Finite Element Method

The finite element method is at the pinnacle of computational structural analysis. Argyris and Clough pioneered its application in structural analysis in the 1960's and its mathematical foundation is the subject of a book by Strang and Fix. Many other textbooks are available on the subject and the method is extended and improved in a variety of old and recent journal papers. All structural analysis programs are based on some form of the finite element method. This document outlines some general features of the finite element method. Several other documents on specific finite elements provide more details.

Derivation

There exist different ways of deriving the finite element method. Some are rather mathematical while others take the stiffness method for truss and beam elements as a starting point. All approaches, however, have the following in common: they include the consideration of equilibrium, kinematics, and material law. In other words, the finite element method addresses the boundary value problem (BVP) of structural analysis. Unfortunately, point-wise satisfaction of all the governing equations is impossible for many practical problems. The finite element method addresses this by approximating either the equilibrium or the kinematic equations. Instead of point-wise satisfaction, the equations are satisfied only on average within each finite element. Naturally, the more elements the more accurate the solution is. In the most popular version of the finite element method the equilibrium requirement is approximated. This is referred to as the displacement-based approach because the displacement field is considered as the unknown. It is approximated by so-called shape functions. Force-based and hybrid finite element methods also exist. In those, the kinematic requirement is approximated and the force field is considered as the unknown and subject to approximation. Regardless of which governing equation is approximated, the derivation of the finite element method starts with either a virtual work principle or an energy principle.

- The popular displacement-based element formulation is based on:
    - the principle of virtual displacements, i.e., the weak form of the BVP, or
    - the principle of minimum potential energy, i.e., the variational form of the BVP.
- The force-based element formulation is based on:
    - the principle of virtual forces, or
    - the principle of complementary potential energy
- The hybrid element formulation, also called the mixed formulation, is based on variational formulations like:
    - the Hu-Washizu variational principle, or
    - the Hellinger-Reissner variational principle

Regardless of approach, integral expressions for the stiffness matrix and load vector—or equivalently the flexibility matrix and the displacement vector—emerge.

Discretization

The fundamental notion of the finite element method is discretization of a continuous boundary value problem. The structure is discretized into finite elements and the unknown field is discretized. In the popular displacement-based finite element method the displacement field is discretized. Specifically, the displacement within each element is described by spatially varying "shape functions," denoted N_i. Each shape function is multiplied by one degree of freedom (DOF), denoted u_i. The displacement at any location within a finite element is the sum of all shape functions multiplied by their DOFs:

$$\tilde{u} = N_i u_i \tag{1}$$

where $\tilde{u}$ is the continuous unknown displacement field within the element and summation over repeated indices is implied. The tilde is utilized to distinguish the unknown displacement field from the nodal deformations u_i. However, the tilde is omitted when this distinction is clear from the context. Often there are several displacement fields on the left-hand side of Eq. (1), in which case Eq. (1) is written:

$$\tilde{\mathbf{u}} = \mathbf{N}\mathbf{u} \tag{2}$$

where matrix notation is introduced in the right-hand and **N** is a vector or matrix of spatially varying shape functions and **u** is the vector of DOFs. The finite element discretization is comparable to the Rayleigh-Ritz method and other energy methods. However, in the finite element method the multiplier of each shape function is a physical displacement or rotation, not a generalized coordinate.

An important consequence of assuming shape functions for the displacement field is that the structure is restrained to deform according to those shapes. This make the structural response too stiff compared to the exact solution. Only when the shape functions contain the solution to the differential equation, i.e., the exact solution is the finite element solution exact. This is possible only for a few structural elements, like trusses and beams.

Generic Expressions for Stiffness Matrix and Load Vector

Other documents contain derivations for specific elements. Here, general expressions for the stiffness matrix and load vector are established. It is selected to base this derivation on the principle of virtual displacements, although other options are possible, as mentioned earlier. First, set the internal virtual work equal to the external virtual work:

$$\int_V \delta\boldsymbol{\varepsilon}^T \boldsymbol{\sigma}\, dV - \int_V \delta\tilde{\mathbf{u}}^T \mathbf{p}\, dV = 0 \tag{3}$$

where, from left to right, the strain tensor, stress tensor, displacement field vector, and vector of forces acting in the displacement are recognized. The material law is written $\boldsymbol{\sigma}=\mathbf{D}\boldsymbol{\varepsilon}$ and substitution yields

$$\int_V \delta\boldsymbol{\varepsilon}^T \mathbf{D}\boldsymbol{\varepsilon}\, dV - \int_V \delta\tilde{\mathbf{u}}^T \mathbf{p}\, dV = 0 \tag{4}$$

The kinematic relationship, which relates strains and the displacement field, is written

$$\boldsymbol{\varepsilon} = \nabla \tilde{\mathbf{u}} \tag{5}$$

where ∇ is a differentiation operator on matrix form. The shape function discretization in Eq. (2) is now introduced and the so-called **B**-matrix is defined:

$$\boldsymbol{\varepsilon} = \nabla \tilde{\mathbf{u}} = \nabla \mathbf{N}\mathbf{u} \equiv \mathbf{B}\mathbf{u} \tag{6}$$

where **B** is a matrix of derivatives of the shape functions. Substitution of Eqs. (2) and (6) into Eq. (4) yields

$$\int_V (\mathbf{B}\delta\mathbf{u})^T \mathbf{D}(\mathbf{B}\mathbf{u})\,dV - \int_V (\mathbf{N}\delta\mathbf{u})^T \mathbf{p}\,dV = 0 \tag{7}$$

Rearranging yields

$$\delta\mathbf{u}^T \left(\left(\int_V \mathbf{B}^T \mathbf{D}\mathbf{B}\,dV \right) \mathbf{u} - \int_V \mathbf{N}^T \mathbf{p}\,dV \right) = 0 \tag{8}$$

Because the virtual displacement field represented by $\delta\mathbf{u}$ is arbitrary the large parenthesis must be zero, hence:

$$\underbrace{\left(\int_V \mathbf{B}^T \mathbf{D}\mathbf{B}\,dV \right)}_{\mathbf{K}} \mathbf{u} = \underbrace{\int_V \mathbf{N}^T \mathbf{p}\,dV}_{\mathbf{F}} \tag{9}$$

where integral expressions for the stiffness matrix, **K**, and load vector, **F**, are identified. Further details are provided in the documents on specific elements. For truss and beam elements the integrals are rather easily evaluated. Conversely, for plane elements and plates and shells it often requires more effort. In particular, the shape functions are sometimes established in a normalized coordinate system, which necessitates integral transformation and the "iso-paramtric" description of the element geometry.

Convergence

If an element satisfies the "convergence conditions" of the finite element method, then the exact solution to the BVP will be obtained by refining the mesh ad infinitum. That is, the more elements are included, the more accurate the solution is. For an element to be convergent, the shape functions must satisfy the following convergence conditions:

Condition #1 (Compatibility)

To ensure finite strain at the element boundaries, it is required that the $(m\text{-}1)^{\text{th}}$ derivative of the displacement field is continuous over the element boundaries, where m is the highest order of differentiation in the integrand of the stiffness matrix, i.e., in the **B**-matrix. It is common to refer to this requirement as "C^{m-1} compatibility" and to call conforming elements "C^{m-1} elements." For example, the beam element is a C^1 element and the Quad4 element is a C^0 element. The former has continuous rotation across element boundaries; the latter only has continuous displacements across element boundaries. A few finite elements deliberately break the compatibility requirement because lack of compatibility provides greater element flexibility, which compensates for the over-estimation of the stiffness introduced by approximate shape functions.

Condition #2 (Completeness)
The element must be able to undergo rigid body motion without producing strain.

Condition #3 (Completeness)
The shape functions must allow the element to be in a state of constant strain.

Patch Test
The patch test is devised to check if an element is convergent. It is carried out as follows:

1. Create a patch of irregularly shaped elements with one free inner node
2. Apply a constant strain pattern or a rigid-body motion displacement pattern
3. Check that the displacement of the inner node is correct, up to computer precision decimals

p and h Refinement
If the mesh is too coarse then the finite element solution is too stiff, and consequently inaccurate. Two strategies are possible to improve the solution: p-refinement and h-refinement. These are explained in the following. Consider a Taylor expansion of the displacement field. The highest order of polynomials included in the shape functions is called p. Because the finite element solution contains polynomial terms up to order p , the error due to omitted terms is of the order p+1:

$$e_{disp} = O\left(h^{p+1}\right) \tag{10}$$

Differentiation reduces the order of accuracy. Hence, because m is the order of differentiation in the B-matrix to obtain the strains from the displacements, the order of the error in the stresses and strains is:

$$e_{strain} = e_{stress} = O(h^{p-m+1}) \tag{11}$$

This leads to the following two strategies to reduce the error:

- *h*-refinement: Refining the mesh by reducing the characteristic element size *h* in the element mesh
- *p*-refinement: Increasing the polynomial order of the shape functions by adding nodes to each element

Reduced Integration

The stiffness matrix integral in Eq. (9) cannot be evaluated analytically for 2D and 3D finite elements like plates and brick elements. Instead, quadrature rules are applied. Depending on the quadrature rule, it may not provide exact results for the polynomial terms contained in the shape functions. For example, Gauss integration provides exact results for integration of polynomials up to order $2n$-1, where n is the number of integration points in one direction. The phrase “full integration” is defined as the integration rule that gives exact results for an undistorted element. “Reduced integration” is anything less than full integration. Apart from reduction in computational cost, reduced integration is sometimes helpful because it makes the element softer, which counteracts the fact that element is made too stiff by the assumption of inexact shape functions. On the other hand, the quadrature order must not be so low that the volume of the element is incorrectly computed. This has to do with retaining the convergence properties of the element; the element must be able to capture the constant strain state exactly.

Mechanisms

Reduced integration and other situations may cause unwanted mechanisms. Other names for a mechanism in this context are zero-energy mode and hourglass mode. These may cause mesh instability, i.e., modes for deformation with artificially low stiffness. In other words, a structure may appear adequately restrained by boundary conditions but may have modes of deformation with very little stiffness. Examples are shown in the document on the Quad4 element. Element mechanisms are detected by performing an eigenvalue test of the element, which reveals the number of zero-energy modes of the element. For this purpose, consider the situation where a displacement-proportional load is applied to the element:

$$\mathbf{Ku} = \mathbf{F} = \lambda \mathbf{u} \tag{12}$$

Rearranging gives

$$(\mathbf{K} - \lambda \mathbf{I})\mathbf{u} = \mathbf{0} \tag{13}$$

which is an eigenvalue problem. The number of eigenvalues is obviously equal to the size of this system of equations, namely the number of degrees of freedom. However, it is the zero eigenvalues that are carefully studied. Zero eigenvalues are associated with rigid-body and potential other zero-energy modes. It is those other modes that the analyst must be aware of. The following procedure is suggested:

1. Determine by manual inspection the number of independent rigid-body displacement and rotation modes that the element should have
2. Solve the eigenvalue problem in Eq. (13) for the element
3. If the number of zero eigenvalues is greater than the number of rigid body modes then the element possesses zero-energy modes that the analyst must be made aware of

Zero-energy modes in an element may be fine as long as they do not result in global mesh instabilities. In practice, potential mesh instabilities are carefully monitored when utilizing elements with zero-energy modes.