Lesson:  
Introduction to the finite element method *  

September 6, 2001

Contents

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
</tr>
<tr>
<td>2</td>
<td>Strong and weak form</td>
</tr>
<tr>
<td>3</td>
<td>Discrete form</td>
</tr>
<tr>
<td>4</td>
<td>Collocation methods</td>
</tr>
<tr>
<td>5</td>
<td>Finite element methods</td>
</tr>
<tr>
<td>6</td>
<td>Numerical integration</td>
</tr>
<tr>
<td>7</td>
<td>Boundary conditions</td>
</tr>
<tr>
<td>8</td>
<td>Solution methods</td>
</tr>
</tbody>
</table>

*This document is a part of the Finite Element Computer Simulation Course for Mechanical Engineering students in Eindhoven
1 Introduction

In this lesson we will illustrate the basic concepts and notation of the Galerkin finite element method by means of an example problem given by a scalar diffusion equation in 2D. More extended information can be found in Baaijens (1999).

2 Strong and weak form

Consider a domain $\Omega \subset \mathbb{R}^2$ with boundary $\Gamma = \Gamma_n \cup \Gamma_d$ such that $\Gamma_n \cap \Gamma_d = \emptyset$. Furthermore we assume $f(\vec{x}) : \Omega \to \mathbb{R}$ to be a given function. The strong form of the diffusion problem is given by the differential equation:

Find $u(\vec{x}) : \Omega \to \mathbb{R}$ that is a solution of:

\[
\begin{cases}
-\vec{\nabla} \cdot \vec{\nabla} u = f & \text{in } \Omega \\
u = g & \text{on } \Gamma_d \\
\frac{\partial u}{\partial n} = h & \text{on } \Gamma_n
\end{cases}
\]

(1)

Here $g(\vec{x})$ is a given function that determines the Dirichlet condition on boundary $\Gamma_d$ and $h(\vec{x})$ is a given function that defines the Neumann condition at boundary $\Gamma_n$.

Remark:

Note that the Neumann condition can also be written as $\vec{\nabla} u \cdot \vec{n} = h$ with $\vec{n}$ the normal on $\Gamma_n$.

In order to derive a weak form of the diffusion problem (1), we define the space of trial functions that satisfy the Dirichlet boundary condition $u = g$ on $\Gamma_d$:

\[U = \{u|u \in H^2(\Omega), u = g \text{ on } \Gamma_d\} \cdot \]

Moreover we define a set of test functions that satisfy the homogeneous Dirichlet boundary conditions $u = 0$ on $\Gamma_d$:

\[W = \{w|w \in L^2(\Omega), u = 0 \text{ on } \Gamma_d\} \cdot \]
Remark:

$L^2(\Omega)$ is the space of functions endowed with the inner product

$$(u, w)_{L^2} = \int_{\Omega} u w d\Omega$$

that are square integrable, i.e.:

$L^2(\Omega) = \{ v | \int_{\Omega} v^2 d\Omega < \infty \}$

$H^2(\Omega)$ is the Hilbert space defined by:

$H^2(\Omega) = \{ v | \int_{\Omega} v^2 d\Omega < \infty \land \int_{\Omega} \nabla v \cdot \nabla v d\Omega < \infty \}$

The corresponding form of (1) is:

Find $u(\vec{x}) \in U$ such that:

$$- \int_{\Omega} \nabla u \cdot \nabla w d\Omega = \int_{\Omega} f w d\Omega \quad \forall w \in W \quad (2)$$

Partial integration of (2) yields the weak form:

$$\int_{\Omega} \nabla u \cdot \nabla w d\Omega = \int_{\Omega} f w d\Omega - \int_{\Gamma_n} \nabla u \cdot \vec{n} w d\Gamma \quad \forall w \in W \quad (3)$$

Substitution of the boundary condition on $\Gamma_n$ yields:

Find $u(\vec{x}) \in U$ such that:

$$\int_{\Omega} \nabla u \cdot \nabla w d\Omega = \int_{\Omega} f w d\Omega - \int_{\Gamma_n} h w d\Gamma \quad \forall w \in W \quad (4)$$
3 Discrete form

We now define a finite dimensional subspace $U^h \subset U$ such that:

$$U^h = \{ u^h | u^h = \sum_{i=1}^{N} u_i \phi_i, u^h = g \text{ on } \Gamma_d \}$$

and a finite dimensional subspace of test functions $W^h \subset W$ such that:

$$W^h = \{ w^h | w^h = \sum_{i=1}^{N} w_i \psi_i, w^h = 0 \text{ on } \Gamma_d \}.$$  

Here $\phi_i(\vec{x})$ are the basis functions of $U^h$ and $\psi_i(\vec{x})$ are the basis functions of $W^h$. The discrete version of the weak form (4) then reads:

$$\int_{\Omega} \sum_{i=1}^{N} u_i \nabla \phi_i \cdot \sum_{j=1}^{N} w_j \nabla \psi_j d\Omega = \int_{\Omega} \sum_{i=1}^{N} f_i \phi_i \sum_{j=1}^{N} w_j \psi_j d\Omega$$

$$- \int_{\Gamma_n} \sum_{i=1}^{N} h_i \phi_i \sum_{j=1}^{N} w_j \psi_j d\Gamma \quad \forall w \in W^h$$  \hspace{1cm} (5)

**Remark:**

Note that also the functions $f(\vec{x})$ and $h(\vec{x})$ are written in terms of the basis functions $\phi_i(\vec{x})$.

After introduction of:

$$u^T = [u_1, \ldots, u_N] \quad \phi^T = [\phi_1, \ldots, \phi_N]$$
$$w^T = [w_1, \ldots, w_N] \quad \psi^T = [\psi_1, \ldots, \psi_N]$$

the discrete set of equations (5) can be written as:

$$\int_{\Omega} u^T \nabla \phi \cdot w^T \nabla \psi d\Omega = \int_{\Omega} f^T \psi^T \phi d\Omega - \int_{\Gamma_n} h^T \phi^T \psi d\Gamma \quad \forall w \in W^h$$

and after some rearrangements:

$$w^T \int_{\Omega} (\nabla \psi) \cdot (\nabla \phi)^T d\Omega u = w^T \int_{\Omega} \psi^T \phi d\Omega f - w^T \int_{\Gamma_n} \psi^T d\Gamma h \quad \forall w \in W^h$$  \hspace{1cm} (6)
Lesson: The finite element method

With introduction of the matrices:
\[
K = \int_\Omega (\vec{\nabla} \psi) \cdot (\vec{\nabla} \phi)^T \, d\Omega, \quad M = \int_\Omega \psi \phi^T \, d\Omega, \quad M^\Gamma = \int_{\Gamma_n} \psi \phi^T \, d\Gamma
\] (7)
this yields:
\[
w^T K u = w^T M f - w^T M^\Gamma h \quad \forall w \in W_h
\]
Since this equation must hold for all \(w\) it is equivalent with
\[
K u = M f - M^\Gamma h
\] (8)
and can be solved if \(\phi\) and \(\psi\) are known and the integrals (7) are not too complex to compute. The solution \(u\) finally will provide the approximate solution \(u_h\) from:
\[
u_h(x) = \sum_{i=1}^N u_i \phi_i(x)
\]

4 Collocation methods

In collocation methods, like the finite element method, the basis functions \(\phi\) and \(\psi\) are restricted to functions that satisfy:
\[
\phi_i(x_j) = \delta_{ij}
\] (9)
with \(\delta_{ij}\) the Cronecker delta and \(x_j\) the collocation points. The advantage is that in this case the parameters \(u_i^h\) are equal to the approximate solution in the collocation points:
\[
u_h(x_j) = \sum_{i=1}^N u_i \phi_i(x_j) = u_j
\] (10)

Remark: A common way to define basis functions that satisfy the condition (9) is to make use of Lagrange interpolation polynomials through the collocation points. In one dimension this yields:
\[
\lambda_i(x) = \frac{\prod_{k \neq i} (x - x_k)}{\prod_{k \neq i} (x_i - x_k)}
\] (11)
with $\prod_{k=1}^{N} a_k = a_1 a_2 \ldots a_N$. In more dimensions combinations of $\lambda_i$ per direction can be used (see Baaijens, 1999)

5 Finite element methods

In order to be able to compute the integrals that appear in the matrices (8), in finite element methods, the domain $\Omega$ is decomposed into $N_e$ subdomains (elements) $\Omega_e$. These subdomains normally are chosen to be lines (in 1D), triangles or rectangles (in 2D) and tetrahedrons and bricks (in 3D). The integrations then can be carried out element by element according to:

$$\int_{\Omega} f d\Omega = \sum_{e=1}^{N_e} \int_{\Omega_e} f d\Omega_e$$

(12)

These element by element integrations can be performed relatively easy if the integrand is a simple function like a locally defined polynomial. To this end the basis functions $\phi_i$ and $\psi_i$ both are chosen to be polynomials $N_i(x)$ of degree smaller than or equal to $n$ defined in each element $\Omega_e$. Equivalent to (6) we obtain:

$$\sum_{e=1}^{N_e} w_e^T \int_{\Omega_e} (\bar{\nabla}N) \cdot (\bar{\nabla}N)^T d\Omega_e u_e = \sum_{e=1}^{N_e} w_e^T \int_{\Omega_e} N N^T d\Omega f_e$$

$$- \sum_{e=1}^{N_e} w_e^T \int_{\Gamma_e} N N^T d\Gamma h_e \quad \forall w \in W^h$$

(13)

With introduction of the element matrices:

$$K_e = \int_{\Omega_e} (\bar{\nabla}N) \cdot (\bar{\nabla}N)^T d\Omega_e, \quad M_e = \int_{\Omega_e} N N^T d\Omega_e, \quad M^T_e = \int_{\Gamma_e} N N^T d\Gamma$$

(14)

this yields

$$\sum_{e=1}^{N_e} w_e^T K_e u_e = \sum_{e=1}^{N_e} w_e^T M_e f_e - \sum_{e=1}^{N_e} w_e^T M^T_e h_e \quad \forall w \in W^h$$
Assembling all local element matrices and vectors in a global system yields:

\[ w^T K u = w^T M f - w^T M^\Gamma h \]  

(15)

In the appendix an example with two degrees of freedom is given.

6 Numerical integration

The element matrices (14) are not always as simple as given in this example and are normally computed by means of a numerical integration rule. To perform this, first the coordinates are transformed to the ones of a 'standard' element:

\[
\int_{\Omega_e} f(\vec{x}) d\Omega = \int_{S_e} f(\vec{\xi}) J(\vec{\xi}) dS
\]

(16)

with \( J(\vec{\xi}) \) the Jacobian of the transformation (see Baaijens, 1999). Finally the integrals are computed with a standard numerical integration (quadrature) rule:

\[ \int_{S_e} f(\vec{\xi}) J(\vec{\xi}) dS = \sum_{k=1}^{n_{int}} f(\vec{\xi}_k) J(\vec{\xi}_k) w_k \]

(17)

With \( \vec{\xi}_k \) the integration points and \( w_k \) the corresponding weight factors.

7 Boundary conditions

In order to incorporate the Dirichlet boundary conditions we must realize that \( w^T \) satisfies the homogeneous boundary conditions. If we partition the matrices according to:

\[
\begin{bmatrix}
  w^T_u & w^T_p \\
  K_{uu} & K_{up} \\
  K_{pu} & K_{pp}
\end{bmatrix}
\begin{bmatrix}
  u_u \\
  u_p
\end{bmatrix}
= \begin{bmatrix}
  w^T_u & w^T_p \\
  M_{uu} & M_{up} \\
  M_{pu} & M_{pp}
\end{bmatrix}
\begin{bmatrix}
  f_u \\
  f_p
\end{bmatrix}
- \begin{bmatrix}
  w^T_u & w^T_p \\
  M^\Gamma_{uu} & M^\Gamma_{up} \\
  M^\Gamma_{pu} & M^\Gamma_{pp}
\end{bmatrix}
\begin{bmatrix}
  h_u \\
  h_p
\end{bmatrix}
\]

(18)

this yields with \( w^T_p = 0^T \) and \( u_p \) known:

\[ K_{uu} u_u = M_{uu} f_u - M^\Gamma_{uu} h_u - K_{up} u_p \]

(19)

This system can be solved by either direct or iterative solvers.
8 Solution methods

8.1 Direct solvers

To obtain a solution of the linear systems that result from a finite element approximation as described above, we have to solve a linear system:

$$Au = f$$  \hspace{1cm} (20)

for the components of column \(u\) where all coefficients in the matrix \(A\) and column \(f\) are known. The most important technique for solving this system is Gauss elimination. First \(A\) is decomposed into a lower triangle \(L\) and upper triangle \(U\):

$$A = LU$$  \hspace{1cm} (21)

Second, the decomposed form is solved as:

$$Ly = f, \quad Uu = y$$  \hspace{1cm} (22)

Since the matrices \(L\) and \(U\) are triangular, only forward and backward substitution is needed to solve (22). There are different ways to compute the \(L\) and \(U\) matrices (Press et al., 1986). A disadvantage of direct methods is that originally sparse matrices (many zeros even in the band or profile of the matrix) result in \(L\) and \(U\) matrices that have the same profile but are not sparse anymore in the profile. Iterative solvers are more suitable in these cases.

8.2 Iterative solvers

There are many iterative solvers available to solve the linear system (20). The structure of these methods is also based on a decomposition but now:

$$A = P - Q$$  \hspace{1cm} (23)

where the preconditioner \(P\) is such that it can be factorized in an efficient way (for instance \(P\) is \(\text{diag}(A)\)). An iterative method then can be built according to:

$$Pu^{n+1} = Qu^n + f$$ \hspace{1cm} or

$$u^{n+1} = P^{-1}Qu^n + P^{-1}f$$ \hspace{1cm} or

$$u^{n+1} = u^n - P^{-1}e^n$$ \hspace{1cm} or
\[ P(u^{n+1} - u^n) = -r^n \]  

where \( r^n \) is the residual of the previous iteration step, i.e.:  
\[ r^n = A u^n - f \]  

It will be clear that the closer the preconditioner \( P \) resembles the original matrix \( A \) the faster this iterative method will converge, but the more difficult it will be to solve \( P(u^{n+1} - u^n) = -r^n \). There are many iterative solvers published that in one or another way try to find an optimized choice of the preconditioner \( P \) (see f.i. Canuto et al., 1988). Note that the matrix \( A \) never has to be inverted and even never has to be stored if the residual \( r = A u - f \) can be computed element by element.

**Appendix**

If we have the set of differential equations:
\[
\begin{bmatrix}
A_{11} & A_{12} \\
A_{21} & A_{22}
\end{bmatrix}
\begin{bmatrix}
u_1 \\
u_2
\end{bmatrix}
=
\begin{bmatrix}
f_1 \\
f_2
\end{bmatrix}
\]  

with \( A_{kl} \) differential operators. The weak form then reads:
\[
\begin{cases}
\int_{\Omega} w A_{11} u_1 d\Omega + \int_{\Omega} w A_{12} u_2 d\Omega = \int_{\Omega} w f_1 d\Omega \\
\int_{\Omega} w A_{21} u_1 d\Omega + \int_{\Omega} w A_{22} u_2 d\Omega = \int_{\Omega} w f_2 d\Omega
\end{cases}
\]  

After introduction of:
\[
w_k = \sum_{i=1}^{N} w_{k,i} \phi_i \quad u_k = \sum_{i=1}^{N} u_{k,i} \phi_i \quad f_k = \sum_{i=1}^{N} f_{k,i} \phi_i \quad k = 1, 2
\]  
\[
w = [w_1, \ldots, w_N]^T \quad u_k = [u_{k,1}, \ldots, u_{k,N}]^T \quad f_k = [f_{k,1}, \ldots, f_{k,N}]^T \quad k = 1, 2
\]  
this can be written as:
\[
\begin{cases}
\int_{\Omega} w^T \phi_i (A_{11}) \phi_i d\Omega + \int_{\Omega} w^T \phi_i (A_{12}) \phi_i d\Omega = \int_{\Omega} w^T \phi_i f_1 d\Omega \\
\int_{\Omega} w^T \phi_i (A_{21}) \phi_i d\Omega + \int_{\Omega} w^T \phi_i (A_{22}) \phi_i d\Omega = \int_{\Omega} w^T \phi_i f_2 d\Omega
\end{cases}
\]
Lesson: The finite element method

This must hold for all \( \tilde{w} \) so:

\[
\begin{align*}
\int_\Omega \phi(A_{11}\phi)^T d\Omega u_1 + \int_\Omega \phi(A_{12}\phi)^T d\Omega u_2 &= \int_\Omega \phi\phi^T d\Omega \tilde{f}_1 \\
\int_\Omega \phi(A_{21}\phi)^T d\Omega u_1 + \int_\Omega \phi(A_{22}\phi)^T d\Omega u_2 &= \int_\Omega \phi\phi^T d\Omega \tilde{f}_2
\end{align*}
\]  

or

\[
\begin{align*}
A^{11}u_1 + A^{12}u_2 &= M_{\tilde{f}_1} \\
A^{21}u_1 + A^{22}u_2 &= M_{\tilde{f}_2}
\end{align*}
\]

with:

\[A^{kl} = \int_\Omega \phi(A_{kl}\phi)^T d\Omega, \quad M = \int_\Omega \phi\phi^T d\Omega,\]

or in finite element context:

\[
\begin{align*}
A^{11}_{e}u_{1,e} + A^{12}_{e}u_{2,e} &= M_{e}\tilde{f}_{1,e} \\
A^{21}_{e}u_{1,e} + A^{22}_{e}u_{2,e} &= M_{e}\tilde{f}_{2,e}
\end{align*}
\]

with:

\[A_{e}^{kl}(i,j) = \int_{\Omega_e} \phi_i(A_{kl}\phi_j)d\Omega_e, \quad M_e = \int_{\Omega_e} \phi_i\phi_jd\Omega_e,\]

In the elements, however the numbering of the unknowns is:

\[\begin{bmatrix} u_{1,1}, u_{2,1}, u_{1,2}, u_{2,2}, \ldots, u_{1,n}, u_{2,n} \end{bmatrix}\]

So the element matrix is:

\[
\begin{align*}
A_{e}(p,q) &= \begin{cases} 
A_{e}^{11}(\frac{p+1}{2}, \frac{q+1}{2}) & p = \text{odd} \\
A_{e}^{12}(\frac{p}{2}, \frac{q+1}{2}) & p = \text{even}
\end{cases} \\
&\begin{cases} 
A_{e}^{21}(\frac{p+1}{2}, \frac{q}{2}) & q = \text{odd} \\
A_{e}^{22}(\frac{p}{2}, \frac{q}{2}) & q = \text{even}
\end{cases}
\end{align*}
\]
References

