \cos\phi)$$

are the components of a first-order tensor in two dimensions.

(b) *That*

$$\begin{pmatrix} x_2^2 & x_1 x_2 \\ x_1 x_2 & x_1^2 \end{pmatrix}$$

is not a tensor of order 2. *To establish that a single element does not transform correctly is sufficient.*

Consider a rotation of the (unprimed) coordinate axes through an angle θ to give the new (primed) axes. Under this rotation,

$$\begin{aligned} x_1 &\to x_1' = x_1 \cos\theta + x_2 \sin\theta, \\ x_2 &\to x_2' = -x_1 \sin\theta + x_2 \cos\theta, \\ x_3 &\to x_3' = x_3, \end{aligned}$$

and the transformation matrix L_{ij} is given by

$$\mathsf{L} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

(a) Denoting $\cos\theta$ and $\sin\theta$ by c and s, respectively, we compare

$$u_1' = x_1' \cos\phi - x_2' \sin\phi = c x_1 \cos\phi + s x_2 \cos\phi + s x_1 \sin\phi - c x_2 \sin\phi$$

with

$$u'_1 = cu_1 + su_2 = cx_1 \cos\phi - cx_2 \sin\phi + sx_1 \sin\phi + sx_2 \cos\phi.$$

These two are equal, showing that the first component transforms correctly. However, this alone is not sufficient; for (u_1, u_2) to be the components of a first-order tensor, *all* components must transform correctly. We therefore also compare the remaining transformed component:

$$u'_2 = x'_1 \sin\phi + x'_2 \cos\phi = cx_1 \sin\phi + sx_2 \sin\phi - sx_1 \cos\phi + cx_2 \cos\phi$$

is to be compared with

$$u'_2 = -su_1 + cu_2 = -sx_1 \cos\phi + sx_2 \sin\phi + cx_1 \sin\phi + cx_2 \cos\phi.$$

These two are also equal, showing that both components do transform correctly and that (u_1, u_2) are indeed the components of a first-order tensor.

We note, in passing, that $u_1 + iu_2$ is the complex vector obtained by rotating the 'base vector' $x_1 + ix_2$ through an angle ϕ in the complex plane:

$$\begin{aligned} u_1 + iu_2 &= e^{i\phi}(x_1 + ix_2) \\ &= (\cos\phi + i\sin\phi)(x_1 + ix_2) \\ &= (x_1 \cos\phi - x_2 \sin\phi) + i(x_1 \sin\phi + x_2 \cos\phi). \end{aligned}$$

In view of this observation, and of the definition of a first-order tensor as a set of objects 'that transform in the same way as a position vector', it is perhaps not surprising to find that the given expressions form the components of a tensor.

(b) Consider the transform of the first element $u_{11} = x_2^2$. This becomes

$$u'_{11} = (x'_2)^2 = (-sx_1 + cx_2)^2 = s^2x_1^2 - 2scx_1x_2 + c^2x_2^2.$$

If it transforms as a component of a tensor, then it must also be the case that

$$u'_{11} = L_{1k}L_{1l}u_{kl} = c^2x_2^2 + csx_1x_2 + scx_1x_2 + s^2x_1^2.$$

But, these two RHSs are not equal, and it follows that the given set of expressions cannot form the components of a tensor of order 2. It is not necessary to consider any more u_{ij}; failure of any one element to transform correctly rules out the possibility of the set being a tensor.

26.3 *In the usual approach to the study of Cartesian tensors the system is considered fixed and the coordinate axes are rotated. The transformation matrix used is therefore that for components relative to rotated coordinate axes. An alternative view is that of taking the coordinate axes as fixed and rotating the components of the system; this is equivalent to reversing the signs of all rotation angles.*

Using this alternative view, determine the matrices representing (a) *a positive rotation of* $\pi/4$ *about the* x*-axis and* (b) *a rotation of* $-\pi/4$ *about the* y*-axis. Determine the initial vector* $\mathbf{r}$ *which, when subjected to* (a) *followed by* (b), *finishes at* $(3, 2, 1)$.

The normal notation for the two rotation matrices would be

$$\mathsf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \text{ and } \mathsf{B} = \begin{pmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{pmatrix},$$

with $\theta = \phi = \pi/4$.

In the alternative view (denoted by ”) they would have the same forms but with $\theta = \phi = -\pi/4$, namely

$$\mathsf{A}” = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix} \text{ and } \mathsf{B}” = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & -1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

The matrix representing (a) followed by (b) in this alternative view is thus

$$\begin{aligned} \mathsf{B}”\mathsf{A}” &= \frac{1}{2} \begin{pmatrix} 1 & 0 & -1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} \sqrt{2} & -1 & -1 \\ 0 & \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & 1 & 1 \end{pmatrix}. \end{aligned}$$

The required point is the solution of

$$\frac{1}{2} \begin{pmatrix} \sqrt{2} & -1 & -1 \\ 0 & \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}.$$

Using the fact that $\mathsf{B}”\mathsf{A}”$ is orthogonal, and therefore its inverse is simply its transpose, this can be solved directly as

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

As a partial check, we compute $|\mathbf{r}_{\text{initial}}|^2 = 8 + (3 - 2\sqrt{2}) + (3 + 2\sqrt{2}) = 14 = 3^2 + 2^2 + 1^2 = |\mathbf{r}_{\text{final}}|^2$, i.e. the length of the vector is unchanged by the rotations, as it should be.

26.5 *Use the quotient law for tensors to show that the array*

$$\begin{pmatrix} y^2 + z^2 - x^2 & -2xy & -2xz \\ -2yx & x^2 + z^2 - y^2 & -2yz \\ -2zx & -2zy & x^2 + y^2 - z^2 \end{pmatrix}$$

forms a second-order tensor.

To test whether the array is a second-order tensor we need to contract it with an arbitrary known second-order tensor. By 'arbitrary tensor' we mean a tensor in which any one component can be made to be the only non-zero component.

Since any second-order tensor can always be written as the sum of a symmetric and an anti-symmetric tensor, and all operations are linear, it will be sufficient to prove the result for one known symmetric tensor and one known antisymmetric tensor. The simplest symmetric second-order tensor S_{ij} is the (symmetric) outer product of the (by definition) first-order tensor (x, y, z) with itself, i.e $S_{ij} = x_i x_j$. Denoting the given array by B_{ij}, we consider

$$\begin{aligned} B_{ij}S_{ij} &= B_{ij}x_i x_j \\ &= x^2(y^2 + z^2 - x^2) + y^2(x^2 + z^2 - y^2) + z^2(x^2 + y^2 - z^2) \\ &\qquad + 2[\,xy(-2xy) + xz(-2xz) + yz(-2yz)\,] \\ &= -2x^2y^2 - 2x^2z^2 - 2y^2z^2 - x^4 - y^4 - z^4 \\ &= -(x^2 + y^2 + z^2)^2 = -|\mathbf{x}|^4. \end{aligned}$$

The term in parentheses in the last line is formally $x_i x_i$, i.e the contracted product of a first-order tensor with itself, and therefore an invariant zero-order tensor. Squaring an invariant or multiplying it by a constant (-1) leaves it as an invariant, leading to the conclusion that $B_{ij}S_{ij}$ is a zero-order tensor.

We now turn to an antisymmetric tensor, where a suitable second-order tensor A_{ij} is the contraction of the third-order tensor ϵ_{ijk} with the first-order tensor x_i. Thus A_{ij} has the form

$$A = \begin{pmatrix} 0 & z & -y \\ -z & 0 & x \\ y & -x & 0 \end{pmatrix}$$

and the contracted tensor is

$$B_{ij}A_{ij} = 0 - 2xyz + 2yxz + 2zyx + 0 - 2xyz - 2yzx + 2xzy + 0 = 0.$$

Now, 0 is an even more obvious invariant than $|\mathbf{x}|^2$ and so $B_{ij}A_{ij}$ is also a zero-order tensor.

Taking the results of the last two paragraphs together, it follows from the quotient law that B_{ij} is a second-order tensor.

26.7 *Use tensor methods to establish that*

$$grad\,\tfrac{1}{2}(\mathbf{u}\cdot\mathbf{u}) = \mathbf{u}\times curl\,\mathbf{u} + (\mathbf{u}\cdot grad)\mathbf{u}.$$

Now use this result and the general divergence theorem for tensors to show that, for a vector field $\mathbf{A}$,

$$\int_S \left[\mathbf{A}(\mathbf{A}\cdot d\mathbf{S}) - \tfrac{1}{2}A^2 d\mathbf{S}\right] = \int_V \left[\mathbf{A}\,div\mathbf{A} - \mathbf{A}\times curl\,\mathbf{A}\right] dV,$$

where S is the surface enclosing the volume V.

We start with the most complicated of the terms in the identity:

$$\begin{aligned}
[\,\mathbf{u}\times(\nabla\times\mathbf{u})\,]_i &= \epsilon_{ijk}u_j(\nabla\times\mathbf{u})_k = \epsilon_{ijk}u_j\,\epsilon_{klm}\frac{\partial u_m}{\partial x_l}\\
&= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})u_j\frac{\partial u_m}{\partial x_l} = u_j\frac{\partial u_j}{\partial x_i} - u_j\frac{\partial u_i}{\partial x_j}\\
&= \frac{1}{2}\frac{\partial}{\partial x_i}(u_ju_j) - (\mathbf{u}\cdot\nabla)u_i = [\,\tfrac{1}{2}\nabla(\mathbf{u}\cdot\mathbf{u}) - (\mathbf{u}\cdot\nabla)\mathbf{u}\,]_i,
\end{aligned}$$

which establishes the first result. To establish the second result we first note that

$$\frac{\partial}{\partial x_j}(A_iA_j) = A_i\frac{\partial A_j}{\partial x_j} + A_j\frac{\partial A_i}{\partial x_j} = [\,\mathbf{A}\nabla\cdot\mathbf{A} + (\mathbf{A}\cdot\nabla)\mathbf{A}\,]_i. \qquad (*)$$

Next we consider the ith component of the integrand on the RHS of the putative equation and use the first result to replace $\mathbf{A}\times(\nabla\times\mathbf{A})$.

$$\begin{aligned}
[\,\mathrm{RHS}\,]_i &= A_i(\nabla\cdot\mathbf{A}) - [\,\mathbf{A}\times(\nabla\times\mathbf{A})\,]_i\\
&= A_i(\nabla\cdot\mathbf{A}) - \tfrac{1}{2}\nabla_iA^2 + (\mathbf{A}\cdot\nabla)\mathbf{A}_i = \frac{\partial}{\partial x_j}(A_iA_j) - \frac{1}{2}\frac{\partial(A^2)}{\partial x_i}, \qquad \text{using } (*).
\end{aligned}$$

We can now integrate this equation over the volume V and apply the divergence theorem for tensors to both terms individually:

$$\begin{aligned}
\int_V [\,\mathrm{RHS}\,]_i\,dV &= \int_V \frac{\partial}{\partial x_j}(A_iA_j)\,dV - \frac{1}{2}\int_V \frac{\partial(A^2)}{\partial x_i}\,dV\\
&= \int_S A_iA_j\,dS_j - \frac{1}{2}\int_S A^2\,dS_i = \int_S \left[\mathbf{A}(\mathbf{A}\cdot d\mathbf{S}) - \tfrac{1}{2}A^2 d\mathbf{S}\right]_i.
\end{aligned}$$

This concludes the proof.

26.9 *The equation*

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

is a more general form of the expression for the determinant of a 3×3 *matrix* A. *This would normally be written as*

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

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

The following is a list of some of the common properties of determinants.

(a) *Determinant of the transpose.* The transpose matrix A^{T} has the same determinant as A itself, i.e.

$$|\mathsf{A}^{\mathrm{T}}| = |\mathsf{A}|.$$

(b) *Interchanging two rows or two columns.* If two rows (or columns) of A are interchanged, its determinant changes sign but is unaltered in magnitude.

(c) *Identical rows or columns.* If any two rows (or columns) of A are identical or are multiples of one another, then it can be shown that $|\mathsf{A}| = 0$.

(d) *Adding a constant multiple of one row (or column) to another.* The determinant of a matrix is unchanged in value by adding to the elements of one row (or column) any fixed multiple of the elements of another row (or column).

(e) *Determinant of a product.* If A and B are square matrices of the same order then

$$|\mathsf{AB}| = |\mathsf{A}||\mathsf{B}| = |\mathsf{BA}|.$$

A simple extension of this property gives, for example,

$$|\mathsf{AB}\cdots\mathsf{G}| = |\mathsf{A}||\mathsf{B}|\cdots|\mathsf{G}| = |\mathsf{A}||\mathsf{G}|\cdots|\mathsf{B}| = |\mathsf{A}\cdots\mathsf{GB}|,$$

which shows that the determinant is invariant to permutations of the matrices in a multiple product.

Use the form given in $(*)$ *to prove the above properties. For definiteness take* $N = 3$, *but convince yourself that your methods of proof would be valid for any positive integer* $N > 1$.

(a) We write the expression for $|\mathsf{A}^{\mathrm{T}}|$ using the given formalism, recalling that $(\mathsf{A}^{\mathrm{T}})_{ij} = (\mathsf{A})_{ji}$. We then contract both sides with ϵ_{lmn}:

$$\begin{aligned}
|\mathsf{A}^{\mathrm{T}}|\epsilon_{lmn} &= A_{il}A_{jm}A_{kn}\epsilon_{ijk},\\
|\mathsf{A}^{\mathrm{T}}|\epsilon_{lmn}\epsilon_{lmn} &= A_{il}A_{jm}A_{kn}\epsilon_{lmn}\epsilon_{ijk},\\
&= |\mathsf{A}|\epsilon_{ijk}\epsilon_{ijk},\\
|\mathsf{A}^{\mathrm{T}}| &= |\mathsf{A}|.
\end{aligned}$$

In the third line we have used the definition of $|\mathsf{A}|$ (with the roles of the sets of dummy variables $\{i, j, k\}$ and $\{l, m, n\}$ interchanged), and in the fourth line, we have cancelled the scalar quantity $\epsilon_{lmn}\epsilon_{lmn} = \epsilon_{ijk}\epsilon_{ijk}$; the value of this scalar is $N(N-1)$, but that is irrelevant here.

(b) Every non-zero term on the RHS of ($*$) contains any particular row index once and only once. The same can be said for the Levi–Civita symbol on the LHS. Thus interchanging two rows is equivalent to interchanging two of the subscripts of ϵ_{lmn} and thereby reversing its sign. Consequently, the whole RHS changes sign and the magnitude of $|\mathsf{A}|$ remains the same, though its sign is changed.

(c) If, say, $A_{pi} = \lambda A_{pj}$, for some particular pair of values i and j and all p, then in the (multiple) summation on the RHS of ($*$) each A_{nk} appears multiplied by (with no summation over i and j)

$$\epsilon_{ijk}A_{li}A_{mj} + \epsilon_{jik}A_{lj}A_{mi} = \epsilon_{ijk}\lambda A_{lj}A_{mj} + \epsilon_{jik}A_{lj}\lambda A_{mj} = 0,$$

since $\epsilon_{ijk} = -\epsilon_{jik}$. Consequently, grouped in this way, all pairs of terms contribute nothing to the sum and $|\mathsf{A}| = 0$.

(d) Consider the matrix B whose m, jth element is defined by $B_{mj} = A_{mj} + \lambda A_{pj}$, where $p \neq m$. The only case that needs detailed analysis is when l, m and n are all different. Since $p \neq m$ it must be the same as either l or n; suppose that $p = l$. The determinant of B is given by

$$\begin{aligned} |\mathsf{B}|\epsilon_{lmn} &= A_{li}(A_{mj} + \lambda A_{lj})A_{nk}\epsilon_{ijk} \\ &= A_{li}A_{mj}A_{nk}\epsilon_{ijk} + \lambda A_{li}A_{lj}A_{nk}\epsilon_{ijk} \\ &= |\mathsf{A}|\epsilon_{lmn} + \lambda 0, \end{aligned}$$

where we have used the row equivalent of the intermediate result obtained for columns in (c). Thus we conclude that $|\mathsf{B}| = |\mathsf{A}|$.

(e) If $\mathsf{X} = \mathsf{AB}$, then

$$|\,\mathsf{X}\,|\epsilon_{lmn} = A_{lx}B_{xi}A_{my}B_{yj}A_{nz}B_{zk}\epsilon_{ijk}.$$

Contract both sides with ϵ_{lmn}:

$$\begin{aligned} |\,\mathsf{X}\,|\epsilon_{lmn}\epsilon_{lmn} &= \epsilon_{lmn}A_{lx}A_{my}A_{nz}\,\epsilon_{ijk}B_{xi}B_{yj}B_{zk} \\ &= \epsilon_{xyz}|\mathsf{A}^{\mathrm{T}}|\epsilon_{xyz}|\mathsf{B}|, \\ \Rightarrow \quad |\,\mathsf{X}\,| &= |\mathsf{A}^{\mathrm{T}}|\,|\mathsf{B}| = |\mathsf{A}|\,|\mathsf{B}|, \quad \text{using result (a).} \end{aligned}$$

To obtain the last line we have cancelled the non-zero scalar $\epsilon_{lmn}\epsilon_{lmn} = \epsilon_{xyz}\epsilon_{xyz}$ from both sides, as we did in the proof of result (a).

The extension to the product of any number of matrices is obvious. Replacing B by CD or by DC and applying the result just proved extends it to a product of three matrices. Extension to any higher number is done in the same way.

26.11 *Given a non-zero vector* $\mathbf{v}$, *find the value that should be assigned to* α *to make*

$$P_{ij} = \alpha v_i v_j \qquad \text{and} \qquad Q_{ij} = \delta_{ij} - \alpha v_i v_j$$

into parallel and orthogonal projection tensors, respectively, i.e. tensors that satisfy, respectively, $P_{ij}v_j = v_i$, $P_{ij}u_j = 0$ *and* $Q_{ij}v_j = 0$, $Q_{ij}u_j = u_i$, *for any vector* $\mathbf{u}$ *that is orthogonal to* $\mathbf{v}$.

Show, in particular, that Q_{ij} *is unique, i.e. that if another tensor* T_{ij} *has the same properties as* Q_{ij} *then* $(Q_{ij} - T_{ij})w_j = 0$ *for* any *vector* $\mathbf{w}$.

Consider

$$P_{ij}v_j = \alpha v_i v_j v_j = \alpha|\mathbf{v}|^2 v_i, \text{ and}$$
$$P_{ij}u_j = \alpha v_i v_j u_j = \alpha v_i(v_j u_j) = 0, \text{ as } \mathbf{u} \text{ is orthogonal to } \mathbf{v}.$$

For $P_{ij}v_j = v_i$ it is clearly necessary that $\alpha = |\mathbf{v}|^{-2}$.

With this choice,

$$Q_{ij}v_j = (\delta_{ij} - \alpha v_i v_j)v_j = v_i - \alpha(v_j v_j)v_i = v_i - |\mathbf{v}|^{-2}(v_j v_j)v_i = 0, \text{ and}$$
$$Q_{ij}u_j = (\delta_{ij} - \alpha v_i v_j)u_j = u_i - \alpha(v_j u_j)v_i = u_i - 0v_i = u_i.$$

Thus the one assigned value for α gives both P_{ij} and Q_{ij} the required properties.

Let $\mathbf{u}^{(1)}$ and $\mathbf{u}^{(2)}$ be any two linearly independent non-zero vectors orthogonal to $\mathbf{v}$. Then *any* vector $\mathbf{w}$ can be expressed as $\lambda\mathbf{v} + \mu\mathbf{u}^{(1)} + \nu\mathbf{u}^{(2)}$.
Now suppose that T_{ij} has the same properties as Q_{ij} and consider

$$\begin{aligned}(Q_{ij} - T_{ij})w_j &= (Q_{ij} - T_{ij})(\lambda v_j + \mu u_j^{(1)} + \nu u_j^{(2)})\\ &= \lambda 0 + \mu u_j^{(1)} + \nu u_j^{(2)} - \lambda T_{ij}v_j - \mu T_{ij}u_j^{(1)} - \nu T_{ij}u_j^{(2)}\\ &= 0 + \mu u_j^{(1)} + \nu u_j^{(2)} - 0 - \mu u_j^{(1)} - \nu u_j^{(2)} = 0.\end{aligned}$$

In this sense, Q_{ij} is unique.

26.13 *In a certain crystal the unit cell can be taken as six identical atoms lying at the corners of a regular octahedron. Convince yourself that these atoms can also be considered as lying at the centres of the faces of a cube and hence that the crystal has cubic symmetry. Use this result to* prove *that the conductivity tensor for the crystal,* σ_{ij}, *must be isotropic.*

It is easiest to start with a cube and then join the centre points of any pair of faces that have a common edge. The network of 12 lines so formed are the edges of a regular octahedron.

The crystal has cubic symmetry and must therefore be invariant under rotations that leave a cube unchanged (apart from the labelling of its corners). One such symmetry operation is rotation (by $2\pi/3$) about a body diagonal; this relabels the axes $O123$ as axes $O3'1'2'$ in the rotated system. The (orthogonal) base-vector transformation matrix S has as its i, jth component the ith component of $\mathbf{e}'_j$ with respect to the basis $\{\mathbf{e}_k\}$. The coordinate transformation matrix L is the transpose of this. For the rotation under consideration,

$$\mathsf{S} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad \mathsf{L} = \mathsf{S}^{\mathrm{T}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

The conductivity tensor is a second-order tensor and so $\sigma'_{ij} = L_{ik}L_{jm}\sigma_{km}$ or, in matrix form,

$$\begin{aligned}
\sigma' &= \mathsf{L}\sigma\mathsf{L}^{\mathrm{T}} \\
&= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \\
&= \begin{pmatrix} \sigma_{22} & \sigma_{23} & \sigma_{21} \\ \sigma_{32} & \sigma_{33} & \sigma_{31} \\ \sigma_{12} & \sigma_{13} & \sigma_{11} \end{pmatrix}.
\end{aligned}$$

This must be the same tensor as σ and so requires that

$$\sigma_{11} = \sigma_{22} = \sigma_{33}; \quad \sigma_{12} = \sigma_{23} = \sigma_{31}; \quad \sigma_{21} = \sigma_{32} = \sigma_{13}.$$

We also note that the transformed tensor is the original one, but with $1 \to 2$, $2 \to 3$ and $3 \to 1$.

Now, restarting from the original situation, consider a rotation of $\pi/2$ about the 3-axis. This clearly carries the $O1$-axis onto the original $O2$-axis and the $O2$-axis onto the original negative $O1$-axis. Therefore, by the substitutions $1 \to 2$, $2 \to -1$ and $3 \to 3$ (where a component changes sign for each minus sign on its subscripts) or by a matrix calculation similar to the previous one, the new transformed conductivity tensor is

$$\sigma' = \begin{pmatrix} \sigma_{22} & -\sigma_{21} & \sigma_{23} \\ -\sigma_{12} & \sigma_{11} & -\sigma_{13} \\ \sigma_{32} & -\sigma_{31} & \sigma_{33} \end{pmatrix}.$$

Again the invariance of σ imposes requirements. In this case,

$$\sigma_{11} = \sigma_{22}; \quad \sigma_{13} = \sigma_{23} = -\sigma_{13}.$$

The last set of equalities requires that $\sigma_{13} = \sigma_{23} = 0$ and hence, by the previous result, that $\sigma_{ij} = 0$ whenever $i \neq j$. Since $\sigma_{11} = \sigma_{22} = \sigma_{33}$, σ is a multiple of the unit matrix, and it follows that σ_{ij} is an isotropic tensor.

Either by direct calculation or by noting that any rotational symmetry of a cube can be represented as an ordered sequence of the two rotations already used, it can be shown that other symmetries do not impose any further constraint on the remaining non-zero elements of the conductivity tensor. Intuitively this must be so, since σ now contains only one free parameter, the common value of σ_{11}, σ_{22} and σ_{33}, and this is required to describe the level of conductivity, which must vary from one crystal to another, and certainly between crystals of different elements.

26.15 *In a certain system of units the electromagnetic stress tensor M_{ij} is given by*

$$M_{ij} = E_iE_j + B_iB_j - \tfrac{1}{2}\delta_{ij}(E_kE_k + B_kB_k),$$

where the electric and magnetic fields, $\mathbf{E}$ *and* $\mathbf{B}$, *are first-order tensors. Show that M_{ij} is a second-order tensor.*

Consider a situation in which $|\mathbf{E}| = |\mathbf{B}|$ but the directions of $\mathbf{E}$ *and* $\mathbf{B}$ *are not parallel. Show that $\mathbf{E} \pm \mathbf{B}$ are principal axes of the stress tensor and find the corresponding principal values. Determine the third principal axis and its corresponding principal value.*

In the calculation of the transformed RHS, $\mathbf{E}$ and $\mathbf{B}$ transform with a single 'L-matrix', but δ_{ij}, being a second-order tensor, requires two. It may simply be noticed that E_kE_k and B_kB_k are scalars and therefore unaltered in the transformation; but, if not, then the orthogonal properties of L, $L_{ik}L_{jk} = \delta_{ij}$ and $L_{ki}L_{kj} = \delta_{ij}$, are needed:

$$\begin{aligned}
M_{ij} &= E_iE_j + B_iB_j - \tfrac{1}{2}\delta_{ij}(E_kE_k + B_kB_k),\\
M'_{ij} &= L_{im}E_mL_{jn}E_n + L_{im}B_mL_{jn}B_n\\
&\qquad - \tfrac{1}{2}L_{ip}L_{jq}\delta_{pq}(L_{kr}E_rL_{ks}E_s + L_{kr}B_rL_{ks}B_s)\\
&= L_{im}L_{jn}(E_mE_n + B_mB_n) - \tfrac{1}{2}L_{ip}L_{jq}\delta_{pq}(\delta_{rs}E_rE_s + \delta_{rs}B_rB_s)\\
&= L_{im}L_{jn}[\,E_mE_n + B_mB_n - \tfrac{1}{2}\delta_{mn}(E_rE_r + B_rB_r)\,]\\
&= L_{im}L_{jn}M_{mn}.
\end{aligned}$$

To obtain the penultimate line we relabelled the dummy suffices p and q as m and n. Thus M_{ij} transforms as a second-order tensor; it is real and symmetric and will therefore have orthogonal eigenvectors.

For the case $|\mathbf{E}| = |\mathbf{B}|$, i.e. $E^2 = B^2$, denote $E_i \pm B_i$ by v_i and consider

$$\begin{aligned}
M_{ij}v_j &= M_{ij}(E_j \pm B_j) \\
&= E_iE_j(E_j \pm B_j) + B_iB_j(E_j \pm B_j) - \tfrac{1}{2}\delta_{ij}(E^2 + B^2)(E_j \pm B_j) \\
&= E_iE^2 \pm E_i(\mathbf{E}\cdot\mathbf{B}) + B_i(\mathbf{B}\cdot\mathbf{E}) \pm B_iB^2 \\
&\qquad - \tfrac{1}{2}(E^2 + B^2)(E_i \pm B_i) \\
&= (E_i \pm B_i)[\,E^2 \pm (\mathbf{E}\cdot\mathbf{B}) - \tfrac{1}{2}\,2E^2\,], \text{ using } E^2 = B^2, \\
&= \pm(\mathbf{E}\cdot\mathbf{B})(E_i \pm B_i) \\
&= \pm(\mathbf{E}\cdot\mathbf{B})v_i.
\end{aligned}$$

This shows that $\mathbf{E}\pm\mathbf{B}$ are eigenvectors of M_{ij} (i.e. its principal axes) with principal values $\pm(\mathbf{E}\cdot\mathbf{B})$.

The third principal axis is orthogonal to both of these and is therefore in the direction

$$(\mathbf{E}+\mathbf{B})\times(\mathbf{E}-\mathbf{B}) = \mathbf{0} + (\mathbf{B}\times\mathbf{E}) - (\mathbf{E}\times\mathbf{B}) - \mathbf{0} = 2(\mathbf{B}\times\mathbf{E}).$$

To determine its principal value, consider

$$\begin{aligned}
M_{ij}(\mathbf{B}\times\mathbf{E})_j &= M_{ij}\epsilon_{jlm}B_lE_m \\
&= E_iE_j\epsilon_{jlm}B_lE_m + B_iB_j\epsilon_{jlm}B_lE_m - \tfrac{1}{2}\delta_{ij}\,2E^2\epsilon_{jlm}B_lE_m \\
&= 0 + 0 - E^2(\mathbf{B}\times\mathbf{E})_i, \text{ since } \epsilon_{jlm}X_lX_j = 0.
\end{aligned}$$

Thus, the third principal value is $-E^2$ (or $-B^2$). This value could have been deduced from the trace of $M_{ij} = E^2 + B^2 - \frac{3}{2}(E^2 + B^2) = -E^2$, since the two eigenvalues found previously are $\pm\mathbf{E}\cdot\mathbf{B}$, which sum to zero. The three eigenvalues together must add up to the trace; hence, the third one is $-E^2$.

26.17 *A rigid body consists of eight particles, each of mass m, held together by light rods. In a certain coordinate frame the particles are at positions*

$$\pm a(3, 1, -1), \qquad \pm a(1, -1, 3), \qquad \pm a(1, 3, -1), \qquad \pm a(-1, 1, 3).$$

Show that, when the body rotates about an axis through the origin, if the angular velocity and angular momentum vectors are parallel then their ratio must be $40ma^2$*,* $64ma^2$ *or* $72ma^2$*.*

Because the particles are symmetrically placed in pairs with respect to the origin, the inertia tensor, given by

$$I_{ij} = \sum_{\text{particles}} m(r^2\delta_{ij} - x_ix_j),$$

will be twice that calulated for the + signs alone. It components are therefore

$$\begin{aligned}
I_{11} &= 2ma^2(2+10+10+10) = 64ma^2,\\
I_{12} = I_{21} &= -2ma^2(3-1+3-1) = -8ma^2,\\
I_{13} = I_{31} &= -2ma^2(-3+3-1-3) = 8ma^2,\\
I_{22} &= 2ma^2(10+10+2+10) = 64ma^2,\\
I_{23} = I_{32} &= -2ma^2(-1-3-3+3) = 8ma^2,\\
I_{33} &= 2ma^2(10+2+10+2) = 48ma^2.
\end{aligned}$$

The resulting tensor is

$$8ma^2 \begin{pmatrix} 8 & -1 & 1 \\ -1 & 8 & 1 \\ 1 & 1 & 6 \end{pmatrix}$$

and its principal moments are $8ma^2\lambda$, where

$$\begin{aligned}
0 &= \begin{vmatrix} 8-\lambda & -1 & 1 \\ -1 & 8-\lambda & 1 \\ 1 & 1 & 6-\lambda \end{vmatrix}\\
&= (8-\lambda)(\lambda^2 - 14\lambda + 47) + (-7+\lambda) + (-9+\lambda)\\
&= (8-\lambda)(\lambda^2 - 14\lambda + 47 - 2)\\
&= (8-\lambda)(\lambda-9)(\lambda-5).
\end{aligned}$$

Thus the principal moments are $40ma^2$, $64ma^2$ and $72ma^2$. As a partial check: $40 + 64 + 72 = 8(8 + 8 + 6)$.

If the angular velocity ω and the angular momentum $\mathbf{J} = I\omega$ are parallel, then the body is rotating about one of its principal axes (the eigenvectors of I); their ratio is the principal moment about that axis and is thus one of the three values calculated above.

26.19 *A block of wood contains a number of thin soft-iron nails (of constant permeability). A unit magnetic field directed eastwards induces a magnetic moment in the block having components* $(3, 1, -2)$*, and similar fields directed northwards and vertically upwards induce moments* $(1, 3, -2)$ *and* $(-2, -2, 2)$ *respectively. Show that all the nails lie in parallel planes.*

The magnetic moment M, the permeability μ and the magnetic field H for iron of constant pemeability are connected by $M_i = \mu_{ij} H_j$. Taking the 1-, 2- and

3-directions as East, North and vertical, μ has the form

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

By adding the first two columns to twice the third one, it can be seen that this matrix has zero determinant. The matrix therefore has at least one zero eigenvalue. (The same conclusion can be reached using the routine method for finding eigenvalues; they are 0, 2 and 6.)

Thus, a field parallel to the eigenvector corresponding to this zero eigenvalue will induce no moment in the block. Physically this means that all the nails lie in planes to which this direction is a normal. To find the direction we solve

$$\begin{pmatrix} 3 & 1 & -2 \\ 1 & 3 & -2 \\ -2 & -2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}.$$

We conclude that all the nails lie at right angles to this direction.

26.21 *For a general isotropic medium, the stress tensor p_{ij} and strain tensors e_{ij} are related by*

$$p_{ij} = \frac{\sigma E}{(1+\sigma)(1-2\sigma)}\, e_{kk}\delta_{ij} + \frac{E}{1+\sigma}\, e_{ij},$$

where E is Young's modulus and σ is Poisson's ratio.

By considering an isotropic body subjected to a uniform hydrostatic pressure (no shearing stress), show that the bulk modulus k, defined by the ratio of the pressure to the fractional decrease in volume, is given by $k = E/[3(1-2\sigma)]$.

Consider a small rectangular parallelepiped, with one corner at the origin and the opposite one at (a_1, a_2, a_3), subjected to a uniform hydrostatic pressure. The isotropy of the pressure means that all forces are normal to the surfaces on which they act and that the stress and strain tensor components $p_{ij} = e_{ij} = 0$ for $i \neq j$. Furthermore, because of the symmetry of the situation, when $i \neq j$, not only is e_{ij} zero, but so are the individual $\partial u_i/\partial x_j$ that are its constituents.

In the current situation, $p_{11} = p_{22} = p_{33} = -p$ and so, writing $\sum_k e_{kk}$ as θ, we have, for each i ($i = 1, 2, 3$) with no summation over i implied, that

$$-p = p_{ii} = \frac{E}{(1+\sigma)(1-2\sigma)}\,[\sigma\theta + (1-2\sigma)e_{ii}].$$

Adding the three equations together gives

$$-3p = \frac{E}{(1+\sigma)(1-2\sigma)}\,[\,3\sigma\theta + (1-2\sigma)\theta\,] = \frac{E\theta}{1-2\sigma}.$$

Now the fractional increase f in the volume of the parallelepiped is given by

$$\frac{1}{a_1a_2a_3}\left(a_1 + \frac{\partial u_1}{\partial x_i}a_i + \cdots\right)\left(a_2 + \frac{\partial u_2}{\partial x_i}a_i + \cdots\right)\left(a_3 + \frac{\partial u_3}{\partial x_i}a_i + \cdots\right) - 1.$$

Since $\partial u_i/\partial x_j = 0$ for $i \neq j$, the only three non-zero first-order terms are

$$f = \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3} = e_{11} + e_{22} + e_{33} = \theta.$$

We conclude that the bulk modulus, k, is given by

$$k = \frac{p}{-f} = \frac{-E\theta}{3(1-2\sigma)}\,\frac{1}{(-\theta)} = \frac{E}{3(1-2\sigma)}.$$

26.23 *A fourth-order tensor T_{ijkl} has the properties*

$$T_{jikl} = -T_{ijkl}, \qquad T_{ijlk} = -T_{ijkl}.$$

Prove that for any such tensor there exists a second-order tensor K_{mn} such that

$$T_{ijkl} = \epsilon_{ijm}\epsilon_{kln}K_{mn}$$

and give an explicit expression for K_{mn}. Consider two (separate) special cases, as follows.

(a) *Given that T_{ijkl} is isotropic and $T_{ijji} = 1$, show that T_{ijkl} is uniquely determined and express it in terms of Kronecker deltas.*

(b) *If now T_{ijkl} has the additional property*

$$T_{klij} = -T_{ijkl},$$

show that T_{ijkl} has only three linearly independent components and find an expression for T_{ijkl} in terms of the vector

$$V_i = -\tfrac{1}{4}\epsilon_{jkl}T_{ijkl}.$$

As K_{mn} is to be a second-order tensor, we need to construct such a tensor from T_{ijkl}. Since the latter is of fourth order, it needs to be contracted n times with a tensor of order $2n-2$ for some positive integer n. In view of the final stated expression for T_{ijkl}, involving $\epsilon_{ijm}\epsilon_{kln}$, i.e. a sixth-order tensor, we try $n = 4$ and, starting from $T_{ijkl} = \epsilon_{ijm}\epsilon_{kln}K_{mn}$, consider

$$\begin{aligned}\epsilon_{pij}\epsilon_{qkl}T_{ijkl} &= \epsilon_{pij}\epsilon_{qkl}\epsilon_{ijm}\epsilon_{kln}K_{mn}\\ &= (\delta_{jj}\delta_{pm} - \delta_{jm}\delta_{pj})(\delta_{ll}\delta_{qn} - \delta_{ln}\delta_{ql})K_{mn}\\ &= (3\delta_{pm} - \delta_{pm})(3\delta_{qn} - \delta_{qn})K_{mn}\\ &= 4K_{pq}.\end{aligned}$$

Clearly, $K_{mn} = \frac{1}{4}\epsilon_{mij}\epsilon_{nkl}T_{ijkl}$ has the required property.

(a) Given that T_{ijkl} is isotropic, and noting that ϵ_{mij} and ϵ_{nkl} are also isotropic, we conclude that K_{mn} must itself be isotropic. It must therefore be some multiple of δ_{mn} (as this is the most general isotropic second-order tensor), i.e. $K_{mn} = \lambda\delta_{mn}$ for one or more values of λ. Thus,

$$\begin{aligned}T_{ijkl} &= \epsilon_{ijm}\,\epsilon_{kln}\lambda\delta_{mn}\\ &= \lambda\epsilon_{ijm}\,\epsilon_{klm}\\ &= \lambda(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}).\end{aligned}$$

Now, since $T_{ijji} = 1$,

$$\begin{aligned}1 &= \lambda(\delta_{ij}\delta_{ji} - \delta_{ii}\delta_{jj})\\ &= \lambda[\,\delta_{ii} - (\delta_{ii})^2\,] = \lambda(3-9) \quad \Rightarrow \quad \lambda = -\tfrac{1}{6}.\end{aligned}$$

We conclude that λ, and therefore also T_{ijkl}, is unique with $T_{ijkl} = \frac{1}{6}(\delta_{il}\delta_{jk} - \delta_{ik}\delta_{jl})$.

(b) To examine the implications of the antisymmetry indicated by $T_{klij} = -T_{ijkl}$, we interchange the pair of dummy suffices $\{i, j\}$ with the pair $\{k, l\}$ to obtain the third line below — and then switch them back again in the fourth line using the antisymmetry:

$$\begin{aligned}K_{mn} &= \tfrac{1}{4}\epsilon_{mij}\,\epsilon_{nkl}\,T_{ijkl},\\ K_{nm} &= \tfrac{1}{4}\epsilon_{nij}\,\epsilon_{mkl}\,T_{ijkl}\\ &= \tfrac{1}{4}\epsilon_{nkl}\,\epsilon_{mij}\,T_{klij}\\ &= -\tfrac{1}{4}\epsilon_{nkl}\,\epsilon_{mij}\,T_{ijkl}\\ &= -K_{mn}.\end{aligned}$$

Thus K_{mn} is antisymmetric. It therefore has zeros on its leading diagonal and only three linearly independent components as non-diagonal elements. Since T_{ijkl} is uniquely defined in terms of K_{mn}, it too has only three linearly independent components.

Now consider

$$\begin{aligned}
\epsilon_{jkl} T_{ijkl} &= \epsilon_{jkl}\, \epsilon_{ijm}\, \epsilon_{kln} K_{mn}, \\
-4V_i &= (\delta_{km}\delta_{li} - \delta_{ki}\delta_{lm})\epsilon_{kln} K_{mn} \\
&= (\epsilon_{min} - \epsilon_{imn}) K_{mn} \\
&= 2\epsilon_{min} K_{mn}.
\end{aligned}$$

To 'invert' this relationship, consider

$$\begin{aligned}
\epsilon_{irs} V_i &= -\tfrac{1}{2}\epsilon_{irs}\, \epsilon_{min} K_{mn} \\
&= -\tfrac{1}{2}(\delta_{rn}\delta_{sm} - \delta_{rm}\delta_{sn}) K_{mn} \\
&= -\tfrac{1}{2}(K_{sr} - K_{rs}) \\
&= K_{rs} \quad \Rightarrow \quad K_{mn} = \epsilon_{pmn} V_p.
\end{aligned}$$

Finally, expressing T_{ijkl}, as given in the question, explicitly in terms of the vector V_i, using the result obtained above, we have

$$\begin{aligned}
T_{ijkl} &= \epsilon_{ijm}\, \epsilon_{kln}\, \epsilon_{pmn} V_p \\
&= (\delta_{in}\delta_{jp} - \delta_{ip}\delta_{jn})\epsilon_{kln} V_p \\
&= \epsilon_{kli} V_j - \epsilon_{klj} V_i.
\end{aligned}$$

26.25 *In a general coordinate system u^i, $i = 1, 2, 3$, in three-dimensional Euclidean space, a volume element is given by*

$$dV = |\mathbf{e}_1\, du^1 \cdot (\mathbf{e}_2\, du^2 \times \mathbf{e}_3\, du^3)|.$$

Show that an alternative form for this expression, written in terms of the determinant g of the metric tensor, is given by

$$dV = \sqrt{g}\, du^1\, du^2\, du^3.$$

Show that under a general coordinate transformation to a new coordinate system u'^i, the volume element dV remains unchanged, i.e. show that it is a scalar quantity.

Working in terms of the Cartesian bases vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$, let

$$\mathbf{e}_m = \lambda_{mx}\, \mathbf{i} + \lambda_{my}\, \mathbf{j} + \lambda_{mz}\, \mathbf{k}, \text{ for } m = 1, 2, 3.$$

Then,

$$\begin{aligned}
\mathbf{e}_2\, du^2 \times\ \mathbf{e}_3\, du^3 = du^2\, du^3(&\lambda_{2x}\lambda_{3y}\, \mathbf{k} - \lambda_{2x}\lambda_{3z}\, \mathbf{j} - \lambda_{2y}\lambda_{3x}\, \mathbf{k} \\
&+ \lambda_{2y}\lambda_{3z}\, \mathbf{i} + \lambda_{2z}\lambda_{3x}\, \mathbf{j} - \lambda_{2z}\lambda_{3y}\, \mathbf{i}),
\end{aligned}$$

and it follows that

$$\begin{aligned}
dV &= \mathbf{e}_1\, du^1 \cdot (\mathbf{e}_2\, du^2 \times \mathbf{e}_3\, du^3) \\
&= du^1\, du^2\, du^3 [\, \lambda_{1x}(\lambda_{2y}\lambda_{3z} - \lambda_{2z}\lambda_{3y}) + \lambda_{1y}(\lambda_{2z}\lambda_{3x} - \lambda_{2x}\lambda_{3z}) \\
&\qquad\qquad\qquad\qquad + \lambda_{1z}(\lambda_{2x}\lambda_{3y} - \lambda_{2y}\lambda_{3x})\,] \\
&= du^1\, du^2\, du^3 \begin{vmatrix} \lambda_{1x} & \lambda_{1y} & \lambda_{1z} \\ \lambda_{2x} & \lambda_{2y} & \lambda_{2z} \\ \lambda_{3x} & \lambda_{3y} & \lambda_{3z} \end{vmatrix} \\
&\equiv du^1\, du^2\, du^3 |\mathsf{A}|, \text{ thus defining } \mathsf{A}.
\end{aligned}$$

Now consider an element of the matrix $\mathsf{A}\mathsf{A}^{\mathrm{T}}$:

$$(\mathsf{A}\mathsf{A}^{\mathrm{T}})_{mn} = \sum_r \mathsf{A}_{mr}\mathsf{A}_{nr} = \lambda_{mx}\lambda_{nx} + \lambda_{my}\lambda_{ny} + \lambda_{mz}\lambda_{nz}.$$

But the elements of the metric tensor are given by

$$\mathsf{g}_{mn} = \mathbf{e}_m \cdot \mathbf{e}_n = \lambda_{mx}\lambda_{nx} + \lambda_{my}\lambda_{ny} + \lambda_{mz}\lambda_{nz}.$$

Hence $\mathsf{A}\mathsf{A}^{\mathrm{T}} = \mathsf{g}$ and, in particular, $|\mathsf{A}|\,|\mathsf{A}^{\mathrm{T}}| = |\,\mathsf{g}\,|$. Since $|\mathsf{A}| = |\mathsf{A}^{\mathrm{T}}|$, it follows that $|\mathsf{A}| = |\,\mathsf{g}\,|^{1/2} = \sqrt{g}$ and

$$dV = du^1\, du^2\, du^3\, |\mathsf{A}| = \sqrt{g}\, du^1\, du^2\, du^3.$$

For a transformation $u'^i = u'^i(u^1, u^2, u^3)$,

$$du'^1\, du'^2\, du'^3 = \left| \frac{\partial u'}{\partial u} \right| du^1\, du^2\, du^3,$$

and the covariant components of the second-order tensor g_{ij} transform as

$$\begin{aligned}
g'_{ij} &= \frac{\partial u^k}{\partial u'^i}\frac{\partial u^l}{\partial u'^j}\, g_{kl}, \\
\Rightarrow \quad g' &= \left| \frac{\partial u}{\partial u'} \right| \left| \frac{\partial u}{\partial u'} \right| g \text{ (on taking determinants)}, \\
\Rightarrow \quad \sqrt{g'} &= \left| \frac{\partial u}{\partial u'} \right| \sqrt{g}.
\end{aligned}$$

Thus, the new volume element is

$$\begin{aligned}
dV' &= \sqrt{g'}\, du'^1\, du'^2\, du'^3 \\
&= \left| \frac{\partial u}{\partial u'} \right| \sqrt{g} \left| \frac{\partial u'}{\partial u} \right| du^1\, du^2\, du^3 \\
&= \sqrt{g}\, du^1\, du^2\, du^3 = dV.
\end{aligned}$$

This shows that dV is a scalar quantity.

26.27 *Find an expression for the second covariant derivative, written in semicolon notation as $v_{i;jk} \equiv (v_{i;j})_{;k}$, of a vector v_i. By interchanging the order of differentiation and then subtracting the two expressions, we define the components $R^l_{\ ijk}$ of the* Riemann tensor *as*

$$v_{i;jk} - v_{i;kj} \equiv R^l_{\ ijk} v_l.$$

Show that in a general coordinate system u^i these components are given by

$$R^l_{\ ijk} = \frac{\partial \Gamma^l_{\ ik}}{\partial u^j} - \frac{\partial \Gamma^l_{\ ij}}{\partial u^k} + \Gamma^m_{\ ik}\Gamma^l_{\ mj} - \Gamma^m_{\ ij}\Gamma^l_{\ mk}.$$

By first considering Cartesian coordinates, show that all the components $R^l_{\ ijk} \equiv 0$ for any *coordinate system in three-dimensional Euclidean space.*
In such a space, therefore, we may change the order of the covariant derivatives without changing the resulting expression.

For the covariant derivative of the covariant components of a vector, we have

$$v_{i;j} = \frac{\partial v_i}{\partial u^j} - \Gamma^k_{\ ij} v_k,$$

where $\Gamma^k_{\ ij}$ is a Christoffel symbol of the second kind. Hence,

$$\begin{aligned}
v_{i;jk} &\equiv (v_{i;j})_{;k} \\
&= \left(\frac{\partial v_i}{\partial u^j} - \Gamma^l_{\ ij} v_l \right)_{;k} \\
&= \frac{\partial}{\partial u^k}\left(\frac{\partial v_i}{\partial u^j} - \Gamma^l_{\ ij} v_l \right) - \Gamma^m_{ik}\left(\frac{\partial v_m}{\partial u^j} - \Gamma^l_{\ mj} v_l \right) \\
&= \frac{\partial^2 v_i}{\partial u^k \partial u^j} - \Gamma^l_{\ ij}\frac{\partial v_l}{\partial u^k} - v_l \frac{\partial \Gamma^l_{\ ij}}{\partial u^k} - \Gamma^m_{ik}\frac{\partial v_m}{\partial u^j} + \Gamma^m_{ik}\Gamma^l_{\ mj} v_l.
\end{aligned}$$

Interchanging subscripts j and k,

$$v_{i;kj} = \frac{\partial^2 v_i}{\partial u^j \partial u^k} - \Gamma^l_{\ ik}\frac{\partial v_l}{\partial u^j} - v_l \frac{\partial \Gamma^l_{\ ik}}{\partial u^j} - \Gamma^m_{ij}\frac{\partial v_m}{\partial u^k} + \Gamma^m_{ij}\Gamma^l_{\ mk} v_l.$$

When these two expressions are subtracted to define the Riemann tensor, the first, second and fourth terms (the second of one with the fourth of the other and vice versa) on the two RHSs cancel in pairs to yield

$$R^l_{\ ijk} v_l \equiv v_{i;jk} - v_{i;kj} = \left(\frac{\partial \Gamma^l_{\ ik}}{\partial u^j} - \frac{\partial \Gamma^l_{\ ij}}{\partial u^k} + \Gamma^m_{ik}\Gamma^l_{\ mj} - \Gamma^m_{ij}\Gamma^l_{\ mk} \right) v_l.$$

Now, in three-dimensional Euclidean space, one possible coordinate system is the Cartesian one. In this system $g = 1$ and all of its derivatives are zero. Thus all Christoffel symbols and their derivatives are zero, as are all components of the Riemann tensor. As *all* the components vanish in this Cartesian coordinate system, they must do so in *any* coordinate system in this space.

26.29 *We may define Christoffel symbols of the first kind by*

$$\Gamma_{ijk} = g_{il}\Gamma^{l}{}_{jk}.$$

Show that these are given by

$$\Gamma_{ijk} = \frac{1}{2}\left(\frac{\partial g_{ki}}{\partial u^j} + \frac{\partial g_{ij}}{\partial u^k} - \frac{\partial g_{jk}}{\partial u^i}\right).$$

By permuting indices, verify that

$$\frac{\partial g_{ij}}{\partial u^k} = \Gamma_{ijk} + \Gamma_{jik}.$$

Using the fact that $\Gamma^{l}{}_{jk} = \Gamma^{l}{}_{kj}$, show that

$$g_{ij;k} \equiv 0,$$

i.e. that the covariant derivative of the metric tensor is identically zero in all coordinate systems.

Starting from Christoffel symbols of the second kind, we have

$$\begin{aligned}
\Gamma_{ijk} &= g_{il}\Gamma^{l}{}_{jk} \\
&= \frac{1}{2}g_{il}g^{ln}\left(\frac{\partial g_{kn}}{\partial u^j} + \frac{\partial g_{nj}}{\partial u^k} - \frac{\partial g_{jk}}{\partial u^n}\right) \\
&= \frac{1}{2}\delta_i^n\left(\frac{\partial g_{kn}}{\partial u^j} + \frac{\partial g_{nj}}{\partial u^k} - \frac{\partial g_{jk}}{\partial u^n}\right) \\
&= \frac{1}{2}\left(\frac{\partial g_{ki}}{\partial u^j} + \frac{\partial g_{ij}}{\partial u^k} - \frac{\partial g_{jk}}{\partial u^i}\right).
\end{aligned}$$

Next, forming the symmetric sum of two Christoffel symbols:

$$\begin{aligned}
\Gamma_{ijk} + \Gamma_{jik} &= \frac{1}{2}\left(\frac{\partial g_{ki}}{\partial u^j} + \frac{\partial g_{ij}}{\partial u^k} - \frac{\partial g_{jk}}{\partial u^i}\right) + \frac{1}{2}\left(\frac{\partial g_{kj}}{\partial u^i} + \frac{\partial g_{ji}}{\partial u^k} - \frac{\partial g_{ik}}{\partial u^j}\right) \\
&= \frac{1}{2}\left(\frac{\partial g_{ij}}{\partial u^k} + \frac{\partial g_{ji}}{\partial u^k}\right) + \frac{1}{2}\left(\frac{\partial g_{ki}}{\partial u^j} - \frac{\partial g_{ik}}{\partial u^j}\right) \\
&\qquad + \frac{1}{2}\left(-\frac{\partial g_{jk}}{\partial u^i} + \frac{\partial g_{kj}}{\partial u^i}\right) \\
&= \frac{\partial g_{ij}}{\partial u^k} + 0 + 0.
\end{aligned}$$

To obtain the last line we have used the fact that the metric tensor is symmetric, $g_{ij} = g_{ji}$.

Further, since $\Gamma^l{}_{jk} = \Gamma^l{}_{kj}$, and therefore $g_{il}\Gamma^l{}_{jk} = g_{il}\Gamma^l{}_{kj}$, we have that $\Gamma_{ijk} = \Gamma_{ikj}$, i.e Christoffel symbols of the first kind are symmetric under the interchange of the last two indices.

Finally, forming the covariant derivative of g_{ij}:

$$
\begin{aligned}
g_{ij;k} &= \frac{\partial}{\partial u^k}(g_{ij}\,\mathbf{e}^i \otimes \mathbf{e}^j) \\
&= \frac{\partial g_{ij}}{\partial u^k}\,\mathbf{e}^i \otimes \mathbf{e}^j + g_{ij}\,\frac{\partial \mathbf{e}^i}{\partial u^k} \otimes \mathbf{e}^j + g_{ij}\,\mathbf{e}^i \otimes \frac{\partial \mathbf{e}^j}{\partial u^k} \\
&= \frac{\partial g_{ij}}{\partial u^k}\,\mathbf{e}^i \otimes \mathbf{e}^j + g_{ij}\,(-\Gamma^i{}_{lk}\mathbf{e}^l) \otimes \mathbf{e}^j + g_{ij}\,\mathbf{e}^i \otimes (-\Gamma^j{}_{mk}\mathbf{e}^m) \\
&= \frac{\partial g_{ij}}{\partial u^k}\,\mathbf{e}^i \otimes \mathbf{e}^j - \Gamma_{jlk}\mathbf{e}^l \otimes \mathbf{e}^j - \Gamma_{imk}\mathbf{e}^i \otimes \mathbf{e}^m, \text{ since } g_{ij} = g_{ji}, \\
&= \left(\frac{\partial g_{ij}}{\partial u^k} - \Gamma_{jik} - \Gamma_{ijk}\right)\mathbf{e}^i \otimes \mathbf{e}^j, \text{ renaming dummy suffices,} \\
&= 0\,(\mathbf{e}^i \otimes \mathbf{e}^j), \text{ from the previous result.}
\end{aligned}
$$

Thus the covariant derivative of the metric tensor is identically zero in all coordinate systems.

27

Numerical methods

***27.1** Use an iteration procedure to find the root of the equation* $40x = \exp x$ *to four significant figures.*

To provide a satisfactory iteration scheme, the equation must be rearranged in the form $x = f(x)$, where $f(x)$ is a slowly varying function of x; we then use $x_{n+1} = f(x_n)$ as the iteration scheme.

In the present case the rearrangement is straightforward, as, by taking logarithms, we can write the equation as $x = \ln 40x$. Since $\ln z$ is a slowly varying function of z, we can take $x_{n+1} = \ln 40x_n$ as the iteration scheme.

We start with the (poor) guess that $x = 1$. The successive values generated by the scheme are (to 5 s.f.)

$$1,\ 3.6889,\ 4.9942,\ 5.2972,\ 5.3560,\ 5.3671,\ 5.3691,\ 5.3696,\ 5.3696,\ \ldots\ .$$

Thus to 4 s.f. we give the answer as $x = 5.370$. In fact, after 15 iterations the calculated value is stable to 10 s.f. at 5.369640395.

27.3 *Show the following results about rearrangement schemes for polynomial equations.*

(a) *That if a polynomial equation* $g(x) \equiv x^m - f(x) = 0$, *where* $f(x)$ *is a polynomial of degree less than m and for which* $f(0) \neq 0$, *is solved using a rearrangement iteration scheme* $x_{n+1} = [f(x_n)]^{1/m}$, *then, in general, the scheme will have only first-order convergence.*

(b) *By considering the cubic equation*

$$x^3 - ax^2 + 2abx - (b^3 + ab^2) = 0$$

for arbitrary non-zero values of a and b, demonstrate that, in special cases, the same rearrangement scheme can give second- (or higher-) order convergence.

(a) If we represent the iteration scheme as $x_{n+1} = F(x_n)$ then the scheme will have only first-order convergence unless $F'(\xi) = 0$, where ξ is the solution to the original equation satisfying $\xi^m = f(\xi)$ or, equivalently, $\xi = F(\xi)$.

In this case $F(x) = [\,f(x)\,]^{1/m}$ and

$$F'(\xi) = \frac{1}{m}[\,f(\xi)\,]^{(1-m)/m} f'(\xi).$$

Since $f(0) \neq 0$, $x = 0$ cannot be one of the solutions ξ of the original equation. Now, $f(\xi) = \xi^m$ and so the first two factors in the expression for $F'(\xi)$ have the value $m^{-1}(\xi^m)^{(1-m)/m} = m^{-1}\xi^{1-m}$. This is neither zero nor infinite and so $F'(\xi)$ can only be zero if $f'(\xi) = 0$; in general this will not be the case and the convergence will be only of first order.

(b) For the given equation $m = 3$ and $f(x) = ax^2 - 2abx + (b^3 + ab^2)$. It follows that $f'(x) = 2ax - 2ab$ and that $f'(x) = 0$ when $x = b$. However, $x = b$, also satisfies the original equation

$$b^3 - ab^2 + 2ab^2 - b^3 - ab^2 = 0,$$

and therefore, in the terminology used in part (a), $\xi = b$ and $F'(\xi) = F'(b) = 0$. This shows that the convergence will be of second (or higher) order.

In fact, further differentiation shows that $F''(b) = 2a/3b^2$ and, as this is non-zero, the convergence is only of second order.

***27.5** Solve the following set of simultaneous equations using Gaussian elimination (including interchange where it is formally desirable):*

$$\begin{aligned} x_1 + 3x_2 + 4x_3 + 2x_4 &= 0, \\ 2x_1 + 10x_2 - 5x_3 + x_4 &= 6, \\ 4x_2 + 3x_3 + 3x_4 &= 20, \\ -3x_1 + 6x_2 + 12x_3 - 4x_4 &= 16. \end{aligned}$$

Since the largest (in magnitude) coefficient of x_1 appears in the final equation, we reorder them to make it first (labelled I) and divide through by -1 to make the coefficient of x_1 positive:

$$3x_1 - 6x_2 - 12x_3 + 4x_4 = -16. \qquad \text{(I)}$$

The first and second equations now have $\frac{1}{3}$ and $\frac{2}{3}$ (respectively) of (I) subtracted from them to eliminate x_1. The third equation does not contain x_1 and so is left unchanged:

$$\begin{aligned} 5x_2 + 8x_3 + \tfrac{2}{3}x_4 &= \tfrac{16}{3}, && \text{(a)} \\ 14x_2 + 3x_3 - \tfrac{5}{3}x_4 &= \tfrac{50}{3}, && \text{(b)} \\ 4x_2 + 3x_3 + 3x_4 &= 20. && \text{(c)} \end{aligned}$$

Equation (b) is now the one with the largest coefficient of x_2, and so we take as the second finalised equation

$$14x_2 + 3x_3 - \tfrac{5}{3}x_4 = \tfrac{50}{3}, \qquad \text{(II)}$$

and subtract the needed fractions of this from (a) and (c) to eliminate x_2 from them:

$$\begin{aligned} \tfrac{97}{14}x_3 + \left(\tfrac{2}{3} + \tfrac{25}{42}\right)x_4 &= \tfrac{16}{3} - \tfrac{250}{42}, && \text{(d)} \\ \left(3 - \tfrac{12}{14}\right)x_3 + \left(3 + \tfrac{20}{42}\right)x_4 &= 20 - \tfrac{200}{42}. && \text{(e)} \end{aligned}$$

Rationalising these two equations we have

$$\begin{aligned} 291x_3 + 53x_4 &= -26, && \text{(d}'\text{)} \equiv \text{(III)} \\ 90x_3 + 146x_4 &= 640. && \text{(e}'\text{)} \end{aligned}$$

Finally, eliminating x_3 from (e$'$) gives

$$\begin{aligned} \left(146 - \tfrac{90}{291}\,53\right)x_4 &= 640 - \tfrac{90}{291}(-26), \\ 37716x_4 &= 188580, \\ x_4 &= 5. \qquad \text{(IV)} \end{aligned}$$

Resubstitution then gives

$$\text{from (III)},\ x_3 = \tfrac{-26-(53\times 5)}{291} = -1,$$

$$\text{from (II)},\ x_2 = \tfrac{1}{14}(\tfrac{50}{3} + \tfrac{5\times 5}{3} - 3(-1)) = 2,$$

$$\text{from (I)},\ x_1 = \tfrac{1}{3}(-16 - (4\times 5) + 12(-1) + 6(2)) = -12,$$

making the solution $x_1 = -12$, $x_2 = 2$, $x_3 = -1$ and $x_4 = 5$.

27.7 *Simultaneous linear equations that result in tridiagonal matrices can sometimes be solved in the same way as three-term recurrence relations. Consider the tridiagonal simultaneous equations*

$$x_{i-1} + 4x_i + x_{i+1} = 3(\delta_{i+1,0} - \delta_{i-1,0}), \quad i = 0, \pm 1, \pm 2, \ldots .$$

Prove that for $i > 0$ the equations have a general solution of the form $x_i = \alpha p^i + \beta q^i$, where p and q are the roots of a certain quadratic equation. Show that a similar result holds for $i < 0$. In each case express x_0 in terms of the arbitrary constants $\alpha, \beta, \ldots$.

Now impose the condition that x_i is bounded as $i \to \pm\infty$ and obtain a unique solution.

We substitute the trial solution $x_i = \alpha p^i + \beta q^i$ into the given equation for $i \geq 2$ and obtain

$$\alpha(p^{i-1} + 4p^i + p^{i+1}) + \beta(q^{i-1} + 4q^i + q^{i+1}) = 3(0-0) = 0;$$

this is satisfied for arbitrary α and β if p and q are the two roots of the quadratic equation $1 + 4r + r^2 = 0$.

Using the same form for $i = 1$, but with these specific values for p and q, we have

$$\begin{aligned} x_0 + 4x_1 + x_2 &= 3(0-1), \\ x_0 + 4(\alpha p + \beta q) + \alpha p^2 + \beta q^2 &= -3, \\ x_0 + \alpha(4p + p^2) + \beta(4q + q^2) &= -3, \\ x_0 + \alpha(-1) + \beta(-1) &= -3. \end{aligned}$$

To obtain the final line we used the fact that both p and q satisfy $4r + r^2 = -1$. Similarly, for $i \leq -1$ the solution is $x_i = \alpha' p^i + \beta' q^i$, with $x_0 - \alpha' - \beta' = +3$.

In addition, for $i = 0$, we have from the original equation that

$$\frac{\alpha'}{p} + \frac{\beta'}{q} + 4x_0 + \alpha p + \beta q = 0.$$

The values of p and q are $-2 \pm \sqrt{4-1}$, with, say, $p = -2 + \sqrt{3}$ and $|p| < 1$, and $q = -2 - \sqrt{3}$ and $|q| > 1$.

Now, the solution is to be bounded as $i \to \pm\infty$. The fact that $|q| > 1$ and the condition at $+\infty$ together require that $\beta = 0$, whilst $|p|^{-1} > 1$ and the condition at $-\infty$ imply that $\alpha' = 0$. We are left with three equations for three unknowns:

$$\begin{aligned} x_0 - \alpha + 3 &= 0, \\ x_0 - \beta' - 3 &= 0, \\ \frac{\beta'}{-2-\sqrt{3}} + 4x_0 + \alpha(-2+\sqrt{3}) &= 0. \end{aligned}$$

We now rearrange the last of these and substitute from the first two:

$$\begin{aligned} \beta' + 4(-2-\sqrt{3})x_0 + \alpha &= 0, \\ \Rightarrow \quad (x_0 - 3) - (8 + 4\sqrt{3})x_0 + (x_0 + 3) &= 0, \end{aligned}$$

and $x_0 = 0,\ \alpha = 3,\ \beta' = -3$. The solution is thus

$$x_i = \begin{cases} 3(-2+\sqrt{3})^i & i \geq 1, \\ 0 & i = 0, \\ -3(-2-\sqrt{3})^i & i \leq -1. \end{cases}$$

The final entry could be written as $-3(-2+\sqrt{3})^{-i}$.

27.9 *Although it can easily be shown, by direct calculation, that*

$$\int_0^\infty e^{-x} \cos(kx)\, dx = \frac{1}{1+k^2},$$

the form of the integrand is also appropriate for a Gauss–Laguerre numerical integration. Using a 5-point formula, investigate the range of values of k for which the formula gives accurate results. At about what value of k do the results become inaccurate at the 1% level?

The integrand is an even function of k and so only positive k need be considered. The points and weights for the 5-point Gauss–Laguerre integration are

x_i	w_i
0.26356 03197	0.52175 56106
1.41340 30591	0.39866 68111
3.59642 57710	0.07594 24497
7.08581 00059	0.00361 17587
12.6408 00844	0.00002 33700

The table below gives the exact and calculated results to four places of decimals, as well as the percentage error in the calculated result. It shows that the error is not more than 1% for $|k|$ less than about 1.1.

k	Exact	Calculated	% error
0.0	1.0000	1.0000	0.0
0.5	0.8000	0.8000	0.0
0.8	0.6098	0.6097	0.0
1.0	0.5000	0.5005	0.1
1.1	0.4525	0.4545	0.4
1.2	0.4098	0.4145	1.1
1.3	0.3717	0.3800	2.2
1.5	0.3077	0.3200	4.0
1.7	0.2571	0.2535	−1.4
2.0	0.2000	0.1184	−40.8
3.0	0.1000	0.1674	67.4

27.11 *Consider the integrals I_p defined by*

$$I_p = \int_{-1}^{1} \frac{x^{2p}}{\sqrt{1-x^2}}\, dx.$$

(a) *By setting $x = \sin\theta$ and using the recurrence relation quoted below, show that I_p has the value*

$$I_p = 2\,\frac{2p-1}{2p}\,\frac{2p-3}{2p-2}\cdots\frac{1}{2}\,\frac{\pi}{2}.$$

Recurrence relation: *If $J(n)$ is defined for a non-negative integer n by*

$$J(n) = \int_0^{\pi/2} \sin^n\theta\, d\theta,$$

then, for $n > 2$,

$$J(n) = \frac{n-1}{n} J(n-2).$$

(b) *Evaluate I_p for $p = 1, 2, \ldots, 6$ using 5- and 6-point Gauss–Chebyshev integration (conveniently run on a spreadsheet such as* Excel*) and compare the results with those in* (a)*. In particular, show that, as expected, the* 5*-point scheme first fails to be accurate when the order of the polynomial numerator* ($2p$) *exceeds* $(2 \times 5) - 1 = 9$*. Likewise, verify that the* 6*-point scheme evaluates I_5 accurately but is in error for I_6.*

(a) Setting $x = \sin\theta$ with $dx = \cos\theta$ converts I_p to

$$I_p = \int_{-\pi/2}^{\pi/2} \frac{\sin^{2p}\theta}{\cos\theta}\cos\theta\, d\theta = 2\int_0^{\pi/2} \sin^{2p}\theta\, d\theta = 2J(2p),$$

using the given definition of $J(n)$. Applying the reduction formula then gives

$$I_p = 2\,\frac{2p-1}{2p}\,\frac{2p-3}{2p-2}\cdots\frac{1}{2}\,\frac{\pi}{2},$$

where we have used the obvious result $J(0) = \pi/2$.

(b) The points and weights needed for a Gauss–Chebyshev integration are given analytically by

$$x_i = \cos\frac{(i-\frac{1}{2})\pi}{n}, \qquad w_i = \frac{\pi}{n}, \qquad \text{for } i = 1, \ldots, n.$$

Here we have to take the cases $n = 5$ and $n = 6$. The following table gives the exact result calculated in (a) and the values obtained using the n-point Gauss–Chebyshev formula.

p	Exact	$n = 5$	$n = 6$
1	1.570796	1.570796	1.570796
2	1.178097	1.178097	1.178097
3	0.981748	0.981748	0.981748
4	0.859029	0.859029	0.859029
5	0.773126	0.766990	0.773126
6	0.708699	0.690291	0.707165

It will be seen that, as stated in the question, the $p = 5$, $n = 5$ and both the $p = 6$ values diverge from the exact result. The discrepancy is of the order of 1% when $p = n$, i.e. when the order of the polynomial in the numerator of I_p first exceeds $2n - 1$.

27.13 *Given a random number η uniformly distributed on $(0, 1)$, determine the function $\xi = \xi(\eta)$ that would generate a random number ξ distributed as*

(a) 2ξ *on* $0 \le \xi < 1$,
(b) $\frac{3}{2}\sqrt{\xi}$ *on* $0 \le \xi < 1$,
(c) $\dfrac{\pi}{4a}\cos\dfrac{\pi\xi}{2a}$ *on* $-a \le \xi < a$,
(d) $\frac{1}{2}\exp(-\mid\xi\mid)$ *on* $-\infty < \xi < \infty$.

For each required distribution $f(t)$ in the range (a, b) we need to determine the cumulative distribution function $F(y) = \int_a^y dt$ and then take $F(y)$ as uniformly distributed on $(0, 1)$. A correctly normalised distribution has $F(b) = 1$. For any given random number η, the corresponding variable, distributed as $f(\xi)$, is $\xi = F^{-1}(\eta)$.

(a) For $f(t) = 2t$,

$$F(y) = \int_0^y 2t\,dt = y^2 \quad \Rightarrow \quad \eta = \xi^2 \quad \Rightarrow \quad \xi = \sqrt{\eta}.$$

(b) For $f(t) = \frac{3}{2}\sqrt{t}$,

$$F(y) = \int_0^y \tfrac{3}{2}\sqrt{t}\,dt = y^{3/2} \quad \Rightarrow \quad \eta = \xi^{3/2} \quad \Rightarrow \quad \xi = \eta^{2/3}.$$

(c) For $f(t) = \dfrac{\pi}{4a}\cos\dfrac{\pi t}{2a}$,

$$F(y) = \frac{\pi}{4a}\int_{-a}^y \cos\left(\frac{\pi t}{2a}\right) dt = \frac{1}{2}\left[\sin\left(\frac{\pi y}{2a}\right) + 1\right],$$

$$\Rightarrow \quad \eta = \frac{1}{2}\left[\sin\left(\frac{\pi \xi}{2a}\right) + 1\right] \quad \Rightarrow \quad \xi = \frac{2a}{\pi}\sin^{-1}(2\eta - 1).$$

(d) For $f(t) = \frac{1}{2}\exp(-|\,t\,|)$,

for $y < 0$, $\quad F(y) = \displaystyle\int_{-\infty}^y \frac{e^t}{2}\,dt = \frac{e^y}{2}$,

$$\text{for } y > 0, \quad F(y) = \int_{-\infty}^0 \frac{e^t}{2}\,dt + \int_0^y \frac{e^{-t}}{2}\,dt = \frac{1}{2} + \frac{1 - e^{-y}}{2} = \frac{1}{2}(2 - e^{-y}).$$

It follows that

$$\eta = \begin{cases} \frac{1}{2}e^{\xi} & \xi \le 0, \\ 1 - \frac{1}{2}e^{-\xi} & \xi > 0, \end{cases} \quad \text{and } \xi = \begin{cases} \ln 2\eta & \eta \le 0.5, \\ -\ln(2 - 2\eta) & 0.5 < \eta < 1. \end{cases}$$

27.15 *Use a Taylor series to solve the equation*

$$\frac{dy}{dx} + xy = 0, \qquad y(0) = 1,$$

evaluating $y(x)$ *for* $x = 0.0$ *to* 0.5 *in steps of* 0.1.

In order to construct the Taylor series we need to find the derivatives $y^{(n)} \equiv$

$d^{(n)}y/dx^n$ up to, say, $n = 6$ and evaluate them at $x = 0$. We will also need $y(0) = 1$. The derivatives are

$$\begin{aligned}
y' &= -xy \quad \Rightarrow \quad y^{(1)}(0) = 0,\\
y^{(2)} &= = -y - xy'(x) = -y + x^2y \quad \Rightarrow \quad y^{(2)}(0) = -1,\\
y^{(3)} &= 2xy + (-1 + x^2)y'(x) = 3xy - x^3y \quad \Rightarrow \quad y^{(3)}(0) = 0,\\
y^{(4)} &= 3y - 3x^2y + (3x - x^3)y'(x)\\
&= 3y - 6x^2y + x^4y \quad \Rightarrow \quad y^{(4)}(0) = 3,\\
y^{(5)} &= -12xy + 4x^3y + (3 - 6x^2 + x^4)y'(x)\\
&= -15xy + 10x^3y - x^5y \quad \Rightarrow \quad y^{(5)}(0) = 0,\\
y^{(6)} &= -15y + 30x^2y - 5x^4y + (-15x + 10x^3 - x^5)y'(x)\\
&= -15y + 45x^2y - 15x^4y + x^6y \quad \Rightarrow \quad y^{(6)}(0) = -15.
\end{aligned}$$

Thus, the Taylor series for an expansion about $x = 0$ is given by

$$\begin{aligned}
y(x) &= 1 - \frac{x^2}{2!} + \frac{3x^4}{4!} - \frac{15x^6}{6!} + \mathrm{O}(x^8)\\
&= 1 - \frac{x^2}{2} + \frac{x^4}{8} - \frac{x^6}{48} + \mathrm{O}(x^8).
\end{aligned}$$

To four significant figures the values of $y(x)$ calculated using this Taylor series are $y(0.1) = 0.9950$, $y(0.2) = 0.9802$, $y(0.3) = 0.9560$, $y(0.4) = 0.9231$ and $y(0.5) = 0.8825$.

For interest, we note that the exact solution of the differential equation, which is separable, is given by

$$\begin{aligned}
\frac{dy}{y} = -x\,dx \quad &\Rightarrow \quad \ln y = -\frac{x^2}{2} + c\\
y(0) = 1 \quad &\Rightarrow \quad c = 0 \quad \Rightarrow \quad y(x) = e^{-x^2/2},
\end{aligned}$$

which has the Taylor series

$$y(x) = 1 - \frac{x^2}{2^1\,1!} + \frac{x^4}{2^2\,2!} - \frac{x^6}{2^3\,3!} + \cdots .$$

As expected, this is the same as that found directly from the differential equation, up to the last term calculated; clearly the next term is $O(x^8)$.

To four significant figures the exact solution and the Taylor expansion give the same values over the given range of x; for $x = 0.6$ they differ by 1 in the fourth decimal place.

***27.17** A more refined form of the Adams predictor–corrector method for solving the first-order differential equation*

$$\frac{dy}{dx} = f(x, y)$$

is known as the Adams–Moulton–Bashforth scheme. At any stage (say the nth) in an Nth-order scheme, the values of x and y at the previous N solution points are first used to predict the value of y_{n+1}. This approximate value of y at the next solution point, x_{n+1}, denoted by $\bar{y}_{n+1}$, is then used together with those at the previous $N-1$ solution points to make a more refined (corrected) estimation of $y(x_{n+1})$. The calculational procedure for a third-order scheme is summarised by the following two equations:

$$\overline{y_{n+1}} = y_n + h(a_1 f_n + a_2 f_{n-1} + a_3 f_{n-2}) \qquad \textit{(predictor)},$$
$$y_{n+1} = y_n + h(b_1 f(x_{n+1}, \overline{y_{n+1}}) + b_2 f_n + b_3 f_{n-1}) \qquad \textit{(corrector)}.$$

(a) *Find Taylor series expansions for f_{n-1} and f_{n-2} in terms of the function $f_n = f(x_n, y_n)$ and its derivatives at x_n.*
(b) *Substitute them into the predictor equation and, by making that expression for $\bar{y}_{n+1}$ coincide with the true Taylor series for y_{n+1} up to order h^3, establish simultaneous equations that determine the values of a_1, a_2 and a_3.*
(c) *Find the Taylor series for f_{n+1} and substitute it and that for f_{n-1} into the corrector equation. Make the corrected prediction for y_{n+1} coincide with the true Taylor series by choosing the weights b_1, b_2 and b_3 appropriately.*
(d) *The values of the numerical solution of the differential equation*

$$\frac{dy}{dx} = \frac{2(1+x)y + x^{3/2}}{2x(1+x)}$$

at three values of x are given in the following table.

x	0.1	0.2	0.3
$y(x)$	0.030628	0.084107	0.150328

Use the above predictor–corrector scheme to find the value of $y(0.4)$ and compare your answer with the accurate value, 0.225577.

(a) ‘Taylor series’ expansions, using increments in x of $-h$ and $-2h$, give

$$f_{n-1} = f_n - hf_n' + \tfrac{1}{2}h^2 f_n'' - \tfrac{1}{6}h^3 f_n^{(3)} + \cdots,$$
$$f_{n-2} = f_n - 2hf_n' + \tfrac{4}{2}h^2 f_n'' - \tfrac{8}{6}h^3 f_n^{(3)} + \cdots.$$

These expansions are not true Taylor series as the only derivatives used are those with respect to x; however, the same is true of all subsequent expansions.

(b) Substitution in the predictor equation gives

$$\begin{aligned}\overline{y_{n+1}} &= y_n + h(a_1 f_n + a_2 f_{n-1} + a_3 f_{n-2}) \\ &= y_n + h[\,(a_1 + a_2 + a_3) f_n + h(-a_2 - 2a_3) f_n' \\ &\qquad + h^2(\tfrac{1}{2}a_2 + 2a_3) f_n'' + \cdots].\end{aligned}$$

Now, the accurate Taylor series for y_{n+1} is

$$y_{n+1} = y_n + hf_n + \tfrac{1}{2}h^2 f_n' + \tfrac{1}{6}h^3 f_n'' + \cdots .$$

To make these two expressions coincide up to order h^3, we need

$$\left.\begin{aligned} a_1 + a_2 + a_3 &= 1 \\ -a_2 - 2a_3 &= \tfrac{1}{2} \\ \tfrac{1}{2}a_2 + 2a_3 &= \tfrac{1}{6} \end{aligned}\right\} a_2 = -\tfrac{4}{3},\ a_3 = \tfrac{5}{12},\ a_1 = \tfrac{23}{12}.$$

(c) In the same way as in part (a),

$$f_{n+1} = f_n + hf_n' + \tfrac{1}{2}h^2 f_n'' + \tfrac{1}{6}h^3 f_n^{(3)} + \cdots ,$$

and substitution in the corrector equation gives

$$\begin{aligned} y_{n+1} &= y_n + h(b_1 f(x_{n+1}, \overline{y_{n+1}}) + b_2 f_n + b_3 f_{n-1}) \\ &= y_n + h[\,b_1 f(x_{n+1}, y_{n+1}) + b_2 f_n + b_3 f_{n-1}\,],\ \text{ to order } h^3 \\ &\equiv y_n + h(b_1 f_{n+1} + b_2 f_n + b_3 f_{n-1}),\ \text{ to order } h^3 \\ &= y_n + h[\,(b_1 + b_2 + b_3) f_n + h(b_1 - b_3) f_n' \\ &\qquad + h^2(\tfrac{1}{2}b_1 + \tfrac{1}{2}b_3) f_n'' + \cdots].\end{aligned}$$

To make this coincide with the accurate Taylor series up to order h^3, we need

$$\left.\begin{aligned} b_1 + b_2 + b_3 &= 1 \\ b_1 - b_3 &= \tfrac{1}{2} \\ \tfrac{1}{2}b_1 + \tfrac{1}{2}b_3 &= \tfrac{1}{6} \end{aligned}\right\} b_1 = \tfrac{5}{12},\ b_3 = -\tfrac{1}{12},\ b_2 = \tfrac{2}{3}.$$

(d) We repeat the given table, indexing it and adding a line giving the values of $f(x, y)$.

n	1	2	3
x_n	0.1	0.2	0.3
$y_n(x_n)$	0.030628	0.084107	0.150328
$f_n(x_n, y_n)$	0.450020	0.606874	0.711756

Now, taking $n = 3$, we apply the predictor formula with the calculated values for the a_i and find $\overline{y_4} = 0.224582$. This allows us to calculate $f(x_4, \overline{y_4})$ as 0.787332. Finally, applying the corrector formula, using the calculated values for the b_i, we find the corrected value $y_4 = 0.225527$. This is to be compared with the accurate value of 0.225577 (and the predicted, but uncorrected, value of 0.224582).

27.19 *To solve the ordinary differential equation*

$$\frac{du}{dt} = f(u, t)$$

for $f = f(t)$, the explicit two-step finite difference scheme

$$u_{n+1} = \alpha u_n + \beta u_{n-1} + h(\mu f_n + \nu f_{n-1})$$

may be used. Here, in the usual notation, h is the time step, $t_n = nh$, $u_n = u(t_n)$ and $f_n = f(u_n, t_n)$; α, β, μ, and ν are constants.

(a) *A particular scheme has $\alpha = 1$, $\beta = 0, \mu = 3/2$ and $\nu = -1/2$. By considering Taylor expansions about $t = t_n$ for both u_{n+j} and f_{n+j}, show that this scheme gives errors of order h^3.*
(b) *Find the values of α, β, μ and ν that will give the greatest accuracy.*

We will need the Taylor expansions of $u_{n\pm1}$ and f_{n-1}. They are given by

$$\begin{aligned} u_{n\pm1} &= u_n \pm hu_n' + \frac{1}{2!}\,h^2u_n'' \pm \frac{1}{3!}\,h^3u_n^{(3)} + \cdots, \\ f_{n-1} &= u_{n-1}' = u_n' - hu_n'' + \frac{1}{2!}\,h^2u_n^{(3)} - \frac{1}{3!}\,h^3u_n^{(4)} + \cdots. \end{aligned}$$

(a) This scheme calculates u_{n+1} as

$$\begin{aligned} u_{n+1} &= u_n + h\left(\frac{3}{2}f_n - \frac{1}{2}f_{n-1}\right) \\ &= u_n + h\left[\frac{3}{2}u_n' - \frac{1}{2}\left(u_n' - hu_n'' + \frac{1}{2!}\,h^2u_n^{(3)} - \frac{1}{3!}\,h^3u_n^{(4)} + \cdots\right)\right]. \end{aligned}$$

This is to be compared with

$$u_{n+1} = u_n + hu_n' + \frac{1}{2!}\,h^2u_n'' + \frac{1}{3!}\,h^3u_n^{(3)} + \cdots.$$

Omitting terms that appear in both expressions, we have

$$\frac{1}{3!}\,h^3u_n^{(3)} + \cdots \approx -\frac{1}{4}\,h^3u_n^{(3)} + \cdots,$$

showing that the error is

$$\left(\frac{1}{3!} + \frac{1}{4}\right)h^3u_n^{(3)} = \frac{5}{12}\,h^3u_n^{(3)} + O(h^4).$$

(b) For the best accuracy we require that

$$u_{n+1} = u_n + hu_n' + \frac{1}{2!}\,h^2u_n'' + \frac{1}{3!}\,h^3u_n^{(3)} + \cdots$$

and

$$\alpha u_n + \beta\left(u_n - hu_n' + \frac{1}{2!}\,h^2u_n'' - \frac{1}{3!}\,h^3u_n^{(3)} + \cdots\right) + h\mu u_n'$$
$$+\, h\nu\left(u_n' - hu_n'' + \frac{1}{2!}\,h^2u_n^{(3)} - \frac{1}{3!}\,h^3u_n^{(4)} + \cdots\right)$$

should match up to as high a positive power of h as possible.

With four parameters available, we can expect to match terms in h^n up to $n = 3$:

$$\begin{aligned}
h^0 &: \ 1 = \alpha + \beta,\\
h^1 &: \ 1 = -\beta + \mu + \nu,\\
h^2 &: \ \tfrac{1}{2} = \tfrac{1}{2}\,\beta - \nu,\\
h^3 &: \ \tfrac{1}{6} = -\tfrac{1}{6}\,\beta + \tfrac{1}{2}\,\nu.
\end{aligned}$$

The final two equations are equivalent to $\beta = 1 + 2\nu$ and $1 + \beta = 3\nu$, yielding $\nu = 2$ and $\beta = 5$; it then follows that $\mu = 4$ and $\alpha = -4$. With this set of values, the finite difference scheme,

$$u_{n+1} = -4u_n + 5u_{n-1} + h(4f_n + 2f_{n-1}),$$

has errors of order h^4.

27.21 *Write a computer program that would solve, for a range of values of λ, the differential equation*

$$\frac{dy}{dx} = \frac{1}{\sqrt{x^2 + \lambda y^2}}, \qquad y(0) = 1,$$

using a third-order Runge–Kutta scheme. Consider the difficulties that might arise when $\lambda < 0$.

The relevant equations for a third-order Runge–Kutta scheme are

$$y_{i+1} = y_i + \tfrac{1}{6}(b_1 + 4b_2 + b_3),$$

where

$$\begin{aligned}
b_1 &= hf(x_i, y_i),\\
b_2 &= hf(x_i + \tfrac{1}{2}h,\ y_i + \tfrac{1}{2}b_1),\\
b_3 &= hf(x_i + h,\ y_i + 2b_2 - b_1).
\end{aligned}$$

The function $f(x, y)$, in this case, is $(x^2 + \lambda y^2)^{-1/2}$.

This calculation can be set up easily on a spreadsheet such as *Excel*, and it is immediately apparent that, with the given boundary value $y(0) = 1$, no significant finesse is needed. For positive values of λ the solution y is a monotonically (and boringly!) increasing function of x with values lying between 1 and ∞, the latter being approached rapidly only when λ is very small. Even with λ as small as 0.01, a step size Δx of 0.1 is adequate unless great precision is needed.

The difficulties that might arise for $\lambda < 0$ do not need much consideration; there is no real solution for *any* negative value of λ. The reason for this is easy to see. At the initial point, $x = 0$, $y = 1$ and λy^2 is negative and so the square root does not yield a real value for the derivative dy/dx.

More interesting results arise if the initial value is given elsewhere than at $x = 0$. For example, if $f(1) = 1$ then a solution can be calculated for negative values of λ greater than about -0.582 and if $f(1) = 2$ then a solution exists for $\lambda > -0.2057$.

27.23 *For some problems, numerical or algebraic experimentation may suggest the form of the complete solution. Consider the problem of numerically integrating the first-order wave equation*

$$\frac{\partial u}{\partial t} + A\frac{\partial u}{\partial x} = 0,$$

in which A is a positive constant. A finite difference scheme for this partial differential equation is

$$\frac{u(p, n+1) - u(p, n)}{\Delta t} + A\frac{u(p, n) - u(p-1, n)}{\Delta x} = 0,$$

where $x = p\Delta x$ and $t = n\Delta t$, with p any integer and n a non-negative integer. The initial values are $u(0, 0) = 1$ and $u(p, 0) = 0$ for $p \neq 0$.

(a) *Carry the difference equation forward in time for two or three steps and attempt to identify the pattern of solution. Establish the criterion for the method to be numerically stable.*
(b) *Suggest a general form for $u(p, n)$, expressing it in generator function form, i.e. as '$u(p, n)$ is the coefficient of s^p in the expansion of $G(n, s)$'.*
(c) *Using your form of solution (or that given in the answers!), obtain an explicit general expression for $u(p, n)$ and verify it by direct substitution into the difference equation.*
(d) *An analytic solution of the original PDE indicates that an initial disturbance propagates undistorted. Under what circumstances would the difference scheme reproduce that behaviour?*

If we write $A\Delta t/\Delta x$ as c, the equation becomes

$$u(p, n+1) - u(p, n) + c[\,u(p, n) - u(p-1, n)\,] = 0,$$

with $u(0, 0) = 1$ and $u(p, 0) = 0$ for $p \neq 0$.

(a) For calculational purposes we rearrange the equation and then substitute trial values:

$$\begin{aligned}
u(p, n+1) &= (1-c)u(p, n) + cu(p-1, n), \qquad\qquad (*)\\
u(0, 1) &= (1-c)u(0, 0) + cu(-1, 0) = 1 - c,\\
u(1, 1) &= (1-c)u(1, 0) + cu(0, 0) = c,\\
u(m, 1) &= (1-c)u(m, 0) + cu(m-1, 0) = 0 \text{ for } m > 1,\\
u(0, 2) &= (1-c)u(0, 1) + cu(-1, 1) = (1-c)^2,\\
u(1, 2) &= (1-c)u(1, 1) + cu(0, 1) = 2c(1-c),\\
u(2, 2) &= (1-c)u(2, 1) + cu(1, 1) = c^2,\\
u(m, 2) &= (1-c)u(m, 1) + cu(m-1, 1) = 0 \text{ for } m > 2.
\end{aligned}$$

By now the pattern is clear, as is the condition for numerical stability, namely $c < 1$.

(b) For the nth time-step, the $n+1$ values of $u(p, n)$, $p = 0, 1, \ldots, n$ appear to be given by the terms in the binomial expansion of $[\,(1-c) + cs\,]^n$. Using the language of generating functions, we would say that '$u(p, n)$ is the coefficient of s^p in the expansion of $[\,(1-c) + cs\,]^n$'.

(c) If this conjecture is correct, then

$$u(p, n) = \frac{n!\,(1-c)^{n-p}\,c^p}{p\,!\,(n-p)!}.$$

Substituting this form into the difference equation $(*)$ yields

$$\frac{(n+1)!\,(1-c)^{n+1-p}\,c^p}{p\,!\,(n+1-p)!} = \frac{(1-c)\,n!\,(1-c)^{n-p}\,c^p}{p\,!\,(n-p)!} + \frac{c\,n!\,(1-c)^{n+1-p}\,c^{p-1}}{(p-1)!\,(n+1-p)!}.$$

Multiplying through by $p\,!\,(n+1-p)!$ and dividing by $n!\,(1-c)^{n+1-p}\,c^p$ gives

$$(n+1) = (n-p+1) + p.$$

This is satisfied for all n and p, showing that the proposed solution satisfies the equation. It also gives $u(0, 0) = 1$, confirming that it is the required solution.

(d) For the special case $c = 1$, the recurrence relation reduces to

$$u(p, n+1) = u(p-1, n),$$

i.e. the disturbance u at the point $p\Delta x$ at time $(n+1)\Delta t$ is exactly the same as that at position $(p-1)\Delta x$ one time-step earlier. In other words, the disturbance propagates undistorted at speed A.

From the point of view of the numerical integration, this situation ($c = 1$ exactly) is both on the edge of instability and unlikely to be realised in practice.

27.25 *Laplace's equation,*

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0,$$

is to be solved for the region and boundary conditions shown in figure 27.1.

Starting from the given initial guess for the potential values V and using the simplest possible form of relaxation, obtain a better approximation to the actual solution. Do not aim to be more accurate than ± 0.5 units and so terminate the process when subsequent changes would be no greater than this.

We start by imposing a coordinate grid symmetrically on the region, so that the initial guess is

$$V(0,1) = V(\pm 1, 1) = 20,$$
$$V(i,2) = 40 \text{ for all } i,$$

and the fixed boundary conditions are

$$V(i,0) = 0 \quad \text{for } |i| < 2,$$
$$V(i,1) = 0 \quad \text{for all } |i| \geq 2,$$
$$V(i,3) = 80 \quad \text{for all } i.$$

On symmetry grounds, we need consider only non-negative values of i.

We now apply the simplest relaxation scheme,

$$V_{i,j} \to \tfrac{1}{4}(V_{i+1,j} + V_{i-1,j} + V_{i,j+1} + V_{i,j-1}),$$

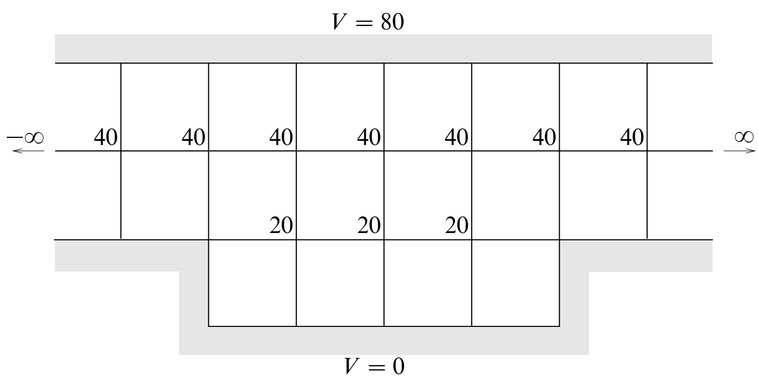

Figure 27.1 Region, boundary values and initial guessed solution values.

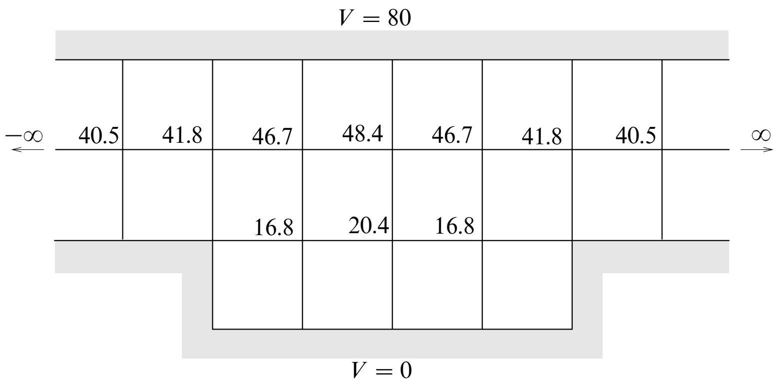

Figure 27.2 The solution to exercise 27.25.

for each point (i, j) that does not lie on the boundaries, where V_{ij} is prescribed and cannot be changed. The very simplest scheme would use only values from the previous iteration, but there is no additional labour involved in using previously calculated values from the current iteration when evaluating the RHS of the relationship. For this scheme the first few iterations produce the following results (to 3 s.f.):

$V_{0,1}$	$V_{1,1}$	$V_{0,2}$	$V_{1,2}$	$V_{2,2}$	$V_{3,2}$
20.0	20.0	40.0	40.0	40.0	40.0
20.0	15.0	45.0	45.0	41.3	40.3
18.8	15.9	47.2	46.1	41.6	40.4
19.8	16.5	48.0	46.5	41.7	40.4
20.2	16.7	48.3	46.7	41.8	40.4
20.4	16.8	48.4	46.8	41.8	40.4

The value at $(0, 1)$ is the one most likely to show the largest change at each iteration, as it is the one 'furthest from the fixed boundaries'. As the most recent changes have been 0.4 and 0.2, the process can be halted at this point, although the monotonic behaviour of values after the second iteration makes it harder to be sure that the differences between the final values and the current ones are within any given range.

The correct self-consistent solution (again to 3 s.f.) has corresponding values 20.6, 16.8, 48.5, 46.8, 41.8 and 40.5. This set of values is reached after nine iterations and is shown in figure 27.2. If the values from the previous iteration (rather than the most recently calculated ones) are used, the same ultimate result is reached (as expected), but about 17 iterations are needed to achieve the same self-consistency.

27.27 *The Schrödinger equation for a quantum mechanical particle of mass m moving in a one-dimensional harmonic oscillator potential $V(x) = kx^2/2$ is*

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + \frac{kx^2\psi}{2} = E\psi.$$

For physically acceptable solutions the wavefunction $\psi(x)$ must be finite at $x = 0$, tend to zero as $x \to \pm\infty$ and be normalised, so that $\int |\psi|^2\,dx = 1$. In practice, these constraints mean that only certain (quantised) values of E, the energy of the particle, are allowed. The allowed values fall into two groups, those for which $\psi(0) = 0$ and those for which $\psi(0) \neq 0$.

Show that if the unit of length is taken as $[\hbar^2/(mk)]^{1/4}$ and the unit of energy as $\hbar(k/m)^{1/2}$ then the Schrödinger equation takes the form

$$\frac{d^2\psi}{dy^2} + (2E' - y^2)\psi = 0.$$

Devise an outline computerised scheme, using Runge–Kutta integration, that will enable you to:

- *determine the three lowest allowed values of E;*
- *tabulate the normalised wavefunction corresponding to the lowest allowed energy.*

You should consider explicitly:

- *the variables to use in the numerical integration;*
- *how starting values near $y = 0$ are to be chosen;*
- *how the condition on ψ as $y \to \pm\infty$ is to be implemented;*
- *how the required values of E are to be extracted from the results of the integration;*
- *how the normalisation is to be carried out.*

We start by setting $x = \alpha y$, where α is the new unit of length; then $d/dx = \alpha^{-1} d/dy$ and

$$-\frac{\hbar^2}{2m}\alpha^{-2}\frac{d^2\psi}{dy^2} + \frac{k\alpha^2 y^2}{2} = E\psi,$$

$$\frac{d^2\psi}{dy^2} - \alpha^4\frac{mk}{\hbar^2}y^2\psi + \alpha^2\frac{2mE}{\hbar^2}\psi = 0.$$

Although, strictly, it should be given a new symbol, we continue to denote the required solution by ψ, now taken as a function of y rather than of x.

Now if α is chosen as $(\hbar^2/mk)^{1/4}$ and E is written as $E = \beta E'$, where $\beta = \hbar(k/m)^{1/2}$, this equation becomes

$$\frac{d^2\psi}{dy^2} + (2E' - y^2)\psi = 0.$$

We note that this is a Sturm–Liouville equation with $p = 1$, $q = -y^2$, unit weight function and eigenvalue $2E'$; we therefore expect its solutions for different values of E' to be orthogonal.

To keep the notation the same as that normally used when describing numerical integration, we rewrite the equation as

$$\frac{d^2y}{dx^2} + (\lambda - x^2)y = 0. \qquad (*)$$

So that this second-order equation can be handled using an R–K routine, it has to be written as two first-order equations using an auxiliary variable. We make the simplest choice of $z \equiv dy/dx$, thus making a (two-component) 'vector' of dependent variables $(y,\ z)^{\mathrm{T}}$ with governing equations

$$\frac{dy}{dx} = z \qquad \text{and} \qquad \frac{dz}{dx} = (x^2 - \lambda)y.$$

The computer program will need to contain a subroutine that, given an input vector $(x,\ y,\ z)^{\mathrm{T}}$, returns an output vector $(dy/dx,\ dz/dx)^{\mathrm{T}}$ calculated as $(z,\ (x^2 - \lambda)y)^{\mathrm{T}}$. This is used to calculate the function $f(x_i,\ u_i)$ that appears on the (four) RHSs of, say, a fourth-order RK routine:

$$u_{i+1} = u_i + \tfrac{1}{6}(c_1 + 2c_2 + 2c_3 + c_4),$$

where

$$\begin{aligned} c_1 &= hf(x_i, u_i), \\ c_2 &= hf(x_i + \tfrac{1}{2}h,\ u_i + \tfrac{1}{2}c_1), \\ c_3 &= hf(x_i + \tfrac{1}{2}h,\ u_i + \tfrac{1}{2}c_2), \\ c_4 &= hf(x_i + h,\ u_i + c_3). \end{aligned}$$

Here, at each stage in the calculation of a one-step advance, u_i stands for y_i and z_i in turn.

Since equation $(*)$ is unchanged under the substitution $x \to -x$, and the boundary conditions, $y \to 0$ at $\pm\infty$, can be considered as both symmetric and antisymmetric, we can expect to find solutions that are either purely symmetric or purely antisymmetric. Consequently, we need only consider positive values of x, starting with $y(0) = 1$ for symmetric solutions and $y(0) = 0$ for antisymmetric ones. What will distinguish one potential solution from another is the value assigned to the initial slope $z(0)$. Clearly, one combination to be avoided is $y(0) = z(0) = 0$; such a computation will 'never get off the ground'. Intuition suggests that the initial slope should be zero if $y(0) = 1$ and non-zero if $y(0) = 0$.

As the formal range of x is infinite, we need to investigate the likely behaviour of a computed solution for large x; we want it to tend to zero for acceptable solutions. For large x equation (∗) approximates to

$$\frac{d^2y}{dx^2} = x^2 y,$$

and if we substitute a trial function $y = e^{\gamma x^n}$ with $n > 0$ we find that

$$\frac{d^2y}{dx^2} = n(n-1)\gamma x^{n-2} e^{\gamma x^n} + n^2\gamma^2 x^{2n-2} e^{\gamma x^n}.$$

The dominant term in this expression is the second one; if this is to match $x^2 y$ then $n = 2$ and $\gamma = \pm\frac{1}{2}$. The case $\gamma = -\frac{1}{2}$ is clearly the one that is required, but, inevitably, even if the appropriate eigenvalue could be hit upon exactly, rounding errors are bound to introduce some of the $\gamma = +\frac{1}{2}$ solution. Thus we could never expect a computed solution actually to tend to zero and remain close to it however many steps are taken.

A more practical way to implement the boundary condition is to require y (and hence necessarily z) to remain within some specified narrow (but empirical) band about zero over, say, the interval $5 < x < 6$ — chosen because $5^n \exp(-5^2/2)$ is less than $\sim 10^{-3}$ for any moderate value of n and we cannot hope to achieve better accuracy than one part in a thousand without using more sophisticated techniques. Thus, in practice, the integration has to be over a finite range.

A crude technique is therefore to run the integration routine from $x = 0$ up to $x = 5$ for a mesh of values for λ (≥ 0) and $z(0)$ (in the ranges discussed above) and so evaluate the solution $v(\lambda, z(0)) = y(5)$. If all v have the same sign and vary smoothly with λ and $z(0)$, then a larger range of λ is indicated. However, if the v have mixed signs, interpolated values of λ and $z(0)$ should be tried, aiming to produce $v(\lambda, z(0)) \approx 0$. When this has been achieved, the test in the previous paragraph should be implemented to give further refinement. A graphical screen display of the calculated solution would be a considerable advantage in following what is happening.

Once a value of λ that results in a solution that approaches and stays near zero over the test range has been found, the corresponding values of y need to be divided by the square root of the value of the integral $\int_{-\infty}^{\infty} y^2\,dx$, so as to normalise the solution; they can then be tabulated. The integral can be evaluated well enough using the trapezium or Simpson's rule formulae over the finite range $0 < x < 5$ and doubling the result.

In order to be reasonably certain of finding the three lowest allowed values of E, the search should start from λ (i.e. $2E$) equal to zero and incremented in amounts $\Delta\lambda$ less than, but not negligible compared with, the average values of x^2 to be covered. The latter are of order unity, and so $\Delta\lambda = 0.1$ is reasonable. The step length h in the x-variable might be chosen in the range 0.01 to 0.1 with the smaller values used when near a potential solution in the $(\lambda, z(0))$ grid.

[As has been indicated in several exercises in previous chapters, the actual eigenvalues λ are 1, 3, 5, ..., $2n+1$, ... and the corresponding solutions are $\exp(-x^2/2)$ multiplied by a Hermite polynomial $H_n(x)$.]

28

Group theory

28.1 *For each of the following sets, determine whether they form a group under the operation indicated (where it is relevant you may assume that matrix multiplication is associative):*

(a) *the integers (mod* 10*) under addition;*
(b) *the integers (mod* 10*) under multiplication;*
(c) *the integers* $1, 2, 3, 4, 5, 6$ *under multiplication (mod* 7*);*
(d) *the integers* $1, 2, 3, 4, 5$ *under multiplication (mod* 6*);*
(e) *all matrices of the form*

$$\begin{pmatrix} a & a-b \\ 0 & b \end{pmatrix},$$

where a *and* b *are integers (mod* 5*) and* $a \neq 0 \neq b$*, under matrix multiplication;*
(f) *those elements of the set in* (e) *that are of order* 1 *or* 2 *(taken together);*
(g) *all matrices of the form*

$$\begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & c & 1 \end{pmatrix},$$

where a*,* b*,* c *are integers, under matrix multiplication.*

In all cases we need to establish whether the prescribed combination law is associative and whether, under it, the set possesses the properties of (i) closure, (ii) having an identity element and (iii) containing an inverse for every element present. If any one of these conditions fails, the set cannot form a group under the given law.

(a) Addition is associative and the set $\{0, 1, 2, \ldots, 9\}$ is closed under addition (mod 10), e.g. $7 + 6 = 3$. The identity is 0 and every element has an inverse, e.g. $(7)^{-1} = 3$. The set does form a group.

(b) For the set $\{0, 1, 2, \ldots, 9\}$ under multiplication the identity can only be 1. However, for any element $X \neq 1$ the set does *not* contain an inverse Y such that $XY = 1$. As a specific example, if $X = 2$ then 0.5 would need to be in the set – but it is not. The set does *not* form a group under multiplication.

(c) Multiplication is associative and the group table would be as below. The entries are calculated by expressing each product modulo 7. For example, $4 \times 5 = 20 = (2 \times 7) + 6 = 6 \pmod 7$

	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	1	3	5
3	3	6	2	5	1	4
4	4	1	5	2	6	3
5	5	3	1	6	4	2
6	6	5	4	3	2	1

This demonstrates (i) closure, (ii) the existence of an identity element (1) and (iii) an inverse for each element (1 appears in every row). The set does form a group.

(d) The set is not closed under multiplication (mod 6) and *cannot* form a group. For example, $2 \times 3 = 0 \pmod 6$ and 0 is not in the given set.

(e) With the associativity of matrix multiplication assumed and $a = b = 1$ yielding a unit element, consider

$$\begin{pmatrix} a & a-b \\ 0 & b \end{pmatrix}\begin{pmatrix} c & c-d \\ 0 & d \end{pmatrix} = \begin{pmatrix} ac & ac - ad + ad - bd \\ 0 & bd \end{pmatrix}$$
$$= \begin{pmatrix} ac & ac - bd \\ 0 & bd \end{pmatrix},$$

implying closure. We also note that interchanging a with c and b with d shows that any two matrices in the set commute.

Since neither a nor b is 0, the determinant of a general matrix in the set is non-zero and its inverse can be constructed as

$$\frac{1}{ab}\begin{pmatrix} b & b-a \\ 0 & a \end{pmatrix} = \begin{pmatrix} a^{-1} & a^{-1} - b^{-1} \\ 0 & b^{-1} \end{pmatrix}.$$

The question then arises as to whether a^{-1} is an integer; in multiplication mod 5 it is. For example, if $a = 3$ then $a^{-1} = 2$ since $3 \times 2 = 6 = 1 \pmod 5$. The full set of values is: $1^{-1} = 1$, $2^{-1} = 3$, $3^{-1} = 2$ and $4^{-1} = 4$.

Thus each inverse is of the required form and a general one can be verified:

$$\frac{1}{ab}\begin{pmatrix} b & b-a \\ 0 & a \end{pmatrix}\begin{pmatrix} a & a-b \\ 0 & b \end{pmatrix} = \frac{1}{ab}\begin{pmatrix} ab & ab-b^2+b^2-ab \\ 0 & ab \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

All four requirements are satisfied and the set of matrices is, in fact, a group.

(f) As always, the only element of order 1 is the unit element. Elements of order 2 must satisfy

$$\begin{pmatrix} a & a-b \\ 0 & b \end{pmatrix}\begin{pmatrix} a & a-b \\ 0 & b \end{pmatrix} = \begin{pmatrix} a^2 & a^2-b^2 \\ 0 & b^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Thus a and b must both be elements whose squares are unity (mod 5); each must be either 1 or 4 [since $4^2 = 16 = 1$ (mod 5)]. The four matrices to consider are thus

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 \\ 0 & 4 \end{pmatrix}, \quad \begin{pmatrix} 4 & 3 \\ 0 & 1 \end{pmatrix}.$$

The identity element is present and, from the way they were defined, each is its own inverse. Only closure remains to be tested. As all matrices in the set commute [see (e) above], we need test only

$$\begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ 0 & 1 \end{pmatrix},$$
$$\begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}\begin{pmatrix} 4 & 3 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 4 \end{pmatrix},$$
$$\begin{pmatrix} 1 & 2 \\ 0 & 4 \end{pmatrix}\begin{pmatrix} 4 & 3 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}.$$

Each product is one of the set of four. So closure is established and the set does form a group – a subgroup, of order 4, of the group in (e).

(g) The product of two such matrices is

$$\begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & c & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ x & 1 & 0 \\ y & z & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ a+x & 1 & 0 \\ b+cx+y & c+z & 1 \end{pmatrix}.$$

Since all elements of the original two matrices are integers, so are all elements of the product and closure is established.

Clearly, $a = b = c = 0$ provides the identity element and, since the determinant of each matrix is 1, inverses can be constructed in the usual way, typically

$$\begin{pmatrix} 1 & 0 & 0 \\ -a & 1 & 0 \\ ac-b & -c & 1 \end{pmatrix}.$$

This is of the correct form as can be verified as follows:

$$\begin{pmatrix} 1 & 0 & 0 \\ -a & 1 & 0 \\ ac-b & -c & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & c & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Thus, assuming associativity, the group property of the set is established.

28.3 *Define a binary operation $\bullet$ on the set of real numbers by*

$$x \bullet y = x + y + rxy,$$

where r is a non-zero real number. Show that the operation $\bullet$ is associative.

Prove that $x \bullet y = -r^{-1}$ if, and only if, $x = -r^{-1}$ or $y = -r^{-1}$. Hence prove that the set of all real numbers excluding $-r^{-1}$ forms a group under the operation $\bullet$.

To demonstrate the associativity we need to show that $x \bullet (y \bullet z)$ is the same thing as $(x \bullet y) \bullet z$. So consider

$$\begin{aligned} x \bullet (y \bullet z) &= x + (y \bullet z) + rx(y \bullet z) \\ &= x + y + z + ryz + rx(y + z + ryz) \\ &= x + y + z + r(yz + xy + xz) + r^2xyz \end{aligned}$$

and

$$\begin{aligned} (x \bullet y) \bullet z &= (x \bullet y) + z + r(x \bullet y)z \\ &= x + y + rxy + z + r(x + y + rxy)z \\ &= x + y + z + r(xy + xz + yz) + r^2xyz. \end{aligned}$$

The two RHSs are equal, showing that the operation $\bullet$ is associative.

Firstly, suppose that $x = -r^{-1}$. Then

$$x \bullet y = -\frac{1}{r} + y + r\left(-\frac{1}{r}\right) y = -\frac{1}{r} + y - y = -\frac{1}{r}.$$

Similarly $y = -r^{-1} \quad \Rightarrow \quad x \bullet y = -r^{-1}$.

Secondly, suppose that $x \bullet y = -r^{-1}$. Then

$$\begin{aligned} x + y + rxy &= -r^{-1}, \\ rx + ry + r^2xy + 1 &= 0, \\ (rx + 1)(ry + 1) &= 0, \\ \Rightarrow \quad \text{either } x = -r^{-1} &\text{ or } y = -r^{-1}. \end{aligned}$$

Thus $x \bullet y = -r^{-1} \iff (x = -r^{-1} \text{ or } y = -r^{-1})$.

If $\mathcal{S} = \{\text{real numbers } \neq -r^{-1}\}$, then

(i) Associativity under $\bullet$ has already been shown.
(ii) If x and y belong to $\mathcal{S}$, then $x \bullet y$ is a real number and, in view of the second result above, is $\neq -r^{-1}$. Thus $x \bullet y$ belongs to $\mathcal{S}$ and the set is closed under the operation $\bullet$.
(iii) For any x belonging to $\mathcal{S}$, $x \bullet 0 = x + 0 + rx0 = x$. Thus 0 is an identity element.
(iv) An inverse x^{-1} of x must satisfy $x \bullet x^{-1} = 0$, i.e.

$$x + x^{-1} + rxx^{-1} = 0 \quad \Rightarrow \quad x^{-1} = -\frac{x}{1 + rx}.$$

This is a real (finite) number since $x \neq -r^{-1}$ and, further, $x^{-1} \neq -r^{-1}$, since if it were we could deduce that $1 = 0$. Thus the set $\mathcal{S}$ contains an inverse for each of its elements.

These four results together show that $\mathcal{S}$ is a group under the operation $\bullet$.

28.5 *The following is a 'proof' that reflexivity is an unnecessary axiom for an equivalence relation.*

Because of symmetry $X \sim Y$ implies $Y \sim X$. Then by transitivity $X \sim Y$ and $Y \sim X$ imply $X \sim X$. Thus symmetry and transitivity imply reflexivity, which therefore need not be separately required.

Demonstrate the flaw in this proof using the set consisting of all real numbers plus the number i. Show by investigating the following specific cases that, whether or not reflexivity actually holds, it cannot be deduced from symmetry and transitivity alone.

(a) $X \sim Y$ *if* $X + Y$ *is real.*
(b) $X \sim Y$ *if* XY *is real.*

Let elements X and Y be drawn from the set $\mathcal{S}$ consisting of the real numbers together with i.

(a) For the definition $X \sim Y$ if $X + Y$ is real, we have

(i) that

$$X \sim Y \quad \Rightarrow \quad X + Y \text{ is real} \quad \Rightarrow \quad Y + X \text{ is real} \quad \Rightarrow \quad Y \sim X,$$

i.e symmetry holds;

(ii) that if $X \sim Y$ then neither X nor Y can be i and, equally, if $Y \sim Z$ then neither Y nor Z can be i. It then follows that $X + Z$ is real and $X \sim Z$, i.e. transitivity holds.

Thus both symmetry and transitivity hold, though it is obvious that $X \nsim X$ if X is i. Thus symmetry and transitivity do *not* necessarily imply reflexivity, showing that the 'proof' is flawed – in this case the proof fails when X is i because there is no distinct 'Y' available, something assumed in the proof.

(b) For the definition $X \sim Y$ if XY is real, we have

(i) that

$$X \sim Y \quad \Rightarrow \quad XY \text{ is real} \quad \Rightarrow \quad YX \text{ is real} \quad \Rightarrow \quad Y \sim X,$$

i.e symmetry holds;

(ii) that if $X \sim Y$ then neither X nor Y is i. Equally, if $Y \sim Z$ then neither Y nor Z is i. It then follows that XZ is real and $X \sim Z$, i.e. transitivity holds.

Thus both symmetry and transitivity hold and, setting Z equal to X, they do imply the reflexivity property for the real elements of $\mathcal{S}$. However, they cannot establish it for the element i – though it happens to be true in this case as i^2 is real.

> ***28.7*** *$\mathcal{S}$ is the set of all 2×2 matrices of the form*
>
> $$A = \begin{pmatrix} w & x \\ y & z \end{pmatrix}, \qquad \text{where } wz - xy = 1.$$
>
> *Show that $\mathcal{S}$ is a group under matrix multiplication. Which element(s) have order 2? Prove that an element A has order 3 if $w + z + 1 = 0$.*

The condition $wz - xy = 1$ is the same as $|\mathsf{A}| = 1$; it follows that the set contains an identity element (with $w = z = 1$ and $x = y = 0$). Moreover, each matrix in $\mathcal{S}$ has an inverse and, since $|\mathsf{A}^{-1}|\,|\mathsf{A}| = |\mathsf{I}| = 1$ implies that $|\mathsf{A}^{-1}| = 1$, the inverses also belong to the set.

If A and B belong to $\mathcal{S}$ then, since $|\mathsf{AB}| = |\mathsf{A}|\,|\mathsf{B}| = 1 \times 1 = 1$, their product also belongs to $\mathcal{S}$, i.e. the set is closed.

These observations, together with the associativity of matrix multiplication establish that the set $\mathcal{S}$ is, in fact, a group under this operation.

If A is to have order 2 then

$$\begin{pmatrix} w & x \\ y & z \end{pmatrix}\begin{pmatrix} w & x \\ y & z \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

i.e. $w^2 + xy = 1, \quad x(w+z) = 0, \quad y(w+z) = 0, \quad xy + z^2 = 1.$

These imply that $w^2 = z^2$ and that either $z = -w$ or $x = y = 0$. If $z = -w$, then both

$$\begin{aligned} w^2 + xy &= 1, &&\text{from the above condition,} \\ \text{and} \quad -w^2 - xy &= 1, &&\text{from } wz - xy = 1. \end{aligned}$$

This is not possible and so we must have $x = y = 0$, implying that w and z are either both $+1$ or both -1. The former gives the identity (of order 1), and so the matrix given by the latter, $\mathsf{A} = -\mathsf{I}$, is the only element in $\mathcal{S}$ of order 2.

If $w + z + 1 = 0$ (as well as $xy = wz - 1$), A^2 can be written as

$$\begin{aligned} \mathsf{A}^2 &= \begin{pmatrix} w^2 + xy & x(w+z) \\ y(w+z) & xy + z^2 \end{pmatrix} \\ &= \begin{pmatrix} w^2 + wz - 1 & -x \\ -y & wz - 1 + z^2 \end{pmatrix} \\ &= \begin{pmatrix} -w - 1 & -x \\ -y & -z - 1 \end{pmatrix}. \end{aligned}$$

Multiplying again by A gives

$$\begin{aligned} \mathsf{A}^3 &= \begin{pmatrix} -w - 1 & -x \\ -y & -z - 1 \end{pmatrix} \begin{pmatrix} w & x \\ y & z \end{pmatrix} \\ &= -\begin{pmatrix} w(w+1) + xy & (w+1)x + xz \\ wy + y(z+1) & xy + z(z+1) \end{pmatrix} \\ &= -\begin{pmatrix} w(w+1) + wz - 1 & x \times 0 \\ y \times 0 & wz - 1 + z(z+1) \end{pmatrix} \\ &= -\begin{pmatrix} (w \times 0) - 1 & 0 \\ 0 & (z \times 0) - 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Thus A has order 3.

28.9 *If $\mathcal{A}$ is a group in which every element other than the identity, I, has order 2, prove that $\mathcal{A}$ is Abelian. Hence show that if X and Y are distinct elements of $\mathcal{A}$, neither being equal to the identity, then the set $\{I, X, Y, XY\}$ forms a subgroup of $\mathcal{A}$.*

Deduce that if $\mathcal{B}$ is a group of order $2p$, with p a prime greater than 2, then $\mathcal{B}$ must contain an element of order p.

If every element of $\mathcal{A}$, apart from I, has order 2, then, for any element X, $X^2 = I$. Consider two elements X and Y and let $XY = Z$. Then

$$\begin{aligned} X^2 Y = XZ \quad &\Rightarrow \quad Y = XZ, \\ XY^2 = ZY \quad &\Rightarrow \quad X = ZY. \end{aligned}$$

It follows that $YX = XZZY$. But, since $XY = Z$, Z must belong to $\mathcal{A}$ and therefore $Z^2 = I$. Substituting this gives $YX = XY$, proving that the group is Abelian.

Consider the set $\mathcal{S} = \{I, X, Y, XY\}$, for which

(i) Associativity holds, since it does for $\mathcal{A}$.
(ii) It is closed, the only products needing non-trivial examinations being $XY\,X = X\,XY = X^2Y = Y$ and $Y\,XY = XY\,Y = XY^2 = X$ (here we have twice used the fact that $\mathcal{A}$, and hence $\mathcal{S}$, is Abelian).
(iii) The identity I is contained in the set.
(iv) Since all elements are of order 2 (or 1), each is its own inverse.

Thus the set is a subgroup of $\mathcal{A}$ of order 4.

Now consider the group $\mathcal{B}$ of order $2p$, where p is prime. Since the order of an element must divide the order of the group, all elements in $\mathcal{B}$ must have order 1 (I only) or 2 or p.

Suppose *all* elements, other than I, have order 2. Then, as shown above, $\mathcal{B}$ must be Abelian and have a subgroup of order 4. However, the order of any subgroup must divide the order of the group and 4 cannot divide $2p$ since p is prime. It follows that the supposition that all elements can be of order 2 is false, and consequently at least one must have order p.

28.11 *Identify the eight symmetry operations on a square. Show that they form a group $\mathcal{D}_4$ (known to crystallographers as 4mm and to chemists as $\mathcal{C}_{4v}$) having one element of order 1, five of order 2 and two of order 4. Find its proper subgroups and the corresponding cosets.*

The operation of leaving the square alone is a trivial symmetry operation, but an important one, as it is the identity I of the group; it has order 1.

Rotations about an axis perpendicular to the plane of the square by $\pi/4$, $\pi/2$ and $3\pi/2$ each take the square into itself. The first and last of these have to be repeated four times to reproduce the effect of I, and so they have order 4. The rotation by $\pi/2$ clearly has order 2.

Reflections in the two axes parallel to the sides of the square and passing through its centre are also symmetry operations, as are reflections in the two principal diagonals of the square; all of these reflections have order 2.

Using the notation indicated in figure 28.1, R being a rotation of $\pi/2$ about an axis perpendicular to the square, we have: I has order 1; R^2, m_1, m_2, m_3, m_4 have order 2; R, R^3 have order 4.

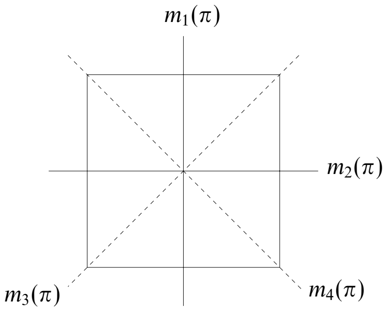

Figure 28.1 The notation for exercise 28.11.

The group multiplication table takes the form

	I	R	R^2	R^3	m_1	m_2	m_3	m_4
I	I	R	R^2	R^3	m_1	m_2	m_3	m_4
R	R	R^2	R^3	I	m_4	m_3	m_1	m_2
R^2	R^2	R^3	I	R	m_2	m_1	m_4	m_3
R^3	R^3	I	R	R^2	m_3	m_4	m_2	m_1
m_1	m_1	m_3	m_2	m_4	I	R^2	R	R^3
m_2	m_2	m_4	m_1	m_3	R^2	I	R^3	R
m_3	m_3	m_2	m_4	m_1	R^3	R	I	R^2
m_4	m_4	m_1	m_3	m_2	R	R^3	R^2	I

Inspection of this table shows the existence of the non-trivial subgroups listed below, and tedious but straightforward evaluation of the products of selected elements of the group with all the elements of any one subgroup provides the cosets of that subgroup. The results are as follows:

subgroup $\{I, R, R^2, R^3\}$ has cosets $\{I, R, R^2, R^3\}$, $\{m_1, m_2, m_3, m_4\}$;

subgroup $\{I, R^2, m_1, m_2\}$ has cosets $\{I, R^2, m_1, m_2\}$, $\{I, R^2, m_3, m_4\}$;

subgroup $\{I, R^2, m_3, m_4\}$ has cosets $\{I, R^2, m_3, m_4\}$, $\{I, R^2, m_1, m_2\}$;

subgroup $\{I, R^2\}$ has cosets $\{I, R^2\}$, $\{R, R^3\}$, $\{m_1, m_2\}$, $\{m_3, m_4\}$;

subgroup $\{I, m_1\}$ has cosets $\{I, m_1\}$, $\{R, m_3\}$, $\{R^2, m_2\}$, $\{R^3, m_4\}$;

subgroup $\{I, m_2\}$ has cosets $\{I, m_2\}$, $\{R, m_4\}$, $\{R^2, m_1\}$, $\{R^3, m_3\}$;

subgroup $\{I, m_3\}$ has cosets $\{I, m_3\}$, $\{R, m_2\}$, $\{R^2, m_4\}$, $\{R^3, m_1\}$;

subgroup $\{I, m_4\}$ has cosets $\{I, m_4\}$, $\{R, m_1\}$, $\{R^2, m_3\}$, $\{R^3, m_2\}$.

> **28.13** *Find the group $\mathcal{G}$ generated under matrix multiplication by the matrices*
> $$\mathsf{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \mathsf{B} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$
> *Determine its proper subgroups, and verify for each of them that its cosets exhaust $\mathcal{G}$.*

Before we can draw up a group multiplication table to search for subgroups, we must determine the multiple products of A and B with themselves and with each other:

$$\mathsf{A}^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathsf{I}.$$

Since $\mathsf{B} = i\mathsf{A}$, it follows that $\mathsf{B}^2 = -I$, that $\mathsf{AB} = i\mathsf{I} = \mathsf{BA}$, and that $\mathsf{B}^3 = -\mathsf{B}$. In brief, A is of order 2, B is of order 4, and A and B commute. The eight distinct elements of the group are therefore: $\mathsf{I}, \mathsf{A}, \mathsf{B}, \mathsf{B}^2, \mathsf{B}^3, \mathsf{AB}, \mathsf{AB}^2$ and AB^3.

The group multiplication table is

	I	A	B	B^2	B^3	AB	AB^2	AB^3
I	I	A	B	B^2	B^3	AB	AB^2	AB^3
A	A	I	AB	AB^2	AB^3	B	B^2	B^3
B	B	AB	B^2	B^3	I	AB^2	AB^3	A
B^2	B^2	AB^2	B^3	I	B	AB^3	A	AB
B^3	B^3	AB^3	I	B	B^2	A	AB	AB^2
AB	AB	B	AB^2	AB^3	A	B^2	B^3	I
AB^2	AB^2	B^2	AB^3	A	B	B^3	I	B
AB^3	AB^3	B^3	A	AB	AB^2	I	B	B^2

By inspection, the subgroups and their cosets are as follows:

$$\begin{aligned}
\{\mathsf{I}, \mathsf{A}\} &: \{\mathsf{I}, \mathsf{A}\}, \{\mathsf{B}, \mathsf{AB}\}, \{\mathsf{B}^2, \mathsf{AB}^2\}, \{\mathsf{B}^3, \mathsf{AB}^3\}; \\
\{\mathsf{I}, \mathsf{B}^2\} &: \{\mathsf{I}, \mathsf{B}^2\}, \{\mathsf{A}, \mathsf{AB}^2\}, \{\mathsf{B}, \mathsf{B}^3\}, \{\mathsf{AB}, \mathsf{AB}^3\}; \\
\{\mathsf{I}, \mathsf{AB}^2\} &: \{\mathsf{I}, \mathsf{AB}^2\}, \{\mathsf{A}, \mathsf{B}^2\}, \{\mathsf{B}, \mathsf{AB}^3\}, \{\mathsf{B}^3, \mathsf{AB}\}; \\
\{\mathsf{I}, \mathsf{B}, \mathsf{B}^2, \mathsf{B}^3\} &: \{\mathsf{I}, \mathsf{B}, \mathsf{B}^2, \mathsf{B}^3\}, \{\mathsf{A}, \mathsf{AB}, \mathsf{AB}^2, \mathsf{AB}^3\}; \\
\{\mathsf{I}, \mathsf{AB}, \mathsf{B}^2, \mathsf{AB}^3\} &: \{\mathsf{I}, \mathsf{AB}, \mathsf{B}^2, \mathsf{AB}^3\}, \{\mathsf{A}, \mathsf{B}, \mathsf{AB}^2, \mathsf{B}^3\}.
\end{aligned}$$

As expected, in each case the cosets exhaust the group, with each element in one and only one coset.

28.15 *Consider the following mappings between a permutation group and a cyclic group.*

(a) *Denote by A_n the subset of the permutation group S_n that contains all the even permutations. Show that A_n is a subgroup of S_n.*
(b) *List the elements of S_3 in cycle notation and identify the subgroup A_3.*
(c) *For each element X of S_3, let $p(X) = 1$ if X belongs to A_3 and $p(X) = -1$ if it does not. Denote by $\mathcal{C}_2$ the multiplicative cyclic group of order 2. Determine the images of each of the elements of S_3 for the following four mappings:*

$$\begin{aligned} \Phi_1 &: S_3 \to \mathcal{C}_2 \qquad X \to p(X), \\ \Phi_2 &: S_3 \to \mathcal{C}_2 \qquad X \to -p(X), \\ \Phi_3 &: S_3 \to A_3 \qquad X \to X^2, \\ \Phi_4 &: S_3 \to S_3 \qquad X \to X^3. \end{aligned}$$

(d) *For each mapping, determine whether the kernel $\mathcal{K}$ is a subgroup of S_3 and, if so, whether the mapping is a homomorphism.*

(a) With A_n as the subset of S_n that contains all the even permutations, we need to demonstrate that it has the four properties that characterise a group:

(i) If X and Y belong to A_n so does XY, as the product of two even permutations is even. This establishes closure.
(ii) From the definition of an even permutation, the identity I belongs to A_n.
(iii) If X belongs to A_n so does X^{-1}, as the number of pair interchanges needed to change from X to I is the same as the number needed to go in the opposite direction. This establishes the existence, within the subset, of an inverse for each member of the subset.
(iv) Associativity follows from that of the group S_n.

Thus A_n does possess the four properties and is a subgroup of S_n.

(b) (1), (123) and (132) belong to A_3. The permutations (12), (13) and (23) do not belong, as each involves only one pair interchange.

(c) With the given definition of $p(X)$,

$$\begin{aligned} p(X) &= 1 \text{ for } X = (1),\ (123),\ (132), \\ p(X) &= -1 \text{ for } X = (12),\ (13),\ (23). \end{aligned}$$

$\mathcal{C}_2$ consists of the two elements $+1$ and -1.

For $\Phi_1 : S_3 \to \mathcal{C}_2,\ X \to p(X)$, elements in A_3 have image $+1$; those not in A_3 have image -1.

For $\Phi_2 : S_3 \to \mathcal{C}_2,\ X \to -p(X)$, elements in A_3 have image -1; those not in A_3 have image $+1$.

For $\Phi_3 : S_3 \to A_3,\ X \to X^2$

$$\begin{aligned}(1) &\to (1)(1) = (1),\\ (123) &\to (123)(123) = (132),\\ (132) &\to (132)(132) = (123),\\ (12) &\to (12)(12) = (1), \quad \text{similarly, (13) and (23).}\end{aligned}$$

For $\Phi_4 : S_3 \to S_3,\ X \to X^3$

$$\begin{aligned}(1) &\to (1)(1) = (1),\\ (123) &\to (123)(123)(123) = (132)(123) = (1),\\ (132) &\to (132)(132)(132) = (123)(132) = (1),\\ (12) &\to (12)(12)(12) = (1)(12) = (12), \quad \text{similarly, (13) and (23).}\end{aligned}$$

(d) For Φ_1, the kernel is the set of elements belonging to A_3 and, as already shown, this is a subgroup of S_3.

We note that the product of two even or two odd permutations is an even permutation, whilst the product of an odd and an even permutation is an odd permutation. We also note that $+1\times+1$ and -1×-1 are both equal to $+1$, whilst $+1\times-1=-1$. Since Φ_1 maps even permutations onto $+1$ and odd permutations onto -1, the preceding observations imply that Φ_1 is a homomorphism.

For Φ_2, the kernel is the set of elements *not* belonging to A_3. Since this set does not contain the identity (1), it cannot be a subgroup of S_3.

For Φ_3 the kernel is the set $\{(1),(12),(13),(23)\}$. Since, for example, $(12)(13) = (132)$, the set is not closed and so cannot form a group. It cannot, therefore, be a subgroup of S_3.

For Φ_4 the kernel is the set $\{(1),(123),(132)\}$, i.e. the subgroup A_3. However, for example, $[(12)(13)]' = (132)' = (1)$, whilst $(12)'(13)' = (12)(13) = (132)$; these two results are not equal, showing that the mapping *cannot* be a homomorphism.

28.17 *The group of all non-singular $n\times n$ matrices is known as the general linear group $GL(n)$ and that with only real elements as $GL(n,\mathbf{R})$. If $\mathbf{R}^*$ denotes the multiplicative group of non-zero real numbers, prove that the mapping $\Phi : GL(n,\mathbf{R}) \to \mathbf{R}^*$, defined by $\Phi(\mathsf{M}) = \det \mathsf{M}$, is a homomorphism.*

Show that the kernel $\mathcal{K}$ of Φ is a subgroup of $GL(n,\mathbf{R})$. Determine its cosets and show that they themselves form a group.

If P and Q are two matrices belonging to $GL(n,\mathbf{R})$ then, under Φ,

$$(\mathsf{PQ})' = |\mathsf{PQ}| = |\mathsf{P}|\,|\mathsf{Q}| = \mathsf{P}'\mathsf{Q}'.$$

Thus Φ is a homomorphism.

The kernel $\mathcal{K}$ of the mapping consists of all matrices in $GL(n, \mathbf{R})$ that map onto the identity in $\mathbf{R}^*$, i.e all matrices whose determinant is 1.

To determine whether $\mathcal{K}$ is a subgroup of the general linear group, let X and Y belong to $\mathcal{K}$. Then, testing $\mathcal{K}$ for the four group-defining properties, we have

(i) $(XY)' = X'Y' = 1 \times 1 = 1$, i.e. XY also belongs to $\mathcal{K}$, showing the closure of the kernel.
(ii) The associative law holds for the elements of $\mathcal{K}$ since it does so for all elements of $GL(n, \mathbf{R})$.
(iii) $|\mathsf{I}| = 1$ and so I belongs to $\mathcal{K}$.
(iv) Since $\mathsf{X}^{-1}\mathsf{X} = \mathsf{I}$, it follows that $|\mathsf{X}^{-1}|\,|\mathsf{X}| = |\mathsf{I}|$ and $|\mathsf{X}^{-1}| \times 1 = 1$. Hence $|\mathsf{X}^{-1}| = 1$ and so X^{-1} also belongs to $\mathcal{K}$.

This completes the proof that $\mathcal{K}$ is a subgroup of $GL(n, \mathbf{R})$.

Two matrices P and Q in $GL(n, \mathbf{R})$ belong to the same coset of $\mathcal{K}$ if

$$\mathsf{Q} = \mathsf{PX}, \text{ where } \mathsf{X} \text{ is some element in } \mathcal{K}.$$

It then follows that

$$|\mathsf{Q}| = |\mathsf{P}|\,|\mathsf{X}| = |\mathsf{P}| \times 1.$$

Thus the requirement for two matrices to be in the same coset is that they have equal determinants.

Let us denote by $\mathcal{C}_i$ the (infinite) set of all matrices whose determinant has the value i; the label i will itself take on an infinite continuum of values, excluding 0. Then,

(i) For *any* $\mathsf{M}_i \in \mathcal{C}_i$ and *any* $\mathsf{M}_j \in \mathcal{C}_j$ we have

$$|\mathsf{M}_i\mathsf{M}_j| = |\mathsf{M}_i|\,|\mathsf{M}_j| = i \times j,$$

implying that we always have $\mathsf{M}_i\mathsf{M}_j \in \mathcal{C}_{(i\times j)}$. Thus the set of cosets is closed, with $\mathcal{C}_i \times \mathcal{C}_j = \mathcal{C}_{(i\times j)}$.
(ii) The associative law holds, since it does so for matrix multiplication in general, and the product of three matrices, and hence its determinant, is independent of the order in which the individual multiplications are carried out.
(iii) Since

$$|\mathsf{M}_i\mathsf{M}_1| = |\mathsf{M}_i|\,|\mathsf{M}_1| = i,$$

$\mathcal{C}_i \times \mathcal{C}_1 = \mathcal{C}_i$, showing that $\mathcal{C}_1$ is an identity element in the set.
(iv) Since

$$|\mathsf{M}_i\mathsf{M}_{1/i}| = |\mathsf{M}_i|\,|\mathsf{M}_{1/i}| = i \times (1/i) = 1,$$

$\mathcal{C}_i \times \mathcal{C}_{1/i} = \mathcal{C}_1$, showing that the set of cosets contains an inverse for each coset.

This completes the proof that the cosets themselves form a group under coset multiplication (and also that $\mathcal{K}$ is a normal subgroup).

28.19 *Given that matrix* M *is a member of the multiplicative group* $GL(3,\mathbf{R})$, *determine, for each of the following additional constraints on* M *(applied separately), whether the subset satisfying the constraint is a subgroup of* $GL(3,\mathbf{R})$:

(a) $\mathsf{M}^T = \mathsf{M}$;
(b) $\mathsf{M}^T\mathsf{M} = \mathsf{I}$;
(c) $|\mathsf{M}| = 1$;
(d) $M_{ij} = 0$ *for* $j > i$ *and* $M_{ii} \neq 0$.

The matrices belonging to $GL(3,\mathbf{R})$ have the general properties that they are non-singular, possess inverses and have real elements. The operation of matrix multiplication is associative, and this will be assumed in the rest of the exercise, in which A and B are general matrices satisfying the various defining constraints.

(a) $\mathsf{M}^T = \mathsf{M}$: the set of symmetric matrices.
Now, for two symmetric matrices A and B,

$$(\mathsf{AB})^{\mathrm{T}} = \mathsf{B}^{\mathrm{T}}\mathsf{A}^{\mathrm{T}} = \mathsf{BA}$$

and this is not equal to AB in general, as matrix multiplication is not necessarily commutative. The set is therefore not closed and cannot form a group; equally it cannot be a subgroup of $GL(3,\mathbf{R})$.

(b) $\mathsf{M}^T\mathsf{M} = \mathsf{I}$: the set of orthogonal matrices.
Clearly, the identity I belongs to the set, and furthermore

$$(\mathsf{M}^{-1})^{\mathrm{T}}\mathsf{M}^{-1} = (\mathsf{M}^{\mathrm{T}})^{-1}\mathsf{M}^{-1} = (\mathsf{M}^{-1})^{-1}\mathsf{M}^{-1} = \mathsf{M}\mathsf{M}^{-1} = \mathsf{I},$$

i.e. if M belongs to the set then so does M^{-1}. Finally,

$$(\mathsf{AB})^{\mathrm{T}}\mathsf{AB} = \mathsf{B}^{\mathrm{T}}\mathsf{A}^{\mathrm{T}}\mathsf{AB} = \mathsf{B}^{\mathrm{T}}\mathsf{I}\,\mathsf{B} = \mathsf{I},$$

showing that the set is closed. This completes the proof that the orthogonal matrices form a subgroup of $GL(3,\mathbf{R})$.

(c) $|\mathsf{M}| = 1$: the set of unimodular matrices.

$$\begin{aligned} |\mathsf{AB}| = |\mathsf{A}|\,|\mathsf{B}| = 1 \times 1 &= 1 \quad \text{closure,} \\ |\,\mathsf{I}\,| &= 1 \quad \text{identity,} \\ \mathsf{M}^{-1}\mathsf{M} = \mathsf{I} \quad \Rightarrow \quad |\mathsf{M}^{-1}|\,|\mathsf{M}| = 1 \quad \Rightarrow \quad |\mathsf{M}^{-1}| &= 1 \quad \text{inverse.} \end{aligned}$$

These three results (and associativity) show that the unimodular matrices do form a subgroup of $GL(3,\mathbf{R})$.

(d) $M_{ij} = 0$ for $j > i$ and $M_{ii} \neq 0$: the set of lower diagonal matrices with non-zero diagonal elements.
Taking first the question of closure, consider the matrix product $\mathsf{C} = \mathsf{AB}$. A typical element of C above the leading diagonal is

$$\begin{aligned} C_{12} &= A_{11}B_{12} + A_{12}B_{22} + A_{13}B_{32} \\ &= A_{11}\,0 + 0\,B_{22} + 0\,A_{32} = 0, \end{aligned}$$

and a typical element on the leading diagonal is

$$\begin{aligned} C_{11} &= A_{11}B_{11} + A_{12}B_{21} + A_{13}B_{31} \\ &= A_{11}B_{11} + 0\,B_{21} + 0\,A_{31} = A_{11}B_{11} \neq 0. \end{aligned}$$

That C_{11} is not equal to zero follows from the fact that neither A_{11} nor B_{11} is zero. Similarly, C_{13} and C_{23} are zero, whilst C_{22} and C_{33} are non-zero. Thus C has all the properties defining the set, and so belongs to it. The set is therefore closed.

Clearly the matrix I_3 belongs to the set, which therefore contains an identity element.

Since no diagonal element is zero, the determinant (which is the product of the diagonal elements for lower triangular matrices) of any member of the set cannot be zero. All members must therefore have inverses. For the matrix

$$\mathsf{A} = \begin{pmatrix} A_{11} & 0 & 0 \\ A_{21} & A_{22} & 0 \\ A_{31} & A_{32} & A_{33} \end{pmatrix},$$

the inverse is given by

$$\mathsf{A}^{-1} = \frac{1}{A_{11}A_{22}A_{33}} \begin{pmatrix} A_{22}A_{33} & 0 & 0 \\ -A_{21}A_{33} & A_{11}A_{33} & 0 \\ A_{21}A_{32} - A_{22}A_{31} & -A_{11}A_{32} & A_{11}A_{22} \end{pmatrix}.$$

We note that each of the diagonal elements of A^{-1} is the product of two non-zero terms, and is therefore itself non-zero. Thus A^{-1} has the correct form for a member of the set – lower diagonal with non-zero diagonal elements – and so belongs to the set, which has now been shown to have all the properties needed to make it a group and hence a subgroup of $GL(3, \mathbf{R})$.

28.21 *Show that $\mathcal{D}_4$, the group of symmetries of a square, has two isomorphic subgroups of order* 4.

The quaternion group Q is the set of elements

$$\{1, -1, i, -i, j, -j, k, -k\},$$

with $i^2 = j^2 = k^2 = -1$, $ij = k$ and its cyclic permutations, and $ji = -k$ and its cyclic permutations. Its multiplication table reads as follows:

	1	-1	i	$-i$	j	$-j$	k	$-k$
1	1	-1	i	$-i$	j	$-j$	k	$-k$
-1	-1	1	$-i$	i	$-j$	j	$-k$	k
i	i	$-i$	-1	1	k	$-k$	$-j$	j
$-i$	$-i$	i	1	-1	$-k$	k	j	$-j$
j	j	$-j$	$-k$	k	-1	1	i	$-i$
$-j$	$-j$	j	k	$-k$	1	-1	$-i$	i
k	k	$-k$	j	$-j$	$-i$	i	-1	1
$-k$	$-k$	k	$-j$	j	i	$-i$	1	-1

Show that there exists a two-to-one homomorphism from the quaternion group Q onto one (and hence either) of the two subgroups of $\mathcal{D}_4$, and determine its kernel.

We first reproduce the multiplication table for $\mathcal{D}_4$:

	I	R	R^2	R^3	m_1	m_2	m_3	m_4
I	I	R	R^2	R^3	m_1	m_2	m_3	m_4
R	R	R^2	R^3	I	m_4	m_3	m_1	m_2
R^2	R^2	R^3	I	R	m_2	m_1	m_4	m_3
R^3	R^3	I	R	R^2	m_3	m_4	m_2	m_1
m_1	m_1	m_3	m_2	m_4	I	R^2	R	R^3
m_2	m_2	m_4	m_1	m_3	R^2	I	R^3	R
m_3	m_3	m_2	m_4	m_1	R^3	R	I	R^2
m_4	m_4	m_1	m_3	m_2	R	R^3	R^2	I

Here R is a rotation by $\pi/2$ in the plane of the square, m_1 and m_2 are reflections in the axes parallel to the sides of the square, and m_3 and m_4 are reflections in the square's diagonals.

As shown in exercise 28.11, $\mathcal{D}_4$ has three proper subgroups of order 4. They are $\{I, R, R^2, R^3\}$, $\mathcal{H}_1 = \{I, R^2, m_1, m_2\}$ and $\mathcal{H}_2 = \{I, R^2, m_3, m_4\}$. The first of these is a cyclic subgroup but the other two are not. The group tables for the latter two, extracted from the one above, are as follows

$\mathcal{H}_1$	I	R^2	m_1	m_2
I	I	R^2	m_1	m_2
R^2	R^2	I	m_2	m_1
m_1	m_1	m_2	I	R^2
m_2	m_2	m_1	R^2	I

$\mathcal{H}_2$	I	R^2	m_3	m_4
I	I	R^2	m_3	m_4
R^2	R^2	I	m_4	m_3
m_3	m_3	m_4	I	R^2
m_4	m_4	m_3	R^2	I

Clearly, these two subgroups are isomorphic with $m_1 \leftrightarrow m_3$ and $m_2 \leftrightarrow m_4$ and the other elements unchanged.

Next, we reproduce the group table for the quaternion group, but with the columns and rows reordered (this does not alter the information it carries):

	1	i	j	k	-1	$-i$	$-j$	$-k$
1	1	i	j	k	-1	$-i$	$-j$	$-k$
i	i	-1	k	$-j$	$-i$	1	$-k$	j
j	j	$-k$	-1	i	$-j$	k	1	$-i$
k	k	j	$-i$	-1	$-k$	$-j$	i	1
-1	-1	$-i$	$-j$	$-k$	1	i	j	k
$-i$	$-i$	1	$-k$	j	i	-1	k	$-j$
$-j$	$-j$	k	1	$-i$	j	$-k$	-1	i
$-k$	$-k$	$-j$	i	1	k	j	$-i$	-1

If we now make the two-to-one mapping

$$\Phi \quad : \quad \pm 1 \to I, \quad \pm i \to R^2, \quad \pm j \to m_1, \quad \pm k \to m_2,$$

each quadrant of the table for $\mathcal{Q}$ becomes identical to that for $\mathcal{H}_1$, showing that Φ is a homomorphism of $\mathcal{Q}$ onto $\mathcal{H}_1$. As $\mathcal{H}_1$ and $\mathcal{H}_2$ are isomorphic there exists a similar homomorphism onto $\mathcal{H}_2$.

Finally, the kernel of either mapping contains those elements of $\mathcal{Q}$ that map onto I, namely 1 and -1.

28.23 *Find* (a) *all the proper subgroups and* (b) *all the conjugacy classes of the symmetry group of a regular pentagon.*

A regular pentagon (see figure 28.2) has rotational symmetries and reflection symmetries about lines that join a vertex to the centre-point of the opposite side. Clearly there are five of the latter, m_i $(i = 1, 2, \ldots, 5)$. If R represents a rotation by $2\pi/5$, then the rotational symmetries are R, R^2, R^3 and R^4. To these must be added the 'do nothing' identity I. The symmetry group of the regular pentagon

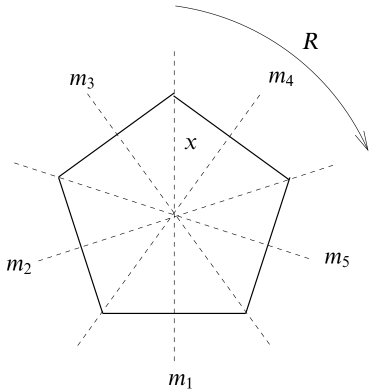

Figure 28.2 The regular pentagon of exercise 28.23.

($\mathcal{C}_{5v}$ in chemical notation) therefore consists of the following ten elements (with their orders):

element	:	I	R	R^2	R^3	R^4	m_1	m_2	m_3	m_4	m_5
order	:	1	5	5	5	5	2	2	2	2	2

(a) As the order of the group is 10, the order of any proper subgroup can only be 2 or 5. As I must be in every subgroup and the order of any element in it must divide the order of the subgroup, it is clear that there is only one subgroup of order 5 and that is $\{I, R, R^2, R^3, R^4\}$. Similarly, there are five subgroups of order 2, namely $\{I, m_i\}$ for m_i $(i = 1, 2, \dots, 5)$.

(b) As always, I is in a class by itself.

We now prove a useful general result about elements in the same conjugacy class: namely, that they have the same order. Let X and Y be in the same class ($Y = g_i X g_i^{-1}$ for some g_i belong to the group) and let X have order m, i.e. $X^m = I$. Then

$$Y^m = g_i X g_i^{-1}\, g_i X g_i^{-1} \cdots g_i X g_i^{-1} = g_i X^m g_i^{-1} = g_i g_i^{-1} = I.$$

This implies that the order of Y divides the order of X. Similarly the order of X divides the order of Y. Therefore X and Y have the same order. Applying this result to the given group, we see that a conjugacy class cannot contain a mixture of rotations and reflections.

We first note the obvious result that $R^p\, R^q\, (R^p)^{-1} = R^q$ for any valid p and q. Next, by considering the effects of various combinations of symmetries on a general point x of the pentagon (as marked in the figure), we find that for any i and j, $(i, j = 1, 2 \dots, 5)$,

$$m_i R m_i = R^4 \quad \text{and} \quad m_j R^4 m_j = R.$$

These results, together with that noted above, imply that R and R^4 constitute a class. Similarly R^2 and R^3 make up a (different) class.

Turning to the reflections, we see that the following chain of results, for example, shows that all five reflections must be in the same class (recall that each reflection is its own inverse):

$$m_1m_2m_1 = m_5, \quad m_3m_5m_3 = m_1, \quad m_2m_1m_2 = m_3, \quad m_1m_3m_1 = m_4.$$

In summary, there are four conjugacy classes and they are I, (R, R^4), (R^2, R^3) and $(m_1, m_2, m_3, m_4, m_5)$.

29

Representation theory

29.1 *A group $\mathcal{G}$ has four elements I, X, Y and Z, which satisfy $X^2 = Y^2 = Z^2 = XYZ = I$. Show that $\mathcal{G}$ is Abelian and hence deduce the form of its character table.*

Show that the matrices

$$\mathsf{D}(I) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \mathsf{D}(X) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$\mathsf{D}(Y) = \begin{pmatrix} -1 & -p \\ 0 & 1 \end{pmatrix}, \qquad \mathsf{D}(Z) = \begin{pmatrix} 1 & p \\ 0 & -1 \end{pmatrix},$$

where p is a real number, form a representation D of $\mathcal{G}$. Find its characters and decompose it into irreps.

Since I necessarily commutes with all other elements, we need only consider products such as XY. Now,

$$\begin{aligned} XYZ = I \quad &\Rightarrow \quad X^2YZ = X \quad \Rightarrow \quad YZ = X, \\ XYZ = I \quad &\Rightarrow \quad XYZ^2 = Z \quad \Rightarrow \quad XY = Z \\ &\Rightarrow \quad XY^2 = ZY \quad \Rightarrow \quad X = ZY. \end{aligned}$$

Thus, $YZ = X = ZY$, showing that Y and Z commute. Similarly, $XY = Z = YX$ and $XZ = Y = ZX$. We conclude that the group is Abelian.

As the group is Abelian, each element is in a class of its own and there are therefore four classes and consequently four irreps $\mathrm{D}^{(\lambda)}$. Since

$$\sum_{\lambda=1}^{4} n_\lambda^2 = g = 4,$$

where n_λ is the dimension of representation λ, the only possibility is that $n_\lambda = 1$ for each λ, i.e the group has four one-dimensional irreps.

As for all sets of irreps, the identity irrep $D^{(1)} = A_1$ must be present and the characters of the others must be orthogonal to the (1,1,1,1) characters of A_1. Further, for each one-dimensional irrep, the identity I must have the character $+1$. The character table must therefore take the form

χ	I	X	Y	Z
A_1	1	1	1	1
$D^{(2)}$	1	1	−1	−1
$D^{(3)}$	1	−1	1	−1
$D^{(4)}$	1	−1	−1	1

For the proposed representation we first need to verify the multiplication properties. Those for $\mathsf{D}(I)$ are immediate. For the others:

$$\mathsf{D}(X)\mathsf{D}(X) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathsf{D}(I),$$

$$\mathsf{D}(Y)\mathsf{D}(Y) = \begin{pmatrix} -1 & -p \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & -p \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathsf{D}(I),$$

$$\mathsf{D}(Z)\mathsf{D}(Z) = \begin{pmatrix} 1 & p \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & p \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathsf{D}(I),$$

$$\begin{aligned}\mathsf{D}(X)\mathsf{D}(Y)\mathsf{D}(Z) &= \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} -1 & -p \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & p \\ 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = \mathsf{D}(I),\end{aligned}$$

Since the defining relationships for I, X, Y and Z are $X^2 = Y^2 = Z^2 = XYZ = I$, these results show that the matrices form a representation of $\mathcal{G}$, whatever the value of p.

Now $\chi(g_i)$ is equal to the trace of $\mathsf{D}(g_i)$ and so the character set for this representation (in the order I, X, Y, Z) is $(2, -2, 0, 0)$. The only rows in the character table that can be added to produce the correct totals for all four elements are the third and fourth. This shows that

$$\mathsf{D} = D^{(3)} \oplus D^{(4)}.$$

29.3 *The quaternion group $\mathcal{Q}$ (see exercise 28.21) has eight elements $\{\pm 1, \pm i, \pm j, \pm k\}$ obeying the relations*

$$i^2 = j^2 = k^2 = -1, \quad ij = k = -ji.$$

Determine the conjugacy classes of $\mathcal{Q}$ and deduce the dimensions of its irreps. Show that $\mathcal{Q}$ is homomorphic to the four-element group $\mathcal{V}$, which is generated by two distinct elements a and b with $a^2 = b^2 = (ab)^2 = I$. Find the one-dimensional irreps of $\mathcal{V}$ and use these to help determine the full character table for $\mathcal{Q}$.

As always, the identity, $+1$, is in a class by itself and, since it commutes with every other element in the group, so is -1.

Now consider all products of the form $X^{-1}iX$:

$$\begin{aligned}
1\,i\,1 &= i, & (-1)\,i\,(-1) &= i,\\
(-i)\,i\,i &= i, & i\,i\,(-i) &= i,\\
(-j)\,i\,j = (-j)\,k &= -i, & j\,i\,(-j) &= (-k)\,(-j) = -i,\\
(-k)\,i\,k = (-k)\,(-j) &= -i, & k\,i\,(-k) &= j\,(-k) = -i.
\end{aligned}$$

These show that i and $-i$ are in the same class. Similarly $\{j, -j\}$ and $\{k, -k\}$ are two other classes. This exhausts the group.

In summary there are five classes, the elements in any one class all having the same order. They are

$$\begin{array}{rccccc}
\text{class :} & \{1\} & \{-1\} & \{\pm i\} & \{\pm j\} & \{\pm k\} \\
\text{order :} & 1 & 2 & 4 & 4 & 4
\end{array}$$

It follows that there are five irreps and, since $\sum_{\lambda=1}^{5} n_\lambda^2 = 8$, they can only be one two-dimensional and four one-dimensional irreps.

Turning to the group $\mathcal{V}$,

$$\begin{aligned}
(ab)^2 = I \quad &\Rightarrow \quad abab^2 = b \quad \Rightarrow \quad aba = b \quad \Rightarrow \quad aba^2 = ba\\
&\Rightarrow \quad ab = ba,
\end{aligned}$$

i.e. a and b commute. Also, it follows that

$$a(ab) = a(ba) = (ab)a \quad \text{and} \quad b(ab) = (ba)b = (ab)b.$$

Thus, all four elements commute, the group $\mathcal{V}$ is Abelian, and each of its elements is in a class of its own. As in exercise 29.1, an Abelian group of order 4 must have four irreps and the character table

$\chi_{\mathcal{V}}$	I	a	b	ab
A_1	1	1	1	1
$D^{(2)}$	1	1	−1	−1
$D^{(3)}$	1	−1	1	−1
$D^{(4)}$	1	−1	−1	1

The multiplication table for the quaternion group is, as given in exercise 21 of chapter 28 of the form

$\mathcal{Q}$	1	-1	i	$-i$	j	$-j$	k	$-k$
1	1	-1	i	$-i$	j	$-j$	k	$-k$
-1	-1	1	$-i$	i	$-j$	j	$-k$	k
i	i	$-i$	-1	1	k	$-k$	$-j$	j
$-i$	$-i$	i	1	-1	$-k$	k	j	$-j$
j	j	$-j$	$-k$	k	-1	1	i	$-i$
$-j$	$-j$	j	k	$-k$	1	-1	$-i$	i
k	k	$-k$	j	$-j$	$-i$	i	-1	1
$-k$	$-k$	k	$-j$	j	i	$-i$	1	-1

If each entry ± 1 is replaced by I, each entry $\pm i$ by a, each entry $\pm j$ by b and each entry $\pm k$ by ab, then this table reduces to four copies of the table

$\mathcal{V}$	I	a	b	ab
I	I	a	b	ab
a	a	I	ab	b
b	b	ab	I	a
ab	ab	b	a	I

which is the group multiplication table for group $\mathcal{V}$. The same conclusion can be reached by replacing each 2×2 block containing only ± 1 by I, each 2×2 block containing only $\pm i$ by a, etc.; this results in a single copy. Both approaches lead to the conclusion that there is a two-to-one homomorphism from $\mathcal{Q}$ onto $\mathcal{V}$.

Since the homomorphism maps all elements of any one conjugacy class of $\mathcal{Q}$ onto the same element of $\mathcal{V}$, the one-dimensional irreps of $\mathcal{Q}$ must be the same as those of $\mathcal{V}$. Further, since both of the classes $\{1\}$ and $\{-1\}$ map onto I in $\mathcal{V}$ they will have common characters in each one-dimensional irrep (1 in every case, as it happens). As shown earlier, there will also be a two-dimensional irrep. Its character for I must be 2 (the dimension of the irrep); the other characters can be determined from the requirement of orthogonality to the characters of the other (one-dimensional) irreps.

The character table for Q therefore has the form

χ_Q	1	-1	$\pm i$	$\pm j$	$\pm k$
A_1	1	1	1	1	1
$D^{(2)}$	1	1	1	-1	-1
$D^{(3)}$	1	1	-1	1	-1
$D^{(4)}$	1	1	-1	-1	1
$D^{(5)}$	2	w	x	y	z

We require that

$$\begin{aligned}
1(1)(2) + 1(1)(w) + 2(1)(x) + 2(1)y + 2(1)z &= 0,\\
1(1)(2) + 1(1)(w) + 2(1)(x) + 2(-1)y + 2(-1)z &= 0,\\
1(1)(2) + 1(1)(w) + 2(-1)(x) + 2(1)y + 2(-1)z &= 0,\\
1(1)(2) + 1(1)(w) + 2(-1)(x) + 2(-1)y + 2(1)z &= 0.
\end{aligned}$$

These equations have the solution $w = -2$, $x = y = z = 0$, thus completing the full character table for Q.

29.5 *The group of pure rotations (excluding reflections and inversions) that take a cube into itself has 24 elements. The group is isomorphic to the permutation group S_4 and hence has the same character table, once corresponding classes have been established. By counting the number of elements in each class, make the correspondences below (the final two cannot be decided purely by counting, and should be taken as given).*

Permutation class type	*Symbol (physics)*	*Action*
(1)	I	*none*
(123)	3	*rotations about a body diagonal*
(12)(34)	2_z	*rotation of π about the normal to a face*
(1234)	4_z	*rotations of $\pm\pi/2$ about the normal to a face*
(12)	2_d	*rotation of π about an axis through the centres of opposite edges*

Given in table 29.1 is the character table for S_4. Reformulate it in terms of the elements of the rotation symmetry group (432 or O) of a cube and use it when answering exercise 29.7.

The rotational symmetries of the cube are as follows.

(i) ‘Do nothing’. There is only one such operation; it is the identity and so corresponds to $(1) \equiv (1)(2)(3)(4)$.

Irrep	Typical element and class size				
	(1)	(12)	(123)	(1234)	(12)(34)
	1	6	8	6	3
A_1	1	1	1	1	1
A_2	1	−1	1	−1	1
E	2	0	−1	0	2
T_1	3	1	0	−1	−1
T_2	3	−1	0	1	−1

Table 29.1 The character table for the permutation group S_4.

(ii) Rotations about a body diagonal. There are four body diagonals and rotations of $2\pi/3$ and $4\pi/3$ are possible about each. Thus there are eight elements and this must correspond to $(123) \equiv (123)(4)$.

(iii) Rotations by π about a face normal. Although there are six faces to the cube, they define only three distinct face normals and hence there are three elements in this set. They therefore correspond to (12)(34).

(iv) Rotations of $\pi/2$ and $3\pi/2$ about a face normal. With three distinct face normals and two possible rotation angles for each, the set contains six elements. These could correspond to $(12) \equiv (12)(3)(4)$ or to (1234).

(v) Rotations of π about axes that join the centres of opposite edges. There are six such axes and hence six elements. As in (iv), these could correspond to $(12) \equiv (12)(3)(4)$ or to (1234).

Taking the identification given in the question for (iv) and (v), the reformulated table (in which only the headings have changed) is given in table 29.2.

Irrep	Typical element and class size				
	I	2_d	3	4_z	2_z
	1	6	8	6	3
A_1	1	1	1	1	1
A_2	1	−1	1	−1	1
E	2	0	−1	0	2
T_1	3	1	0	−1	−1
T_2	3	−1	0	1	−1

Table 29.2 The character table for the symmetry group 432 or O.

We note that the rotational symmetries of a cube can, alternatively, be characterised by the effects they have on the orientations in space of its four body diagonals. For example, a rotation of π about a face normal interchanges them in pairs – represented in cycle notation by the form (12)(34). Using this formulation of the symmetry group, the assignments for (iv) and (v) are unambiguous.

29.7 *In a certain crystalline compound, a thorium atom lies at the centre of a regular octahedron of six sulphur atoms at positions $(\pm a, 0, 0)$, $(0, \pm a, 0)$, $(0, 0, \pm a)$. These can be considered as being positioned at the centres of the faces of a cube of side $2a$. The sulphur atoms produce at the site of the thorium atom an electric field that has the same symmetry group as a cube (*432 *or* O*).*

The five degenerate d-electron orbitals of the thorium atom can be expressed, relative to any arbitrary polar axis, as

$$(3\cos^2\theta - 1)f(r), \qquad e^{\pm i\phi}\sin\theta\cos\theta f(r), \qquad e^{\pm 2i\phi}\sin^2\theta f(r).$$

A rotation about that polar axis through an angle ϕ' in effect changes ϕ to $\phi - \phi'$. Use this to show that the character of the rotation in a representation based on the orbital wavefunctions is given by

$$1 + 2\cos\phi' + 2\cos 2\phi'$$

and hence that the characters of the representation, in the order of the symbols given in exercise 29.5, is 5, −1, 1, −1, 1. *Deduce that the five-fold degenerate level is split into two levels, a doublet and a triplet.*

The electric field at the thorium atom has the symmetries of group 432 and the d-electron orbitals are

$$\begin{aligned}
\psi_1 &= (3\cos^2\theta - 1)f(r),\\
\psi_{2,3} &= e^{\pm i\phi}\sin\theta\cos\theta f(r),\\
\psi_{4,5} &= e^{\pm 2i\phi}\sin^2\theta f(r).
\end{aligned}$$

Taking the ψ_i ($i = 1, 2, \ldots, 5$) as a basis, the representation of a rotation by ϕ' is a 5×5 matrix whose diagonal elements are equal to the factor by which each basis function is multiplied when that function is subjected to the rotation.

ψ_1 does not depend upon ϕ and so is unaltered; its entry is 1.

$\psi_{2,3}$ become $\sin\theta\cos\theta f(r)\, e^{\pm i\phi \mp i\phi'}$; their entries are $e^{-i\phi'}$ and $e^{i\phi'}$.

$\psi_{4,5}$ become $\sin^2\theta f(r)\, e^{\pm 2i\phi \mp 2i\phi'}$; their entries are $e^{-2i\phi'}$ and $e^{i2\phi'}$.

The trace of the representative matrix, and therefore the character of the rotation, is thus

$$\chi = 1 + e^{-i\phi'} + e^{i\phi'} + e^{-2i\phi'} + e^{2i\phi'} = 1 + 2\cos\phi' + 2\cos 2\phi'.$$

For the symmetry elements in the group 432, the corresponding rotation angles and characters are as follows:

Symmetry	ϕ'	χ
I	0	$1+2+2=5$
3	$\pm 2\pi/3$	$1+2(-\frac{1}{2})+2(-\frac{1}{2})=-1$
2_z	π	$1+2(-1)+2(1)=1$
4_z	$\pm\pi/2$	$1+2(0)+2(-1)=-1$
2_d	π	$1+2(-1)+2(1)=1$

Rewriting these results in a form similar to that in which the character table of 432 has been previously presented, we have

Symmetry	I	2_d	3	$4z$	$2z$
Character, χ	5	1	-1	-1	1

We now compare this with table 29.2, compiled in exercise 29.5, and see, or calculate using the equation

$$m_\mu = \frac{1}{g}\sum_X \left[\chi^{(\mu)}(X)\right]^* \chi(X) = \frac{1}{g}\sum_i c_i \left[\chi^{(\mu)}(X_i)\right]^* \chi(X_i),$$

that this character set is the direct sum of those for the two dimensional irrep E and the three-dimensional irrep T_1, as given in that table:

	I	2_d	3	4_z	2_z
E	2	0	-1	0	2
T_1	3	1	0	-1	-1
χ	5	1	-1	-1	1

The n (mixed) orbitals transforming according to any particular n-dimensional irrep will all have the same energy, but, barring accidental coincidences, it will be a different energy from that corresponding to a different irrep. Thus the five-fold degenerate level is split into a doublet (E) and a triplet (T_1).

29.9 *The hydrogen atoms in a methane molecule* CH_4 *form a perfect tetrahedron with the carbon atom at its centre. The molecule is most conveniently described mathematically by placing the hydrogen atoms at the points* $(1,1,1)$, $(1,-1,-1)$, $(-1,1,-1)$ *and* $(-1,-1,1)$. *The symmetry group to which it belongs, the tetrahedral group (*$\bar{4}3m$ *or* T_d*), has classes typified by* I, 3, 2_z, m_d *and* $\bar{4}_z$, *where the first three are as in exercise 29.5,* m_d *is a reflection in the mirror plane* $x - y = 0$ *and* $\bar{4}_z$ *is a rotation of* $\pi/2$ *about the* z*-axis followed by an inversion in the origin. A reflection in a mirror plane can be considered as a rotation of* π *about an axis perpendicular to the plane, followed by an inversion in the origin.*

The character table for the group $\bar{4}3m$ *is very similar to that for the group* 432, *and has the form shown in table 29.3. By following the steps given below, determine how many different internal vibration frequencies the* CH_4 *molecule has.*

(a) *Consider a representation based on the twelve coordinates* x_i, y_i, z_i *for* $i = 1, 2, 3, 4$. *For those hydrogen atoms that transform into themselves, a rotation through an angle* θ *about an axis parallel to one of the coordinate axes gives rise in this natural representation to the diagonal elements* 1 *for the corresponding coordinate and* $2\cos\theta$ *for the two orthogonal coordinates. If the rotation is followed by an inversion then these entries are multiplied by* -1. *Atoms not transforming into themselves give a zero diagonal contribution. Show that the characters of the natural representation are* 12, 0, 0, 0, 2 *and hence that its expression in terms of irreps is*

$$A_1 \oplus E \oplus T_1 \oplus 2T_2.$$

(b) *The irreps of the bodily translational and rotational motions are included in this expression and need to be identified and removed. Show that when this is done it can be concluded that there are three different internal vibration frequencies in the* CH_4 *molecule. State their degeneracies and check that they are consistent with the expected number of normal coordinates needed to describe the internal motions of the molecule.*

(a) We consider each type of rotation in turn and determine how many of the hydrogen atoms are transformed into themselves, i.e. do not change position.

Under I all twelve atoms retain their original positions and so $\chi(I) = 12$.

For the symmetry 3 the rotation angle is $\pm 2\pi/3$ and for each such rotation one atom retains its original place. However, it contributes $1 + 2\cos(2\pi/3) = 1 + 2(-\frac{1}{2}) = 0$ and so $\chi(3) = 0$.

For the symmetries 2_z and $\bar{4}_z$ no atoms retain their original places and the corresponding characters are both 0.

Irreps	Typical element and class size I 1	3 8	2_z 3	$\bar{4}_z$ 6	m_d 6	Functions transforming according to irrep
A_1	1	1	1	1	1	$x^2+y^2+z^2$
A_2	1	1	1	-1	-1	
E	2	-1	2	0	0	$(x^2-y^2, 3z^2-r^2)$
T_1	3	0	-1	1	-1	(R_x, R_y, R_z)
T_2	3	0	-1	-1	1	(x, y, z); (xy, yz, zx)

Table 29.3 The character table for the symmetry group $\bar{4}3m$.

Finally, for m_d, a reflection in one of the six mirror planes, there are two atoms that lie in any one of the planes (with the other two atoms placed symmetrically, one on either side of it). Thus two atoms are unchanged. As explained in the question, a reflection in a mirror plane can be considered as a rotation of π about an axis perpendicular to the plane, followed by an inversion in the origin; the latter gives rise to an additional factor of -1. As a result, each of the two atoms contributes $(-1)(1+2\cos\pi)=1$ to the character of m_d. Thus $\chi(m_d)=2$ and the full character set of the natural representation is $(12, 0, 0, 0, 2)$.

It then follows that

$$m_{A_1} = \frac{1(1)(12)+8(1)(0)+3(1)(0)+6(1)(0)+6(1)(2)}{24} = 1,$$

$$m_{A_2} = \frac{1(1)(12)+8(1)(0)+3(1)(0)+6(-1)(0)+6(-1)(2)}{24} = 0,$$

$$m_{E} = \frac{1(2)(12)+8(-1)(0)+3(2)(0)+6(0)(0)+6(0)(2)}{24} = 1,$$

$$m_{T_1} = \frac{1(3)(12)+8(0)(0)+3(-1)(0)+6(1)(0)+6(-1)(2)}{24} = 1,$$

$$m_{T_2} = \frac{1(3)(12)+8(0)(0)+3(-1)(0)+6(-1)(0)+6(1)(2)}{24} = 2.$$

Thus the irreps present in this representation are

$$A_1 \oplus E \oplus T_1 \oplus 2T_2.$$

(b) Bodily translation of the centre of mass of the molecule is included in this representation since the representation allows all coordinates to vary independently. From table 29.3, the set (x, y, z) transforms as T_2 and so this motion corresponds to one of the two T_2 irreps found above.

Equally a rigid-body rotation of the molecule *about* its centre of mass is included; from the table (R_x, R_y, R_z) transform as T_1 and so this rotation, which contains no internal vibrations, accounts for the T_1 irrep.

After these two irreps are removed we are left with the irreps of the internal vibrations, which are A_1, E, and T_2. They are, respectively, one-, two- and three-dimensional irreps and therefore the corresponding vibration frequencies have degeneracies of 1, 2 and 3. This gives a total of six internal coordinates, in accordance with the twelve original ones, less the three translational coordinates of the centre of mass and the three coordinates needed to specify the direction of the axis of a rigid-body rotation.

***29.11** Use the results of exercise 28.23 to find the character table for the dihedral group $\mathcal{D}_5$, the symmetry group of a regular pentagon.*

As shown in exercise 28.23, the group $\mathcal{D}_5$ has ten elements and four classes: $\{I\}$, $\{R, R^4\}$, $\{R^2, R^3\}$ and $\{m_1, m_2, m_3, m_4, m_5\}$. Here R is a rotation through $2\pi/5$.

Since there are ten elements and four classes, and hence four irreps, we must have that the dimensionalities of the irreps satisfy

$$n_1^2 + n_2^2 + n_3^2 + n_4^2 = 10.$$

This has only one (non-zero) integral solution, $n_1 = n_2 = 1$ and $n_3 = n_4 = 2$. The identity irrep, A_1, must be one of the one-dimensional irreps, and the character table must have the form

Irrep	I	R, R^4	R^2, R^3	m_i $(i = 1, 5)$
A_1	1	1	1	1
A_2	1	a	b	c
E_1	2	d	e	f
E_2	2	g	h	j

For A_2 we must have both $1 + 2|a|^2 + 2|b|^2 + 5|c|^2 = 10$ (summation rule) and $1 + 2a + 2b + 5c = 0$ (orthogonality with A_1). Since the m_i have order 2 and A_2 is one-dimensional, c can only be a second root of unity, i.e. either of 1 or -1. The only solution to these simultaneous equations, even allowing a and b to be complex (but restricted to each being a fifth root of unity), is $a = b = 1$ and $c = -1$.

For E_1 (and similarly for E_2),

$$4 + 2|d|^2 + 2|e|^2 + 5|f|^2 = 10.$$

Arguing as previously, we conclude that, because E is two-dimensional, f can only be the sum of two values which are either $+1$ or -1. Hence, only 0 and ± 2 are possible, and the values ± 2 are impossible in this case. This conclusion can be confirmed by noting that the E_1 character set has to be orthogonal to those for both A_1 and A_2. So both

$$1(1)2 + 2(1)d + 2(1)e + 5(1)f = 0$$

and $$1(1)2 + 2(1)d + 2(1)e + 5(-1)f = 0,$$

implying that $f = 0$.

We are left with

$$|d|^2 + |e|^2 = 3 \text{ and } 1 + d + e = 0.$$

This clearly has no integer solutions, but we attempt to find real solutions before considering complex ones. If d is real then e must also be real. Substituting $e = -1 - d$ into the first equation gives the quadratic equation

$$d^2 + (-1 - d)^2 = 3 \quad \Rightarrow \quad d^2 + d - 1 = 0 \quad \Rightarrow \quad d = \frac{-1 \pm \sqrt{5}}{2}.$$

If d is taken as $\frac{1}{2}(-1 + \sqrt{5})$ (the golden mean!) then $e = \frac{1}{2}(-1 - \sqrt{5})$, the other root of the quadratic. This completes the character set for E_1. That for E_2 is obtained by setting $j = 0$ and assigning $\frac{1}{2}(-1 + \sqrt{5})$ to h and $\frac{1}{2}(-1 - \sqrt{5})$ to g; this can be confirmed by checking the orthogonality relation

$$1(2)(2) + 2[\tfrac{1}{2}(-1 + \sqrt{5})\,\tfrac{1}{2}(-1 - \sqrt{5})] \\ +2[\tfrac{1}{2}(-1 - \sqrt{5})\,\tfrac{1}{2}(-1 + \sqrt{5})] + 5(0)(0) = 4 - 2 - 2 + 0 = 0.$$

We also note that, for example,

$$\frac{-1 + \sqrt{5}}{2} = \exp\left(\frac{2\pi i}{5}\right) + \exp\left(\frac{4 \times 2\pi i}{5}\right) = 2\cos\frac{2\pi}{5} = 0.6180,$$

i.e. d and h are each the sum of two fifth roots of unity. The same applies to e and g. The final character table reads

Irrep	I	R, R^4	R^2, R^3	m_i $(i = 1, 5)$
A_1	1	1	1	1
A_2	1	1	1	-1
E_1	2	$\frac{1}{2}(-1 + \sqrt{5})$	$-\frac{1}{2}(1 + \sqrt{5})$	0
E_2	2	$-\frac{1}{2}(1 + \sqrt{5})$	$\frac{1}{2}(-1 + \sqrt{5})$	0

***29.13** Further investigation of the crystalline compound considered in exercise 29.7 shows that the octahedron is not quite perfect but is elongated along the* $(1,1,1)$ *direction with the sulphur atoms at positions* $\pm(a+\delta,\delta,\delta)$, $\pm(\delta,a+\delta,\delta)$, $\pm(\delta,\delta,a+\delta)$, *where* $\delta \ll a$. *This structure is invariant under the (crystallographic) symmetry group 32 with three two-fold axes along directions typified by* $(1,-1,0)$. *The latter axes, which are perpendicular to the* $(1,1,1)$ *direction, are axes of two-fold symmetry for the perfect octahedron. The group 32 is really the three-dimensional version of the group 3m and has the same character table. That for 3m is*

$3m$	I	A, B	C, D, E
A_1	1	1	1
A_2	1	1	-1
E	2	-1	0

Use this to show that, when the distortion of the octahedron is included, the doublet found in exercise 29.7 is unsplit but the triplet breaks up into a singlet and a doublet.

The perfect octahedron is invariant under the operations of group 432, whose character table is as follows:

Irrep	Typical element and class size				
	I	2_d	3	4_z	2_z
	1	6	8	6	3
A_1	1	1	1	1	1
A_2	1	-1	1	-1	1
E	2	0	-1	0	2
T_1	3	1	0	-1	-1
T_2	3	-1	0	1	-1

The distorted octahedron is invariant only under the operations of the smaller group 32, whose character table is

$3m$	I	R, R^2	m_i
A_1	1	1	1
A_2	1	1	-1
E	2	-1	0

Here R is a rotation through $2\pi/3$ and its class corresponds to the class denoted by '3' in group 432. The reflection symmetries correspond to rotations by π when considered as operations in three dimensions (as opposed to in a plane); thus they correspond to the class 2_d. We are thus concerned with the first three classes in the 432 table, but with the second and third interchanged as compared with the 32 table.

Using the order in the 32 table, E has the characters $(2, -1, 0)$. This two-dimensional irrep also appears in the 432 table, and so the corresponding doublet level in the thorium atom is not split as a result of the distortion of the sulphur octahedron. However, the triplet level, whose components transform as T_1, will be affected by the distortion. The irrep T_1 does not appear in the 32 table but has to be made up from E and A_1; in terms of character sets

$$(3, 0, 1) = (2, -1, 0) + (1, 1, 1).$$

In physical terms, the triplet state in thorium is split by the distorted electric field due to the sulphur atoms into a doublet and a singlet.

30

Probability

30.1 *By shading or numbering Venn diagrams, determine which of the following are valid relationships between events. For those that are, prove the relationship using de Morgan's laws.*

(a) $\overline{(\bar{X} \cup Y)} = X \cap \bar{Y}$.
(b) $\bar{X} \cup \bar{Y} = \overline{(X \cup Y)}$.
(c) $(X \cup Y) \cap Z = (X \cup Z) \cap Y$.
(d) $X \cup \overline{(Y \cap Z)} = (X \cup \bar{Y}) \cap \bar{Z}$.
(e) $X \cup \overline{(Y \cap Z)} = (X \cup \bar{Y}) \cup \bar{Z}$.

For each part of this question we refer to the corresponding part of figure 30.1.

(a) This relationship is correct as both expressions define the shaded region that is both inside X and outside Y.

(b) This relationship is *not* valid. The LHS specifies the whole sample space *apart from* the region marked with the heavy shading. The RHS defines the region that is lightly shaded. The unmarked regions of X and Y are included in the former but not in the latter.

(c) This relationship is *not* valid. The LHS specifies the sum of the regions marked 2, 3 and 4 in the figure, whilst the RHS defines the sum of the regions marked 1, 3 and 4.

(d) This relationship is *not* valid. On the LHS, $\overline{Y \cap Z}$ is the whole sample space apart from regions 3 and 4. So $X \cup \overline{(Y \cap Z)}$ consists of all regions except for region 3. On the RHS, $X \cup \bar{Y}$ contains all regions except 3 and 7. The events $\bar{Z}$ contain regions 1, 6, 7 and 8 and so $(X \cup \bar{Y}) \cap \bar{Z}$ consists of regions 1, 6 and 8. Thus regions 2, 4, 5 and 7 are in one specification but not in the other.

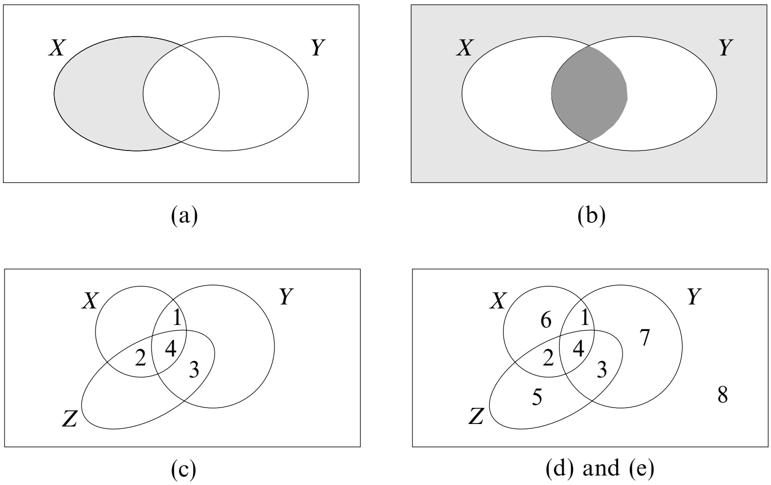

Figure 30.1 The Venn diagrams used in exercise 30.1.

(e) This relationship is valid. The LHS is as found in (d), namely all regions except for region 3. The RHS consists of the union (as opposed to the intersection) of the two subregions found in (d) and thus contains those regions found in either or both of $X \cup \bar{Y}$ (1, 2, 4, 5, 6 and 8) and $\bar{Z}$ (1, 6, 7 and 8). This covers all regions except region 3 – in agreement with those found for the LHS.

For the two valid relationships, their proofs using de Morgan's laws are:

(a) $$\overline{(\bar{X} \cup Y)} = \bar{\bar{X}} \cap \bar{Y} = X \cap \bar{Y},$$

(e) $$X \cup \overline{(Y \cap Z)} = X \cup (\bar{Y} \cup \bar{Z}) = (X \cup \bar{Y}) \cup \bar{Z}.$$

30.3 *A and B each have two unbiased four-faced dice, the four faces being numbered* 1, 2, 3 *and* 4. *Without looking, B tries to guess the sum x of the numbers on the bottom faces of A's two dice after they have been thrown onto a table. If the guess is correct B receives x^2 euros, but if not he loses x euros.*

Determine B's expected gain per throw of A's dice when he adopts each of the following strategies:

(a) *he selects x at random in the range $2 \le x \le 8$;*
(b) *he throws his own two dice and guesses x to be whatever they indicate;*
(c) *he takes your advice and always chooses the same value for x. Which number would you advise?*

We first calculate the probabilities $p(x)$ and the corresponding gains $g(x) = p(x)x^2 - [1 - p(x)]x$ for each value of the total x. Expressing both in units of 1/16, they are as follows:

x	2	3	4	5	6	7	8
$p(x)$	1	2	3	4	3	2	1
$g(x)$	−26	−24	−4	40	30	0	−56

(a) If B's guess is random in the range $2 \leq x \leq 8$ then his expected return is

$$\frac{1}{16}\frac{1}{7}(-26 - 24 - 4 + 40 + 30 + 0 - 56) = -\frac{40}{112} = -0.36 \text{ euros.}$$

(b) If he picks by throwing his own dice then his distribution of guesses is the same as that of $p(x)$ and his expected return is

$$\frac{1}{16}\frac{1}{16}[1(-26) + 2(-24) + 3(-4) + 4(40) + 3(30) + 2(0) + 1(-56)]$$
$$= \frac{108}{256} = 0.42 \text{ euros.}$$

(c) As is clear from the tabulation, the best return of $40/16 = 2.5$ euros is expected if B always chooses '5' as his guess. Of course, you should not advise him but offer to take his place!

30.5 *Two duellists, A and B, take alternate shots at each other, and the duel is over when a shot (fatal or otherwise!) hits its target. Each shot fired by A has a probability α of hitting B, and each shot fired by B has a probability β of hitting A. Calculate the probabilities P_1 and P_2, defined as follows, that A will win such a duel: P_1, A fires the first shot; P_2, B fires the first shot.*

If they agree to fire simultaneously, rather than alternately, what is the probability P_3 that A will win, i.e. hit B without being hit himself?

Each shot has only two possible outcomes, a hit or a miss. P_1 is the probability that A will win when it is his turn to fire the next shot, and he is still able to do so (event W). There are three possible outcomes of the first two shots: C_1, A hits with his shot; C_2, A misses but B hits; C_3, both miss. Thus

$$\begin{aligned} P_1 &= \sum_i \Pr(C_i)\Pr(W|C_i) \\ &= [\,\alpha \times 1\,] + [\,(1-\alpha)\beta \times 0\,] + [\,(1-\alpha)(1-\beta) \times P_1\,] \\ \Rightarrow \quad P_1 &= \frac{\alpha}{\alpha + \beta - \alpha\beta}. \end{aligned}$$

When B fires first but misses, the situation is the one just considered. But if B

hits with his first shot then clearly A's chances of winning are zero. Since these are the only two possible outcomes of B's first shot, we can write

$$P_2 = [\beta \times 0] + [(1-\beta) \times P_1] \quad \Rightarrow \quad P_2 = \frac{(1-\beta)\alpha}{\alpha + \beta - \alpha\beta}.$$

When both fire at the same time there are four possible outcomes D_i to the first round: D_1, A hits and B misses; D_2, B hits but A misses; D_3, they both hit; D_4, they both miss. If getting hit, even if you manage to hit your opponent, does not count as a win, then

$$\begin{aligned} P_3 &= \sum_i \Pr(D_i)\Pr(W|D_i) \\ &= [\alpha(1-\beta) \times 1] + [(1-\alpha)\beta \times 0] + [\alpha\beta \times 0] + \; [(1-\alpha)(1-\beta) \times P_3]. \end{aligned}$$

This can be rearranged as

$$P_3 = \frac{\alpha(1-\beta)}{\alpha + \beta - \alpha\beta} = P_2.$$

Thus the result is the same as if B had fired first. However, we also note that if all that matters to A is that B is hit, whether or not he is hit himself, then the third bracket takes the value $\alpha\beta \times 1$ and P_3 takes the same value as P_1.

30.7 *A tennis tournament is arranged on a straight knockout basis for 2^n players, and for each round, except the final, opponents for those still in the competition are drawn at random. The quality of the field is so even that in any match it is equally likely that either player will win. Two of the players have surnames that begin with 'Q'. Find the probabilities that they play each other*

(a) *in the final,*
(b) *at some stage in the tournament.*

Let p_r be the probability that *before* the rth round the two players are both still in the tournament (and, by implication, have not met each other). Clearly, $p_1 = 1$.

Before the rth round there are 2^{n+1-r} players left in. For both 'Q' players to still be in before the $(r+1)$th round, Q_1 must avoid Q_2 in the draw and both must win their matches. Thus

$$p_{r+1} = \frac{2^{n+1-r} - 2}{2^{n+1-r} - 1} \left(\frac{1}{2}\right)^2 p_r.$$

(a) The probability that they meet in the final is p_n, given by

$$\begin{aligned}
p_n &= 1\,\frac{2^n-2}{2^n-1}\,\frac{1}{4}\,\frac{2^{n-1}-2}{2^{n-1}-1}\,\frac{1}{4}\,\cdots\,\frac{2^2-2}{2^2-1}\,\frac{1}{4} \\
&= \left(\frac{1}{4}\right)^{n-1} 2^{n-1}\left[\frac{(2^{n-1}-1)(2^{n-2}-1)\cdots(2^1-1)}{(2^n-1)(2^{n-1}-1)\cdots(2^2-1)}\right] \\
&= \left(\frac{1}{4}\right)^{n-1} 2^{n-1}\,\frac{1}{2^n-1} \\
&= \frac{1}{2^{n-1}(2^n-1)}.
\end{aligned}$$

(b) The more general solution to the recurrence relation derived above is

$$\begin{aligned}
p_r &= 1\,\frac{2^n-2}{2^n-1}\,\frac{1}{4}\,\frac{2^{n-1}-2}{2^{n-1}-1}\,\frac{1}{4}\,\cdots\,\frac{2^{n+2-r}-2}{2^{n+2-r}-1}\,\frac{1}{4} \\
&= \left(\frac{1}{4}\right)^{r-1} 2^{r-1}\left[\frac{(2^{n-1}-1)(2^{n-2}-1)\cdots(2^{n+1-r}-1)}{(2^n-1)(2^{n-1}-1)\cdots(2^{n+2-r}-1)}\right] \\
&= \left(\frac{1}{2}\right)^{r-1} \frac{2^{n+1-r}-1}{2^n-1}.
\end{aligned}$$

Before the rth round, if they are both still in the tournament, the probability that they will be drawn against each other is $(2^{n-r+1}-1)^{-1}$. Consequently, the chance that they will meet at *some* stage is

$$\begin{aligned}
\sum_{r=1}^{n} p_r\,\frac{1}{2^{n-r+1}-1} &= \sum_{r=1}^{n}\left(\frac{1}{2}\right)^{r-1}\frac{2^{n+1-r}-1}{2^n-1}\,\frac{1}{2^{n-r+1}-1} \\
&= \frac{1}{2^n-1}\sum_{r=1}^{n}\left(\frac{1}{2}\right)^{r-1} \\
&= \frac{1}{2^n-1}\,\frac{1-(\frac{1}{2})^n}{1-\frac{1}{2}} = \frac{1}{2^{n-1}}.
\end{aligned}$$

This same conclusion can also be reached in the folowing way. The probability that Q_1 *is not* put out of (i.e. wins) the tournament is $(\frac{1}{2})^n$. It follows that the probability that Q_1 *is* put out is $1-(\frac{1}{2})^n$ and that the player responsible is Q_2 with probability $[\,1-(\frac{1}{2})^n\,]/(2^n-1) = 2^{-n}$. Similarly, the probability that Q_2 is put out and that the player responsible is Q_1 is also 2^{-n}. These are exclusive events but cover all cases in which Q_1 and Q_2 meet during the tournament, the probability of which is therefore $2 \times 2^{-n} = 2^{n-1}$.

30.9 *An electronics assembly firm buys its microchips from three different suppliers; half of them are bought from firm X, whilst firms Y and Z supply 30% and 20%, respectively. The suppliers use different quality-control procedures and the percentages of defective chips are 2%, 4% and 4% for X, Y and Z, respectively. The probabilities that a defective chip will fail two or more assembly-line tests are 40%, 60% and 80%, respectively, whilst all defective chips have a 10% chance of escaping detection. An assembler finds a chip that fails only one test. What is the probability that it came from supplier X?*

Since the number of tests failed by a defective chip are mutually exclusive outcomes (0, 1 or ≥ 2), a chip supplied by X has a probability of failing just one test given by $0.02(1-0.1-0.4) = 0.010$. The corresponding probabilities for chips supplied by Y and Z are $0.04(1-0.1-0.6) = 0.012$ and $0.04(1-0.1-0.8) = 0.004$, respectively.

Using '1' to denote failing a single test, Bayes' theorem gives the probability that the chip was supplied by X as

$$\begin{aligned}\Pr(X|1) &= \frac{\Pr(1|X)\,\Pr(X)}{\Pr(1|X)\,\Pr(X) + \Pr(1|Y)\,\Pr(Y) + \Pr(1|Z)\,\Pr(Z)}\\ &= \frac{0.010 \times 0.5}{0.010 \times 0.5 + 0.012 \times 0.3 + 0.004 \times 0.2} = \frac{50}{94}.\end{aligned}$$

30.11 *A boy is selected at random from amongst the children belonging to families with n children. It is known that he has at least two sisters. Show that the probability that he has $k-1$ brothers is*

$$\frac{(n-1)!}{(2^{n-1}-n)(k-1)!(n-k)!},$$

for $1 \leq k \leq n-2$ and zero for other values of k. Assume that boys and girls are equally likely.

The boy has $n-1$ siblings. Let A_j be the event that $j-1$ of them are brothers, i.e. his family contains j boys and $n-j$ girls. The probability of event A_j is

$$\Pr(A_j) = \frac{{}^{n-1}C_{j-1}\left(\frac{1}{2}\right)^{n-1}}{\sum_{j=1}^{n}\ {}^{n-1}C_{j-1}\left(\frac{1}{2}\right)^{n-1}} = \frac{(n-1)!}{2^{n-1}(j-1)!(n-j)!}.$$

If B is the event that the boy has at least two sisters, then

$$\Pr(B|A_j) = \begin{cases} 1 & 1 \leq j \leq n-2, \\ 0 & n-1 \leq j \leq n. \end{cases}$$

Now we apply Bayes' theorem to give the probability that he has $k-1$ brothers:

$$\Pr(A_k|B) = \frac{1\ \Pr(A_k)}{\sum_{j=1}^{n-2} 1\ \Pr(A_j)},$$

for $1 \le k \le n-2$. The denominator of this expression is the sum $1 = (\frac{1}{2}+\frac{1}{2})^{n-1} = \sum_{j=1}^{n}\ {}^{n-1}C_{j-1}\left(\frac{1}{2}\right)^{n-1}$, but omitting the $j = n-1$ and the $j = n$ terms, and so is equal to

$$1 - \frac{(n-1)!}{2^{n-1}(n-2)!\,1!} - \frac{(n-1)!}{2^{n-1}(n-1)!\,0!} = \frac{1}{2^{n-1}}\left[2^{n-1} - (n-1) - 1\right].$$

Thus,

$$\Pr(A_k|B) = \frac{(n-1)!}{2^{n-1}(k-1)!(n-k)!}\,\frac{2^{n-1}}{2^{n-1}-n} = \frac{(n-1)!}{(2^{n-1}-n)(k-1)!(n-k)!},$$

as given in the question.

30.13 *A set of $2N+1$ rods consists of one of each integer length $1, 2, \ldots, 2N, 2N+1$. Three, of lengths a, b and c, are selected, of which a is the longest. By considering the possible values of b and c, determine the number of ways in which a non-degenerate triangle (i.e. one of non-zero area) can be formed (i) if a is even, and (ii) if a is odd. Combine these results appropriately to determine the total number of non-degenerate triangles that can be formed with the $2N+1$ rods, and hence show that the probability that such a triangle can be formed from a random selection (without replacement) of three rods is*

$$\frac{(N-1)(4N+1)}{2(4N^2-1)}.$$

Rod a is the longest of the three rods. As no two are the same length, let $a > b > c$. To form a non-degenerate triangle we require that $b + c > a$, and, in consequence, $4 \le a \le 2N+1$.

(i) With a even. Consider each b ($< a$) in turn and determine how many values of c allow a triangle to be made:

b	Values of c	Number of c values
$a-1$	$2, 3, \cdots, a-2$	$a-3$
$a-2$	$3, 4, \cdots, a-3$	$a-5$
$\cdots$	$\cdots$	$\cdots$
$\frac{1}{2}a+1$	$\frac{1}{2}a$	1

Thus, there are $1 + 3 + 5 + \cdots + (a-3)$ possible triangles when a is even.

(ii) A table for odd a is similar, except that the last line will read $b = \frac{1}{2}(a+3)$, $c = \frac{1}{2}(a-1)$ or $\frac{1}{2}(a+1)$, and the number of c values $= 2$. Thus there are $2 + 4 + 6 + \cdots + (a-3)$ possible triangles when a is odd.

To find the total number $n(N)$ of possible triangles, we group together the cases $a = 2m$ and $a = 2m+1$, where $m = 1, 2, \ldots, N$. Then,

$$\begin{aligned}
n(N) &= \sum_{m=2}^{N} [\,1 + 3 + \cdots + (2m-3)\,] + [\,2 + 4 + \cdots + (2m+1-3)\,] \\
&= \sum_{m=2}^{N} \sum_{k=1}^{2m-2} k = \sum_{m=2}^{N} \tfrac{1}{2}(2m-2)(2m-1) = \textstyle\sum_{m=2}^{N} 2m^2 - 3m + 1 \\
&= 2\left[\,\tfrac{1}{6}N(N+1)(2N+1) - 1\,\right] - 3\left[\,\tfrac{1}{2}N(N+1) - 1\,\right] + N - 1 \\
&= \frac{N}{6}[\,2(N+1)(2N+1) - 9(N+1) + 6\,] \\
&= \frac{N}{6}(4N^2 - 3N - 1) = \frac{N}{6}(4N+1)(N-1).
\end{aligned}$$

The number of ways that three rods can be drawn at random (without replacement) is $(2N+1)(2N)(2N-1)/3!$ and so the probability that they can form a triangle is

$$\frac{N(4N+1)(N-1)}{6}\,\frac{3!}{(2N+1)(2N)(2N-1)} = \frac{(N-1)(4N+1)}{2(4N^2-1)},$$

as stated in the question.

30.15 *The duration (in minutes) of a telephone call made from a public call-box is a random variable T. The probability density function of T is*

$$f(t) = \begin{cases} 0 & t < 0, \\ \frac{1}{2} & 0 \le t < 1, \\ ke^{-2t} & t \ge 1, \end{cases}$$

where k is a constant. To pay for the call, 20 *pence has to be inserted at the beginning, and a further* 20 *pence after each subsequent half-minute. Determine by how much the average cost of a call exceeds the cost of a call of average length charged at* 40 *pence per minute.*

From the normalisation of the PDF, we must have

$$1 = \int_0^{\infty} f(t)\,dt = \frac{1}{2} + \int_1^{\infty} ke^{-2t}\,dt = \frac{1}{2} + \frac{ke^{-2}}{2} \quad \Rightarrow \quad k = e^2.$$

The average length of a call is given by

$$\begin{aligned}\bar{t} &= \int_0^1 t\,\frac{1}{2}\,dt + \int_1^\infty t\,e^2 e^{-2t}\,dt \\ &= \frac{1}{2}\,\frac{1}{2} + \left[\frac{te^2 e^{-2t}}{-2}\right]_1^\infty + \int_1^\infty \frac{e^2\,e^{-2t}}{2}\,dt = \frac{1}{4} + \frac{1}{2} + \frac{e^2}{2}\left[\frac{e^{-2t}}{-2}\right]_1^\infty = \frac{3}{4} + \frac{1}{4} = 1.\end{aligned}$$

Let $p_n = \Pr\{\frac{1}{2}(n-1) < t < \frac{1}{2}n\}$. The corresponding cost is $c_n = 20n$. Clearly, $p_1 = p_2 = \frac{1}{4}$ and, for $n > 2$,

$$p_n = e^2 \int_{(n-1)/2}^{n/2} e^{-2t}\,dt = e^2 \left[\frac{e^{-2t}}{-2}\right]_{(n-1)/2}^{n/2} = \frac{1}{2}e^2(e-1)e^{-n}.$$

The average cost of a call is therefore

$$\bar{c} = 20\left[\frac{1}{4} + 2\,\frac{1}{4} + \sum_{n=3}^{\infty} n\frac{1}{2}e^2(e-1)e^{-n}\right] = 15 + 10e^2(e-1)\sum_{n=3}^{\infty} ne^{-n}.$$

Now, the final summation might be recognised as part of an arithmetico-geometric series whose sum can be found from the standard formula

$$S = \frac{a}{1-r} + \frac{rd}{(1-r)^2},$$

with $a = 0$, $d = 1$ and $r = e^{-1}$, or could be evaluated directly by noting that as a geometric series,

$$\sum_{n=0}^{\infty} e^{-nx} = \frac{1}{1-e^{-x}}.$$

Differentiating this with respect to x and then setting $x = 1$ gives

$$-\sum_{n=0}^{\infty} ne^{-nx} = -\frac{e^{-x}}{(1-e^{-x})^2} \quad\Rightarrow\quad \sum_{n=0}^{\infty} ne^{-n} = \frac{e^{-1}}{(1-e^{-1})^2}.$$

From either method it follows that

$$\begin{aligned}\sum_{n=3}^{\infty} ne^{-n} &= \frac{e}{(e-1)^2} - e^{-1} - 2e^{-2} \\ &= \frac{e - e + 2 - e^{-1} - 2 + 4e^{-1} - 2e^{-2}}{(e-1)^2} = \frac{3e^{-1} - 2e^{-2}}{(e-1)^2}.\end{aligned}$$

The total charge therefore exceeds that of a call of average length (1 minute) charged at 40 pence per minute by the amount (in pence)

$$15 + 10e^2(e-1)\frac{3e^{-1} - 2e^{-2}}{(e-1)^2} - 40 = \frac{10(3e-2) - 25e + 25}{e-1} = \frac{5e+5}{e-1} = 10.82.$$

30.17 *If the scores in a cup football match are equal at the end of the normal period of play, a 'penalty shoot-out' is held in which each side takes up to five shots (from the penalty spot) alternately, the shoot-out being stopped if one side acquires an unassailable lead (i.e. has a lead greater than its opponents have shots remaining). If the scores are still level after the shoot-out a 'sudden death' competition takes place.*

In sudden death each side takes one shot and the competition is over if one side scores and the other does not; if both score, or both fail to score, a further shot is taken by each side, and so on. Team 1, which takes the first penalty, has a probability p_1, which is independent of the player involved, of scoring and a probability q_1 $(= 1 - p_1)$ of missing; p_2 and q_2 are defined likewise.

Let $\Pr(i : x, y)$ *be the probability that team i has scored x goals after y attempts, and* $f(M)$ *be the probability that the shoot-out terminates after a* total *of M shots.*

(a) *Prove that the probability that 'sudden death' will be needed is*

$$f(11+) = \sum_{r=0}^{5} ({}^5C_r)^2 (p_1p_2)^r (q_1q_2)^{5-r}.$$

(b) *Give reasoned arguments (preferably without first looking at the expressions involved) which show that, for* $N = 3, 4, 5$,

$$f(M = 2N) = \sum_{r=0}^{2N-6} \left\{ \begin{array}{l} p_2 \Pr(1 : r, N) \Pr(2 : 5 - N + r, N - 1) \\ \quad + q_2 \Pr(1 : 6 - N + r, N) \Pr(2 : r, N - 1) \end{array} \right\}$$

and, for $N = 3, 4$,

$$f(M = 2N + 1) = \sum_{r=0}^{2N-5} \left\{ \begin{array}{l} p_1 \Pr(1 : 5 - N + r, N) \Pr(2 : r, N) \\ \quad + q_1 \Pr(1 : r, N) \Pr(2 : 5 - N + r, N). \end{array} \right\}$$

(c) *Give an explicit expression for* $\Pr(i : x, y)$ *and hence show that if the teams are so well matched that* $p_1 = p_2 = 1/2$ *then*

$$f(2N) = \sum_{r=0}^{2N-6} \left(\frac{1}{2^{2N}} \right) \frac{N!(N-1)!6}{r!(N-r)!(6-N+r)!(2N-6-r)!},$$

$$f(2N+1) = \sum_{r=0}^{2N-5} \left(\frac{1}{2^{2N}} \right) \frac{(N!)^2}{r!(N-r)!(5-N+r)!(2N-5-r)!}.$$

(d) *Evaluate these expressions to show that, expressing* $f(M)$ *in units of* 2^{-8},

M	6	7	8	9	10	11+
$f(M)$	8	24	42	56	63	63

Give a simple explanation of why $f(10) = f(11+)$.

(a) For 'sudden death' to be needed the scores must be equal after ten shots, five from each side. A score of r goals each has a probability

$$\left({}^5C_r\, p_1^r\, q_1^{5-r} \right) \times \left({}^5C_r\, p_2^r\, q_2^{5-r} \right),$$

and the total probability that the scores are equal after ten shots is obtained by summing this over all possible values of r ($r = 0, 1, \ldots, 5$). Thus

$$f(11+) = \sum_{r=0}^{5} ({}^5C_r)^2 (p_1 p_2)^r (q_1 q_2)^{5-r}.$$

(b) For the shoot-out to terminate after $2N$ shots (≤ 10 shots), one team must be $6 - N$ goals ahead and team 2 must just have taken the last shot.

(i) If team 1 won, it was because team 2 failed with their Nth shot and team 1 must have been $6 - N$ goals ahead before the final shot was taken. The probability for this is $q_2 \Pr(1 : 6 - N + r, N) \Pr(2 : r, N - 1)$.
(ii) If team 2 won, it must have been successful with its last shot and, before it, must have been $5 - N$ goals ahead. The probability for this is $p_2 \Pr(1 : r, N) \Pr(2 : 5 - N + r, N - 1)$.

This type of finish can only arise if $N > 5 - N$, i.e. $N = 3$, 4 or 5. Further, since in $\Pr(i : x, y)$ we must have $x \leq y$, the range for r is determined, from (i), by $6 - N + r \leq N$ and, from (ii), by $5 - N + r \leq N - 1$; these both give $0 \leq r \leq 2N - 6$. Thus

$$f(M = 2N) = \sum_{r=0}^{2N-6} \left\{ \begin{array}{l} p_2 \Pr(1 : r, N) \Pr(2 : 5 - N + r, N - 1) \\ \quad + q_2 \Pr(1 : 6 - N + r, N) \Pr(2 : r, N - 1) \end{array} \right\}.$$

For $M = 2N + 1$, the shoot-out terminates after team 1's $(N + 1)$th shot, which must have been successful if it wins, or unsuccessful if team 2 wins.

(i) If team 1 wins, it must now be $6 - N$ goals ahead, i.e. it was $5 - N$ goals ahead before its successful $(N + 1)$th shot. This has probability $p_1 \Pr(1 : 5 - N + r, N) \Pr(2 : r, N)$.
(i) If team 2 wins, it must have been $5 - N$ goals ahead, before team 1's unsuccessful $(N + 1)$th shot. The probability for this is $q_1 \Pr(1 : r, N) \Pr(2 : 5 - N + r, N)$.

This type of ending can only occur if $N > 5 - N$ and $2N + 1 \leq 10$, i.e. $N = 3$ or 4. Arguing as before, we see that both (i) and (ii) require $5 - N + r \leq N$, i.e. $0 \leq r \leq 2N - 5$. Thus

$$f(M = 2N + 1) = \sum_{r=0}^{2N-5} \left\{ \begin{array}{l} p_1 \Pr(1 : 5 - N + r, N) \Pr(2 : r, N) \\ \quad + q_1 \Pr(1 : r, N) \Pr(2 : 5 - N + r, N) \end{array} \right\}.$$

(c) As in part (a), $\Pr(i : x, y)$ is given by the binomial distribution as

$$\Pr(i : x, y) = {}^yC_x\, p_i^x\, q_i^{y-x}.$$

We now set $p_1 = p_2 = q_1 = q_2 = \frac{1}{2}$ and calculate

$$\begin{aligned}
f(2N) &= \sum_{r=0}^{2N-6} \left[\frac{1}{2}\, {}^{N}C_r \; {}^{N-1}C_{5-N+r} \left(\frac{1}{2}\right)^{N} \left(\frac{1}{2}\right)^{N-1} \right. \\
&\qquad\qquad \left. + \; \frac{1}{2}\, {}^{N}C_{6-N+r} \; {}^{N-1}C_r \left(\frac{1}{2}\right)^{N} \left(\frac{1}{2}\right)^{N-1} \right] \\
&= \sum_{r=0}^{2N-6} \left(\frac{1}{2^{2N}}\right) \left[\frac{N!\,(N-1)!}{r!\,(N-r)!\,(5-N+r)!\,(2N-6-r)!} \right. \\
&\qquad\qquad \left. + \; \frac{N!\,(N-1)!}{(6-N+r)!\,(2N-6-r)!\,r!\,(N-1-r)!} \right] \\
&= \sum_{r=0}^{2N-6} \left(\frac{1}{2^{2N}}\right) \frac{N!\,(N-1)!\,[6-N+r+N-r]}{r!\,(N-r)!\,(6-N+r)!\,(2N-6-r)!} \\
&= \sum_{r=0}^{2N-6} \left(\frac{1}{2^{2N}}\right) \frac{N!\,(N-1)!\,6}{r!\,(N-r)!\,(6-N+r)!\,(2N-6-r)!}.
\end{aligned}$$

The value of $f(2N+1)$ is found in a similar way. But, since $p_1 = p_2 = q_1 = q_2 = \frac{1}{2}$, the two terms contributing to it for any particular value of r are equal and each has the value

$$\frac{1}{2}\; {}^{N}C_{5-N+r} \left(\frac{1}{2}\right)^{N} \; {}^{N}C_r \left(\frac{1}{2}\right)^{N}.$$

When these terms are added and then summed over r we obtain

$$f(2N+1) = \sum_{r=0}^{2N-5} \left(\frac{1}{2^{2N}}\right) \frac{(N!)^2}{r!(N-r)!(5-N+r)!(2N-5-r)!}.$$

(d) Evaluating these expressions for the allowed values of N, that is 3, 4 and 5 for $f(2N)$, and 3 and 4 for $f(2N+1)$, is straightforward but somewhat tedious. The results, as given in the question, are

M	6	7	8	9	10	11+
$f(M)$	8	24	42	56	63	63

Here $f(M)$ is expressed in units of 2^{-8}. As expected, these probabilities add up to unity, and it can be seen that sudden death is needed in about one-quarter of such shoot-outs.

The equality of $f(10)$ and f(11+) is simply explained by the fact that, if the shoot-out has not been settled by then, team 2 is just as likely ($p_2 = \frac{1}{2}$) to take it into sudden death by scoring with its fifth shot as it is to lose it ($q_2 = \frac{1}{2}$) by missing.

30.19 *A continuous random variable X has a probability density function $f(x)$; the corresponding cumulative probability function is $F(x)$. Show that the random variable $Y = F(X)$ is uniformly distributed between 0 and 1.*

We first note that, as $F(x)$ is a cumulative probability density function, it has values $F(-\infty) = 0$ and $F(\infty) = 1$ and that $y = F(x)$ has a single-valued inverse $x = x(y)$.

With $Y = F(X)$, we have from the standard result for the distribution of single-valued inverse functions that

$$g(Y) = f(X(Y)) \left| \frac{dX}{dY} \right| .$$

However, in this particular case of Y being the cumulative probability function of X, we can evaluate $|dX/dY|$ more explicitly. This is because

$$\frac{dY}{dX} = \frac{d}{dX} F(X) = \frac{d}{dX} \int_{-\infty}^{X} f(u)\, du = f(X),$$

and is non-negative. So,

$$g(Y) = f(X(Y)) \left| \frac{dX}{dY} \right| = \frac{dY}{dX} \left| \frac{dX}{dY} \right| = 1.$$

This shows that Y is uniformly distributed on $(0, 1)$.

30.21 *This exercise is about interrelated binomial trials.*

(a) *In two sets of binomial trials T and t, the probabilities that a trial has a successful outcome are P and p, respectively, with corresponding probabilites of failure of $Q = 1 - P$ and $q = 1 - p$. One 'game' consists of a trial T, followed, if T is successful, by a trial t and then a further trial T. The two trials continue to alternate until one of the T-trials fails, at which point the game ends. The score S for the game is the total number of successes in the t-trials. Find the PGF for S and use it to show that*

$$E[S] = \frac{Pp}{Q}, \qquad V[S] = \frac{Pp(1 - Pq)}{Q^2}.$$

(b) *Two normal unbiased six-faced dice A and B are rolled alternately starting with A; if A shows a 6 the experiment ends. If B shows an odd number no points are scored, one point is scored for a 2 or a 4, and two points are awarded for a 6. Find the average and standard deviation of the score for the experiment and show that the latter is the greater.*

This is a situation in which the score for the game is a variable length sum, the length N being determined by the outcome of the T-trials. The probability that $N = n$ is given by $h_n = P^n Q$, since n T-trials must succeed and then be followed by a failing T-trial. Thus the PGF for the length of each 'game' is given by

$$\chi_N(t) \equiv \sum_{n=0}^{\infty} h_n t^n = \sum_{n=0}^{\infty} P^n Q t^n = \frac{Q}{1-Pt}.$$

For each permitted Bernoulli t-trial, $X_i = 1$ with probability p and $X_i = 0$ with probability q; its PGF is thus $\Phi_X(t) = q + pt$. The score for the game is $S = \sum_{i=1}^{N} X_i$ and its PGF is given by the compound function

$$\begin{aligned}\Xi_S(t) &= \chi_N(\Phi_X(t)) \\ &= \frac{Q}{1-P(q+pt)},\end{aligned}$$

in which the PGF for a single t-trial forms the argument of the PGF for the length of each 'game'.

It follows that the mean of S is found from

$$\Xi_S'(t) = \frac{QPp}{(1-Pq-Ppt)^2} \quad \Rightarrow \quad E[\,S\,] = \Xi_S'(1) = \frac{QPp}{(1-P)^2} = \frac{Pp}{Q}.$$

To calculate the variance of S we need to find $\Xi_S''(1)$. This second derivative is

$$\Xi_S''(t) = \frac{2QP^2p^2}{(1-Pq-Ppt)^3} \quad \Rightarrow \quad \Xi_S''(1) = \frac{2P^2p^2}{Q^2}.$$

The variance is therefore

$$\begin{aligned}V[\,S\,] &= \Xi_S''(1) + \Xi_S'(1) - [\,\Xi_S'(1)\,]^2 \\ &= \frac{2P^2p^2}{Q^2} + \frac{Pp}{Q} - \frac{P^2p^2}{Q^2} \\ &= \frac{Pp(Pp+Q)}{Q^2} = \frac{Pp(P-Pq+Q)}{Q^2} = \frac{Pp(1-Pq)}{Q^2}.\end{aligned}$$

(b) For die A: $P = \frac{5}{6}$ and $Q = \frac{1}{6}$ giving $\chi_N(t) = 1/(6-5t)$.

For die B: $\Pr(X = 0) = \frac{3}{6}$, $\Pr(X = 1) = \frac{2}{6}$ and $\Pr(X = 2) = \frac{1}{6}$ giving $\Phi_X(t) = (3 + 2t + t^2)/6$.

The PGF for the game score S is thus

$$\Xi_S(t) = \frac{1}{6 - \frac{5}{6}(3+2t+t^2)} = \frac{6}{21-10t-5t^2}.$$

We need to evaluate the first two derivatives of $\Xi_S(t)$ at $t = 1$, as follows:

$$\Xi_S'(t) = \frac{-6(-10-10t)}{(21-10t-5t^2)^2} = \frac{60+60t}{(21-10t-5t^2)^2}$$

$$\Rightarrow \quad E[S] = \Xi_S'(1) = \frac{120}{6^2} = \frac{10}{3} = 3.33,$$

$$\Xi_S''(t) = \frac{60}{(21-10t-5t^2)^2} - \frac{2(60+60t)(-10-10t)}{(21-10t-5t^2)^3}$$

$$\Rightarrow \quad \Xi_S''(1) = \frac{60}{36} - \frac{2(120)(-20)}{(6)^3} = \frac{215}{9}.$$

Substituting the calculated values gives $V[S]$ as

$$V[S] = \frac{215}{9} + \frac{10}{3} - \left(\frac{10}{3}\right)^2 = \frac{145}{9},$$

from which it follows that

$$\sigma_S = \sqrt{V[S]} = 4.01, \text{ i.e. greater than the mean.}$$

30.23 *A point P is chosen at random on the circle $x^2 + y^2 = 1$. The random variable X denotes the distance of P from $(1, 0)$. Find the mean and variance of X and the probability that X is greater than its mean.*

With O as the centre of the unit circle and Q as the point $(1, 0)$, let OP make an angle θ with the x-axis OQ. The random variable X then has the value $2\sin(\theta/2)$ with θ uniformly distributed on $(0, 2\pi)$, i.e.

$$f(x)\,dx = \frac{1}{2\pi}\,d\theta.$$

The mean of X is given straightforwardly by

$$\langle X\rangle = \int_0^2 Xf(x)\,dx = \int_0^{2\pi} 2\sin\left(\frac{\theta}{2}\right)\frac{1}{2\pi}\,d\theta = \frac{1}{\pi}\left[-2\cos\frac{\theta}{2}\right]_0^{2\pi} = \frac{4}{\pi}.$$

For the variance we have

$$\sigma_X^2 = \langle X^2\rangle - \langle X\rangle^2 = \int_0^{2\pi} 4\sin^2\left(\frac{\theta}{2}\right)\frac{1}{2\pi}\,d\theta - \frac{16}{\pi^2} = \frac{4}{2\pi}\,\frac{1}{2}\,2\pi - \frac{16}{\pi^2} = 2 - \frac{16}{\pi^2}.$$

When $X = \langle X\rangle = 4/\pi$, the angle $\theta = 2\sin^{-1}(2/\pi)$ and so

$$\Pr(X > \langle X\rangle) = \frac{2\pi - 4\sin^{-1}\dfrac{2}{\pi}}{2\pi} = 0.561.$$

30.25 *The number of errors needing correction on each page of a set of proofs follows a Poisson distribution of mean μ. The cost of the first correction on any page is α and that of each subsequent correction on the same page is β. Prove that the average cost of correcting a page is*

$$\alpha + \beta(\mu - 1) - (\alpha - \beta)e^{-\mu}.$$

Since the number of errors on a page is Poisson distributed, the probability of n errors on any particular page is

$$\Pr(n \text{ errors}) = p_n = e^{-\mu}\,\frac{\mu^n}{n!}.$$

The average cost per page, found by averaging the corresponding cost over all values of n, is

$$\begin{aligned} c &= 0\,p_0 + \alpha p_1 + \sum_{n=2}^{\infty}[\,\alpha + (n-1)\beta\,]p_n \\ &= \alpha\mu e^{-\mu} + (\alpha - \beta)\sum_{n=2}^{\infty} p_n + \beta\sum_{n=2}^{\infty} np_n. \end{aligned}$$

Now, $\sum_{n=0}^{\infty} p_n = 1$ and, for a Poisson distribution, $\sum_{n=0}^{\infty} np_n = \mu$. These can be used to evaluate the above, once the $n = 0$ and $n = 1$ terms have been removed. Thus

$$\begin{aligned} c &= \alpha\mu e^{-\mu} + (\alpha - \beta)(1 - e^{-\mu} - \mu e^{-\mu}) + \beta(\mu - 0 - \mu e^{-\mu}) \\ &= \alpha + \beta(\mu - 1) + e^{-\mu}(\alpha\mu - \alpha + \beta - \mu\alpha + \mu\beta - \mu\beta) \\ &= \alpha + \beta(\mu - 1) + e^{-\mu}(\beta - \alpha), \end{aligned}$$

as given in the question.

30.27 *Show that for large r the value at the maximum of the PDF for the gamma distribution of order r with parameter λ is approximately $\lambda/\sqrt{2\pi(r-1)}$.*

The gamma distribution takes the form

$$f(x) = \frac{\lambda}{\Gamma(r)}\,(\lambda x)^{r-1}e^{-\lambda x}$$

and its maximum will occur when $y(x) = x^{(r-1)}e^{-\lambda x}$ is maximal. This requires

$$0 = \frac{dy}{dx} = (r-1)x^{(r-2)}e^{-\lambda x} - \lambda x^{(r-1)}e^{-\lambda x} \quad \Rightarrow \quad \lambda x = r - 1.$$

The maximum value is thus

$$\gamma_{\max}(r) = \frac{\lambda}{\Gamma(r)}\,(r-1)^{(r-1)}\,e^{-(r-1)}.$$

Now, using Stirling's approximation,

$$\Gamma(n+1) = n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \quad \text{for large } n,$$

we obtain

$$\begin{aligned}\gamma_{\max}(r) &\approx \frac{\lambda}{\sqrt{2\pi(r-1)}}\,\frac{e^{(r-1)}}{(r-1)^{(r-1)}}\,(r-1)^{(r-1)}\,e^{-(r-1)}\\ &= \frac{\lambda}{\sqrt{2\pi(r-1)}}.\end{aligned}$$

30.29 *The probability distribution for the number of eggs in a clutch is Po(λ), and the probability that each egg will hatch is p (independently of the size of the clutch). Show by direct calculation that the probability distribution for the number of chicks that hatch is Po(λp).*

Clearly, to determine the probability that a clutch produces k chicks, we must consider clutches of size n, for all $n \geq k$, and for each such clutch find the probability that exactly k of the n chicks do hatch. We then average over all n, weighting the results according to the distribution of n.

The probability that k chicks hatch from a clutch of size n is ${}^nC_k p^k q^{n-k}$, where $q = 1-p$. The probability that the clutch is of size n is $e^{-\lambda}\lambda^n/n!$. Consequently, the overall probability of k chicks hatching from a clutch is

$$\begin{aligned}\Pr(k \text{ chicks}) &= \sum_{n=k}^{\infty} e^{-\lambda}\,\frac{\lambda^n}{n!}\;{}^nC_k\,p^k\,q^{n-k}\\ &= e^{-\lambda}p^k\lambda^k\sum_{n=k}^{\infty}\frac{(\lambda q)^{n-k}}{n!}\,\frac{n!}{k!\,(n-k)!}, \quad \text{set } n-k=m,\\ &= e^{-\lambda}\frac{(\lambda p)^k}{k!}\sum_{m=0}^{\infty}\frac{(\lambda q)^m}{m!}\\ &= e^{-\lambda}\,\frac{(\lambda p)^k}{k!}\,e^{\lambda q}\\ &= \frac{e^{-\lambda p}(\lambda p)^k}{k!},\end{aligned}$$

since $q = 1-p$. Thus Pr(k chicks) is distributed as a Poisson distribution with parameter $\mu = \lambda p$.

30.31 *Under EU legislation on harmonisation, all kippers are to weigh* 0.2000 kg *and vendors who sell underweight kippers must be fined by their government. The weight of a kipper is normally distributed with a mean of* 0.2000 kg *and a standard deviation of* 0.0100 kg. *They are packed in cartons of* 100 *and large quantities of them are sold.*

Every day a carton is to be selected at random from each vendor and tested according to one of the following schemes, which have been approved for the purpose.

(a) *The entire carton is weighed and the vendor is fined* 2500 *euros if the average weight of a kipper is less than* 0.1975 kg.
(b) *Twenty-five kippers are selected at random from the carton; the vendor is fined* 100 *euros if the average weight of a kipper is less than* 0.1980 kg.
(c) *Kippers are removed one at a time, at random, until one has been found that weighs* more *than* 0.2000 kg*; the vendor is fined* $4n(n-1)$ *euros, where* n *is the number of kippers removed.*

Which scheme should the Chancellor of the Exchequer be urging his government to adopt?

For these calculations we measure weights in grammes.

(a) For this scheme we have a normal distribution with mean $\mu = 200$ and s.d. $\sigma = 10$. The s.d. for a carton is $\sqrt{100}\,\sigma = 100$ and the mean weight is 20000. There is a penalty if the weight of a carton is less than 19750. This critical value represents a standard variable of

$$Z = \frac{19750 - 20000}{100} = -2.5.$$

The probability that $Z < -2.5 = 1 - \Phi(2.5) = 1 - 0.9938 = 0.0062$. Thus the average fine per carton tested on this scheme is $0.0062 \times 2500 = 15.5$ euros.

(b) For this scheme the general parameters are the same but the mean weight of the sample measured is 5000 and its s.d is $\sqrt{25}\,(10) = 50$. The Z-value at which a fine is imposed is

$$Z = \frac{(198 \times 25) - 5000}{50} = -1.$$

The probability that $Z < -1.0 = 1 - \Phi(1.0) = 1 - 0.8413 = 0.1587$. Thus the average fine per carton tested on this scheme is $0.1587 \times 100 = 15.9$ euros.

(c) This scheme is a series of Bernoulli trials in which the probability of success is $\frac{1}{2}$ (since half of all kippers weigh more than 200 and the distribution is normal). The probability that it will take n kippers to find one that passes the test is

$q^{n-1}p = (\frac{1}{2})^n$. The expected fine is therefore

$$f = \sum_{n=2}^{\infty} 4n(n-1)\left(\frac{1}{2}\right)^n = 4\,\frac{2\,(\frac{1}{4})}{(\frac{1}{2})^3} = 16 \text{ euros.}$$

The expression for the sum was found by twice differentiating the sum of the geometric series $\sum r^n$ with respect to r, as follows:

$$\begin{aligned}\sum_{n=0}^{\infty} r^n = \frac{1}{1-r} \quad &\Rightarrow \quad \sum_{n=1}^{\infty} nr^{n-1} = \frac{1}{(1-r)^2}\\ &\Rightarrow \quad \sum_{n=2}^{\infty} n(n-1)r^{n-2} = \frac{2}{(1-r)^3}\\ &\Rightarrow \quad \sum_{n=2}^{\infty} n(n-1)r^{n} = \frac{2r^2}{(1-r)^3}.\end{aligned}$$

There is, in fact, little to choose between the schemes on monetary grounds; no doubt political considerations, such as the current unemployment rate, will decide!

30.33 *A practical-class demonstrator sends his twelve students to the storeroom to collect apparatus for an experiment, but forgets to tell each which type of component to bring. There are three types, A, B and C, held in the stores (in large numbers) in the proportions 20%, 30% and 50%, respectively, and each student picks a component at random. In order to set up one experiment, one unit each of A and B and two units of C are needed. Let* $\Pr(N)$ *be the probability that at least N experiments can be setup.*

(a) *Evaluate* $\Pr(3)$.
(b) *Find an expression for* $\Pr(N)$ *in terms of k_1 and k_2, the numbers of components of types A and B, respectively, selected by the students. Show that* $\Pr(2)$ *can be written in the form*

$$\Pr(2) = (0.5)^{12} \sum_{i=2}^{6} {}^{12}C_i\ (0.4)^i \sum_{j=2}^{8-i} {}^{12-i}C_j\ (0.6)^j.$$

(c) *By considering the conditions under which no experiments can be set up, show that* $\Pr(1) = 0.9145$.

(a) To make three experiments possible the twelve components picked must be three each of A and B and six of C. The probability of this is given by the multinomial distribution as

$$\Pr(3) = \frac{(12)!}{3!\,3!\,6!}\,(0.2)^3(0.3)^3(0.5)^6 = 0.06237.$$

(b) Let the numbers of A, B and C selected be k_1, k_2 and k_3, respectively, and consider when *at least* N experiments can be set up. We have the obvious inequalities $k_1 \geq N$, $k_2 \geq N$ and $k_3 \geq 2N$. In addition $k_3 = 12 - k_1 - k_2$, implying that $k_2 \leq 12 - 2N - k_1$. Further, k_1 cannot be greater than $12 - 3N$ if at least N experiments are to be set up, as each requires three other components that are not of type A. These inequalities set the limits on the acceptable values of k_1 and k_2 (k_3 is not a third independent variable). Thus $\Pr(N)$ is given by

$$\sum_{k_1 \geq N}^{12-3N} \sum_{k_2 \geq N}^{12-2N-k_1} \frac{(12)!}{k_1!\,k_2!\,(12-k_1-k_2)!}\,(0.2)^{k_1}\,(0.3)^{k_2}\,(0.5)^{12-k_1-k_2}.$$

The answer to part (a) is a particular case of this with $N = 3$, when each summation reduces to a single term.

For $N = 2$ the expression becomes

$$\begin{aligned}
\Pr(2) &= \sum_{k_1 \geq 2}^{6} \sum_{k_2 \geq 2}^{8-k_1} \frac{(12)!}{k_1!\,k_2!\,(12-k_1-k_2)!}\,(0.2)^{k_1}\,(0.3)^{k_2}\,(0.5)^{12-k_1-k_2} \\
&= (0.5)^{12} \sum_{i=2}^{6} \sum_{j=2}^{8-i} \frac{(12)!\,(0.2/0.5)^i}{i!\,(12-i)!}\,\frac{(12-i)!\,(0.3/0.5)^j}{j!\,(12-i-j)!} \\
&= (0.5)^{12} \sum_{i=2}^{6} {}^{12}C_i\,(0.4)^i \sum_{j=2}^{8-i} {}^{12-i}C_j\,(0.6)^j.
\end{aligned}$$

(c) No experiment can be set up if any one of the following four events occurs: $A_1 = (k_1 = 0)$, $A_2 = (k_2 = 0)$, $A_3 = (k_3 = 0)$ and $A_4 = (k_3 = 1)$. The probability for the union of these four events is given by

$$\Pr(A_1 \cup A_2 \cup A_3 \cup A_4) = \sum_{i=1}^{4} \Pr(A_i) - \sum_{i,j} \Pr(A_i \cap A_j) + \cdots.$$

The probabilities $\Pr(A_i)$ are straightforward to calculate as follows:

$$\begin{aligned}
\Pr(A_1) &= (1-0.2)^{12}, &\quad \Pr(A_2) &= (1-0.3)^{12}, \\
\Pr(A_3) &= (1-0.5)^{12}, &\quad \Pr(A_4) &= {}^{12}C_1(1-0.5)^{12}(0.5).
\end{aligned}$$

The calculation of the probability for the intersection of two events is typified by

$$\begin{aligned}
&\Pr(A_1 \cap A_2) = [\,1 - (0.2 + 0.3)\,]^{12} \\
\text{and }\ &\Pr(A_1 \cap A_4) = {}^{12}C_1[\,1 - (0.2 + 0.5)\,]^{11}(0.5)^1.
\end{aligned}$$

A few trial evaluations show that these are of order 10^{-4} and can be ignored by comparison with the larger terms in the first sum, which are (after rounding)

$$\sum_{i=1}^{4} \Pr(A_i) = (0.8)^{12} + (0.7)^{12} + (0.5)^{12} + 12(0.5)^{11}(0.5)$$
$$= 0.0687 + 0.0138 + 0.0002 + 0.0029 = 0.0856.$$

Since the probability of no experiments being possible is 0.0856, it follows that $\Pr(1) = 0.9144$.

> ***30.35*** *The continuous random variables X and Y have a joint PDF proportional to $xy(x-y)^2$ with $0 \le x \le 1$ and $0 \le y \le 1$. Find the marginal distributions for X and Y and show that they are negatively correlated with correlation coefficient $-\frac{2}{3}$.*

This PDF is clearly symmetric between x and y. We start by finding its normalisation constant c:

$$\int_0^1 \int_0^1 c(x^3y - 2x^2y^2 + xy^3)\,dx\,dy = c\left(\frac{1}{4}\frac{1}{2} - 2\frac{1}{3}\frac{1}{3} + \frac{1}{2}\frac{1}{4}\right) = \frac{c}{36}.$$

Thus, we must have that $c = 36$.

The marginal distribution for x is given by

$$\begin{aligned} f(x) &= 36\int_0^1 (x^3y - 2x^2y^2 + xy^3)\,dy \\ &= 36(\tfrac{1}{2}x^3 - \tfrac{2}{3}x^2 + \tfrac{1}{4}x) \\ &= 18x^3 - 24x^2 + 9x, \end{aligned}$$

and the mean of x by

$$\mu_X = \bar{x} = \int_0^1 (18x^4 - 24x^3 + 9x^2)\,dx = \frac{18}{5} - \frac{24}{4} + \frac{9}{3} = \frac{3}{5}.$$

By symmetry, the marginal distribution and the mean for y are $18y^3 - 24y^2 + 9y$ and $\frac{3}{5}$, repectively.

To calculate the correlation coefficient we also need the variances of x and y and their covariance. The variances, obviously equal, are given by

$$\begin{aligned} \sigma_X^2 &= \int_0^1 x^2(18x^3 - 24x^2 + 9x)\,dx - (\tfrac{3}{5})^2 \\ &= \frac{18}{6} - \frac{24}{5} + \frac{9}{4} - \frac{9}{25} \\ &= \frac{900 - 1440 + 675 - 108}{300} = \frac{9}{100}. \end{aligned}$$

The standard deviations σ_X and σ_Y are therefore both equal to 3/10.

The covariance is calculated next; it is given by

$$\begin{aligned}
\text{Cov}[X, Y] &= \langle XY \rangle - \mu_X \mu_Y \\
&= 36 \int_0^1 \int_0^1 (x^4 y^2 - 2x^3 y^3 + x^2 y^4)\, dx\, dy - \frac{3}{5}\frac{3}{5} \\
&= \frac{36}{5 \times 3} - \frac{72}{4 \times 4} + \frac{36}{3 \times 5} - \frac{9}{25} \\
&= \frac{12}{5} - \frac{9}{2} + \frac{12}{5} - \frac{9}{25} \\
&= \frac{120 - 225 + 120 - 18}{50} = -\frac{3}{50}.
\end{aligned}$$

Finally,

$$\text{Corr}[X, Y] = \frac{\text{Cov}[X, Y]}{\sigma_X\, \sigma_Y} = \frac{-\frac{3}{50}}{\frac{3}{10}\,\frac{3}{10}} = -\frac{2}{3}.$$

30.37 *Two continuous random variables X and Y have a joint probability distribution*

$$f(x, y) = A(x^2 + y^2),$$

where A is a constant and $0 \le x \le a$, $0 \le y \le a$. *Show that X and Y are negatively correlated with correlation coefficient* $-15/73$. *By sketching a rough contour map of* $f(x, y)$ *and marking off the regions of positive and negative correlation, convince yourself that this (perhaps counter-intuitive) result is plausible.*

The calculations of the various parameters of the distribution are straightforward (see exercise 30.35). The parameter A is determined by the normalisation condition:

$$1 = \int_0^a \int_0^a A(x^2 + y^2)\, dx\, dy = A\left(\frac{a^4}{3} + \frac{a^4}{3}\right) \quad \Rightarrow \quad A = \frac{3}{2a^4}.$$

The two expectation values required are given by

$$\begin{aligned}
E[X] &= \int_0^a \int_0^a Ax(x^2 + y^2)\, dx\, dy \\
&= \frac{3}{2a^4}\left(\frac{a^5}{4 \times 1} + \frac{a^5}{2 \times 3}\right) = \frac{5a}{8}, \qquad (E[Y] = E[X]), \\
E[X^2] &= \int_0^a \int_0^a Ax^2(x^2 + y^2)\, dx\, dy \\
&= \frac{3}{2a^4}\left(\frac{a^6}{5 \times 1} + \frac{a^6}{3 \times 3}\right) = \frac{7a^2}{15}.
\end{aligned}$$

Hence the variance, calculated from the general result $V[X] = E[X^2] - (E[X])^2$, is

$$V[X] = \frac{7a^2}{15} - \left(\frac{5a}{8}\right)^2 = \frac{73}{960}a^2,$$

and the standard deviations are given by

$$\sigma_X = \sigma_Y = \sqrt{\frac{73}{960}}\,a.$$

To obtain the correlation coefficient we need also to calculate the following:

$$\begin{aligned} E[XY] &= \int_0^a \int_0^a Axy(x^2 + y^2)\,dx\,dy \\ &= \frac{3}{2a^4}\left(\frac{a^6}{4 \times 2} + \frac{a^6}{2 \times 4}\right) = \frac{3a^2}{8}. \end{aligned}$$

Then the covariance, given by $\text{Cov}[X, Y] = E[XY] - E[X]E[Y]$, is evaluated as

$$\text{Cov}[X, Y] = \frac{3}{8}a^2 - \frac{5a}{8}\frac{5a}{8} = -\frac{a^2}{64}.$$

Combining this last result with the standard deviations calculated above, we then obtain

$$\text{Corr}[X, Y] = \frac{-(a^2/64)}{\sqrt{\frac{73}{960}}\,a\,\sqrt{\frac{73}{960}}\,a} = -\frac{15}{73}.$$

As the means of both X and Y are $\frac{5}{8}a = 0.62a$, the areas of the square of side a for which $X - \mu_X$ and $Y - \mu_Y$ have the same sign (i.e. regions of positive correlation) are about $(0.62)^2 \approx 39\%$ and $(0.38)^2 \approx 14\%$ of the total area of the square. The regions of negative correlation occupy some 47% of the square.

However, $f(x, y) = A(x^2 + y^2)$ favours the regions where one or both of x and y are large and close to unity. Broadly speaking, this gives little weight to the region in which both X and Y are less than their means, and so, although it is the largest region in area, it contributes relatively little to the overall correlation. The two (equal area) regions of negative correlation together outweigh the smaller high probability region of positive correlation in the top right-hand corner of the square; the overall result is a net negative correlation coefficient.

30.39 *Show that, as the number of trials n becomes large but $np_i = \lambda_i$, $i = 1, 2, \ldots, k-1$, remains finite, the multinomial probability distribution,*

$$M_n(x_1, x_2, \ldots, x_k) = \frac{n!}{x_1!x_2!\cdots x_k!} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k},$$

can be approximated by a multiple Poisson distribution with $k-1$ factors:

$$M'_n(x_1, x_2, \ldots, x_{k-1}) = \prod_{i=1}^{k-1} \frac{e^{-\lambda_i}\lambda_i^{x_i}}{x_i!}.$$

(Write $\sum_i^{k-1} p_i = \delta$ and express all terms involving subscript k in terms of n and δ, either exactly or approximately. You will need to use $n! \approx n^{\epsilon}[(n-\epsilon)!]$ and $(1-a/n)^n \approx e^{-a}$ for large n.)

(a) *Verify that the terms of M'_n add up to unity when summed over all possible values of the random variables $x_1, x_2, \ldots, x_{k-1}$.*

(b) *If $k = 7$ and $\lambda_i = 9$ for all $i = 1, 2, \ldots, 6$, estimate, using the appropriate Gaussian approximation, the chance that at least three of $x_1, x_2, \ldots, x_6$ will be 15 or greater.*

The probabilities p_i are not all independent, and $p_k = 1 - \sum' p_i$, where, for compactness and typographical clarity, we denote $\sum_{i=1}^{k-1}$ by $\sum'$. We further write $\sum' p_i$ as δ. In the same way, we denote $\sum' x_i$ by ϵ and can write $x_k = n - \epsilon$.

Now, as $n \to \infty$ with $p_i \to 0$, whilst the product np_i remains finite and equal to λ_i, we will have that $\delta \to 0$, $n\delta \to \sum' \lambda_i$ and $(n-\epsilon)/n \to 1$. Making these replacements in the factors that contain subscript k gives

$$\begin{aligned}
M_n(x_1, x_2, \ldots, x_k) & \\
&= \frac{n!}{x_1!\,x_2!\cdots x_{k-1}!\,(n-\epsilon)!}\, p_1^{x_1}\, p_2^{x_2} \cdots p_{k-1}^{x_{k-1}}\,(1-\delta)^{n-\epsilon} \\
&\approx \frac{n^{\epsilon}\,(n-\epsilon)!}{x_1!\,x_2!\cdots x_{k-1}!\,(n-\epsilon)!}\, p_1^{x_1}\, p_2^{x_2} \cdots p_{k-1}^{x_{k-1}} \left(1 - \frac{n\delta}{n}\right)^{n-\epsilon} \\
&= \frac{n^{x_1+x_2+\cdots+x_{k-1}}}{x_1!\,x_2!\cdots x_{k-1}!}\, p_1^{x_1}\, p_2^{x_2} \cdots p_{k-1}^{x_{k-1}} \left(1 - \frac{n\delta}{n}\right)^{n-\epsilon} \\
&\to \frac{\lambda_1^{x_1}\,\lambda_2^{x_2}\cdots\lambda_{k-1}^{x_{k-1}}}{x_1!\,x_2!\cdots x_{k-1}!}\, e^{-(\lambda_1+\lambda_2+\cdots+\lambda_{k-1})} \\
&= \prod_{i=1}^{k-1} \frac{e^{-\lambda_i}\lambda_i^{x_i}}{x_i!},
\end{aligned}$$

i.e. as $n \to \infty$ $M_n(x_1, x_2, \ldots, x_k)$ can be approximated by the direct product of $k-1$ separate Poisson distributions.

(a) Since the modified expression $M_n'(x_1, x_2, \ldots, x_{k-1})$ consists of this multiple product of factors, the summation between 0 and ∞ over any particular variable, x_j say, can be carried out separately, with the factors not involving x_j treated as constant multipliers. A typical sum is

$$\sum_{x_j=0}^{\infty} \frac{e^{-\lambda_j} \lambda_j^{x_j}}{x_j!} = e^{-\lambda_j} e^{\lambda_j} = 1.$$

When all the summations have been carried out,

$$\sum_{\text{all } x_i} M_n'(x_1, x_2, \ldots, x_{k-1}) = (1)^{k-1} = 1.$$

(b) The Gaussian approximation to each Poisson distribution Po(9) is $N(9, 9)$, for which the standard variable is given by

$$Z = \frac{X - 9}{\sqrt{9}}.$$

Thus the probability that one of the x_i will exceed 15 (after including a continuity correction) is

$$\Pr(x_i \geq 15) = \Pr\left(Z > \frac{14.5 - 9}{3}\right) = 1 - \Phi(1.833) = 1 - 0.966 = 0.0334.$$

That (any) three of them should do so has probability

$$^6C_3(0.0334)^3 = 20 \times 3.726\ 10^{-5} = 7.5 \times 10^{-4}.$$

The probabilities that 4, 5 or 6 of the x_i will exceed 15 make negligible additions to this, which is already an approximation in any case.

31

Statistics

> **31.1** *A group of students uses a pendulum experiment to measure g, the acceleration of free fall, and obtains the following values (in* $\mathrm{m\,s^{-2}}$*):* 9.80, 9.84, 9.72, 9.74, 9.87, 9.77, 9.28, 9.86, 9.81, 9.79, 9.82. *What would you give as the best value and standard error for g as measured by the group*?

We first note that the reading of 9.28 m s^{-2} is so far from the others that it is almost certainly in error and should not be used in the calculation. The mean of the ten remaining values is 9.802 and the standard deviation of the sample about its mean is 0.04643. After including Bessel's correction factor, the estimate of the population s.d. is $\sigma = 0.0489$, leading to a s.d. in the measured value of the mean of $0.0489/\sqrt{10} = 0.0155$. We therefore give the best value and standard error for g as 9.80 ± 0.02 m s^{-2}.

> **31.3** *The following are the values obtained by a class of* 14 *students when measuring a physical quantity x:* 53.8, 53.1, 56.9, 54.7, 58.2, 54.1, 56.4, 54.8, 57.3, 51.0, 55.1, 55.0, 54.2, 56.6.
>
> (a) *Display these results as a histogram and state what you would give as the best value for x.*
> (b) *Without calculation, estimate how much reliance could be placed upon your answer to* (a).
> (c) *Databooks give the value of x as* 53.6 *with negligible error. Are the data obtained by the students in conflict with this*?

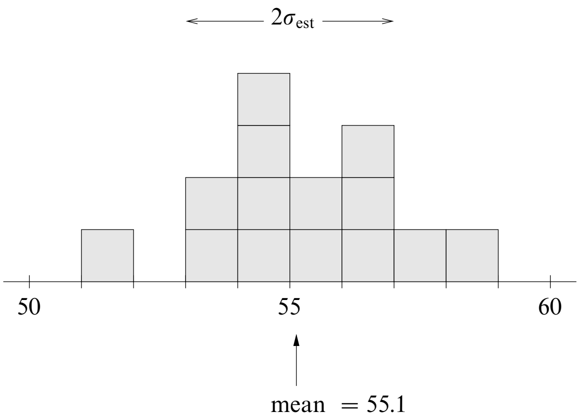

Figure 31.1 Histogram of the data in exercise 31.3.

(a) The histogram in figure 31.1 shows no reading that is an obvious mistake and there is no reason to suppose other than a Gaussian distribution. The best value for x is the arithmetic mean of the fourteen values given, i.e. 55.1.

(b) We note that eleven values, i.e. approximately two-thirds of the fourteen readings, lie within ± 2 bins of the mean. This estimates the s.d for the population as 2.0 and gives a standard error in the mean of $\approx 2.0/\sqrt{14} \approx 0.6$.

(c) Within the accuracy we are likely to achieve by estimating σ for the sample by eye, the value of Student's t is $(55.1 - 53.6)/0.6$, i.e. about 2.5. With fourteen readings there are 13 degrees of freedom. From standard tables for the Student's t-test, $C_{13}(2.5) \approx 0.985$. It is therefore likely at the $2 \times 0.015 = 3\%$ significance level that the data are in conflict with the accepted value.

[Numerical analysis of the data, rather than a visual estimate, gives the lower value 0.51 for the standard error in the mean and implies that there is a conflict between the data and the accepted value at the 1.0% significance level.]

31.5 *Measured quantities x and y are known to be connected by the formula*

$$y = \frac{ax}{x^2 + b},$$

where a and b are constants. Pairs of values obtained experimentally are

x:	2.0	3.0	4.0	5.0	6.0
y:	0.32	0.29	0.25	0.21	0.18.

Use these data to make best estimates of the values of y that would be obtained for (a) $x = 7.0$, *and* (b) $x = -3.5$. *As measured by fractional error, which estimate is likely to be the more accurate?*

In order to use this limited data to best advantage when estimating a and b graphically, the equation needs to be arranged in the linear form $v = mu + c$, since a straight-line graph is much the easiest form from which to extract parameters. The given equation can be arranged as

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

which is represented by a line with slope a^{-1} and intercept b/a when x^2 is used as the independent variable and x/y as the dependent one. The required tabulation is:

x	2.0	3.0	4.0	5.0	6.0
y	0.32	0.29	0.25	0.21	0.18
x^2	4.0	9.0	16.0	25.0	36.0
x/y	6.25	10.34	16.00	23.81	33.33

Plotting these data as a graph for $0 \leq x^2 \leq 40$ produces a straight line (within normal plotting accuracy). The line has a slope

$$\frac{1}{a} = \frac{28.1 - 2.7}{30.0 - 0.0} = 0.847 \quad \Rightarrow \quad a = 1.18.$$

The intercept is at $x/y = 2.7$, and, as this is equal to b/a, it follows that $b = 2.7 \times 1.18 = 3.2$. In fractional terms this is not likely to be very accurate as $b \ll x^2$ for all but two of the x-values used.

(a) For $x = 7.0$, the estimated value of y is

$$y = \frac{1.18 \times 7.0}{49.0 + 3.2} = 0.158.$$

(b) For $x = -3.5$, the estimated value of y is

$$y = \frac{1.18 \times (-3.5)}{12.25 + 3.2} = -0.267.$$

Although as a graphical extrapolation estimate (b) is further removed from the measured values, it is likely to be the more accurate because, using the fact that $y(-x) = -y(x)$, it is effectively obtained by (visual) *inter*polation amongst measured data rather than by *extra*polation from it.

31.7 *A population contains individuals of k types in equal proportions. A quantity X has mean μ_i amongst individuals of type i and variance σ^2, which has the same value for all types. In order to estimate the mean of X over the whole population, two schemes are considered; each involves a total sample size of nk. In the first the sample is drawn randomly from the whole population, whilst in the second* (stratified sampling) *n individuals are randomly selected from each of the k types.*

Show that in both cases the estimate has expectation

$$\mu = \frac{1}{k}\sum_{i=1}^{k}\mu_i,$$

but that the variance of the first scheme exceeds that of the second by an amount

$$\frac{1}{k^2 n}\sum_{i=1}^{k}(\mu_i - \mu)^2.$$

(i) For the first scheme the estimator $\hat{\mu}$ has expectation

$$\langle\hat{\mu}\rangle = \frac{1}{nk}\sum_{j=1}^{nk}\langle x_j\rangle,$$

where

$$\langle x_j\rangle = \frac{1}{k}\sum_{i=1}^{k}\mu_i \text{ for all } j,$$

since the k types are in equal proportions in the population. Thus,

$$\langle\hat{\mu}\rangle = \frac{1}{nk}\sum_{j=1}^{nk}\frac{1}{k}\sum_{i=1}^{k}\mu_i = \frac{1}{k}\sum_{i=1}^{k}\mu_i = \mu.$$

The variance of $\hat{\mu}$ is given by

$$\begin{aligned} V[\hat{\mu}] &= \frac{1}{n^2k^2}\, nk\, V[x] \\ &= \frac{1}{nk}\left(\langle x^2\rangle - \mu^2\right) \\ &= \frac{1}{nk}\left(\frac{1}{k}\sum_{i=1}^{k}\langle x_i^2\rangle - \mu^2\right), \end{aligned}$$

again since the k types are in equal proportions in the population.

Now we use the relationship $\sigma^2 = \langle x_i^2\rangle - \mu_i^2$ to replace $\langle x_i^2\rangle$ for each type, noting that σ^2 has the same value in each case. The expression for the variance becomes

$$\begin{aligned}
V[\hat{\mu}] &= \frac{1}{nk}\left[\frac{1}{k}\sum_{i=1}^{k}(\mu_i^2+\sigma^2)-\mu^2\right]\\
&= \frac{\sigma^2-\mu^2}{nk}+\frac{1}{nk^2}\sum_{i=1}^{k}(\mu_i-\mu+\mu)^2\\
&= \frac{\sigma^2-\mu^2}{nk}+\frac{1}{nk^2}\sum_{i=1}^{k}\left[\,(\mu_i-\mu)^2+2\mu(\mu_i-\mu)+\mu^2\,\right]\\
&= \frac{\sigma^2-\mu^2}{nk}+\frac{1}{nk^2}\sum_{i=1}^{k}(\mu_i-\mu)^2+0+\frac{k\mu^2}{nk^2}\\
&= \frac{\sigma^2}{nk}+\frac{1}{nk^2}\sum_{i=1}^{k}(\mu_i-\mu)^2.
\end{aligned}$$

(ii) For the second scheme the calculations are more straightforward. The expectation value of the estimator $\hat{\mu}=(nk)^{-1}\sum_{i=1}^{k}\langle x_i\rangle$ is

$$\langle\hat{\mu}\rangle = \frac{1}{nk}\sum_{i=1}^{k}n\mu_i = \frac{1}{k}\sum_{i=1}^{k}\mu_i = \mu,$$

whilst the variance is given by

$$V[\hat{\mu}] = \frac{1}{n^2k^2}\sum_{i=1}^{k}V[\langle x_i\rangle] = \frac{1}{n^2k^2}\sum_{i=1}^{k}n\sigma_i^2 = \frac{1}{k^2}\frac{k\sigma^2}{n} = \frac{\sigma^2}{kn},$$

since $\sigma_i^2=\sigma^2$ for all i.

Comparing the results from (i) and (ii), we see that the variance of the estimator in the first scheme is larger by

$$\frac{1}{nk^2}\sum_{i=1}^{k}(\mu_i-\mu)^2.$$

31.9 *Each of a series of experiments consists of a large, but unknown, number n ($\gg 1$) of trials, in each of which the probability of success p is the same, but also unknown. In the ith experiment, $i = 1, 2, \ldots, N$, the total number of successes is x_i ($\gg 1$). Determine the log-likelihood function.*

Using Stirling's approximation to $\ln(n-x)$*, show that*

$$\frac{d\,\ln(n-x)}{dn} \approx \frac{1}{2(n-x)} + \ln(n-x),$$

and hence evaluate $\partial({}^nC_x)/\partial n$*.*

By finding the (coupled) equations determining the ML estimators $\hat{p}$ and $\hat{n}$, show that, to order n^{-1}, they must satisfy the simultaneous 'arithmetic' and 'geometric' mean constraints

$$\hat{n}\hat{p} = \frac{1}{N}\sum_{i=1}^{N} x_i \quad \text{and} \quad (1-\hat{p})^N = \prod_{i=1}^{N}\left(1 - \frac{x_i}{\hat{n}}\right).$$

The likelihood function for these N Bernoulli trials is given by

$$L(\mathbf{x}; n, p) = \prod_{i=1}^{N} {}^nC_{x_i}\, p^{x_i}\,(1-p)^{n-x_i}$$

and the corresponding log-likelihood function is

$$\ln L = \sum_{i=1}^{N} \ln\, {}^nC_{x_i} + \ln p \sum_{i=1}^{N} x_i + \ln(1-p)\left[Nn - \sum_{i=1}^{N} x_i\right].$$

The binomial coefficient depends upon n and so we need to determine $\partial({}^nC_x)/\partial n$. To do so, we first consider the derivative of $n!$. Stirling's approximation to $n!$ is

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n, \qquad \text{for large } n.$$

The derivative of n^n is found by setting $y = n^n$ and proceeding as follows:

$$\ln y = n \ln n \quad \Rightarrow \quad \frac{1}{y}\frac{dy}{dn} = \ln n + \frac{n}{n} \quad \Rightarrow \quad \frac{dy}{dn} = n^n(1 + \ln n).$$

It follows that

$$\begin{aligned}
\frac{d(n!)}{dn} &= \sqrt{2\pi}\left[\frac{1}{2\sqrt{n}}\left(\frac{n}{e}\right)^n + \frac{\sqrt{n}}{e^n}\,n^n(1+\ln n) - \sqrt{n}\,n^n\,e^{-n}\right] \\
&= \sqrt{2\pi}\left[\frac{1}{2\sqrt{n}}\left(\frac{n}{e}\right)^n + \frac{\sqrt{n}}{e^n}\,n^n \ln n\right] \\
&= \sqrt{2\pi n}\left(\frac{n}{e}\right)^n\left[\frac{1}{2n} + \ln n\right] = n!\left(\frac{1}{2n} + \ln n\right).
\end{aligned}$$

An immediate consequence of this is

$$\frac{d(\ln n!)}{dn} = \frac{1}{n!}\frac{d(n!)}{dn} = \frac{1}{2n} + \ln n.$$

We now return to the log-likelihood function, the first term of which is

$$\sum_{i=1}^{N} \ln {}^nC_{x_i} = \sum_{i=1}^{N} [\ln n! - \ln x_i! - \ln(n - x_i)!],$$

with, for large n, a partial derivative with respect to n of

$$\begin{aligned}
&\sum_{i=1}^{N} \left[\frac{1}{2n} + \ln n - 0 - \frac{1}{2(n - x_i)} - \ln(n - x_i) \right] \\
&= \sum_{i=1}^{N} \left[\ln \frac{n}{n - x_i} - \frac{x_i}{2n(n - x_i)} \right].
\end{aligned}$$

We are now in a position to find the partial derivatives of the log-likelihood function with respect to p and n and equate each of them to zero, thus yielding the equations $\hat{p}$ and $\hat{n}$ must satisfy.

Firstly, differentiating with respect to p gives

$$\begin{aligned}
\frac{\partial(\ln L)}{\partial p} &= \frac{1}{p}\sum_{i=1}^{N} x_i - \frac{1}{(1-p)} \left[Nn - \sum_{i=1}^{N} x_i \right] = 0, \\
\frac{N\hat{n}}{1 - \hat{p}} &= \left(\frac{1}{\hat{p}} + \frac{1}{1 - \hat{p}} \right) \sum_{i=1}^{N} x_i, \\
\frac{1}{\hat{p}} \sum_{i=1}^{N} x_i = N\hat{n} \quad &\Rightarrow \quad \hat{n}\hat{p} = \frac{1}{N}\sum_{i=1}^{N} x_i.
\end{aligned}$$

Secondly, differentiation with respect to n yields

$$\frac{\partial(\ln L)}{\partial n} = \sum_{i=1}^{N} \left[\ln \frac{n}{n - x_i} - \frac{x_i}{2n(n - x_i)} \right] + N \ln(1 - p) = 0.$$

For large n (and, consequently, large x_i), the first term in the square brackets is of zero-order in n whilst the second is of order n^{-1}. Ignoring the second term and recalling that $\ln 1 = 0$, the equation is equivalent to

$$(1 - \hat{p})^N \prod_{i=1}^{N} \frac{\hat{n}}{\hat{n} - x_i} = 1 \quad \Rightarrow \quad (1 - \hat{p})^N = \prod_{i=1}^{N} \left(1 - \frac{x_i}{\hat{n}} \right).$$

31.11 *According to a particular theory, two dimensionless quantities X and Y have equal values. Nine measurements of X gave values of* 22, 11, 19, 19, 14, 27, 8, 24 *and* 18, *whilst seven measured values of Y were* 11, 14, 17, 14, 19, 16 *and* 14. *Assuming that the measurements of both quantities are Gaussian distributed with a common variance, are they consistent with the theory? An alternative theory predicts that $Y^2 = \pi^2 X$; are the data consistent with this proposal?*

On the hypothesis that $X = Y$ and both quantities have Gaussian distributions with a common variance, we need to calculate the value of t given by

$$t = \frac{\bar{w} - \omega}{\hat{\sigma}} \left(\frac{N_1 N_2}{N_1 + N_2} \right)^{1/2},$$

where $\bar{w} = \bar{x}_1 - \bar{x}_2$, $\omega = \mu_1 - \mu_2 = 0$ and

$$\hat{\sigma} = \left[\frac{N_1 s_1^2 + N_2 s_2^2}{N_1 + N_2 - 2} \right]^{1/2}.$$

The nine measurements of X have a mean of 18.0 and a value for s^2 of 33.33. The corresponding values for the seven measurements of Y are 15.0 and 5.71. Substituting these values gives

$$\hat{\sigma} = \left[\frac{9 \times 33.33 + 7 \times 5.71}{9 + 7 - 2} \right]^{1/2} = 4.93,$$

$$t = \frac{18.0 - 15.0 - 0}{4.93} \left(\frac{9 \times 7}{9 + 7} \right)^{1/2} = 1.21.$$

This variable follows a Student's t-distribution for $9 + 7 - 2 = 14$ degrees of freedom. Interpolation in standard tables gives $C_{14}(1.21) \approx 0.874$, showing that a larger value of t could be expected in about $2 \times (1 - 0.874) = 25\%$ of cases. Thus no inconsistency between the data and the first theory has been established.

For the second theory we are testing Y^2 against $\pi^2 X$; the former will not be Gaussian distributed and the two distributions will not have a common variance. Thus the best we can do is to compare the difference between the two expressions, evaluated with the mean values of X and Y, against the estimated error in that difference.

The difference in the expressions is $(15.0)^2 - 18.0\pi^2 = 47.3$. The error in the difference between the functions of Y and X is given approximately by

$$\begin{aligned} V(Y^2 - \pi^2 X) &= (2Y)^2\, V[\,Y\,] + (\pi^2)^2\, V[X\,] \\ &= (30.0)^2 \frac{5.71}{7-1} + (\pi^2)^2 \frac{33.33}{9-1} \\ &= 1262 \quad \Rightarrow \quad \sigma \approx 35.5. \end{aligned}$$

The difference is thus about $47.3/35.5 = 1.33$ standard deviations away from the theoretical value of 0. The distribution will not be truly Gaussian but, if it were, this figure would have a probability of being exceeded in magnitude some $2 \times (1 - 0.908) = 18\%$ of the time. Again no inconsistency between the data and theory has been established.

31.13 *The χ^2 distribution can be used to test for correlations between characteristics of sampled data. To illustrate this consider the following problem.*

During an investigation into possible links between mathematics and classical music, pupils at a school were asked whether they had preferences (a) *between mathematics and english, and* (b) *between classical and pop music. The results are given below.*

	Classical	*None*	*Pop*
Mathematics	23	13	14
None	17	17	36
English	30	10	40

By computing tables of expected numbers, based on the assumption that no correlations exist, and calculating the relevant values of χ^2, determine whether there is any evidence for

(a) *a link between academic and musical tastes, and*
(b) *a claim that pupils either had preferences in both areas or had no preference.*

You will need to consider the appropriate value for the number of degrees of freedom to use when applying the χ^2 test.

We first note that there were 200 pupils taking part in the survey. Denoting no academic preference between mathematics and english by NA and no musical preference by NM, we draw up an enhanced table of the actual numbers m_{XY} of preferences for the various combinations that also shows the overall probabilities p_X and p_Y of the three choices in each selection.

	C	NM	P	Total	p_X
M	23	13	14	50	0.25
NA	17	17	36	70	0.35
E	30	10	40	80	0.40
Total	70	40	90	200	
p_Y	0.35	0.20	0.45		

(a) If we now assume the (null) hypothesis that there are no correlations in the data and that any apparent correlations are the result of statistical fluctuations, then the expected number of pupils opting for the combination X and Y is $n_{\rm XY} = 200 \times p_{\rm X} \times p_{\rm Y}$. A table of $n_{\rm XY}$ is as follows:

	C	NM	P	Total
M	17.5	10	22.5	50
NA	24.5	14	31.5	70
E	28	16	36	80
Total	70	40	90	200

Taking the standard deviation as the square root of the expected number of votes for each particular combination, the value of χ^2 is given by

$$\chi^2 = \sum_{\text{all XY combinations}} \left(\frac{n_i - m_i}{\sqrt{n_i}} \right)^2 = 12.3.$$

For an $n \times n$ correlation table (here $n = 3$), the $(n-1) \times (n-1)$ block of entries in the upper left-hand can be filled in arbitrarily. But, as the totals for each row and column are predetermined, the remaining $2n - 1$ entries are not arbitrary. Thus the number of degrees of freedom (d.o.f.) for such a table is $(n-1)^2$, here 4 d.o.f. From tables, a χ^2 of 12.3 for 4 d.o.f. makes the assumed hypothesis less than 2% likely, and so it is almost certain that a correlation between academic and musical tastes does exist.

(b) To investigate a claim that pupils either had preferences in both areas or had no preference, we must combine expressed preferences for classical or pop into one set labelled PM meaning 'expressed a musical preference'; similarly for academic subjects. The correlation table is now a 2×2 one and will have only one degree of freedom. The actual and expected ($n_{\rm XY} = 200 p_{\rm X} p_{\rm Y}$) data tables are

	PM	NM	Total	$p_{\rm X}$
PA	107	23	130	0.65
NA	53	17	70	0.35
Total	160	40	200	
$p_{\rm Y}$	0.80	0.20		

	PM	NM	Total
PA	104	26	130
NA	56	14	70
Total	160	40	200

The value of χ^2 is

$$\chi^2 = \frac{(-3)^2}{104} + \frac{(3)^2}{26} + \frac{(3)^2}{56} + \frac{(-3)^2}{14} = 1.24.$$

This is close to the expected value (1) of χ^2 for 1 d.o.f. and is neither too big nor too small. Thus there is no evidence for the claim (or for any tampering with the data!).

31.15 *A particle detector consisting of a shielded scintillator is being tested by placing it near a particle source whose intensity can be controlled by the use of absorbers. It might register counts even in the absence of particles from the source because of the cosmic ray background.*

The number of counts n registered in a fixed time interval as a function of the source strength s is given as:

source strength s:	0	1	2	3	4	5	6
counts n:	6	11	20	42	44	62	61

At any given source strength, the number of counts is expected to be Poisson distributed with mean

$$n = a + bs,$$

where a and b are constants. Analyse the data for a fit to this relationship and obtain the best values for a and b together with their standard errors.

(a) *How well is the cosmic ray background determined?*
(b) *What is the value of the correlation coefficient between a and b? Is this consistent with what would happen if the cosmic ray background were imagined to be negligible?*
(c) *Do the data fit the expected relationship well? Is there any evidence that the reported data 'are too good a fit'?*

Because in this exercise the independent variable s takes only consecutive integer values, we will use it as a label i and denote the number of counts corresponding to $s = i$ by n_i. As the data are expected to be Poisson distributed, the best estimate of the variance of each reading is equal to the best estimate of the reading itself, namely the actual measured value. Thus each reading n_i has an error of $\sqrt{n_i}$, and the covariance matrix N takes the form $N = \text{diag}(n_0, n_1, \ldots, n_6)$, i.e. it is diagonal, but not a multiple of the unit matrix.

The expression for χ^2 is

$$\chi^2(a,b) = \sum_{i=0}^{6} \left(\frac{n_i - a - bi}{\sqrt{n_i}} \right)^2 \qquad (*).$$

Minimisation with respect to a and b gives the simultaneous equations

$$0 = \frac{\partial \chi^2}{\partial a} = -2 \sum_{i=0}^{6} \frac{n_i - a - bi}{n_i},$$

$$0 = \frac{\partial \chi^2}{\partial b} = -2 \sum_{i=0}^{6} \frac{i(n_i - a - bi)}{n_i}.$$

As is shown more generally in textbooks on numerical computing (e.g. William H. Press *et al.*, *Numerical Recipes in C*, 2nd edn (Cambridge: Cambridge University Press, 1996), Sect. 15.2), these equations are most conveniently solved by defining the quantities

$$S \equiv \sum_{i=0}^{6} \frac{1}{n_i}, \quad S_x \equiv \sum_{i=0}^{6} \frac{i}{n_i}, \quad S_y \equiv \sum_{i=0}^{6} \frac{n_i}{n_i},$$

$$S_{xx} \equiv \sum_{i=0}^{6} \frac{i^2}{n_i}, \quad S_{xy} \equiv \sum_{i=0}^{6} \frac{in_i}{n_i}, \quad \Delta \equiv SS_{xx} - (S_x)^2.$$

With these definitions (which correspond to the quantities calculated and accessibly stored in most calculators programmed to perform least-squares fitting), the solutions for the best estimators of a and b are

$$\hat{a} = \frac{S_{xx}S_y - S_xS_{xy}}{\Delta},$$

$$\hat{b} = \frac{S_{xy}S - S_xS_y}{\Delta},$$

with variances and covariance given by

$$\sigma_a^2 = \frac{S_{xx}}{\Delta}, \quad \sigma_b^2 = \frac{S}{\Delta}, \quad \mathrm{Cov}(a, b) = -\frac{S_x}{\Delta}.$$

The computed values of these quantities are: $S = 0.38664$; $S_x = 0.53225$; $S_y = 7$; $S_{xx} = 1.86221$; $S_{xy} = 21$; $\Delta = 0.43671$.

From these values, the best estimates of $\hat{a}$, $\hat{b}$ and the variances σ_a^2 and σ_b^2 are

$$\hat{a} = 4.2552, \quad \hat{b} = 10.061, \quad \sigma_a^2 = 4.264, \quad \sigma_b^2 = 0.8853.$$

The covariance is $\mathrm{Cov}(a, b) = -1.2187$, giving estimates for a and b of

$$a = 4.3 \pm 2.1 \qquad \text{and} \qquad b = 10.06 \pm 0.94,$$

with a correlation coefficient $r_{ab} = -0.63$.

(a) The cosmic ray background must be present, since $n(0) \neq 0$, but its value of about 4 is uncertain to within a factor of 2.

(b) The correlation between a and b is negative and quite strong. This is as expected since, if the cosmic ray background represented by a were reduced towards zero, then b would have to be increased to compensate when fitting to the measured data for non-zero source strengths.

(c) A measure of the goodness-of-fit is the value of χ^2 achieved using the best-fit values for a and b. Direct resubstitution of the values found into (∗) gives $\chi^2 = 4.9$. If the weight of a particular reading is taken as the square root of the predicted (rather than the measured) value, then χ^2 rises slightly to 5.1. In either case the result is almost exactly that 'expected' for 5 d.o.f. – neither too good nor too bad.

There are five degrees of freedom because there are seven data points and two parameters have been chosen to give a best fit.

***31.17** The following are the values and standard errors of a physical quantity $f(\theta)$ measured at various values of θ (in which there is negligible error):*

θ	0	$\pi/6$	$\pi/4$	$\pi/3$
$f(\theta)$	3.72 ± 0.2	1.98 ± 0.1	-0.06 ± 0.1	-2.05 ± 0.1

θ	$\pi/2$	$2\pi/3$	$3\pi/4$	π
$f(\theta)$	-2.83 ± 0.2	1.15 ± 0.1	3.99 ± 0.2	9.71 ± 0.4

Theory suggests that f should be of the form $a_1 + a_2 \cos\theta + a_3 \cos 2\theta$. Show that the normal equations for the coefficients a_i are

$$\begin{aligned}
481.3a_1 + 158.4a_2 - 43.8a_3 &= 284.7,\\
158.4a_1 + 218.8a_2 + 62.1a_3 &= -31.1,\\
-43.8a_1 + 62.1a_2 + 131.3a_3 &= 368.4.
\end{aligned}$$

(a) *If you have matrix inversion routines available on a computer, determine the best values and variances for the coefficients a_i and the correlation between the coefficients a_1 and a_2.*
(b) *If you have only a calculator available, solve for the values using a Gauss–Seidel iteration and start from the approximate solution $a_1 = 2$, $a_2 = -2$, $a_3 = 4$.*

Assume that the measured data have uncorrelated errors. The quoted errors are not all equal and so the covariance matrix N, whilst being diagonal, will not be a multiple of the unit matrix; it will be

$$\mathsf{N} = \text{diag}(0.04, 0.01, 0.01, 0.01, 0.04, 0.01, 0.04, 0.16).$$

Using as base functions the three functions $h_1(\theta) = 1$, $h_2(\theta) = \cos\theta$ and $h_3(\theta) = \cos 2\theta$, we calculate the elements of the 8×3 response matrix $\mathsf{R}_{ij} = h_j(\theta_i)$. To save space we display its 3×8 transpose:

$$\mathsf{R}^{\mathrm{T}} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0.866 & 0.707 & 0.500 & 0 & -0.500 & -0.707 & -1 \\ 1 & 0.500 & 0 & -0.500 & -1 & -0.500 & 0 & 1 \end{pmatrix}$$

Then

$$\mathsf{R}^{\mathrm{T}}\mathsf{N}^{-1} = \begin{pmatrix} 25 & 100 & 100 & 100 & 25 & 100 & 25 & 6.25 \\ 25 & 86.6 & 70.7 & 50 & 0 & -50 & -17.7 & -6.25 \\ 25 & 50.0 & 0 & -50.0 & -25 & -50 & 0 & 6.25 \end{pmatrix}$$

and

$$\mathsf{R}^{\mathrm{T}}\mathsf{N}^{-1}\mathsf{R} = \mathsf{R}^{\mathrm{T}}\mathsf{N}^{-1}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0.866 & 0.500 \\ 1 & 0.707 & 0 \\ 1 & 0.500 & -0.500 \\ 1 & 0 & -1 \\ 1 & -0.500 & -0.500 \\ 1 & -0.707 & 0 \\ 1 & -1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 481.25 & 158.35 & -43.75 \\ 158.35 & 218.76 & 62.05 \\ -43.75 & 62.05 & 131.25 \end{pmatrix}.$$

From the measured values,

$$\mathsf{f} = (3.72,\ 1.98,\ -0.06,\ -2.05,\ -2.83,\ 1.15,\ 3.99,\ 9.71)^{\mathrm{T}},$$

we need to calculate $\mathsf{R}^{\mathrm{T}}\mathsf{N}^{-1}\mathsf{f}$, which is given by

$$\begin{pmatrix} 25 & 100 & 100 & 100 & 25 & 100 & 25 & 6.25 \\ 25 & 86.6 & 70.7 & 50 & 0 & -50 & -17.7 & -6.25 \\ 25 & 50.0 & 0 & -50 & -25 & -50 & 0 & 6.25 \end{pmatrix} \begin{pmatrix} 3.72 \\ 1.98 \\ -0.06 \\ -2.05 \\ -2.83 \\ 1.15 \\ 3.99 \\ 9.71 \end{pmatrix},$$

i.e. $(284.7,\ -31.08,\ 368.44)^{\mathrm{T}}$.

The vector of LS estimators of a_i satisfies $\mathsf{R}^{\mathrm{T}}\mathsf{N}^{-1}\mathsf{R}\hat{\mathsf{a}} = \mathsf{R}^{\mathrm{T}}\mathsf{N}^{-1}\mathsf{f}$. Substituting the forms calculated above into the two sides of the equality gives the set of equations stated in the question.

(a) Machine (or manual!) inversion gives

$$(\mathsf{R}^{\mathrm{T}}\mathsf{N}^{-1}\mathsf{R})^{-1} = 10^{-3}\begin{pmatrix} 3.362 & -3.177 & 2.623 \\ -3.177 & 8.282 & -4.975 \\ 2.623 & -4.975 & 10.845 \end{pmatrix}.$$

From this (covariance matrix) we can calculate the standard errors on the a_i from the square roots of the terms on the leading diagonal as ± 0.058, ± 0.091 and ± 0.104. We can further calculate the correlation coefficient r_{12} between a_1 and a_2 as

$$r_{12} = \frac{-3.177 \times 10^{-3}}{0.058 \times 0.091} = -0.60.$$

The best values for the a_i are given by the result of multiplying the column matrix $(284.7,\ -31.08,\ 368.44)^{\mathrm{T}}$ by the above inverted matrix. This yields $(2.022,\ -2.944,\ 4.897)^{\mathrm{T}}$ to give the best estimates of the a_i as

$$a_1 = 2.02 \pm 0.06, \quad a_2 = -2.99 \pm 0.09, \quad a_3 = 4.90 \pm 0.10.$$

(b) Denote the given set of equations by $\mathsf{Aa} = \mathsf{b}$ and start by dividing each equation by the quantity needed to make the diagonal elements of A each equal to unity; this produces $\mathsf{Ca} = \mathsf{d}$. Then, writing $\mathsf{C} = \mathsf{I} - \mathsf{F}$ yields the basis of the iteration scheme,

$$\mathsf{a}_{n+1} = \mathsf{F}\mathsf{a}_n + \mathsf{d}.$$

We use only the simplest form of Gauss–Seidel iteration (with no separation into upper and lower diagonal matrices).

The explicit form of $\mathsf{Ca} = \mathsf{d}$ is

$$\begin{pmatrix} 1 & 0.3290 & -0.0909 \\ 0.7239 & 1 & 0.2836 \\ -0.3333 & 0.4728 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 0.5916 \\ -0.1421 \\ 2.8072 \end{pmatrix}$$

and

$$\mathsf{F} = \begin{pmatrix} 0 & -0.3290 & 0.0909 \\ -0.7239 & 0 & -0.2836 \\ 0.3333 & -0.4728 & 0 \end{pmatrix}.$$

Starting with the approximate solution $a_1 = 2$, $a_2 = -2$, $a_3 = 4$ gives as the result of the first ten iterations

a_1	a_2	a_3
2.000	−2.000	4.000
1.613	−2.724	4.419
1.890	−2.563	4.633
1.856	−2.824	4.649
1.943	−2.804	4.761
1.947	−2.899	4.781
1.980	−2.907	4.827
1.987	−2.944	4.842
2.000	−2.953	4.861
2.005	−2.969	4.870
2.011	−2.975	4.879

This final set of values is in close agreement with that obtained by direct inversion; in fact, after eighteen iterations the values agree exactly to three significant figures. Of course, using this method makes it difficult to estimate the errors in the derived values.

31.19 *The F-distribution $h(F)$ for the ratio F of the variances of two samples of sizes N_1 and N_2 drawn from populations with a common variance is*

$$\left(\frac{n_1}{n_2}\right)^{n_1/2} \frac{F^{(n_1-2)/2}}{B\left(\frac{n_1}{2},\frac{n_2}{2}\right)} \left(1+\frac{n_1}{n_2}F\right)^{-(n_1+n_2)/2},$$

where, to save space, we have written N_1-1 as n_1 and N_2-1 as n_2.

Verify that the F-distribution $P(F)$ is symmetric between the two data samples, i.e. that it retains the same form but with N_1 and N_2 interchanged, if F is replaced by $F'=F^{-1}$. Symbolically, if $P'(F')$ is the distribution of F' and $P(F)=\eta(F,N_1,N_2)$, then $P'(F')=\eta(F',N_2,N_1)$.

We first write $F^{-1}=F'$ with $|dF|=|dF'|/F'^2$ and rewrite $h(F)$ as

$$\left(\frac{n_1}{n_2}\right)^{n_1/2} \frac{(F')^{-(n_1-2)/2}}{B\left(\frac{n_1}{2},\frac{n_2}{2}\right)} \left(1+\frac{n_1}{n_2\,F'}\right)^{-(n_1+n_2)/2}$$

$$= \left(\frac{n_1}{n_2}\right)^{n_1/2} \frac{(F')^{-(n_1-2)/2}}{B\left(\frac{n_1}{2},\frac{n_2}{2}\right)} (F')^{(n_1+n_2)/2} \left(\frac{n_2}{n_1}\right)^{(n_1+n_2)/2} \left(\frac{F'\,n_2}{n_1}+1\right)^{-(n_1+n_2)/2}$$

$$= \left(\frac{n_2}{n_1}\right)^{n_2/2} \frac{(F')^{(n_2+2)/2}}{B\left(\frac{n_1}{2},\frac{n_2}{2}\right)} \left(1+\frac{n_2\,F'}{n_1}\right)^{-(n_1+n_2)/2}.$$

Further,

$$h(F)\,|dF| = \left(\frac{n_2}{n_1}\right)^{n_2/2} \frac{(F')^{(n_2+2)/2}}{B\left(\frac{n_1}{2},\frac{n_2}{2}\right)} \left(1+\frac{n_2\,F'}{n_1}\right)^{-(n_1+n_2)/2} \frac{1}{F'^2}\,|dF'|$$

$$= \left(\frac{n_2}{n_1}\right)^{n_2/2} \frac{(F')^{(n_2-2)/2}}{B\left(\frac{n_2}{2},\frac{n_1}{2}\right)} \left(1+\frac{n_2\,F'}{n_1}\right)^{-(n_1+n_2)/2} |dF'|.$$

In the last step we have made use of the symmetry of the beta function $B(x,y)$ with respect to its arguments. To express the final result in the usual F-distribution form, we need to restore n_1 to N_1-1 and n_2 to N_2-1, but the symmetry between the data samples has already been demonstrated.