Norwegian University of Science and Technology
Department of Mathematical Sciences

Contact during the exam:
Elena Celledoni, tlf. 735 93541 mobile 48238584

EXAM in TMA4212

Monday 12th August 2013
Time: 9:00–13:00

Allowed aids: Approved simple pocket calculator. All written and handwritten material from the course.

Learning outcome **L6** and **L3** have been tested through the project work[1].

Problem 1 *Learning outcome* **L2, L7**
Given that α is a nonnegative real number, consider the differential equation

$$-u'' + u = f(x), \quad \text{for } x \in (0,1),$$

subject to the boundary conditions

$$u(0) = 0, \qquad \alpha u(1) + u'(1) = 0.$$

a) State the weak formulation of the problem.

b) Using continuous piecewise linear basis functions on a uniform subdivision of $[0,1]$ into elements of size $h = \frac{1}{n}$, $n \geq 2$, write down the finite element approximation to this problem and show that this has a unique solution. Expand u^h in terms of the standard piecewise linear finite element basis functions (hat functions) φ_i, $i = 1, 2, \ldots, n$, by writing

$$u^h(x) = \sum_{i=1}^{n} U_i \varphi_i(x)$$

to obtain a system of linear equations for the vector of unknowns $(U_1, \ldots, U_n)^T$.

[1] The learning outcome is published on the course webpage and on the official description of the course.

c) Suppose $\alpha = 0$, $f(x) \equiv 1$ and $h = \frac{1}{3}$. Solve the resulting system of linear equations and compute the corresponding numerical solution.

d) Compare the numerical solution with the exact solution u by finding the values of the error function $u_h - u$ in the nodes of the finite element discretization. You can find u assuming $u(x) = c_1 e^{-x} + c_2 e^{x} + 1$, and determining c_1 and c_2 imposing the boundary conditions.

Solution (a)

We multiply the equation by a test function $v \in H^1(0,1)$ such that $v(0) = 0$, and integrate between 0 and 1:

$$-\int_0^1 u'' v\, dx + \int_0^1 uv\, dx = \int_0^1 fv\, dx.$$

Integrating by parts and using the boundary conditions (using $v(0) = 0$ and $u(0) = 0$, $\alpha u(1) + u'(1) = 0$) we obtain

$$\alpha u(1)v(1) + \int_0^1 (u'v' + uv)\, dx = \langle f, v\rangle,$$

where $\langle \cdot, \cdot \rangle$ denotes the $L_2(0,1)$ inner product. We then define the bilinear map

$$A(u,v) := \alpha u(1)v(1) + \int_0^1 (u'v' + uv)\, dx,$$

which is symmetric and positive definite[2]. So the weak formulation is:

Find $u \in H^1(0,1)$ satisfying $u(0) = 0$ and $\alpha u(1) + u'(1) = 0$, such that

$$A(u,v) = \langle f, v\rangle, \qquad v \in H^1(0,1) \quad v(0) = 0.$$

Solution (b)

We consider the basis functions n piecewise linear basis functions $\varphi_1, \ldots, \varphi_n$ such that $\varphi_j(x_j) = 1$, $j = 1, \ldots, n$:

$$\varphi_j(x) = \begin{cases} \frac{x - x_{j-1}}{h}, & x \in [x_{j-1}, x_j), \\ \frac{x_{j+1} - x}{h}, & x \in [x_j, x_{j+1}], \\ 0 & \text{otherwise}, \end{cases} \qquad j = 1, \ldots, n-1$$

[2]

- $A(u,v) = A(v,u)$
- $A(v,v) = v(1)^2 + \int_0^1 (v')^2 + v^2\, dx > 0$ if v is not equal to zero almost everywhere, because v is absolutely continuous.

and

$$\varphi_n = \begin{cases} \frac{x-x_{n-1}}{h}, & x \in [x_{n-1}, x_n], \\ 0 & \text{otherwise.} \end{cases}$$

The finite element space is defined by

$$S_h := \{w_h \in H^1 \mid w_h = \sum_{j=1}^{n} w_j \varphi_j, \},$$

and the Galerkin method is

$$\text{Find}\, u_h \in S_h \,\text{s.t.}\, A(u_h, v_h) = \langle f, v_h \rangle, \quad \forall v_h \in S_h.$$

This is equivalent to

$$\text{Find}\, u_h \in S_h \,\text{s.t.}\, A(u_h, \varphi_i) = \langle f, \varphi_i \rangle, \quad i = 1, \ldots, n.$$

Expanding u^h by means of the basis functions

$$u_h(x) = \sum_{j=1}^{n} U_j \varphi_j(x)$$

and substituting in the Galerkin method we get

$$\text{Find}\, u_h = \sum_{j=1}^{n} U_j \varphi_j(x) \,\text{s.t.} \sum_{j=1}^{n} M_{i,j} U_j = \langle f, \varphi_i \rangle, \quad j = 1, \ldots, n,$$

where the entries

$$M_{i,j} := A(\varphi_j, \varphi_i), \quad i, j = 1, \ldots, n$$

form the matrix M of a linear system with unknowns $\mathbf{x} := (U_1, \ldots, U_n)^T$:

$$M\mathbf{x} = \mathbf{b},$$

where $\mathbf{b} := (\langle f, \varphi_1 \rangle, \ldots, \langle f, \varphi_n \rangle)^T$. The matrix M is symmetric (because of the symmetry of the bilinear map A) and positive definite. Infact for any $\mathbf{v} \in \mathbf{R}^n$, $\mathbf{v} \neq 0$,

$$\mathbf{v}^T M \mathbf{v} = A\big(\sum_{i=1}^{n} v_i \varphi_i, \sum_{j=1}^{n} v_j \varphi_j \big)$$

and since $v(x) = \sum_{j=1}^{n} v_j \varphi_j \in H^1$ is non-zero almost everywhere then

$$\mathbf{v}^T M \mathbf{v} = A(v, v) > 0,$$

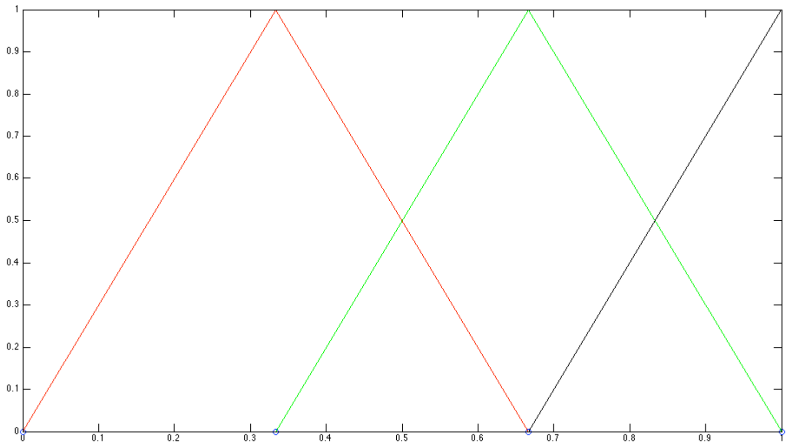

Figure 1: Basis functions

and M is positive definite (implying that it is also non-singular). This means that the system has a unique solution $(U_1, \dots, U_n)^T$ and therefore the solution of the Galerkin method

$$u^h(x) = \sum_{j=1}^{n} U_j \varphi_j(x)$$

exists and is unique.

Solution (c)

In Figure 1 the three basis functions φ_1 (red), φ_2 (green) and φ_3 (black) are depicted and

$$u_h = \sum_{j=1}^{3} U_j \varphi_j.$$

The linear system consists of a 3×3 matrix with entries $M_{i,j} := A(\varphi_j, \varphi_i)$. To compute these entries we will use

$$\varphi_j'(x) = \begin{cases} \frac{1}{h}, & x \in [x_{j-1}, x_j), \\ -\frac{1}{h}, & x \in [x_j, x_{j+1}], \\ 0 & \text{otherwise}, \end{cases} \qquad j = 1, 2$$

and

$$\varphi_3' = \begin{cases} \frac{1}{h}, & x \in [x_2, x_3], \\ 0 & \text{otherwise}. \end{cases}$$

We assume $f \equiv 1$, $\alpha = 0$. We get

$$A(\varphi_1, \varphi_1) = \int_0^1 (\varphi')^2 \, dx + \int_0^1 (\varphi)^2 \, dx = \frac{2}{h} + \frac{2}{3}h = \frac{56}{9}.$$

$A(\varphi_2, \varphi_2) = A(\varphi_1, \varphi_1)$ and

$$A(\varphi_3, \varphi_3) = \frac{1}{h} + h\frac{1}{3}.$$

Similarly

$$A(\varphi_1, \varphi_2) = \frac{1}{6}h - \frac{1}{h} = A(\varphi_2, \varphi_3).$$

So

$$M := \begin{bmatrix} \frac{2}{3}h + \frac{2}{h} & \frac{1}{6}h - \frac{1}{h} & 0 \\ \frac{1}{6}h - \frac{1}{h} & \frac{2}{3}h + \frac{2}{h} & \frac{1}{6}h - \frac{1}{h} \\ 0 & \frac{1}{6}h - \frac{1}{h} & \frac{1}{3}h + \frac{1}{h} \end{bmatrix}, \quad \mathbf{b} := h \begin{bmatrix} 1 \\ 1 \\ \frac{1}{2} \end{bmatrix}$$

Solving the linear system we get $U_1 = 0.2039$, $U_2 = 0.3177$, $U_3 = 0.3543$. Figure 2 reports a comparison of the exact (red curve) and the numerical solution obtained with the finite element method (blue line). The exact solution is

$$u(x) = c_1 e^{-x} + c_2 e^{x} + 1,$$

with $c_1 = -\frac{e^2}{1+e^2}$, $c_2 = -\frac{1}{1+e^2}$.

Solution (d) The error in the nodes of the finite element discretization is $|U_1 - u(h)| = 0.2039 - 0.2025 = 0.0014$, $|U_2 - u(2h)| = 0.3177 - 0.3156 = 0.0021$, and $|U_3 - u(3h)| = 0.3543 - 0.3519 = 0.0023$.

Problem 2 *Learning outcome* **L1, L4, L5, L7**

Consider the linear PDE

$$u_t + u_{xxx} = 0, \quad x \in [0, 1], \quad t \geq 0,$$

with periodic boundary conditions. Consider the grid $x_m = hm$, $h = 1/M$, $m = 1, \ldots, M$. Discretize with central finite differences in space and the Forward Euler method in time (forward differences in time).

Use the following central differences approximation of the third derivative

$$u_{xxx}|_{(x_m,t)} = \frac{u(x_{m+3}, t) - 3u(x_{m+1}, t) + 3u(x_{m-1}, t) - u(x_{m-3}, t)}{8h^3} + \mathcal{O}(h^2).$$

a) Perform a Von Neumann stability analysis of the method.

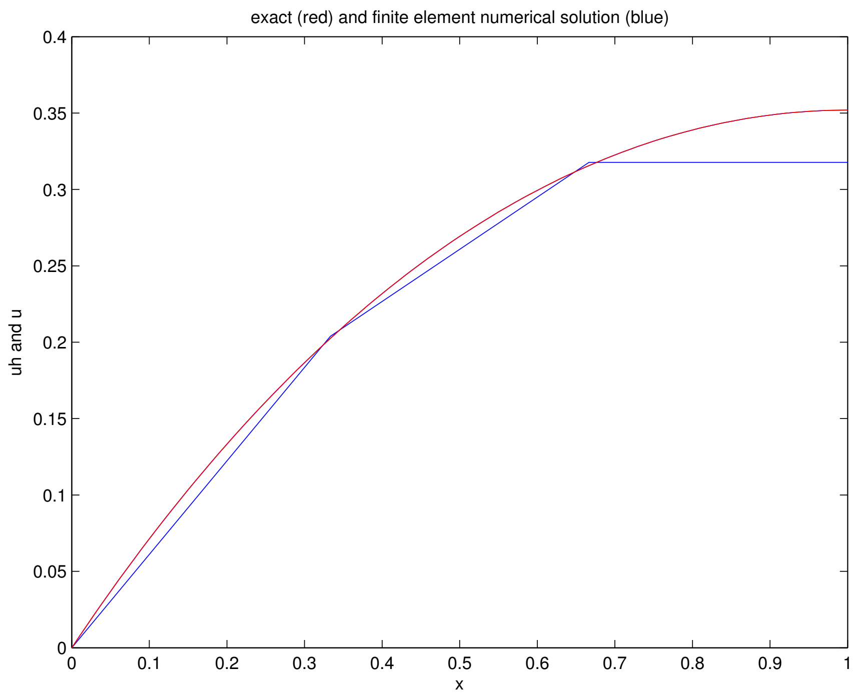

Figure 2: Exact (red) and numerical (FEM) solution (blue)

b) Consider now a Crank-Nicolson method obtained applying central differences in space (as in point (a)) and the trapezoidal rule in time. Perform a stability analysis of the method following the techniques explained in chapter 5 of the note (see page 55 chapter 5.6 of the note). Find under which restrictions on h and k the method is stable.

Hint: You might use the fact that the chosen discretization of the third derivative operator is a skew-symmetric matrix (and circulant). See also exercise (c) below.

c) Consider now the method obtained applying the Backward Euler method in time (backward differences in time). The linear system to be solved at each time-step is of the form

$$(I - kA)\, x = b,$$

with A $M \times M$ circulant and skew-symmetric. (See the appendix on circulant matrices). Consider the Jacobi iterative method to solve this linear system. Assume M and h are fixed and find under which conditions on k the Jacobi iteration is guaranteed to converge.

Solution (a)

Using central differences in space and backward Euler in time we obtain

$$U_m^{n+1} = U_m^n + \frac{k}{8h^3}\left(U_{m-3}^n - 3U_{m-1}^n + 3U_{m+1}^n - U_{m+3}^n\right),$$

by assuming $U_m^n = \xi^n e^{i\beta x_m}$, and $i = \sqrt{-1}$, and substituting in the method we obtain

$$\xi = 1 - i\alpha,$$

with

$$\alpha = \frac{k}{8h^3}(\sin(3\beta h) - 3\sin(\beta h)).$$

Finally we obtain

$$\xi^*\xi = |\xi|^2 = (1+i\alpha)(1-i\alpha) = 1+\alpha^2,$$

which cannot be is less than or equal to one for all β. This means that the method is never Von Neumann stable.

Solution (b)

The Crank-Nicolson method is

$$U_m^{n+1} = U_m^n + \frac{k}{16h^3}\left(U_{m-3}^n - 3U_{m-1}^n + 3U_{m+1}^n - U_{m+3}^n + U_{m-3}^{n+1} - 3U_{m-1}^{n+1} + 3U_{m+1}^{n+1} - U_{m+3}^{n+1}\right).$$

We can write the method in the compact form

$$\mathbf{U}^{n+1} = \mathbf{U}^n + \frac{k}{2}(A\mathbf{U}^n + A\mathbf{U}^{n+1}),$$

where A is the cyclic matrix with first row

$$\mathbf{a}^T = \frac{1}{8h^3}\left[\begin{array}{cccccccccc} 0 & 3 & 0 & -1 & 0 & \dots & 0 & 1 & 0 & -3 \end{array}\right] \in \mathbf{R}^M,$$

with eigenvalues $\sqrt{M}\,\Omega^H\mathbf{a}$. The method becomes

$$\mathbf{U}^{n+1} = (I - \frac{k}{2}A)^{-1}(I + \frac{k}{2}A)\mathbf{U}^n.$$

A sufficient condition for stability for a finite difference method given in the form

$$\mathbf{U}^{n+1} = B\mathbf{U}^n + \mathbf{g}$$

is that

$$\|B\| \leq 1$$

for some norm $\|\cdot\|$ (which we choose here to be the 2-norm). In our case $\mathbf{g} = 0$ and

$$B = (I - \frac{k}{2}A)^{-1}(I + \frac{k}{2}A)$$

and can be diagonalized as follows

$$B = \Omega^H (I - \frac{k}{2}\Lambda)^{-1}(I + \frac{k}{2}\Lambda)\Omega$$

so

$$\|B\|_2 = \|(I - \frac{k}{2}\Lambda)^{-1}(I + \frac{k}{2}\Lambda)\| = \max_{j=1,\dots,M} \left| \frac{1 + \frac{k}{2}\lambda_j}{1 - \frac{k}{2}\lambda_j} \right|$$

here λ_j are the eigenvalues of a skew-symmetric matrix and are therefore pure imaginary and appearing in conjugate pairs, so

$$\|B\|_2 = 1.$$

In conclusion the method is always stable.

Solution (c)

The Jacobi iteration method for this problem is

$$x^{\ell+1} = kA\,x^{\ell} + b, \qquad \ell = 0, 1, \dots$$

and a necessary and sufficient condition for convergence of this iteration is

$$\rho(kA) < 1,$$

where ρ denotes the spectral radius. The eigenvalues λ_l, $l = 1, \dots, M$ of kA are the components of the vector

$$k\sqrt{M}\,\Omega^H \mathbf{a}.$$

Multiplying out the columns of Ω^H and $\mathbf{a}$ and using $\alpha_l := \frac{l-1}{M}$ we obtain

$$\begin{aligned}
\lambda_l &= \frac{k}{8h^3}\left(3\exp(2\pi i\,\alpha_l) - \exp(2\pi i\,3\alpha_l) + \exp(2\pi i\,((l-1) - 3\alpha_l)) - 3\exp(2\pi i\,((l-1) - \alpha_l)\right) \\
&= \frac{k}{8h^3} 2i\,\left(3\sin(\omega_l) - \sin(3\omega_l)\right), \quad \omega_l := 2\pi\alpha_l, \\
&= \frac{k}{8h^3} 8i\sin(\omega_l)^3
\end{aligned}$$

From which we deduce

$$\rho(kA) \leq \frac{k}{h^3}$$

and to ensure $\rho(kA) < 1$ it is sufficient to require $k < h^3$.

Appendix

A circulant matrix is a matrix of the type

$$C = \begin{bmatrix} c_0 & c_1 & c_2 & \cdots & c_{M-2} & c_{M-1} \\ c_{M-1} & c_0 & c_1 & c_2 & \ddots & c_{M-2} \\ c_{M-2} & c_{M-1} & c_0 & c_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & c_1 \\ c_1 & \cdots & \cdots & \cdots & c_{M-1} & c_0 \end{bmatrix},$$

and the first row in the matrix is

$$\mathbf{c}^T = [c_0, c_1, c_2, \ldots, c_{M-2}, c_{M-1}],$$

and it is determining the whole matrix. Circulant matrices can be diagonalized using the Fourier matrix Ω, where

$$\Omega_{k,l} = \frac{1}{\sqrt{M}} \exp\left(2\pi i \cdot (k-1)(l-1)/M\right), \quad i = \sqrt{-1}$$

and $C = \Omega^H \Lambda \Omega$ with Λ a diagonal matrix, and $\Omega^H \Omega = I$. The eigenvalues of C, i.e. the diagonal elements of Λ, are given by

$$\sqrt{M} \cdot \Omega^H \mathbf{c}.$$