Q1

(a) $$\boldsymbol{M} = \boldsymbol{U\Sigma V}^T$$

Recall that for a symmetric matrix $\boldsymbol{S}$

$$\boldsymbol{S} = \boldsymbol{U\Lambda U}^T$$

where the columns of $\boldsymbol{U}$ are the normalised eigenvectors of $\boldsymbol{S}$, and $\boldsymbol{\Lambda} = \text{diag}(\lambda_1,...,\lambda_n)$.

$\boldsymbol{MM}^T$ is symmetric.

$$\boldsymbol{MM}^T = \boldsymbol{U\Sigma V}^T\left(\boldsymbol{U\Sigma V}^T\right)^T = \boldsymbol{U\Sigma V}^T\boldsymbol{V\Sigma}^T\boldsymbol{U}^T = \boldsymbol{U\Sigma\Sigma}^T\boldsymbol{U}^T$$

Hence the columns of $\boldsymbol{U}$ are the normalised eigenvectors of $\boldsymbol{MM}^T$, and

$$\boldsymbol{\Sigma\Sigma}^T = \boldsymbol{\Lambda} = \text{diag}(\lambda_1,...,\lambda_n)$$

where $\lambda_1,...,\lambda_n$ are the eigenvalues of $\boldsymbol{MM}^T$.

$$\therefore \quad \boldsymbol{\Sigma} = \text{diag}\left(\sqrt{\lambda_1},...,\sqrt{\lambda_n}\right)$$

$\boldsymbol{M}^T\boldsymbol{M}$ is also symmetric.

$$\boldsymbol{M}^T\boldsymbol{M} = \left(\boldsymbol{U\Sigma V}^T\right)^T\boldsymbol{U\Sigma V}^T = \boldsymbol{V\Sigma}^T\boldsymbol{U}^T\boldsymbol{U\Sigma V}^T = \boldsymbol{V\Sigma}^T\boldsymbol{\Sigma V}^T$$

Hence the columns of $\boldsymbol{V}$ are the normalised eigenvectors of $\boldsymbol{M}^T\boldsymbol{M}$.

Note that non-zero eigenvalues of $\boldsymbol{MM}^T$ and $\boldsymbol{M}^T\boldsymbol{M}$ are the same.

Note that eigenvalues of $\boldsymbol{M}^T\boldsymbol{M}$ are non-negative, i.e.

$$\boldsymbol{M}^T\boldsymbol{Mx} = \lambda\boldsymbol{x} \quad \Rightarrow \quad \boldsymbol{x}^T\boldsymbol{M}^T\boldsymbol{Mx} = \lambda\boldsymbol{x}^T\boldsymbol{x}$$

$$\therefore \quad \|\boldsymbol{Mx}\|_2^2 = \lambda\|\boldsymbol{x}\|_2^2$$

$\|\boldsymbol{Mx}\|_2^2$ and $\|\boldsymbol{x}\|_2^2$ are both > 0 by definition, therefore $\lambda \geq 0$, which implies that $\boldsymbol{\Sigma}$ is real. [40%]

(b) $\boldsymbol{M}_1$: invalid because $\Sigma_2 < 0$.

$\boldsymbol{M}_2$: valid, rank = 2.

$\boldsymbol{M}_3$: valid, rank = 1.

$\boldsymbol{M}_4$: invalid, $\boldsymbol{\Sigma}$ is not diagonal.

$\boldsymbol{M}_5$: invalid, $\boldsymbol{U}$ is not orthogonal. [20%]

(c)(i) $\kappa_2 = \dfrac{9}{1.2 \times 10^{-9}} = 7.5 \times 10^9$

This is large, so the solution will be sensitive to small perturbations in the right-hand side and round-off error. [10%]

(ii) The exact rank is 3, but since $\Sigma_3$ is small, effective rank is 2. [5%]

(iii) Drop small singular values in $\boldsymbol{M} \to \boldsymbol{M}^*$.

$$\boldsymbol{M}^* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 9 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix}$$

$$\therefore \quad \boldsymbol{U} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}; \quad \boldsymbol{\Sigma} = \begin{bmatrix} 9 & 0 \\ 0 & 3 \end{bmatrix}; \quad \boldsymbol{V}^T = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix}$$

$$\boldsymbol{M}^{*-1} = \boldsymbol{V}\boldsymbol{\Sigma}^{-1}\boldsymbol{U}^T \quad \Rightarrow \quad \boldsymbol{x} = \boldsymbol{V}\boldsymbol{\Sigma}^{-1}\boldsymbol{U}^T\boldsymbol{b}$$

$$\therefore \quad \boldsymbol{x} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/9 & 0 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 7 \\ 15 \\ 6 \end{bmatrix}$$

$$\therefore \quad \boldsymbol{x} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/9 & 0 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} 7 \\ 15 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 7/9 \\ 5 \end{bmatrix} = \begin{bmatrix} \frac{7}{9\sqrt{2}} - \frac{5}{\sqrt{2}} \\ \frac{7}{9\sqrt{2}} + \frac{5}{\sqrt{2}} \\ 0 \end{bmatrix}$$

This is the best fit that minimises $\|\boldsymbol{x}\|_2$. [25%]

**Assessor's comments:**

37 attempts, Average mark 12.2/20, Maximum 20, Minimum 1.
Not a popular question, attempted by well under half the candidates. Attempts ranged from the perfect to the derisory.
Few attempts at part (a) were totally rigorous, although many were on the right lines.
The most common error was the failure to spot that $\boldsymbol{U}$ for $\boldsymbol{M}_5$ in part (b) was not orthogonal.
Most candidates who attempted (c)(iii) had a good idea of how to find $x$ but quite a few made errors in inverting the matrices or in matrix multiplication.

Q2

(a) Maximising the discharge capacity is equivalent to minimising $p$

$$\therefore \quad \text{Minimise} \quad p = b + 2h\csc\phi = f$$

$$\text{subject to} \quad hb + h^2\cot\phi = A$$

$$\text{and} \quad h \geq 0; \quad b \geq 0; \quad \phi \geq 0$$

[10%]

(b)

$$hb = A - h^2\cot\phi$$

$$\therefore \quad b = \frac{A}{h} - h\cot\phi$$

$$\therefore \quad \text{Minimise} \quad p = \frac{A}{h} - h\cot\phi + 2h\csc\phi = f$$

[10%]

(c) Newton's Method:

$$\underline{x}_{k+1} = \underline{x}_k - \underline{\underline{H}}(\underline{x}_k)^{-1}\underline{\nabla f}(\underline{x}_k)$$

Noting that $\frac{d}{d\phi}\cot\phi = -\csc^2\phi$ and $\frac{d}{d\phi}\csc\phi = -\csc\phi\cot\phi$

Here
$$\underline{\nabla f} = \begin{bmatrix} \frac{\partial f}{\partial h} \\ \frac{\partial f}{\partial \phi} \end{bmatrix} = \begin{bmatrix} -\frac{A}{h^2} - \cot\phi + 2\csc\phi \\ h\csc^2\phi - 2h\csc\phi\cot\phi \end{bmatrix}$$

$$\underline{\underline{H}} = \begin{bmatrix} \frac{2A}{h^3} & \csc^2\phi - 2\csc\phi\cot\phi \\ \csc^2\phi - 2\csc\phi\cot\phi & 2h\left(-\csc^2\phi\cot\phi + \csc\phi\cot^2\phi + \csc^3\phi\right) \end{bmatrix}$$

For $\underline{x}_1 = \begin{bmatrix} 5 \\ \pi/4 \end{bmatrix}$
$$\underline{\nabla f}(\underline{x}_1) = \begin{bmatrix} -2.1716 \\ -4.1421 \end{bmatrix}$$

$$\underline{\underline{H}}(\underline{x}_1) = \begin{bmatrix} 1.6 & -0.82843 \\ -0.82843 & 22.4264 \end{bmatrix} \Rightarrow \left\|\underline{\underline{H}}(\underline{x}_1)\right\| = 35.196$$

$$\therefore \quad \underline{x}_2 = \begin{bmatrix} 5 \\ \pi/4 \end{bmatrix} - \frac{1}{35.196}\begin{bmatrix} 22.4264 & 0.82843 \\ 0.82843 & 1.6 \end{bmatrix}\begin{bmatrix} -2.1716 \\ -4.1421 \end{bmatrix} = \begin{bmatrix} 6.4812 \\ 1.0248 \end{bmatrix}$$

[45%]

(d) At the optimum $\underline{\nabla f} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$$\therefore \quad -\frac{A}{h^2} - \cot\phi + 2\csc\phi = 0 \qquad (2.1)$$

and
$$h\csc^2\phi - 2h\csc\phi\cot\phi = 0 \qquad (2.2)$$

$$\therefore \quad \csc\phi = 2\cot\phi$$

$$\therefore \quad \frac{1}{\sin\phi} = 2\frac{\cos\phi}{\sin\phi}$$

$$\therefore \quad \cos\phi = \tfrac{1}{2} \quad \Rightarrow \quad \phi = \frac{\pi}{3} = 1.04720$$

Rearranging equation (2.1)

$$h^2 = \frac{A}{2\csc\phi - \cot\phi}$$

$$\therefore \quad h^2 = \frac{100}{2\csc(\pi/3) - \cot(\pi/3)} = \frac{100}{\frac{4}{\sqrt{3}} - \frac{1}{\sqrt{3}}} = \frac{100}{\sqrt{3}}$$

$$\therefore \quad h = \sqrt{\frac{100}{\sqrt{3}}} = 7.5984 \text{ m}$$

At this point $\underline{\underline{H}} = \begin{bmatrix} 0.4559 & 0 \\ 0 & 17.548 \end{bmatrix}$

This is clearly (by inspection) positive definite, therefore this is indeed a minimum.

Newton's Method seems to be converging on this point. If $f$ was quadratic Newton's Method would converge in one iteration. In this case $f$ is not quadratic so convergence takes longer. [35%]

**Assessor's comments:**

93 attempts, Average mark 12.0/20, Maximum 20, Minimum 6.
A very popular question attempted by all but six candidates. There were many reasonable attempts but very few really good ones due to a litany of different errors.
Many answers were undermined by poor differentiation skills, poor algebra, poor attention to detail, transcription errors and calculator errors.
Many candidates overlooked the need for non-negativity bounds in part (a) despite the hint in part (b).
Bewilderingly, more than one candidate failed to show the required result in part (b) and then proceeded to try to do the rest of the question using their erroneous version of $f$.
It is apparent that several candidates were unfamiliar with the definitions of $\csc\phi$, $\sec\phi$ and $\cot\phi$.
Some candidates did not appreciate that $\frac{\partial^2 f}{\partial h \partial \phi}$ must equal $\frac{\partial^2 f}{\partial \phi \partial h}$ and tried to apply Newton's method with an asymmetric Hessian.
Several candidates apparently did not know how to invert a 2×2 matrix.
There were several suggestions of using Lagrange multipliers in part (d) rather than simply exploiting results from part (c) in applying the optimality criteria for an unconstrained minimum.
Several candidates found the correct solution in part (d) but failed to check that the Hessian was positive definite there.

Q3

(a) The identification of constraints $g_3$ to $g_6$ is very straightforward.

$g_1$:
$$T = \frac{2\pi f p}{3\sin\alpha}\left(R_1^3 - R_2^3\right) \geq 15{,}000 \text{ N cm}$$

Substituting for $p$ $$T = \frac{2\pi f}{3\sin\alpha}\frac{F}{\pi\left(R_1^2 - R_2^2\right)}\left(R_1^3 - R_2^3\right) \geq 15{,}000 \text{ N cm}$$

$$\therefore \quad T = \frac{2f}{3\sin\alpha}\frac{F}{\left(R_1^2 - R_2^2\right)}\left(R_1^3 - R_2^3\right) \geq 15{,}000 \text{ N cm}$$

For the given values of $\alpha$, $f$ and $F$

$$\therefore \quad T = 500\frac{\left(R_1^3 - R_2^3\right)}{\left(R_1^2 - R_2^2\right)} \geq 15{,}000 \text{ N cm}$$

Noting that $R_1^3 - R_2^3 = (R_1 - R_2)\left(R_1^2 + R_1R_2 + R_2^2\right)$ and that $R_1^2 - R_2^2 = (R_1 - R_2)(R_1 + R_2)$

$$\therefore \quad T = \frac{500}{(R_1 + R_2)}\left(R_1^2 + R_1R_2 + R_2^2\right) \geq 15{,}000 \text{ N cm}$$

$$\therefore \quad R_1^2 + R_1R_2 + R_2^2 \geq 30(R_1 + R_2) \qquad (3.1)$$

This is a symmetric function in $R_1$ and $R_2$. When equality holds, this line passes through (0,30) and (30,0). There is only one possible candidate in the figure (as shown below). For the inequality to be satisfied, we need to be to the right of this line.

$g_2$:
$$p = \frac{F}{\pi\left(R_1^2 - R_2^2\right)} \leq \frac{15}{4\pi} \text{ N cm}^{-2}$$

For the given value of $F$

$$\therefore \quad p = \frac{750}{\pi\left(R_1^2 - R_2^2\right)} \leq \frac{15}{4\pi} \text{ N cm}^{-2}$$

$$\therefore \quad 200 \leq R_1^2 - R_2^2 \qquad (3.2)$$

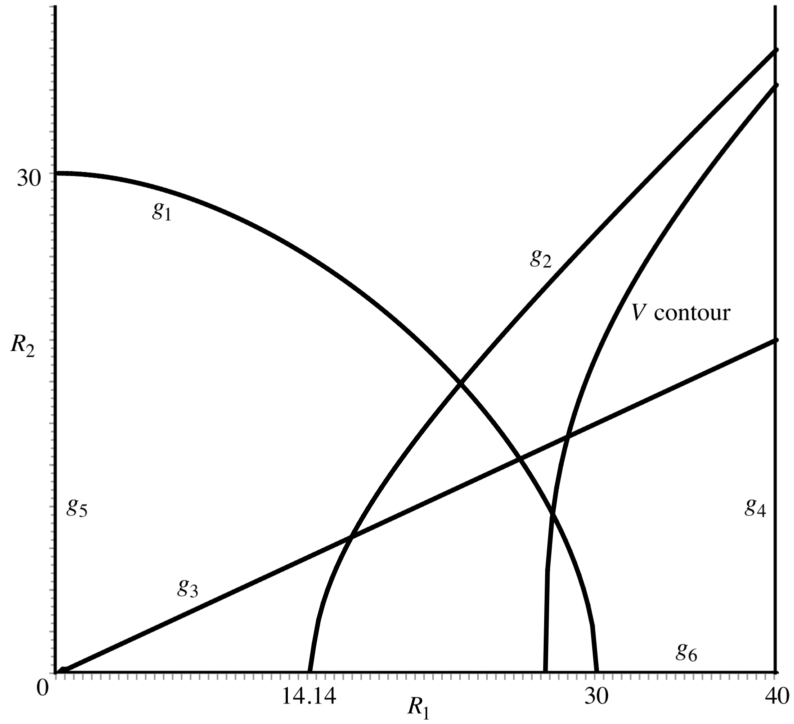
When equality holds, this defines a hyperbola passing through $\left(\sqrt{200},0\right) = (14.14,0)$, resulting in the identification shown. For the inequality to be satisfied, we need to be to the right of this line.

Contours of $V$ are lines where

$$V = \tfrac{1}{3}\pi\cot\alpha\left(R_1^3 - R_2^3\right) = \text{const}$$

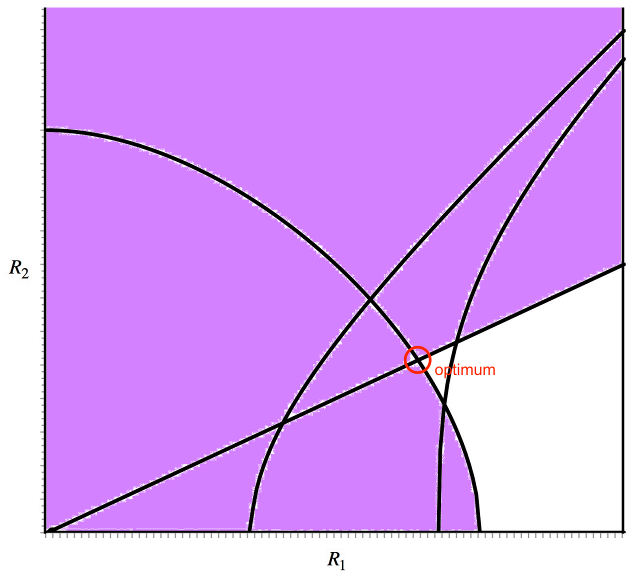
$$\therefore \quad R_2^3 = R_1^3 - \frac{3V\tan\alpha}{\pi}$$

In $R_1$-$R_2$ space such lines are a bit like hyperbolae (e.g. asymptotically convergent on $R_2 = R_1$ for large $R_1$) but have a steeper gradient at $R_2 = 0$, leading to the identification shown.




[40%]

(b) The constraints render infeasible the areas shaded in the figure below:




Contours of $V$ move to the left as $V$ decreases, thus the optimum lies at the intersection of $g_1$ and $g_3$.

On $g_3$ $R_2 = \frac{1}{2}R_1$, so, using the equation (3.1) version of $g_1$:

$$R_1^2 + \tfrac{1}{2}R_1^2 + \tfrac{1}{4}R_1^2 = 30(R_1 + \tfrac{1}{2}R_1)$$

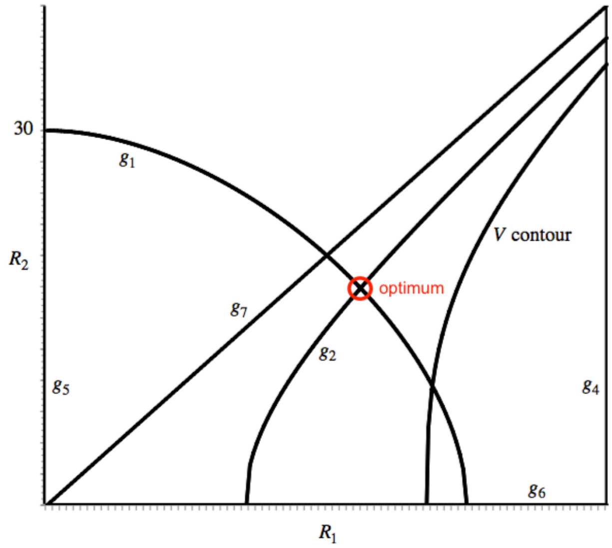
$$\therefore \quad \tfrac{7}{4}R_1^2 = 45R_1 \quad \Rightarrow \quad R_1 = \frac{180}{7} = 25.71\,\text{cm}$$

$$\therefore \quad R_2 = \tfrac{1}{2}R_1 = \frac{180}{14} = 12.86\,\text{cm}$$

$$\therefore \quad V = \tfrac{1}{3}\pi\cot\alpha\left(R_1^3 - R_2^3\right) = \tfrac{1}{3}\pi\sqrt{3}\left[\left(\frac{180}{7}\right)^3 - \left(\frac{180}{14}\right)^3\right] = 26.98 \times 10^3\ \text{cm}^3$$

[25%]

(c) If $g_3$ is eliminated, $g_7$ is unnecessary because it is rendered redundant by $g_2$ – any solution that violates $g_7$ also violates $g_2$ (but not vice versa). This is particularly obvious if $g_7$ is plotted:




The optimum will now lie at the intersection of $g_1$ and $g_2$. [15%]

(d) If $T = \dfrac{\pi f p R_2}{\sin\alpha}\left(R_1^2 - R_2^2\right)$, then substituting for $p$

$$T = \frac{\pi f R_2}{\sin\alpha}\frac{F}{\pi\left(R_1^2 - R_2^2\right)}\left(R_1^2 - R_2^2\right) = \frac{f R_2 F}{\sin\alpha}$$

For the given values of $\alpha$, $f$ and $F$

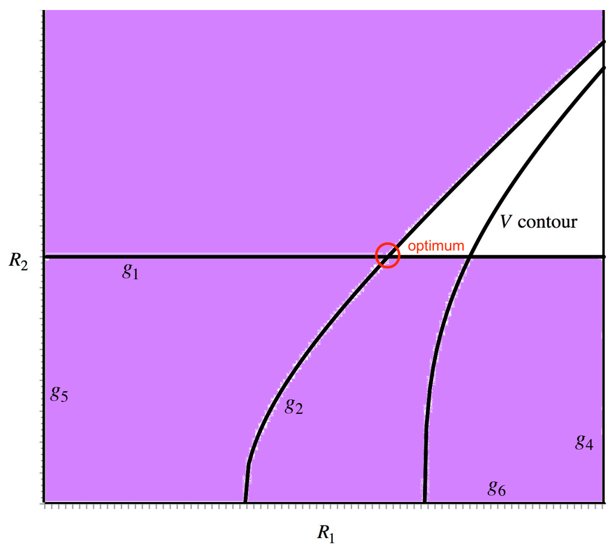
$$\therefore \quad T = 750 R_2 \; \geq \; 15{,}000\ \text{N cm} \quad \Rightarrow \quad R_2 \; \geq \; 20\ \text{cm}$$

This is a new version of constraint $g_1$. The feasible space is now as shown below, with the optimum occurring at the intersection of $g_2$ and the new $g_1$.

Obviously $R_2 = 20$ cm. Using the equation (3.2) version of $g_2$:

$$R_1^2 = 200 + R_2^2 = 200 + 20^2 \quad \Rightarrow \quad R_1 = \sqrt{600} = 24.49\ \text{cm}$$

$$\therefore \quad V = \tfrac{1}{3}\pi\cot\alpha\left(R_1^3 - R_2^3\right) = \tfrac{1}{3}\pi\sqrt{3}\left[600^{\frac{3}{2}} - 20^3\right] = 12.15\times 10^3\ \text{cm}^3$$




[20%]

**Assessor's comments:**

75 attempts, Average mark 11.6/20, Maximum 20, Minimum 1.
A popular question attempted by 76% of candidates.
The external examiner expressed concerns that this question would prove to be too easy. These concerns proved to be unfounded. Although there were a number of excellent attempts, the general standard of answers was disappointing, not least because many of the failings were in fundamental mathematical skills rather than optimization knowledge.
Part (a), which was expected to be straightforward, proved surprisingly difficult for many candidates. Several lost marks by simply not identifying $g_4$, $g_5$ and $g_6$ on the figure provided, despite this being explicitly required by the question. A surprising number of other candidates got $g_5$ and $g_6$ the wrong way round.
Many candidates could not distinguish the equations of a circle and a hyperbola. More understandably, others had problems distinguishing the line representing $g_2$ from that representing a contour of $V$.
Several candidates overlooked the fact that $T$ is a function of $p$, which is, in turn, a function of $R_1$ and $R_2$.
Many candidates had problems visualising the feasible region, even when they had identified all the lines correctly, with consequent knock-on effects for later parts of the question. These were sometimes compounded by a failure to recognise where $V$ would be minimised, even when the line representing the $V$ contour was known. It was particularly common to identify the wrong side of the $g_1$ constraint as feasible.

Q4

(a) The transition matrix is

$$\mathbf{P} = \begin{bmatrix} 1-a & a & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0.2 & 0 & 0.3 & 0.5 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

By inspection, the stationary distribution must be

$$\boldsymbol{\pi} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

(unless $a = 0$, in which case $\boldsymbol{\pi} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \end{bmatrix}$) [15%]

(b) Using the notation that $q_i$ is the expected time to first visit State 5 given a start in State *i*, the equations associated with this are:

$$\begin{aligned} q_1 &= (1-a)(1+q_1) + a(1+q_2) \\ q_2 &= 1 + q_3 \\ q_3 &= 1 + q_4 \\ q_4 &= 0.5 + 0.3(1+q_4) + 0.2(1+q_2) \end{aligned}$$

It is then possible to write:

$$\begin{aligned} q_2 &= 2 + q_4 \\ 0.7q_4 &= 1 + 0.2q_2 \end{aligned}$$

$$\therefore \quad 0.7q_4 = 1 + 0.2(2+q_4) \quad \Rightarrow \quad 0.5q_4 = 1.4$$

$$\therefore \quad q_4 = 2.8 \quad \text{and} \quad q_2 = 4.8$$

Finally, solving for $q_1$:

$$q_1 = (1-a)(1+q_1) + a(1+q_2)$$

$$\therefore \quad aq_1 = 1 + aq_2$$

$$\therefore \quad q_1 = \frac{1}{a} + q_2 = \frac{1}{a} + 4.8$$

[30%]

(c) The distribution after *n* steps is

$$\boldsymbol{\pi}^{(n)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \end{bmatrix} \mathbf{P}^n$$

[10%]

(d)(i) As a new item enters State 1 every step, the distribution is now changed to

$$\boldsymbol{\pi}^{(N)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \end{bmatrix} \sum_{i=1}^{N} \mathbf{P}^i$$

This is a geometric progression but based on vectors. Using the standard proof for a GP yields:

$$\boldsymbol{\pi}^{(N)} - \boldsymbol{\pi}^{(N)}\mathbf{P} = \begin{bmatrix}1 & 0 & 0 & 0 & 0\end{bmatrix}\left(\sum_{i=1}^{N}\mathbf{P}^{i} - \sum_{i=1}^{N}\mathbf{P}^{i+1}\right) = \begin{bmatrix}1 & 0 & 0 & 0 & 0\end{bmatrix}\left(\mathbf{P}^{1} - \mathbf{P}^{N+1}\right)$$

$$\therefore \quad \boldsymbol{\pi}^{(N)}(\mathbf{I} - \mathbf{P}) = \begin{bmatrix}1 & 0 & 0 & 0 & 0\end{bmatrix}\left(\mathbf{P} - \mathbf{P}^{N+1}\right)$$

$$\therefore \quad \boldsymbol{\pi}^{(N)} = \begin{bmatrix}1 & 0 & 0 & 0 & 0\end{bmatrix}\left(\mathbf{P} - \mathbf{P}^{N+1}\right)(\mathbf{I} - \mathbf{P})^{-1}$$

Hence $\pi = \begin{bmatrix}1 & 0 & 0 & 0 & 0\end{bmatrix}$

$$\mathbf{A} = \mathbf{P}^{N+1}$$

$$\mathbf{B} = \mathbf{I} - \mathbf{P}$$

[30%]

(ii) “a” influences the probability of a single item entering the process. Thus the number of items stored in State 1 is determined by this. Again, this can be written as a geometric progression. As $N$ becomes large the system will converge on a steady state, as once an item enters the process the process it is not influenced by “a”. [15%]

**Assessor’s comments:**

90 attempts, Average mark 12.1/20, Maximum 19, Minimum 6.
A very popular question attempted by all but 9 candidates. Parts (a), (b) and (c) were generally done well, but part (d) baffled most candidates.
All but one candidate found the correct transition matrix. Surprisingly, a large number of candidates could not identify the stationary distribution in part (a) by inspection.
In part (b) some candidates erroneously thought that the expected time for an item to first enter State 5 should be an integer. Most candidates set up the equations governing the waiting times correctly but many then made small slips in solving them.
In part (d) most candidates missed the fact that if the process had been running for $N$ steps, $N$ items would have entered the process – one each step – and instead sought futilely to cast their answer to part (c) in the form required.
Few answers to (d)(ii) used the mark scheme to inform how much discussion was expected.