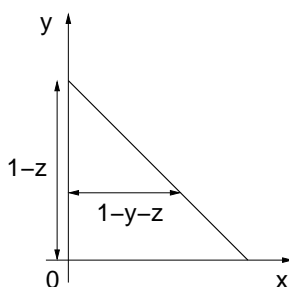Paper 7: Mathematical Methods

Solutions to 2002 Tripos Paper

1. *Evaluation of integrals, Jacobians.*

   (a) Consider a horizontal $(x\text{–}y)$ slice through the tetrahedron:

   If we evaluate the x integral first, then the y integral, then the z integral, the limits are

   $$0 \le z \le 1 \,, \ \ 0 \le y \le 1 - z \,, \ \ 0 \le x \le 1 - y - z \;.$$

   The integral can therefore be written as

   $$\int_{z=0}^{1} \left[\int_{y=0}^{1-z} \left(\int_{x=0}^{1-y-z} yz \; dx \right) dy \right] dz$$

   Starting with the x integral, we have

   $$\int_{0}^{1-y-z} yz \; dx = [yzx]_{0}^{1-y-z} = yz - zy^2 - yz^2 = yz(1 - z) - zy^2 \;.$$

   Continuing with the y integral, we have

   $$\begin{aligned}
   \int_{0}^{1-z} \left(yz(1 - z) - zy^2 \right) dy &= \left[\frac{1}{2}z(1 - z)y^2 - \frac{1}{3}zy^3 \right]_{0}^{1-z} \\
   &= \frac{1}{2}z(1 - z)(1 - z)^2 - \frac{1}{3}z(1 - z)^3 = \frac{1}{6}z(1 - z)^3 \\
   &= \frac{1}{6}z - \frac{1}{2}z^2 + \frac{1}{2}z^3 - \frac{1}{6}z^4 \;.
   \end{aligned}$$

   Finally, the integral in z is

   $$\int_{0}^{1} \left(\frac{1}{6}z - \frac{1}{2}z^2 + \frac{1}{2}z^3 - \frac{1}{6}z^4 \right) dz = \left[\frac{1}{12}z^2 - \frac{1}{6}z^3 + \frac{1}{8}z^4 - \frac{1}{30}z^5 \right]_{0}^{1} = \frac{1}{120} \;. \qquad [10]$$

   1

(b) The $\mathbf{i}$ component of $\mathbf{J}$ is parallel to S, while the $\mathbf{k}$ component is perpendicular to S, so the flux integral becomes

$$\int_S xy \, dx \, dy \ .$$

To simplify the region of integration, we change to a polar coordinate system centred at $(1, 1)$, so

$$x = 1 + r \cos \theta \ , \quad y = 1 + r \sin \theta \ .$$

The limits of the area S are then $0 \le r \le 1$ and $-\pi \le \theta \le -\pi/2$. The Jacobian of the transformation is given by

$$J = \left| \begin{array}{cc} \partial x/\partial r & \partial x/\partial \theta \\ \partial y/\partial r & \partial y/\partial \theta \end{array} \right| = r \cos^2 \theta + r \sin^2 \theta = r \ .$$

The flux integral is therefore

$$\begin{aligned}
\int_S xy \, dx \, dy &= \int_{\theta=-\pi}^{-\pi/2} \int_{r=0}^{1} (1 + r \cos \theta)(1 + r \sin \theta) \, r \, dr \, d\theta \\
&= \int_{\theta=-\pi}^{-\pi/2} \int_{r=0}^{1} \left(r + r^2(\cos \theta + \sin \theta) + r^3 \sin \theta \cos \theta \right) \, dr \, d\theta \\
&= \int_{-\pi}^{-\pi/2} \left(\frac{1}{2} + \frac{1}{3}(\cos \theta + \sin \theta) + \frac{1}{8} \sin 2\theta \right) d\theta \\
&= \left[\frac{\theta}{2} + \frac{1}{3}(\sin \theta - \cos \theta) - \frac{1}{16} \cos 2\theta \right]_{-\pi}^{-\pi/2} \\
&= \left(-\frac{\pi}{4} + \frac{1}{3}(-1 - 0) - \frac{1}{16}(-1) \right) - \left(-\frac{\pi}{2} + \frac{1}{3}(0 + 1) - \frac{1}{16}(1) \right) \\
&= \frac{\pi}{4} - \frac{13}{24} \ .
\end{aligned}$$

[10]

Examiner's remarks: A popular question on volume and flux integrals that was particularly well answered by many. The principal error was that of minor slips in mathematical manipulation. An error of more concern was in carelessness in defining coordinates and angles. A small, but worrying, number of candidates believed that they were calculating the volume of a tetrahedron in (a).

2. *Irrotational vector fields and scalar potentials.*

(a) An irrotational vector field $\mathbf{u}$ has zero curl everywhere. The line integral $\oint_c \mathbf{u}.\mathbf{dl}$ is zero around any closed path c, while the open line integral $\int_a^b \mathbf{u}.\mathbf{dl}$ depends only on the end points a and b and not on the path taken between these points. For every irrotational vector field $\mathbf{u}$ there exists a scalar potential ϕ such that $\mathbf{u} = \nabla \phi$. The open line integral $\int_a^b \mathbf{u}$ is equal to $\phi(b) - \phi(a)$.

2

From the above, it follows that $\nabla \times \nabla \phi = \mathbf{0}$ for any differentiable scalar field ϕ. This can be proved by expanding in Cartesian coordinates:

$$\nabla \times \nabla \phi = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ \partial\phi/\partial x & \partial\phi/\partial y & \partial\phi/\partial z \end{vmatrix} = \begin{bmatrix} \partial^2\phi/\partial y\partial z - \partial^2\phi/\partial z\partial y \\ -(\partial^2\phi/\partial x\partial z - \partial^2\phi/\partial z\partial x) \\ \partial^2\phi/\partial x\partial y - \partial^2\phi/\partial y\partial x \end{bmatrix} = \mathbf{0} \ ,$$

since the cross-derivatives are equal for any differentiable scalar field ϕ. [6]

(b) (i) Converting from Cartesian to to cylindrical polar coordinates, we have $x\mathbf{i}+y\mathbf{j} = r\mathbf{e}_r$, and $x^2 + y^2 = r^2$. Hence

$$\mathbf{F} = \frac{r\mathbf{e}_r}{r^2} = \frac{\mathbf{e}_r}{r} = r^n\mathbf{e}_r \quad \text{where } n = -1.$$ [2]

(ii) To show that $\mathbf{F}$ is irrotational, we compute its curl in cylindrical polar coordinates:

$$\nabla \times \mathbf{F} = \frac{1}{r} \begin{vmatrix} \mathbf{e}_r & r\mathbf{e}_\theta & \mathbf{e}_z \\ \partial/\partial r & \partial/\partial \theta & \partial/\partial z \\ r^{-1} & 0 & 0 \end{vmatrix} = \mathbf{0} \ .$$

So $\mathbf{F}$ is indeed irrotational. Its scalar potential ϕ can be found as follows:

$$\nabla\phi \ = \ \mathbf{e}_r\frac{\partial\phi}{\partial r} + \frac{\mathbf{e}_\theta}{r}\frac{\partial\phi}{\partial \theta} + \mathbf{e}_z\frac{\partial\phi}{\partial z} = \mathbf{F} = r^{-1}\mathbf{e}_r$$

$$\Rightarrow \frac{\partial\phi}{\partial r} \ = \ r^{-1} \Leftrightarrow \phi = \ln r + c \ , \quad \text{for any constant } c.$$

Since ϕ depends on r only, the equipotential surfaces will be surfaces of constant r, which are simply concentric cylinders around the z-axis.
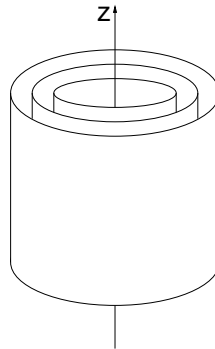

[5]

(iii) Since $\mathbf{F}$ is irrotational, we can compute the line integral by evaluating the scalar potential at the end points. The starting point at $t = 0$ has Cartesian coordinates $(1, 0, 0)$ and cylindrical radius $r = 1$. The finishing point at $t = 1$ has Cartesian coordinates $(2, 1, \sin 1)$ and cylindrical radius $r = \sqrt{5}$. The difference between the start and end potentials is therefore $\ln \sqrt{5} - \ln 1 = \ln \sqrt{5} = 0.805$. [2]

(c) A solenoidal vector field has zero divergence everywhere. Taking the divergence of $r^n \mathbf{e}_r$, we get

$$\nabla.(r^n \mathbf{e}_r) = \frac{1}{r}\frac{\partial(r.r^n)}{\partial r} = \begin{cases} \frac{1}{r}(n+1)r^n & \text{if } n \neq -1 \\ \frac{1}{r} \times 0 & \text{if } n = -1 \end{cases} .$$

So, apart from a singularity at $r = 0$ (the z-axis), the field $r^n \mathbf{e}_r$ with $n = -1$ is solenoidal: all other values of n give fields that are not solenoidal.

Consider the cube bounded by the planes $|x| = 1$, $|y| = 1$ and $|z| = 1$, and also the cylinder with length 2 and radius 1, centred on the origin and aligned along the z-axis: a top view of these two surfaces is shown below.
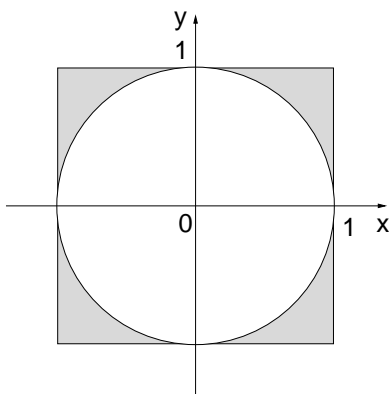
Since the divergence of $\mathbf{F}$ is zero everywhere inside the shaded volume, Gauss's law tells us that the flux of $\mathbf{F}$ into the shaded volume is the same as the flux out of the shaded volume. This implies that the flux out of the cube is the same as the flux out of the cylinder. It is straightforward to evaluate $\int_S \mathbf{F}.\mathbf{dA}$ over the cylinder, since $\mathbf{F}$ is perpendicular to the cylinder at all points on its surface, so the integral is simply the surface area of the cylinder multiplied by $|\mathbf{F}|$. Furthermore, $|\mathbf{F}| = r^{-1} = 1$ at the surface of the cylinder. Hence

$$\int_S \mathbf{F}.\mathbf{dA} = 1 \times 2\pi \times 1 \times 2 = 4\pi . \qquad [5]$$

Examiner's remarks: A popular and well answered question on vector calculus. The most common error was found in (c), where most did not appreciate the importance of the singularity at the origin and said that the flux integral was zero. The line integral in (b)(iii) was most easily calculated making use of the scalar potential in (b)(ii), but most candidates attempted the integration directly.

3. *Partial differential equations, the diffusion equation.*

(a) Heat flow is proportional to ∇T, which is simply $\partial T/\partial x$ in the one-dimensional case. Since the bar is insulated at its ends, there is no heat flow at these points and so $\partial T/\partial x = 0$ at $x = 0$ and $x = L$. $\qquad [2]$

(b) Assume the solution is of the form $T(x,t) = F(x)G(t)$. Then $\partial T/\partial t = F\dot{G}$ and $\partial^2 T/\partial x^2 = F''G$. Substituting into the diffusion equation, we obtain $F\dot{G} = \alpha F''G$, and rearranging this we obtain

$$\frac{\dot{G}}{\alpha G} = \frac{F''}{F} \ .$$

Since the left hand side depends only on t and the right hand side depends only on x, it follows that both sides must equal a constant, k say. We thus obtain two ordinary differential equations for F and G:

$$F'' - kF = 0 \quad \text{and} \quad \dot{G} - \alpha kG = 0 \ .$$

The boundary conditions $\partial T/\partial x = 0$ at $x = 0$ and $x = L$ shed some light on the sign of k. Since $\partial T/\partial x = F'(x)G(t)$, it follows that $F'(x) = 0$ at $x = 0$ and $x = L$: the alternative, that $G(t) = 0$, leads to the uninteresting solution $T(x,t) = 0$. If k is positive, the ordinary differential equation for F has a weighted exponential solution, which will satisfy the boundary conditions for $F'(x)$ only for the uninteresting weighting of zero. We therefore deduce that k is negative, $k = -p^2$ say, and so

$$F(x) = A \cos px + B \sin px \Rightarrow F'(x) = -Ap \sin px + Bp \cos px \ .$$

If $F'(x) = 0$ at $x = 0$, B must be zero. Furthermore, considering the remaining $\sin px$ term, if $F'(x) = 0$ at $x = L$, $p = n\pi/L$ for any integer n. The ordinary differential equation for G can now be written $\dot{G} + \lambda_n^2 G = 0$, where $\lambda_n = p\sqrt{\alpha} = n\pi\sqrt{\alpha}/L$. The solution is $G(t) = Be^{-\lambda_n^2 t}$ for any constant B. Bringing everything together, we have

$$T_n(x,t) = F(x)G(t) = A_n \cos \frac{n\pi x}{L} e^{-\lambda_n^2 t}, \quad \text{where } \lambda_n = \frac{n\pi\sqrt{\alpha}}{L} \ .$$

If T_n satisfies the diffusion equation, then so does an infinite sum of T_n over all integers n. The general solution is therefore

$$T(x,t) = \sum_{n=-\infty}^{\infty} T_n(x,t) = \sum_{n=0}^{\infty} A_n \cos \frac{n\pi x}{L} e^{-\lambda_n^2 t} \ ,$$

where we have ignored the $n < 0$ terms, since they are the same as the $n > 0$ terms. [10]

(c) For $t = 0$, we have

$$T(x,0) = \sum_{n=0}^{\infty} A_n \cos \frac{n\pi x}{L} \ .$$

This is an even Fourier series: we need to find the coefficients A_n which result in the given square wave in the range 0 to L. If we differentiate the Fourier series for a triangular wave in the electrical data book, we obtain an even Fourier series for a square wave of peak-to-peak amplitude $8/T$ and period T:

$$\frac{dg}{dt} = \frac{8}{\pi^2} \sum_{n=1}^{\infty} (-1)^{n+1} \omega_0 \frac{\cos(2n-1)\omega_0 t}{(2n-1)} \ , \text{where } \omega_0 = \frac{2\pi}{T} \ .$$

5

If we equate t with x and T with $2L$, scale the amplitude by $100T/8$ and add an offset of 150 (the average of the given square wave), then this is exactly what we want for $T(x, 0)$. So

$$T(x, 0) = 150 + \frac{200}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{(2n-1)} \cos \frac{(2n-1)\pi x}{L} = 150 + \frac{200}{\pi} \cos \frac{\pi x}{L} - \ldots$$

Thus $A_0 = 150$, $A_1 = 200/\pi$ and $A_2 = 0$. As $t \to \infty$, all the terms in the expression for $T(x, t)$ decay to zero apart from the A_0 term, which is what we'd expect: the temperature eventually levels out everywhere to the average initial temperature. The higher harmonics decay faster than the lower harmonics, so at some intermediate value of t we can approximate $T(x, t)$ by the A_0 and A_1 terms alone.

very large t   intermediate t   [8]
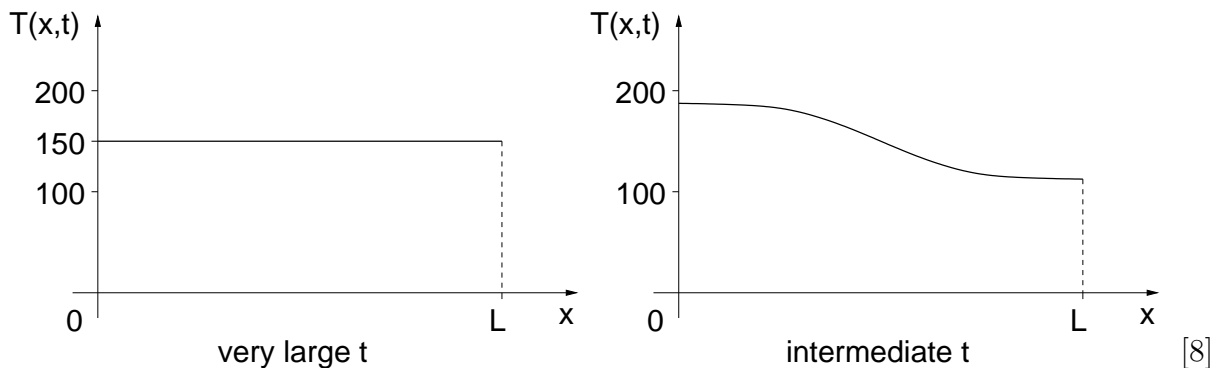
Examiner's remarks: This was the least popular question in Section A and the worst handled (though there were several excellent answers). Part (a) was surprisingly poorly answered, suggesting that many candidates had no physical view of the problem they were solving. Much the same might be said of the answers to (c). In solving the PDE, many candidates chose solutions arbitrarily to ensure they matched the given answer, with no attempt to provide an explicit argument for their choice: this is disturbing.

4. *LU decomposition, the fundamental subspaces of a matrix, least squares.*

(a) We can compute the **LU** decomposition by performing Gaussian elimination on **A**, keeping track of **L** and **U** as we go:

$$\mathbf{A} = \begin{bmatrix} 3 & -1 & 4 \\ 6 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & -1 & 4 \\ 6 & 2 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 3 & -1 & 4 \\ 0 & 4 & -7 \end{bmatrix} = \mathbf{LU}$$

There are two pivots, so the rank of **A** is two.   [4]

6

(b) The dimensions of the row and column spaces are both equal to the rank of the matrix, in this case two. The dimensions of the row space and the left nullspace add up to two, the number of rows: hence the dimension of the left nullspace is zero. The dimensions of the column space and the nullspace add up to three, the number of columns: hence the dimension of the nullspace is one. [4]

(c) The column space is given by the columns of $\mathbf{A}$ that have pivots. These are $[\, 3 \; 6 \,]^T$ and $[\, -1 \; 2 \,]^T$.

The row space is given by the rows of $\mathbf{U}$ that have pivots. These are $[\, 3 \;\; -1 \; 4 \,]^T$ and $[\, 0 \; 4 \;\; -7 \,]^T$.

The nullspace is given by the solution of $\mathbf{Ux} = \mathbf{0}$, which is found by setting the free variable to 1 and solving by back-substitution:

$$\begin{bmatrix} 3 & -1 & 4 \\ 0 & 4 & -7 \end{bmatrix} \begin{bmatrix} -3/4 \\ 7/4 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

So the nullspace is $[\, -3 \; 7 \; 4 \,]^T$.

The left nullspace has zero dimension, as noted in (b). So, the left nullspace contains only the zero vector $[\, 0 \; 0 \,]^T$. [6]

(d) $\mathbf{A}^T \bar{\mathbf{x}}$ lies somewhere in the column space of $\mathbf{A}^T$ (the row space of $\mathbf{A}$). The error vector lies in the left nullspace of $\mathbf{A}^T$ (the nullspace of $\mathbf{A}$):

$$\text{error vector } = \begin{bmatrix} -3 \\ 7 \\ 4 \end{bmatrix} - \mathbf{A}^T \bar{\mathbf{x}} = \lambda \begin{bmatrix} -3 \\ 7 \\ 4 \end{bmatrix} \Leftrightarrow \mathbf{A}^T \bar{\mathbf{x}} = \mu \begin{bmatrix} -3 \\ 7 \\ 4 \end{bmatrix}$$

So we have established that $\mathbf{A}^T \bar{\mathbf{x}}$ lies in the row space of $\mathbf{A}$ and the nullspace of $\mathbf{A}$. Since these two subspaces are orthogonal, the only solution for $\mathbf{A}^T \bar{\mathbf{x}}$ is the zero vector $[\, 0\,0\,0 \,]^T$.

We can now deduce $\bar{\mathbf{x}}$. Since $\mathbf{A}^T \bar{\mathbf{x}} = \mathbf{0}$, $\bar{\mathbf{x}}$ lies in the nullspace of $\mathbf{A}^T$ which is the same as the left nullspace of $\mathbf{A}$. We have already established in (b) that the left nullspace of $\mathbf{A}$ is the zero vector, so $\bar{\mathbf{x}} = [\, 0 \;\; 0 \,]^T$. [6]

Examiner's remarks: This question asked the candidates to explore the properties of a 2×3 matrix. Almost all found the $\mathbf{LU}$ decomposition correctly, but it was downhill from there. About half the candidates thought the rank of the matrix was 1 (clearly indicating that they did not understand what the rank of a matrix is), while fewer than half managed to compute bases for the four fundamental subspaces. In part (d), only a couple of candidates solved the least squares problem by considering the orthogonality of the subspaces, some others got the right answer by tedious algebraic expansion, while most did not have a clue.

5. *Eigenvalues, eigenvectors, linear difference equations.*

(a) By definition, we know that $\mathbf{A}\mathbf{u}_k = \lambda_k \mathbf{u}_k$, for $k \in \{1 \ldots n\}$. Concatenating all n of these equations, we obtain

$$\mathbf{A}\,[\mathbf{u}_1\ \mathbf{u}_2\ \ldots\ \mathbf{u}_n] = [\lambda_1 \mathbf{u}_1\ \lambda_2 \mathbf{u}_2\ \ldots\ \lambda_n \mathbf{u}_n] \Leftrightarrow \mathbf{A}\mathbf{U} = \mathbf{U}\Lambda$$

Finally, post-multiplying both sides by $\mathbf{U}^{-1}$, we get $\mathbf{A} = \mathbf{U}\Lambda\mathbf{U}^{-1}$. [4]

(b) (i) The difference equations may be written in the form $\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k$, where $\mathbf{x}_k = [\,y_k\ z_k\,]^T$ and

$$\mathbf{A} = \begin{bmatrix} 0.7 & 0.1 \\ 0.3 & 0.9 \end{bmatrix}$$ [2]

(ii) From the recurrence relation $\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k$, it follows that $\mathbf{x}_1 = \mathbf{A}\mathbf{x}_0$, $\mathbf{x}_2 = \mathbf{A}\mathbf{A}\mathbf{x}_0$, and so on, giving the general solution $\mathbf{x}_k = \mathbf{A}^k \mathbf{x}_0$. Substituting the diagonal factorisation from (a), we obtain

$$\mathbf{x}_k = \mathbf{U}\Lambda\mathbf{U}^{-1}\mathbf{U}\Lambda\mathbf{U}^{-1}\mathbf{U}\Lambda\mathbf{U}^{-1}\ldots\mathbf{U}\Lambda\mathbf{U}^{-1}\mathbf{x}_0 = \mathbf{U}\Lambda^k\mathbf{U}^{-1}\mathbf{x}_0$$ [2]

(iii) The characteristic equation for $\mathbf{A}$ is

$$\begin{vmatrix} (0.7 - \lambda) & 0.1 \\ 0.3 & (0.9 - \lambda) \end{vmatrix} = 0 \Leftrightarrow (0.7 - \lambda)(0.9 - \lambda) - 0.03 = 0$$

$$\Leftrightarrow \lambda^2 - 1.6\lambda + 0.6 = 0 \Leftrightarrow \lambda = 1 \text{ or } \lambda = 0.6$$

Now solve for the eigenvectors, starting with $\lambda = 1$. Set $u_2 = 1$ and solve for u_1 by back-substitution:

$$\begin{bmatrix} 0.7 & 0.1 \\ 0.3 & 0.9 \end{bmatrix}\begin{bmatrix} 1/3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 1 \end{bmatrix}$$

So the unit eigenvector corresponding to $\lambda = 1$ is $[\,0.3162\ 0.9487\,]^T$. Repeating for $\lambda = 0.6$, we get

$$\begin{bmatrix} 0.7 & 0.1 \\ 0.3 & 0.9 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \end{bmatrix} = 0.6 \times \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

So the unit eigenvector corresponding to $\lambda = 0.6$ is $[\,-0.7071\ 0.7071\,]^T$.

The solution to the difference equations is therefore

$$\begin{aligned}
\mathbf{x}_k &= \begin{bmatrix} 0.3162 & -0.7071 \\ 0.9487 & 0.7071 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0.6 \end{bmatrix}^k \begin{bmatrix} 0.3162 & -0.7071 \\ 0.9487 & 0.7071 \end{bmatrix}^{-1}\begin{bmatrix} 0 \\ 3 \end{bmatrix} \\
&= \begin{bmatrix} 0.3162 & -0.7071 \\ 0.9487 & 0.7071 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0.6 \end{bmatrix}^k \begin{bmatrix} 0.7906 & 0.7906 \\ -1.061 & 0.3536 \end{bmatrix}\begin{bmatrix} 0 \\ 3 \end{bmatrix}
\end{aligned}$$

8

In the limit $k \to \infty$, we have

$$\mathbf{x}_k \to \begin{bmatrix} 0.3162 & -0.7071 \\ 0.9487 & 0.7071 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0.7906 & 0.7906 \\ -1.061 & 0.3536 \end{bmatrix} \begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 0.75 \\ 2.25 \end{bmatrix} \qquad [8]$$

(c) The eigenvalues of $\mathbf{A}^{-1}$ are 1.0 and 1.667 (the reciprocals of the eigenvalues of $\mathbf{A}$). The eigenvectors of $\mathbf{A}^{-1}$ are the same as the eigenvectors of $\mathbf{A}$. The solution of the difference equation is therefore

$$\mathbf{x}_k = \mathbf{U} \begin{bmatrix} 1 & 0 \\ 0 & 1.667 \end{bmatrix}^k \mathbf{U}^{-1}\mathbf{x}_0$$

So, unless $\mathbf{x}_0 = \mathbf{0}$, $|\mathbf{x_k}| \to \infty$ as $k \to \infty$ (since one of the eigenvalues is greater than 1). If $\mathbf{x}_0 = \mathbf{0}$, $\mathbf{x}_k = \mathbf{0}$ for all k. $\qquad [4]$

Examiner's remarks: This straightforward question asked the candidates to apply simple eigenvector techniques to analyse a system of two linear difference equations (describing a Markov process, incidentally). While some candidates scored close to full marks, a disappointing number struggled with even the simplest of calculations, like finding the eigenvalues and eigenvectors of a 2×2 matrix. Fewer than half the candidates were able to prove the fundamental relationship $\mathbf{A} = \mathbf{U \Lambda U}^{-1}$. Only a handful successfully negotiated a few lines of algebraic expansion to find the limiting values of the difference equations.

6. *Sampling and reconstruction.*

(a) The signal must be sampled at angular frequency $\omega \geq 2\omega_m$, so the limiting value of T is π/ω_m. The spectrum of $x_s(t)$ is a scaled and repeated version of the spectrum of $x(t)$, with repetition frequency $2\pi/T$. Assuming (arbitrarily) a triangular spectrum for $x(t)$, bandlimited to $|\omega| < \omega_m$, we have:

$\qquad [4]$
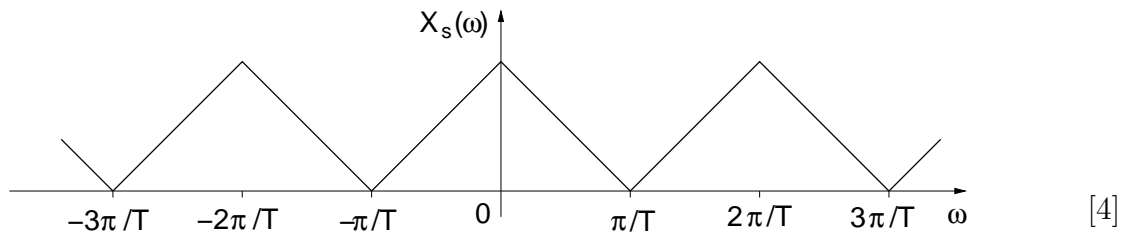
(b) The Fourier transform of $x_s(t)$ is given by

$$\begin{aligned} X_s(\omega) &= \int_{-\infty}^{\infty} x_s(t)e^{-i\omega t}\,dt = \int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} x(t)\delta(t - nT)e^{-i\omega t}\,dt \\ &= \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} x(t)\delta(t - nT)e^{-i\omega t}\,dt = \sum_{n=-\infty}^{\infty} x(nT)e^{-i\omega nT} \ , \end{aligned}$$

9

by the sifting property of delta functions. The ideal reconstruction filter has frequency response

$$H(\omega) = \begin{cases} T & \text{for } |\omega| \leq \pi/T \\ 0 & \text{otherwise} \end{cases}$$

so, taking the inverse Fourier transform of $H(\omega)X_s(\omega)$ we can recover the original signal as follows:

$$
\begin{aligned}
x(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} H(\omega)X_s(\omega)e^{i\omega t}\,d\omega = \frac{T}{2\pi}\int_{-\pi/T}^{\pi/T} X_s(\omega)e^{i\omega t}\,d\omega \\
&= \frac{T}{2\pi}\int_{-\pi/T}^{\pi/T}\sum_{n=-\infty}^{\infty} x(nT)e^{-i\omega nT}e^{i\omega t}\,d\omega = \frac{T}{2\pi}\sum_{n=-\infty}^{\infty} x(nT)\int_{-\pi/T}^{\pi/T} e^{i\omega(t-nT)}\,d\omega \\
&= \frac{T}{2\pi}\sum_{n=-\infty}^{\infty} x(nT)\left[\frac{e^{i\omega(t-nT)}}{i(t-nT)}\right]_{-\pi/T}^{\pi/T} = \frac{T}{2\pi}\sum_{n=-\infty}^{\infty} x(nT)\left(\frac{e^{i\pi(t-nT)/T} - e^{-i\pi(t-nT)/T}}{i(t-nT)}\right) \\
&= \frac{T}{2\pi}\sum_{n=-\infty}^{\infty} x(nT)\frac{2i\sin\pi(t-nT)/T}{i(t-nT)} = \sum_{n=-\infty}^{\infty} x(nT)\frac{\sin\pi(t-nT)/T}{\pi(t-nT)/T} \\
&= \sum_{n=-\infty}^{\infty} x(nT)\operatorname{sinc}\frac{\pi(t-nT)}{T}\ .
\end{aligned}
$$

[8]

(c) The reconstruction formula in (b) allows the original signal to be derived from its samples by summing a number of sinc pulses, with one pulse for each sample. This is infeasible for real-time applications, since the reconstructed signal at time t depends on an *infinite* sum over all past *and future* samples. The sinc function is also rather expensive to compute.

$x_0(t)$ is a zero-order hold reconstruction, found by passing the weighted impulse train $x_s(t)$ through a filter with the following impulse and frequency responses (the frequency response is a sinc function, as given in the electrical data book):
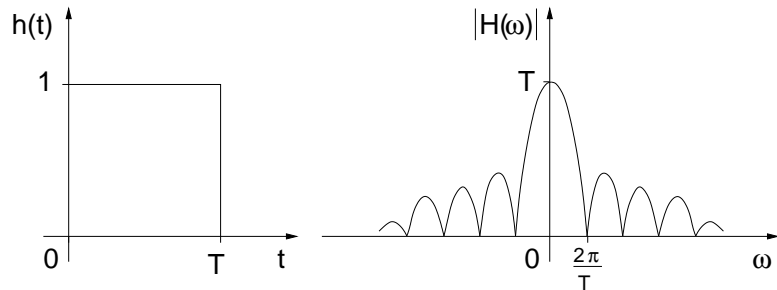
$x_1(t)$ is a first-order hold reconstruction, found by passing the weighted impulse train $x_s(t)$ through a filter with the following impulse and frequency responses (the frequency response is a sinc^2 function, as given in the electrical data book):
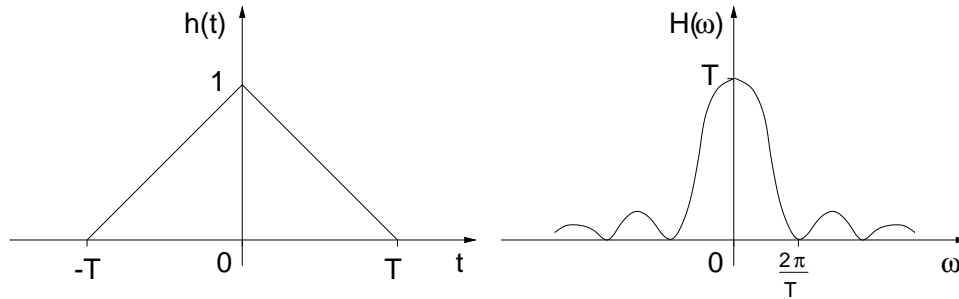
In practice, the first-order hold reconstruction is implemented with a T second delay, so that the filter's impulse response starts at $t = 0$ and not $t = -T$. [8]

Examiner's remarks: This question tested the candidate's knowledge of sampling and reconstruction techniques. The mathematics involved finding a simple forward Fourier transform, then a slightly more tricky (in that it required five lines of algebra) inverse Fourier transform. The level of understanding of this material was extremely poor. Most candidates could not even deduce the maximum sampling period to avoid aliasing. When asked to sketch the spectrum of the sampled signal, around half the candidates sketched a graph of something (it was hard to tell what) plotted against time. Only a handful of candidates made a serious attempt at the inverse Fourier transform. The last part of the question asked candidates to describe filters for sample-and-hold and linear interpolation filters. Unfortunately, hardly anyone even attempted this part of the question.

7. *Fourier transforms, probability density functions.*

   (a) (i) $x(t)$ can be written as the product of two time-domain signals:

$$x(t) = \cos \omega_0 t \times H(t - T/2) \;,$$

where $H(t)$ is a unit pulse of width T centred on the origin. The convolution theorem tells us that

$$X(\omega) = \frac{1}{2\pi} \mathcal{F}(\cos \omega_0 t) * \mathcal{F}(H(t - T/2)) \;.$$

Now, from the electrical data book we know that

$$\mathcal{F}(\cos \omega_0 t) = \pi \left[\delta(\omega - \omega_0) + \delta(\omega + \omega_0) \right]$$

and from the shift in time theorem we have

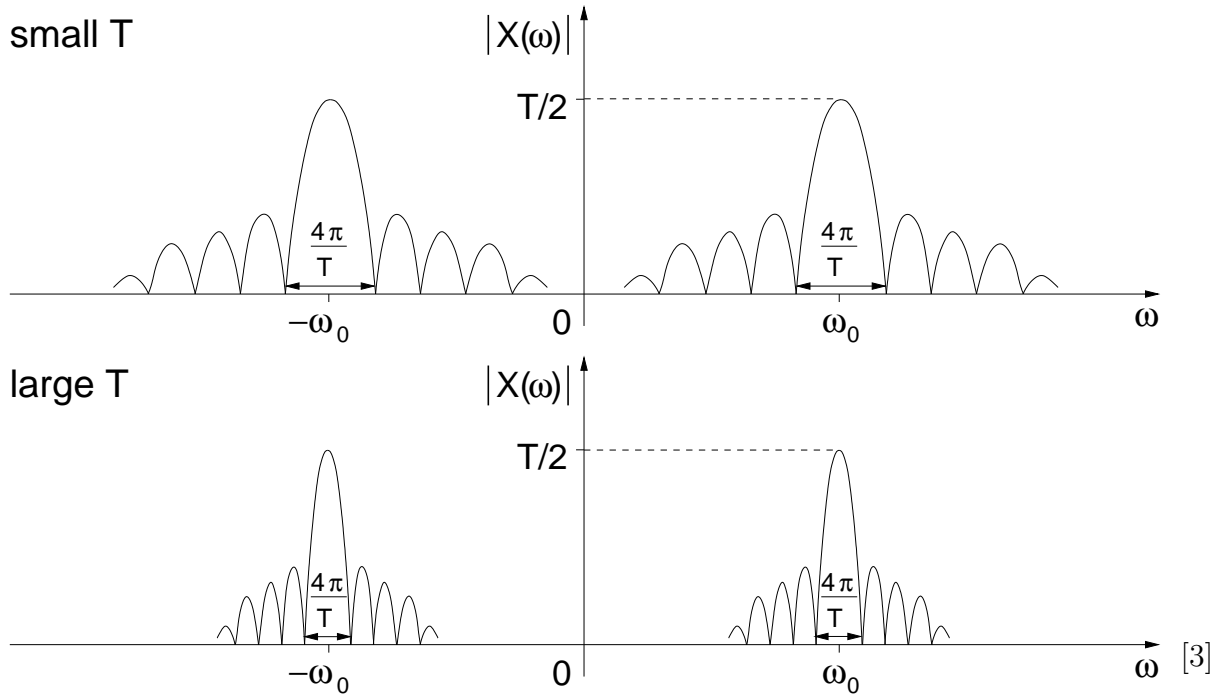$$\mathcal{F}(H(t - T/2)) = e^{-i\omega T/2} \mathcal{F}(H(t)) \;.$$

Finally, from the electrical data book we know that

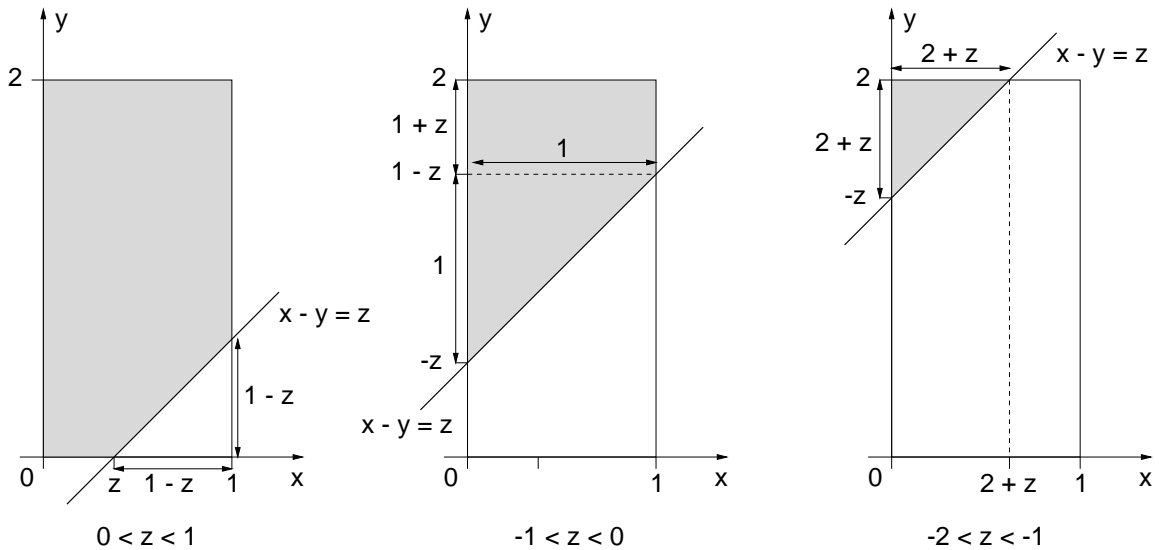$$\mathcal{F}(H(t)) = T \mathrm{sinc}\left(\frac{\omega T}{2} \right) \;.$$

11

Combining this information, we have

$$X(\omega) = \frac{1}{2\pi}\,\pi\,[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)] * e^{-i\omega T/2}\,T\mathrm{sinc}\left(\frac{\omega T}{2}\right)$$

$$= \frac{T}{2}\left[e^{-i(\omega-\omega_0)T/2}\,\mathrm{sinc}\left(\frac{(\omega - \omega_0)T}{2}\right) + e^{-i(\omega+\omega_0)T/2}\,\mathrm{sinc}\left(\frac{(\omega + \omega_0)T}{2}\right)\right] \ . \quad [7]$$

(ii) The modulus of the spectrum looks like this for large and small T:

(b) (i) The values of X and Y are uniformly distributed inside a 1×2 rectangle.

12

For the case $0 \leq z \leq 1$:

$$
\begin{aligned}
F(z) &= \mathrm{P}(Z < z) = \frac{2 - \frac{1}{2}(1-z)^2}{2} = \frac{1}{2}\left(\frac{3}{2} + z - \frac{z^2}{2}\right) \\
\Rightarrow f(z) &= \frac{dF(z)}{dz} = \frac{1}{2}(1-z) \ .
\end{aligned}
$$

For the case $-1 \leq z \leq 0$:

$$
\begin{aligned}
F(z) &= \mathrm{P}(Z < z) = \frac{(1+z) + \frac{1}{2}}{2} = \frac{1}{2}\left(\frac{3}{2} + z\right) \\
\Rightarrow f(z) &= \frac{dF(z)}{dz} = \frac{1}{2} \ .
\end{aligned}
$$

For the case $-2 \leq z \leq -1$:

$$
\begin{aligned}
F(z) &= \mathrm{P}(Z < z) = \frac{\frac{1}{2}(2+z)^2}{2} = \frac{1}{2}\left(2 + 2z + \frac{z^2}{2}\right) \\
\Rightarrow f(z) &= \frac{dF(z)}{dz} = \frac{1}{2}(2+z) \ .
\end{aligned}
$$

The overall probability density function looks like this:
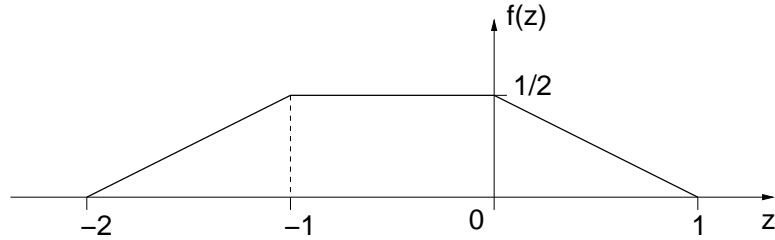
We have arrived at a valid probability density function, since the area under the curve is one (by inspection). [8]

(ii) From the sketch above, it is clear that $\mathrm{E}(Z) = -0.5$. This can also be deduced by $\mathrm{E}(Z) = \mathrm{E}(X) - \mathrm{E}(Y) = 0.5 - 1 = -0.5$. [2]

Examiner's remarks: The first part of this question asked the candidates to calculate the Fourier transform of a truncated cosine wave (demonstrating spectral leakage). Recognising that the truncated cosine wave is the product of an infinite cosine wave and a finite duration pulse, the maths involved copying a couple of formulae from the data book and performing a trivial convolution. However, almost all candidates attempted to find the Fourier transform by the direct route, and most got hopelessly lost. The second part of the question asked the candidates to calculate the probability density function (pdf) of $Z = X - Y$, where X and Y are

independent random variables uniformly distributed over different ranges. Very few candidates arrived at the correct answer, with some even suggesting that a pdf could take negative values. Even more worrying, only about a quarter of the candidates could calculate the mean of Z, given the means of X and Y.

8. *Probability.*

(a) (i) The number of faulty items in a sample of 10 would be well modelled by the Binomial distribution if:

- the probability of selecting a faulty item is p and remains the same as we select more items;

- each selection is independent.

Neither of these conditions is strictly true, since we are dealing with a finite population of size 500. Thus, the probability of selecting a faulty item depends on the previous selections. However, with a large population and a relatively small sample, we would expect the Binomial distribution to be a good approximation. [4]

(ii) The exact probability that the batch is shipped can be calculated by considering combinations.

P(no faulty items in sample of 10) =
$$\frac{\text{number of ways of choosing 10 non-faulty items from batch}}{\text{number of ways of choosing 10 items from batch}}$$
$$= \frac{^{495}C_{10}}{^{500}C_{10}} = \frac{495!/(485! \times 10!)}{500!/(490! \times 10!)} = \frac{495!/485!}{500!/490!}$$
$$= \frac{495 \times 494 \times \ldots \times 487 \times 486}{500 \times 499 \times \ldots \times 492 \times 491} = \frac{490 \times 489 \times 488 \times 487 \times 486}{500 \times 499 \times 498 \times 497 \times 496} = 0.9035 \qquad [4]$$

(iii) The Binomial approximation is much quicker to calculate. If we assume that the number of faulty items X in a sample is distributed according to $X \sim \mathrm{B}(10, 0.01)$, then
$$\mathrm{P}(X = 0) = (0.99)^{10} = 0.9044 \qquad [2]$$

(b) (i) The central limit theorem states that a combination of n identically distributed, independent random variables has a distribution which becomes Normal as n becomes large, irrespective of how the typical individual is distributed.

We can use the central limit theorem to estimate the probability of different waiting times for later appointments. This probability depends on the accumulated consultation times of all earlier appointments. Even though each consultation time is uniformly distributed, the sum of n consultation times will tend to be Normally distributed as n becomes large. [3]

(ii) If I have the 9.15am appointment, then I will have to wait for more than two minutes if the 9am consultation takes longer than 17 minutes. The time taken for the 9am consultation is uniformly distributed in the range 12 to 19 minutes, so

$$P(\text{wait more than two minutes}) = 2/7 = 0.286 \qquad [2]$$

(iii) If I have the 5.30pm appointment, then the time I have to wait depends on the accumulated consultation times of the 34 previous appointments. Each consultation time is uniformly distributed with mean 15.5 minutes and variance $49/12$ (see the mathematics data book for the mean and variance of a uniform distribution). Assuming consecutive consultation times are independent, the time X for 34 consultations will have mean $34 \times 15.5 = 527$ minutes, and variance $34 \times (49/12) = 138.8$ minutes. The central limit theorem tells us that the distribution of X will be approximately Normal, so $X \sim N(527, \sqrt{138.8})$. I will have to wait more than two minutes if X exceeds 512 minutes (8 hours and 32 minutes).

We now normalise the limiting time of 512 minutes, to see where it lies on the standard Normal curve:

$$\frac{512 - 527}{\sqrt{138.8}} = -1.273 \quad \text{and} \quad \phi(-1.273) = 1 - 0.899$$

So there is an 89.9% chance that I will have to wait more than two minutes for my 5.30pm appointment. The principal assumption we have made is that the doctor never has to wait for the patient: we have assumed that the patient is always in the waiting room, even if the doctor is ahead of schedule. [5]

Examiner's remarks: The first part of this question tested the candidates' understanding of the Binomial distribution and its application to sampling problems. The second part concerned the central limit theorem and the use of the Normal distribution. Understanding of this material was generally good. Most candidates could manipulate the Binomial distribution, though many thought it a perfect model for sampling-without-replacement scenarios. Most could quote the central limit theorem and knew more or less how to use the Normal tables, though the usual lack of attention to detail meant that few managed to get the right answer.

Andrew Gee
June 2002