Theory of Finite Element Method (FEM)

FEM formulation

$$\frac{d^2u}{dx^2} + 4u = 8x^2$$

subjected to boundary conditions

$$u(0) = u\left(\frac{\pi}{4}\right) = 0$$

using the weak formulation.

The first step in developing the weak formulation is to assume a weight function and a trial function. Take $U(x)$ as the trial function and $\phi(x)$ as the weight function. We discuss the exact forms of these functions later. The trial function $U(x)$ forms a solution to the ODE. Therefore, if we substitute $U(x)$ in (22), the resulting equation gives the residual:

$$R = \left|\frac{d^2U}{dx^2} + 4U - 8x^2\right| .$$

Subsequent steps really amount to the minimization of this residual. The minimization process starts by evaluating the *weighted residual*. To evaluate the weighted residual, multiply (24) by $\phi(x)$ and integrate over the domain (i.e. $0 \le x \le \pi/4$).

$$R(x) = \int_0^{\pi/4} \left(\phi \frac{d^2U}{dx^2} + 4\phi U - 8\phi x^2 \right) dx \,.$$

Using integration by parts, one can simplify above equation to obtain

$$R(x) = \int_0^{\pi/4} \left(4\phi U - \frac{d\phi}{dx}\frac{dU}{dx} \right) dx - 8 \int_0^{\pi/4} \phi x^2 dx \,.$$

To advance further we need to make some crucial assumptions. Since we are free to assign any function to $U(x)$ and $\phi(x)$ as far as they agree with the boundary conditions, we assume $U(x) = \phi(x)$. This is known as Galerkin's method. If $U(x) \neq \phi(x)$, then it gives the Rayleigh-Ritz formulation. We have to select an algebraic function of x to satisfy the boundary conditions $u(0) = u(\pi/4) = 0$.

$$\varphi(x) = U(x) = \phi(x) = u_1\varphi_1 + u_2\varphi_2 + \cdots + u_N\varphi_N = \sum_{i=0} u_i\varphi_i \,.$$

We assume N functions as follows.

$$\varphi_1 = x\left(\frac{\pi}{4} - x\right), \qquad \varphi_2 = x^2\left(\frac{\pi}{4} - x\right), \quad \ldots, \quad \varphi_N = x^N\left(\frac{\pi}{4} - x\right)$$

This selection satisfies the boundary conditions regardless the number of terms included in the series. Since $U(x) = \phi(x)$, the weighted residual becomes

$$R(x) = \int_0^{\pi/4} \frac{1}{2}\left(4\varphi^2 - \left(\frac{d\varphi}{dx}\right)^2\right) dx - 8\int_0^{\pi/4} \phi x^2 dx\,.$$

$$\frac{d\varphi}{dx} = \sum u_i \left(\frac{\pi}{4} i x^{i-1} - (i+1)x^i\right)$$

Therefore we have

$$R(u_i) = 2\sum u_i u_j \left(\frac{\pi}{4}\right)^{i+j+3}\left(\frac{1}{i+j+1} - \frac{2}{i+j+2} + \frac{1}{i+j+3}\right)$$

$$-\frac{1}{2}\sum u_j u_j \left(\frac{\pi}{4}\right)^{i+j+1}\left(\frac{ij}{i+j-1} - \frac{2ij+i+j}{i+j} + \frac{(i+1)(j+1)}{i+j+1}\right)$$

$$-8\sum u_i \left(\frac{\pi}{4}\right)^{i+4}\left(\frac{1}{i+3} - \frac{1}{i+4}\right), \quad i,j = 1,2,3,\ldots,N$$

Then we minimize the residual by taking derivatives of $\mathbf{R}$ w.r.t $\mathbf{u}_j$. For predetermined number N, this results in N algebraic equations that have to be solved simultaneously.

$$\frac{dR(u_i)}{du_j} = 0 .$$

For $N = 2$ there are only two unknowns; u_1 and u_2. It produces two linear equations.

$$0.122u_1 + 0.048u_2 = -0.120 ,$$
$$0.048u_1 + 0.033u_2 = -0.063 .$$

In matrix form

$$\begin{bmatrix} 0.122 & 0.048 \\ 0.048 & 0.033 \end{bmatrix} \begin{Bmatrix} u_1 \\ u_2 \end{Bmatrix} = \begin{Bmatrix} -0.120 \\ -0.063 \end{Bmatrix} .$$

It resembles the general form

$$[K]\{u\} = \{L\} ,$$

$$u_1 = -0.554579 ,$$
$$u_2 = -1.112560 .$$

Therefore the solution to (22) is

$$u(x) = -0.554579x\left(\frac{\pi}{4} - x\right) - 1.11256x^2\left(\frac{\pi}{4} - x\right)$$

If we go one step further by assuming $N = 3$, then we get three braic equations with three unknowns; u_1, u_2 and u_3. The resulting equation is

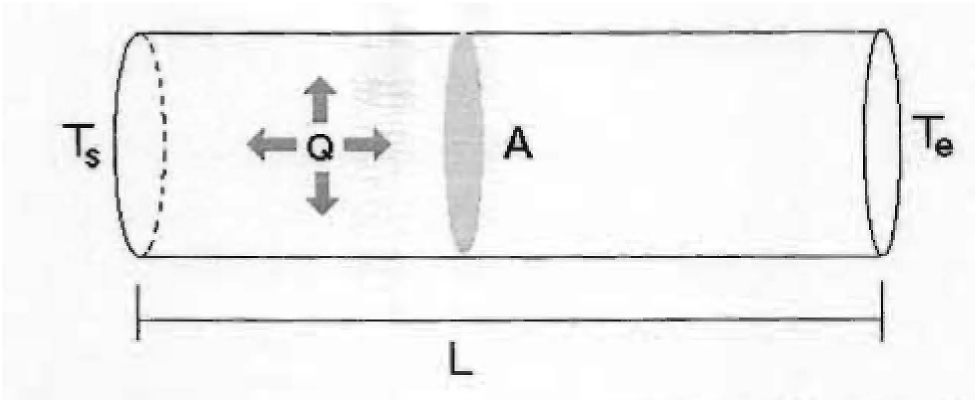
$$\begin{bmatrix} 0.1216 & 0.0477 & 0.0228 \\ 0.0477 & 0.0328 & 0.0200 \\ 0.0228 & 0.0200 & 0.0139 \end{bmatrix} \begin{Bmatrix} u_1 \\ u_2 \\ u_3 \end{Bmatrix} = \begin{Bmatrix} -0.120 \\ -0.630 \\ -0.351 \end{Bmatrix}.$$

$$u_1 = -0.588, \qquad u_2 = -0.838, \qquad u_3 = -0.349.$$

Therefore the new solution for (22) becomes

$$u(x) = -0.588x\left(\frac{\pi}{4} - x\right) - 0.838x^2\left(\frac{\pi}{4} - x\right) - 0.349x^3\left(\frac{\pi}{4} - x\right)$$

Axial heat transfer along an insulated rod



$$\frac{d}{dx}\left(k\frac{dT}{dx}\right)+Q=0; \qquad 0\leq x\leq 1$$

$$k = 3.3 \text{ J/°Cms},$$
$$Q = 10 \text{ J/sm},$$
$$T|_{x=0} = 1,$$
$$q|_{x=1} = 1.25 \text{ J/m}^2\text{s}.$$

Step 1. Variational formulation

This PDE is the strong form of the equation for heat conduction within a cylinder. The first step in FEM is to derive the weak form of the equations.

$$\int_0^1 w \left[\frac{d}{dx}\left(k\frac{dT}{dx}\right) + Q\right] dx = 0 .$$

Integrating by parts (using the divergence theorem in 1-D) we obtain

$$\int_0^1 \left(\frac{dw}{dk} k \frac{dT}{dx}\right) dx = \left[wk\frac{dT}{dx}\right]_0^1 + \int_0^1 wQdx .$$

From heat transfer theory, Fourier's law gives the heat flux across a unit cross section is given by Fourier's law $q = -k\frac{dT}{dx}$. Therefore,

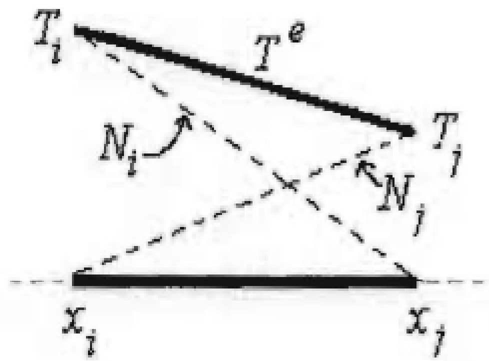
$$\int_0^1 \left(\frac{dw}{dk} k \frac{dT}{dx}\right) dx = -[wq]_0^1 + \int_0^1 wQdx$$

Step 2. Discretization and choice of polynomials

It is obvious that we are going to use 1-D elements. We can have simplex elements for simplicity, i.e. linear polynomials to approximate the unknowns.

$$T = a + bx$$

$$T_i = a + bx_i\,,$$
$$T_j = a + bx_j\,.$$



Solving for a and b gives

$$a = \frac{1}{l}(T_i x_j - T_j x_i)\,,$$

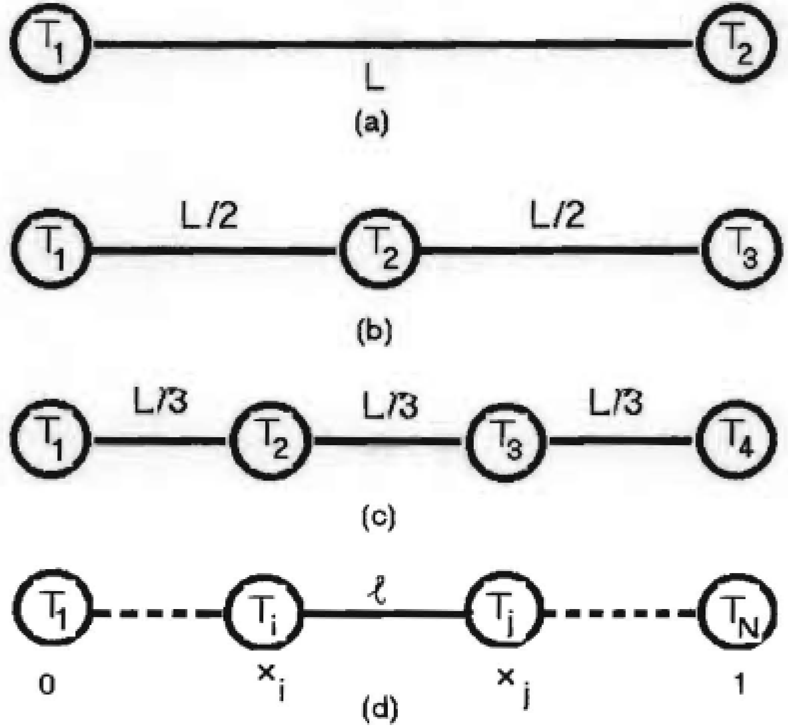
$$b = \frac{1}{l}(T_j - T_i)\,,$$

$$T^e = \frac{1}{l}(x_j - x)T_i + \frac{1}{l}(x - x_i)T_j$$

N_i and N_j are known as the shape functions
$N_i = 1$ at $x = x_i$ and $N_i = 0$ at $x = x_j$.
$N_j = 1$ at $x = x_j$ and $N_j = 0$ at $x = x_i$.

$$T^e = N_i T_i + N_j T_j$$

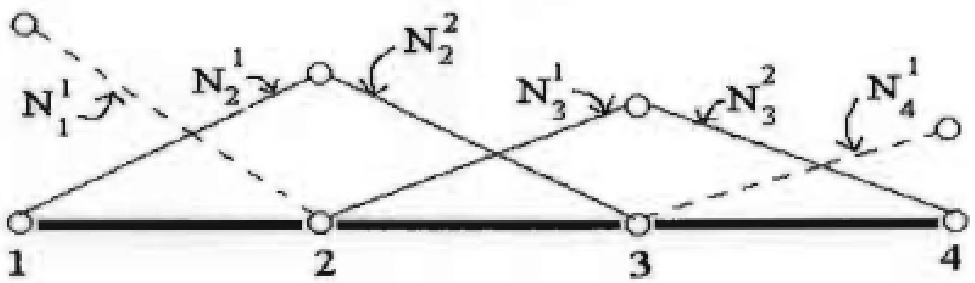
Discretization of cylinder



Global shape function

If we consider the first element there are two local shape functions: N_1^1 which is associated with node 1 and N_2^1 associated with node 2. For the second element again we have a local shape function associated with node 2 defined as N_2^2. Each global shape function is zero elsewhere except in the elements associated with the corresponding nodes. This enables us to define the global temperature variation.

$$T = \sum T^e = N_1 T_1 + \cdots + N_i T_i + \cdots N_N T_N = \sum_{j=1}^{N} N_j T_j$$



Step 3. Assembling the element equations to form the global problem

$$\underbrace{\int_0^1 \left[k\frac{d}{dx}\left(\sum N_i T_i\right)\frac{d}{dx}\left(\sum N_j T_j\right)\right]dx}_{\mathbf{I_k}}$$

$$= -\underbrace{\left[q\sum N_i T_i\right]_0^1}_{} + \underbrace{\int_0^1 q\sum N_i T_i dx}_{}$$

$$I_k(T)$$

$$= \int_{x_i}^{x_j}\left[\frac{d}{dx}N_1T_1\left(\frac{d}{dx}N_1T_1 + \frac{d}{dx}N_2T_2 + \frac{d}{dx}N_3T_3 + \frac{d}{dx}N_4T_4\right)\right]dx$$

$$+ \int_{x_i}^{x_j}\left[\frac{d}{dx}N_2T_2\left(\frac{d}{dx}N_1T_1 + \frac{d}{dx}N_2T_2 + \frac{d}{dx}N_3T_3 + \frac{d}{dx}N_4T_4\right)\right]dx$$

$$+ \int_{x_i}^{x_j}\left[\frac{d}{dx}N_3T_3\left(\frac{d}{dx}N_1T_1 + \frac{d}{dx}N_2T_2 + \frac{d}{dx}N_3T_3 + \frac{d}{dx}N_4T_4\right)\right]dx$$

$$+ \int_{x_i}^{x_j}\left[\frac{d}{dx}N_4T_4\left(\frac{d}{dx}N_1T_1 + \frac{d}{dx}N_2T_2 + \frac{d}{dx}N_3T_3 + \frac{d}{dx}N_4T_4\right)\right]dx$$

Minimization

In the minimization process we differentiate $\mathbf{I}_k(T)$ w.r.t. each T_i and set each derivative to zero. This procedure generates a number of equations equal to number of nodes.

$$\int_{x_i}^{x_j} \frac{dN_1}{dx}\frac{dN_1}{dx}T_1 dx + \int_{x_i}^{x_j} \frac{dN_1}{dx}\frac{dN_2}{dx}T_2 dx + \int_{x_i}^{x_j} \frac{dN_1}{dx}\frac{dN_3}{dx}T_3 dx + \int_{x_i}^{x_j} \frac{dN_1}{dx}\frac{dN_4}{dx}T_4 dx$$

Likewise, there will be three more equations. In matrix form it gives the stiffness matrix.

$$
\begin{aligned}
&\frac{d}{dT} I_k(T) \\
&= K \\
&= \begin{bmatrix}
\int_{x_i}^{x_j} k \frac{dN_1}{dx} \frac{dN_1}{dx} dx & \int_{x_i}^{x_j} k \frac{dN_1}{dx} \frac{dN_2}{dx} dx & \int_{x_i}^{x_j} k \frac{dN_1}{dx} \frac{dN_3}{dx} dx & \int_{x_i}^{x_j} k \frac{dN_1}{dx} \frac{dN_4}{dx} dx \\
\int_{x_i}^{x_j} k \frac{dN_2}{dx} \frac{dN_1}{dx} dx & \int_{x_i}^{x_j} k \frac{dN_2}{dx} \frac{dN_2}{dx} dx & \int_{x_i}^{x_j} k \frac{dN_2}{dx} \frac{dN_3}{dx} dx & \int_{x_i}^{x_j} k \frac{dN_2}{dx} \frac{dN_4}{dx} dx \\
\int_{x_i}^{x_j} k \frac{dN_3}{dx} \frac{dN_1}{dx} dx & \int_{x_i}^{x_j} k \frac{dN_3}{dx} \frac{dN_2}{dx} dx & \int_{x_i}^{x_j} k \frac{dN_3}{dx} \frac{dN_3}{dx} dx & \int_{x_i}^{x_j} k \frac{dN_3}{dx} \frac{dN_4}{dx} dx \\
\int_{x_i}^{x_j} k \frac{dN_4}{dx} \frac{dN_1}{dx} dx & \int_{x_i}^{x_j} k \frac{dN_4}{dx} \frac{dN_2}{dx} dx & \int_{x_i}^{x_j} k \frac{dN_4}{dx} \frac{dN_3}{dx} dx & \int_{x_i}^{x_j} k \frac{dN_4}{dx} \frac{dN_4}{dx} dx
\end{bmatrix}
\end{aligned}
$$

$$f_b = -\begin{bmatrix} qN_1|_0^1 \\ qN_2|_0^1 \\ qN_3|_0^1 \\ qN_4|_0^1 \end{bmatrix} \quad \text{and } f_s = \begin{bmatrix} \int_0^1 QN_1 dx \\ \int_0^1 QN_2 dx \\ \int_0^1 QN_3 dx \\ \int_0^1 QN_4 dx \end{bmatrix}.$$

The compact equation is $[K]\{x\} = \{L\}$ where $\mathbf{F} = \mathbf{f}_b + \mathbf{f}_s$. The column matrix $\mathbf{f}_b$ contains the boundary terms and $\mathbf{f}_s$ contain the source terms. $\mathbf{x}$ is the vector of unknowns (nodal temperatures in our case). Components in $\mathbf{L}$, $\mathbf{f}_b$ and $\mathbf{f}_s$ have to be evaluated elementwise.

Step 4. Numerical manipulation

As we formulated the global problem in Step 3, the rest is down to matrix manipulation to evaluate the unknowns. As the first step we have to evaluate the components K_{mn} of stiffness matrix $\mathbf{K}$. K_{1n} corresponds to node 1. Therefore N_1^1 and N_2^1 are the only nonzero global shape functions.

$$K_{11} = \int_0^{0.33} k \frac{dN_1^1}{dx} \frac{dN_1^1}{dx} dx$$

$$= \int_0^{0.33} 3.3 \left(-\frac{1}{0.33}\right) \left(-\frac{1}{0.33}\right) dx = 10\,,$$

$$K_{12} = \int_0^{0.33} k \frac{dN_1^1}{dx} \frac{dN_2^1}{dx} dx$$

$$= \int_0^{0.33} 3.3 \left(-\frac{1}{0.33}\right) \left(\frac{1}{0.33}\right) dx = -10\,,$$

$$K_{13} = K_{14} = 0\,.$$

In evaluating terms in the second row we immediately make use of the symmetry of the matrix.

$$K_{21} = K_{12} = -10\,.$$

Upon evaluating the K_{22} we run into a problem — which shape function to use N_2^1 or N_2^2? The solution is simple. Since those two functions are defined over two elements, we have to integrate relevant function over the appropriate element considering the limits from 0 to 0.66 (or more generally $2l$).

$$\begin{aligned}
K_{22} = \int_0^{0.66} k\frac{dN_2}{dx}\frac{dN_2}{dx}dx &= \int_0^{0.33} k\frac{dN_2^1}{dx}\frac{dN_2^1}{dx}dx + \int_{0.33}^{0.66} k\frac{dN_2^2}{dx}\frac{dN_2^2}{dx}dx \\
&= \int_0^{0.33} 3.3\left(\frac{1}{0.33}\right)\left(\frac{1}{0.33}\right)dx + \int_{0.33}^{0.66} 3.3\left(-\frac{1}{0.33}\right)\left(-\frac{1}{0.33}\right)dx \\
&= 20\,.
\end{aligned}$$

K_{23} involves functions defined only over the second element. Hence N_2^2 and N_3^1 are to be considered.

$$K_{23} = \int_{0.33}^{0.66} k \frac{dN_2^2}{dx} \frac{dN_3^1}{dx} dx$$
$$= \int_{0.33}^{0.66} 3.3 \left(-\frac{1}{0.33}\right) \left(\frac{1}{0.33}\right) dx = -10\,.$$

$K_{24} = 0$ since N_4 does not share node 2. $K_{31} = 0$ according to the same line of reasoning. $K_{32} = K_{23} = -10$ and K_{34} is to be evaluated in the same manner we evaluated K_{23}. Again $K_{41} = K_{42} = 0$ as the shape functions do not share the node in question. $K_{43} = K_{34}$ by symmetry. K_{33} and K_{44} are to be evaluated in the same way we evaluated the K_{22}, considering the relevant shape function over the relevant domain. The completed stiffness matrix is given below.

$$K = \begin{bmatrix} 10 & -10 & 0 & 0 \\ -10 & 20 & -10 & 0 \\ 0 & -10 & 20 & -10 \\ 0 & 0 & -10 & 10 \end{bmatrix}$$

This is the famous tridiagonal matrix in FEM. In this case it is only 4×4 since we have only four nodes. With full modeling, one would get a huge, sparse matrix of few thousands of components, yet still banded.

The next step is to evaluate the components in $\mathbf{f}_b$ and $\mathbf{f}_s$. In evaluating the terms in $\mathbf{f}_b$ it is important to identify only N_1 and N_4 remain nonzero at $x = 0$ and $x = 1$. In fact $N_1 = N_4 = 1$ at $x = 0$ and $x = 1$. All other shape functions are zero as far as start and end points are concerned. Therefore

$$f_b = -\begin{bmatrix} qN_1|_0^1 \\ qN_2|_0^1 \\ qN_3|_0^1 \\ qN_4|_0^1 \end{bmatrix} = \begin{bmatrix} q_{x=0} \\ 0 \\ 0 \\ -1.25 \end{bmatrix}.$$

In computing the terms in $\mathbf{f}_s$, the integrals are to be evaluated taking into consideration where the global shape functions are defined.

$$f_s = \begin{bmatrix} \int_0^1 QN_1 dx \\ \int_0^1 QN_2 dx \\ \int_0^1 QN_3 dx \\ \int_0^1 QN_4 dx \end{bmatrix} = \begin{bmatrix} \int_0^{0.33} QN_1^1 dx \\ \int_0^{0.33} QN_2^1 dx + \int_{0.33}^{0.66} QN_2^2 dx \\ \int_{0.33}^{0.66} QN_2^1 dx + \int_{0.66}^{1.0} QN_2^2 dx \\ \int_{0.66}^{1.0} QN_4^1 dx \end{bmatrix} = \begin{bmatrix} 1.65 \\ 3.3 \\ 3.3 \\ 1.65 \end{bmatrix}$$

Matrix equation and solution

$$\begin{bmatrix} 10 & -10 & 0 & 0 \\ -10 & 20 & -10 & 0 \\ 0 & -10 & 20 & -10 \\ 0 & 0 & -10 & 10 \end{bmatrix} \begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \end{bmatrix} = \begin{bmatrix} q_{x=0} \\ 0 \\ 0 \\ -1.25 \end{bmatrix} + \begin{bmatrix} 1.65 \\ 3.3 \\ 3.3 \\ 1.65 \end{bmatrix}$$

$$\begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1.7 \\ 2.01 \\ 2.11 \end{bmatrix}.$$