

DEGREE OF MASTER OF SCIENCE
MATHEMATICAL MODELLING AND SCIENTIFIC COMPUTING

B1 Numerical Linear Algebra and Numerical Solution
of Differential Equations

HILARY TERM 2020
FRIDAY, 17 JANUARY 2020, 9.30am to 12.00pm

This exam paper contains two sections. You may attempt as many questions as you like but you must answer at least one question in each section. Your best answer in each section will count, along with your next best two answers, making a total of four answers.

*Please start the answer to each question in a new booklet.
All questions will carry equal marks.*

Do not turn this page until you are told that you may do so

Section A: Numerical Solution of Differential Equations

1. Consider the initial-value problem $y' = f(x, y)$, $y(0) = 1$, where f is a real-valued twice continuously differentiable function of its arguments such that $|\frac{\partial f}{\partial y}(x, y)| \leq L$ for all $(x, y) \in \mathbb{R}^2$, where L is a positive real number. Suppose further that the unique solution y of this initial-value problem is a three times continuously differentiable function of x on the interval $[0, 1]$ of the real line. Let N be a positive integer, $h := 1/N$, $x_n := nh$ for $n = 0, 1, \dots, N$, and let y_n be an approximation to $y(x_n)$, $n = 0, 1, \dots, N$, defined successively by the explicit one-step method

$$y_{n+1} := y_n + hf(x_n + \beta h, y_n + \beta hf(x_n, y_n)), \quad n = 0, 1, \dots, N-1, \quad y_0 := 1,$$

where $\beta \in [0, 1]$ is a parameter.

- (a) [5 marks] Show that the method is consistent for any value of β .
 (b) [12 marks] Show that the consistency error T_n of the method can be expressed as

$$T_n = h \left(\frac{1}{2} - \beta \right) y''(x_n) + \mathcal{O}(h^2).$$

Deduce that if $\beta \neq \frac{1}{2}$ then the method is first-order accurate, and that if $\beta = \frac{1}{2}$ then it is at least second-order accurate.

Show further by deriving a bound on the global error of the method in terms of the consistency error that if $\beta = \frac{1}{2}$ then the order of convergence of the method is at least 2.

- (c) [8 marks] Apply the method, with $\beta = \frac{1}{2}$, to the initial-value problem $y' = y$, $y(0) = 1$, and show that

$$y_n = \left(1 + h + \frac{1}{2}h^2 \right)^n, \quad n = 0, 1, \dots, N.$$

By using the identity $a^n - b^n = (a - b) \sum_{k=0}^{n-1} a^{n-k-1} b^k$, where a and b are arbitrary real numbers and n is a positive integer, show that

$$y(x_n) - y_n \leq \frac{1}{3!} h^2 x_n e^{x_n}, \quad n = 0, 1, \dots, N.$$

Show further that

$$y(x_n) - y_n \geq \frac{1}{3!} h^2 x_n, \quad n = 0, 1, \dots, N.$$

Hence deduce that for $\beta = 1/2$ the order of convergence of the method is equal to 2.

2. Consider the ordinary differential equation $y' = f(x, y)$, where f is a real-valued continuous function defined for all $(x, y) \in \mathbb{R}^2$, and let $x_0, y_0 \in \mathbb{R}$.
- (a) [2 marks] State the general form of a linear k -step method for the numerical solution of the initial-value problem $y' = f(x, y)$, $y(x_0) = y_0$ on the mesh $\{x_n : x_n = x_0 + nh, n = 0, 1, \dots\}$ of uniform spacing $h > 0$.
- (b) [6 marks] Define the *consistency error* of a linear k -step method. What is meant by saying that a linear k -step method is *consistent*? What is meant by saying that a linear multistep method is *second-order accurate*?
- (c) [6 marks] What is meant by saying that a linear k -step method is *zero-stable*? Formulate an equivalent characterisation of zero-stability in terms of the roots of a certain polynomial of degree k .
- (d) [11 marks] Consider the three-parameter family of linear two-step methods defined by

$$y_{n+2} - ay_{n+1} + by_n = hc f_{n+2},$$

where $f_j = f(x_j, y_j)$, and a, b and c are real numbers. Show that there exists a unique choice of a, b and c such that the method is second-order accurate; show further that, for these values of a, b and c , the method is second-order convergent.

[If Dahlquist's Theorem is used, it must be stated carefully.]

3. Consider the initial-value problem

$$\frac{\partial u}{\partial t} + u = \frac{\partial^2 u}{\partial x^2}, \quad -\infty < x < \infty, \quad 0 < t \leq T,$$

$$u(x, 0) = u_0(x), \quad -\infty < x < \infty,$$

where T is a fixed real number, and u_0 is a real-valued continuous function of $x \in (-\infty, \infty)$.

- (a) [5 marks] Formulate the θ scheme for the numerical solution of this initial-value problem on a mesh with uniform spacings $\Delta x > 0$ and $\Delta t = T/M$ in the x and t co-ordinate directions, respectively, where M is a positive integer. You should state the scheme so that $\theta = 1$ corresponds to the implicit (backward) Euler scheme.
- (b) [10 marks] Let U_j^m denote the θ -scheme-approximation to $u(j\Delta x, m\Delta t)$, $0 \leq m \leq M$, $j \in \mathbb{Z}$, where \mathbb{Z} denotes the set of all integers. Let $\|U^m\|_{\ell_\infty} := \max_{j \in \mathbb{Z}} |U_j^m|$, and suppose that $\|U^0\|_{\ell_\infty}$ is finite. Show that if $\theta \in [0, 1]$ then

$$\|U^m\|_{\ell_\infty} \leq \left(\frac{1 - (1 - \theta)\Delta t}{1 + \theta\Delta t} \right)^m \|U^0\|_{\ell_\infty}$$

for all m , $1 \leq m \leq M$, provided that $A(\theta)\Delta t \leq \frac{(\Delta x)^2}{2 + (\Delta x)^2}$, where $A(\theta)$ is a constant, depending on the choice of θ , which you should determine.

Deduce that the implicit (backward) Euler scheme is *unconditionally stable* in the $\|\cdot\|_{\ell_\infty}$ norm. Show, further, that the Crank–Nicolson scheme is *conditionally stable* in the $\|\cdot\|_{\ell_\infty}$ norm and state the condition on Δt and Δx that ensures stability.

- (c) [10 marks] Let U_j^m denote the θ -scheme-approximation to $u(j\Delta x, m\Delta t)$, $0 \leq m \leq M$, $j \in \mathbb{Z}$, where \mathbb{Z} denotes the set of all integers. Let $\|U^m\|_{\ell_2} := \left(\Delta x \sum_{j \in \mathbb{Z}} |U_j^m|^2 \right)^{1/2}$ and suppose that $\|U^0\|_{\ell_2}$ is finite. Show that if $\theta \in [\frac{1}{2}, 1]$, then

$$\|U^m\|_{\ell_2} \leq \|U^0\|_{\ell_2}$$

for all m , $1 \leq m \leq M$, for any Δt and Δx .

Now, suppose that $\theta \in [0, \frac{1}{2})$. Show that $\|U^m\|_{\ell_2} \leq \|U^0\|_{\ell_2}$ for all m , $1 \leq m \leq M$, provided that $B(\theta)\Delta t \leq \frac{2(\Delta x)^2}{4 + (\Delta x)^2}$, where $B(\theta)$ is a constant, depending on the choice of θ , which you should determine.

Deduce that the implicit (backward) Euler scheme and the Crank–Nicolson scheme are *unconditionally stable* in the $\|\cdot\|_{\ell_2}$ norm.

4. Consider the finite difference mesh $\mathcal{M} := \{(x_j, t_m) : j = 0, 1, \dots, J, m = 0, 1, \dots, M\}$, where $x_j := j\Delta x$ and $t_m := m\Delta t$, with $\Delta x := 1/J$, $\Delta t := T/M$, $J \geq 2$, $M \geq 1$, and $T > 0$.

(a) [5 marks] Formulate the explicit Euler scheme on \mathcal{M} for the numerical solution of the initial-boundary-value problem

$$u_t = \kappa u_{xx}, \quad x \in (0, 1), \quad t \in (0, T];$$

$$u(0, t) = A(t), \quad u(1, t) = B(t), \quad t \in (0, T]; \quad u(x, 0) = u_0(x), \quad x \in [0, 1],$$

where κ is a positive real number, A, B are continuous real-valued functions defined on $[0, T]$, and u_0 is a continuous real-valued function defined on $[0, 1]$ with $u_0(0) = A(0)$ and $u_0(1) = B(0)$.

(b) [10 marks] Show that U_j^m , the approximation to $u(x_j, t_m)$ computed from the explicit Euler scheme, is bounded above by U_{\max} , where

$$U_{\max} := \max \left\{ \max_{0 \leq m \leq M} A(t_m), \max_{0 \leq m \leq M} B(t_m), \max_{0 \leq j \leq J} u_0(x_j) \right\},$$

provided that a stability condition of the form

$$0 < \mu \leq \mu_0$$

is satisfied with $\mu := \frac{\kappa \Delta t}{(\Delta x)^2}$, where μ_0 is a positive real number, independent of Δt and Δx , which you should determine.

(c) [10 marks] Write down the recurrence relation satisfied by the *global error* at the mesh points, defined by

$$e_j^m := u(x_j, t_m) - U_j^m.$$

Assuming that the initial and boundary conditions for the explicit Euler scheme are exact, and Δt and Δx are such that $0 < \mu \leq \mu_0$, derive a bound on

$$E^m := \max_{0 \leq j \leq J} |e_j^m|, \quad 0 \leq m \leq M,$$

in terms of

$$T^m := \max_{1 \leq j \leq J-1} |T_j^m|, \quad 0 \leq m \leq M,$$

where T_j^m is the *consistency error* of the explicit Euler scheme at the mesh point (x_j, t_m) .

Section B: Numerical Linear Algebra

5. In this question m and n are positive integers, and $A \in \mathbb{R}^{m \times n}$ denotes a generic matrix.

(a) [1 mark] Define the *Frobenius norm* $\|A\|_F$.

(b) [2 marks] Give the definition of an *orthogonal matrix*. What is $\|P\|_F$ when $P \in \mathbb{R}^{k \times k}$ is an orthogonal matrix?

(c) [4 marks] Show that $\|QAZ\|_F = \|A\|_F$ whenever $Q \in \mathbb{R}^{m \times m}$ and $Z \in \mathbb{R}^{n \times n}$ are orthogonal matrices.

(d) [3 marks] What does it mean to say that $A = U\Sigma V^T$ is a *Singular Value Decomposition* of A ? (You may assume, and do not need to prove, that such a decomposition exists). What are the *singular values*? Express $\|A\|_F$ in terms of the singular values of A .

(e) [10 marks] What are the singular values of

$$B = \begin{bmatrix} 2 & 0 \\ 0 & 3 \\ 1 & 0 \end{bmatrix} ?$$

What are the eigenvalues of BB^T ?

Calculate an *LU factorisation* of BB^T and hence or otherwise write down the set of all solutions to

$$BB^T \mathbf{x} = \begin{bmatrix} 2 \\ 9 \\ 1 \end{bmatrix}.$$

(f) [5 marks] A *generalized inverse* of a matrix $A \in \mathbb{R}^{m \times n}$ is a matrix $C \in \mathbb{R}^{n \times m}$ such that $ACA = A$. Identify a candidate for C in terms of the Singular Value Decomposition of A .

6. (a) [4 marks] Calculate the first two iterate vectors \mathbf{x}_1 and \mathbf{x}_2 obtained by applying Jacobi iteration for the linear system

$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

starting from $\mathbf{x}_0 = [0, 0, 0]^T$.

(b) [2 marks] What is *relaxed Jacobi iteration* with relaxation parameter θ for a linear system $A\mathbf{x} = \mathbf{b}$?

(c) [2 marks] State, without proof, a necessary and sufficient condition for the convergence of the sequence of iterates generated from the simple iteration

$$M\mathbf{x}_{k+1} = N\mathbf{x}_k + \mathbf{b}, \quad A = M - N, \quad A, M \text{ nonsingular}$$

for any choice of \mathbf{x}_0 and any fixed vector \mathbf{b} . What is the *iteration matrix*?

(d) [10 marks] Prove that if Jacobi iteration generates a convergent sequence of iterates for a particular system $A\mathbf{x} = \mathbf{b}$ and any \mathbf{x}_0 , then relaxed Jacobi iteration must also generate a convergent sequence of iterates for the same system and any \mathbf{x}_0 whenever $0 < \theta \leq 1$.

For a particular linear system if the Jacobi iteration has all real eigenvalues with the biggest being $3/4$ and the smallest being $-1/2$, then what is the optimal value of θ which will give fastest convergence of the corresponding relaxed Jacobi iteration? Is this convergence faster than that for the original Jacobi iteration?

What might be the advantage of other choices of θ in the context of *multigrid* iteration?

(e) [7 marks] In a 2-grid iteration applied to $Ax = b$ where $A \in \mathbb{R}^{N \times N}$, the *prolongation operator* $\mathcal{P} : \mathbb{R}^n \rightarrow \mathbb{R}^N$ with $n < N$ is represented by the matrix $P \in \mathbb{R}^{N \times n}$ and the *restriction operator* $\mathcal{R} : \mathbb{R}^N \rightarrow \mathbb{R}^n$ by $P^T \in \mathbb{R}^{n \times N}$. The 2-grid iteration matrix is

$$G := (I - M^{-1}A) \left(A^{-1} - P\bar{A}^{-1}P^T \right) A(I - M^{-1}A)$$

where $\bar{A} = P^T A P$ is the *Galerkin coarse grid operator* and $I - M^{-1}A$ is the iteration matrix associated to the simple iteration employed for smoothing; 1 step of this smoothing iteration is used for pre-smoothing and 1 step for post-smoothing. If $A = A^T$ and $M = M^T$, prove that G can be written as

$$I - M_{MG}^{-1}A$$

where $M_{MG} = M_{MG}^T$. For a system with $A = A^T$, does this hold when relaxed Jacobi iteration is employed for smoothing?