Part II

Numerical Analysis II

Paper 1, Section II

41C Numerical Analysis

Consider the diffusion equation in 2D on a square domain $(x,y) \in [0,1]^2$

$$\frac{\partial u}{\partial t}(x,y,t) = \nabla^2 u(x,y,t), \tag{1}$$

where $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$ is the Laplacian. We assume zero Dirichlet boundary conditions $u(x,0,t) = u(x,1,t) = u(0,y,t) = u(1,y,t) = 0$ for all $t \geqslant 0$.

We discretize the domain $[0,1]^2$ by a regular grid (ih, jh) where $0 \leqslant i,j \leqslant m+1$ and $h = 1/(m+1)$.

(a) Show that if we discretize the Laplacian operator ∇^2 by the five-point finite-difference scheme, we get an ordinary differential equation of the form

$$\frac{d\mathbf{u}(t)}{dt} = \frac{1}{h^2}(A_x + A_y)\mathbf{u}(t) \qquad \mathbf{u}(t) \in \mathbb{R}^{m^2}, \tag{2}$$

where $\mathbf{u}_{ij} \approx u(ih, jh)$, and A_x and A_y are two matrices of size $m^2 \times m^2$ that correspond respectively to discretizations of $\partial^2/\partial x^2$ and $\partial^2/\partial y^2$. You should verify that your matrices A_x and A_y commute, i.e., $A_x A_y = A_y A_x$.

(b) Consider the following time-stepping scheme for (2), where $k > 0$ is the time step and $\mu = k/h^2$:

$$\begin{cases} \mathbf{u}^{n+1/2} &= \mathbf{u}^n + \mu A_y \mathbf{u}^{n+1/2} \\ \mathbf{u}^{n+1} &= \mathbf{u}^{n+1/2} + \mu A_x \mathbf{u}^{n+1/2}. \end{cases}$$

(i) Explain why $\mathbf{u}^{n+1}$ can be computed from $\mathbf{u}^n$ using at most $\mathcal{O}(m^2)$ arithmetic operations.

(ii) Show that $\mathbf{u}^{n+1} = C\mathbf{u}^n$ for some matrix C that you should make explicit. Deduce conditions on μ for the method to be stable.

[*Hint: For (ii), you can use the fact that A_x and A_y are diagonalizable in the same orthogonal basis of eigenvectors $(\mathbf{v}^{(p,q)})_{1 \leqslant p,q \leqslant m}$ where $\mathbf{v}^{(p,q)} \in \mathbb{R}^{m^2}$, and that $A_x \mathbf{v}^{(p,q)} = \lambda_p \mathbf{v}^{(p,q)}$ and $A_y \mathbf{v}^{(p,q)} = \lambda_q \mathbf{v}^{(p,q)}$ and $\lambda_p = -4\sin^2(p\pi h/2)$.*]

(c) We consider the following modified discretization method to compute $\mathbf{u}^{n+1}$ from $\mathbf{u}^n$:

$$\begin{cases} \widetilde{\mathbf{u}}^{n+1/2} &= \mathbf{u}^n + \mu A_y \widetilde{\mathbf{u}}^{n+1/2} \\ \widetilde{\mathbf{u}}^{n+1} &= \widetilde{\mathbf{u}}^{n+1/2} + \mu A_x \widetilde{\mathbf{u}}^{n+1/2} \\ \mathbf{u}^{n+1} &= \widetilde{\mathbf{u}}^{n+1} + \mu A_x (\mathbf{u}^{n+1} - \mathbf{u}^n). \end{cases}$$

By writing the method as $\mathbf{u}^{n+1} = D\mathbf{u}^n$ for some matrix D, and analyzing the eigenvalues of D, show that this method is stable for any choice of $\mu > 0$.

Paper 2, Section II

41C Numerical Analysis

Consider the variable coefficient advection equation

$$\frac{\partial u}{\partial t}(x,t) + c(x)\frac{\partial u}{\partial x}(x,t) = 0 \tag{1}$$

where $x \in (-\infty, \infty)$ and $t \geqslant 0$. Assume that $c(x) > 0$ is 2-periodic, i.e., $c(x+2) = c(x)$. We will seek a 2-periodic solution $u(x,t)$ that satisfies $u(x,t) = u(x+2,t)$ for all t.

(a) Assume $c(x)$ has a finite decomposition in a Fourier basis

$$c(x) = \sum_{n=-d}^{d} \widehat{c}_n e^{i\pi n x} \qquad (\widehat{c}_i = 0 \text{ for } |i| > d).$$

Give an expression for $\widehat{c}_n$ in terms of $c(x)$. Using the fact that $c(x) > 0$ for all x, show that the $(2d+1) \times (2d+1)$ matrix $[\widehat{c}_{n-m}]_{-d \leqslant n,m \leqslant d}$ is Hermitian positive definite.

(b) We seek a solution $u(x,t)$ of (1) of the form

$$u(x,t) = \sum_{n=-d}^{d} \widehat{u}_n(t) e^{i\pi n x}.$$

Let $\widehat{\mathbf{u}}(t) = (\widehat{u}_n(t))_{|n| \leqslant d} \in \mathbb{C}^{2d+1}$. Applying the spectral method to (1) derive an ODE of the form

$$\frac{d\widehat{\mathbf{u}}(t)}{dt} = i\pi B \widehat{\mathbf{u}}(t) \tag{2}$$

for some matrix B of size $(2d+1) \times (2d+1)$ that you should specify.

(c) Explain why the eigenvalues of B are all real. Deduce that the explicit Euler discretization of (2) is unstable.

[*Hint: you can assume, without proof, that if P and Q are two Hermitian matrices and P is positive definite, then the eigenvalues of PQ are all real.*]

(d) Consider the case $c(x) = 2 + \cos(\pi x) - (1/2)\sin(\pi x)$ and $d = 1$. Form the matrix B and compute its eigenvalues.

Paper 3, Section II

40C Numerical Analysis

Let A be an $n \times n$ real symmetric positive definite matrix and consider the linear system of equations $A\mathbf{x} = \mathbf{b}$, with $\mathbf{b}, \mathbf{x} \in \mathbb{R}^n$. Let $F(\mathbf{x}) = (1/2)\mathbf{x}^T A\mathbf{x} - \mathbf{b}^T\mathbf{x}$.

(a) Define the *steepest descent method with exact line search* to minimize F. Show that for the 2×2 linear system

$$A = \begin{pmatrix} 1 & 0 \\ 0 & \gamma \end{pmatrix}, \quad \mathbf{b} = \mathbf{0} \in \mathbb{R}^2 \quad (\gamma > 1), \tag{1}$$

with the starting point $\mathbf{x}^{(0)} = (\gamma, 1)$, the k-th iterate of this method satisfies

$$\frac{\|\mathbf{x}^{(k)} - \mathbf{x}^*\|_2}{\|\mathbf{x}^{(0)} - \mathbf{x}^*\|_2} = \left(\frac{\kappa - 1}{\kappa + 1}\right)^k \tag{2}$$

where κ is the *condition number* of A that you should define.

Define the *conjugate gradient method.* If the conjugate gradient method is applied to this example, at most how many iterations will be needed to reach $\mathbf{x}^*$?

(b) Return to the case of general $n \times n$ A as specified at the beginning of the question. The *heavy-ball method* to minimize $F(\mathbf{x})$ is defined by the following iterations

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha \nabla F(\mathbf{x}^{(k)}) + \beta(\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}), \tag{3}$$

for some constants $\alpha, \beta > 0$, with the initial point $\mathbf{x}^{(0)} = 0$. Show that $\mathbf{r}^{(k)} \in \mathcal{K}_k(A, \mathbf{b})$ where $\mathbf{r}^{(k)} = \mathbf{b} - A\mathbf{x}^{(k)}$ is the residual at the kth iterate, and $\mathcal{K}_k(A, \mathbf{b})$ is the kth Krylov subspace of A with respect to $\mathbf{b}$.

(c) Let $\mathbf{e}^{(k)} = \mathbf{x}^* - \mathbf{x}^{(k)}$ be the error for the iterates of the heavy-ball method. Show that we can find a matrix M of size $2n \times 2n$ such that

$$\begin{pmatrix} \mathbf{e}^{(k+1)} \\ \mathbf{e}^{(k)} \end{pmatrix} = M \begin{pmatrix} \mathbf{e}^{(k)} \\ \mathbf{e}^{(k-1)} \end{pmatrix}.$$

Your matrix M should be explicit, and depend only on A, α and β. Assuming A is diagonal, show that M can be made block diagonal with 2×2 blocks by an appropriate permutation of its rows and columns (i.e. there is a permutation matrix P such that PMP^T is block diagonal).

(d) Compute the spectral radius of M for the particular A and $\mathbf{b}$ given in (1) and the choice $\alpha = 1/\gamma$ and $\beta = (1 - \sqrt{1/\gamma})^2$. Compare your result with the rate in (2) when $\gamma \gg 1$. [*Hint: To simplify the algebra you may find it helpful to write α in terms of β.*]

Paper 4, Section II

40C Numerical Analysis

(a) Define the *Rayleigh quotient* of a matrix $A \in \mathbb{R}^{n\times n}$ at a vector $\mathbf{x} \in \mathbb{R}^n$. Describe the method of *Rayleigh quotient iteration* to compute an eigenvalue of a matrix.

In the remainder of the question $A \in \mathbb{R}^{n\times n}$ and $\lambda \in \mathbb{R}$ is a simple eigenvalue of A. $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, with $\|\mathbf{u}\|_2 = \|\mathbf{v}\|_2 = 1$, are respectively the left and right eigenvectors of A associated with the eigenvalue λ. We define $s(\lambda) = 1/|\mathbf{u}^T\mathbf{v}|$ to be the *sensitivity* of the eigenvalue λ.

When A is to be regarded as depending on a parameter t the notation $A(t)$ will be used, with corresponding use of $\lambda(t)$, $\mathbf{u}(t)$ and $\mathbf{v}(t)$.

(b) Let $E \in \mathbb{R}^{n\times n}$ be a perturbation matrix and let $\lambda(t)$ be an eigenvalue of $A(t) = A(0) + tE$ with $t \in \mathbb{R}$. Assuming $\lambda(t)$ is differentiable at $t = 0$, show that

$$|\lambda'(0)| \leqslant \frac{\|E\|_2}{|\mathbf{u}(0)^T\mathbf{v}(0)|}, \tag{1}$$

where $\|E\|_2$ is the operator norm of E.

[*Hint: consider* $\mathbf{u}(0)^T A(t)\mathbf{v}(t)$.]

(c) What can you say about the sensitivity $s(\lambda)$ if A is a symmetric matrix? More generally, what can you say if A is a normal matrix?

(d) Let

$$A = \begin{pmatrix} \lambda_1 & 1 & & \\ & \lambda_2 & 1 & \\ & & \ddots & 1 \\ & & & \lambda_n \end{pmatrix},$$

where $\lambda_1 = 1$, and $\lambda_i = 1 - 1/i$ for $i \geqslant 2$. Show that for the eigenvalue $\lambda = \lambda_1 = 1$, the sensitivity $s(\lambda)$ is at least $n!$.

(e) Consider applying Rayleigh quotient iterations to compute the eigenvalue λ of a matrix A. Upon termination of the algorithm, we obtain $\tilde{\mathbf{v}} \in \mathbb{R}^n$, $\|\tilde{\mathbf{v}}\|_2 = 1$ and $\tilde{\lambda} \in \mathbb{R}$ such that

$$\|A\tilde{\mathbf{v}} - \tilde{\lambda}\tilde{\mathbf{v}}\|_2 = \epsilon$$

where ϵ is the machine precision. Show that $|\tilde{\lambda} - \lambda| \lesssim \epsilon s(\lambda)$.

[*Hint: construct a perturbation matrix E such that* $(A + E)\tilde{\mathbf{v}} = \tilde{\lambda}\tilde{\mathbf{v}}$ *and use the approximation* $|\lambda(1) - \lambda(0)| \approx |\lambda'(0)|$.]

Paper 1, Section II
41C Numerical Analysis

(a) Let $H \in \mathbb{R}^{n\times n}$ be diagonalisable. Show that the sequence defined by $\mathbf{z}^{(k+1)} = H\mathbf{z}^{(k)}$ converges to 0 for all initial vectors $\mathbf{z}^{(0)} \in \mathbb{C}^n$ if, and only if, $\rho(H) < 1$ where $\rho(H)$ is the spectral radius of H.

Let $A \in \mathbb{R}^{n\times n}$ be a symmetric positive definite matrix, and let $\mathbf{b} \in \mathbb{R}^n$.

(b) Prove that the solution to $A\mathbf{x} = \mathbf{b}$ is the unique minimiser of the function $f(\mathbf{x}) = (1/2)\mathbf{x}^T A\mathbf{x} - \mathbf{b}^T\mathbf{x}$.

(c) The *steepest descent method* with constant step size α is defined by

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha\nabla f(\mathbf{x}^{(k)}).$$

Applying the method to the function f given in (b), write down the iterations explicitly in terms of A and $\mathbf{b}$. Under what conditions on α does the sequence $\mathbf{x}^{(k)}$ converge to $A^{-1}\mathbf{b}$?

(d) Consider the steepest descent method with *exact line search*, where at each iteration k, the constant $\alpha = \alpha^{(k)}$ is chosen so that $f(\mathbf{x}^{(k+1)})$ is as small as possible. Give an explicit expression for the step size $\alpha^{(k)}$. Show that, in this case, the residuals $\mathbf{r}^{(k)} = \mathbf{b} - A\mathbf{x}^{(k)}$ satisfy $(\mathbf{r}^{(k)})^T\mathbf{r}^{(k+1)} = 0$ for all k.

Paper 2, Section II
41C Numerical Analysis

(a) Consider a linear recurrence relation

$$\sum_{k=r}^{s} a_k u_{m+k}^{n+1} = \sum_{k=r}^{s} b_k u_{m+k}^{n} \qquad n \geqslant 0, \;\; m \in \mathbb{Z}\,,$$

where (a_k) and (b_k) are fixed coefficients.

(i) Show that if we define the Fourier transform of $\boldsymbol{u}^n = (u_m^n)_{m\in\mathbb{Z}}$ by $\widehat{u^n}(\theta) = \sum_{m\in\mathbb{Z}} e^{-im\theta} u_m^n$, then the linear recurrence relation takes the form

$$\widehat{u^{n+1}}(\theta) = H(\theta)\widehat{u^n}(\theta)\,,$$

where $H(\theta)$ is a function that you should specify.

(ii) Show that the sequence $(\boldsymbol{u}^n)_{n\geqslant 0}$ is bounded in the ℓ_2 norm, for all $\boldsymbol{u}^0$, if and only if $|H(\theta)| \leqslant 1$ for all $\theta \in [-\pi, \pi]$.

[*You may assume Parseval's identity:*

$$\|u\|_{\ell_2}^2 = \sum_{m\in\mathbb{Z}} |u_m|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |\widehat{u}(\theta)|^2 \, d\theta. \qquad]$$

(b) Consider the following three recurrence relations:

(i) $u_m^{n+1} = u_m^n + \mu(u_m^n - u_{m-1}^n)$

(ii) $u_m^{n+1} = \frac{1}{2}\mu(1+\mu)u_{m-1}^n + (1-\mu^2)u_m^n - \frac{1}{2}\mu(1-\mu)u_{m+1}^n$

(iii) $u_m^{n+1} - \frac{1}{2}(\mu-\alpha)(u_{m-1}^{n+1} - 2u_m^{n+1} + u_{m+1}^{n+1}) = u_m^n + \frac{1}{2}(\mu+\alpha)(u_{m-1}^n - 2u_m^n + u_{m+1}^n)$

where $n \in \mathbb{N}$ is the time discretization index, $m \in \mathbb{Z}$ is the spatial discretization index, $\mu \geqslant 0$ is the Courant number, and, for (iii), $\alpha \geqslant 0$ is a parameter. In each case give an expression for the amplification factor $H(\theta)$, and deduce the set of values μ (and α for (iii)) for which we have stability.

Paper 3, Section II
40C Numerical Analysis

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix with real eigenvalues $\lambda_1, \ldots, \lambda_n$ ordered by their magnitudes in nonincreasing order, $|\lambda_1| \geqslant |\lambda_2| \geqslant \ldots \geqslant |\lambda_n|$.

(a) Define the *power method* to compute the leading eigenvalue of A. Show that, under suitable assumptions, the iterates $(\mathbf{x}_k)$ of the power method satisfy

$$r(\mathbf{x}_k) - \lambda_1 = O(|\lambda_2/\lambda_1|^{2k})$$

as $k \to \infty$, where $r(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}/\mathbf{x}^T\mathbf{x}$ is the Rayleigh quotient.

(b) Let

$$A = \begin{bmatrix} 5 & 1 & 3 \\ 1 & 7 & 1 \\ 3 & 1 & 5 \end{bmatrix}$$

to which we apply the power method with starting vector $\mathbf{x}_0 = (1/\sqrt{2}, -1/\sqrt{2}, 0)$. Compute $\mathbf{x}_k$ and $r(\mathbf{x}_k)$ explicitly, and find the limit value $\lim_{k\to\infty} r(\mathbf{x}_k)$. Compare with the result in (a) and comment. [*Hint: The eigenvalues of A are 9, 6 and 2.*]

(c) Define the *inverse iteration with shift*, and describe (without proof) the convergence of the method, clearly stating the assumptions.

Paper 4, Section II
40C Numerical Analysis

(a) State and prove the *Gershgorin circle theorem.*

(b) Consider the diffusion equation on the square $[0,1]^2$

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(a(x,y)\frac{\partial}{\partial x}u(x,y)\right) + \frac{\partial}{\partial y}\left(a(x,y)\frac{\partial}{\partial y}u(x,y)\right),$$

where $0 < a(x,y) < a_{\max}$ for all $(x,y) \in [0,1]^2$ is the diffusion coefficient, and with Dirichlet boundary conditions $u(x,y,t) = 0$ for (x,y) on the boundary of $[0,1]^2$.

Consider a uniform grid of size $M \times M$ with step $h = 1/(M+1)$ and let $u_{i,j} = u(ih, jh)$ for $1 \leqslant i \leqslant M$ and $1 \leqslant j \leqslant M$.

(i) Using finite differences, show that the right-hand side of the diffusion equation can be discretised by an expression of the form

$$\frac{1}{h^2}(\alpha u_{i-1,j} + \beta u_{i+1,j} + \gamma u_{i,j-1} + \delta u_{i,j+1} - (\alpha+\beta+\gamma+\delta)u_{i,j})$$

for some $\alpha, \beta, \gamma, \delta$ which you should specify, and which depend on i, j and the diffusion coefficient. Show that the error of this discretisation is $O(h^2)$.

(ii) The time derivative is discretised using a forward Euler scheme with a time step $\Delta t = k$. Use Gershgorin's theorem, clearly justifying all your steps, to show that the resulting scheme is stable when $0 < \mu \leqslant 1/(4a_{\max})$, where $\mu = k/h^2$ is the Courant number.

Paper 1, Section II

41E Numerical Analysis

Let $A \in \mathbb{R}^{n \times n}$ with $n > 2$ and define $\operatorname{Spec}(A) = \{\lambda \in \mathbb{C} \mid A - \lambda I \text{ is not invertible}\}$. The QR algorithm for computing $\operatorname{Spec}(A)$ is defined as follows. Set $A_0 = A$. For $k = 0, 1, \ldots$ compute the QR factorization $A_k = Q_k R_k$ and set $A_{k+1} = R_k Q_k$. (Here Q_k is an $n \times n$ orthogonal matrix and R_k is an $n \times n$ upper triangular matrix.)

(a) Show that A_{k+1} is related to the original matrix A by the similarity transformation $A_{k+1} = \bar{Q}_k^T A \bar{Q}_k$, where $\bar{Q}_k = Q_0 Q_1 \cdots Q_k$ is orthogonal and $\bar{Q}_k \bar{R}_k$ is the QR factorization of A^{k+1} with $\bar{R}_k = R_k R_{k-1} \cdots R_0$.

(b) Suppose that A is symmetric and that its eigenvalues satisfy

$$|\lambda_1| < |\lambda_2| < \cdots < |\lambda_{n-1}| = |\lambda_n| \,.$$

Suppose, in addition, that the first two canonical basis vectors are given by $\mathbf{e}_1 = \sum_{i=1}^n b_i \mathbf{w}_i$, $\mathbf{e}_2 = \sum_{i=1}^n c_i \mathbf{w}_i$, where $b_i \neq 0$, $c_i \neq 0$ for $i = 1, \ldots, n$ and $\{\mathbf{w}_i\}_{i=1}^n$ are the normalised eigenvectors of A.

Let $B_k \in \mathbb{R}^{2 \times 2}$ be the 2×2 upper left corner of A_k. Show that $d_{\mathrm{H}}\big(\operatorname{Spec}(B_k), S\big) \to 0$ as $k \to \infty$, where $S = \{\lambda_n\} \cup \{\lambda_{n-1}\}$ and d_{H} denotes the Hausdorff metric

$$d_{\mathrm{H}}(X, Y) = \max\left\{ \sup_{x \in X} \inf_{y \in Y} |x - y|, \sup_{y \in Y} \inf_{x \in X} |x - y| \right\}, \qquad X, Y \subset \mathbb{C}\,.$$

[*Hint: You may use the fact that for real symmetric matrices U, V we have* $d_{\mathrm{H}}\big(\operatorname{Spec}(U), \operatorname{Spec}(V)\big) \leqslant \|U - V\|_2$.]

Paper 2, Section II

41E Numerical Analysis

(a) Let $\mathbf{x} \in \mathbb{R}^N$ and define $\mathbf{y} \in \mathbb{R}^{2N}$ by

$$y_n = \begin{cases} x_n, & 0 \leqslant n \leqslant N-1 \\ x_{2N-n-1}, & N \leqslant n \leqslant 2N-1. \end{cases}$$

Let $\mathbf{Y} \in \mathbb{C}^{2N}$ be defined as the discrete Fourier transform (DFT) of $\mathbf{y}$, i.e.

$$Y_k = \sum_{n=0}^{2N-1} y_n \omega_{2N}^{nk}, \quad \omega_{2N} = \exp\left(-\pi i/N\right), \quad 0 \leqslant k \leqslant 2N-1.$$

Show that

$$Y_k = 2\omega_{2N}^{-k/2} \sum_{n=0}^{N-1} x_n \cos\left[\frac{\pi}{N}\left(n+\frac{1}{2}\right)k\right], \quad 0 \leqslant k \leqslant 2N-1.$$

(b) Define the discrete cosine transform (DCT) $\mathcal{C}_N : \mathbb{R}^N \to \mathbb{R}^N$ by

$$\mathbf{z} = \mathcal{C}_N \mathbf{x}, \text{ where } z_k = \sum_{n=0}^{N-1} x_n \cos\left[\frac{\pi}{N}\left(n+\frac{1}{2}\right)k\right], \quad k = 0, \ldots, N-1\,.$$

For $N = 2^p$ with $p \in \mathbb{N}$, show that, similar to the Fast Fourier Transform (FFT), there exists an algorithm that computes the DCT of a vector of length N, where the number of multiplications required is bounded by $CN \log N$, where C is some constant independent of N.

[*You may not assume that the FFT algorithm requires $\mathcal{O}(N \log N)$ multiplications to compute the DFT of a vector of length N. If you use this, you must prove it.*]

Paper 3, Section II

40E Numerical Analysis

Consider discretisation of the diffusion equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}\,, \qquad 0 \leqslant t \leqslant 1\,, \tag{$*$}$$

by the Crank–Nicholson method:

$$u_m^{n+1} - \tfrac{1}{2}\mu(u_{m-1}^{n+1} - 2u_m^{n+1} + u_{m+1}^{n+1}) = u_m^n + \tfrac{1}{2}\mu(u_{m-1}^n - 2u_m^n + u_{m+1}^n)\,, \quad n = 0\,, \ldots, N\,, \tag{$\dagger$}$$

where $\mu = \frac{k}{h^2}$ is the Courant number, h is the step size in the space discretisation, $k = \frac{1}{N+1}$ is the step size in the time discretisation, and $u_m^n \approx u(mh, nk)$, where $u(x,t)$ is the solution of $(*)$. The initial condition $u(x,0) = u_0(x)$ is given.

(a) Consider the Cauchy problem for $(*)$ on the whole line, $x \in \mathbb{R}$ (thus $m \in \mathbb{Z}$), and derive the formula for the amplification factor of the Crank–Nicholson method $(\dagger)$. Use the amplification factor to show that the Crank–Nicholson method is stable for the Cauchy problem for all $\mu > 0$.

[You may quote basic properties of the Fourier transform mentioned in lectures, but not the theorem on sufficient and necessary conditions on the amplification factor to have stability.]

(b) Consider $(*)$ on the interval $0 \leqslant x \leqslant 1$ (thus $m = 1, \ldots, M$ and $h = \frac{1}{M+1}$) with Dirichlet boundary conditions $u(0,t) = \phi_0(t)$ and $u(1,t) = \phi_1(t)$, for some sufficiently smooth functions ϕ_0 and ϕ_1. Show directly (without using the Lax equivalence theorem) that, given sufficient smoothness of u, the Crank–Nicholson method is convergent, for any $\mu > 0$, in the norm defined by $\|\boldsymbol{\eta}\|_{2,h} = \left(h \sum_{m=1}^{M} |\eta_m|^2\right)^{1/2}$ for $\boldsymbol{\eta} \in \mathbb{R}^M$.

[You may assume that the Trapezoidal method has local order 3, and that the standard three-point centred discretisation of the second derivative (as used in the Crank–Nicholson method) has local order 2.]

Paper 4, Section II

40E Numerical Analysis

(a) Show that if A and B are real matrices such that both A and $A - B - B^T$ are symmetric positive definite, then the spectral radius of $H = -(A - B)^{-1}B$ is strictly less than 1.

(b) Consider the Poisson equation $\nabla^2 u = f$ (with zero Dirichlet boundary condition) on the unit square, where f is some smooth function. Given $m \in \mathbb{N}$ and an equidistant grid on the unit square with stepsize $h = 1/(m+1)$, the standard five-point method is given by

$$u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1} - 4u_{i,j} = h^2 f_{i,j}\,, \qquad i, j = 1, \ldots, m\,, \qquad (*)$$

where $f_{i,j} = f(ih, jh)$ and $u_{0,j} = u_{m+1,j} = u_{i,0} = u_{i,m+1} = 0$. Equation $(*)$ can be written as a linear system $Ax = b$, where $A \in \mathbb{R}^{m^2 \times m^2}$ and $b \in \mathbb{R}^{m^2}$ both depend on the chosen ordering of the grid points.

Use the result in part (a) to show that the Gauss–Seidel method converges for the linear system $Ax = b$ described above, regardless of the choice of ordering of the grid points.

[*You may quote convergence results – based on the spectral radius of the iteration matrix – mentioned in the lecture notes.*]

Paper 1, Section II

41E Numerical Analysis

Let $A \in \mathbb{R}^{n \times n}$ be a real symmetric matrix with distinct eigenvalues $\lambda_1 < \lambda_2 < \cdots < \lambda_n$ and a corresponding orthonormal basis of real eigenvectors $\{\boldsymbol{w}_i\}_{i=1}^n$. Given a unit norm vector $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$, and a set of parameters $s_k \in \mathbb{R}$, consider the inverse iteration algorithm

$$(A - s_k I)\,\boldsymbol{y} = \boldsymbol{x}^{(k)}, \qquad \boldsymbol{x}^{(k+1)} = \boldsymbol{y}/\|\boldsymbol{y}\|, \qquad k \geqslant 0\,.$$

(a) Let $s_k = s = \text{const}$ for all k. Assuming that $\boldsymbol{x}^{(0)} = \sum_{i=1}^{n} c_i \boldsymbol{w}_i$ with all $c_i \neq 0$, prove that

$$s < \lambda_1 \quad \Rightarrow \quad \boldsymbol{x}^{(k)} \to \boldsymbol{w}_1 \quad \text{or} \quad \boldsymbol{x}^{(k)} \to -\boldsymbol{w}_1 \quad (k \to \infty)\,.$$

Explain briefly what happens to $\boldsymbol{x}^{(k)}$ when $\lambda_m < s < \lambda_{m+1}$ for some $m \in \{1, 2, \ldots, n-1\}$, and when $\lambda_n < s$.

(b) Let $s_k = (A\boldsymbol{x}^{(k)}, \boldsymbol{x}^{(k)})$ for $k \geqslant 0$. Assuming that, for some k, some $a_i \in \mathbb{R}$ and for a small ϵ,

$$\boldsymbol{x}^{(k)} = c^{-1}\Big(\boldsymbol{w}_1 + \epsilon \sum_{i \geqslant 2} a_i \boldsymbol{w}_i\Big)\,,$$

where c is the appropriate normalising constant. Show that $s_k = \lambda_1 - K\epsilon^2 + \mathcal{O}(\epsilon^4)$ and determine the value of K. Hence show that

$$\boldsymbol{x}^{(k+1)} = c_1^{-1}\Big(\boldsymbol{w}_1 + \epsilon^3 \sum_{i \geqslant 2} b_i \boldsymbol{w}_i + \mathcal{O}(\epsilon^5)\Big),$$

where c_1 is the appropriate normalising constant, and find expressions for b_i.

Paper 2, Section II

40E Numerical Analysis

(a) For $A \in \mathbb{R}^{n\times n}$ and nonzero $\boldsymbol{v} \in \mathbb{R}^n$, define the *m-th Krylov subspace* $K_m(A, \boldsymbol{v})$ of $\mathbb{R}^n$. Prove that if A has n linearly independent eigenvectors with at most s distinct eigenvalues, then

$$\dim K_m(A, \boldsymbol{v}) \leqslant s \quad \forall m\,.$$

(b) Define the term *residual* in the conjugate gradient (CG) method for solving a system $A\boldsymbol{x} = \boldsymbol{b}$ with a symmetric positive definite A. State two properties of the method regarding residuals and their connection to certain Krylov subspaces, and hence show that, for any right-hand side $\boldsymbol{b}$, the method finds the exact solution after at most s iterations, where s is the number of distinct eigenvalues of A.

(c) The preconditioned CG-method $PAP^T\widehat{\boldsymbol{x}} = P\boldsymbol{b}$ is applied for solving $A\boldsymbol{x} = \boldsymbol{b}$, with

$$A = \begin{bmatrix} 2 & 1 & & \\ 1 & 2 & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & 2 \end{bmatrix}, \quad P^{-1} = Q = \begin{bmatrix} 1 & & & \\ 1 & 1 & & \\ & \ddots & \ddots & \\ & & 1 & 1 \end{bmatrix}.$$

Prove that the method finds the exact solution after two iterations at most.

(d) Prove that, for any symmetric positive definite A, we can find a preconditioner P such that the preconditioned CG-method for solving $A\boldsymbol{x} = \boldsymbol{b}$ would require only one step. Explain why this preconditioning is of hardly any use.

Paper 3, Section II

40E Numerical Analysis

(a) Give the definition of a *normal* matrix. Prove that if A is normal, then the (Euclidean) matrix ℓ_2-norm of A is equal to its spectral radius, i.e., $\|A\|_2 = \rho(A)$.

(b) The advection equation

$$u_t = u_x\,, \qquad 0 \leqslant x \leqslant 1, \qquad 0 \leqslant t < \infty,$$

is discretized by the Crank–Nicolson scheme

$$u_m^{n+1} - u_m^n = \tfrac{1}{4}\mu(u_{m+1}^{n+1} - u_{m-1}^{n+1}) + \tfrac{1}{4}\mu(u_{m+1}^n - u_{m-1}^n), \qquad m = 1, 2, \ldots, M\,, \quad n \in \mathbb{Z}_+\,.$$

Here, $\mu = \frac{k}{h}$ is the Courant number, with $k = \Delta t$, $h = \Delta x = \frac{1}{M+1}$, and u_m^n is an approximation to $u(mh, nk)$.

Using the eigenvalue analysis and carefully justifying each step, determine conditions on $\mu > 0$ for which the method is stable. [*Hint: All $M \times M$ Toeplitz anti-symmetric tridiagonal (TAT) matrices have the same set of orthogonal eigenvectors, and a TAT matrix with the elements $a_{j,j} = a$ and $a_{j,j+1} = -a_{j,j-1} = b$ has the eigenvalues $\lambda_k = a + 2\mathrm{i}b\cos\frac{\pi k}{M+1}$ where $\mathrm{i} = \sqrt{-1}$.*]

(c) Consider the same advection equation for the Cauchy problem ($x \in \mathbb{R},\ 0 \leqslant t \leqslant T$). Now it is discretized by the two-step leapfrog scheme

$$u_m^{n+1} = \mu\,(u_{m+1}^n - u_{m-1}^n) + u_m^{n-1}.$$

Applying the Fourier technique, find the range of $\mu > 0$ for which the method is stable.

Paper 4, Section II

40E Numerical Analysis

(a) For a function $f = f(x, y)$ which is real analytic in $\mathbb{R}^2$ and 2-periodic in each variable, its Fourier expansion is given by the formula

$$f(x,y) = \sum_{m,n\in\mathbb{Z}} \widehat{f}_{m,n} e^{i\pi mx + i\pi ny}, \qquad \widehat{f}_{m,n} = \frac{1}{4}\int_{-1}^{1}\int_{-1}^{1} f(t,\theta) e^{-i\pi mt - i\pi n\theta}\, dt d\theta\,.$$

Derive expressions for the Fourier coefficients of partial derivatives f_x, f_y and those of the product $h(x, y) = f(x, y)g(x, y)$ in terms of $\widehat{f}_{m,n}$ and $\widehat{g}_{m,n}$.

(b) Let $u(x, y)$ be the 2-periodic solution in $\mathbb{R}^2$ of the general second-order elliptic PDE

$$(au_x)_x + (au_y)_y = f,$$

where a and f are both real analytic and 2-periodic, and $a(x, y) > 0$. We impose the normalisation condition $\int_{-1}^{1}\int_{-1}^{1} u\, dxdy = 0$ and note from the PDE $\int_{-1}^{1}\int_{-1}^{1} f\, dxdy = 0$.

Construct explicitly the infinite-dimensional linear algebraic system that arises from the application of the Fourier spectral method to the above equation, and explain how to truncate this system to a finite-dimensional one.

(c) Specify the truncated system for the unknowns $\{\widehat{u}_{m,n}\}$ for the case

$$a(x,y) = 5 + 2\cos\pi x + 2\cos\pi y\,,$$

and prove that, for any ordering of the Fourier coefficients $\{\widehat{u}_{m,n}\}$ into one-dimensional array, the resulting system is symmetric and positive definite. [You may use the Gershgorin theorem without proof.]

Paper 4, Section II

39C Numerical Analysis

For a 2-periodic analytic function f, its Fourier expansion is given by the formula

$$f(x) = \sum_{n=-\infty}^{\infty} \widehat{f}_n e^{i\pi n x}, \qquad \widehat{f}_n = \frac{1}{2}\int_{-1}^{1} f(t)e^{-i\pi n t}\,dt\,.$$

(a) Consider the two-point boundary value problem

$$-\frac{1}{\pi^2}(1+2\cos\pi x)u'' + u = 1 + \sum_{n=1}^{\infty}\frac{2}{n^2+1}\cos\pi n x, \qquad -1 \leqslant x \leqslant 1\,,$$

with periodic boundary conditions $u(-1) = u(1)$. Construct explicitly the infinite dimensional linear algebraic system that arises from the application of the Fourier spectral method to the above equation, and explain how to truncate the system to a finite-dimensional one.

(b) A rectangle rule is applied to computing the integral of a 2-periodic analytic function h:

$$\int_{-1}^{1} h(t)\,dt \approx \frac{2}{N}\sum_{k=-N/2+1}^{N/2} h\left(\frac{2k}{N}\right). \qquad (*)$$

Find an expression for the error $e_N(h) :=$ LHS $-$ RHS of $(*)$, in terms of $\widehat{h}_n$, and show that $e_N(h)$ has a spectral rate of decay as $N \to \infty$. [In the last part, you may quote a relevant theorem about $\widehat{h}_n$.]

Paper 2, Section II

39C Numerical Analysis

The Poisson equation on the unit square, equipped with zero boundary conditions, is discretized with the 9-point scheme:

$$-\tfrac{10}{3}u_{i,j} + \tfrac{2}{3}(u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1}) \\ + \tfrac{1}{6}(u_{i+1,j+1} + u_{i+1,j-1} + u_{i-1,j+1} + u_{i-1,j-1}) = h^2 f_{i,j}\,,$$

where $1 \leqslant i, j \leqslant m$, $u_{i,j} \approx u(ih, jh)$, and (ih, jh) are the grid points with $h = \frac{1}{m+1}$. We also assume that $u_{0,j} = u_{i,0} = u_{m+1,j} = u_{i,m+1} = 0$.

(a) Prove that all $m \times m$ tridiagonal symmetric Toeplitz (TST-) matrices

$$H = [\beta, \alpha, \beta] := \begin{bmatrix} \alpha & \beta & & \\ \beta & \alpha & \ddots & \\ & \ddots & \ddots & \beta \\ & & \beta & \alpha \end{bmatrix} \in \mathbb{R}^{m \times m} \tag{1}$$

share the same eigenvectors $\boldsymbol{q}_k$ with the components $(\sin jk\pi h)_{j=1}^m$ for $k = 1, ..., m$. Find expressions for the corresponding eigenvalues λ_k for $k = 1, ..., m$. Deduce that $H = QDQ^{-1}$, where $D = \operatorname{diag}\{\lambda_k\}$ and Q is the matrix $(\sin ij\pi h)_{i,j=1}^m$.

(b) Show that, by arranging the grid points (ih, jh) into a one-dimensional array by columns, the 9-points scheme results in the following system of linear equations of the form

$$A\boldsymbol{u} = \boldsymbol{b} \quad \Leftrightarrow \quad \begin{bmatrix} B & C & & \\ C & B & \ddots & \\ & \ddots & \ddots & C \\ & & C & B \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \vdots \\ \boldsymbol{u}_m \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \\ \boldsymbol{b}_m \end{bmatrix}, \tag{2}$$

where $A \in \mathbb{R}^{m^2 \times m^2}$, the vectors $\boldsymbol{u}_k, \boldsymbol{b}_k \in \mathbb{R}^m$ are portions of $\boldsymbol{u}, \boldsymbol{b} \in \mathbb{R}^{m^2}$, respectively, and B, C are $m \times m$ TST-matrices whose elements you should determine.

(c) Using $\boldsymbol{v}_k = Q^{-1}\boldsymbol{u}_k$, $\boldsymbol{c}_k = Q^{-1}\boldsymbol{b}_k$, show that (2) is equivalent to

$$\begin{bmatrix} D & E & & \\ E & D & \ddots & \\ & \ddots & \ddots & E \\ & & E & D \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \\ \vdots \\ \boldsymbol{v}_m \end{bmatrix} = \begin{bmatrix} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \\ \vdots \\ \boldsymbol{c}_m \end{bmatrix}, \tag{3}$$

where D and E are diagonal matrices.

(d) Show that, by appropriate reordering of the grid, the system (3) is reduced to m uncoupled $m \times m$ systems of the form

$$\Lambda_k \widehat{v}_k = \widehat{c}_k, \qquad k = 1, \ldots, m.$$

Determine the elements of the matrices Λ_k.

Paper 3, Section II

40C Numerical Analysis

The diffusion equation

$$u_t = u_{xx}, \qquad 0 \leqslant x \leqslant 1, \quad t \geqslant 0,$$

with the initial condition $u(x, 0) = \phi(x)$, $0 \leqslant x \leqslant 1$, and boundary conditions $u(0, t) = u(1, t) = 0$, is discretised by $u_m^n \approx u(mh, nk)$ with $k = \Delta t$, $h = \Delta x = 1/(1 + M)$. The Courant number is given by $\mu = k/h^2$.

(a) The system is solved numerically by the method

$$u_m^{n+1} = u_m^n + \mu \left(u_{m-1}^n - 2u_m^n + u_{m+1}^n\right), \qquad m = 1, 2, ..., M, \quad n \geqslant 0\,.$$

Prove directly that $\mu \leqslant 1/2$ implies convergence.

(b) Now consider the method

$$au_m^{n+1} - \tfrac{1}{4}(\mu - c)\left(u_{m-1}^{n+1} - 2u_m^{n+1} + u_{m+1}^{n+1}\right) = au_m^n + \tfrac{1}{4}(\mu + c)\left(u_{m-1}^n - 2u_m^n + u_{m+1}^n\right)\,,$$

where a and c are real constants. Using an eigenvalue analysis and carefully justifying each step, determine conditions on μ, a and c for this method to be stable.

[*You may use the notation* $[\beta, \alpha, \beta]$ *for the tridiagonal matrix with* α *along the diagonal, and* β *along the sub- and super-diagonals and use without proof any relevant theorems about such matrices.*]

Paper 1, Section II

40C Numerical Analysis

(a) Describe the *Jacobi method* for solving a system of linear equations $A\boldsymbol{x} = \boldsymbol{b}$ as a particular case of splitting, and state the criterion for its convergence in terms of the iteration matrix.

(b) For the case when

$$A = \begin{bmatrix} 1 & \alpha & \alpha \\ \alpha & 1 & \alpha \\ \alpha & \alpha & 1 \end{bmatrix},$$

find the exact range of the parameter α for which the Jacobi method converges.

(c) State the *Householder-John theorem* and deduce that the Jacobi method converges if A is a symmetric positive-definite tridiagonal matrix.

Paper 4, Section II

40E Numerical Analysis

The inverse discrete Fourier transform $\mathcal{F}_n^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ is given by the formula

$$\boldsymbol{x} = \mathcal{F}_n^{-1}\boldsymbol{y}, \qquad \text{where} \qquad x_\ell = \sum_{j=0}^{n-1} \omega_n^{j\ell} y_j, \quad \ell = 0, \ldots, n-1.$$

Here, $\omega_n = \exp \frac{2\pi i}{n}$ is the primitive root of unity of degree n and $n = 2^p$, $p = 1, 2, \ldots$.

(a) Show how to assemble $\boldsymbol{x} = \mathcal{F}_{2m}^{-1}\boldsymbol{y}$ in a small number of operations if the Fourier transforms of the even and odd parts of $\boldsymbol{y}$,

$$\boldsymbol{x}^{(\mathrm{E})} = \mathcal{F}_m^{-1}\boldsymbol{y}^{(\mathrm{E})}, \qquad \boldsymbol{x}^{(\mathrm{O})} = \mathcal{F}_m^{-1}\boldsymbol{y}^{(\mathrm{O})},$$

are already known.

(b) Describe the Fast Fourier Transform (FFT) method for evaluating $\boldsymbol{x}$, and draw a diagram to illustrate the method for $n = 8$.

(c) Find the cost of the FFT method for $n = 2^p$ (only multiplications count).

(d) For $n = 4$ use the FFT method to find $\boldsymbol{x} = \mathcal{F}_n^{-1}\boldsymbol{y}$ when:

$$\text{(i)} \quad \boldsymbol{y} = (1, -1, 1, -1), \qquad \text{(ii)} \quad \boldsymbol{y} = (1, 1, -1, -1).$$

Paper 2, Section II

40E Numerical Analysis

The Poisson equation $\frac{d^2u}{dx^2} = f$ in the unit interval $[0,1]$, with $u(0) = u(1) = 0$, is discretised with the formula

$$u_{i-1} - 2u_i + u_{i+1} = h^2 f_i, \qquad 1 \leqslant i \leqslant n,$$

where $u_0 = u_{n+1} = 0$, $h = (n+1)^{-1}$, the grid points are at $x = ih$ and $u_i \approx u(ih)$.

(a) Write the above system of equations in the vector form $A\boldsymbol{u} = \boldsymbol{b}$ and describe the relaxed Jacobi method with relaxation parameter ω for solving this linear system.

(b) For $\boldsymbol{x}^*$ and $\boldsymbol{x}^{(\nu)}$ being the exact and the iterated solution, respectively, let $\boldsymbol{e}^{(\nu)} := \boldsymbol{x}^{(\nu)} - \boldsymbol{x}^*$ be the error and H_ω be the iteration matrix, so that

$$\boldsymbol{e}^{(\nu+1)} = H_\omega \, \boldsymbol{e}^{(\nu)} .$$

Express H_ω in terms of the matrix A and the relaxation parameter ω. Using the fact that for any $n \times n$ Toeplitz symmetric tridiagonal matrix, the eigenvectors $\boldsymbol{v}_k$ $(k = 1, \ldots, n)$ have the form:

$$\boldsymbol{v}_k = (\sin ikx)_{i=1}^n, \qquad x = \tfrac{\pi}{n+1} \, ,$$

find the eigenvalues $\lambda_k(A)$ of A. Hence deduce the eigenvalues $\lambda_k(\omega)$ of H_ω.

(c) For A as above, let

$$\boldsymbol{e}^{(\nu)} = \sum_{k=1}^{n} a_k^{(\nu)} \boldsymbol{v}_k$$

be the expansion of the error with respect to the eigenvectors $(\boldsymbol{v}_k)$ of H_ω.

Find the range of the parameter ω which provides convergence of the method for any n, and prove that, for any such ω, the rate of convergence $\boldsymbol{e}^{(\nu)} \to 0$ is not faster than $(1 - c/n^2)^\nu$ when n is large.

(d) Show that, for an appropriate range of ω, the high frequency components $a_k^{(\nu)}$ $(\frac{n+1}{2} \leqslant k \leqslant n)$ of the error $\boldsymbol{e}^{(\nu)}$ tend to zero much faster than the rate obtained in part (c). Determine the optimal parameter ω_* which provides the largest supression of the high frequency components per iteration, and find the corresponding attenuation factor μ_* assuming n is large. That is, find the least μ_ω such that $|a_k^{(\nu+1)}| \leqslant \mu_\omega |a_k^{(\nu)}|$ for $\frac{n+1}{2} \leqslant k \leqslant n$.

Paper 1, Section II

40E Numerical Analysis

(a) Suppose that A is a real $n \times n$ matrix, and $\boldsymbol{w} \in \mathbb{R}^n$ and $\lambda_1 \in \mathbb{R}$ are given so that $A\boldsymbol{w} = \lambda_1 \boldsymbol{w}$. Further, let S be a non-singular matrix such that $S\boldsymbol{w} = c\boldsymbol{e}^{(1)}$, where $\boldsymbol{e}^{(1)}$ is the first coordinate vector and $c \neq 0$.

Let $\widehat{A} = SAS^{-1}$. Prove that the eigenvalues of A are λ_1 together with the eigenvalues of the bottom right $(n-1) \times (n-1)$ submatrix of $\widehat{A}$.

Explain briefly how, given a vector $\boldsymbol{w}$, an orthogonal matrix S such that $S\boldsymbol{w} = c\boldsymbol{e}^{(1)}$ can be constructed.

(b) Suppose that A is a real $n \times n$ matrix, and two linearly independent vectors $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n$ are given such that the linear subspace $L\{\boldsymbol{v}, \boldsymbol{w}\}$ spanned by $\boldsymbol{v}$ and $\boldsymbol{w}$ is invariant under the action of A, i.e.,

$$x \in L\{\boldsymbol{v}, \boldsymbol{w}\} \quad \Rightarrow \quad Ax \in L\{\boldsymbol{v}, \boldsymbol{w}\}.$$

Denote by V an $n \times 2$ matrix whose two columns are the vectors $\boldsymbol{v}$ and $\boldsymbol{w}$, and let S be a non-singular matrix such that $R = SV$ is upper triangular:

$$SV = S \times \begin{bmatrix} v_1 & w_1 \\ v_2 & w_2 \\ v_3 & w_3 \\ \vdots & \vdots \\ v_n & w_n \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} \\ 0 & r_{22} \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix}. \qquad (*)$$

Again, let $\widehat{A} = SAS^{-1}$. Prove that the eigenvalues of A are the eigenvalues of the top left 2×2 submatrix of $\widehat{A}$ together with the eigenvalues of the bottom right $(n-2) \times (n-2)$ submatrix of $\widehat{A}$.

Explain briefly how, for given vectors $\boldsymbol{v}$, $\boldsymbol{w}$, an orthogonal matrix S which satisfies $(*)$ can be constructed.

Paper 3, Section II

41E Numerical Analysis

The diffusion equation for $u(x,t)$:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \qquad x \in \mathbb{R}, \quad t \geqslant 0,$$

is solved numerically by the difference scheme

$$u_m^{n+1} = u_m^n + \tfrac{3}{2}\mu(u_{m-1}^n - 2u_m^n + u_{m+1}^n) - \tfrac{1}{2}\mu(u_{m-1}^{n-1} - 2u_m^{n-1} + u_{m+1}^{n-1})\,.$$

Here $\mu = \frac{k}{h^2}$ is the Courant number, with $k = \Delta t$, $h = \Delta x$, and $u_m^n \approx u(mh, nk)$.

(a) Prove that, as $k \to 0$ with constant μ, the local error of the method is $\mathcal{O}(k^2)$.

(b) Applying the Fourier stability analysis, show that the method is stable if and only if $\mu \leqslant \frac{1}{4}$. [*Hint: If a polynomial $p(x) = x^2 - 2\alpha x + \beta$ has real roots, then those roots lie in $[a,b]$ if and only if $p(a)p(b) \geqslant 0$ and $\alpha \in [a,b]$.*]

(c) Prove that, for the same equation, the leapfrog scheme

$$u_m^{n+1} = u_m^{n-1} + 2\mu(u_{m-1}^n - 2u_m^n + u_{m+1}^n)$$

is unstable for any choice of $\mu > 0$.

38A Numerical Analysis

The Poisson equation $\nabla^2 u = f$ in the unit square $\Omega = [0,1] \times [0,1]$, equipped with the zero Dirichlet boundary conditions on $\partial\Omega$, is discretized with the nine-point formula:

$$\begin{aligned}\Gamma_9[u_{i,j}] := -\frac{10}{3}\, u_{i,j} + \frac{2}{3}\,(u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1}) \\ +\frac{1}{6}\,(u_{i+1,j+1} + u_{i+1,j-1} + u_{i-1,j+1} + u_{i-1,j-1}) = h^2 f_{i,j},\end{aligned}$$

where $1 \leqslant i,j \leqslant m$, $u_{i,j} \approx u(ih, jh)$, and (ih, jh) are the grid points with $h = \frac{1}{m+1}$.

(i) Find the order of the local truncation error $\eta_{i,j}$ of the approximation.

(ii) Prove that the order of the truncation error is smaller if f satisfies the Laplace equation $\nabla^2 f = 0$.

(iii) Show that the modified nine-point scheme

$$\begin{aligned}\Gamma_9[u_{i,j}] &= h^2 f_{i,j} + \frac{1}{12} h^2 \Gamma_5[f_{i,j}] \\ &:= h^2 f_{i,j} + \frac{1}{12} h^2 (f_{i+1,j} + f_{i-1,j} + f_{i,j+1} + f_{i,j-1} - 4 f_{i,j})\end{aligned}$$

has a truncation error of the same order as in part (ii).

(iv) Let $(u_{i,j})_{i,j=1}^m$ be a solution to the $m^2 \times m^2$ system of linear equations $A\mathbf{u} = \mathbf{b}$ arising from the modified nine-point scheme in part (iii). Further, let $u(x,y)$ be the exact solution and let $e_{i,j} := u_{i,j} - u(ih, jh)$ be the error of approximation at grid points. Prove that there exists a constant c such that

$$\|\mathbf{e}\|_2 := \left[\sum_{i,j=1}^m |e_{i,j}|^2\right]^{1/2} < ch^3, \qquad h \to 0.$$

[*Hint: The nine-point discretization of* $\nabla^2 u$ *can be written as*

$$\Gamma_9[u] = (\Gamma_5 + \tfrac{1}{6}\Delta_x^2 \Delta_y^2)\, u = (\Delta_x^2 + \Delta_y^2 + \tfrac{1}{6}\Delta_x^2\Delta_y^2)\, u\,,$$

where $\Gamma_5[u] = (\Delta_x^2 + \Delta_y^2) u$ *is the five-point discretization and*

$$\begin{aligned}\Delta_x^2\, u(x,y) &:= u(x-h, y) - 2u(x,y) + u(x+h, y), \\ \Delta_y^2\, u(x,y) &:= u(x, y-h) - 2u(x,y) + u(x, y+h)\,.\end{aligned}$$]

[*Hint: The matrix* A *of the nine-point scheme is symmetric, with the eigenvalues*

$$\lambda_{k,\ell} = -\,4\sin^2\tfrac{k\pi h}{2} - 4\sin^2\tfrac{\ell\pi h}{2} + \tfrac{8}{3}\sin^2\tfrac{k\pi h}{2}\sin^2\tfrac{\ell\pi h}{2}\,, \qquad 1 \leqslant k,\ell \leqslant m\,.$$]

Paper 1, Section II

39A Numerical Analysis

State the Householder–John theorem and explain how it can be used in designing iterative methods for solving a system of linear equations $A\mathbf{x} = \mathbf{b}$. [You may quote other relevant theorems if needed.]

Consider the following iterative scheme for solving $A\mathbf{x} = \mathbf{b}$. Let $A = L + D + U$, where D is the diagonal part of A, and L and U are the strictly lower and upper triangular parts of A, respectively. Then, with some starting vector $\mathbf{x}^{(0)}$, the scheme is as follows:

$$(D + \omega L)\mathbf{x}^{(k+1)} = \big[(1-\omega)D - \omega U\big]\mathbf{x}^{(k)} + \omega \mathbf{b}\,.$$

Prove that if A is a symmetric positive definite matrix and $\omega \in (0, 2)$, then, for any $\mathbf{x}^{(0)}$, the above iteration converges to the solution of the system $A\mathbf{x} = \mathbf{b}$.

Which method corresponds to the case $\omega = 1$?

Paper 3, Section II

39A Numerical Analysis

Let A be a real symmetric $n \times n$ matrix with real and distinct eigenvalues $0 = \lambda_1 < \cdots < \lambda_{n-1} = 1 < \lambda_n$ and a corresponding orthogonal basis of normalized real eigenvectors $(\mathbf{w}_i)_{i=1}^n$.

To estimate the eigenvector $\mathbf{w}_n$ of A whose eigenvalue is λ_n, the power method with shifts is employed which has the following form:

$$\mathbf{y} = (A - s_k I)\mathbf{x}^{(k)}, \qquad \mathbf{x}^{(k+1)} = \mathbf{y}/\|\mathbf{y}\|\,, \qquad s_k \in \mathbb{R}, \qquad k = 0, 1, 2, \ldots$$

Three versions of this method are considered:

(i) no shift: $s_k \equiv 0$;

(ii) single shift: $s_k \equiv \frac{1}{2}$;

(iii) double shift: $s_{2\ell} \equiv s_0 = \frac{1}{4}(2+\sqrt{2})$, $s_{2\ell+1} \equiv s_1 = \frac{1}{4}(2-\sqrt{2})$.

Assume that $\lambda_n = 1 + \epsilon$, where $\epsilon > 0$ is very small, so that the terms $\mathcal{O}(\epsilon^2)$ are negligible, and that $\mathbf{x}^{(0)}$ contains substantial components of all the eigenvectors.

By considering the approximation after $2m$ iterations in the form

$$\mathbf{x}^{(2m)} = \pm\mathbf{w}_n + \mathcal{O}(\rho^{2m}) \quad (m \to \infty),$$

find ρ as a function of ϵ for each of the three versions of the method.

Compare the convergence rates of the three versions of the method, with reference to the number of iterations needed to achieve a prescribed accuracy.

Paper 4, Section II
39A Numerical Analysis

(a) The diffusion equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad 0 \leqslant x \leqslant 1, \quad 0 \leqslant t \leqslant T,$$

is approximated by the Crank–Nicolson scheme

$$u_m^{n+1} - \frac{1}{2}\mu\left(u_{m-1}^{n+1} - 2u_m^{n+1} + u_{m+1}^{n+1}\right) = u_m^n + \frac{1}{2}\mu\left(u_{m-1}^n - 2u_m^n + u_{m+1}^n\right),$$

with $m = 1, \ldots, M$. Here $\mu = k/h^2$, $k = \Delta t$, $h = \Delta x = \frac{1}{M+1}$, and u_m^n is an approximation to $u(mh, nk)$. Assuming that $u(0,t) = u(1,t) = 0$, show that the above scheme can be written in the form

$$B\mathbf{u}^{n+1} = C\mathbf{u}^n, \quad 0 \leqslant n \leqslant T/k - 1,$$

where $\mathbf{u}^n = [u_1^n, \ldots, u_M^n]^T$ and the real matrices B and C should be found. Using matrix analysis, find the range of $\mu > 0$ for which the scheme is stable.

[*Hint: All Toeplitz symmetric tridiagonal (TST) matrices have the same set of orthogonal eigenvectors, and a TST matrix with the elements* $a_{i,i} = a$ *and* $a_{i,i\pm1} = b$ *has the eigenvalues* $\lambda_k = a + 2b\cos\frac{\pi k}{M+1}$.]

(b) The wave equation

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}, \qquad x \in \mathbb{R}, \qquad t \geqslant 0,$$

with given initial conditions for u and $\partial u/\partial t$, is approximated by the scheme

$$u_m^{n+1} - 2u_m^n + u_m^{n-1} = \mu(u_{m+1}^n - 2u_m^n + u_{m-1}^n),$$

with the Courant number now $\mu = k^2/h^2$. Applying the Fourier technique, find the range of $\mu > 0$ for which the method is stable.

Paper 4, Section II

38B Numerical Analysis

(a) Describe an implementation of the *power method* for determining the eigenvalue of largest modulus and its associated eigenvector for a matrix that has a unique eigenvalue of largest modulus.

Now let A be a real $n \times n$ matrix with distinct eigenvalues satisfying $|\lambda_n| = |\lambda_{n-1}|$ and $|\lambda_n| > |\lambda_i|$, $i = 1, \ldots, n-2$. The power method is applied to A, with an initial condition $\mathbf{x}^{(0)} = \sum_{i=1}^{n} c_i \mathbf{w}_i$ such that $c_{n-1}c_n \neq 0$, where $\mathbf{w}_i$ is the eigenvector associated with λ_i. Show that the power method does not converge. Explain why $\mathbf{x}^{(k)}$, $\mathbf{x}^{(k+1)}$ and $\mathbf{x}^{(k+2)}$ become linearly dependent as $k \to \infty$.

(b) Consider the following variant of the power method, called the two-stage power method, applied to the matrix A of part (a):

0. Pick $\mathbf{x}^{(0)} \in \mathbb{R}^n$ satisfying $\|\mathbf{x}^{(0)}\| = 1$. Let $0 < \varepsilon \ll 1$. Set $k = 0$ and $\mathbf{x}^{(1)} = A\mathbf{x}^{(0)}$.

1. Calculate $\mathbf{x}^{(k+2)} = A\mathbf{x}^{(k+1)}$ and calculate α, β that minimise
$$f(\alpha, \beta) = \|\mathbf{x}^{(k+2)} + \alpha\mathbf{x}^{(k+1)} + \beta\mathbf{x}^{(k)}\|.$$

2. If $f(\alpha, \beta) \leqslant \varepsilon$, solve $\lambda^2 + \alpha\lambda + \beta = 0$ and let the roots be λ_1 and λ_2. They are accepted as eigenvalues of A, and the corresponding eigenvectors are estimated as $\mathbf{x}^{(k+1)} - \lambda_2\mathbf{x}^{(k)}$ and $\mathbf{x}^{(k+1)} - \lambda_1\mathbf{x}^{(k)}$.

3. Otherwise, divide $\mathbf{x}^{(k+2)}$ and $\mathbf{x}^{(k+1)}$ by the current value of $\|\mathbf{x}^{(k+1)}\|$, increase k by 1 and return to Step **1.**

Explain the justification behind Step **2** of the algorithm.

(c) Let $n = 3$, and suppose that, for a large value of k, the two-stage power method yields the vectors

$$\mathbf{y}_k = \mathbf{x}^{(k)} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{y}_{k+1} = A\mathbf{x}^{(k)} = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}, \quad \mathbf{y}_{k+2} = A^2\mathbf{x}^{(k)} = \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}.$$

Find two eigenvalues of A and the corresponding eigenvectors.

Paper 2, Section II

38B Numerical Analysis

(a) The advection equation

$$u_t = u_x, \quad 0 \leqslant x \leqslant 1,\ t \geqslant 0$$

is discretised using an equidistant grid with stepsizes $\Delta x = h$ and $\Delta t = k$. The spatial derivatives are approximated with central differences and the resulting ODEs are approximated with the trapezoidal rule. Write down the relevant difference equation for determining (u_m^{n+1}) from (u_m^n). What is the name of this scheme? What is the local truncation error?

The boundary condition is periodic, $u(0,t) = u(1,t)$. Explain briefly how to write the discretised scheme in the form $B\mathbf{u}^{n+1} = C\mathbf{u}^n$, where the matrices B and C, to be identified, have a circulant form. Using matrix analysis, find the range of $\mu = \Delta t/\Delta x$ for which the scheme is stable. [Standard results may be used without proof if quoted carefully.]

[*Hint: An $n \times n$* circulant *matrix has the form*

$$A = \begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_{n-1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_1 \\ a_1 & \cdots & a_{n-1} & a_0 \end{pmatrix}.$$

All such matrices have the same set of eigenvectors $\mathbf{v}_\ell = \left(\omega^{j\ell}\right)_{j=0}^{n-1}$, $\ell = 0, 1, \ldots, n-1$, where $\omega = e^{2\pi i/n}$, and the corresponding eigenvalues are $\lambda_\ell = \sum_{k=0}^{n-1} a_k \omega^{k\ell}$.]

(b) Consider the advection equation on the unit square

$$u_t = a u_x + b u_y, \quad 0 \leqslant x, y \leqslant 1,\ t \geqslant 0\,,$$

where u satisfies doubly periodic boundary conditions, $u(0,y) = u(1,y)$, $u(x,0) = u(x,1)$, and $a(x,y)$ and $b(x,y)$ are given doubly periodic functions. The system is discretised with the Crank–Nicolson scheme, with central differences for the space derivatives, using an equidistant grid with stepsizes $\Delta x = \Delta y = h$ and $\Delta t = k$. Write down the relevant difference equation, and show how to write the scheme in the form

$$\mathbf{u}^{n+1} = (I - \tfrac{1}{4}\mu A)^{-1}(I + \tfrac{1}{4}\mu A)\mathbf{u}^n\,, \qquad (*)$$

where the matrix A should be identified. Describe how $(*)$ can be approximated by Strang splitting, and explain the advantages of doing so.

[*Hint: Inversion of the matrix B in part (a) has a similar computational cost to that of a tridiagonal matrix.*]

Paper 1, Section II

38B Numerical Analysis

(a) Consider the periodic function

$$f(x) = 5 + 2\cos\left(2\pi x - \frac{\pi}{2}\right) + 3\cos(4\pi x)$$

on the interval $[0,1]$. The N-point discrete Fourier transform of f is defined by

$$F_N(n) = \frac{1}{N}\sum_{k=0}^{N-1} f_k\, \omega_N^{-nk}, \quad n = 0, 1, \ldots, N-1, \qquad (*)$$

where $\omega_N = e^{2\pi i/N}$ and $f_k = f(k/N)$. Compute $F_4(n)$, $n = 0, \ldots, 3$, using the Fast Fourier Transform (FFT). Compare your result with what you get by computing $F_4(n)$ directly from $(*)$. Carefully explain all your computations.

(b) Now let f be any analytic function on $\mathbb{R}$ that is periodic with period 1. The continuous Fourier transform of f is defined by

$$\hat{f}_n = \int_0^1 f(\tau)\, e^{-2\pi i n\tau}\, d\tau\,, \quad n \in \mathbb{Z}\,.$$

Use integration by parts to show that the Fourier coefficients $\hat{f}_n$ decay spectrally.

Explain what it means for the discrete Fourier transform of f to approximate the continuous Fourier transform with *spectral accuracy*. Prove that it does so.

What can you say about the behaviour of $F_N(N-n)$ as $N \to \infty$ for fixed n?

Paper 3, Section II

38B Numerical Analysis

(a) Define the Jacobi and Gauss–Seidel iteration schemes for solving a linear system of the form $A\mathbf{u} = \mathbf{b}$, where $\mathbf{u}, \mathbf{b} \in \mathbb{R}^M$ and $A \in \mathbb{R}^{M \times M}$, giving formulae for the corresponding iteration matrices H_J and H_{GS} in terms of the usual decomposition $A = L_0 + D + U_0$.

Show that both iteration schemes converge when A results from discretization of Poisson's equation on a square with the five-point formula, that is when

$$A = \begin{bmatrix} S & I & & & \\ I & S & I & & \\ & \ddots & \ddots & \ddots & \\ & & I & S & I \\ & & & I & S \end{bmatrix}, \qquad S = \begin{bmatrix} -4 & 1 & & & \\ 1 & -4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -4 & 1 \\ & & & 1 & -4 \end{bmatrix} \in \mathbb{R}^{m \times m} \qquad (*)$$

and $M = m^2$. [You may state the Householder–John theorem without proof.]

(b) For the matrix A given in $(*)$:

(i) Calculate the eigenvalues of H_J and deduce its spectral radius $\rho(H_J)$.

(ii) Show that each eigenvector $\mathbf{q}$ of H_{GS} is related to an eigenvector $\mathbf{p}$ of H_J by a transformation of the form $q_{i,j} = \alpha^{i+j} p_{i,j}$ for a suitable value of α.

(iii) Deduce that $\rho(H_{GS}) = \rho^2(H_J)$. What is the significance of this result for the two iteration schemes?

[*Hint: You may assume that the eigenvalues of the matrix A in $(*)$ are*

$$\lambda_{k,\ell} = -4\left(\sin^2\frac{x}{2} + \sin^2\frac{y}{2}\right), \quad \textit{where } x = \frac{k\pi h}{m+1}, \; y = \frac{\ell\pi h}{m+1}, \quad k, \ell = 1, \ldots, m,$$

with corresponding eigenvectors $\mathbf{v} = (v_{i,j})$, $\quad v_{i,j} = \sin ix \sin jy$, $\quad i, j = 1, \ldots, m$.]

Paper 4, Section II
37E Numerical Analysis

(a) Define the mth *Krylov space* $K_m(A, v)$ for $A \in \mathbb{R}^{n\times n}$ and $0 \neq v \in \mathbb{R}^n$. Letting δ_m be the dimension of $K_m(A, v)$, prove the following results.

(i) There exists a positive integer $s \leqslant n$ such that $\delta_m = m$ for $m \leqslant s$ and $\delta_m = s$ for $m > s$.

(ii) If $v = \sum_{i=1}^{s'} c_i w_i$, where w_i are eigenvectors of A for distinct eigenvalues and all c_i are nonzero, then $s = s'$.

(b) Define the term *residual* in the conjugate gradient (CG) method for solving a system $Ax = b$ with symmetric positive definite A. Explain (without proof) the connection to Krylov spaces and prove that for any right-hand side b the CG method finds an exact solution after at most t steps, where t is the number of distinct eigenvalues of A. [You may use without proof known properties of the iterates of the CG method.]

Define what is meant by preconditioning, and explain two ways in which preconditioning can speed up convergence. Can we choose the preconditioner so that the CG method requires only one step? If yes, is it a reasonable method for speeding up the computation?

Paper 2, Section II

37E Numerical Analysis

(a) The boundary value problem $-\Delta u + cu = f$ on the unit square $[0,1]^2$ with zero boundary conditions and scalar constant $c > 0$ is discretised using finite differences as

$$-u_{i-1,j} - u_{i+1,j} - u_{i,j-1} - u_{i,j+1} + 4u_{i,j} + ch^2 u_{i,j} = h^2 f(ih, jh), \qquad i,j = 1, \ldots, m,$$

with $h = 1/(m+1)$. Show that for the resulting system $Au = b$, for a suitable matrix A and vectors u and b, both the Jacobi and Gauss–Seidel methods converge. [You may cite and use known results on the discretised Laplace operator and on the convergence of iterative methods.]

Define the Jacobi method with relaxation parameter ω. Find the eigenvalues $\lambda_{k,l}$ of the iteration matrix H_ω for the above problem and show that, in order to ensure convergence for all h, the condition $0 < \omega \leqslant 1$ is necessary.

[*Hint: The eigenvalues of the discretised Laplace operator in two dimensions are* $4\left(\sin^2 \frac{\pi k h}{2} + \sin^2 \frac{\pi l h}{2}\right)$ *for integers* k, l.]

(b) Explain the components and steps in a multigrid method for solving the Poisson equation, discretised as $A_h u_h = b_h$. If we use the relaxed Jacobi method within the multigrid method, is it necessary to choose $\omega \neq 1$ to get fast convergence? Explain why or why not.

Paper 3, Section II
38E Numerical Analysis

(a) Given the finite-difference recurrence

$$\sum_{k=r}^{s} a_k u_{m+k}^{n+1} = \sum_{k=r}^{s} b_k u_{m+k}^{n}, \quad m \in \mathbb{Z},\ n \in \mathbb{Z}^+,$$

that discretises a Cauchy problem, the *amplification factor* is defined by

$$H(\theta) = \left(\sum_{k=r}^{s} b_k e^{ik\theta} \right) \Bigg/ \left(\sum_{k=r}^{s} a_k e^{ik\theta} \right).$$

Show how $H(\theta)$ acts on the Fourier transform $\hat{u}^n$ of u^n. Hence prove that the method is stable if and only if $|H(\theta)| \leqslant 1$ for all $\theta \in [-\pi, \pi]$.

(b) The two-dimensional diffusion equation

$$u_t = u_{xx} + c u_{yy}$$

for some scalar constant $c > 0$ is discretised with the forward Euler scheme

$$u_{i,j}^{n+1} = u_{i,j}^{n} + \mu(u_{i+1,j}^{n} - 2u_{i,j}^{n} + u_{i-1,j}^{n} + c u_{i,j+1}^{n} - 2c u_{i,j}^{n} + c u_{i,j-1}^{n}).$$

Using Fourier stability analysis, find the range of values $\mu > 0$ for which the scheme is stable.

Paper 1, Section II
38E Numerical Analysis

(a) The diffusion equation

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(a(x)\frac{\partial u}{\partial x}\right) \quad \text{in} \quad 0 \leqslant x \leqslant 1, \quad t \geqslant 0,$$

with the initial condition $u(x,0) = \phi(x)$ in $0 \leqslant x \leqslant 1$ and zero boundary conditions at $x = 0$ and $x = 1$, is solved by the finite-difference method

$$u_m^{n+1} = u_m^n + \mu\left[a_{m-\frac{1}{2}}u_{m-1}^n - (a_{m-\frac{1}{2}} + a_{m+\frac{1}{2}})u_m^n + a_{m+\frac{1}{2}}u_{m+1}^n\right], \qquad m = 1, 2, \ldots, M,$$

where $\mu = k/h^2,\ k = \Delta t,\ h = 1/(M+1),\ u_m^n \approx u(mh, nk)$, and $a_\alpha = a(\alpha h)$.

Assuming that the function a and the exact solution are sufficiently smooth, prove that the exact solution satisfies the numerical scheme with error $O(h^3)$ for constant μ.

(b) For the problem in part (a), assume that there exist $0 < a_- < a_+ < \infty$ such that $a_- \leqslant a(x) \leqslant a_+$ for all $0 \leqslant x \leqslant 1$. State (without proof) the Gershgorin theorem and prove that the method is stable for $0 < \mu \leqslant 1/(2a_+)$.

Paper 4, Section II

39D Numerical Analysis

Let A be a real symmetric $n \times n$ matrix with n distinct real eigenvalues $\lambda_1 < \lambda_2 < \cdots < \lambda_n$ and a corresponding orthogonal basis of normalized real eigenvectors $\{\mathbf{w}_i\}_{i=1}^n$.

(i) Let $s \in \mathbb{R}$ satisfy $s < \lambda_1$. Given a unit vector $\mathbf{x}^{(0)} \in \mathbb{R}^n$, the iteration scheme

$$\begin{aligned}(A - sI)\mathbf{y} &= \mathbf{x}^{(k)}, \\ \mathbf{x}^{(k+1)} &= \mathbf{y}/\|\mathbf{y}\|,\end{aligned}$$

generates a sequence of vectors $\mathbf{x}^{(k+1)}$ for $k = 0, 1, 2, \ldots$. Assuming that $\mathbf{x}^{(0)} = \sum c_i \mathbf{w}_i$ with $c_1 \neq 0$, prove that $\mathbf{x}^{(k)}$ tends to $\pm\,\mathbf{w}_1$ as $k \to \infty$. What happens to $\mathbf{x}^{(k)}$ if $s > \lambda_1$? [Consider all cases.]

(ii) Describe how to implement an inverse-iteration algorithm to compute the eigenvalues and eigenvectors of A, given some initial estimates for the eigenvalues.

(iii) Let $n = 2$. For iterates $\mathbf{x}^{(k)}$ of an inverse-iteration algorithm with a fixed value of $s \neq \lambda_1, \lambda_2$, show that if

$$\mathbf{x}^{(k)} = (\mathbf{w}_1 + \epsilon_k \mathbf{w}_2)/(1 + \epsilon_k^2)^{1/2},$$

where $|\epsilon_k|$ is small, then $|\epsilon_{k+1}|$ is of the same order of magnitude as $|\epsilon_k|$.

(iv) Let $n = 2$ still. Consider the iteration scheme

$$s_k = \left(\mathbf{x}^{(k)}, A\mathbf{x}^{(k)}\right), \qquad (A - s_k I)\mathbf{y} = \mathbf{x}^{(k)}, \qquad \mathbf{x}^{(k+1)} = \mathbf{y}/\|\mathbf{y}\|$$

for $k = 0, 1, 2, \ldots$, where $(\ ,\ )$ denotes the inner product. Show that with this scheme $|\epsilon_{k+1}| = |\epsilon_k|^3$.

Paper 2, Section II

39D Numerical Analysis

Consider the one-dimensional advection equation

$$u_t = u_x\,, \quad -\infty < x < \infty\,, \quad t \geqslant 0\,,$$

subject to an initial condition $u(x,0) = \varphi(x)$. Consider discretization of this equation with finite differences on an equidistant space-time $\{(mh, nk),\ m \in \mathbb{Z},\ n \in \mathbb{Z}^+\}$ with step size $h > 0$ in space and step size $k > 0$ in time. Define the *Courant number* μ and explain briefly how such a discretization can be used to derive numerical schemes in which solutions $u_m^n \approx u(mh, nk)$, $m \in \mathbb{Z}$ and $n \in \mathbb{Z}^+$ satisfy equations of the form

$$\sum_{i=r}^{s} a_i u_{m+i}^{n+1} = \sum_{i=r}^{s} b_i u_{m+i}^{n}\,, \tag{1}$$

where the coefficients a_i, b_i are independent of m, n.

(i) Define the *order* of a numerical scheme such as (1). Define what a *convergent* numerical scheme is. Explain the notion of *stability* and state the Lax equivalence theorem that connects convergence and stability of numerical schemes for linear partial differential equations.

(ii) Consider the following example of (1):

$$u_m^{n+1} = u_m^n + \frac{\mu}{2}(u_{m+1}^n - u_{m-1}^n) + \frac{\mu^2}{2}(u_{m+1}^n - 2u_m^n + u_{m-1}^n)\,. \tag{2}$$

Determine conditions on μ such that the scheme (2) is stable and convergent. What is the order of this scheme?

Paper 3, Section II

40D Numerical Analysis

Consider the linear system

$$Ax = b\,, \tag{1}$$

where $A \in \mathbb{R}^{n\times n}$ and $b, x \in \mathbb{R}^n$.

(i) Define the *Jacobi iteration method with relaxation parameter* ω for solving (1).

(ii) Assume that A is a symmetric positive-definite matrix whose diagonal part D is such that the matrix $2D - A$ is also positive definite. Prove that the relaxed Jacobi iteration method always converges if the relaxation parameter ω is equal to 1.

(iii) Let A be the tridiagonal matrix with diagonal elements $a_{ii} = \alpha$ and off-diagonal elements $a_{i+1,i} = a_{i,i+1} = \beta$, where $0 < \beta < \frac{1}{2}\alpha$. For which values of ω (expressed in terms of α and β) does the relaxed Jacobi iteration method converge? What choice of ω gives the optimal convergence speed?

[You may quote without proof any relevant results about the convergence of iterative methods and about the eigenvalues of matrices.]

Paper 1, Section II

40D Numerical Analysis

(i) Consider the numerical approximation of the boundary-value problem

$$u'' = f\,, \qquad u : [0,1] \to \mathbb{R}\,,$$

$$u(0) = \varphi_0\,, \quad u(1) = \varphi_1\,,$$

where φ_0, φ_1 are given constants and f is a given smooth function on $[0,1]$. A grid $\{x_1, x_2, \ldots, x_N\}$, $N \geqslant 3$, on $[0,1]$ is given by

$$x_1 = \alpha_1 h\,, \qquad x_i = x_{i-1} + h \text{ for } i = 2, \ldots, N-1\,, \qquad x_N = 1 - \alpha_2 h\,,$$

where $0 < \alpha_1, \alpha_2 < 1$, $\alpha_1 + \alpha_2 = 1$ and $h = 1/N$. Derive finite-difference approximations for $u''(x_i)$, for $i = 1, \ldots, N$, using at most one neighbouring grid point of x_i on each side. Hence write down a numerical scheme to solve the problem, displaying explicitly the entries of the system matrix A in the resulting system of linear equations $Au = b$, $A \in \mathbb{R}^{N\times N}$, $u, b \in \mathbb{R}^N$. What is the overall order of this numerical scheme? Explain briefly one strategy by which the order could be improved with the same grid.

(ii) Consider the numerical approximation of the boundary-value problem

$$\nabla^2 u = f\,, \qquad u : \Omega \to \mathbb{R}\,,$$

$$u(x) = 0 \;\text{ for all } x \in \partial\Omega\,,$$

where $\Omega \subset \mathbb{R}^2$ is an arbitrary, simply connected bounded domain with smooth boundary $\partial\Omega$, and f is a given smooth function. Define the *9-point formula* used to approximate the Laplacian. Using this formula and an equidistant grid inside Ω, define a numerical scheme for which the system matrix is symmetric and negative definite. Prove that the system matrix of your scheme has these properties for all choices of ordering of the grid points.

Paper 4, Section II

39C Numerical Analysis

Consider the solution of the two-point boundary value problem

$$(2-\sin \pi x)u'' + u = 1, \quad -1 \leqslant x \leqslant 1,$$

with periodic boundary conditions at $x = -1$ and $x = 1$. Construct explicitly the linear algebraic system that arises from the application of a spectral method to the above equation.

The Fourier coefficients of u are defined by

$$\hat{u}_n = \frac{1}{2} \int_{-1}^{1} u(\tau) e^{-i\pi n\tau} \, d\tau.$$

Prove that the computation of the Fourier coefficients for the truncated system with $-N/2+1 \leqslant n \leqslant N/2$ (where N is an even and positive integer, and assuming that $\hat{u}_n = 0$ outside this range of n) reduces to the solution of a tridiagonal system of algebraic equations, which you should specify.

Explain the term *convergence with spectral speed* and justify its validity for the derived approximation of u.

Paper 2, Section II

39C Numerical Analysis

Consider the advection equation $u_t = u_x$ on the unit interval $x \in [0,1]$ and $t \geqslant 0$, where $u = u(x,t)$, subject to the initial condition $u(x,0) = \varphi(x)$ and the boundary condition $u(1,t) = 0$, where φ is a given smooth function on $[0,1]$.

(i) We commence by discretising the advection equation above with finite differences on the equidistant space-time grid $\{(m\Delta x, n\Delta t),\ m = 0,\dots,M+1,\ n = 0,\dots,T\}$ with $\Delta x = 1/(M+1)$ and $\Delta t > 0$. We obtain an equation for $u_m^n \approx u(m\Delta x, n\Delta t)$ that reads

$$u_m^{n+1} = u_m^n + \frac{1}{2}\mu(u_{m+1}^n - u_{m-1}^n), \quad m = 1,\dots,M,\ n \in \mathbb{Z}^+,$$

with the condition $u_0^n = 0$ for all $n \in \mathbb{Z}^+$ and $\mu = \Delta t/\Delta x$.

What is the order of approximation (that is, the order of the local error) in space and time of the above discrete solution to the exact solution of the advection equation? Write the scheme in matrix form and deduce for which choices of μ this approximation converges to the exact solution. State (without proof) any theorems you use. [You may use the fact that for a tridiagonal $M \times M$ matrix

$$\begin{pmatrix} \alpha & \beta & 0 & 0 \\ -\beta & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \beta \\ 0 & 0 & -\beta & \alpha \end{pmatrix}$$

the eigenvalues are given by $\lambda_\ell = \alpha + 2i\beta\cos\frac{\ell\pi}{M+1}$.]

(ii) How does the order change when we replace the central difference approximation of the first derivative in space by forward differences, that is $u_{m+1}^n - u_m^n$ instead of $(u_{m+1}^n - u_{m-1}^n)/2$? For which choices of μ is this new scheme convergent?

(iii) Instead of the approximation in (i) we consider the following method for numerically solving the advection equation,

$$u_m^{n+1} = \mu(u_{m+1}^n - u_{m-1}^n) + u_m^{n-1},$$

where we additionally assume that u_m^1 is given. What is the order of this method for a fixed μ?

Paper 3, Section II

40C Numerical Analysis

(i) Suppose that A is a real $n \times n$ matrix, and that $\mathbf{w} \in \mathbb{R}^n$ and $\lambda_1 \in \mathbb{R}$ are given so that $A\mathbf{w} = \lambda_1 \mathbf{w}$. Further, let S be a non-singular matrix such that $S\mathbf{w} = c\mathbf{e}_1$, where $\mathbf{e}_1$ is the first coordinate vector and $c \neq 0$. Let $\widehat{A} = SAS^{-1}$. Prove that the eigenvalues of A are λ_1 together with the eigenvalues of the bottom right $(n-1) \times (n-1)$ submatrix of $\widehat{A}$.

(ii) Suppose again that A is a real $n \times n$ matrix, and that two linearly independent vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ are given such that the linear subspace $\mathcal{L}\{\mathbf{v}, \mathbf{w}\}$ spanned by $\mathbf{v}$ and $\mathbf{w}$ is invariant under the action of A, that is

$$x \in \mathcal{L}\{\mathbf{v}, \mathbf{w}\} \quad \Rightarrow \quad Ax \in \mathcal{L}\{\mathbf{v}, \mathbf{w}\}.$$

Denote by V an $n \times 2$ matrix whose two columns are the vectors $\mathbf{v}$ and $\mathbf{w}$, and let S be a non-singular matrix such that $R = SV$ is upper triangular, that is

$$R = SV = S \times \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \\ \vdots & \vdots \\ v_n & w_n \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} \\ 0 & r_{22} \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix}.$$

Again, let $\widehat{A} = SAS^{-1}$. Prove that the eigenvalues of A are the eigenvalues of the top left 2×2 submatrix of $\widehat{A}$ together with the eigenvalues of the bottom right $(n-2) \times (n-2)$ submatrix of $\widehat{A}$.

Paper 1, Section II

40C Numerical Analysis

Let

$$A(\alpha) = \begin{pmatrix} 1 & \alpha & \alpha \\ \alpha & 1 & \alpha \\ \alpha & \alpha & 1 \end{pmatrix}, \quad \alpha \in \mathbb{R}.$$

(i) For which values of α is $A(\alpha)$ positive definite?

(ii) Formulate the Gauss–Seidel method for the solution $\mathbf{x} \in \mathbb{R}^3$ of a system

$$A(\alpha)\mathbf{x} = \mathbf{b},$$

with $A(\alpha)$ as defined above and $\mathbf{b} \in \mathbb{R}^3$. Prove that the Gauss–Seidel method converges to the solution of the above system whenever A is positive definite. [You may state and use the Householder–John theorem without proof.]

(iii) For which values of α does the Jacobi iteration applied to the solution of the above system converge?

Paper 4, Section II
39D Numerical Analysis

(i) Formulate the conjugate gradient method for the solution of a system $A\mathbf{x} = \mathbf{b}$ with $A \in \mathbb{R}^{n\times n}$ and $\mathbf{b} \in \mathbb{R}^n$, $n > 0$.

(ii) Prove that if the conjugate gradient method is applied in exact arithmetic then, for any $\mathbf{x}^{(0)} \in \mathbb{R}^n$, termination occurs after at most n iterations.

(iii) The polynomial $p(x) = x^m + \sum_{i=0}^{m-1} c_i x^i$ is the *minimal polynomial* of the $n \times n$ matrix A if it is the polynomial of lowest degree that satisfies $p(A) = 0$. [*Note that* $m \leqslant n$.] Give an example of a 3×3 symmetric positive definite matrix with a quadratic minimal polynomial.

Prove that (in exact arithmetic) the conjugate gradient method requires at most m iterations to calculate the exact solution of $A\mathbf{x} = \mathbf{b}$, where m is the degree of the minimal polynomial of A.

Paper 2, Section II
39D Numerical Analysis

(i) The diffusion equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad 0 \leqslant x \leqslant 1,\ t \geqslant 0\,,$$

with the initial condition $u(x,0) = \phi(x)$, $0 \leqslant x \leqslant 1$, and with zero boundary conditions at $x = 0$ and $x = 1$, can be solved numerically by the method

$$u_m^{n+1} = u_m^n + \mu(u_{m-1}^n - 2u_m^n + u_{m+1}^n), \quad m = 1, 2, \ldots, M,\ n \geqslant 0\,,$$

where $\Delta x = 1/(M+1)$, $\mu = \Delta t/(\Delta x)^2$, and $u_m^n \approx u(m\Delta x, n\Delta t)$. Prove that $\mu \leqslant 1/2$ implies convergence.

(ii) By discretising the diffusion equation and employing the same notation as in (i) above, determine [**without** using Fourier analysis] conditions on μ and the constant α such that the method

$$u_m^{n+1} - \frac{1}{2}(\mu - \alpha)(u_{m-1}^{n+1} - 2u_m^{n+1} + u_{m+1}^{n+1}) = u_m^n + \frac{1}{2}(\mu + \alpha)(u_{m-1}^n - 2u_m^n + u_{m+1}^n)$$

is stable.

Paper 3, Section II

40D Numerical Analysis

The inverse discrete Fourier transform $\mathcal{F}_n^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ is given by the formula

$$\mathbf{x} = \mathcal{F}_n^{-1}\mathbf{y}, \quad \text{where} \quad x_l = \sum_{j=0}^{n-1} \omega_n^{jl} y_j, \quad l = 0, \ldots, n-1\,.$$

Here, $\omega_n = \exp(2\pi i/n)$ is the primitive root of unity of degree n and $n = 2^p$, $p = 1, 2, \ldots$

(i) Show how to assemble $\mathbf{x} = \mathcal{F}_{2m}^{-1}\mathbf{y}$ in a small number of operations if the Fourier transforms of the even and odd parts of $\mathbf{y}$,

$$\mathbf{x}^{(E)} = \mathcal{F}_m^{-1}\mathbf{y}^{(E)}, \quad \mathbf{x}^{(O)} = \mathcal{F}_m^{-1}\mathbf{y}^{(O)}\,,$$

are already known.

(ii) Describe the Fast Fourier Transform (FFT) method for evaluating $\mathbf{x}$, and draw a relevant diagram for $n = 8$.

(iii) Find the costs of the FFT method for $n = 2^p$ (only multiplications count).

(iv) For $n = 4$ use the FFT method to find $\mathbf{x} = \mathcal{F}_4^{-1}\mathbf{y}$ when:

(a) $\mathbf{y} = (1, 1, -1, -1)$,
(b) $\mathbf{y} = (1, -1, 1, -1)$.

Paper 1, Section II

40D Numerical Analysis

The Poisson equation $u_{xx} = f$ in the unit interval $\Omega = [0,1]$, $u = 0$ on $\partial\Omega$ is discretised with the formula

$$u_{i-1} + u_{i+1} - 2u_i = h^2 f_i \,,$$

where $1 \leqslant i \leqslant n$, $u_i \approx u(ih)$ and ih are the grid points.

(i) Define the above system of equations in vector form $A\mathbf{u} = \mathbf{b}$ and describe the relaxed Jacobi method with relaxation parameter ω for solving this linear system. For $\mathbf{x}^*$ and $\mathbf{x}^{(\nu)}$ being the exact solution and the iterated solution respectively, let $\mathbf{e}^{(\nu)} = \mathbf{x}^{(\nu)} - \mathbf{x}^*$ be the error and H_ω the iteration matrix, so that

$$\mathbf{e}^{(\nu+1)} = H_\omega \mathbf{e}^{(\nu)} \,.$$

Express H_ω in terms of the matrix A, the diagonal part D of A and ω, and find the eigenvectors $\mathbf{v}_k$ and the eigenvalues $\lambda_k(\omega)$ of H_ω.

(ii) For A as above, let

$$\mathbf{e}^{(\nu)} = \sum_{k=1}^{n} a_k^{(\nu)} \mathbf{v}_k$$

be the expansion of the error with respect to the eigenvectors of H_ω. Derive conditions on ω such that the method converges for any n, and prove that, for any such ω, the rate of convergence of $\mathbf{e}^{(\nu)} \to 0$ is not faster than $(1 - c/n^2)^\nu$.

(iii) Show that, for some ω, the high frequency components ($\frac{n+1}{2} \leqslant k \leqslant n$) of the error $\mathbf{e}^{(\nu)}$ tend to zero much faster than $(1 - c/n^2)^\nu$. Determine the optimal parameter ω_* which provides the largest suppression of the high frequency components per iteration, and find the corresponding attenuation factor μ_* (i.e., the least μ_ω such that $|a_k^{(\nu+1)}| \leqslant \mu_\omega |a_k^{(\nu)}|$ for $\frac{n+1}{2} \leqslant k \leqslant n$).

Paper 1, Section II

40A Numerical Analysis

The nine-point method for the Poisson equation $\nabla^2 u = f$ (with zero Dirichlet boundary conditions) in a square, reads

$$\frac{2}{3}\left(u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1}\right) + \frac{1}{6}\left(u_{i-1,j-1} + u_{i-1,j+1} + u_{i+1,j-1} + u_{i+1,j+1}\right) - \frac{10}{3}u_{i,j} = h^2 f_{i,j}, \quad i,j = 1,\ldots,m,$$

where $u_{0,j} = u_{m+1,j} = u_{i,0} = u_{i,m+1} = 0$, for all $i,j = 0,\ldots,m+1$.

(i) By arranging the two-dimensional arrays $\{u_{i,j}\}_{i,j=1,\ldots,m}$ and $\{f_{i,j}\}_{i,j=1,\ldots,m}$ into column vectors $u \in \mathbb{R}^{m^2}$ and $b \in \mathbb{R}^{m^2}$ respectively, the linear system above takes the matrix form $Au = b$. Prove that, regardless of the ordering of the points on the grid, the matrix A is symmetric and negative definite.

(ii) Formulate the Jacobi method with relaxation for solving the above linear system.

(iii) Prove that the iteration converges if the relaxation parameter ω is equal to 1.

[You may quote without proof any relevant result about convergence of iterative methods.]

Paper 2, Section II

39A Numerical Analysis

Let $A \in \mathbb{R}^{n\times n}$ be a real matrix with n linearly independent eigenvectors. The eigenvalues of A can be calculated from the sequence $x^{(k)}$, $k = 0, 1, \ldots$, which is generated by the power method

$$x^{(k+1)} = \frac{Ax^{(k)}}{\|Ax^{(k)}\|},$$

where $x^{(0)}$ is a real nonzero vector.

(i) Describe the asymptotic properties of the sequence $x^{(k)}$ in the case that the eigenvalues λ_i of A satisfy $|\lambda_i| < |\lambda_n|$, $i = 1, \ldots, n-1$, and the eigenvectors are of unit length.

(ii) Present the implementation details for the power method for the setting in (i) and define the Rayleigh quotient.

(iii) Let A be the 3×3 matrix

$$A = \lambda I + P, \quad P = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

where λ is real and nonzero. Find an explicit expression for A^k, $k = 1, 2, 3, \ldots$.

Let the sequence $x^{(k)}$ be generated by the power method as above. Deduce from your expression for A^k that the first and second components of $x^{(k+1)}$ tend to zero as $k \to \infty$. Further show that this implies $Ax^{(k+1)} - \lambda x^{(k+1)} \to 0$ as $k \to \infty$.

Paper 3, Section II
39A Numerical Analysis

(i) The difference equation

$$u_i^{n+1} = u_i^n + \frac{3}{2}\mu\left(u_{i-1}^n - 2u_i^n + u_{i+1}^n\right) - \frac{1}{2}\mu\left(u_{i-1}^{n-1} - 2u_i^{n-1} + u_{i+1}^{n-1}\right),$$

where $\mu = \Delta t/(\Delta x)^2$, is the basic equation used in the second-order Adams–Bashforth method and can be employed to approximate a solution of the diffusion equation $u_t = u_{xx}$. Prove that, as $\Delta t \to 0$ with constant μ, the local error of the method is $O(\Delta t)^2$.

(ii) By applying the Fourier stability test, show that the above method is stable if and only if $\mu \leqslant 1/4$.

(iii) Define the leapfrog scheme to approximate the diffusion equation and prove that it is unstable for every choice of $\mu > 0$.

Paper 4, Section II
39A Numerical Analysis

(i) Consider the Poisson equation

$$\nabla^2 u = f\,, \quad -1 \leqslant x, y \leqslant 1\,,$$

with the periodic boundary conditions

$$\begin{aligned} u(-1,y) &= u(1,y)\,, \quad u_x(-1,y) = u_x(1,y)\,, \quad -1 \leqslant y \leqslant 1\,, \\ u(x,-1) &= u(x,1)\,, \quad u_y(x,-1) = u_y(x,1)\,, \quad -1 \leqslant x \leqslant 1 \end{aligned}$$

and the normalization condition

$$\int_{-1}^{1}\int_{-1}^{1} u(x,y)\; dx\; dy = 0\,.$$

Moreover, f is analytic and obeys the periodic boundary conditions $f(-1,y) = f(1,y)$, $f(x,-1) = f(x,1)$, $-1 \leqslant x, y \leqslant 1$.

Derive an explicit expression of the approximation of a solution u by means of a spectral method. Explain the term *convergence with spectral speed* and state its validity for the approximation of u.

(ii) Consider the second-order linear elliptic partial differential equation

$$\nabla \cdot (a \nabla u) = f\,, \quad -1 \leqslant x, y \leqslant 1\,,$$

with the periodic boundary conditions and normalization condition specified in (i). Moreover, a and f are given by

$$a(x,y) = \cos(\pi x) + \cos(\pi y) + 3\,, \quad f(x,y) = \sin(\pi x) + \sin(\pi y)\,.$$

[Note that a is a positive analytic periodic function.]

Construct explicitly the linear algebraic system that arises from the implementation of a spectral method to the above equation.

Paper 1, Section II

39A Numerical Analysis

(a) State the Householder–John theorem and explain its relation to the convergence analysis of splitting methods for solving a system of linear equations $Ax = b$ with a positive definite matrix A.

(b) Describe the Jacobi method for solving a system $Ax = b$, and deduce from the above theorem that if A is a symmetric positive definite *tridiagonal* matrix,

$$A = \begin{bmatrix} a_1 & c_1 & & & \\ c_1 & a_2 & c_2 & & \mathbf{0} \\ & \ddots & \ddots & \ddots & \\ \mathbf{0} & & c_{n-2} & a_{n-1} & c_{n-1} \\ & & & c_{n-1} & a_n \end{bmatrix},$$

then the Jacobi method converges.

[*Hint: At the last step, you may find it useful to consider two vectors* $x = (x_1, x_2, \ldots, x_n)$ *and* $y = ((-1)x_1, (-1)^2 x_2, \ldots, (-1)^n x_n)$.]

Paper 2, Section II

39A Numerical Analysis

The inverse discrete Fourier transform $\mathcal{F}_n^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ is given by the formula

$$\mathbf{x} = \mathcal{F}_n^{-1}\,\mathbf{y}\,, \qquad \text{where} \qquad x_\ell = \sum_{j=0}^{n-1} \omega_n^{j\ell}\, y_j\,, \qquad \ell = 0, \ldots, n-1\,.$$

Here, $\omega_n = \exp(2\pi i/n)$ is the primitive root of unity of degree n, and $n = 2^p$, $p = 1, 2, \ldots$.

(1) Show how to assemble $\mathbf{x} = \mathcal{F}_{2m}^{-1}\,\mathbf{y}$ in a small number of operations if we already know the Fourier transforms of the even and odd portions of $\mathbf{y}$:

$$\mathbf{x}^{(\mathrm{E})} = \mathcal{F}_m^{-1}\,\mathbf{y}^{(\mathrm{E})}\,, \qquad \mathbf{x}^{(\mathrm{O})} = \mathcal{F}_m^{-1}\,\mathbf{y}^{(\mathrm{O})}\,.$$

(2) Describe the Fast Fourier Transform (FFT) method for evaluating $\mathbf{x}$ and draw a relevant diagram for $n = 8$.

(3) Find the costs of the FFT for $n = 2^p$ (only multiplications count).

(4) For $n = 4$, using the FFT technique, find

$$\mathbf{x} = \mathcal{F}_4^{-1}\,\mathbf{y}\,, \qquad \text{for} \qquad \mathbf{y} = [1, 1, -1, -1]\,, \qquad \text{and} \qquad \mathbf{y} = [1, -1, 1, -1]\,.$$

Paper 3, Section II

39A Numerical Analysis

The Poisson equation $\nabla^2 u = f$ in the unit square $\Omega = [0,1] \times [0,1]$, $u = 0$ on $\partial\Omega$, is discretized with the five-point formula

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

where $1 \leqslant i,\ j \leqslant M$, $u_{i,j} \approx u(ih, jh)$ and (ih, jh) are grid points.

Let $u(x,y)$ be the exact solution, and let $e_{i,j} = u_{i,j} - u(ih, jh)$ be the error of the five-point formula at the (i,j)th grid point. Justifying each step, prove that

$$\|\mathbf{e}\| = \left[\sum_{i,j=1}^{M} |\,e_{i,j}|^2\right]^{1/2} \leqslant ch \quad \text{for sufficiently small} \quad h > 0\,,$$

where c is some constant independent of h.

Paper 4, Section II

39A Numerical Analysis

An s-stage explicit Runge–Kutta method of order p, with constant step size $h > 0$, is applied to the differential equation $y' = \lambda y,\ t \geqslant 0$.

(a) Prove that

$$y_{n+1} = P_s(\lambda h)\, y_n\,.$$

where P_s is a polynomial of degree s.

(b) Prove that the order p of any s-stage explicit Runge–Kutta method satisfies the inequality $p \leqslant s$ and, for $p = s$, write down an explicit expression for P_s.

(c) Prove that no explicit Runge–Kutta method can be A-stable.

Paper 1, Section II

39B Numerical Analysis

(i) Define the Jacobi method with relaxation for solving the linear system $A\boldsymbol{x} = \boldsymbol{b}$.

(ii) For $\boldsymbol{x}^*$ and $\boldsymbol{x}^{(\nu)}$ being the exact and the iterated solution, respectively, let $\boldsymbol{e}^{(\nu)} := \boldsymbol{x}^{(\nu)} - \boldsymbol{x}^*$ be the error and H_ω the iteration matrix, so that

$$\mathbf{e}^{(\nu+1)} = H_\omega \mathbf{e}^{(\nu)}\,.$$

Express H_ω in terms of the matrix A, its diagonal part D and the relaxation parameter ω, and find the eigenvectors $\boldsymbol{v}_k$ and the eigenvalues $\lambda_k(\omega)$ of H_ω for the $n \times n$ tridiagonal matrix

$$A = \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix}.$$

[*Hint: The eigenvectors and eigenvalues of A are*

$$(\boldsymbol{u}_k)_i = \sin\tfrac{ki\pi}{n+1}, \quad i = 1, \ldots, n, \qquad \lambda_k(A) = 4\sin^2\tfrac{k\pi}{2(n+1)}, \qquad k = 1, \ldots, n\,.]$$

(iii) For A as above, let

$$\boldsymbol{e}^{(\nu)} = \sum_{k=1}^{n} a_k^{(\nu)} \boldsymbol{v}_k$$

be the expansion of the error with respect to the eigenvectors $(\boldsymbol{v}_k)$ of H_ω.

Find the range of parameter ω which provides convergence of the method for any n, and prove that, for any such ω, the rate of convergence $\boldsymbol{e}^{(\nu)} \to 0$ is not faster than $(1 - c/n^2)^\nu$.

(iv) Show that, for some ω, the high frequency components ($\frac{n+1}{2} \leqslant k \leqslant n$) of the error $\boldsymbol{e}^{(\nu)}$ tend to zero much faster. Determine the optimal parameter ω_* which provides the largest suppression of the high frequency components per iteration, and find the corresponding attenuation factor μ_* (i.e. the least μ_ω such that $|a_k^{(\nu+1)}| \leqslant \mu_\omega |a_k^{(\nu)}|$ for $\frac{n+1}{2} \leqslant k \leqslant n$).

Paper 2, Section II

39B Numerical Analysis

The Poisson equation $\nabla^2 u = f$ in the unit square $\Omega = [0,1] \times [0,1]$, equipped with appropriate boundary conditions on $\partial\Omega$, is discretized with the nine-point formula:

$$\Gamma_9[u_{m,n}] := -\frac{10}{3}\,u_{m,n} + \frac{2}{3}\,(u_{m+1,n} + u_{m-1,n} + u_{m,n+1} + u_{m,n-1}) \\ +\frac{1}{6}\,(u_{m+1,n+1} + u_{m+1,n-1} + u_{m-1,n+1} + u_{m-1,n-1}) = h^2 f_{m,n},$$

where $1 \leqslant m, n \leqslant M$, $u_{m,n} \approx u(mh, nh)$, and (mh, nh) are grid points.

(i) Find the local error of approximation.

(ii) Prove that the error is smaller if f happens to satisfy the Laplace equation $\nabla^2 f = 0$.

(iii) Hence show that the modified nine-point scheme

$$\begin{aligned} \Gamma_9[u_{m,n}] &= h^2 f_{m,n} + \frac{1}{12} h^2 \Gamma_5[f_{m,n}] \\ &:= h^2 f_{m,n} + \frac{1}{12} h^2 (f_{m+1,n} + f_{m-1,n} + f_{m,n+1} + f_{m,n-1} - 4 f_{m,n}) \end{aligned}$$

has the same smaller error as in (ii).

[*Hint. The nine-point discretization of* $\nabla^2 u$ *can be written as*

$$\Gamma_9[u] = (\Gamma_5 + \tfrac{1}{6}\Delta_x^2\Delta_y^2)\,u = (\Delta_x^2 + \Delta_y^2 + \tfrac{1}{6}\Delta_x^2\Delta_y^2)\,u$$

where $\Gamma_5[u] = (\Delta_x^2 + \Delta_y^2)u$ *is the five-point discretization and*

$$\begin{aligned} \Delta_x^2\,u(x,y) &:= u(x-h,y) - 2u(x,y) + u(x+h,y), \\ \Delta_y^2\,u(x,y) &:= u(x,y-h) - 2u(x,y) + u(x,y+h)\,.\,] \end{aligned}$$

Paper 3, Section II

39B Numerical Analysis

Prove that all Toeplitz tridiagonal $M \times M$ matrices A of the form

$$A = \begin{bmatrix} a & b & & & \\ -b & a & b & & \\ & \ddots & \ddots & \ddots & \\ & & -b & a & b \\ & & & -b & a \end{bmatrix}$$

share the same eigenvectors $(\boldsymbol{v}^{(k)})_{k=1}^{M}$, with the components $\boldsymbol{v}_m^{(k)} = i^m \sin \frac{km\pi}{M+1}$, $m = 1, \ldots, M$, where $i = \sqrt{-1}$, and find their eigenvalues.

The advection equation

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x}, \quad 0 \leqslant x \leqslant 1, \quad 0 \leqslant t \leqslant T,$$

is approximated by the Crank–Nicolson scheme

$$u_m^{n+1} - u_m^n = \frac{1}{4}\mu \left(u_{m+1}^{n+1} - u_{m-1}^{n+1}\right) + \frac{1}{4}\mu \left(u_{m+1}^{n} - u_{m-1}^{n}\right),$$

where $\mu = \frac{\Delta t}{(\Delta x)^2}$, $\Delta x = \frac{1}{M+1}$, and u_m^n is an approximation to $u(m\Delta x, n\Delta t)$. Assuming that $u(0,t) = u(1,t) = 0$, show that the above scheme can be written in the form

$$B\boldsymbol{u}^{n+1} = C\boldsymbol{u}^n, \quad 0 \leqslant n \leqslant T/\Delta t - 1,$$

where $\boldsymbol{u}^n = [u_1^n, \ldots, u_M^n]^T$ and the real matrices B and C should be found. Using matrix analysis, find the range of μ for which the scheme is stable. [Fourier analysis is not acceptable.]

Paper 4, Section II

39B Numerical Analysis

(a) For the s-step s-order Backward Differentiation Formula (BDF) for ordinary differential equations,

$$\sum_{m=0}^{s} a_m y_{n+m} = h f_{n+s},$$

express the polynomial $\rho(w) = \sum_{m=0}^{s} a_m w^m$ in a convenient explicit form.

(b) Prove that the interval $(-\infty, 0)$ belongs to the linear stability domain of the 2-step BDF method.

1/II/38C **Numerical Analysis**

The Poisson equation $\nabla^2 u = f$ in the unit square $\Omega = [0,1] \times [0,1]$, with zero boundary conditions on $\partial\Omega$, is discretized with the nine-point formula

$$\frac{10}{3} u_{m,n} - \frac{2}{3}\left(u_{m+1,n} + u_{m-1,n} + u_{m,n+1} + u_{m,n-1}\right)$$
$$- \frac{1}{6}\left(u_{m+1,n+1} + u_{m+1,n-1} + u_{m-1,n+1} + u_{m-1,n-1}\right) = -h^2 f_{m,n},$$

where $1 \leqslant m, n \leqslant M$, $u_{m,n} \approx u(mh, nh)$, and (mh, nh) are grid points.

(a) Prove that, for any ordering of the grid points, the method can be written as $A\mathbf{u} = \mathbf{b}$ with a symmetric positive-definite matrix A.

(b) Describe the Jacobi method for solving a linear system of equations, and prove that it converges for the above system.

[You may quote without proof the corollary of the Householder–John theorem regarding convergence of the Jacobi method.]

2/II/38C **Numerical Analysis**

The advection equation

$$u_t = u_x, \qquad x \in \mathbb{R}, \quad t \geqslant 0,$$

is solved by the leapfrog scheme

$$u_m^{n+1} = \mu\left(u_{m+1}^n - u_{m-1}^n\right) + u_m^{n-1},$$

where $n \geqslant 1$ and $\mu = \Delta t / \Delta x$ is the Courant number.

(a) Determine the local error of the method.

(b) Applying the Fourier technique, find the range of $\mu > 0$ for which the method is stable.

3/II/38C **Numerical Analysis**

(a) A numerical method for solving the ordinary differential equation

$$y'(t) = f(t, y), \qquad t \in [0, T], \qquad y(0) = y_0,$$

generates for every $h > 0$ a sequence $\{y_n\}$, where y_n is an approximation to $y(t_n)$ and $t_n = nh$. Explain what is meant by the convergence of the method.

(b) Prove from first principles that if the function f is sufficiently smooth and satisfies the Lipschitz condition

$$|f(t, x) - f(t, y)| \leqslant \lambda |x - y|, \qquad x, y \in \mathbb{R}, \qquad t \in [0, T],$$

for some $\lambda > 0$, then the trapezoidal rule

$$y_{n+1} = y_n + \frac{1}{2} h \left[f(t_n, y_n) + f(t_{n+1}, y_{n+1}) \right]$$

converges.

4/II/39C **Numerical Analysis**

Let $A \in \mathbb{R}^{n \times n}$ be a real matrix with n linearly independent eigenvectors. When calculating eigenvalues of A, the sequence $\mathbf{x}^{(k)}$, $k = 0, 1, 2, \ldots$, is generated by the power method $\mathbf{x}^{(k+1)} = A\mathbf{x}^{(k)} / \|A\mathbf{x}^{(k)}\|$, where $\mathbf{x}^{(0)}$ is a real nonzero vector.

(a) Describe the asymptotic properties of the sequence $\mathbf{x}^{(k)}$, both in the case where the eigenvalues λ_i of A satisfy $|\lambda_i| < |\lambda_n|$, $i = 1, \ldots, n-1$, and in the case where $|\lambda_i| < |\lambda_{n-1}| = |\lambda_n|$, $i = 1, \ldots, n-2$. In the latter case explain how the (possibly complex-valued) eigenvalues λ_{n-1}, λ_n and their corresponding eigenvectors can be determined.

(b) Let $n = 3$, and suppose that, for a large k, we obtain the vectors

$$\mathbf{y}_k = \mathbf{x}_k = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{y}_{k+1} = A\mathbf{x}_k = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}, \quad \mathbf{y}_{k+2} = A^2\mathbf{x}_k = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}.$$

Find two eigenvalues of A and their corresponding eigenvectors.

1/II/38C Numerical Analysis

(a) For a numerical method to solve $y' = f(t, y)$, define the linear stability domain and state when such a method is A-stable.

(b) Determine all values of the real parameter a for which the Runge–Kutta method

$$\begin{aligned} k_1 &= f\Big(t_n + (\tfrac{1}{2} - a)h,\ y_n + h\big[\tfrac{1}{4}k_1 + (\tfrac{1}{4} - a)k_2\big]\Big), \\ k_2 &= f\Big(t_n + (\tfrac{1}{2} + a)h,\ y_n + h\big[(\tfrac{1}{4} + a)k_1 + \tfrac{1}{4}k_2\big]\Big), \\ y_{n+1} &= y_n + \tfrac{1}{2}h(k_1 + k_2) \end{aligned}$$

is A-stable.

2/II/38C Numerical Analysis

(a) State the Householder–John theorem and explain how it can be used to design iterative methods for solving a system of linear equations $Ax = b$.

(b) Let $A = L + D + U$ where D is the diagonal part of A, and L and U are, respectively, the strictly lower and strictly upper triangular parts of A. Given a vector b, consider the following iterative scheme:

$$(D + \omega L)x^{(k+1)} = (1 - \omega)Dx^{(k)} - \omega U x^{(k)} + \omega b\,.$$

Prove that if A is a symmetric positive definite matrix, and $\omega \in (0, 2)$, then the above iteration converges to the solution of the system $Ax = b$.

3/II/38C **Numerical Analysis**

(a) Prove that all Toeplitz symmetric tridiagonal $M \times M$ matrices

$$A = \begin{bmatrix} a & b & 0 & \cdots & 0 \\ b & a & b & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & b & a & b \\ 0 & \cdots & 0 & b & a \end{bmatrix}$$

share the same eigenvectors $(v^{(k)})_{k=1}^{M}$ with components $v_i^{(k)} = \sin \dfrac{ki\pi}{M+1}$, $i = 1, \ldots, M$, and eigenvalues to be determined.

(b) The diffusion equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad 0 \leqslant x \leqslant 1, \quad 0 \leqslant t \leqslant T,$$

is approximated by the Crank–Nicolson scheme

$$u_m^{n+1} - \tfrac{1}{2}\mu \left(u_{m-1}^{n+1} - 2u_m^{n+1} + u_{m+1}^{n+1} \right) = u_m^n + \tfrac{1}{2}\mu \left(u_{m-1}^n - 2u_m^n + u_{m+1}^n \right) \ ,$$
$$\text{for} \qquad m = 1, \ldots, M,$$

where $\mu = \Delta t/(\Delta x)^2$, $\Delta x = 1/(M+1)$, and u_m^n is an approximation to $u(m\Delta x, n\Delta t)$. Assuming that $u(0,t) = u(1,t) = 0$, $\forall t$, show that the above scheme can be written in the form

$$Bu^{n+1} = Cu^n, \qquad 0 \leqslant n \leqslant (T/\Delta t) - 1$$

where $u^n = [u_1^n, \ldots, u_M^n]^{\mathsf{T}}$ and the real matrices B and C should be found. Using matrix analysis, find the range of μ for which the scheme is stable. [*Do not use Fourier analysis.*]

4/II/39C **Numerical Analysis**

(a) Suppose that A is a real $n \times n$ matrix, and that $w \in \mathbb{R}^n$ and $\lambda_1 \in \mathbb{R}$ are given so that $Aw = \lambda_1 w$. Further, let S be a non-singular matrix such that $Sw = ce^{(1)}$, where $e^{(1)}$ is the first coordinate vector and $c \neq 0$. Let $\widehat{A} = SAS^{-1}$. Prove that the eigenvalues of A are λ_1 together with the eigenvalues of the bottom right $(n-1) \times (n-1)$ submatrix of $\widehat{A}$.

(b) Suppose again that A is a real $n \times n$ matrix, and that two linearly independent vectors $v, w \in \mathbb{R}^n$ are given such that the linear subspace $L\{v, w\}$ spanned by v and w is invariant under the action of A, i.e.,

$$x \in L\{v, w\} \quad \Rightarrow \quad Ax \in L\{v, w\}.$$

Denote by V an $n \times 2$ matrix whose two columns are the vectors v and w, and let S be a non-singular matrix such that $R = SV$ is upper triangular, that is,

$$R = SV = S \times \begin{bmatrix} v_1 & w_1 \\ v_2 & w_2 \\ v_3 & w_3 \\ \vdots & \vdots \\ v_n & w_n \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} \\ 0 & r_{22} \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix}.$$

Again let $\widehat{A} = SAS^{-1}$. Prove that the eigenvalues of A are the eigenvalues of the top left 2×2 submatrix of $\widehat{A}$ together with the eigenvalues of the bottom right $(n-2) \times (n-2)$ submatrix of $\widehat{A}$.

1/II/38C Numerical Analysis

(a) Define the Jacobi method with relaxation for solving the linear system $Ax = b$.

(b) Let A be a symmetric positive definite matrix with diagonal part D such that the matrix $2D - A$ is also positive definite. Prove that the iteration always converges if the relaxation parameter ω is equal to 1.

(c) Let A be the tridiagonal matrix with diagonal elements $a_{ii} = 1$ and off-diagonal elements $a_{i+1,i} = a_{i,i+1} = 1/4$. Prove that convergence occurs if ω satisfies $0 < \omega \leqslant 4/3$. Explain briefly why the choice $\omega = 1$ is optimal.

[*You may quote without proof any relevant result about the convergence of iterative methods and about the eigenvalues of matrices.*]

2/II/38C Numerical Analysis

In the unit square the Poisson equation $\nabla^2 u = f$, with zero Dirichlet boundary conditions, is being solved by the five-point formula using a square grid of mesh size $h = 1/(M+1)$,

$$u_{i,j-1} + u_{i,j+1} + u_{i-1,j} + u_{i+1,j} - 4u_{i,j} = h^2 f_{i,j} \ .$$

Let $u(x,y)$ be the exact solution, and let $e_{i,j} = u_{i,j} - u(ih, jh)$ be the error of the five-point formula at the (i,j)th grid point. Justifying each step, prove that

$$\left[\sum_{i,j=1}^{M} |e_{i,j}|^2 \right]^{1/2} \leqslant ch, \quad h \to 0 \ ,$$

where c is some constant.

3/II/38C **Numerical Analysis**

(a) For the equation $y' = f(t, y)$, consider the following multistep method with s steps,

$$\sum_{i=0}^{s} \rho_i y_{n+i} = h \sum_{i=0}^{s} \sigma_i f(t_{n+i},\, y_{n+i}) \ ,$$

where h is the step size and ρ_i, σ_i are specified constants with $\rho_s = 1$. Prove that this method is of order p if and only if

$$\sum_{i=0}^{s} \rho_i Q(t_{n+i}) = h \sum_{i=0}^{s} \sigma_i Q'(t_{n+i})$$

for any polynomial Q of degree $\leqslant p$. Deduce that there is no s-step method of order $2s+1$.

[*You may use the fact that, for any* a_i, b_i*, the Hermite interpolation problem*

$$Q(x_i) = a_i, \quad Q'(x_i) = b_i, \qquad i = 0, \ldots, s$$

is uniquely solvable in the space of polynomials of degree $2s+1$.]

(b) State the Dahlquist equivalence theorem regarding the convergence of a multistep method. Determine all the values of the real parameter $a \neq 0$ for which the multistep method

$$y_{n+3} + (2a-3)[y_{n+2} - y_{n+1}] - y_n = ha\,[f_{n+2} + f_{n+1}]$$

is convergent, and determine the order of convergence.

4/II/39C **Numerical Analysis**

The difference equation

$$u_m^{n+1} = u_m^n + \tfrac{3}{2}\mu\left(u_{m-1}^n - 2u_m^n + u_{m+1}^n\right) - \tfrac{1}{2}\mu\left(u_{m-1}^{n-1} - 2u_m^{n-1} + u_{m+1}^{n-1}\right) ,$$

where $\mu = \Delta t/(\Delta x)^2$, is used to approximate a solution of the diffusion equation $u_t = u_{xx}$.

(a) Prove that, as $\Delta t \to 0$ with constant μ, the local error of the method is $\mathcal{O}(\Delta t)^2$.

(b) Applying the Fourier stability test, show that the method is stable if and only if $\mu \leqslant \frac{1}{4}$.

1/II/38A Numerical Analysis

Let

$$\frac{\mu}{4}u_{m-1}^{n+1} + u_m^{n+1} - \frac{\mu}{4}u_{m+1}^{n+1} = -\frac{\mu}{4}u_{m-1}^{n} + u_m^{n} + \frac{\mu}{4}u_{m+1}^{n},$$

where n is a positive integer and m ranges over all integers, be a finite-difference method for the advection equation

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x}, \qquad -\infty < x < \infty, \quad t \geqslant 0.$$

Here $\mu = \frac{\Delta t}{\Delta x}$ is the Courant number.

(a) Show that the local error of the method is $O((\Delta x)^3)$.

(b) Determine the range of $\mu > 0$ for which the method is stable.

2/II/38A Numerical Analysis

Define a *Krylov subspace* $\mathcal{K}_n(A, v)$.

Let d_n be the dimension of $\mathcal{K}_n(A, v)$. Prove that the sequence $\{d_m\}_{m=1,2,\ldots}$ increases monotonically. Show that, moreover, there exists an integer k with the following property: $d_m = m$ for $m = 1, 2, \ldots, k$, while $d_m = k$ for $m \geqslant k$. Assuming that A has a full set of eigenvectors, show that k is equal to the number of eigenvectors of A required to represent the vector v.

3/II/38A Numerical Analysis

Consider the Runge–Kutta method

$$\begin{aligned} k_1 &= f(y_n), \\ k_2 &= f(y_n + (1-a)hk_1 + ahk_2), \\ y_{n+1} &= y_n + \frac{h}{2}(k_1 + k_2) \end{aligned}$$

for the solution of the scalar ordinary differential equation $y' = f(y)$. Here a is a real parameter.

(a) Determine the order of the method.

(b) Find the range of values of a for which the method is A-stable.

4/II/39A **Numerical Analysis**

An $n \times n$ *skew-symmetric* matrix A is converted into an upper-Hessenberg form B, say, by Householder reflections.

(a) Describe each step of the procedure and observe that B is tridiagonal. Your algorithm should take advantage of the special form of A to reduce the number of computations.

(b) Compare the cost (counting only products and looking only at the leading term) of converting a skew-symmetric and a symmetric matrix to an upper-Hessenberg form using Householder reflections.