UPPSALA UNIVERSITY
Department of Information Technology
Division of Scientific Computing
Solutions to exam in Finite element methods 2010-12-21

Stefan Engblom

Question 1

(a) It is clear that a homogeneous Dirichlet condition is used at $x = 0$ since the basis function φ_0 is not present and since there are no modifications of the first element of the load-vector. Define therefore $V^0 = \{w;\ \|w\| + \|w_x\| < \infty, w(0) = 0\}$ where we will understand the usual $L^2(I)$-inner product and -norm. The variational formulation reads: *find* $u \in V^0$ *s.t.* $(v_x, u_x) + (v, u) = (v, f) + \beta v(1)$ *for* $\forall v \in V^0$. For the basis functions given, $\varphi_N(1) = 1$, explaining the final element of the vector d.

Integrating by parts "backwards" we get using the Dirichlet condition on v, $(v, f)+\beta v(1) = (v, -u_{xx} + u) + v(1)u_x(1)$. Hence the strong formulation is *find* u *s.t.* $-u_{xx} + u = f$ *for* $x \in I$ *with* $u(0) = 0$ *and* $u_x(1) = \beta$.

A different and less constructive solution is to simply guess a strong formulation and then derive the FEM to see that it matches the one given.

(b) For a generic element $I = [x_1, x_2]$ the local mass-matrix is defined by $M^I_{ij} = (\varphi_i, \varphi_j)$ for $i,\ j = 1, 2$. The assembly of the global mass-matrix is carried out by adding the contributions from all local mass-matrices according to the numbering of the element they belong to. In 1D, one can show that $M^I = h/6 \times [2\ 1; 1\ 2]$ where $h = x_2 - x_1$ is the element size.

(c) Multiplying the discrete FEM with ξ^T we get the relation $\|U_x\|^2 + \|U\|^2 = (U, f) \le \|U\|\|f\| \le \|U\|^2/2 + \|f\|^2/2$ using the Cauchy-Schwartz inequality and the hint. Hence $\|U_x\|^2 \le \|U_x\|^2 + \|U\|^2/2 \le \|f\|^2/2$. A slightly sharper estimate comes from using the inequality given, but with $a = \sqrt{2}\|U\|$ and $b = \|f\|/\sqrt{2}$ instead. One then obtains $\|U_x\| \le \|f\|/2$. The same sharp estimate is obtained from $\|U\|\|f\| \ge \|U_x\|^2 + \|U\|^2 \ge$ *(hint)* $2\|U\|\|U_x\|$ after dividing through with $\|U\|$. It is also possible (but less sharp) to use the Poincaré inequality here since $U(0) = 0$. A derivation reads: $\|U_x\|^2 \le \|U_x\|^2 + \|U\|^2 \le \|U\|\|f\| \le$ *(Poincaré)* $C\|U_x\|\|f\|$ and dividing through with $\|U_x\|$.

Question 2

Note: the fact that the heat-capacity κ is supposed to be a positive constant is missing.

(a) Multiplying with a test-function $v \in V_0 := \{w;\ \|w\| + \|\nabla w\| < \infty, w|_{\partial\Omega} = 0\}$ we get using Green's formula that $(v, u_t) = -\kappa(\nabla v, \nabla u)$ in the $L^2(\Omega)$-inner product and induced norm (defined also for vector-valued functions by $(F, G) := \int_\Omega F \cdot G\, dx$). Given the triangulation $\mathcal{K}$, define $V_{h,0} := \{w;\ w$ piecewise linear and continuous on $\mathcal{K}, w|_{\partial\Omega} = 0\}$. Using the standard basis $\{\varphi_j\}_{j=1}^N$ for $V_{h,0}$ we get the FEM $M\xi_t = -\kappa A\xi$ for $t > 0$. The initial data can be obtained from a straightforward projection as $M\xi_0 = d$ (solved once!). Here $M_{ij} = (\varphi_i, \varphi_j)$, $A_{ij} = (\nabla\varphi_i, \nabla\varphi_j)$, and $d_i = (\varphi_i, u_0)$. Using the Euler backward method results in the linear system of equations $(M + k\kappa A)\xi_{n+1} = M\xi_n$ with time-step k for $n = 0, 1, \ldots$

(b) The $L^2(\Omega)$-error is expected to behave as $O(k) + O(h^2)$, so in order for both terms to become 100 times smaller, our best guess is that $(k, h) \to (k/100, h/10)$ should suffice.

(c) Multiplying the PDE with the exact solution u and integrating using Green's formula we get $(u, u_t) = d/dt\,\|u\|^2/2 = -\kappa\|\nabla u\|^2 \le 0$ which proves the first assertion. The same strategy in the discrete case yields $\xi^T_{n+1}(M + k\kappa A)\xi_{n+1} = \xi^T_{n+1}M\xi_n$ which can be understood as $\|U_{n+1}\|^2 + k\kappa\|\nabla U_{n+1}\|^2 = (U_{n+1}, U_n) \le \|U_{n+1}\|\|U_n\|$ by Cauchy-Schwartz.

Hence $\|U_{n+1}\|^2 \leq \|U_{n+1}\|^2 + k\kappa\|\nabla U_{n+1}\|^2 \leq \|U_{n+1}\|\|U_n\|$ and so $\|U_{n+1}\| \leq \|U_n\|$ as claimed. *Extra:* the energy method can in fact be applied to the initial data as well. Then we get $\|U_0\|^2 = (U_0, u_0) \leq \|U_0\|\|u_0\|$ so that in fact $\|U_n\| \leq \cdots \|U_0\| \leq \|u_0\|$.

Question 3

(a) A T-matrix that will do is (note the convention of numbering nodes in a triangle counterclockwise)

$$T = \begin{bmatrix} 1 & 1 & 2 & 4 & 4 & 4 & 5 & 7 & 7 \\ 2 & 4 & 5 & 5 & 7 & 6 & 9 & 9 & 8 \\ 4 & 3 & 4 & 7 & 6 & 3 & 7 & 8 & 6 \end{bmatrix}.$$

The sparsity pattern of the mass-matrix can be obtained by inspecting the connectivity of the nodes in the mesh. Nodes (i, j) connected by an edge implies that $M_{ij} \neq 0$.

$$M \sim \left[\begin{array}{ccc|ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ \hline 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{array}\right].$$

(b) See Figure 1. An example of a hanging node is the one on the edge 1-4 *before* it is connected to node 3.

(c) FEM can be used in the industry already at the earlier stages of construction work. For instance, from a suggested CAD-geometry of some new detail in a car one can build a computational mesh (triangulation) and then test how it will deform when heated. Another typical design question might be how large the forces will be when an outer pressure or tension is applied given certain material properties. In this way suggestions as to improve the construction can be obtained *before* actually building the detail and/or performing any experiments. Other physical properties that can be similarly tested are flow of air or liquids through channels or cavities, response to electromagnetic waves, chemical reactions and the like.

(d) For example, one can

- Check that all boundary conditions are fulfilled.
- Check that the scaling of the output is reasonable and compare reasonably well with physical intuition.
- Try to solve the same problem again but with a finer mesh (using also a smaller time-step in the case of a time-dependent PDE).
- Compare how well the FEM agrees with an *analytical solution*. For the problem at hand, usually no analytical solution exists but one can try using a simpler geometry and/or simpler boundary conditions, simpler form of PDE...
- Check some property of the solution which is known exactly. For instance, check that the total energy remains or that the total amount of substance (or whatever) is preserved.
- If an *a posteriori* error estimate is available, an adaptive solver might be used to enforce a certain degree of accuracy.

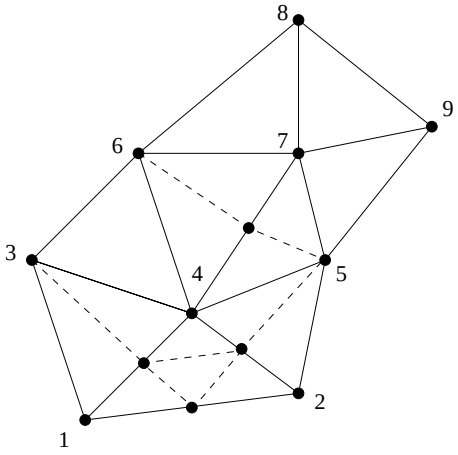

Figure 1: The mesh after uniform refinement of triangle 1-2-4 and splitting the edge 4-7, ensuring also a valid triangulation by connecting hanging nodes appropriately. The new nodes have not yet received any numbers.

- Solve the problem anew but with altered boundary conditions and/or initial data. Does the new solution change in the way we expect?

(and there are probably more examples here).

Question 4

(a) Homogeneous Dirichlet boundary conditions means that $u(0) = u(1) = 0$. Multiplying the PDE by a test-function v and integrating by parts using the BCs we readily get $\varepsilon(v_x, u_x) = (v, f)$. Define $V_0 = \{w;\ \|w\| + \|w_x\| < \infty, w(0) = w(1) = 0\}$. The variational form (VF) is "Find $u \in V_0$ s.t. $\varepsilon(v_x, u_x) = (v, f)$ for $\forall v \in V_0$". Defining $V_{h,0} = \{w;\ w$ piecewise linear and continuous on the mesh of $I, w(0) = w(1) = 0\}$ we see that $\{\varphi_i\}_{i=1}^{N-1}$ is a basis of $V_{h,0}$. The FEM is just "Find $U \in V_{h,0}$ s.t. $\varepsilon(v_x, U_x) = (v, f)$ for $\forall v \in V_{h,0}$". Using the ansatz $U = \sum_{j=1}^{N-1} \varphi_j \xi_j$ we get the fully discrete FEM $\varepsilon A\xi = b$ in terms of $A_{ij} = (\varphi_i', \varphi_j')$, and $b_i = (\varphi_i, f)$ for $i, j = 1 \dots N-1$.

(b) By subtracting (FEM) from (VF) we get the Galerkin orthogonality $(v_x, u_x - U_x) = 0$ for $\forall v \in V_{h,0}$. Defining the error to be $e = u - U$ we have that $\|e_x\|^2 = (u_x - U_x, u_x - U_x) = (u_x - v_x, u_x - U_x) + (v_x - U_x, u_x - U_x)$. For $v, U \in V_{h,0}$ we therefore get using Galerkin orthogonality that $\|e_x\|^2 = (u_x - v_x, u_x - U_x) \le \|u_x - v_x\|\|e_x\|$ from Cauchy-Schwartz. Hence $\|e_x\| \le \|u_x - v_x\|$ for $\forall v \in V_{h,0}$. That is, in the sense of the energy norm, U is a best approximation in $V_{h,0}$.

(c) From Galerkin orthogonality we get $E := \varepsilon\|e_x\|^2 = \varepsilon(e_x, e_x - (\pi e)_x) = \sum_i \varepsilon(e_x, e_x - (\pi e)_x)_i$ where the sum is over elements $i = 1 \dots N-1$ and the inner product is the one over element I_i. Integrating by parts and using the fact that the interpolation π is exact at the endpoints of I_i we get $E = \sum_i \varepsilon(-e_{xx}, e - \pi e)_i = \sum_i (f + \varepsilon U_{xx}, e - \pi e)_i$ by using the PDE. From Cauchy-Schwartz we get $E \le \sum_i \|f + \varepsilon U_{xx}\|_i \|e - \pi e\|_i \le \sum_i \|f + \varepsilon U_{xx}\|_i C h_i \|e_x\|_i \le C(\sum_i \|f + \varepsilon U_{xx}\|_i^2 h_i^2)^{1/2} (\sum_i \|e_x\|_i^2)^{1/2}$ using the suggested interpolation estimate and the (discrete) Cauchy-Schwartz. To the right we recognizes $\|e_x\|$ which finishes the derivation.

(d) One typically starts from some suitable coarse mesh and computes the first FE-solution U. Then the *a posteriori* estimate is computed (per element) and the elements i for which $R_i(U)$ is large are refined. A typical refinement criterion is to refine elements j where $R_j(U) \ge \lambda \max_i R_i(U)$ and where $\lambda \in (0, 1)$ is a constant. To stop the algorithm one can check two subsequent solutions against each-other, one can check the total residual, or one can stop when the total computing time is too long (or something similar).

Question 5

Note: the facts that the wave-speed ϵ is supposed to be a constant and that f is time-independent are missing from the statement of the question.

(a) Multiplying with a test-function $v \in V := \{w;\ \|w\|+\|\nabla w\| < \infty\}$ and integrating using Green's formula and the homogeneous Neumann conditions we get $(v, u_{tt}) = -\epsilon(\nabla v, \nabla u)+ (v, f)$ in terms of the $L^2(\Omega)$-inner product (defined as usual also for vector-valued functions by $(F, G) := \int_\Omega F \cdot G\, dx$). Define $V_h = \{w;\ w$ piecewise linear and continuous on $\mathcal{K}\}$ and let $\{\varphi_j\}_{j=1}^N$ be a basis. A semi-discrete FEM is now $M\xi_{tt} = -\epsilon A\xi + b$ in terms of the FE-solution $U = \sum_j \varphi_j \xi_j(t)$, $M_{ij} = (\varphi_i, \varphi_j)$, $A_{ij} = (\nabla\varphi_i, \nabla\varphi_j)$, and $b_i = (\varphi_i, f)$. The initial data can be obtained by projection as $M\xi(0) = c$ and $M\xi_t(0) = d$ with $c_i = (\varphi_i, u_0)$ and $d_i = (\varphi_i, v_0)$. To discretize in time, define the variable $\eta = \xi_t$ so that $M\eta(0) = d$ and

$$\begin{bmatrix} M & 0 \\ 0 & M \end{bmatrix} \begin{bmatrix} \xi_t \\ \eta_t \end{bmatrix} = \begin{bmatrix} 0 & M \\ -\epsilon A & 0 \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix} + \begin{bmatrix} 0 \\ b \end{bmatrix}.$$

The trapezoidal rule is the forward difference for the left side and an average for the right side. Hence

$$\begin{bmatrix} M & 0 \\ 0 & M \end{bmatrix} \begin{bmatrix} \xi_{n+1} - \xi_n \\ \eta_{n+1} - \eta_n \end{bmatrix} = \frac{k}{2} \begin{bmatrix} 0 & M \\ -\epsilon A & 0 \end{bmatrix} \begin{bmatrix} \xi_{n+1} + \xi_n \\ \eta_{n+1} + \eta_n \end{bmatrix} + k \begin{bmatrix} 0 \\ b \end{bmatrix}.$$

Equivalently,

$$\begin{bmatrix} M & -k/2\ M \\ \epsilon k/2\ A & M \end{bmatrix} \begin{bmatrix} \xi_{n+1} \\ \eta_{n+1} \end{bmatrix} = \begin{bmatrix} M & k/2\ M \\ -\epsilon k/2\ A & M \end{bmatrix} \begin{bmatrix} \xi_n \\ \eta_n \end{bmatrix} + k \begin{bmatrix} 0 \\ b \end{bmatrix},$$

in terms of the time-step k. The initial data are given as before by $M\xi_0 = c$ and $M\eta_0 = d$.

(b) Multiply the PDE with u_t and integrate using Green's formula and the homogeneous Neumann condition. We get $(u_t, u_{tt}) = -\epsilon(\nabla u_t, \nabla u)$. Dropping a factor of $1/2$ this can be written as $d/dt\, \|u_t\|^2 = -\epsilon\, d/dt\, \|\nabla u\|^2$. Hence upon integration over 0 to T we get $\|u_t(T)\|^2 - \|v_0\|^2 = -\epsilon\|\nabla u(T)\|^2 + \epsilon\|\nabla u_0\|^2$ which is the stated equality. The result can be understood as a kind of energy conservation for the wave equation under consideration.