# Mathematics

FINITE ELEMENT ANALYSIS AS COMPUTATION

What the textbooks don't teach you about finite element analysis

Chapter 2: Paradigms and some approximate solutions

Gangan Prathap
Director
NISCAIR, S.V. Marg
New Delhi - 110016

## Contents

# Chapter 2

## Paradigms and some approximate solutions

### 2.1 Introduction

We saw in the preceding chapter that a continuum problem in structural or solid mechanics can either be described by a set of partial differential equations and boundary conditions or as a functional $\Pi$ based on the energy principle whose extremum describes the equilibrium state of the problem. Now, a continuum has infinitely many material points and therefore has infinitely many degrees of freedom. Thus, a solution is complete only if analytical functions can be found for the displacement and stress fields, which describe these states, exactly everywhere in the domain of the problem. It is not difficult to imagine that such solutions can be found only for a few problems.

We also saw earlier that the Rayleigh-Ritz (RR) and finite element (fem) approaches offer ways in which approximate solutions can be achieved without the need to solve the differential equations or boundary conditions. This is managed by performing the discretization operation directly on the functional. Thus, a real problem with an infinitely large number of degrees of freedom is replaced with a computational model having a finite number of degrees of freedom. In the RR procedure, the solution is approximated by using a finite number of admissible functions $f_i$ and a finite number of unknown constants $a_i$ so that the approximate displacement field is represented by a linear combination of these functions using the unknown constants. In the fem process, this is done in a piecewise manner - over each sub-region (element) of the structure, the displacement field is approximated by using shape functions $N_i$ within each sub-region and nodal degrees of freedom $u_i$ at nodes strategically located so that they connect the elements together without generating gaps or overlaps. The functional now becomes a function of the degrees of freedom ($a_i$ or $u_i$ as the case may be). The equilibrium configuration is obtained by applying the criterion that $\Pi$ must be stationary with respect to the degrees of freedom.

It is assumed that this solution process of seeking the stationary or extremum point of the discretized functional will determine the unknown constants such that these will combine together with the admissible or shape functions to represent some aspect of the problem to some "best" advantage. Which aspect this actually is has been a matter of some intellectual speculation. Three competing paradigms present themselves.

It is possible to believe that by "best" we mean that the functions tend to satisfy the differential equations of equilibrium and the stress boundary conditions more and more closely as more terms are added to the RR series or more elements are added to the structural mesh. The second school of thought believes that it is the displacement field, which is approximated to greater accuracy with improved idealization. The "aliasing" paradigm, which will be critically discussed later, belongs to this school. It follows from this that stresses, which are computed as derivatives of the approximate displacement fields, will be less accurate. In this book however, we will seek to establish the currency of a third paradigm - that the RR or fem process actually seeks to find to best advantage, the state of stress or strain in the structure. In this scenario, the displacement fields are computed from these "best-fit" stresses as a consequence.

Before we enter into a detailed examination of the merits or faults of each of these paradigms, we shall briefly introduce a short statement on what is meant by the use of the term "paradigm" in the present context. We shall follow this by examining a series of simple approximations to the cantilever bar problem but with more and more complex loading schemes to see how the overall picture emerges.

## 2.2 What is a "paradigm"?

Before we proceed further it may be worthwhile to state what we mean by a *paradigm* here. This is a word that is uncommon to the vocabulary of a trained engineer. The dictionary meaning of paradigm is *pattern* or *model* or *example*. This does not convey much in our context. Here, we use the word in the greatly enlarged sense in which the philosopher T. S. Kuhn introduced it in his classic study on scientific progress [2.1]. In this sense, a *paradigm* is a **"framework of suppositions as to what constitutes problems, theories and solutions. It can be a collection of metaphysical assumptions, heuristic models, commitments, values, hunches, which are all shared by a scientific community and which provide the conceptual framework within which they can recognize problems and solve them [2.2]. The *aliasing* and *best-fit* paradigms can be thought of as two competing scenarios, which attempt to explain how the finite element method computes strains and stresses. Our task will therefore be to establish which paradigm has greater explanatory power and range of application.**

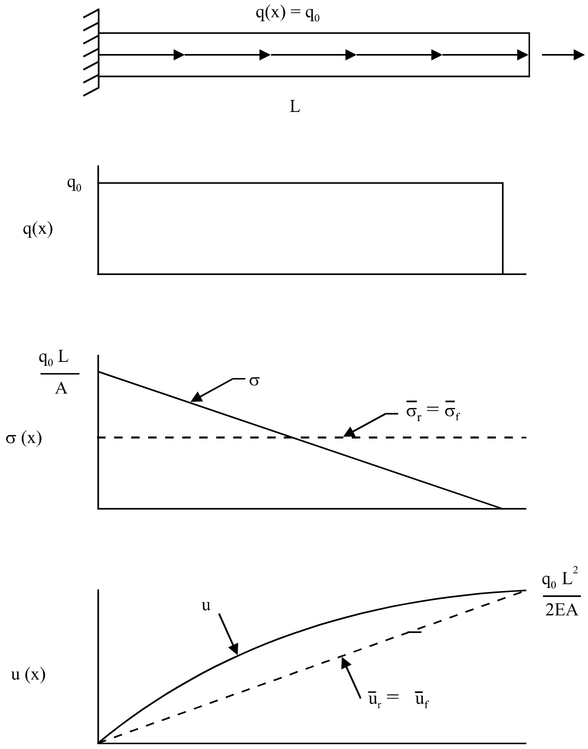


Fig. 2.1 Bar under uniformly distributed axial load

### 2.3 Bar under uniformly distributed axial load

Consider a cantilever bar subjected to a uniformly distributed axial load of intensity $q_0$ per unit length (Fig. 2.1). Starting with the differential equation of equilibrium, it is easy to show that the analytical solution to the problem is

$$u(x) = (q_0/EA)\ (Lx - x^2/2) \tag{2.1a}$$

$$\sigma(x) = (q_0/A)(L - x) \tag{2.1b}$$

Consider a one-term RR solution based on $\bar{u}_r = a_1 x$, where the subscript *r* denotes the use of the RR approach. It is left to the reader to show that the approximate solution obtained is

$$\bar{u}_r(x) = (q_0/EA)\ (Lx/2) \tag{2.2a}$$

$$\bar{\sigma}_r(x) = (q_0/A)\ (L/2) \tag{2.2b}$$

We now compare this with an fem solution based on a two-noded linear element. The solution obtained will be

$$\bar{u}_f(x) = (q_0/EA)\ (Lx/2) \tag{2.3a}$$

$$\bar{\sigma}_f(x) = (q_0/A)\ (L/2) \tag{2.3b}$$

We see that the RR and fem solutions are identical. This is to be expected because the fem solution is effectively an RR solution. We can note also the curious coincidence where all three solutions predict the same displacement at the tip. However, from Fig. 2.1 we can see that the $\bar{u}_r$ and $\bar{u}_f$ are approximate solutions to *u*. It is also clear from Fig. 2.1 that $\bar{\sigma}_r = \bar{\sigma}_f = \sigma$ at the mid-point of the beam. It is possible to speculate that $\bar{\sigma}_r$ and $\bar{\sigma}_f$ bear some unique relationship to the true variation $\sigma$.

Consider next what will happen if two equal length linear (i.e., two-noded) bar elements are used to model the bar. It is left to the reader to work out that the solution described in Fig. 2.2 will be obtained. First, we must note that the distributed axial load is consistently lumped at the nodes. Thus the physical load system that the fem equations are solving is not that described in Fig. 2.1 or Fig. 2.2 as $\sigma$. Instead, we must think of a substitute stairstep distribution $\sigma_f$, produced by the consistent load lumping process (see Fig. 2.2) which is sensed by the fem stiffness matrix. Now, a solution of the set of algebraic equations will result in $\bar{\sigma}_f$ and $\bar{u}_f$ as the fem solution.

We see once again that the nodal predictions are exact. This is only a coincidence for this particular type of problem and nothing more can be read into this fact. More striking is the observation that the stresses computed by the finite element system now approximates the original true stress in a stairstep fashion.

It also seems reasonable to conclude that within each element, the true state of stress is captured by the finite element stress in a "best-fit" sense. In other words, we can generalize from Figs. 2.1 and 2.2, that the finite element stress magnitudes are being computed according to some precise rule.

Also, there is the promise that by carefully understanding what this rule is, it will be possible to derive some unequivocal guidelines as to where the stresses are accurate. In this example, where an element capable of yielding constant stresses is used to model a problem where the true stresses vary linearly, the centroid of the element yields exact predictions. As we take up further examples later, this will become more firmly established.

A cursory comparison of Figs. 2.1 and 2.2 also indicates that in a general sense, the approximate displacements are more accurate than the approximate stresses. It seems compelling now to argue that this is so because the approximate displacements emerge as "discretized" integrals of the stresses or strains and for that reason, are more accurate than the stresses.

## 2.4 Bar under linearly varying distributed axial load

We now take up a slightly more difficult problem. The cantilever bar has a load distributed in a linearly varying fashion (Fig. 2.3). It is easy to establish that the exact stress distribution in this case will be quadratic in nature. We can write it as

$$\sigma(x) = \left(q_0\ L^2 \big/ 8A\right)\ \left(4/3 - 2\xi - \left(1 - 3\xi^2\right)\Big/3\right) \tag{2.4}$$

Some interesting features about this equation can be noted. A dimensionless coordinate, $\xi = 2x/L-1$ has been chosen so that it will also serve as a natural coordinate system taking on values *-1* and *1* at the ends of the single three-node bar element shown as modeling the entire bar in Fig.2.3. We have also very curiously expanded the quadratic variation using the terms *1*, $\xi$, $(1-3\xi^2)$. These can be identified with the Legendre polynomials and its relevance to the treatment here will become more obvious as we proceed further.

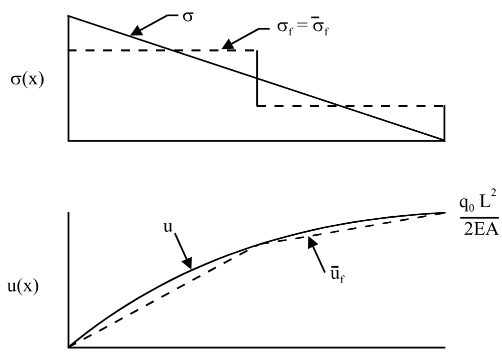


Fig. 2.2 Two element model of bar under uniformly distributed axial load.

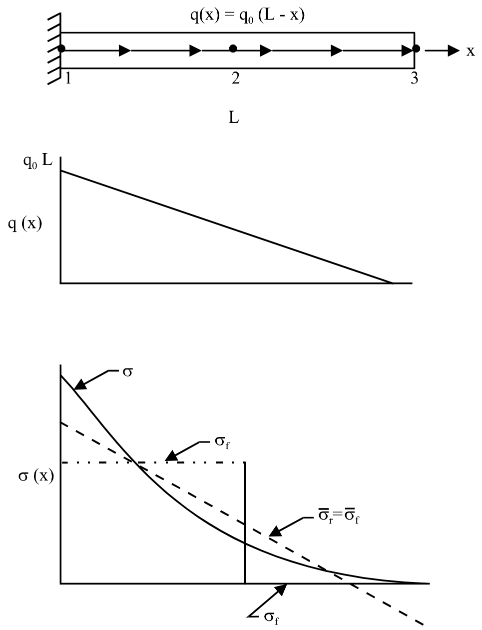


Fig. 2.3 Bar under linearly varying axial load

We shall postpone the first obvious approximation, that of using a one-term series $\bar{u}_r = a_1 x$ till later. For now, we shall consider a two-term series $\bar{u}_r = a_1 x + a_2 x^2$. This is chosen so that the essential boundary conditions at $x=0$ is satisfied. No attempt is made to satisfy the force boundary condition at $x=L$. The reader can verify by carrying out the necessary algebra associated with the RR process that the solution obtained will yield an approximate stress pattern given by

$$\bar{\sigma}_r(x) = \left(q_0 L^2 / 8A\right)\left(4/3 - 2\xi\right) \tag{2.5}$$

This is plotted on Fig. 2.3 as the dashed line. A comparison of Equations (2.4) and (2.5) reveals an interesting fact - only the first two Legendre polynomial terms are retained. Taking into account the fact that the Legendre polynomials are orthogonal, what this means is that in this problem, we have obtained $\bar{\sigma}_r$ in a manner that seems to satisfy the following integral condition:

$$\int_{-1}^{1} \delta\bar{\sigma}_r \left(\bar{\sigma}_r - \sigma\right) d\xi = 0 \tag{2.6}$$

That is, the RR procedure has determined a $\bar{\sigma}_r$ that is a "best-fit" of the true state of stress $\sigma$ in the sense described by the orthogonality condition in Equation (2.6). This is a result anticipated from our emerging results to the various exercises we have conducted so far. We

have not been able to derive it from any general principle that this must be so for stronger reasons than shown here up till now.

Let us now proceed to an fem solution. It is logical to start here with the three-node element that uses the quadratic shape functions,

$$N_1 = \xi(\xi - 1)/2, \qquad N_2 = (1 - \xi^2) \qquad N_3 = \xi(\xi + 1)/2.$$

We first compute the consistent loads to be placed $P_i = \int N_i q\,dx$

at the nodes due to the distributed load using .

$$P_1 = q_0 L^2/6,$$

This results in the following scheme of loads at the three nodes identified in Fig. 2.3:

$P_2 = q_0 L^2/3$ and $P_3 = 0$

The resulting load configuration can be represented in the form of a stress system shown as $\sigma_f$, represented by the dashed-dotted lines in Fig. 2.3. Thus, any fem discretization automatically replaces a smoothly varying stress system by a step-wise system as shown by $\sigma_f$ in Figs. 2.2 and 2.3. It is this step-wise system that the finite element solution $\bar{\sigma}_f$ responds to. If the finite element computations are actually performed using the stiffness matrix for the three-node element and the consistent load vector, it turns out, as the reader can assure himself, that the computed fem stress, pattern will be

$$\bar{\sigma}_f(x) = \left(q_0 L^2/8A\right)\ (4/3 - 2\xi) \tag{2.7}$$

This turns out to be exactly the same as the $\bar{\sigma}_r$ computed by the RR process in Equation (2.5). At first sight, this does not seem to be entirely unexpected. Both the RR process and the fem process here have started out with quadratic admissible functions for the displacement fields. This implies that both have the capability to represent linear stress fields exactly or more complicated stress fields by an approximate linear field that is in some sense a best approximation. On second thought however, there is some more subtlety to be taken care of. In the RR process, the computed $\bar{\sigma}_r$ was responding to a quadratically varying system $\sigma$ (see Fig. 2.3). We could easily establish through Equation (2.6) that $\bar{\sigma}_r$ responded to $\sigma$ in a "best-fit" manner. However, in the fem process, the load system that is being used is the $\sigma_f$ system, which varies in the stairstep fashion. The question confronting us now is, in what manner did $\bar{\sigma}_f$ respond to $\sigma_f$ - is it also consistent with the "best-fit" paradigm?

Let us now assume an unknown field $\bar{\bar{\sigma}} = c_0 + c_1\xi$ which is a "best-fit" of the stairstep field given by $\sigma_f = 2q_0 L^2/3$ in $0 < x < L/2$ and $\sigma_f = 0$ in $L/2 < x < L$. We shall determine the constants $c_0$ and $c_1$ so that the "best-fit" variation shown below is satisfied:

$$\int_{-1}^{0} \delta\bar{\bar{\sigma}}\left(\bar{\bar{\sigma}} - 2q_0 L^2/3\right) d\xi \ + \ \int_{0}^{1} \delta\bar{\bar{\sigma}}\left(\bar{\bar{\sigma}} - 0\right) d\xi = 0 \tag{2.8}$$

The reader can work out that this leads to

$$\bar{\bar{\sigma}}(x) = \left(q_0 L^2 / 8A\right) \left(4/3 - 2\xi\right), \tag{2.9}$$

which is identical to the result obtained in Equation (2.7) by carrying out the finite element process. In other words, the fem process follows exactly the "best-fit" description of computing stress fields. Another important lesson to be learnt from this exercise is that the consistent lumping process preserves the "best-fit" nature of the stress representation and subsequent prediction. Thus, $\bar{\sigma}_f$ is a "best-fit" of both σ and $\sigma_f$! Later, we shall see how the best-fit nature gets disturbed if the loads are not consistently derived.

It again seems reasonable to argue that the nodal displacements computed directly from the stiffness equations from which the stress field $\bar{\sigma}_f$ has been processed can actually be thought of as being "integrated" from the "best-fit" stress approximation. It seems possible now to be optimistic about our ability to ask and answer questions like: What kind of approximating fields are best to start with? and, How good are the computed results?

Before we pass on, it is useful for future reference to note that the approximate solutions $\bar{\sigma}_r$ or $\bar{\sigma}_f$ intersect the exact solution $\sigma$ at two points. A comparison of Equation (2.4) with Equations (2.5) and (2.7) indicates that these are the points where the quadratic Legendre polynomial, $\left(1 - 3\xi^2\right)$ vanishes, i.e., at $\xi = \pm 1/\sqrt{3}$ . Such points are well known in the literature of the finite element method as optimal stress points or Barlow points, named after the person who first detected such points. Our presentation shows clearly why such points exist, and why in this problem, where a quadratic stress state is sought to be approximated by a linear stress state, these points are at $\xi = \pm 1/\sqrt{3}$ . Curiously, these are also the points used in what is called the two-point rule for Gauss-Legendre numerical integration. Very often, the two issues are confused. Our interpretation so far shows that the fact that such points are sampling points of the Gauss-Legendre rule has nothing to do with the fact that similar points arise as points of optimal stress recovery. The link between them is that in both, the Legendre polynomials play a decisive role - the zeros of the Legendre polynomials define the optimal points for numerical integration and these points also help determine where the "best-fit" approximation coincides with a polynomial field which is exactly one order higher.

We shall now return to the linear Ritz admissible function, $\bar{u}_r = a_1 x$ , to see if it operates in the best-fit sense. This would be identical to using a single two-node bar element to perform the same function. Such a field is capable of representing only a constant stress. This must now approximate the quadratically varying stress field *σ(x)* given by Equation (2.4). This gives us an opportunity to observe what happens to the optimal stress point, whether one exists, and whether it can be easily identified to coincide with a Gauss integration point.

Again, the algebra is very simple and is omitted here. One can show that the one-term approximate solution would lead to the following computed stress:

$$\bar{\sigma}_r\ (x) = \left(q_0 L^2 / 8A\right) \left(4/3\right) \tag{2.10}$$

What becomes obvious by comparing this with the true stress *σ(x)* in Equation (2.4) and the computed stress from the two-term solution, $\bar{\sigma}_r(x)$ in Equation (2.5) is that the one-term solution corresponds to only the constant part of the Legendre polynomial expansion! Thus, given the orthogonal nature of the Legendre polynomials, we can conclude that we have obtained the "best-fit" state of stress even here. Also, it is clear that the optimal stress point is not easy to identify to coincide with any of the points corresponding to the various Gauss integration rules. The optimal point here is given by $\xi = 1 - \sqrt{4/3}$ .

## 2.5 The aliasing paradigm

We have made out a very strong case for the best-fit paradigm. Let us now examine the merits, if any, of the competing paradigms. The argument that fem procedures look to satisfy the differential equations and boundary conditions does not seem compelling enough to warrant further discussion. However, the belief that finite elements seek to determine nodal displacements accurately was the basis for Barlow's original derivation of optimal points [2.3] and is also the basis for what is called the "aliasing" paradigm [2.4].

The term *aliasing* is borrowed from sample data theory where it is used to describe the misinterpretation of a time signal by a sampling device. An original sine wave is represented in the output of a sampling device by an altered sine wave of lower frequency - this is called the *alias* of the true signal. This concept can be extended to finite element discretization - the sample data points are now the values of the displacements at the nodes and the alias is the function, which interpolates the displacements within the element from the nodal displacements. We can recognize at once that Barlow [2.3] developed his theory of optimal points using an identical idea - the term *substitute function* is used instead of alias.

Let us now use the aliasing concept to derive the location of the optimal points, as Barlow did earlier [2.3], or as MacNeal did more recently [2.4]. We assume here that the finite element method seeks discretized displacement fields, which are substitutes or aliases of the true displacement fields by sensing the nodal displacements directly. We can compare this with the "best-fit" interpretation where the fem is seen to seek discretized strain/stress fields, which are the substitutes/aliases of the true strain/stress fields in a "best fit" or "best approximation" sense. It is instructive now to see how the alternative paradigm, the "displacement aliasing" approach leads to subtle differences in interpreting the relationship between the Barlow points and the Gauss points.

Table 2.1 Barlow and Gauss points for one-dimensional case

| p | Nodes at | $u$ | $\bar{u}$ | $\varepsilon$ | $\bar{\varepsilon}$ | Gauss points | Barlow points | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | 'best-fit' | aliasing |
| 1 | $\pm 1$ | $\xi^2$ | $\xi$ | $\xi$ | $1$ | $0$ | $0$ | $0$ |
| 2 | $0, \pm 1$ | $\xi^3$ | $\xi^2$ | $\xi^2$ | $\xi$ | $\pm 1/\sqrt{3}$ | $\pm 1/\sqrt{3}$ | $\pm 1/\sqrt{3}$ |
| 3 | $\pm 1/3, \pm 1$ | $\xi^4$ | $\xi^3$ | $\xi^3$ | $\xi^2$ | $0, \pm\sqrt{3/5}$ | $0, \pm\sqrt{3/5}$ | $0, \pm\sqrt{5/3}$ |

$1, \xi .., \xi^4$ indicate polynomial orders

### *2.5.1 A one dimensional problem*

We again take up the simplest problem, a bar under axial loading. We shall assume that the bar is replaced by a single element of varying polynomial order for its basis function (i.e. having varying no. of equally spaced nodes). Thus, from Table 2.1, we see that *p=1,2,3* correspond to basis functions of linear, quadratic and cubic order, implying that the corresponding elements have *2,3,4* nodes respectively. These elements are therefore capable of representing a constant, linear and quadratic state of strain/stress, where strain is taken to be the first derivative of the displacement field. We shall adopt the following notation: The true displacement, strain and stress field will be designated by $u, \varepsilon$ and $\sigma$. The discretized displacement, strain and stress field will be designated by $\bar{u}$, $\bar{\varepsilon}$ and $\bar{\sigma}$. The *aliased* displacement, strain and stress field will be designated by $u^a$, $\varepsilon^a$ and $\sigma^a$. Nodal displacements will be represented by $u_i$.

We shall examine three scenarios. In the simplest, Scenario a, the true displacement field $u$ is exactly one polynomial order higher than what the finite element is capable of representing - we shall see that the Barlow points can be determined exactly in terms of the Gauss points only for this case. In Scenario b, we consider the case where the true field $u$ is two orders higher than the discretized field $\bar{u}$. The definition of an identifiable optimal point becomes difficult. In both Scenarios a and b, we assume that the rigidity of the bar is a constant, i.e., $\sigma = D\ \varepsilon$. In Scenario c, we take up a case where the rigidity can vary, i.e., $\sigma = D(\xi)\ \varepsilon$. We shall see that once again, it becomes difficult to identify the optimal points by any simple rule.

We shall now take for granted that the best-fit rule operates according to the orthogonality condition expressed in Equation (2.6) and that it can be used interchangeably for stresses and strains. We shall designate the optimal points determined by the aliasing algorithm as $\xi_a$, the Barlow points (aliasing), and the points determined by the best-fit algorithm as $\xi_b$, the Barlow points (best- fit). Note that $\xi_a$ are the points established by Barlow [2.3] and MacNeal [2.4], while $\xi_b$ will correspond to the points given in References 2.6 and 2.7. Note that the natural coordinate system $\xi$ is being used here for convenience.

Thus, for Scenarios a and b, this leads to

$$\int \delta\bar{\varepsilon}^T \left(\bar{\varepsilon} - \varepsilon\right)\ dV = 0 \tag{2.11}$$

whereas for Scenario c, it becomes

$$\int \delta\bar{\varepsilon}^T\ D(\xi)\left(\bar{\varepsilon} - \varepsilon\right)\ dV = 0 \tag{2.12}$$

Note that we can consider Equation (2.11) as a special case of the orthogonality condition in Equation (2.12) with the weight function $D(\xi)=1$. This case corresponds to one in which a straightforward application of Legendre polynomials can be made. This point was observed very early by Moan [2.5]. In this case, one can determine the points where $\bar{\varepsilon} = \varepsilon$ as those corresponding to points, which are the zeros of the Legendre polynomials. See Table 2.2 for a list of unnormalised Legendre polynomials. We shall show below that in Equation (2.11), the points of minimum error are the sampling points of the Gauss Legendre integration rule only if $\bar{\varepsilon}$ is exactly one polynomial order lower than $\varepsilon$. And in Equation (2.12), the optimal points no longer depend on the nature of the Legendre polynomials, making it difficult to identify the optimal points.

***Scenario a***

We shall consider fem solutions using a linear (two-node), a quadratic (three-node) and a cubic (four-node) element. The true displacement field is taken to be one order higher than the discretized field in each case.

Linear element *(p = 1)*

$u$ = *quadratic* = $b_o + b_1\,\xi + b_2\,\xi^2$ (2.13)

$$\varepsilon = linear = u,_{\xi} = b_1 + 2b_2\,\xi = \sum_{s=0}^{p=1} \varepsilon_s\, P_s\,(\xi) \quad (2.14)$$

Note that we have written $\varepsilon$ in terms of the Legendre polynomials for future convenience. Note also that we have simplified the algebra by assuming that strains can be written as derivatives in the natural co-ordinate system. It is now necessary to work out how the algebra differs for the *aliasing* and *best-fit* approaches.

*Aliasing:* At $\xi_i = \pm 1$, $u_i^a = u_i$ ; then points where $\varepsilon^a = \varepsilon$ are given by $\xi_a=0$. (The algebra is very elementary and is left to the reader to work out). Thus, the Barlow point (aliasing) is $\xi_a = 0$, for this case.

*Best-fit:* $\bar{u}$ =*linear*, is undetermined at first. Let $\bar{\varepsilon} = c_0$, as the element is capable of representing only a constant strain. Equation (2.11) will now give $\bar{\varepsilon} = c_0 = b_1$. Thus, the optimal point is $\xi_b = 0$, the point where the Legendre polynomial $P_1(\xi) = \xi$ vanishes. Therefore, the Barlow point (best-fit) for this example is $\xi_b = 0$.

Table 2.2 The Legendre polynomials $P_i$

| **Order of polynomial *i*** | **Polynomial $P_i$** |
|---|---|
| **0** | **1** |
| **1** | $\xi$ |
| **2** | **$(1-3\xi^2)$** |
| **3** | **$(3\xi-5\xi^3)$** |
| **4** | **$(3-30\xi^2+35\xi^4)$** |

<u>Quadratic</u> <u>element</u> *(p = 2)*

$u$ = *cubic* = $b_0 + b_1\,\xi + b_2\,\xi^2 + b_3\,\xi^3$

$\varepsilon$ = *quadratic* = $u,_{\xi}$ (2.15)

$$= (b_1 + b_3) + 2b_2\,\xi - b_3\left(1 - 3\xi^2\right) = \sum_{s=0} \varepsilon_s P_s\,(\xi) \quad (2.16)$$

*Aliasing:* At $\xi_i = 0, \pm 1$, $u_i^a = u_i$ ;; then points where $\varepsilon^a = \varepsilon$ are given by $\xi_a = \pm 1/\sqrt{3}$. (Again, the algebra is left to the reader to work out). Thus, the Barlow points (aliasing) are $\xi_a = \pm 1/\sqrt{3}$, for this case.

*Best-fit:* $\bar{u}$ = *quadratic*. Let $\bar{\varepsilon} = c_0 + c_1\xi$ as this element is capable of representing a linear strain. Equation (2.11) will now give $\bar{\varepsilon} = (b_1 + b_3) + 2b_2\,\xi$. Thus, the optimal points are

$\xi_b = \pm 1/\sqrt{3}$, the points where the Legendre polynomial $P_2(\xi) = (1 - 3\xi^2)$ vanishes. Therefore, the Barlow points (best-fit) for this example are $\xi_b = \pm 1/\sqrt{3}$.

Note that in these two examples, i.e. for the linear and quadratic elements, the Barlow points from both schemes coincide with the Gauss points (the points where the corresponding Legendre polynomials vanish). In our next example we will find that this will not be so.

Cubic element *(p = 3)*

$$u = quartic = b_0 + b_1\xi + b_2\xi^2 + b_3\xi^3 + b_4\xi^4$$
(2.17)

$$\varepsilon = cubic = u_{,\xi}$$

$$= (b_1 + b_3) + (2b_2 + 12b_4/5)\xi - b_3(1 - 3\xi^2) - 4b_4/5(3\xi - 5\xi^3)$$

$$= \sum_{s=0}^{p=3} \varepsilon_s P_s(\xi)$$
(2.18)

*Aliasing:* At $\xi_i = \pm 1/3, \ \pm 1, \ u_i^a = u_i$; then points where $\varepsilon^a = \varepsilon$ are given by $\xi_a = 0, \pm\sqrt{5}/3$. Thus, the Barlow points (aliasing) are $\xi_a = 0, \pm\sqrt{5}/3$ for this case. Note that the points where the Legendre polynomial $P_3(\xi) = (3\xi - 5\xi^3)$ vanishes are $\xi_0 = 0, \sqrt{3/5}$ !

*Best-fit:* $\bar{u}$ = *cubic*. Let $\bar{\varepsilon} = c_0 + c_1\xi + c_2(1 - 3\xi^2)$, as this element is capable of representing a quadratic strain. Equation (2.11) will now give $\bar{\varepsilon} = (b_1 + b_3) + (2b_2 + 12b_4/5)\xi - b_3(1 - 3\xi^2)$. Thus, the Barlow points (best-fit) for this example are $\xi_b = 0, \sqrt{3/5}$, the points where the Legendre polynomial $P_3(\xi) = (3\xi - 5\xi^3)$ vanishes.

Therefore, we have an example where the aliasing paradigm does not give the correct picture about the way the finite element process computes strains. However, the best-fit paradigm shows that as long as the discretized strain is one order lower than the true strain, the corresponding Gauss points are the optimal points. Table 2.1 summarizes the results obtained so far.

The experience of this writer and many of his colleagues is that the best -fit model, is the one that corresponds to reality. If one were to actually solve a problem where the true strain varies cubically using a 4-noded element, which offers a discretized strain which is of quadratic order, the points of optimal strain actually coincide with the Gauss points.

***Scenario b:***

So far, we have examined simple scenarios where the true strain is exactly one polynomial order higher than the discretized strain with which it is replaced. If $P_p(\xi)$, denoting the Legendre polynomial of order *p*, describes the order by which the true strain exceeds the discretized strain, the simple rule is that he optimal points are obtained as $P_p(\xi_b) = 0$. These are therefore the set of *p* Gauss points at which the Legendre polynomial of order *p* vanishes. Consider now a case where the true strain is two orders higher than the discretized strain - e.g. a quadratic element *(p = 2)* modeling a region where the strain and stress field vary cubically. Thus, we have,

$$\varepsilon = (b_1 + b_3) + (2b_2 + 12b_4/5)\xi - b_3\left(1 - 3\xi^2\right) - 4b_4/5\left(3\xi - 5\xi^3\right)$$

$$\bar{\varepsilon} = c_0 + c_1\xi$$

(2.19)

Equation (2.11) allows us to determine the coefficient $c_i$ in terms of $b_i$; it turns out that

$$\bar{\varepsilon} = (b_1 + b_3) + (2b_2 + 12b_4/5)\,\xi\,, \qquad (2.20)$$

a representation made very easy by the fact that the Legendre polynomials area orthogonal and that therefore $\bar{\varepsilon}$ can be obtained from $\varepsilon$ by simple inspection. It is not however a simple matter to determine whether the optimal points coincide with other well-known points like the Gauss points. In this example, we have to seek the zeros of

$$b_3\left(1 - 3\xi^2\right) + 4b_4/5\left(3\xi - 5\xi^3\right) \qquad (2.21)$$

Since $b_3$ and $b_4$ are arbitrary, depending on the problem, it is not possible to seek universally valid points where this would vanish, unlike in the case of Scenario a earlier. Therefore, in such cases, it is not worthwhile to seek points of optimal accuracy. It is sufficient to acknowledge that the finite element procedure yields strains $\bar{\varepsilon}$ , which are the most reasonable one can obtain in the circumstances.

***Scenario c***

So far, we have confined attention to problems where $\sigma$ is related to $\varepsilon$ through a simple constant rigidity term. Consider an exercise where (the one-dimensional bar again) the rigidity varies because the cross-sectional area varies or because the modulus of elasticity varies or both, i.e. $\sigma = D(\xi)\varepsilon$. The orthogonality condition that governs this case is given by Equation (2.12). Thus, it may not be possible to determine universally valid Barlow points *a priori* if $D(\xi)$ varies considerably.

**2.6 The `best-fit' rule from a variational theorem**

Our investigation here will be complete in all respects if the basis for the best-fit paradigm can be derived logically from a basic principle. In fact, some recent work [2.6,2.7] shows how by taking an enlarged view of the variational basis for the displacement type fem approach we will be actually led to the conclusion that strains or stresses are always sought in the best-fit manner.

The `best-fit' manner in which finite elements compute strains can be shown to follow from an interpretation uses the Hu-Washizu theorem. To see how we progress from the continuum domain to the discretized domain, we will find it most convenient to develop the theory from the generalized Hu-Washizu theorem [2.8] rather than the minimum potential theorem. As we have seen earlier, these theorems are the most basic statements that can be made about the laws of nature. The minimum potential theorem corresponds to the conventional energy theorem. However, for applications to problems in structural and solid mechanics, Hu proposed a generalized theorem, which had somewhat more flexibility [2.8]. Its usefulness came to be recognized when one had to grapple with some of the problems raised by finite element modeling. One such puzzle is the rationale for the "best-fit" paradigm.

Let the continuum linear elastic problem have an exact solution described by the displacement field $u$, strain field $\varepsilon$ and stress field $\sigma$. (We project that the strain field $\varepsilon$ is derived from the displacement field through the strain-displacement gradient operators of the theory of elasticity and that the stress field is derived from the strain field through the constitutive laws.) Let us now replace the continuum domain by a discretized domain and describe the computed state to be defined by the quantities $\bar{u}$, $\bar{\varepsilon}$ and $\bar{\sigma}$, where again, we take that the strain fields and stress fields are computed from the strain-displacement and constitutive relationships. It is clear that $\bar{\varepsilon}$ is an approximation of the true strain field $\varepsilon$. What the Hu-Washizu theorem does, following the interpretation given by Fraejis de Veubeke [2.9], is to introduce a dislocation potential to augment the usual total potential. This dislocation potential is based on a third independent stress field $\bar{\bar{\sigma}}$ which can be considered to be the Lagrange multiplier removing the lack of compatibility appearing between the true strain field $\varepsilon$ and the discretized strain field $\bar{\varepsilon}$. Note that, $\bar{\bar{\sigma}}$ is now an approximation of $\bar{\sigma}$. The three-field Hu-Washizu theorem can be stated as

$$\delta\pi = 0 \tag{2.22}$$

where

$$\pi = \int \left\{ \bar{\sigma}^T \bar{\varepsilon}/2 + \bar{\bar{\sigma}}^T (\varepsilon - \bar{\varepsilon}) + P \right\} dv \tag{2.23}$$

and $P$ is the potential energy of the prescribed loads.

In the simpler minimum total potential principle, which is the basis for the derivation of the displacement type finite element formulation in most textbooks, only one field (i.e., the displacement field $u$) is subject to variation. However, in this more general three field approach, all the three fields are subject to variation and leads to three sets of equations which can be grouped and classified as follows:

| Variation on | Nature | Equation | |
|---|---|---|---|
| $\bar{u}$ | Equilibrium | $\nabla \bar{\bar{\sigma}}$ + terms from $P = 0$ | (2.24a) |
| $\bar{\bar{\sigma}}$ | Orthogonality (Compatibility) | $\int \delta\bar{\bar{\sigma}}^T (\bar{\varepsilon} - \varepsilon)\, dv = 0$ | (2.24b) |
| $\bar{\varepsilon}$ | Orthogonality (Equilibrium) | $\int \delta\bar{\varepsilon}^T (\bar{\bar{\sigma}} - \bar{\sigma})\, dv = 0$ | (2.24c) |

Equation (2.24a) shows that the variation on the displacement field $u$ requires that the independent stress field $\bar{\bar{\sigma}}$ must satisfy the equilibrium equations ( $\nabla$ signifies the operators that describe the equilibrium condition. Equation (2.24c) is a variational condition to restore the equilibrium imbalance between $\bar{\sigma}$ and $\bar{\bar{\sigma}}$. In the displacement type formulation, we choose $\bar{\bar{\sigma}} = \bar{\sigma}$. This satisfies the orthogonality condition seen in Equation (2.24c) identically. This leaves us with the orthogonality condition in Equation (2.24b). We can now argue that this tries to restore the compatibility imbalance between the exact strain field ε and the discretized strain field $\bar{\varepsilon}$ . In the displacement type formulation this can be stated as,

$$\int \delta\bar{\sigma}^T \left(\bar{\varepsilon} - \varepsilon\right) dV = 0 \qquad (2.25)$$

Thus we see very clearly, that the strains computed by the finite element procedure are a variationally correct (in a sense, a least squares correct) `best approximation' of the true state of strain.

**2.7 What does the finite element method do?**

To understand the motivation for the aliasing paradigm, it will help to remember that it was widely believed that the finite element method sought approximations to the displacement fields and that the strains/stresses were computed by differentiating these fields. Thus, elements were believed to be "capable of representing the nodal displacements in the field to a good degree of accuracy." Each finite element samples the displacements at the nodes, and internally, within the element, the displacement field is interpolated using the basis functions. The strain fields are computed from these using a process that involves differentiation. It is argued further that as a result, displacements are more accurately computed than the strain and stress field. This follows from the generally accepted axiom that derivatives of functions are less accurate than the original functions. It is also argued that strains/stresses are usually most inaccurate at the nodes and that they are of greater accuracy near the element centers - this, it is thought, is a consequence of the mean value theorem for derivatives.

However, we have demonstrated convincingly that in fact, the Ritz approximation process, and the displacement type fem which can be interpreted as a piecewise Ritz procedure, do exactly the opposite - it is the strain fields which are computed, almost independently as it were within each element. This can be derived in a formal way. Many attempts have been made to give expression to this idea, (e.g., Barlow [2.1] and Moan [2.7]), but it appears that the most intellectually satisfying proof can be arrived at by starting with a mixed principle known as the Hu-Washizu theorem [2.8]. This proof has been taken up in greater detail in Reference 2.3 and was reviewed only briefly in the preceding section. Having said that the Ritz type procedures determine strains, it follows that the displacement fields are then constructed from this in an integral sense. The stiffness equation actually reflecting this integration process and the continuity of fields across element boundaries and suppression of the field values at domain edges being reflected by the imposition of boundary conditions. It must therefore be argued that displacements are more accurate than strains because integrals of smooth functions are generally more accurate than the original data. We have thus turned the whole argument on its head.

## 2.8 Conclusions

In this chapter, we postulated a few models to explain how the displacements type fem works. We worked out a series of simple problems of increasing complexity to establish that our conjecture that strains and stresses appear in a "best-fit" sense could be verified (falsified, in the Popperian sense) by carefully designed numerical experiments.

An important part of this exercise depended on our careful choice and use of various stress terms. Thus terms like $\sigma$ and $\sigma_f$ was actual physical states that were sought to be modeled. The stress terms $\bar{\sigma}_r$ and $\bar{\sigma}_f$ were the quantities that emerged in what we can call the "first-order tradition" analysis in the language of Sir Karl Popper [2.10] - where the RR or fem operations are mechanically carried out using functional approximation and finite element stiffness equations respectively. We noticed certain features that seemed to relate these computed stresses to the true system they were modeling in a predictable or repeatable manner. We then proposed a mechanism to explain how this could have taken place. Our bold conjecture, after examining these numerical experiments, was to propose that it is effectively seeking a best-fit state.

To confirm that this conjecture is scientifically coherent and complete, we had to enter into a "second-order tradition" exercise. We assumed that this is indeed the mechanism that is operating behind the scenes and derived predicted quantities $\bar{\bar{\sigma}}_r$ that will result from the best-fit paradigm when this was applied to the true state of stress. These predicted quantities turned out to be exactly the same as the quantities computed by the RR and fem procedures. In this manner, we could satisfy ourselves that the "best-fit" paradigm had successfully survived a falsification test.

Another important step we took was to prove that the "best-fit" paradigm was neither gratuitous nor fortuitous. In fact, we could also establish that this could be derived from more basic principles - in this regard, the generalized theorem of Hu was found valuable to determine that the best-fit paradigm had a rational basis. In subsequent chapters, we shall use this foundation to examine other features of the finite element method.

One important conclusion we can derive from the best-fit paradigm is that, an interpolation field for the stresses $\bar{\sigma}$ (or stress resultants as the case may be) which is of higher order than the strain fields $\varepsilon$. On which it must `do work' in the energy or virtual work principle is actually self-defeating because the higher order terms cannot be `sensed'. This is precisely the basis for de Veubeke's famous limitation principle [2.9], that *'it is useless to look for a better solution by injecting additional degrees of freedom in the stresses.'* We can see that one cannot get stresses, which are of higher order than are reflected in the strain expressions.

## 2.9 References


2.1 T. S. Kuhn, *The Structure of Scientific Revolution*, University of Chicago Press, 1962.

2.2 S. Dasgupta, Understanding design: Artificial intelligence as an explanatory paradigm, SADHANA, 19, 5-21, 1994.

2.3. J. Barlow, Optimal stress locations in finite element models, *Int. J. Num. Meth. Engng.* 10, 243-251 (1976).

2.4. R. H. MacNeal, *Finite Elements: Their Design and Performance,* Marcel Dekker, NY, 1994.

2.5 T. Moan, On the local distribution of errors by finite element approximations, *Theory and Practice in Finite Element Structural Analysis. Proceedings of the 1973 Tokyo Seminar on Finite Element Analysis,* Tokyo, 43-60, 1973.

2.6. G. Prathap, *The Finite Element Method in Structural Mechanics*, Kluwer Academic Press, Dordrecht, 1993.

2.7. G. Prathap, A variational basis for the Barlow points, Comput. Struct. 49, 381-383, 1993.

2.8. H. C. Hu, On some variational principles in the theory of elasticity and the theory of plasticity, *Scientia Sinica,* 4, 33-54 (1955).

2.9. B. F. de Veubeke, Displacement and equilibrium models in the finite element method, in *Stress Analysis,* Ellis Horwood, England, 1980.

2.10. B. Magee, *Popper*, Fontana Press, London, 1988.