

**Numerical Analysis Preliminary Exam**  
1 TO 4 PM, MAY 21, 2026

INSTRUCTIONS: Submit solutions to four (and no more) of the following six problems. Please start each problem on a new page. You MUST prove your conclusions or show a counter-example for all problems unless otherwise noted. Write your Student ID on your exam; **do not write your name on your exam.**

**Problem 1: Rootfinding**

Consider the following iteration

$$x_{n+1} = \ln \frac{1}{\epsilon} - \ln x_n, \quad x_0 = \ln \frac{1}{\epsilon}.$$

A fixed point of this iteration is a value  $x = \alpha$  such that

$$\epsilon\alpha = e^{-\alpha}.$$

Find an interval of values of  $\epsilon > 0$  for which the iteration converges.

**Solution:** It's easy to show that  $|g'(\alpha)| < 1$  for  $0 < \epsilon < 1/e$ , but this only guarantees that an iteration started sufficiently close will converge for this range of  $\epsilon$ . Since the problem specifies an initial condition and asks whether the iteration will converge from this initial condition, another approach is needed.

To show convergence for this specific initial condition we rely on a lemma from the numerics textbook Atkinson (2.5): If  $a \leq g(x) \leq b$  for  $y \in [a, b]$  and the iteration function  $g$  is Lipschitz continuous on  $[a, b]$  with constant  $K < 1$  then the iteration will converge to the fixed point for every initial condition in  $[a, b]$ .

First note that  $g'(x) = -1/x$ , and for a continuously-differentiable function on a compact interval, the Lipschitz constant is just the max of  $|g'|$  on that interval. Fortunately  $|g'(x)| < 1$  for all  $x > 1$ , so any interval with  $a > 1$  will work for the Lipschitz constant. But  $a = 1$  will not work.

Next note that  $g$  is monotone decreasing for  $x \geq 1$ , so  $1 < a \leq x \leq b$  implies  $g(b) \leq g(x) \leq g(a)$ . In order to get a contraction we need  $a$  and  $b$  such that

$$a \leq g(b) \text{ and } g(a) \leq b.$$

We can choose  $b$  using  $1 < a$  and monotonicity, which implies  $g(a) < g(1) = -\ln \epsilon$ . So set  $b = -\ln \epsilon$ . This also fortunately implies that our initial condition is in the interval. (Note that a solution with  $b = 1/(\epsilon e)$  will not work because then  $g(b) = 1$ , but if  $a = 1$  then we do not have the necessary Lipschitz condition.)

Now that we have  $b$  we can just set  $a = g(b) = -\ln \epsilon - \ln \ln \epsilon^{-1}$ . Note that this only satisfies  $1 < a$  when  $\epsilon < 1/e$ .

At this point we can conclude that the iteration satisfies the conditions of the theorem and therefore converges to the fixed point for  $0 < \epsilon < 1/e$ . But what happens at  $\epsilon = 1/e$ ? Just because the

conditions of the theorem are not met does not mean the iteration does not converge. As it turns out, when  $\epsilon = 1/e$  the initial guess is at the exact fixed point, so the iteration converges there too. We conclude that the iteration converges to the fixed point for

$$0 < \epsilon \leq \frac{1}{e}.$$

**Problem 2: Interpolation & Approximation**

(a) Let  $p$  be a trigonometric polynomial

$$p(x) = \sum_{n=-M}^M c_n e^{inx}$$

that interpolates a function  $f(x)$  at the points

$$x_m = 2\pi \frac{m}{2M+1}, \quad m = 0, \dots, 2M$$

i.e.  $p(x_m) = f(x_m)$  for  $m = 0, \dots, 2M$ . Derive an explicit expression for the coefficients  $c_n$  in terms of the values of  $f$  at the interpolation nodes.

(b) Suppose that  $f$  is a  $2\pi$  periodic function with uniformly-convergent Fourier series

$$f(x) = \sum_{-\infty}^{\infty} \hat{f}_k e^{ikx}.$$

Derive an explicit expression for the error  $\hat{f}_n - c_n$  in terms of the Fourier coefficients  $\hat{f}_n$  of  $f$ .

**Solution:** (a) The linear system for the interpolation coefficients takes the form

$$\mathbf{F}\mathbf{c} = \mathbf{f}$$

where

$$\mathbf{F}_{mn} = e^{2\pi i \frac{m}{2M+1} n} \text{ and } f_m = f(x_m).$$

The matrix  $\mathbf{F}$  is a discrete Fourier transform matrix with the property

$$\mathbf{F}^* \mathbf{F} = (2M+1)\mathbf{I}$$

where  $*$  denotes the complex-conjugate transpose. The interpolation coefficients are therefore

$$\mathbf{c} = \frac{1}{2M+1} \mathbf{F}^* \mathbf{f}$$

or in component form

$$c_n = \frac{1}{2M+1} \sum_{m=0}^{2M} f(x_m) e^{-ix_m n}.$$

(b) Insert the series representation of  $f$  into the above expression for  $c_n$  and simplify

$$\begin{aligned} c_n &= \frac{1}{2M+1} \sum_{m=0}^{2M} \left[ e^{-ix_m n} \sum_{-\infty}^{\infty} \hat{f}_k e^{ikx_m} \right] = \frac{1}{2M+1} \sum_{m=0}^{2M} \sum_{-\infty}^{\infty} \hat{f}_k e^{i(k-n)x_m} = \frac{1}{2M+1} \sum_{-\infty}^{\infty} \sum_{m=0}^{2M} \hat{f}_k e^{i(k-n)x_m} \\ &= \frac{1}{2M+1} \sum_{-\infty}^{\infty} \sum_{m=0}^{2M} \hat{f}_k e^{2\pi i(k-n) \frac{m}{2M+1}} \end{aligned}$$

The inner sum is geometric

$$\sum_{m=0}^{2M} \hat{f}_k e^{2\pi i(k-n) \frac{m}{2M+1}} = \sum_{m=0}^{2M} \hat{f}_k (e^{2\pi i(k-n) \frac{1}{2M+1}})^m$$

when  $k - n$  is an integer multiple of  $2M + 1$  the sum is exactly  $2M + 1$ , else it is zero. Letting  $k - n = q(2M + 1)$  for  $q \in \mathbb{Z}$  we find

$$c_n = \sum_{q=-\infty}^{\infty} \hat{f}_{n+q(2M+1)} = \dots + \hat{f}_{n-(2M+1)} + \hat{f}_n + \hat{f}_{n+(2M+1)} + \dots$$

The desired error formula is therefore

$$\hat{f}_n - c_n = - \sum_{q \in \mathbb{Z}, q \neq 0} \hat{f}_{n+q(2M+1)}.$$

This illustrates the phenomenon of aliasing.

### Problem 3: Quadrature

The following is a generalization of numerical quadrature (as it uses evaluations of  $f(x)$  and  $f'(x)$ ) to approximate integrals of smooth functions on the interval  $(-1, 1)$

$$I[f] = \int_{-1}^1 f(x)dx \approx Q[f] = \sum_{j=0}^n \omega_j f(x_j) + \eta_j f'(x_j)$$

(a) Consider the case  $n = 1$ , corresponding to evaluation at two points  $x_0, x_1 \in [-1, 1]$ . Assuming we have already chosen our quadrature nodes  $x_0$  and  $x_1$ , write down a system of equations aimed at *integrating a polynomial of the highest degree possible  $q$  exactly*.

(b) Given a choice of  $x_0, x_1$ , consider the Hermite interpolant basis of cubic polynomials  $\{H_0, H_1, K_0, K_1\}$  satisfying

$$\begin{aligned} H_0(x_0) &= 1 & H_0(x_1) &= 0 & H'_0(x_0) &= 0 & H'_0(x_1) &= 0 \\ H_1(x_0) &= 0 & H_1(x_1) &= 1 & H'_1(x_0) &= 0 & H'_1(x_1) &= 0 \\ K_0(x_0) &= 0 & K_0(x_1) &= 0 & K'_0(x_0) &= 1 & K'_0(x_1) &= 0 \\ K_1(x_0) &= 0 & K_1(x_1) &= 0 & K'_1(x_0) &= 0 & K'_1(x_1) &= 1 \end{aligned}$$

Applying the quadrature rule to each member of this basis (or by some other method), find formulas for  $\omega_0, \omega_1, \eta_0, \eta_1$ .

(c) Using the fact that you can integrate polynomials up to degree  $q$  exactly (or by some other method), write down an explicit expression for the quadrature error  $E[f] = I[f] - Q[f]$ ; simplify as much as you can. Your expression should not include the integral of  $f$ .

### Solution:

(a) Given that we have 4 unknown quadrature weights, we have 4 degrees of freedom and so, we write equations to integrate polynomials up to degree  $q = 3$  exactly. Applying the proposed quadrature to the monomial basis  $\{1, x, x^2, x^3\}$  gives us 4 linear equations:

$$\begin{aligned} \omega_0 + \omega_1 &= \int_{-1}^1 dx = 2 \\ \omega_0 x_0 + \omega_1 x_1 + \eta_0 + \eta_1 &= \int_{-1}^1 x dx = 0 \\ \omega_0 x_0^2 + \omega_1 x_1^2 + \eta_0(2x_0) + \eta_1(2x_1) &= \int_{-1}^1 x^2 dx = 2/3 \\ \omega_0 x_0^3 + \omega_1 x_1^3 + \eta_0(3x_0^2) + \eta_1(3x_1^2) &= \int_{-1}^1 x^3 dx = 0 \end{aligned}$$

Or in matrix-vector form,

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ x_0 & x_1 & 1 & 1 \\ x_0^2 & x_1^2 & 2x_0 & 2x_1 \\ x_0^3 & x_1^3 & 3x_0^2 & 3x_1^2 \end{bmatrix} \begin{bmatrix} \omega_0 \\ \omega_1 \\ \eta_0 \\ \eta_1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 2/3 \\ 0 \end{bmatrix}$$

- (b) The Hermite basis is a basis of  $P_3([-1, 1])$  that is tailored to Hermite polynomial interpolation (which uses the very same data we are now using for quadrature). Given what we know about standard quadrature weights (they are integrals of the Lagrange interpolant basis), we can conjecture the same will be true in this case for the Hermite basis (this is not a proof; it merely builds intuition guiding our next step). Applying the quadrature and using that it is exact for all cubic polynomials:

$$\begin{aligned} Q[H_0] &= \omega_0 = \int_{-1}^1 H_0(x) dx \\ Q[H_1] &= \omega_1 = \int_{-1}^1 H_1(x) dx \\ Q[K_0] &= \eta_0 = \int_{-1}^1 K_0(x) dx \\ Q[K_1] &= \eta_1 = \int_{-1}^1 K_1(x) dx \end{aligned}$$

This provides formulas for the quadrature weights that are sufficient to answer the question. To obtain a more explicit representation let

$$z = \frac{x - x_0}{x_1 - x_0}, \quad x = x_0 + z(x_1 - x_0).$$

Then

$$\begin{aligned} \int_{-1}^1 \phi(x) dx &= (x_1 - x_0) \int_{z_0}^{z_1} \phi(x_0 + z(x_1 - x_0)) dz \\ z_0 &= \frac{-1 - x_0}{x_1 - x_0}, \quad z_1 = \frac{1 - x_0}{x_1 - x_0}. \end{aligned}$$

The Hermite basis is simple with respect to  $z$ :

$$H_0(x(z)) = (z-1)^2(2z+1), \quad H_1(x(z)) = (3-2z)z^2, \quad K_0(x(z)) = (-1+z)^2z, \quad K_1(x(z)) = (-1+z)z^2$$

Integrating these yields the explicit formulas

$$\begin{aligned} \omega_0 &= \int_{-1}^1 H_0(x) dx = \frac{2(3x_0x_1^2 + x_0 - x_1^3 + x_1)}{(x_0 - x_1)^3} \\ \omega_1 &= \int_{-1}^1 H_1(x) dx = \frac{2(x_0^3 - 3x_0^2x_1 - x_0 - x_1)}{(x_0 - x_1)^3} \\ \eta_0 &= \int_{-1}^1 K_0(x) dx = \frac{2(3x_0x_1^2 + x_0 + 2x_1)}{3(x_0 - x_1)^3} \\ \eta_1 &= \int_{-1}^1 K_1(x) dx = \frac{2(3x_0^2x_1 + 2x_0 + x_1)}{3(x_0 - x_1)^3}. \end{aligned}$$

- (c) Since the quadrature and the integral are linear operators on functions, the same arguments used to show error bounds assuming a quadrature rule is exact for polynomials up to a certain degree  $q$  (e.g. via Taylor expansion) apply in this case. Alternatively, we could use

approximation results from Hermite interpolation to also write down an error bound. Either way, given that this quadrature integrates polynomials of the same degree as Simpson's, we could also conjecture we will get an expression for the error of the form

$$E[f] = C f^{(4)}(\eta)(b-a)^5$$

In either case, we have the following:

$$f(x) = p_3(x) + \frac{f^{(4)}(\eta_x)}{4!} \Psi(x)$$

where  $p_3(x)$  is a cubic polynomial (e.g. 2-pt Hermite interpolant), and  $\Psi(x)$  is a polynomial of degree 4 that is non-negative on  $[-1, 1]$  ( $\Psi(x) = (x - x_0)^2(x - x_1)^2$  for Hermite).

$$\begin{aligned} E[f] &= I[f] - Q[f] \\ &= I\left[\frac{f^{(4)}(\eta_x)}{4!} \Psi(x)\right] - Q\left[\frac{f^{(4)}(\eta_x)}{4!} \Psi(x)\right] \\ &= \frac{f^{(4)}(\eta)}{4!} I[\Psi]. \end{aligned}$$

For general  $x_0, x_1$  the integral is

$$I[\Psi] = \frac{2}{5} + \frac{2}{3}(4x_0x_1 + x_1^2 + x_0^2(1 + 3x_1^2)).$$

here we use the MVT to "extract" the fourth derivative out of the integral, and the fact that  $Q[f] = Q[p_3]$  for the Hermite interpolant (and so the quadrature for the residual term is zero).

**Problem 4: Linear Algebra**

Let  $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$  symmetric matrices. Given  $\mathbf{C} \in \mathbb{R}^{n \times n}$ , we pose the *Sylvester Equation*:

$$\mathbf{A}\mathbf{X} - \mathbf{X}\mathbf{B} = \mathbf{C}$$

for unknown matrix  $\mathbf{X} \in \mathbb{R}^{n \times n}$ . This equation shows up in problems in control theory, Lyapunov stability in dynamical systems, among others.

- (a) Explain why, given our assumptions, we can compute eigendecompositions:

$$\mathbf{A} = \mathbf{U}\mathbf{R}\mathbf{U}^T \quad \mathbf{B} = \mathbf{V}\mathbf{S}\mathbf{V}^T$$

with  $\mathbf{U}, \mathbf{V}$  orthogonal matrices (their columns are orthonormal bases of eigenvectors) and  $\mathbf{R}, \mathbf{S}$  real, diagonal matrices (diagonal entries are eigenvalues). Write a simple description of the most efficient algorithm you know to compute these decompositions.

- (b) Plugging in both eigendecompositions (or by some other method), show the Sylvester equation is equivalent to

$$\mathbf{R}\mathbf{Y} - \mathbf{Y}\mathbf{S} = \mathbf{D}$$

with  $\mathbf{Y} = \mathbf{U}^T \mathbf{X} \mathbf{V}$  and  $\mathbf{D} = \mathbf{U}^T \mathbf{C} \mathbf{V}$ .

- (c) By evaluating the  $k$ -th column of the identity in (b), write down a formula for the  $k$ -th column of  $\mathbf{Y}$ ,  $\mathbf{y}_k$ . Show that this equation has a unique solution if and only if  $\mathbf{A}$  and  $\mathbf{B}$  *do not have any eigenvalues in common*.

**Solution:**

- (a) We have assumed our matrices are *real and symmetric*. By the spectral theorem, that means there is an orthonormal basis of eigenvectors for each one, and that they have real eigenvalues; this shows we can compute the eigendecompositions as described. An algorithm we know to compute it is the *QR iteration*, with basic iteration:

$$\begin{aligned} [\mathbf{Q}_k, \mathbf{R}_k] &= \text{QR}(\mathbf{A}_k) \\ \mathbf{A}_{k+1} &= \mathbf{R}_k \mathbf{Q}_k \end{aligned}$$

To accelerate it, we first compute a unitarily similar tridiagonal  $\mathbf{A}_0$  (e.g. via Householder reflections), and can also accelerate the QR iteration using Rayleigh shifts. When the iteration converges to a diagonal matrix, the diagonal entries of this matrix are the eigenvalues of  $\mathbf{A}$ . To find eigenvectors requires accumulating the product of the  $\mathbf{Q}_k$  matrices; the *rows* of the resulting product matrix are the eigenvectors.

- (b) We plug in the eigendecompositions on the Sylvester equations and use the fact that  $\mathbf{U}, \mathbf{V}$  are unitary:

$$\begin{aligned}
\mathbf{A}\mathbf{X} - \mathbf{X}\mathbf{B} &= \mathbf{C} \\
\mathbf{U}\mathbf{R}\mathbf{U}^T\mathbf{X} - \mathbf{X}\mathbf{V}\mathbf{S}\mathbf{V}^T &= \mathbf{C} \\
\mathbf{R}(\mathbf{U}^T\mathbf{X}\mathbf{V}) - (\mathbf{U}^T\mathbf{X}\mathbf{V})\mathbf{S} &= (\mathbf{U}^T\mathbf{C}\mathbf{V}) \\
\mathbf{R}\mathbf{Y} - \mathbf{Y}\mathbf{S} &= \mathbf{D}
\end{aligned}$$

in which we multiply by  $\mathbf{U}^T$  on the left and  $\mathbf{V}$  on the right to obtain our result.

- (c) We evaluate the  $k$ -th column for this equation, and we use the fact that multiplication by a diagonal matrix from the left scales rows and from the right scales columns. So,

$$\begin{aligned}
\mathbf{R}\mathbf{Y}(:, k) - \mathbf{Y}\mathbf{S}(:, k) &= \mathbf{D}(:, k) \\
\mathbf{R}\mathbf{Y}(:, k) - s_{kk}\mathbf{Y}(:, k) &= \mathbf{D}(:, k) \\
(\mathbf{R} - s_{kk}\mathbf{I})\mathbf{y}_k &= \mathbf{d}_k \\
(r_{ii} - s_{kk})\mathbf{y}_k(i) &= \mathbf{d}_k(i)
\end{aligned}$$

with  $\mathbf{d}_k = \mathbf{D}(:, k)$ . It is clear that we can only solve for all columns of  $\mathbf{Y}$  (and then find  $\mathbf{X}$ ) if and only if  $r_{ii} - s_{kk}$  is non-zero for all  $(i, k)$ . That is: if  $\mathbf{A}$  and  $\mathbf{B}$  do not share any eigenvalues in common.

**Problem 5: Ordinary Differential Equations**

(a) Write the ordinary differential equation

$$y''' + 2y'' + y' + 2y = 0$$

with initial conditions

$$y(0) = 1, y'(0) = 0, y''(0) = 0$$

as a first-order system, and find its characteristic polynomial.

(b) Sketch the absolute stability region for the second-order explicit Runge-Kutta method (find its intersections with real and imaginary axes and include these in your plot),

$$\begin{aligned} k_1 &= hf(t_n, y_n) \\ k_2 &= hf\left(t_n + \frac{h}{2}, y_n + \frac{k_1}{2}\right) \\ y_{n+1} &= y_n + k_2 \end{aligned}$$

(c) For the fourth-order explicit Runge Kutta method

$$\begin{aligned} k_1 &= hf(t_n, y_n) \\ k_2 &= hf\left(t_n + \frac{h}{2}, y_n + \frac{k_1}{2}\right) \\ k_3 &= hf\left(t_n + \frac{h}{2}, y_n + \frac{k_2}{2}\right) \\ k_4 &= hf(t_n + h, y_n + k_3) \\ y_{n+1} &= y_n + \frac{1}{6}k_1 + \frac{1}{3}k_2 + \frac{1}{3}k_3 + \frac{1}{6}k_4 \end{aligned}$$

write equations to determine the absolute stability region and use the equations to check that points  $z = 2i$  and  $z = -2$  are inside it.

(d) For the ODE from (a), should you use the method from parts (b) or (c)? Explain your answer.

**Solution:**

(a) Setting

$$\begin{cases} u = y \\ v = y' \\ w = y'' \end{cases}$$

we obtain

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix}_t = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -1 & -2 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = A \begin{pmatrix} u \\ v \\ w \end{pmatrix}$$

with the initial condition

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix}_{t=0} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

We can either compute eigenvalues of the matrix  $A$  or write the characteristic polynomial directly to obtain

$$z^3 + 2z^2 + z + 2 = (z + 2)(z^2 + 1) = 0$$

so that the eigenvalues of  $A$  are  $\lambda_1 = -2$  and  $\lambda_{2,3} = \pm i$ .

(b) Solving the test problem

$$\begin{cases} y' = \lambda y & \lambda \in \mathbb{C}, \\ y(0) = y_0 \end{cases}$$

using the second order explicit Runge-Kutta method, we obtain

$$\begin{cases} k_1 = h\lambda y_n \\ k_2 = h\lambda (y_n + k_1/2) = h\lambda y_n + \frac{1}{2} (h\lambda)^2 y_n \\ y_{n+1} = \left(1 + h\lambda + \frac{1}{2} (h\lambda)^2\right) y_n \end{cases}$$

(c) Setting  $z = h\lambda$ , the absolute value of the amplification factor

$$1 + z + \frac{1}{2}z^2$$

should be less or equal to 1. Finding intersection with the real axis, we solve

$$1 + r + \frac{1}{2}r^2 = 1$$

yielding  $r = 0$  and  $r = -2$ . Finding intersection with the imaginary axis, we set  $z = it$  and solve

$$1 + it - \frac{1}{2}t^2 = 1.$$

We have  $t = 0$  as the only real solution so that the region of absolute stability has only one intersection with the imaginary axis at zero.

Setting  $z = h\lambda$ , for the fourth order explicit Runge-Kutta method, we have

$$\begin{aligned} k_1 &= zy_n \\ k_2 &= zy_n + \frac{1}{2}zk_1 \\ k_3 &= zy_n + \frac{1}{2}zk_2 \\ k_4 &= zy_n + zk_3 \\ y_{n+1} &= y_n + \frac{1}{6}k_1 + \frac{1}{3}k_2 + \frac{1}{3}k_3 + \frac{1}{6}k_4 \end{aligned}$$

We obtain

$$\begin{aligned} k_1 &= zy_n \\ k_2 &= zy_n + \frac{1}{2}z^2y_n \\ k_3 &= zy_n + \frac{1}{2}z^2y_n + \frac{1}{4}z^3y_n \\ k_4 &= zy_n + z^2y_n + \frac{1}{2}z^3y_n + \frac{1}{4}z^4y_n \end{aligned}$$

and

$$y_{n+1} = \left( 1 + z + \frac{1}{2}z^2 + \frac{1}{6}z^3 + \frac{1}{24}z^4 \right) y_n.$$

The amplification factor is

$$q(z) = 1 + z + \frac{1}{2}z^2 + \frac{1}{6}z^3 + \frac{1}{24}z^4,$$

and  $q(-2) = \frac{1}{3}$  and  $q(2i) = -\frac{1}{3} + \frac{2}{3}i$ , both less than 1 by absolute value.

- (d) The second order explicit Runge-Kutta method cannot be used since the eigenvalues  $\lambda_{2,3} = \pm i$  (for any choice of the step size  $h$ ) are outside the absolute stability region of the method. The fourth order method can be used as its absolute stability region includes a part of the imaginary axis.

### Problem 6: Partial Differential Equations

(a) Show that the explicit Dufort-Frankel method

$$\frac{u_j^{m+1} - u_j^{m-1}}{2h_t} = \frac{c}{h_x^2} \left( u_{j+1}^m - u_j^{m+1} - u_j^{m-1} + u_{j-1}^m \right),$$

for the heat equation

$$u_t = cu_{xx}, \quad c > 0, \quad t \geq 0,$$

with periodic boundary conditions is second order in both space and time variables, is conditionally consistent (i.e. convergence depends on the ratio of time/space step sizes). Here  $u_j^m$  is the value at the space step  $j$  and time step  $m$ ,  $h_x$  is the step size in space and  $h_t$  is the step size in time variables.

(b) Analyze the stability of the Dufort-Frankel method: Provide an algebraic equation whose solution determines stability, and state the condition on the solution of this algebraic equation that ensures that the method is stable.

**Solution:** (a) Taylor expanding the differences and simplifying leaves truncation error

$$\partial_t u_j^m - c \partial_x^2 u_j^m - c \frac{h_x^2}{12} \partial_x^4 u_j^m + c \frac{h_t^2}{h_x^2} \partial_t^2 u_j^m + \frac{h_t^2}{6} \partial_t^3 u_j^m + \mathcal{O}(h_x^4 + h_t^4).$$

The first two terms clearly cancel due to the PDE. This leaves truncation error with leading terms proportional to

$$h_x^2, h_t^2, \text{ and } \frac{h_t^2}{h_x^2}.$$

In order for the method to be consistent we need to refine the space/time grid in such a way that

$$\lim_{h_x, h_t \rightarrow 0} \frac{h_t}{h_x} = 0$$

and the overall order of the method will be determined by the rate at which this ratio approaches zero. Suppose that  $h_t = h_x$ ; then the method is not convergent because the term in the truncation error proportional to  $(h_t/h_x)^2$  does not approach zero as the grid is refined. In order to be consistent we need  $h_t$  to go to zero faster than  $h_x$ . In order to be second-order, as claimed in the problem, we need

$$\frac{h_t}{h_x} = \mathcal{O}(h_x)$$

which requires

$$h_t = \mathcal{O}(h_x^2).$$

(b) Denoting

$$k = \frac{2ch_t}{h_x^2},$$

the explicit scheme can be re-written

$$u_j^{m+1} = \frac{1-k}{1+k} u_j^{m-1} + \frac{k}{1+k} (u_{j+1}^m + u_{j-1}^m).$$

Apply the scheme to  $r^m e^{iph_x j}$ , where  $r$  is amplification factor, to obtain

$$r^{m+1} e^{iph_x j} = \frac{1-k}{1+k} r^{m-1} e^{iph_x j} + \frac{k}{1+k} r^m e^{iph_x j} (e^{iph_x} + e^{-iph_x})$$

or

$$r^2 = \frac{1-k}{1+k} + r \frac{2k}{1+k} \cos(ph_x).$$

The method is stable when both roots are inside the unit disk in the complex plane.