4H Numerical Linear Algebra & PDE's MATH4041
Epiphany Term: Solutions

1. From a theorem in lectures it is known that the Jacobi and Gauss-Seidel iterates converge for diagonally dominant matrices, thus both iterates converge when $|\rho| < 1$. Which converges faster though?! We know that $\boldsymbol{e}^{(k)} = \boldsymbol{x}^{(k)} - \boldsymbol{x} = M^k \boldsymbol{e}^{(0)}$, thus $\|M\|$ will give the speed of convergence. Noting that the eigenvalues of

$$M_J = \begin{pmatrix} 0 & \rho \\ \rho & 0 \end{pmatrix} \qquad M_{GS} = \begin{pmatrix} 0 & \rho \\ 0 & -\rho^2 \end{pmatrix}$$

are $\pm\rho$ and $0, -\rho^2$ respectively, we conclude $\|M_J\|_2 = |\rho|$ and $\|M_{GS}\|_2 = \rho^2$, i.e. the Gauss-Seidel method is much better.

2. Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ where $ad \neq 0$, calculating the iteration matrix for the Jacobi and Gauss-Seidel methods:

$$M_J = \begin{pmatrix} 0 & \frac{b}{a} \\ \frac{c}{d} & 0 \end{pmatrix} \qquad M_{GS} = \begin{pmatrix} 0 & \frac{b}{a} \\ 0 & -\frac{bc}{ad} \end{pmatrix}.$$

The eigenvalues of the iteration matrices are $\pm\sqrt{\frac{bc}{ad}}$ and $0, -\frac{bc}{ad}$ respectively. Thus, in both cases we require that $|\frac{bc}{ad}| < 1$ for convergence. The Gauss-Seidel will converge faster since the magnitude of the largest eigenvalue is smaller.

3. The matrix given in the question is diagonally dominant, so both Jacobi and Gauss-Seidel iterations are known to converge. The Jacobi and Gauss-Seidel iteration for the equation are

$$\begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \\ x_3^{(k+1)} \end{pmatrix} = \frac{1}{10} \begin{pmatrix} 15 - x_2^{(k)} - x_3^{(k)} \\ 24 - x_1^{(k)} - x_3^{(k)} \\ 33 - x_1^{(k)} - x_2^{(k)} \end{pmatrix} \text{ and } \begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \\ x_3^{(k+1)} \end{pmatrix} = \frac{1}{10} \begin{pmatrix} 15 - x_2^{(k)} - x_3^{(k)} \\ 24 - x_1^{(k+1)} - x_3^{(k)} \\ 33 - x_1^{(k+1)} - x_2^{(k+1)} \end{pmatrix}.$$

If $\boldsymbol{x}^{(0)} = (0, 0, 0)^T$ then the table of Jacobi and Gauss-Seidel iterations are:

$$\boldsymbol{x}^{(k)} : \quad \begin{pmatrix} 1.5 \\ 2.4 \\ 3.3 \end{pmatrix} \begin{pmatrix} 0.93 \\ 1.92 \\ 2.91 \end{pmatrix} \begin{pmatrix} 1.017 \\ 2.016 \\ 3.015 \end{pmatrix} \begin{pmatrix} 0.9969 \\ 1.9968 \\ 2.9967 \end{pmatrix} \begin{pmatrix} 1.00065 \\ 2.00064 \\ 3.00063 \end{pmatrix} \begin{pmatrix} 0.999873 \\ 1.999872 \\ 2.999871 \end{pmatrix}$$

$$\boldsymbol{x}^{(k)} : \quad \begin{pmatrix} 1.5 \\ 2.25 \\ 2.925 \end{pmatrix} \begin{pmatrix} 0.9825 \\ 2.00925 \\ 3.000825 \end{pmatrix} \begin{pmatrix} 0.9989925 \\ 2.00001825 \\ 3.000098925 \end{pmatrix} \begin{pmatrix} 0.9999882825 \\ 1.99999127925 \\ 3.000002043825 \end{pmatrix};$$

both appear to be converging to $(1, 2, 3)^T$. Using the inequality

$$\|\boldsymbol{x} - \boldsymbol{x}^{(k)}\|_\infty \leqslant \frac{\|M_J\|_\infty^k}{1 - \|M_J\|_\infty} \|\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)}\|_\infty$$

where M_J is the Jacobi iteration matrix, we can ensure that the Jacobi iterations are accurate to 6 decimal places by enforcing the inequality

$$\frac{\|M_J\|_\infty^k}{1-\|M_J\|_\infty}\|\boldsymbol{x}^{(1)}-\boldsymbol{x}^{(0)}\|_\infty \leqslant 5\times 10^{-7}.$$

A calculation reveals that $\|M_J\|_\infty = 1/5$ and $\|\boldsymbol{x}^{(1)}-\boldsymbol{x}^{(0)}\| = 33/10$ thus we require

$$\frac{5}{4\times 5^k}\times\frac{33}{10}\leqslant 5\times 10^{-7} \iff 5^k \geqslant \frac{33\times 10^7}{40} \iff k\geqslant 10.$$

4. The iteration Jacobi and Gauss-Seidel iteration matrices for A are:

$$M_J = \begin{pmatrix} 0 & a & 0 \\ a & 0 & a \\ 0 & a & 0 \end{pmatrix} \qquad M_{GS} = \begin{pmatrix} 0 & a & 0 \\ 0 & -a^2 & a \\ 0 & a^3 & -a^2 \end{pmatrix}.$$

The eigenvalues of M_J are $0, \pm\sqrt{2a^2}$ and of M_{GS} are $0, 0, -2a^2$. The Jacobi and Gauss-Seidel methods converge/diverge if the magnitude of the eigenvalues are $< / \geqslant 1$, i.e. we have convergence if $a < 1/\sqrt{2}$ and divergence if $a \geqslant 1/\sqrt{2}$.
Since $a < 1/\sqrt{2}$ the eigenvalues of the iteration matrix in the Gauss-Seidel method will be smaller and hence the method converges faster than Jacobi's method.
With $\boldsymbol{x}^{(0)} = (0,0,0)^T$ the Jacobi and Gauss-Seidel iterates are:

$$\boldsymbol{x}^{(k)}: \quad \begin{pmatrix} 2 \\ 4 \\ 4 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \begin{pmatrix} 1.5 \\ 3.0 \\ 3.5 \end{pmatrix} \begin{pmatrix} 0.5 \\ 1.5 \\ 2.5 \end{pmatrix} \begin{pmatrix} 1.25 \\ 2.50 \\ 3.25 \end{pmatrix} \begin{pmatrix} 0.75 \\ 1.75 \\ 2.75 \end{pmatrix} \begin{pmatrix} 1.125 \\ 2.250 \\ 3.125 \end{pmatrix} \begin{pmatrix} 0.875 \\ 1.875 \\ 2.875 \end{pmatrix} \begin{pmatrix} 1.0625 \\ 2.1250 \\ 3.0625 \end{pmatrix}$$

$$\boldsymbol{x}^{(k)}: \quad \begin{pmatrix} 2.0 \\ 3.0 \\ 2.5 \end{pmatrix} \begin{pmatrix} 0.50 \\ 2.50 \\ 2.75 \end{pmatrix} \begin{pmatrix} 0.750 \\ 2.250 \\ 2.875 \end{pmatrix} \begin{pmatrix} 0.8750 \\ 2.1250 \\ 2.9375 \end{pmatrix} \begin{pmatrix} 0.93750 \\ 2.06250 \\ 2.96875 \end{pmatrix} \begin{pmatrix} 0.968750 \\ 2.031250 \\ 2.984375 \end{pmatrix} \begin{pmatrix} 0.9843750 \\ 2.0156250 \\ 2.9921875 \end{pmatrix}$$

Clearly the convergence of the Gauss-Seidel iteration to $(1,2,3)^T$ is superior.

5. The Gauss-Seidel method for $A\boldsymbol{x} = \boldsymbol{b}$ is

$$(D+L)\boldsymbol{x}^{(k+1)} = \boldsymbol{b} - U\boldsymbol{x}^{(k)} \iff B = -(D+L)^{-1}U \text{ and } \boldsymbol{c} = (D+L)^{-1}\boldsymbol{b}.$$

and $\qquad D\boldsymbol{x}^{(k+1)} = \boldsymbol{b} + D\boldsymbol{x}^{(k)} - (D+U)\boldsymbol{x}^{(k)} - L\boldsymbol{x}^{(k+1)} \iff \boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + D^{-1}(\boldsymbol{b} - (D+U)\boldsymbol{x}^{(k)} - L\boldsymbol{x}^{(k+1)})$

The successive relaxation formula is

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \omega D^{-1}[\boldsymbol{b} - (D+U)\boldsymbol{x}^{(k)} - L\boldsymbol{x}^{(k+1)}] \iff (I+\omega D^{-1}L)\boldsymbol{x}^{(k+1)} = (I-\omega D^{-1}(D+U))\boldsymbol{x}^{(k)} + \omega D^{-1}\boldsymbol{b}$$

so that $M_\omega = (I+\omega D^{-1}L)^{-1}(I-\omega D^{-1}(D+U))$ and $\boldsymbol{d} = \omega(I+\omega D^{-1}L)^{-1}D^{-1}\boldsymbol{b}$. Let

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Longrightarrow (I+\omega D^{-1}L)^{-1} = \begin{pmatrix} 1 & 0 \\ -\omega c/d & 1 \end{pmatrix}, \ (I-\omega D^{-1}(D+U)) = \begin{pmatrix} 1-\omega & -\omega b/a \\ 0 & 1-\omega \end{pmatrix}$$

$$\Longrightarrow M_\omega = \begin{pmatrix} 1-\omega & -\omega b/a \\ -(1-\omega)\omega c/d & 1-\omega+\omega^2 bc/(ad) \end{pmatrix}$$

Thus

$$0 = \det(M_\omega - \lambda I) = (1-\omega-\lambda)^2 - \frac{\omega^2 bc}{ad}\lambda, \ 0 = \det(B-\mu I) = \mu^2 - \frac{bc}{ad}\mu \Longrightarrow \mu = 0, \frac{bc}{ad}.$$

Thus $(\lambda - 1 + \omega)^2 = \lambda\omega^2\mu$ where λ is an eigenvalue of M_ω and μ is the largest in modulus eigenvalues of B.
The definition for the asymoptotic rate of convergence comes from the fact that

$$\boldsymbol{e}^{(k)} = \boldsymbol{x} - \boldsymbol{x}^{(k)} = M^k \boldsymbol{e}^{(0)}$$

and M^k will converge to 0 at approximately a rate of $\rho(M)$. Defining $\rho(B) = |\mu| = 1 - \varepsilon$ and taking $\omega = 1.5$

$$0 = \lambda^2 - \underbrace{(2(1-\omega)+\omega^2\mu)}_{=1.25-2.25\varepsilon}\lambda + \underbrace{(1-\omega)^2}_{=0.25}$$

$$\lambda = \frac{1.25 - 2.25\varepsilon \pm \sqrt{(1.25-2.25\varepsilon)^2 - 1}}{2} = \frac{1.25 - 2.25\varepsilon \pm 0.75\sqrt{1 - 10\varepsilon + 9\varepsilon^2}}{2}$$
$$= \frac{1.25 - 2.25\varepsilon \pm 0.75(1 - 5\varepsilon + O(\varepsilon^2))}{2}$$

so that $\rho(M) = 1 - 3\varepsilon + O(\varepsilon^2)$. Thus the asymptotic rate of the SOR formula is $-\log(1 - 3\varepsilon + O(\varepsilon^2)) \approx 3\varepsilon$, three times better than the Gauss-Seidel iteration.

6. The iteration matrices for the Jacobi and Gauss-Seidel methods are

$$M_J = \frac{1}{4}\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} \text{ and } M_{GS} = \frac{1}{4}\begin{pmatrix} 0 & -1 & -1 & 0 \\ 0 & -\frac{1}{4} & -\frac{1}{4} & -1 \\ 0 & -\frac{1}{4} & -\frac{1}{4} & -1 \\ 0 & -\frac{1}{8} & -\frac{1}{8} & -\frac{1}{2} \end{pmatrix}.$$

The characteristic equations are $\lambda^2(\lambda^2 - 1/4)$ and $\lambda^3(\lambda + 1/4)$ respectively. Thus

$$-\log\rho(M_J) = -\log 0.5 \approx 0.6931 \text{ and } -\log\rho(M_{GS}) = -\log 0.25 \approx 1.386$$

7. The system

$$\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

is to be solved by an iterative method, starting with $x_1^{(0)} = 0 = x_2^{(0)}$. The Jacobi and Gauss-Seidel iterations are respectively

$$\begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \end{pmatrix} = \begin{pmatrix} 1 + x_2^{(k)} \\ 3 - x_1^{(k)} \end{pmatrix} \text{ and } \begin{pmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \end{pmatrix} = \begin{pmatrix} 1 + x_2^{(k)} \\ 3 - x_1^{(k+1)} \end{pmatrix}$$

so that

$$\begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \end{pmatrix} : \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \ldots \qquad \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \end{pmatrix} : \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \ldots$$

Also note that the eigenvalues of $M_J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and $M_{GS} = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}$ are $\pm i$ and $0, -1$ respectively, so that $\rho(M_{GS}) = \rho(M_J) = 1$ and neither iteration will converge. The SOR iteration, $0 < \omega < 2$, converges if the eigenvalues of the iteration matrix,

$$M_\omega = (I + \omega D^{-1}L)^{-1}((1-\omega)I - \omega D^{-1}U) = \begin{pmatrix} 1-\omega & \omega \\ \omega^2 - \omega & 1 - \omega - \omega^2 \end{pmatrix},$$

are smaller than one in modulus. Using question 10.1 $\mu = -1$, or computing directly, we need to find $\lambda_1,\ \lambda_2$ which solve

$$0 = \det(M_\omega - \lambda I) = (\lambda - 1 + \omega)^2 + \lambda\omega^2 = \lambda^2 - (2(1-\omega) - \omega^2)\lambda + (1-\omega)^2$$

$$\implies \lambda_1, \lambda_2 = \frac{2(1-\omega) - \omega^2 \pm \sqrt{(2(1-\omega)-\omega^2)^2 - 4(1-\omega)^2}}{2} = \frac{2(1-\omega) - \omega^2 \pm \omega\sqrt{\omega^2 + 4\omega - 4}}{2}.$$

Thus, computing $|\lambda_i|$ we have two cases to consider when $\omega^2 + 4\omega - 4 < 0$ and $\geqslant 0$. Notice $\omega^2 + 4\omega - 4 = 0$ iff $\omega = -2 \pm 2\sqrt{2}$. For $\omega \in (0, -2 + 2\sqrt{2})$ the roots are complex and

$$|\lambda_1|^2 = |\lambda_2|^2 = \frac{\overbrace{(2(1-\omega) - \omega^2)^2}^{\omega^4 - 4\omega^2(1-\omega) + 4(1-\omega)^2} + \omega^2(4 - 4\omega - \omega^2)}{4} = (1-\omega)^2 \implies \rho(M_\omega) = |1-\omega|.$$

For $\omega \in [-2 + 2\sqrt{2}, 2)$ we want

$$-1 < \frac{2(1-\omega) - \omega^2 \pm \omega\sqrt{\omega^2 + 4\omega - 4}}{2} < 1 \iff -2 + \omega + \frac{\omega^2}{2} < \pm\frac{\omega\sqrt{\omega^2 + 4\omega - 4}}{2} < \omega + \frac{\omega^2}{2}$$

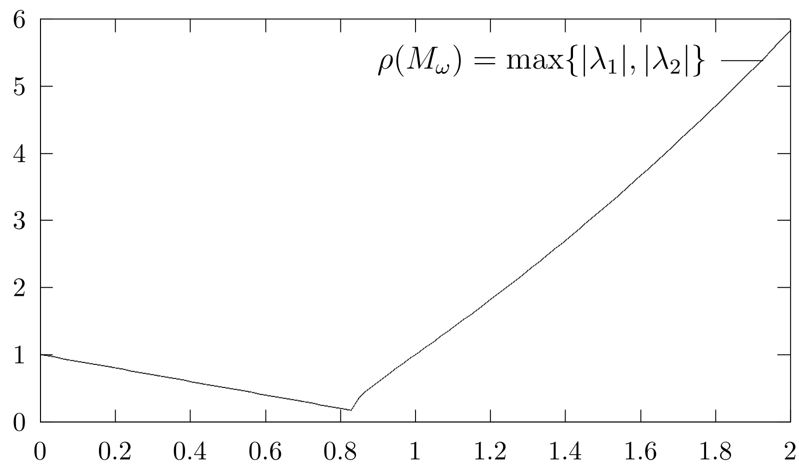
That is

$$\frac{\omega\sqrt{\omega^2 + 4\omega - 4}}{2} < \omega + \frac{\omega^2}{2} \implies \omega^2 + 4\omega - 4 < (2+\omega)^2 = \omega^2 + 4\omega + 4 \iff -4 < 4!$$

And

$$-2+\omega+\frac{\omega^2}{2} < -\frac{\omega\sqrt{\omega^2 + 4\omega - 4}}{2} \implies \omega^2(\omega^2+4\omega-4) < (4-2\omega-\omega^2)^2 = \omega^4+4\omega^3-4\omega^2-16\omega+16 \iff 0 < 16(1-\omega).$$

Thus any $\omega \in (0, 1)$ will do. Plotting the graph of $\rho(M_\omega)$



Notice the discontinuity where the discriminant changes sign. A reasonable value to take for ω is 0.5 in which case $\rho(M_\omega) = 0.5$. The best value to take for ω is $-2 + 2\sqrt{2}$.

8. Let λ and $\boldsymbol{e}$ be an eigenvalue/vector of $M_J = -D^{-1}(L + U)$. Thus premultiplication by $\boldsymbol{e}^H D$ and rearranging yields

$$\lambda \boldsymbol{e} = -D^{-1}(L+U)\boldsymbol{e} \iff \lambda \boldsymbol{e}^H D\boldsymbol{e} = -\boldsymbol{e}^H(L+U)\boldsymbol{e} = \boldsymbol{e}^H D\boldsymbol{e} - \boldsymbol{e}^H A\boldsymbol{e}$$

$$\iff \lambda = 1 - \frac{2\boldsymbol{e}^H A\boldsymbol{e}}{2\boldsymbol{e}^H D\boldsymbol{e}} = 1 - \frac{2\boldsymbol{e}^H A\boldsymbol{e}}{\boldsymbol{e}^H(2D-A)\boldsymbol{e} + \boldsymbol{e}^H A\boldsymbol{e}}.$$

Since A and $2D - A$ are symmetric positive definite matrices, see lecture notes,

$$0 < \boldsymbol{e}^H A\boldsymbol{e} < \boldsymbol{e}^H(2D-A)\boldsymbol{e} + \boldsymbol{e}^H A\boldsymbol{e}$$

so that $-1 < \lambda < 1$ and therefore $\rho(M_J) < 1$. Since the eigenvalues of the Jacobi iteration matrix are smaller than one, the Jacobi iteration will converge.

9. The modified Jacobi iteration for the linear system $A\boldsymbol{x} = \boldsymbol{b}$ is given by

$$D\boldsymbol{x}^{(k+1)} = \omega\boldsymbol{b} + (1-\omega)D\boldsymbol{x}^{(k)} - \omega(L+U)\boldsymbol{x}^{(k)} = \omega\boldsymbol{b} + (D - \omega\underbrace{(D+L+U)}_{=A})\boldsymbol{x}^{(k)}$$

$$\iff \boldsymbol{x}^{(k+1)} = \omega D^{-1}\boldsymbol{b} + (I - \omega D^{-1}A)\boldsymbol{x}^{(k)},$$

i.e. $M_{MJ} = I - \omega D^{-1}A$. If the iteration $\{\boldsymbol{x}^{(k)}\}$ converges to $\boldsymbol{x}$ it satisfies

$$\boldsymbol{x} = \omega D^{-1}\boldsymbol{b} + (I - \omega D^{-1}A)\boldsymbol{x} \iff \omega D^{-1}A\boldsymbol{x} = \omega D^{-1}\boldsymbol{b} \iff A\boldsymbol{x} = \boldsymbol{b}.$$

Let $\{\lambda_i\}_{i=1}^n$ be the set of eigenvalues of the Jacobi iteration matrix, i.e.

$$-D^{-1}(L+U)\boldsymbol{e}_i = \lambda_i\boldsymbol{e}_i \text{ thus } \mu_i\boldsymbol{e}_i = M_{MJ}\boldsymbol{e}_i = \boldsymbol{e}_i - \omega D^{-1}(D+L+U)\boldsymbol{e}_i = (1 - \omega + \omega\lambda_i)\boldsymbol{e}_i,$$

i.e. $\mu_i = 1 - \omega(1 - \lambda_i)$. If all the eigenvalues λ_i are real then

$$1 - \omega(1 - \underline{\lambda}) \leqslant \mu_i \leqslant 1 - \omega(1 - \bar{\lambda})$$

so that the greatest magnitude of μ_i's may be minimised by taking ω so that the upper and lower bound have the same value in magnitude

$$-(1-\omega(1-\underline{\lambda})) = 1-\omega(1-\bar{\lambda}) \Longleftrightarrow \omega = \frac{2}{2-(\bar{\lambda}+\underline{\lambda})}$$

For

$$A = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix}, M_J = \frac{1}{2}\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix}.$$

The eigenvalues of M_J are $-1, \frac{1}{2}, \frac{1}{2}$ so that $\rho(M_J) = 1$ and the Jacobi process does not converge. Since $\bar{\lambda} = \frac{1}{2}$ and $\underline{\lambda} = -1$ taking $\omega = 2/(2-\frac{1}{2}+1) = 0.8$ $\rho(M_{MJ}) = 1-\omega(1-\bar{\lambda}) = 0.6$ so that the modified Jacobi iterates will converge.

10. Since u is analytic, it agrees with its Taylor series expansion about (jh, nk), hence

$$u_j^{n+1} = u(jh,(n+1)k) = u + ku_t + \frac{k^2}{2!}u_{tt} + \frac{k^3}{3!}u_{ttt} + \frac{k^4}{4!}u_{tttt} + \cdots$$

so that rearranging

$$\frac{u_j^{n+1}-u_j^n}{k} = \frac{u(jh,(n+1)k)-u(jh,nk)}{k} = u_t + \frac{k}{2!}u_{tt} + \frac{k^2}{3!}u_{ttt} + \frac{k^3}{4!}u_{tttt} + \cdots$$

Consider the Taylor series expansion about (jh, nk) of

$$u_{j\pm1}^n = u \pm hu_x + \frac{h^2}{2!}u_{xx} \pm \frac{h^3}{3!}u_{xxx} + \frac{h^4}{4!}u_{xxxx} +$$

Hence subtracting the "+"terms disappear

$$u_{j+1}^n - u_{j-1}^n = 2hu_x + 2\frac{h^3}{3!}u_{xxx} + 2\frac{h^5}{5!}u_{xxxxx} +$$

Hence

$$\frac{u_{j+1}^n - u_{j-1}^n}{2h} = u_x + \frac{h^2}{3!}u_{xxx} + \frac{h^4}{5!}u_{xxxxx} + \cdots.$$

11. Calculating the truncation error and noting that $u_t = u_{xx}$, $k = \frac{h^2}{6}$, $u_{tt} = (u_{xx})_t = (u_t)_{xx} = (u_{xx})_{xx} = u_{xxxx}$, $u_{ttt} = u_{xxxxxx}$ and

$$\begin{aligned} T_j^n &= \frac{u_j^{n+1}-u_j^n}{k} - \frac{1}{h^2}\left[u_{j+1}^n - 2u_j^n + u_{j-1}^n\right] \\ &= \left(u_t + \frac{k}{2!}u_{tt} + \frac{k^2}{3!}u_{ttt} + \cdots\right) - \left(u_{xx} + \frac{h^2}{12}u_{xxxx} + \frac{h^4}{360}u_{xxxxxx} + \cdots\right) \\ &= (u_t - u_{xx}) + (\frac{h^2}{6\times 2!}u_{tt} - \frac{h^2}{12}u_{xxxx}) + \frac{h^4}{36\times 3!}u_{xxxxxx} - \frac{h^4}{360}u_{xxxxxx} + \cdots \\ &= O(h^4) \end{aligned}$$

12. (a) The truncation error is

$$
\begin{aligned}
T_j^n &= \frac{u_j^{n+1}-u_j^n}{k} - \frac{\delta^2 u_j^n}{h^2} - a\left[\frac{u_{j+1}^n - u_{j-1}^n}{2h}\right] \\
&= u_t + \frac{k}{2!}u_{tt} + \cdots - (u_{xx} + \frac{h^2}{12}u_{xxxx} + \cdots) \\
&\quad -a(u_x + \frac{h^2}{3!}u_{xxx} + \cdots). \\
&= (u_t - u_{xx} - au_x) + \frac{k}{2!}u_{tt} - (\frac{h^2}{12}u_{xxxx} + \cdots) \\
&\quad -a(\frac{h^2}{3!}u_{xxx} + \cdots). \\
&= O(k) + O(h^2)
\end{aligned}
$$

and hence $T_j^n \to 0$ as $h, k \to 0$, so the scheme is consistent.

(b) Assuming that u is analytic

$$
\begin{aligned}
u_{j+1}^{n+1} - u_j^{n+1} &= u + hu_x + ku_t + \frac{h^2}{2!}u_{xx} + hku_{xt} + \frac{k^2}{2!}u_{tt} \\
&\quad + \frac{h^3}{3!}u_{xxx} + 3\frac{h^2k}{3!}u_{xxt} + 3\frac{hk^2}{3!}u_{xtt} + \frac{k^3}{3!}u_{ttt} + \cdots \\
&\quad - \left(u + ku_t + \frac{k^2}{2!}u_{tt} + \frac{k^3}{3!}u_{ttt} + \cdots\right) \\
&= hu_x + \frac{h^2}{2!}u_{xx} + hku_{xt} + h\left[\frac{h^2}{3!}u_{xxx} + 3\frac{hk}{3!}u_{xxt} + 3\frac{k^2}{3!}u_{xtt}\right] + \cdots
\end{aligned}
$$

and also

$$
u_{j-1}^n = u - hu_x + \frac{h^2}{2!}u_{xx} - \frac{h^3}{3!}u_{xxx} + \cdots
$$

hence on noting that $u_{tt} = (-au_x)_t = -au_{xt}$, the truncation error is given by

$$
\begin{aligned}
T_j^n &= \frac{u_j^{n+1}-u_j^n}{k} + \frac{a}{2}\left[\frac{u_{j+1}^{n+1} - u_j^{n+1}}{h} + \frac{u_j^n - u_{j-1}^n}{h}\right] \\
&= u_t + \frac{k}{2!}u_{tt} + \frac{k^2}{3!}u_{ttt} + \cdots + \frac{a}{2}\left[u_x + \frac{h}{2!}u_{xx} + ku_{xt}\right. \\
&\quad \left. + \frac{h^2}{3!}u_{xxx} + 3\frac{hk}{3!}u_{xxt} + 3\frac{k^2}{3!}u_{xtt} + \cdots + u_x - \frac{h}{2!}u_{xx} + \frac{h^2}{3!}u_{xxx} + \cdots\right] \\
&= (u_t + au_x) + \frac{k}{2!}(u_{tt} + au_{xt}) + \frac{k^2}{3!}u_{ttt} \\
&\quad + \frac{a}{2}\left[\frac{h^2}{3!}u_{xxx} + 3\frac{hk}{3!}u_{xxt} + 3\frac{k^2}{3!}u_{xtt}\right] + \cdots \\
&= O(k^2) + O(h^2)
\end{aligned}
$$

and hence the truncation error converges to 0 as $h, k \to 0$.

13. (a) Noting that

$$
\begin{aligned}
& u_{j+1}^n - 2((1-\theta)u_j^{n-1} + \theta u_j^{n+1}) + u_{j-1}^n \\
= \;& u + hu_x + \frac{h^2}{2!}u_{xx} + \frac{h^3}{3!}u_{xxx} + \frac{h^4}{4!}u_{xxxx} + \frac{h^5}{5!}u_{xxxxx} + O(h^6) \\
& -2\left((1-\theta)(u - ku_t + \frac{k^2}{2!}u_{tt} - \frac{k^3}{3!}u_{ttt} + \frac{k^4}{4!}u_{tttt} + O(k^5))\right. \\
& \left. +\theta\left(u + ku_t + \frac{k^2}{2!}u_{tt} + \frac{k^3}{3!}u_{ttt} + \frac{k^4}{4!}u_{tttt} + O(k^5)\right)\right) \\
& +u - hu_x + \frac{h^2}{2!}u_{xx} - \frac{h^3}{3!}u_{xxx} + \frac{h^4}{4!}u_{xxxx} - \frac{h^5}{5!}u_{xxxxx} + O(h^6) \\
= \;& 2\left[\frac{h^2}{2!}u_{xx} + \frac{h^4}{4!}u_{xxxx} + O(h^6)\right] \\
& -2\left(k(2\theta-1)u_t + \frac{k^2}{2!}u_{tt} + \frac{k^3}{3!}(2\theta-1)u_{ttt} + O(k^4)\right)
\end{aligned}
$$

On noting that $2\mu = 1$ and $u_t - u_{xx} = 0$, the truncation error, T_j^n, for the first discretization is

$$
\begin{aligned}
T_j^n &= \frac{u_j^{n+1} - u_j^{n-1}}{2k} - \frac{u_{j+1}^n - 2((1-\theta)u_j^{n-1} + \theta u_j^{n+1}) + u_{j-1}^n}{h^2} \\
&= \frac{k^2}{3!}u_{ttt} + O(k^4) - \left[\left(\frac{2h^2}{4!}u_{xxxx} + O(h^4)\right)\right. \\
&\quad \left. - \left((2\theta-1)u_t + \frac{k}{2!}u_{tt} + \frac{k^2}{3!}(2\theta-1)u_{ttt} + O(k^3)\right)\right] \\
&= (2\theta-1)u_{xx} + (-\frac{2h^2}{4!} + \frac{k}{2!})u_{xxxx} + \frac{k^2}{3!}2\theta u_{xxxxxx} + O(k^3) + O(h^4)
\end{aligned}
$$

Hence for $\theta = 1/2$, $T_j^n = O(h^2) + O(k) = O(h^2) \to 0$ as $k,\ h \to 0$?

(b) We start by calculating

$$
\begin{aligned}
& u_{j+1}^n - 2u_j^{n+1} + u_{j-1}^n \\
= \;& u + hu_x + \frac{h^2}{2!}u_{xx} + \frac{h^3}{3!}u_{xxx} + \frac{h^4}{4!}u_{xxxx} + \cdots - 2\left(u + ku_t + \frac{k^2}{2!}u_{tt} + \cdots\right) \\
& u - hu_x + \frac{h^2}{2!}u_{xx} - \frac{h^3}{3!}u_{xxx} + \frac{h^4}{4!}u_{xxxx} + \cdots \\
= \;& \frac{2h^2}{2!}u_{xx} + \frac{2h^4}{4!}u_{xxxx} + \cdots - 2\left(ku_t + \frac{k^2}{2!}u_{tt} + \cdots\right)
\end{aligned}
$$

hence the truncation error is, on noting that $u_t = u_{xx}$,

$$\begin{aligned}
T_j^n &= \frac{u_j^{n+1} - u_j^n}{k} - \frac{u_{j+1}^n - 2u_j^{n+1} + u_{j-1}^n}{h^2} \\
&= u_t + ku_{tt} + \frac{k^2}{2!}u_{ttt} + \cdots \\
&\quad - \left(u_{xx} + \frac{2h^2}{4!}u_{xxxx} + \cdots - 2\left(\mu u_t + \frac{k}{2!}\mu u_{tt} + \cdots \right) \right) \\
&= \left(k - \frac{2h^2}{4!} \right) u_{xxxx} + \frac{k^2}{2!}u_{xxxxxx} + 2\mu u_{xx} + k\mu u_{xxxx} + \cdots
\end{aligned}$$

which converges to zero when $\mu \to 0$.

14. (a) Consider the j'th row where $j = 2, \cdots, m-2$ with the ansatz suggested:

$$\begin{aligned}
&a_j x_{j-1}^k + d_j x_j^k + c_j x_{j+1}^k \\
&= a\sin(\frac{k\pi(j-1)}{m+1}) + d\sin(\frac{k\pi j}{m+1}) + a\sin(\frac{k\pi(j+1)}{m+1}) \\
&= \left[2a\cos(\frac{k\pi}{m+1}) + d \right] \sin(\frac{k\pi j}{m+1}) = \lambda_k x_j^k
\end{aligned}$$

When $j = 1$ (we introduce $x_0^k = 0 = \sin(\frac{k\pi 0}{m+1})$)

$$d_1 x_1^k + c_1 x_2^k = a_1 x_0^k + d_1 x_1^k + c_1 x_2^k = \left[2a\cos(\frac{k\pi}{m+1}) + d \right] \sin(\frac{k\pi}{m+1}) = \lambda_k x_1^k.$$

Similarly, with $j = m$ (we introduce $x_{m+1} = 0 = \sin(\frac{k\pi(m+1)}{m+1})$)

$$a_m x_{m-1}^k + d_m x_m^k + c_m x_{m+1}^k = \left[2a\cos(\frac{k\pi}{m+1}) + d \right] \sin(\frac{k\pi m}{m+1}) = \lambda_k x_m^k.$$

Hence, $\boldsymbol{x}^k$ is an eigenvector of A with eigenvector given by λ_k.
Since the eigenvalues are distinct, the eigenvectors form a basis for $\mathbb{R}^m$.

(b) This was done last term, but is included here for completeness

15. To prove that $\|\boldsymbol{V}\|_\infty = \sup_{j\in\mathbb{Z}} |V_j|$ defines a norm on S we need to check the key properties. Obviously $\|\boldsymbol{V}\|_\infty$ is non-negative and

$$\|\boldsymbol{V}\|_\infty = 0 \Longleftrightarrow |V_j| = 0 \ \forall \ j \in \mathbb{Z} \Longleftrightarrow V_j = 0 \ \forall j \in \mathbb{Z}.$$

Secondly

$$\|\lambda \boldsymbol{V}\|_\infty = \sup_{j\in\mathbb{Z}} |\lambda||V_j| = |\lambda| \sup_{j\in\mathbb{Z}} |V_j| = |\lambda| \|\boldsymbol{V}\|_\infty$$

Finally, for all $j \in \mathbb{Z}$

$$|U_j + V_j| \leqslant |U_j| + |V_j| \leqslant \sup_{j\in\mathbb{Z}} |U_j| + \sup_{j\in\mathbb{Z}} |V_j| = \|U\|_\infty + \|V\|_\infty$$

and hence taking the sup over all $j \in \mathbb{Z}$ yields the triangle inequality.
For the proposed norm,

$$\sum_{j\in\mathbb{Z}} h|V_j|^2 < \infty$$

the first two properties follow easily. The triangle inequality is slightly more difficult. We start by proving the triangle inequality for a finite sum. Define

$$(\boldsymbol{V}, \boldsymbol{W}) = \sum_{|j|\leqslant n} hV_jW_j.$$

this is clearly an inner-product. Let $\boldsymbol{W} \neq \boldsymbol{0}$ (if it is zero, the Cauchy-Schwarz inequality is trivial) and choose n sufficiently large so that $\sum_{|j|\leqslant n} hW_j^2 \neq 0$. Consider

$$0\leqslant(\boldsymbol{V} + \lambda\boldsymbol{W}, \boldsymbol{V} + \lambda\boldsymbol{W}) = \sum_{|j|\leqslant n} h(V_j + \lambda W_j)^2 = \sum_{|j|\leqslant n} hV_j^2 + 2\lambda \sum_{|j|\leqslant n} hV_jW_j + \lambda^2 \sum_{|j|\leqslant n} hW_j^2.$$

This is smallest when

$$0 = \frac{\mathrm{d}}{\mathrm{d}\lambda}(\boldsymbol{V} + \lambda\boldsymbol{W}, \boldsymbol{V} + \lambda\boldsymbol{W}) = 2\sum_{|j|\leqslant n} hV_jW_j + 2\lambda\sum_{|j|\leqslant n} hW_j^2 \Longrightarrow \lambda = -\frac{\sum_{|j|\leqslant n} hV_jW_j}{\sum_{|j|\leqslant n} hW_j^2}.$$

Hence, taking λ to be that given above,

$$0\leqslant \sum_{|j|\leqslant n} hV_j^2 - \frac{\left(\sum_{|j|\leqslant n} hV_jW_j\right)^2}{\sum_{|j|\leqslant n} hW_j^2} \Longrightarrow \left(\sum_{|j|\leqslant n} hV_jW_j\right)^2 \leqslant \sum_{|j|\leqslant n} hV_j^2 \sum_{|j|\leqslant n} hW_j^2.$$

Now starting with a finite sum and using the Cauchy-Schwarz inequality

$$\begin{aligned}
\sum_{|j|\leqslant n} h|U_j + V_j|^2 &= \sum_{|j|\leqslant n} hU_j^2 + 2\sum_{|j|\leqslant n} hU_jV_j + \sum_{|j|\leqslant n} hV_j^2 \\
&\leqslant \left[\left(\sum_{|j|\leqslant n} hU_j^2\right)^{1/2} + \left(\sum_{|j|\leqslant n} hV_j^2\right)^{1/2}\right]^2 \leqslant \left[\left(\sum_{j\in\mathbb{Z}} hU_j^2\right)^{1/2} + \left(\sum_{j\in\mathbb{Z}} hV_j^2\right)^{1/2}\right]^2
\end{aligned}$$

Hence letting $n \to \infty$ we get the result on taking a square-root.

16. Set $u(x,t) = \mathrm{e}^{(-\pi^2+1)t}\sin \pi x$, noting that

$$u_t = (-\pi^2 + 1)u, \quad u_{xx} = -\pi^2 u, \Longrightarrow u_t - u_{xx} - u = 0$$

The other two properties follow trivially.
Assume[1] the solution to the finite difference scheme has the form

$$U_j^n = g^n \sin(mj\pi h) \quad m, j = 1, \cdots, J-1$$

where $h = 1/J$, note boundary conditions are satisfied. Then since

$$\begin{aligned} \delta^2 U_j^n &= U_{j+1}^n - 2U_j^n + U_{j-1}^n \\ &= g^n(\sin((j+1)\pi mh) - 2\sin(j\pi mh) + \sin((j-1)\pi mh)) \\ &= 2g^n \left[\cos(\pi mh) - 1\right] \sin(j\pi mh) \end{aligned}$$

it follows that for $j = 1, \cdots, J-1$

$$g^{n+1}\sin(j\pi mh) = (1 - 4\mu \sin^2(\frac{\pi mh}{2}) + k) g^n \sin(j\pi mh)$$

and so

$$g^{n+1} = (1 - 4\mu \sin^2 \frac{\pi mh}{2} + k) g^n \Longrightarrow g^n = (1 - 4\mu \sin^2 \frac{\pi mh}{2} + k)^n g^0.$$

Since $g = 1 - 4\mu \sin^2 \frac{\pi mh}{2} + k$, i.e. it is dependent on k. For instability[2], we need that $|g^n| \to \infty$ and $k \to \infty$ such that nk is constant for some m. First we note that for all $m = 1, \cdots, J-1$

$$\begin{aligned} 1 - 4\mu \sin^2 \frac{\pi mh}{2} + k \leqslant 1 \quad &\Longleftrightarrow \quad k \leqslant 4\mu \sin^2 \frac{\pi mh}{2} \Longleftrightarrow k \leqslant \frac{4k}{h^2} \sin^2 \frac{\pi mh}{2} \\ &\Longleftrightarrow \quad 1 \leqslant \frac{4}{h^2} \sin^2 \frac{\pi mh}{2} \Longleftrightarrow 1 \leqslant \frac{4 \sin^2 \frac{\pi mh}{2}}{h^2}. \end{aligned}$$

Noting that

$$\lim_{h \to 0} \frac{4 \sin^2 \frac{\pi mh}{2}}{h^2} = m^2 \pi^2 > 1.$$

We conclude that the above inequality to be true for h sufficiently small, that is $1 - 4\mu \sin^2 \frac{\pi mh}{2} + k \leqslant 1$, and no instability.
As for the other inequality suppose $\mu > \mu_\star = (2+k)/4$. Let $\varepsilon > 0$ satisfy $\mu = \frac{2+k+\varepsilon}{4}$, then taking m to be the nearest integer to $J/2$ it follows that

$$-(1 - 4\mu \sin^2 \frac{\pi mh}{2} + k) - 1 = (2 + k + \varepsilon) \sin^2 \frac{\pi mh}{2} - 2 - k \gtrsim \varepsilon$$

Hence $|g^n| \gtrsim |1 + \varepsilon|^n \to \infty$.

[1]It isn't difficult to show that $\sin(mj\pi h)$ are eigenvectors for the computational matrix *and* that they form an orthogonal basis

[2]Notice that from the main Theorem, stability with respect to the $\|\cdot\|_2$ is equivalent to

$$|g| \leqslant 1 + Ck, \quad \forall\, \xi \in \left[-\frac{\pi}{h}, \frac{\pi}{h}\right].$$

To prove that a scheme is not stable with respect to the $\|\cdot\|_2$ this is equivalent to proving that

$$|g| > 1 + Ck, \quad \text{for some } \xi \in \left[-\frac{\pi}{h}, \frac{\pi}{h}\right]$$

as $k \to 0$ $\forall$ fixed constants C. Note that if g is independent of k then it is sufficient to show that $|g| > 1$.

17. The θ-method for solving $u_t = u_{xx}$ subject to initial condition $u(x,0) = u^0(x)$ is

$$\frac{1}{k}(U_j^{n+1} - U_j^n) = \frac{1}{h^2}\left[\theta\delta^2 U_j^{n+1} + (1-\theta)\delta^2 U_j^n\right], \quad U_j^0 = u^0(jh).$$

Noting that

$$\begin{aligned}
\delta^2 u_j^{n+1} &= u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1} \\
&= u + hu_x + ku_t + \frac{h^2}{2}u_{xx} + hku_{xt} + \frac{k^2}{2}u_{tt} \\
&\quad + \frac{1}{3!}(h^3 u_{xxx} + 3h^2 k u_{xxt} + 3hk^2 u_{xtt} + k^3 u_{ttt}) + \cdots \\
&\quad - 2(u + ku_t + \frac{k^2}{2}u_{tt} + \frac{k^3}{3!}u_{ttt} + \cdots) \\
&\quad + u - hu_x + ku_t + \frac{h^2}{2}u_{xx} - hku_{xt} + \frac{k^2}{2}u_{tt} \\
&\quad + \frac{1}{3!}(-h^3 u_{xxx} + 3h^2 k u_{xxt} - 3hk^2 u_{xtt} + k^3 u_{ttt}) + \cdots \\
&= h^2 u_{xx} + \frac{2}{3!}(3h^2 k u_{xxt} + k^3 u_{ttt}) + \frac{2}{4!}(h^4 u_{xxxx} + 6h^2k^2 u_{xxtt} + k^4 u_{tttt}) \\
&\quad + \frac{2}{5!}(5h^4 k u_{xxxxt} + 10h^2 k^3 u_{xxttt} + k^5 u_{ttttt}) + \cdots .
\end{aligned}$$

Hence on noting that $u_t = u_{xx}$, the truncation error is given by

$$\begin{aligned}
T_j^n &= \frac{1}{k}(u_j^{n+1} - u_j^n) - \frac{1}{h^2}\left[\theta\delta^2 u_j^{n+1} + (1-\theta)\delta^2 u_j^n\right] \\
&= u_t + \frac{k}{2!}u_{tt} + \frac{k^2}{3!}u_{ttt} + \cdots \\
&\quad - \theta\left[u_{xx} + \frac{2}{3!}(3ku_{xxt} + k^2\mu u_{ttt}) + \frac{2}{4!}(h^4 u_{xxxx} + 6h^2k^2 u_{xxtt} + k^4 u_{tttt})\right. \\
&\quad \left. + \frac{2}{5!}(5h^4 k u_{xxxxt}\cdots)\right] - (1-\theta)\left[u_{xx} + \frac{h^2}{12}u_{xxxx} + \cdots\right] \\
&= (\frac{k}{2!} - \theta k - \frac{h^2}{12}(1-\theta) - \theta\frac{2}{4!}h^4)u_{xxxx} \\
&\quad + (\frac{k^2}{3!} - k^2\mu\theta\frac{2}{3!} + \frac{2}{4!}6h^2k^2 + \frac{2}{5!}5h^4 k)u_{xxxxxx} + \cdots \\
&= O(k) + O(h^2)
\end{aligned}$$

When $\theta = \frac{1}{2}$, it is clear that $T_j^n = O(k^2) + O(h^2)$

18. Suppose that $U_j^n = g^n \mathrm{e}^{ijh}$ then substituting this into the propose scheme yields

$$\begin{aligned}
(g^{n+1} - g^n)\mathrm{e}^{ijh} &= \mu(g^n \mathrm{e}^{i(j+1)h} - 2g^{n+1}\mathrm{e}^{ijh} + g^n \mathrm{e}^{i(j-1)h}) \\
&= 2\mu(g^n \cos(h) - g^{n+1})\mathrm{e}^{ijh}
\end{aligned}$$

hence we obtain

$$(1+2\mu)g^{n+1} = (1+2\mu\cos(h))g^n \Longrightarrow g^n = \left(\frac{2\mu\cos(h)+1}{1+2\mu}\right)$$

We have stability

$$\Longleftrightarrow -(1+2\mu)\leqslant 2\mu\cos h + 1 \leqslant 1+2\mu \Longleftrightarrow \mu(-1-\cos h)\leqslant 1 \text{ and } \cos h\leqslant 1$$

both of which hold.
From problem **??**b we know that $|T_j^n| \to 0$ we require that $\mu \to 0$ as $h, k \to 0$. Hence under such a condition from Lemma 2.1 we have convergence.
Suppose that $h = \frac{1}{j}$, then we should choose $k = h^{2+\varepsilon}$ for some $\varepsilon > 0$ as $J \to \infty$ to ensure convergence and the rate of convergence will be $O(h^{\varepsilon})$.

19. Suppose that $U_j^n = g^n e^{ij\xi}$. First note that

$$\delta^2 U_j^n = -4\sin^2(\tfrac{\xi}{2})g^n e^{ij\xi}$$

and

$$U_{j+1}^n - U_{j-1}^n = g^n(e^{i(j+1)\xi} - e^{i(j-1)\xi}) = 2i\sin\xi g^n e^{ij\xi}$$

then substituting the ansatz into the finite difference scheme

$$\frac{U_j^{n+1} - U_j^n}{k} = \frac{\delta^2 U_j^n}{h^2} + a\left[\frac{U_{j+1}^n - U_{j-1}^n}{2h}\right]$$

yields

$$(g^{n+1} - g^n)e^{ij\xi} = -4\mu\sin^2(\tfrac{\xi}{2})g^n e^{ij\xi} + a\lambda i\sin\xi g^n e^{ij\xi}$$

hence

$$g^{n+1} = (1-4\mu\sin^2(\tfrac{\xi}{2}) + a\lambda i\sin\xi)g^n \Longrightarrow g^n = (1-4\mu\sin^2\tfrac{\xi}{2} + 2a\lambda i\sin(\tfrac{\xi}{2})\cos(\tfrac{\xi}{2}))^n g^0.$$

Noting the independence of g on k, to ensure stability we require that $|g|\leqslant 1$. Noting that $\mu\leqslant\frac{1}{2}$ and $a^2\lambda^2\leqslant 2\mu$, it follows that

$$\begin{aligned} |g|^2 &= |1-4\mu\sin^2\tfrac{\xi}{2} + 2a\lambda i\sin(\tfrac{\xi}{2})\cos(\tfrac{\xi}{2})|^2 = (1-4\mu\sin^2(\tfrac{\xi}{2}))^2 + 4a^2\lambda^2\sin^2\tfrac{\xi}{2}\cos^2\tfrac{\xi}{2} \\ &= 1 + 4\mu\sin^2(\tfrac{\xi}{2})(-2 + 4\mu\sin^2(\tfrac{\xi}{2}) + \tfrac{a^2\lambda^2}{\mu}\cos^2(\tfrac{\xi}{2})) \\ &\leqslant 1 + 4\mu\sin^2(\tfrac{\xi}{2})(-2 + 2\sin^2(\tfrac{\xi}{2}) + 2\cos^2(\tfrac{\xi}{2})) = 1. \end{aligned}$$

Note that $a^2\lambda^2\leqslant 2\mu\leqslant 1 \Longrightarrow |a|\lambda\leqslant 1$. [3]

[3] If we were not given the conditions how would we derive a condition? Obviously, we need

$$4\mu\sin^2(\tfrac{\xi}{2}) + \tfrac{a^2\lambda^2}{\mu}\cos^2(\tfrac{\xi}{2}) = (4\mu - \tfrac{a^2\lambda^2}{\mu})\sin^2(\tfrac{\xi}{2}) + \tfrac{a^2\lambda^2}{\mu}\leqslant 2$$

to hold for all ξ. Hence, we require that

$$\begin{cases} 4\mu\leqslant 2 & \text{if } \mu\geqslant |a|\lambda/2, \\ \frac{a^2\lambda^2}{\mu}\leqslant 2 & \text{if } \mu\leqslant |a|\lambda/2. \end{cases}$$

The question set of type Section A, the extra bit I have just done is of type Section B.

20. Define $U_0 = U_J = 0$, then the j'th row of the equation is

$$\begin{aligned} U_j^{n+1} &= U_j^n + \mu[\theta(U_{j+1}^{n+1} - 2U_j^{n+1} + U_{j-1}^{n+1}) + (1-\theta)(U_{j+1}^n - 2U_j^n + U_{j-1}^n)] + kU_j^n. \\ &\iff -\mu\theta U_{j-1}^{n+1} + (1+2\mu\theta)U_j^{n+1} - \mu\theta U_{j+1}^{n+1} \\ &\qquad = (1-\theta)\mu U_{j-1}^n + (1-2(1-\theta)\mu + k)U_j^n + (1-\theta)\mu U_{j+1}^n \end{aligned}$$

and hence

$$M_1 = \begin{pmatrix} 1+2\mu\theta & -\mu\theta & 0 & \cdots & 0 \\ -\mu\theta & 1+2\mu\theta & -\mu\theta & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -\mu\theta & 1+2\mu\theta & -\mu\theta \\ 0 & \cdots & 0 & -\mu\theta & 1+2\mu\theta \end{pmatrix}$$

and

$$M_2 = \begin{pmatrix} 1-2(1-\theta)\mu+k & (1-\theta)\mu & 0 & \cdots & 0 \\ (1-\theta)\mu & 1-2(1-\theta)\mu+k & (1-\theta)\mu & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & (1-\theta)\mu & 1-2(1-\theta)\mu+k & (1-\theta)\mu \\ 0 & \cdots & 0 & (1-\theta)\mu & 1-2(1-\theta)\mu+k \end{pmatrix}$$

21. Noting that $u_t = -(au)_x$

$$u_{tt} = (-(au)_x)_t = (-(au)_t)_x = -(au_t)_x = (a(au_x))_x$$

and the Taylor series about (jh, nk) is

$$u_j^{n+1} = u(jh, (n+1)k) = u + ku_t + \frac{k^2}{2!}u_{tt} + O(k^3)$$

The Lax-Wendroff scheme is

$$U_j^{n+1} = U_j^n + k \times \frac{1}{2h}\left[-a_{j+1}U_{j+1}^n + a_{j-1}U_{j-1}^n\right] + \frac{k^2}{2!} \times \frac{1}{h^2}\left[\delta[a_j\delta(a_jU_j^n)]\right]$$

where

$$\begin{aligned} \delta[a_j\delta(a_jU_j^n)] &= \delta[a_j(a_{j+1/2}U_{j+1/2}^n - a_{j-1/2}U_{j-1/2}^n)] \\ &= a_{j+1/2}a_{j+1}U_{j+1}^n - (a_{j+1/2} + a_{j-1/2})a_jU_j^n + a_{j-1/2}a_{j-1}U_{j-1}^n. \end{aligned}$$

Noting that

$$\begin{aligned} a_{j\pm1}u_{j+1}^n &= au \pm h(a_xu + au_x) + \frac{h^2}{2!}(a_{xx}u + 2a_xu_x + au_{xx}) + \cdots \\ &\implies \frac{a_{j-1}u_{j-1}^n - a_{j+1}u_{j+1}^n}{2h} = -(au)_x + O(h^2) \end{aligned}$$

and

$$\begin{aligned}\delta(a_j u_j^n) &= h(au)_x + O(h^3)\\ &\Longrightarrow \left[\delta[a_j\delta(a_j u_j^n)]\right] = \delta[ha(au)_x + O(h^3)] = h^2(a(au)_x)_x + O(h^4)\end{aligned}$$

Hence

$$\begin{aligned}T_j^n &= \frac{1}{k}\left[u_j^{n+1} - u^n - k\times\frac{1}{2h}\left[-a_{j+1}u_{j+1}^n + a_{j-1}u_{j-1}^n\right] - \frac{k^2}{2!}\times\frac{1}{h^2}\left[\delta[a_j\delta(a_j u_j^n)]\right]\right]\\ &= \frac{1}{k}\left[ku_t + \frac{k^2}{2!}u_{tt} + O(k^2) - k(-(au)_x + O(h^2)) - \frac{k^2}{2!}((a(au)_x)_x + O(h^2))\right]\\ &= O(k^2) + O(h^2).\end{aligned}$$

22. The scheme of which I talk is

$$U_j^{n+1} - U_j^n + \frac{a\lambda}{2}(U_{j+1}^n - U_{j-1}^n) - \frac{a^2\lambda^2}{2}\delta^2 U_j^n = 0 \qquad U_j^0 = 0\ j\geqslant 0 \qquad U_j^0 = 1\ j<0.$$

Hence

$$U_j^1 = \begin{cases} 1 & \text{if } j\leqslant -2\\ 1 - \frac{a\lambda}{2}(1+a\lambda) & \text{if } j=-1\\ (a\lambda-1)\frac{a\lambda}{2} & \text{if } j=0\\ 0 & \text{if } j\geqslant 1\end{cases}$$

If the artificial diffusion were not present, then

$$U_j^1 = \begin{cases} 1 & \text{if } j\leqslant -2\\ 1 - \frac{a\lambda}{2} & \text{if } j=-1\\ -\frac{a\lambda}{2} & \text{if } j=0\\ 0 & \text{if } j\geqslant 1\end{cases}$$

and hence the artificial diffusion solution is immediately smoother with no spikes.

23. Noting that

$$T_j^n = \frac{u_j^n - u_j^{n-2}}{2k} + \frac{1}{2h}\left(a_{j+1}u_{j+1}^{n-1} - a_{j-1}u_{j-1}^{n-1}\right) = u_t + O(k) + (au)_x + O(h^2) + \lambda O(k) = O(k) + O(h^2)$$

we deduce consistency. Let the CFL condition holds $|a|\lambda\leqslant 1$. Since a constant and assuming the ansatz $U_j^n = g^n e^{ij\xi}$ it follows in the usual way that

$$g^2 = 1 - 2ia\lambda\sin\xi g \Longleftrightarrow g^2 + 2ia\lambda\sin\xi g - 1 = 0 \Longleftrightarrow g = ia\lambda\sin\xi \pm\sqrt{1 - a^2\lambda^2\sin^2\xi}.$$

Thus

$$|g|^2 = 1 - a^2\lambda^2\sin^2\xi + a^2\lambda^2\sin^2\xi = 1$$

and the scheme is stable.

24. We begin by integration the problem over $(x_i, x_{i+1}) \times (y_j, y_{j+1})$ then

$$
\begin{aligned}
0 &= \int_{x_i}^{x_{i+1}} \int_{y_j}^{y_{j+1}} \left[\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right] \mathrm{d}y\mathrm{d}x \\
&= \int_{y_j}^{y_{j+1}} \left[\frac{\partial u}{\partial x}(x_{i+1}, y) - \frac{\partial u}{\partial x}(x_i, y)\right] \mathrm{d}y + \int_{x_i}^{x_{i+1}} \left[\frac{\partial u}{\partial y}(x, y_{j+1}) - \frac{\partial u}{\partial y}(x, y_j)\right] \mathrm{d}x \\
&\approx \frac{h}{h}\Bigg[\Big(u(x_{i+1} + \frac{1}{2}h, y_j + \frac{1}{2}h) - u(x_{i+1} - \frac{1}{2}h, y_j + \frac{1}{2}h)\Big) \\
&\quad - \Big(u(x_i + \frac{1}{2}h, y_j + \frac{1}{2}h) - u(x_i - \frac{1}{2}h, y_j + \frac{1}{2}h)\Big) \\
&\quad + \Big(u(x_i + \frac{1}{2}h, y_{j+1} + \frac{1}{2}h) - u(x_i + \frac{1}{2}h, y_{j+1} - \frac{1}{2}h)\Big) \\
&\quad - \Big(u(x_i + \frac{1}{2}h, y_j + \frac{1}{2}h) - u(x_i + \frac{1}{2}h, y_j - \frac{1}{2}h)\Big)\Bigg]
\end{aligned}
$$

which leads to the five-point difference operator:

$$
0 = -4U^{i+1/2,j+1/2} + U^{i+3/2,j+1/2} + U^{i-1/2,j+1/2} + U^{i+1/2,j+3/2} + U^{i+1/2,j-1/2}
$$

25. Since both of the approximations for u_t and u_x are second-order in time and space, respectively, I expect the method to be second order in both space and time.

$$
\begin{pmatrix}
1 & \frac{\lambda}{4} & 0 & \cdots & 0 \\
-\frac{\lambda}{4} & 1 & \frac{\lambda}{4} & \cdots & \vdots \\
0 & \ddots & \ddots & \ddots & 0 \\
\vdots & \ddots & -\frac{\lambda}{4} & 1 & \frac{\lambda}{4} \\
0 & \cdots & 0 & -\frac{\lambda}{4} & 1
\end{pmatrix}
\boldsymbol{U}^{n+1} =
\begin{pmatrix}
1 & \frac{\lambda}{4} & 0 & \cdots & 0 \\
-\frac{\lambda}{4} & 1 & \frac{\lambda}{4} & \cdots & \vdots \\
0 & \ddots & \ddots & \ddots & 0 \\
\vdots & \ddots & -\frac{\lambda}{4} & 1 & \frac{\lambda}{4} \\
0 & \cdots & 0 & -\frac{\lambda}{4} & 1
\end{pmatrix}
\boldsymbol{U}^{n} +
\begin{pmatrix}
\frac{\lambda}{4}(f((n+1)k) - f(nk)) \\
0 \\
\vdots \\
0 \\
\frac{\lambda}{4}(g(nk) - g((n+1)k))
\end{pmatrix}
$$

26. Let $U_{i,j} \approx u(ih, jh)$. Define

$$
U_{i,j} = g(ih, jh) \quad \text{if either } i = 0, n \text{ or } j = 0, n
$$

this deals with the boundary conditions. Otherwise, we approximate the equation at interior nodes using the usual approximation for second derivatives

$$
-\frac{1}{h^2}\left(U_{i+1,j} - 2U_{i,j} + U_{i-1,j}\right) - \frac{1}{h^2}\left(U_{i,j+1} - 2U_{i,j} + U_{i,j-1}\right) = f_{i,j}
$$

where $f_{i,j} = f(ih, jh)$. Which we can rewrite as

$$
-U_{i,j-1} - U_{i+1,j} + 4U_{i,j} - U_{i-1,j} - U_{i,j+1} = h^2 f_{i,j}
$$

Hence we arrive at the system of equations

$$\begin{pmatrix} D & -I & 0 & \cdots & 0 \\ -I & D & -I & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -I & D & -I \\ 0 & \cdots & 0 & -I & D \end{pmatrix} \boldsymbol{U} = h^2 \boldsymbol{f}$$

where I is the $(J-1)\times(J-1)$ identity matrix, 0 is the $(J-1)\times(J-1)$ zero matrix and D is the $(J-1)\times(J-1)$ matrix

$$D = \begin{pmatrix} 4 & -1 & 0 & \cdots & 0 \\ -1 & 4 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 4 & -1 \\ 0 & \cdots & 0 & -1 & 4 \end{pmatrix}.$$

27. The three points lie in a plane and order $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$ in anti-clockwise order. Define $\boldsymbol{x}_1$ as the origin and the vectors $\boldsymbol{x}_2 - \boldsymbol{x}_1$ and $\boldsymbol{x}_3 - \boldsymbol{x}_1$ to lie in the x-y plane. Hence $(\boldsymbol{x}_j - \boldsymbol{x}_i) \wedge (\boldsymbol{x}_k - \boldsymbol{x}_i)$ gives a vector in the positive z direction. Note that

$$(0,0,1)^T \cdot (\boldsymbol{x}_j - \boldsymbol{x}_i) \wedge (\boldsymbol{x}_k - \boldsymbol{x}_i) = \begin{vmatrix} 0 & 0 & 1 \\ x_{2,1} - x_{1,1} & x_{2,2} - x_{1,2} & 0 \\ x_{3,1} - x_{1,1} & x_{3,2} - x_{1,2} & 0 \end{vmatrix}$$

gives the volume of the parallepiped with edges given by the vectors $\boldsymbol{x}_2 - \boldsymbol{x}_1$, $\boldsymbol{x}_2 - \boldsymbol{x}_1$ and $(0,0,1)^T$ which is also the area of the parallelogram base. Now, the area of the parallelogram is twice that defined by the triangle with corners $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ and $\boldsymbol{x}_3$ hence is

$$\begin{aligned} 2|\tau| &= \begin{vmatrix} 0 & 0 & 1 \\ x_{2,1} - x_{1,1} & x_{2,2} - x_{1,2} & 0 \\ x_{3,1} - x_{1,1} & x_{3,2} - x_{1,2} & 0 \end{vmatrix} \\ &\xrightarrow[\text{transpose } 3\times 3 \text{ matrix}]{\text{transpose } 2\times 2 \text{ matrix}} \begin{vmatrix} 0 & x_{2,1} - x_{1,1} & x_{2,2} - x_{1,2} \\ 0 & x_{3,1} - x_{1,1} & x_{3,2} - x_{1,2} \\ 1 & 0 & 0 \end{vmatrix} \\ &\xrightarrow[CA_{13}(x_{1,2})]{CA_{12}(x_{1,1})} \begin{vmatrix} 0 & x_{2,1} - x_{1,1} & x_{2,2} - x_{1,2} \\ 0 & x_{3,1} - x_{1,1} & x_{3,2} - x_{1,2} \\ 1 & x_{1,1} & x_{1,2} \end{vmatrix} \xrightarrow[RA_{31}(1)]{RA_{32}(1)} \begin{vmatrix} 1 & x_{2,1} & x_{2,2} \\ 1 & x_{3,1} & x_{3,2} \\ 1 & x_{1,1} & x_{1,2} \end{vmatrix} \\ &\xrightarrow[P_{23}]{P_{13}} \begin{vmatrix} 1 & x_{1,1} & x_{1,2} \\ 1 & x_{2,1} & x_{2,2} \\ 1 & x_{3,1} & x_{3,2} \end{vmatrix} \xrightarrow[P_{23}]{P_{13}} \begin{vmatrix} 1 & x_{3,1} & x_{3,2} \\ 1 & x_{1,1} & x_{1,2} \\ 1 & x_{2,1} & x_{2,2} \end{vmatrix} \end{aligned}$$

28.

29.

30.

31.

32.

33.

34. The element stiffness matrix is

$$A^\tau = \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

35.

36. After applying the conditions at $(0,0)$, $(0,1)$, $(1,0)$, $(0,\frac{1}{2})$, $(\frac{1}{2},0)$ and $(\frac{1}{2},\frac{1}{2})$ we are left with a matrix to invert. There a unique solution if and only if the determinant of the matrix is non-zero.

$$\begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{4} \\ 1 & \frac{1}{2} & 0 & 0 & \frac{1}{4} & 0 \\ 1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{vmatrix} = \begin{vmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{4} \\ \frac{1}{2} & 0 & 0 & \frac{1}{4} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{vmatrix} \underset{A_{25}(-\frac{1}{2})}{\overset{A_{24}(-\frac{1}{2})}{\longrightarrow}} \begin{vmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{4} \\ 0 & 0 & 0 & -\frac{1}{4} & 0 \\ 0 & \frac{1}{2} & \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \end{vmatrix}$$

$$= -\begin{vmatrix} 1 & 0 & 0 & 1 \\ \frac{1}{2} & 0 & 0 & \frac{1}{4} \\ 0 & 0 & -\frac{1}{4} & 0 \\ \frac{1}{2} & \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \end{vmatrix} = \frac{1}{4} \begin{vmatrix} 1 & 0 & 1 \\ \frac{1}{2} & 0 & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{vmatrix} \underset{A_{13}(-\frac{1}{2})}{\overset{A_{12}(-\frac{1}{2})}{\longrightarrow}} \frac{1}{4} \begin{vmatrix} 1 & 0 & 1 \\ 0 & 0 & -\frac{1}{4} \\ 0 & \frac{1}{4} & -\frac{1}{4} \end{vmatrix} = \frac{1}{64}$$

37.

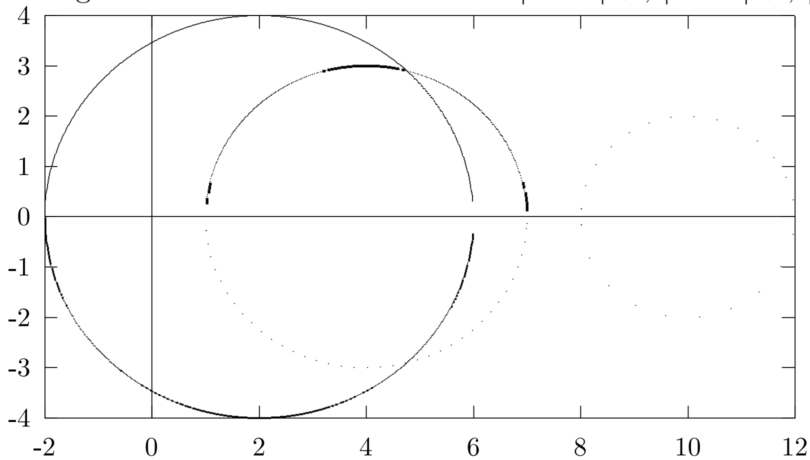
38.

39.

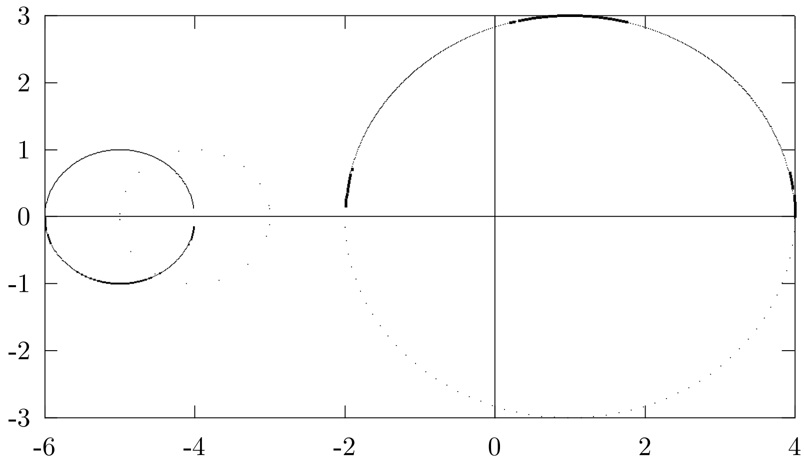
40.

41.

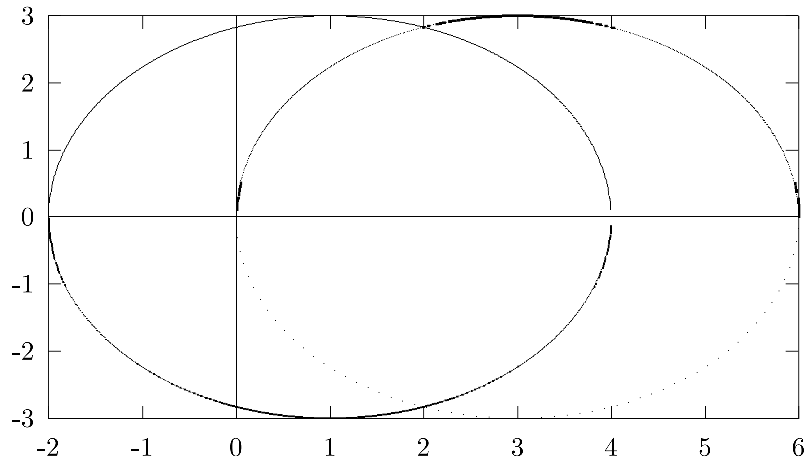
42. A. The Gerschgorin discs for the matrix A are $|z-2|\leqslant 4$, $|z-4|\leqslant 3$, $|z-10|\leqslant 2$, $z\in\mathbb{C}$.



B. The Gerschgorin discs for the matrix B are $|z+5|\leqslant 1$, $|z-1|\leqslant 3$, $|z+4|\leqslant 1$, $z\in\mathbb{C}$.



C. The Gerschgorin discs for the matrix C are $|z-3|\leqslant 3$, $|z-1|\leqslant 3$ both twice, $z\in\mathbb{C}$.



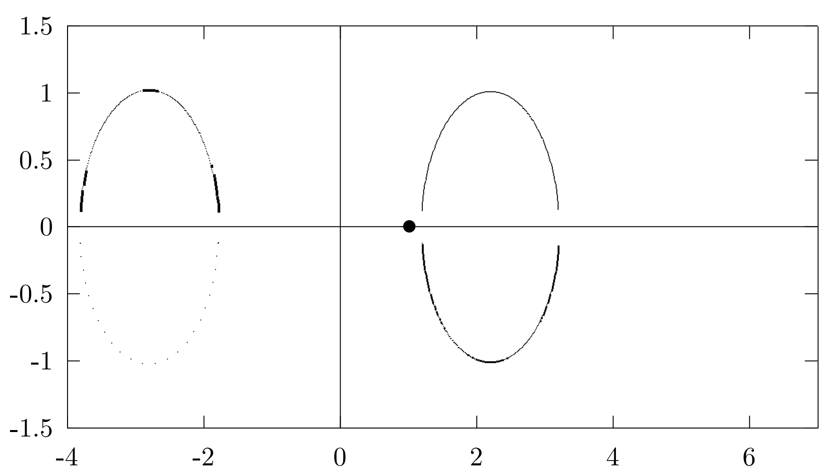
43. The Gerschgorin discs are $|z-0.9|\leqslant 0.03$, $|z-2.2|\leqslant 0.02$ and $|z+2.8|\leqslant 0.03$ where $z\in\mathbb{C}$. These discs do not intersect. Introducing the similarity transofrmation

$$P=\begin{pmatrix} k & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} \text{ we find } B=P^{-1}AP=\begin{pmatrix} 0.9 & 0.01k^{-1} & 0.02k^{-1}\\ -0.01k & 2.2 & 0.01\\ 0.01k & 0.02 & -2.8\end{pmatrix}$$

so that the Gerschgorin discs are $|z-0.9|\leqslant 0.03k^{-1}$, $|z-2.2|\leqslant 0.01(1+k)$ and $|z+2.8|\leqslant 0.01(2+k)$ where $z\in\mathbb{C}$. The discs do not intersect as long as

$$0.9+0.03k^{-1}<2.2-0.01(1+k) \text{ and } -2.8+0.01(2+k)<0.9-0.03k^{-1}$$

In the picture below we have set $k=100$



The first inequality is true for $k \leqslant 128.97\cdots$ and the second one is true when $k \leqslant 367.99\cdots$. Thus taking $k = 128$ we get the improved bound of $|\lambda - 0.9| < 2.35 \times 10^{-4}$.
Introducing a similarity transformation to dialate the $|z - 2.2|$ disc we find

$$0.9 + 0.01(2 + k) < 2.2 - 0.02k^{-1} \text{ and } -2.8 + 0.01(2k + 1) < 2.2 - 0.02k^{-1}$$

which are both true when $k \leqslant 127$, thus taking $k = 127$ we get the improved bound of $|\lambda - 2.2| < 1.58 \times 10^{-4}$.
And finally in an anologous manner, taking $k = 368$ we get the improved bound of $|\lambda + 2.8| < 8.16 \times 10^{-5}$.

44. The Gerschgorin discs are

$$|\lambda_1 - (0.9 + 10^{-6})| \leqslant 6 \times 10^{-6},\ |\lambda_2 - (0.4 + 5 \times 10^{-6})| \leqslant 2 \times 10^{-6},\ |\lambda_3 - (0.2 + 3 \times 10^{-6})| \leqslant 3 \times 10^{-6}$$

which do not intersect, so we can use a similarity transformation, as suggested,

$$D_1^{-1} B D_1 = \begin{pmatrix} 0.9 & 0 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0 & 0.2 \end{pmatrix} + 10^{-5} \begin{pmatrix} 0.1 & 4 \times 10^{-6} & -2 \times 10^{-6} \\ -10^4 & 0.5 & 0.1 \\ 2 \times 10^4 & 0.1 & 0.3 \end{pmatrix}$$

so that the Gerschgorin discs are

$$|\lambda_1 - (0.9 + 10^{-6})| \leqslant 6 \times 10^{-11},\ |\lambda_2 - (0.4 + 5 \times 10^{-6})| \leqslant 0.1 + 10^{-6},\ |\lambda_3 - (0.2 + 3 \times 10^{-6})| \leqslant 0.2 + 10^{-6}.$$

The disc centred on $0.9 + 10^{-6}$ is still disconnected from the others, so we obtain an improved bound. The remainder of the question works through in exactly the same fashion with the improved bounds being

$$|\lambda_2 - (0.4 + 5 \times 10^{-6})| \leqslant 2 \times 10^{-11} \text{ and } |\lambda_3 - (0.2 + 3 \times 10^{-6})| \leqslant 3 \times 10^{-11}.$$

45. Using MATLAB

k	1	2	3
$\boldsymbol{x}^{(k)}$	$(-0.5789, 1, -0.5789)^T$	$(-0.5775, 1, -0.5775)^T$	$-0.5774, 1, -0.5774)^T$
S^k	4	3.75	3.73

it appears S^k is converging to the dominant eigenvalue which is probably 3.7 to 3 d.p. In fact the largest eigenvalue is 3.732 to 3 d.p.

46. Take $\boldsymbol{x}^{(0)} = (1, 0, 0)^T$ and $\boldsymbol{v} = (1, 1, 1)^T$. Again using MATLAB

k	1	2	3	4	5
$\boldsymbol{x}^{(k)}$	$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 0.5 \\ 0.8333 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 0.4516 \\ 0.8065 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 0.4459 \\ 0.8025 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 0.4451 \\ 0.8020 \\ 1 \end{pmatrix}$
S^k	3	4.667	5	5.043	5.048

Using all of the digits available in Aitken's acceleration, we found the limit of S_k to be 5.049 to 3 d.p.

47. Using the power method with $\boldsymbol{x}^{(0)} = (1,1,1,0)^T = \boldsymbol{v}$ to compute the largest eigenvalue we found

k	1	2	3	4	5	6
$\boldsymbol{x}^{(k)}$	$\begin{pmatrix}1\\0.7857\\0.7857\\0.6429\end{pmatrix}$	$\begin{pmatrix}1\\0.9337\\0.9337\\0.8619\end{pmatrix}$	$\begin{pmatrix}1\\0.9757\\0.9757\\0.9518\end{pmatrix}$	$\begin{pmatrix}1\\0.9918\\0.9918\\0.9837\end{pmatrix}$	$\begin{pmatrix}1\\0.9973\\0.9973\\0.9945\end{pmatrix}$	$\begin{pmatrix}1\\0.9991\\0.9991\\0.9982\end{pmatrix}$
S_k	12	14.417	14.751	14.919	14.93	14.991

It is easy to see spot that the eigenvector is $(1,1,1,1)^T$ and $\lambda_1 = 15$. Now considering

$$\begin{pmatrix} -9 & 4 & 4 & 1 \\ 4 & -9 & 1 & 4 \\ 4 & 1 & -9 & 4 \\ 1 & 4 & 4 & -9 \end{pmatrix}$$

we use the power method again, with the same choice for $\boldsymbol{x}^{(0)}$ and $\boldsymbol{v}$.

k	1	2	3	4	5	6	7
$\boldsymbol{x}^{(k)}$	$\begin{pmatrix}-0.1111\\-0.4444\\-0.4444\\1\end{pmatrix}$	$\begin{pmatrix}-0.1228\\0.5614\\0.5614\\-1\end{pmatrix}$	$\begin{pmatrix}0.34383\\-0.6719\\-0.6719\\1\end{pmatrix}$	$\begin{pmatrix}-0.5324\\0.7662\\0.762\\-1\end{pmatrix}$	$\begin{pmatrix}0.67964\\-0.8398\\-0.840\\1\end{pmatrix}$	$\begin{pmatrix}-0.7870\\0.8935\\0.893\\-1\end{pmatrix}$	$\begin{pmatrix}0.86132\\-0.9307\\-0.931\\1\end{pmatrix}$
S_k	-3	-12.667	-13.368	-14.031	-14.597	-15.039	-15.360

using Aitken's acceleration we find that an estimate is -16, so that the smallest eigenvalue of the original matrix is $\lambda_4 = -1$.

48. Let A be a 3×3 matrix with eigenvalues $\{\lambda_i\}$. We assume that A has eigenvectors $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, $\boldsymbol{u}_3$ which form a basis for $\mathbb{R}^3$. Note that $\lambda \neq \lambda_1$, say, as it is only an approximation. Thus

$$(1)\quad \boldsymbol{y}^{(0)} = \sum_{i=1}^{3} \alpha_i \boldsymbol{u}_i, \quad (2)\quad \boldsymbol{z}_1 = (A-\lambda I)^{-1}\boldsymbol{y}_0 = \sum_{i=1}^{3} \alpha_i (A-\lambda I)^{-1}\boldsymbol{u}_i = \sum_{i=1}^{3} \alpha_i \frac{1}{\lambda_i - \lambda}\boldsymbol{u}_i$$

$$(3)\quad \boldsymbol{y}_1 = \boldsymbol{z}_1/\|\boldsymbol{z}_1\|_\infty, \quad (4)\quad \boldsymbol{z}_2 = \frac{\sum_{i=1}^{3} \alpha_i/(\lambda_i-\lambda)^2\boldsymbol{u}_i}{\|\sum_{i=1}^{3} \alpha_i/(\lambda_i-\lambda)\boldsymbol{u}_i\|_\infty} \quad (5)\quad \boldsymbol{y}_2 = \frac{\sum_{i=1}^{3} \alpha_i/(\lambda_i-\lambda)^2\boldsymbol{u}_i}{\|\sum_{i=1}^{3} \alpha_i/(\lambda_i-\lambda)^2\boldsymbol{u}_i\|}$$

Thus we can prove by induction that

$$\boldsymbol{y}_n = \frac{\sum_{i=1}^{3} \alpha_i \frac{1}{(\lambda_i-\lambda)^n}\boldsymbol{u}_i}{\|\sum_{i=1}^{3} \alpha_i \frac{1}{(\lambda_i-\lambda)^n}\boldsymbol{u}_i\|} \times \frac{|\lambda_1-\lambda|^n}{|\lambda_1-\lambda|^n} \Longrightarrow \lim_{n\to\infty} \boldsymbol{y}_n = \frac{\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|}$$

where we have assumed that $|\lambda_1 - \lambda| < 1$. Using two iterations of this method with $\boldsymbol{y}^{(0)} = (1,1,1)^T$

$\lambda = 8.99$:			$\lambda = 4.01$:			$\lambda = -5.99$:		
k	1	2	k	1	2	k	1	2
$\boldsymbol{z}_k$	$\begin{pmatrix} 33.1552 \\ 33.2889 \\ -33.5560 \end{pmatrix}$	$\begin{pmatrix} 99.3377 \\ 99.3366 \\ -99.3353 \end{pmatrix}$	$\boldsymbol{z}_k$	$\begin{pmatrix} -99.8999 \\ 0.0002 \\ -100.1001 \end{pmatrix}$	$\begin{pmatrix} 99.9001 \\ 0.0002 \\ 99.8999 \end{pmatrix}$	$\boldsymbol{z}_k$	$\begin{pmatrix} 33.4557 \\ -66.6444 \\ -33.2555 \end{pmatrix}$	$\begin{pmatrix} -50.0165 \\ 100.0334 \\ 50.0168 \end{pmatrix}$
$\boldsymbol{y}_k$	$\begin{pmatrix} 0.9881 \\ 0.9920 \\ -1.0000 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$	$\boldsymbol{y}_k$	$\begin{pmatrix} -0.9980 \\ 0.0000 \\ -1.0000 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$	$\boldsymbol{y}_k$	$\begin{pmatrix} 0.5020 \\ -1.0000 \\ -0.4990 \end{pmatrix}$	$\begin{pmatrix} -0.5 \\ 1 \\ 0 \end{pmatrix}$

49. The final eigenvector estimate in Qu. **??** was $(-0.5774, 1, -0.5774)^T$ thus the Rayleigh quotient is

$$(-0.5774, 1, -0.5774)\begin{pmatrix} 2 & -1 & 0 \\ -2 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}\begin{pmatrix} -0.5774 \\ 1 \\ -0.5774 \end{pmatrix} \times \frac{1}{\|(-0.5774, 1, -0.5774)\|_2^2} \approx 3.732$$

to 3 d.p. which is is a very good estimate of the largest eigenvalue obtained in Qu. **??**

50. The Rayleigh quotient of the eigenvector $(-2, 1, k)^T$ is

$$(-2, 1, k)\begin{pmatrix} 1 & 2 & \sqrt{2} \\ 2 & 3 & 0 \\ \sqrt{2} & 0 & 1 \end{pmatrix}\begin{pmatrix} -2 \\ 1 \\ k \end{pmatrix} \times \frac{1}{\|(-2, 1, k)^T\|_2^2} = \frac{k^2 - 4\sqrt{2}k - 1}{k^2 + 5} = \lambda(k)$$

since we are told this is a minimum, it follows that $\lambda'(k) = 0$, i.e.

$$\frac{12\,k + 4\,\sqrt{2}k^2 - 20\,\sqrt{2}}{(k^2 + 5)^2} = 0 \iff k = \sqrt{2} \text{ or } -\frac{5}{\sqrt{2}}$$

Noting that $\lambda(\sqrt{2}) = -1$ and $\lambda(-5/\sqrt{2}) = -17/35$ it follows that the eigenvector is $(-2, 1, \sqrt{2})^T$.

51. Consider the problem

$$\min_{\rho \in \mathbb{R}} \|A\boldsymbol{u} - \rho\boldsymbol{u}\|_2^2 =: \mathcal{F}(\rho).$$

This will be minimized when $\mathcal{F}'(\rho) = 0$. Thus

$$\mathcal{F}(\rho) = \|A\boldsymbol{u} - \rho\boldsymbol{u}\|_2^2 = (\boldsymbol{u}^T \underbrace{A^T}_{=A} - \rho\boldsymbol{u}^T)(A\boldsymbol{u} - \rho\boldsymbol{u}) = \boldsymbol{u}^T A^2\boldsymbol{u} - 2\rho\boldsymbol{u}^T A\boldsymbol{u} + \rho^2\boldsymbol{u}^T\boldsymbol{u}$$

$$\implies \mathcal{F}'(\rho) = 0 \iff \rho = \boldsymbol{u}^T A\boldsymbol{u}/\boldsymbol{u}^T\boldsymbol{u}$$