JASS 05

Seminar: Interplay of Mathematical Modeling and Numerical Simulation

An Introduction to Finite Element Methods

Niko Manopulo

May 4, 2005

Abstract

The Finite Element Methods (FEM) are nowadays one of the most frequently used computational methods in solving scientific and engineering problems. This success is mainly due to the fact that FEM are able to reflect the original mathematical model in a very natural way. This paper aims to navigate through different points of view towards FEM in an introductory level, by trying to make clear the strict connection between the mathematical model and the FEM discretization. A simple example from mechanics is selected and solved by using the rigorous mathematical formulation of FEM.

PART I Introduction and Basic Concepts

1 Computational Methods

With the aid of increasing computational power of microprocessors and parallel and distributed systems, science and engineering is more and more based on computer simulations. Scientists and engineers model reality using mathematical tools and then use computers to compute solutions for the given problem. This process can be summarized in the following picture.



Fig. 1. A simplified view of the physical simulation process [1]

The first step of the simulation process is *idealization*. Scientists and engineers spend considerable effort on analyzing the physical system and trying to discover mathematical relationships that describe its behavior. This most of the time results in ordinary and partial differential equations (e.g. Navier – Stokes Equations). The solution of these equations is most of the time impossible to carry out with analytical methods therefore they have to be mapped from a continuous (infinite dimensional) space into a discrete (finite dimensional) space. This process is called *discretization* and FEM enter the picture during this stage. Once the continuous model is mapped to its discrete counterpart, the *solution* of the system can be found with methods for solving linear systems of equations.

2 The Finite Element Methods

The Finite Element Methods were first invented by structural engineers, who based themselves on a strictly physical basis. However mathematicians later discovered that FEM methods could be classified as a subset of the Galerkin Methods for the solution of PDE's. This way the method gained a broader mathematical foundation which extended its use to many engineering problems. Nevertheless this difference in the engineering and mathematics points of view resulted in two different interpretations which also affects the way the method is used in practice.

Physical Interpretation:

The continuus physical model is divided into finite pieces called elements and laws of nature are applied on the generic element. The results are then recombined to represent the continuum.

Mathematical Interpretation:

The differentiational equation representing the system is converted into a variational form and solved by the linear combination of a finite set of trial functions.

2.1 FEM Notation

As the name suggests the FEM treat the continuous problem domain as a collection of individual *finite elements*. The problem parameters are defined on each of the nodes of a typical element. Let us now have a look to the key definitions of the FEM notation.

- *Dimensionality:* The elements can be defined differently depending on the problem context. Dimensionality indeed expresses wether the element has 1, 2 or 3 space dimensions.
- *Nodal Points:* Every element is described by its nodal points. Frequently the nodal points are chosen to be the corners of the element. However in case of non linear geometries nodal points are also defined on the edges.
- *Geometry:* This term is used to describe the domain on which finite element discretization needs to be applied. It can be smooth an regular (e.g. a rectangular plate), or complex (e.g. surface of a machine part). The geometry is defined by the placements of the nodal points.



Fig. 2. Typical Finite Element Geometires [1]

• Degrees of Freedom: The degree of freedom is the number of ways in which the original problem domain can change its state. In the case of the continous problem domain, the DOF is infinite, because problem characteristics can be defined in each point on the domain. In the discrete FEM domain, instead, the DOF is limited by the number elements, because problem characteristics can only be defined on the nodal points.

• *Nodal Forces:* A set of nodal forces (or any other actions depending on the problem) are defined on each nodal point. From the mathematical point of view this corresponds to the non-homogeneous right hand side of the governing DE.

3 Mechanical Approach

As hinted in the introductory part of this paper, there are different approaches to the setup of a Finite Element Method. In this section we will describe a simple mechanical problem, aiming to derive the discretized FEM equations by using the principles of Mechanics of Materials (MoM). We decided to start discovering FEM this way, because most of the terms and concepts in the mathematical formulation (which we will treat in the next part), find its origins in concepts of MoM. Furthermore understanding the connection between the physical model and the mathematical formulation is very valuable, as it enables us to have bigger control on the problem treated.



3.1 Formulation of a Bar Member

Fig. 3. A fixed-free bar member and relative notation [2]

In this section we will treat a one dimensional bar member, fixed at one end and free on the other (see Fig. 3). The member is axially loaded by a distributed load q(x) and point end load P. All the parameters shown on the picture are givens except for the axial strain u(x) which is the unknown. The Principle of Virtual Work will be applied for the solution of the problem. Let us start by defining the *internal strain energy* of the member.

3.1.1 Strain Energy

The stress strain relationship in a bar is given by the Hooke's Law:

$$\sigma(x) = E\varepsilon(x) \tag{1}$$

Where:

$$\varepsilon(x) = \frac{du}{dx} = u' \tag{2}$$

The strain energy density is defined as follows:

$$\Theta = \frac{1}{2}\sigma(x)\varepsilon(x) \tag{3}$$

To obtain the strain energy we must integrate the energy density over the whole volume. Furthermore we can use the definitions we made in Fig. 3 to obtain the following relationship between the unknown u and strain energy U.

$$U = \frac{1}{2} \int_{V} \sigma \varepsilon \, dV = \frac{1}{2} \int_{0}^{L} p \varepsilon \, dx = \frac{1}{2} \int_{0}^{L} (EAu')u' \, dx$$

$$U = \frac{1}{2} \int_{0}^{L} u' EAu' \, dx$$

$$\tag{4}$$

Note that the *rigidity* EA is most of the time constant and can be taken out of the integral. However as we are treating the general case, E(x) and A(x) can be functions of x. The *external work* on the system can be defined in a similar way.

3.1.2 External Work

Work is defined as *Force* x *Displacement*. Therefore we have to identify the external forces on the system and integrate their product with the displacement over the whole volume. Two kinds of external forces are acting on the considered system.

- 1. The distributed load q(x)
- 2. The point end load P.

The distributed load q(x) is continous and therfore integrable. However the end point load P is singular and it has to be treated with the delta dirac function. Nevertheless for our purposes it can be included in q(x) keeping this way the setup of the problem simple.

The *external work* is defined as follows:

$$W = \int_0^L q u \, dx \tag{5}$$

3.1.2 The Minimum Potential Energy Principle

For a system to be in equilibrium internal energy and external energy have to equal each other. However as we will not be treating the original continuous system but the discrete approximation, this equality is unlikely to hold. Instead we will define the following *Total Potential Energy* (TPE) functional which will have to be minimized.

$$\Pi = U - W$$
or
$$\Pi[u(x)] = U[u(x)] - W[u(x)]$$
(6)

According to the principles of variational calculus, a functional is minimized when its variation equals to zero.

$$\partial \Pi = \delta U - \delta W = 0 \tag{7}$$

3.1.3 TPE Discretization

Equation (7) can be used for the derivation of the Finite Element equations. Let us first discretize our problem domain into five equal bars as shown in **Fig. 4**.



Fig. 4. FEM discretization of bar member[2].

Functionals are scalars therefore the TPE can be written as the sum of individual members

$$\Pi = \Pi^{(1)} + \Pi^{(2)} + \Pi^{(3)} + \dots + \Pi^{(N)}$$
(8)

Taking the variation with respect to u:

$$\partial \Pi = \partial \Pi^{(1)} + \partial \Pi^{(2)} + \partial \Pi^{(3)} + \dots + \partial \Pi^{(N)} = 0$$
⁽⁹⁾

For equation (9) to be true each term must be zero. Therfore the following equation can be written for the generic element denoted by e.

$$\partial \Pi^{(e)} = \delta U^{(e)} - \delta W^{(e)} = 0 \tag{10}$$

The next step consists of deriving the FEM equations using the already derived integrals for U and W. However we want to stop here with the mechanical point of view to switch to the Mathematical Formulation, where we will derive similar equations in a more general manner and consider their solution.

PART II Mathematical Formulation

4 Weighted Residual Methods

The class of differential equations containing also the one dimensional bar described above are given as follows :

$$L[u] = -\frac{d}{dx}(p(x)\frac{du}{dx}) + z(x)u = q(x), \qquad 0 < x < 1$$
(11)

$$u(0) = u(1) = 0.$$

It follows that:

$$L[u] - q = 0 \tag{12}$$

Multiplying this by a *weight function* v and integrating over the whole domain we obtain:

$$\int_{0}^{1} (L[u] - q)v \, dx = (v, L[u] - q) = 0$$
(12a)

For the inner product to exist v must be "square integrable" Therefore:

$$v \in L^2(0,1) \tag{13}$$

Equation (12a) is called *variational form*.

We can replace u and v in the formula with their approximation function i.e.

$$u(x) \approx U(x) = \sum_{j=1}^{N} c_j \phi_j(x)$$

$$v(x) \approx V(x) = \sum_{j=1}^{N} d_j \psi_j(x)$$
(14)

The functions ϕ_j and ψ_j are of our choice and are meant to be suitable to the particular problem. For example the choice of sine and cosine functions satisfy boundary conditions hence it could be a good choice. U is called *trial function* and V is called *test function*. Because the differential operator L[u] is second order, u and consequently U must be two times continuously differentiable.

$$u \in C^2(0,1) \Rightarrow U \in C^2(0,1)$$

Therefore we can see U as an element of a finite-dimensional subspace S^N of the infinitedimensional function space $C^{e}(0,1)$.

$$U \in S^N(0,1) \subset C^2(0,1)$$

The same way

$$V \in \hat{S}^{N}(0,1) \subset L^{2}(0,1)$$

Now we can replace u and v in equation (12a) with their approximations U and V to obtain

$$(V,r) = 0, \qquad \forall V \in \hat{S}^{N}$$

$$r(x) = L[U] - q \qquad (15)$$

r(x) is called the residual and it is always orthogonal to V. This is because their inner product is equal to zero.

4.1 Galerkin's Method

If we substitute the approximate test function V of equation 14 into equation 12a and exchanging summation and integration we obtain

$$\sum_{j=1}^{N} d_{j}(\psi_{j}, L[U] - q) = 0 \quad \forall d_{j}, \qquad j = 1, 2, ..., N$$
(16)

By definition the above equation should be satisfied for all choices of d_j . This is only possible if the inner product valses

$$(\psi_j, L[U] - q) = 0$$
 $j = 1, 2, ..., N$ (17)

The next step is finding a suitable test function. One obvious choice would be taking the test functions equal the trial functions. In this case the equation 17 would become

$$(\phi_j, L[U] - q) = 0$$
 $j = 1, 2, ..., N$ (18)

This method was proposed by Galerkin and therefore it is called Galerkin's Method. The eq. (18) is also called the *strong form* because the test space in this case has more continouity than necessary.

4.2 The Weak Form

As said in the previous section the solution space of the problem has more continouity than necessary. Furthermore the strong form is not symmetrical which is physically contradicting. To resolve these issues one can integrate the equation by parts to obtain

$$\int_{0}^{1} v[-(pu')'+zu-q]dx = \int_{0}^{1} (v' pu'+vzu-vq)dx - vpa' \Big|_{0}^{*} = 0$$

The last term is equal to zero because of the boundary conditions. This new form of the problem can be rewritten as

$$A(v,u) - (v,q) = 0$$
(19)

with

$$A(v,u) = \int_0^1 (v' pu' + vzu) dx$$

Integrating by parts we eliminated the second derivatives of u from the problem. This way we now need less continuity then before. This is why eq. (19) is called the *weak form*. A(v,u) is commonly called *Strain Energy* as it corresponds to the energy stored in mechanical systems due to deformation.

The integration however also changed the space in which v has to be defined. In fact now that v' is introduced, there is a need for more continuity than just L^2 . For keeping the symmetry the solution space can be defined as bounded by the values of

$$A(u, u) = \int_{0}^{1} (p(u')^{2} + zu^{2}) dx$$

As p and z are necessarily smooth functions the above requirement can be translated to the following

$$\int_{0}^{1} (u')^2 + u^2 \, dx$$

Functions obeying this rule belong to the so called *Sobolev Space* and they are denoted by H_i . We also require v and u to satisfy boundary conditions so we denote the resulting space as H_0^1 . With this remark the final form of the problem becomes

$$A(v,u) = (v,q) \qquad \forall v \in H_0^1$$
(20)

Finally substituting again the approximate solutions U and V we obtain:

$$U, V \in S_0^N \subset H_0^1$$
$$A(V, U) = (V, q) \qquad \forall V \in S_0^N$$

More explicitly substituting U and V (remember we chose them to have the same base) and swapping summations and integrals we get to the final stage which are the Finite Element Equations.

$$\sum_{k=1}^{N} c_k A(\phi_j, \phi_k) = (\phi_j, q), \qquad j = 1, 2, ..., N$$
⁽²¹⁾

As it can be seen this a linear system of equations with the only unknown being c_k . The equations can be solved using any method for solving LSE and the approximation to u will be found by constructing U with the given parameters c_k .

4.3 Connection to the physical model

The system of equation (21) concludes the mathematical derivation of the FEM equations. Let us now turn back and compare the results of our mathematical derivation with the results obtained in the mechanical approach.

4.3.1 Strain Energy:

We had derived in section 3 the strain energy as

$$U = \frac{1}{2} \int_0^L u' E A u' dx$$

The mathematical formulation instead produced

$$A(v,u) = \int_0^1 (v' pu' + vzu) dx$$

At first sight the two expressions seem to be different. However a closer look reveals that they actually lead to the same concept. The first thing to note is that the starting point of the two different analyses is not the same. In fact for sake of generality the mathematical derivation has been carried out with a general differential equation describing many different situations including the axial strain of a bar. The parameter z in fact has no counterpart in the mechanical derivation, as it stands for internal sources of energy (e.g. heat). The parameter p stands for the resistance of the body to deformation which is translated in the mechanical approach by EA. Finally as FEM uses the Galerkin Method uand v are chosen from the same space. It is easy to see, with these information, that the two expressions are equivalent.

4.3.2 External Work:

The two different expressions obtained for the external work are clearly equivalent if v is used instead of u.

$$W = \int_0^L qu \, dx = (q, u)$$

4.3.3 The Weak Form:

Finally comparing the weak form equations obtained by the two analysis we clearly see that they describe the same main principle: "The variation of internal energy has to be equal to the variation of the external work".

$$\delta \Pi = \delta U - \delta W = 0$$
 $A(v, u) - (v, q) = 0$

With this comparison we conclude the second part of this paper which was dedicated to the mathematical derivation of the FEM equations. In the next section we will discuss an example and we will transform the already derived FEM equations into matrix equations.

PART III Matrix Form of the Problem

5 Finite Element Discretization

In this section we will try to solve an example ODE by using the Finite Element Method. Although the example is quite simple in nature it contains many of the fundamental FEM concepts.

Let us take the initial value problem with constant coefficients

$$-pu''+zu = q(x), \qquad 0 < x < 1$$

 $p, z \ge 0$
 $u(0) = u(1) = 0.$

and divide its domain into N subintervals

$$0 = x_0 < x_1 < \ldots < x_N = 1$$

Each subinterval $(x_{j-1}, x_j), j = 1: N$ is called a *finite element*.

The trial functions can be chosen arbitrarily as long as they fulfill the requirements of the space in which they are defined, however it is wise to select a function basis with *local* support. This means that the functions should get the value 1 in the jth node and 0 in all

other nodes. As we will see this way we obtain tridiagonal matrices instead of dense matrices, which make the solution of the problem less computationally intensive. A suitable trial basis for this problem is the linear *hat function* defined as follows:



Linear combinations of these functions construct the unknown approximation U(x)

$$U(x) = \sum_{j=1}^{N} c_j \phi_j(x)$$

Let us now go back to the weak form derived in PART II and substitute this trial basis into the equation

$$\sum_{k=1}^{N} c_k A(\phi_j, \phi_k) = (\phi_j, q), \qquad j = 1, 2, ..., N$$

We should now consider a typical finite element and derive the matrix equations for it. Afterwards we will combine the single elements to obtain the global matrix equations. The approximation function U(x), on a typical element, takes the following form:

$$U(x) = c_{j-1}\phi_{j-1}(x) + c_j\phi_j(x) \qquad x \in [x_{j-1}, x_j]$$

Written in matrix notation the above equation looks like

$$U(x) = [c_{j-1} \ c_j] \begin{bmatrix} \phi_{j-1}(x) \\ \phi_j(x) \end{bmatrix} = [\phi_{j-1}(x) \ \phi_j(x)] \begin{bmatrix} c_{j-1} \\ c_j \end{bmatrix} \qquad x \in [x_{j-1}, x_j]$$

The same can be done also for V(x)

$$V(x) = \begin{bmatrix} d_{j-1} & d_j \end{bmatrix} \begin{bmatrix} \phi_{j-1}(x) \\ \phi_j(x) \end{bmatrix} = \begin{bmatrix} \phi_{j-1}(x) & \phi_j(x) \end{bmatrix} \begin{bmatrix} d_{j-1} \\ d_j \end{bmatrix} \qquad x \in [x_{j-1}, x_j]$$

The formula requires derivatives of U and V which can be computed as follows

$$U'(x) = \begin{bmatrix} c_{j-1} & c_j \end{bmatrix} \begin{bmatrix} -1/h_j \\ 1/h_j \end{bmatrix} = \begin{bmatrix} -1/h_j & 1/h_j \end{bmatrix} \begin{bmatrix} c_{j-1} \\ c_j \end{bmatrix} \qquad x \in [x_{j-1}, x_j]$$
$$h_j = x_j - x_{j-1}$$
$$V'(x) = \begin{bmatrix} d_{j-1} & d_j \end{bmatrix} \begin{bmatrix} -1/h_j \\ 1/h_j \end{bmatrix} = \begin{bmatrix} -1/h_j & 1/h_j \end{bmatrix} \begin{bmatrix} d_{j-1} \\ d_j \end{bmatrix} \qquad x \in [x_{j-1}, x_j]$$

The calculation of the integrals is rather complicated; it is therefore worthwhile to divide the formula into smaller manageable parts by defining

$$\sum_{j=1}^{N} [A_{j}(V,U) - (V,q)_{j}] = 0$$

$$A_{j}(V,U) = A_{j}^{S}(V,U) + A_{j}^{M}(V,U)$$

$$A_{j}^{S}(V,U) = \int_{x_{j-1}}^{x_{j}} pV'U'dx$$

$$A_{j}^{M}(V,U) = \int_{x_{j-1}}^{x_{j}} zVUdx$$

$$(V,q) = \int_{x_{j-1}}^{x_{j}} Vqdx$$
ation of $A_{j}^{S}(V,U)$

Let us start with the computation of $A_j^s(V,U)$

$$A_{j}^{S}(V,U) = \begin{bmatrix} d_{j-1} & d_{j} \end{bmatrix} \left(\int_{x_{j-1}}^{x_{j}} p \begin{bmatrix} -1/h_{j} \\ 1/h_{j} \end{bmatrix} \begin{bmatrix} -1/h_{j} & 1/h_{j} \end{bmatrix} dx \begin{bmatrix} c_{j-1} \\ c_{j} \end{bmatrix} \right)$$
$$A_{j}^{S}(V,U) = \begin{bmatrix} d_{j-1} & d_{j} \end{bmatrix} \left(\int_{x_{j-1}}^{x_{j}} \frac{p}{h_{j}^{2}} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} dx \begin{bmatrix} c_{j-1} \\ c_{j} \end{bmatrix} = \begin{bmatrix} d_{j-1} & d_{j} \end{bmatrix} K_{j} \begin{bmatrix} c_{j-1} \\ c_{j} \end{bmatrix}$$

$$K_j = \frac{p}{h_j} \begin{bmatrix} -1 & 1\\ 1 & -1 \end{bmatrix}$$

The matrix K_j is called *element stiffness matrix*.

In a similar way one can obtain $A_j^M(V,U)$

$$A_{j}^{M}(V,U) = \begin{bmatrix} d_{j-1} & d_{j} \end{bmatrix} \left(\int_{x_{j-1}}^{x_{j}} z \begin{bmatrix} \phi_{j-1} \\ \phi_{j} \end{bmatrix} \begin{bmatrix} \phi_{j-1} & \phi_{j} \end{bmatrix} dx \right) \begin{bmatrix} c_{j-1} \\ c_{j} \end{bmatrix}$$
$$A_{j}^{S}(V,U) = \begin{bmatrix} d_{j-1} & d_{j} \end{bmatrix} M_{j} \begin{bmatrix} c_{j-1} \\ c_{j} \end{bmatrix}$$
$$M_{j} = \frac{zh_{j}}{6} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

The matrix M_j is called *element mass matrix*.

The external work matrix cannot be evaluated for an arbitrary function q, we will therefore use a linear interpolant of q.

$$q(x) \approx q_{j-1}\phi_{j-1}(x) + q_j\phi_j(x), \qquad x \in [x_{j-1}, x_j]$$

Using the above approximation the integral is readily computed to give

$$(V,q)_{j} \approx \int_{x_{j-1}}^{x_{j}} [d_{j-1},d_{j}] \begin{bmatrix} \phi_{j-1} \\ \phi_{j} \end{bmatrix} [\phi_{j-1},\phi_{j}] \begin{bmatrix} q_{j-1} \\ q_{j} \end{bmatrix} dx = [d_{j-1},d_{j}] l_{j}$$
$$l_{j} = \frac{h_{j}}{6} \begin{bmatrix} 2q_{j-1}+q_{j} \\ q_{j-1}+2q_{j} \end{bmatrix}$$

where l_j is called *element load vector*.

Now the task is to assemble the elements into the whole system in fact we have to sum each integral over all the elements. For doing so we can extend the dimension of each element matrix to N and then put the 2x2 matrix at the appropriate position inside it. With all matrices and vectors having the same dimension the summation looks like

$$\sum_{j=1}^{N} A_{j}^{S} = \mathbf{d}^{T} \mathbf{K} \mathbf{c} \quad c = \begin{bmatrix} c_{1} \\ c_{2} \\ \dots \\ c_{N-1} \end{bmatrix} \qquad d = \begin{bmatrix} d_{1} \\ d_{2} \\ \dots \\ d_{N-1} \end{bmatrix} \qquad \mathbf{K} = \frac{p}{h} \begin{bmatrix} 2 & -1 \\ -1 & 2 & -1 \\ & -1 & 2 & -1 \\ & & \dots & \dots \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{bmatrix}$$

Doing the same for the Mass Matrix and for the Load Vector

$$\sum_{j=1}^{N} A_{j}^{M} = \mathbf{d}^{\mathsf{T}} \mathbf{M} \mathbf{c} \qquad \mathbf{M} = \frac{zh}{6} \begin{bmatrix} 4 & 1 & & \\ 1 & 4 & 1 & & \\ & 1 & 4 & 1 & \\ & & 1 & 4 & 1 \\ & & & 1 & 4 \end{bmatrix}$$
$$\sum_{j=1}^{N} (V, q)_{j} = \mathbf{d}^{\mathsf{T}} \mathbf{l} \qquad \mathbf{I} = \frac{h}{6} \begin{bmatrix} q_{0} + 4q_{1} + q_{2} & & \\ q_{1} + 4q_{2} + q_{3} & & \\ & & \\ q_{N-2} + 4q_{N-1} + q_{N} \end{bmatrix}$$

Substituting this Matrix form of the expressions in

$$\sum_{j=1}^{N} [A_{j}(V,U) - (V,q)_{j}] = 0$$

we obtain the following set of linear equations

$$\mathbf{d}^{\mathrm{T}}[(\mathbf{K}+\mathbf{M})\mathbf{c}-\mathbf{l}]=\mathbf{0}$$

This has to be satisfied for all choices of d therefore which means that the expression in the squared brackets has to vanish

$(\mathbf{K}+\mathbf{M})\mathbf{c}-\mathbf{l}=\mathbf{0}$

This last matrix equation rapresents the linear system of equations which has to be solved in order to get the coefficients for the approximate solution U(x).

6 References

PART I is based on [1] and [2]. PART II and PART III are based on [3]

[1] Felippa C., FEM Modeling: Introduction

http://caswww.colorado.edu/courses.d/IFEM.d/IFEM.Ch06.d/IFEM.Ch06.pdf

- [2] Felippa C., Variational Formulation of Bar Element <u>http://caswww.colorado.edu/courses.d/IFEM.d/IFEM.Ch11.d/IFEM.Ch11.pdf</u>
- [3] Flaherty, J., E., Eaton A., <u>http://www.cs.rpi.edu/~flaherje/FEM/fem1.ps</u>
- [4] Strang, G., Fix, G., J., An Analysis of the Finite Element Method, Prentice-Hall,1973