Finite element analysis and the stress correspondence paradigm

GANGAN PRATHAP

National Aerospace Laboratories, Bangalore 560 017, India, and
Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560 064, India

Abstract. The underlying mechanics of the finite element method as applied to structural analysis is explored in paradigmatic terms. It is shown that the stress correspondence paradigm has the most explanatory power and that it can be axiomatized from a very basic principle, the Hu–Washizu theorem, which is a variation of the least action principle. Numerical experiments are presented to show that the predictions based on analytical quantification from the stress correspondence paradigm are verifiable.

Keywords. Finite element analysis; stress correspondence paradigm; displacement correspondence paradigm; structural analysis; axiomatization; philosophy of science.

1. Introduction

Finite Element Analysis (FEA) is a remarkable technological and commercial success originating out of collective acts of human ingenuity, skill and craft. A hundred thousand or more engineers, technicians, teachers and students routinely use finite element analysis packages (of which there are nearly 1500 codes ranging from small dedicated in-house programmes to large general purpose mega-line codes) in design, analysis, teaching or study environments. There are billions of dollars worth of installed software and hardware dedicated to finite element analysis all over the world and perhaps billions of dollars are spent on analysis costs using this software every year. The primary archival literature has grown rapidly and at the last count there were more than 50,000 papers on the subject (excluding papers on fluid mechanics), and nearly 3800 papers on it are published annually (Mackerle 1995). There are about 400 textbooks and primers, about 400 conference proceedings and perhaps thousands of handbooks, course notes and documentation manuals.

The Finite Element Method (FEM) offers an excellent example of how a body of knowledge first emerges out of the ingenious art and craft of practising engineers, then takes shape as modes and lines of rational enquiry are set up, and then is finally shown to have a scientific basis. FEM is now formally over thirty-five years old (the terminology 'finite

element' having being coined in 1960). As it is understood both by its own practitioners and by all laymen who are, even if only remotely, aware of its scope and potential, FEM is an approximate method of solving problems that arise in science and engineering. In fact, it originated and grew as such, by intuition and inspired guess, by hard work and trial and error. Its origins can be traced to aeronautical and civil engineering practice, mainly from the point of view of structural engineering. Today, it can be used, with clever variations, to solve a wide variety of problems in science and engineering.

However, my experience with students and teachers and technicians and engineers over two decades of interaction is that many if not most are oblivious to the basic principles that drive the method. All of them can understand the 'first-order' tradition of FEM — what goes into the packages; what comes out; how to interpret results and so on. But few could put a finger on why the method does what it does; this 'second-order' tradition, to borrow a phrase from the exquisitely crafted philosophy of Sir Karl Popper (Magee 1988), of critically discussing the myths and the metaphysics of the method is available to a very few.

To understand why this is so, let us briefly review the stages through which the discipline grew. The earliest and largely technological stages are what we can call the 'hands-on' and 'handle-turning' stages of the enterprise – the design, use and re-design of algorithms and software on a trial and error basis (hands-on experience) and the drudgery-filled computation (handle-turning) phase of validation and finally production run analysis. But Art gave way to Science very slowly and very reluctantly. This sadly neglected and unsung third stage to the whole exercise, the 'hand-waving' stage as one may call it, is where the myths and superstitions of the method are created and then resolved as Science. All science is myth-making just as religion is, said Karl Popper. But scientific myths are different because one adopts a critical, argumentative attitude to these myths so that the myths "change, [changing] in the direction of giving a better and better account of the world" (Karl Popper). In this paper, we shall explore one such 'hand-waving' aspect of FEM, to learn how the method works, from the point of view of steady progression of myth to better myth.

There are very good reasons why the emergence of the Science of FEM was in fitful and uncertain steps. It may serve us well to realize that the finite element method has progressed as far as it did precisely because there was more 'art' and 'engineering' and little mathematical rigour and less 'science' in the early years of its development. The invention of the method by engineers in very intuitive ways was the heroic phase of the subject, led entirely by bold pioneers. As Robert M Pirsig described it so graphically in his *Zen and the art of motorcycle maintenance*: "Pioneers [are] invariably, by their nature, mess makers. They go forging ahead, seeing only their noble, distant goal, and never notice any of the crud and debris they leave behind. Someone else gets to clean that up." Now that the noble, distant goal has been fully realized, it's the right time to clean up. This would be the main task of this paper.

The particular issue that we shall grapple with here is: What does the finite element method set out to do in structural analysis? Does it first compute displacements in a given structural domain and then derive stresses from these, as is the most commonly held view? I call this the *displacement correspondence (DC) paradigm*. Or does it intrinsically sample stresses first and produce displacement fields only in a secondary manner — the

stress correspondence (SC) paradigm that is being proposed here. The complete burden of proof, as it is carried by rational argument, is presented here to demonstrate that one paradigm has more explanatory power than the other.

2. Paradigms, some approximate solutions, and derivation from a basic principle

2.1 Introduction

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 Π based on the energy principle whose extremum describes the equilibrium state of the problem. 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 know 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 exactly. 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 degrees of freedom α_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 (α_i or u_i as the case may be). The equilibrium configuration is obtained by applying the criterion that Π 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 displacements which are approximated to greater accuracy with improved idealization. The displacement correspondence paradigm belongs to this school. It follows from this that stresses which are computed as derivatives of the approximate displacement fields will be less accurate. Here, however, we will seek to establish the currency of a third paradigm – that the RR or FEM process

actually seeks to determine to best advantage, the state of stress or strain in the structure. In this stress correspondence paradigm, 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 the present context. Here, we use the word in the greatly enlarged sense in which the philosopher T S Kuhn (1962) introduced it in his classic study on scientific progress. 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 (Dasgupta 1994). The DC and SC paradigms can be thought of as two competing scenarios which attempt to explain how the finite element method computes displacements, strains and stresses. Our task will therefore be to establish which paradigm has greater explanatory power and range of application. Before we take up this task, let us work out a few simple problems. This is the usual epistemological sequence in which learning and experience reinforce our acceptance of one paradigm over the other.

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 (figure 1). Starting with the differential equation of equilibrium, it is easy to show that the analytical solution to the problem is

$$u(x) = (q_0/EA)(Lx - x^2/2),$$

(1a)

$$\sigma(x) = (q_0/A)(L - x).$$

(1b)

Consider a one-term RR solution based on $\bar{u}_r = \alpha_1 x$, where the subscript r denotes the use of the RR approach. It can be shown that the approximate solution obtained is
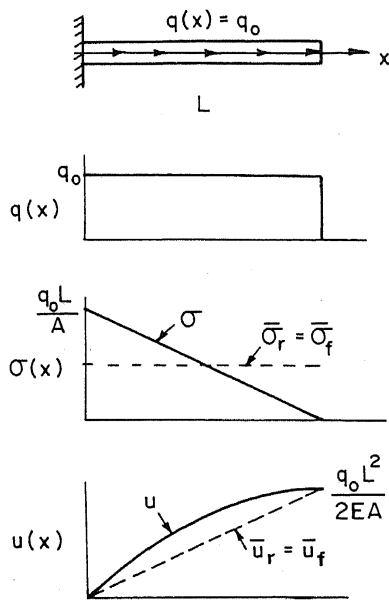
$$\bar{u}_r(x) = (q_0/EA)(Lx/2),$$

(2a)

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

(2b)

An FEM solution based on a two-noded linear element produces

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

(3a)

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

(3b)

Figure 1. Bar under uniformly distributed axial load – one-term RR and one two-node element solution.

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 may also note the curious coincidence where all three solutions predict the same displacement at the tip. However, from figure 1 we can see that the $\bar{u}_r$ and $\bar{u}_f$ are approximate solutions to u. It is also clear from figure 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 σ.

Consider next what will happen if two equal length linear (i.e., two-noded) bar elements are used to model the bar. The solution described in figure 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 figures 1 or 2 as σ. Instead, we must think of a substitute stairstep distribution σ_f, produced by the consistent load lumping process 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
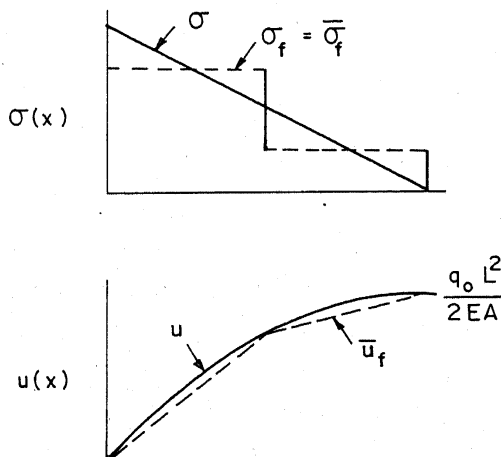
Figure 2. Bar under uniformly distributed axial load – two two-node element solution.

the observation that the stresses computed by the finite element system now approximate 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 figures 1 and 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 figures 1 and 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, appear 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 (figure 3). The exact stress distribution in this case will be quadratic in nature:

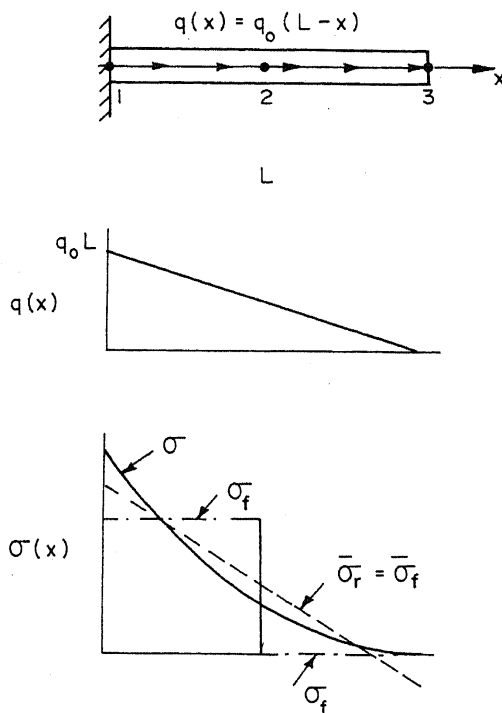$$\sigma(x) = (q_0 L^2/8A)(4/3 - 2\xi - (1 - 3\xi^2)/3). \tag{4}$$

Some interesting features about this equation can be noted down. 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 modelling the entire bar in figure 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.

We shall postpone the first obvious approximation, that of using a one-term series $\bar{u}_r = \alpha_1 x$ till later. For now, we shall consider a two-term series $\bar{u}_r = \alpha_1 x + \alpha_2 x^2$. This is chosen so that the essential boundary condition at $x = 0$ is satisfied. No attempt is made to satisfy the force boundary condition at $x = L$. By carrying out the necessary algebra associated with the RR process the solution obtained will yield an approximate stress pattern given by

$$\bar{\sigma}_r(x) = (q_0 L^2/8A)(4/3 - 2\xi). \tag{5}$$

This is plotted in figure 3 as the dashed line. A comparison of (4) and (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:

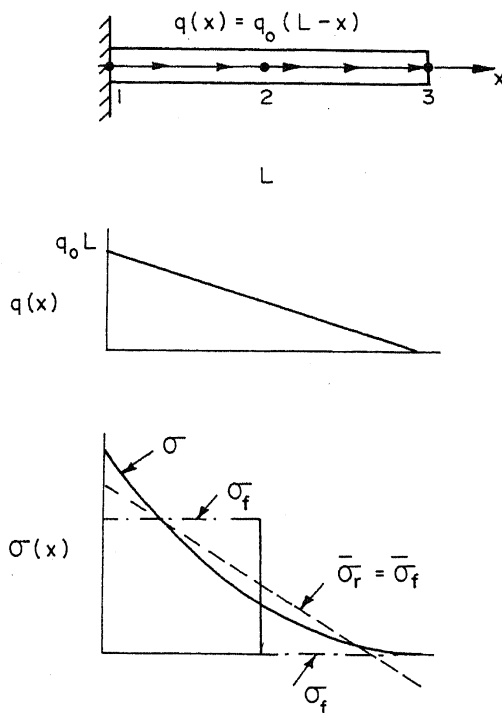$$\int_{-1}^{1} \delta\bar{\sigma}_r(\bar{\sigma}_r - \sigma)d\xi = 0. \tag{6}$$

Figure 3. Bar under linearly varying axial load – two-term RR and one three-node element solution.

That is, the RR procedure has determined a $\overline{\sigma}_r$ that is a 'best-fit' of the true state of stress σ in the sense described by the orthogonality condition in (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 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, $N_1 = \xi(\xi - 1)/2$, $N_2 = (1 - \xi^2)$ and $N_3 = \xi(\xi + 1)/2$. We first compute the consistent loads to be placed at the nodes due to the distributed load using $P_i = \int N_i q \, dx$. This results in the following scheme of loads at the three nodes identified in figure 3: $P_1 = q_0 L^2/6$, $P_2 = q_0 L^2/3$ and $P_3 = 0$. The resulting load configuration can be represented in the form of a stress system shown as σ_f, represented by the chain-dotted lines in figure 3. Thus, any FEM discretization automatically replaces a smoothly varying stress system by a step-wise system as shown by σ_f in figures 2 and 3. It is this step-wise system that the finite element solution $\overline{\sigma}_f$ 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

$$\overline{\sigma}_f(x) = (q_0 L^2/8A)(4/3 - 2\xi). \tag{7}$$

This is exactly the same as the $\overline{\sigma}_r$ computed by the RR process in (5). At first sight, this does not seem to be entirely unexpected. Both the RR and the FEM processes 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 $\overline{\sigma}_r$ was responding to a quadratically varying system σ (see figure 3). We

Figure 3. Bar under linearly varying axial load – two-term RR and one three-node element solution.

That is, the RR procedure has determined a $\overline{\sigma}_r$ that is a 'best-fit' of the true state of stress σ in the sense described by the orthogonality condition in (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 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, $N_1 = \xi(\xi - 1)/2$, $N_2 = (1 - \xi^2)$ and $N_3 = \xi(\xi + 1)/2$. We first compute the consistent loads to be placed at the nodes due to the distributed load using $P_i = \int N_i q \, dx$. This results in the following scheme of loads at the three nodes identified in figure 3: $P_1 = q_0 L^2/6$, $P_2 = q_0 L^2/3$ and $P_3 = 0$. The resulting load configuration can be represented in the form of a stress system shown as σ_f, represented by the chain-dotted lines in figure 3. Thus, any FEM discretization automatically replaces a smoothly varying stress system by a step-wise system as shown by σ_f in figures 2 and 3. It is this step-wise system that the finite element solution $\overline{\sigma}_f$ 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

$$\overline{\sigma}_f(x) = (q_0 L^2/8A)(4/3 - 2\xi). \tag{7}$$

This is exactly the same as the $\overline{\sigma}_r$ computed by the RR process in (5). At first sight, this does not seem to be entirely unexpected. Both the RR and the FEM processes 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 $\overline{\sigma}_r$ was responding to a quadratically varying system σ (see figure 3). We

could easily establish through (6) that $\bar{\sigma}_r$ responded to σ in a 'best-fit' manner. However, in the FEM process, the load system that is being used is the σ_f system which varies in the stairstep fashion. The question confronting us now is, in what manner did $\bar{\sigma}_f$ respond to σ_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 $\sigma_f = 2q_0L^2/3$ 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:

$$\int_{-1}^{0} \delta\bar{\bar{\sigma}}(\bar{\bar{\sigma}} - 2q_0L^2/3)\mathrm{d}\xi + \int_{0}^{1} \delta\bar{\bar{\sigma}}(\bar{\bar{\sigma}} - 0)\mathrm{d}\xi = 0. \tag{8}$$

It can be worked out that this leads to

$$\bar{\bar{\sigma}}(x) = (q_0L^2/8A)(4/3 - 2\xi), \tag{9}$$

which is identical to the result obtained in (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 σ_f!

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. Note that the approximate solutions $\bar{\sigma}_r$ or $\bar{\sigma}_f$ intersect the exact solution σ at two points. A comparison of (4) with (5) and (7) indicate that these are the points where the quadratic Legendre polynomial, $(1 - 3\xi^2)$, 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. 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 $\xi = \pm 1/\sqrt{3}$.

We shall now return to the linear Ritz admissible function, $\bar{u}_r = \alpha_1 x$, 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 $\sigma(x)$ given by equation (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:

$$\bar{\sigma}_r(x) = (q_0L^2/8A)(4/3). \tag{10}$$

What becomes obvious by comparing this with the true stress $\sigma(x)$ in (4) and the computed stress from the two-term solution, $\bar{\sigma}_r(x)$ in (5) is that the one-term solution corresponds to the constant part only 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 $\xi = 1 - (4/3)^{1/2}$.

Table 1. The conceptual frameworks for the displacement correspondence (DC) and stress correspondence (SC) paradigms.

Displacement correspondence paradigm	Stress correspondence paradigm
Displacements at nodes are matched	Stresses at optimal points are matched
Stresses computed as derivatives from displacements	Displacements "integrated" from stresses
Differentiation produces functions which are less accurate than original functions	Integration produces functions which are more accurate than original functions
∴ displacements are accurate; stresses are less accurate.	∴ stresses are accurate; displacements are more accurate (on average).
∴ if there are points where stresses are very accurate, this is a consequence of the mean value theorem	∴ if there are points where stresses are very accurate, this is because stresses are approximated in a "best-fit" sense of true stresses.

2.5 *The DC, SC and aliasing paradigms*

The preceding examples have been so selectively chosen that we seem to have made out a very strong water-tight case for the SC paradigm. Let us however examine the claims and merits if any of the competing DC paradigm for this problem. 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 the original derivation of optimal points (Barlow 1976) – the term *substitute function* is used instead of alias, and is also the basis for what is called the 'aliasing' paradigm (MacNeal 1994). We recognize that this is precisely what we mean by the DC paradigm here.

It is helpful to use the aliasing metaphor to explain what happens in finite element analysis (FEA). 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 (as MacNeal argued) or stresses at optimal points (as argued here) and the alias is the function which interpolates the displacements within the element from the nodal displacements or stresses in the element from the stresses at the optimal points respectively. The FEA can be imagined to be a sampler or processor which translates the true input to an approximate or computed output.

Table 1 summarizes the competing conceptual frameworks presented by the displacement and stress correspondence paradigms. Both paradigms seem heuristically very

Table 2. Barlow and Gauss points for one-dimensional case.

p	Node locations	u	$\bar{u}$	ϵ	$\bar{\epsilon}$	Gauss points	Barlow points	
							SC	DC
1	± 1	ξ^2	ξ	ξ	1	0	0	0
2	$0, \pm 1$	ξ^3	ξ^2	ξ^2	ξ	$\pm 1/\sqrt{3}$	$\pm 1/\sqrt{3}$	$\pm 1/\sqrt{3}$
3	$\pm 1/3, \pm 1$	ξ^4	ξ^3	ξ^3	ξ^2	$0, \pm(3/5)^{1/2}$	$0, \pm(3/5)^{1/2}$	$0, \pm\sqrt{5}/3$

$1, \zeta, \ldots, \zeta^4$ indicate polynomial orders from constant to quartic

appealing. Which would one prefer? The DC paradigm is the universally accepted belief. What is required now to establish that the SC paradigm is superior?

Let us now use the DC concept to derive the location of the optimal points, as Barlow did in 1976, or as MacNeal did more recently in 1994. 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 SC 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 DC approach leads to subtle differences in interpreting the relationship between the Barlow points and the Gauss points.

2.5a *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, 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 fields will be designated by u, ϵ and σ. The discretized displacement, strain and stress fields will be designated by $\bar{u}$, $\bar{\epsilon}$ and $\bar{\sigma}$. The DC displacement, strain and stress fields will be designated by u^d, ϵ^d and σ^d. Nodal displacements will be represented by u_i.

We shall examine a simple scenario where the true displacement field u is exactly one polynomial order higher than what the finite element is capable of representing – the Barlow points can be determined exactly in terms of the Gauss points only for this case.

We shall now take for granted that the best-fit rule operates according to the orthogonality condition expressed in (6) and that it can be used interchangeably for stresses and strains. We shall designate the optimal points determined by the DC algorithm as ξ_d, the Barlow points (DC), and the points determined by the SC algorithm as ξ_s, the Barlow points (SC). Note that ξ_d are the points established by Barlow (1976) and MacNeal (1994), while ξ_s will correspond to the points given in Prathap (1993). The natural coordinate system ξ is used here for convenience.

Table 3. The Legendre polynomials P_i.

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

Thus, for the present case, the use of the SC paradigm leads to

$$\int \delta\bar{\epsilon}^T (\bar{\epsilon} - \epsilon)dV = 0. \tag{11}$$

This case corresponds to one in which a straightforward application of Legendre polynomials can be made. In this case, one can determine the points where $\bar{\epsilon} = \epsilon$ as those corresponding to points which are the zeros of the Legendre polynomials. See table 3 for a list of unnormalised Legendre polynomials. We shall show below that in (11), the points of minimum error are the sampling points of the Gauss-Legendre integration rule if $\bar{\epsilon}$ is exactly one polynomial order lower than ϵ.

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 $(p = 1)$

$$u = \text{quadratic} = b_0 + b_1\xi + b_2\xi^2,$$

$$\epsilon = \text{linear} = u_{,\xi} = b_1 + 2b_2\xi = \sum_{i=0}^{p=1} \epsilon_i P_i(\xi).$$

Note that we have written ϵ 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 DC and SC approaches.

DC: At $\xi_i = \pm 1$, $u_i^d = u_i$; then points where $\epsilon^d = \epsilon$ are given by $\xi_d = 0$. Thus, the Barlow point (DC) is $\xi_d = 0$, for this case.

SC: $\bar{u} = linear$, is undetermined at first. Let $\bar{\epsilon} = c_0$, as the element is capable of representing only a constant strain. Equation (11) will now give $\bar{\epsilon} = c_0 = b_i$. Thus, the optimal point is $\xi_s = 0$, the point where the Legendre polynomial $P_1(\xi) = \xi$ vanishes. Therefore, the Barlow point (SC) for this example is $\xi_s = 0$.

Quadratic element ($p = 2$)

$$u = \text{cubic} = b_0 + b_1\xi + b_2\xi^2 + b_3\xi^3,$$

$$\epsilon = \text{quadratic} = u_{,\xi}$$

$$= (b_1 + b_3) + 2b_2\xi - b_3(1 - 3\xi^2) = \sum_{i=0}^{p=2} \epsilon_i P_i(\xi).$$

DC: At $\xi_i = 0, \pm 1$, $u_i^d = u_i$; then points where $\epsilon^d = \epsilon$ are given by $\xi_d = \pm 1/\sqrt{3}$. Thus, the Barlow points (DC) are $\xi_d = \pm 1/\sqrt{3}$, for this case.

SC: $\bar{u} = quadratic$. Let $\bar{\epsilon} = c_0 + c_1\xi$, as this element is capable of representing a linear strain. Equation (5) will now give $\bar{\epsilon} = (b_1 + b_3) + 2b_2\xi$. Thus, the optimal points are $\xi_s = \pm 1/\sqrt{3}$, the points where the Legendre polynomial $P_2(\xi) = (1 - 3\xi^2)$ vanishes. Therefore, the Barlow points (SC) for this example are $\xi_s = \pm 1/\sqrt{3}$.

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). This also explains why for a very long time, the DC paradigm remained the prevailing wisdom. In our next example we will find that this is not so anymore.

Cubic element ($p = 3$)

$$u = \text{quartic} = b_0 + b_1\xi + b_2\xi^2 + b_3\xi^3 + b_4\xi^4,$$

$$\epsilon = \text{cubic} = u_{,\xi}$$

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

$$= \sum_{i=0}^{p=3} \epsilon_i P_i(\xi).$$

DC: At $\xi_i = \pm 1/3, \pm 1$, $u_i^d = u_i$; then points where $\epsilon^d = \epsilon$ are given by $\xi_d = 0$, $\pm\sqrt{5}/3$. Thus, the Barlow points (DC) are $\xi_d = 0, \pm\sqrt{5}/3$, for this case. Note that the points where the Legendre polynomial $P_s(\xi) = (3\xi - 5\xi^3)$ vanishes are $\xi_G = 0, (3/5)^{1/2}$!

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

Therefore, we have an example where the DC paradigm does not give the correct picture about the way the finite element process computes strains. However, the SC 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 summarizes the results obtained so far.

Our experience is that the SC model is the one that corresponds to reality – that 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, and not as predicted by the DC paradigm (see § 3.1 below).

2.6 Axiomatization :The 'best-fit' nature of the SC paradigm from a variational theorem

Our investigation here will be complete in all respects if the best-fit nature of stress correspondence can be deduced logically and quantitatively from a basic principle. This is the process known as *axiomatization* – the motivation is now to compress the paradigm or derive it out of a single or minimal set of axioms or fundamental principles. In fact, some recent work (Prathap 1993) shows that 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. With the help of hindsight we know that a similar axiomatization from a basic principle does not seem to be possible for the DC paradigm.

The 'best-fit' manner in which finite elements compute strains can be shown to follow from an interpretation using 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 (Hu 1955) rather than the minimum potential theorem. These theorems belong to a family of most basic statements (the least action principles) that can be made about the laws of nature, of matter, motion and energy. The minimum potential theorem corresponds to the conventional energy theorem. However, for applications to problems in structural and solid mechanics, Hu (1955) proposed a generalized theorem which had somewhat more flexibility. Its usefulness came to be recognized when one had to grapple with some of the problems raised by finite element modelling. 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 ϵ and stress field σ (we project that the strain field ϵ 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 $\bar{u}$, $\bar{\epsilon}$ 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 $\bar{\epsilon}$ is an approximation of the true strain field ϵ. What the Hu–Washizu theorem does 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 ϵ and the discretized strain field $\bar{\epsilon}$. Note that $\bar{\bar{\sigma}}$ is now an approximation of σ. The three-field Hu–Washizu theorem can be stated as

$$\delta \pi = 0, \tag{12}$$

where

$$\pi = \int \{\overline{\sigma}^T \overline{\epsilon}/2 + \overline{\overline{\sigma}}^T (\epsilon - \overline{\epsilon})\} dV + P, \tag{13}$$

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 text-books, only one field (i.e., the displacement field u), is subject to variation. However, in this more general three-field approach, all three fields are subject to variation and this leads to three sets of equations which can be grouped and classified as follows.

Variation on	Nature	Equation	
u	Equilibrium	$\nabla \overline{\overline{\sigma}} + \text{terms from } P = 0,$	(14a)
$\overline{\overline{\sigma}}$	Orthogonality (Compatibility)	$\int \delta \overline{\overline{\sigma}}^T (\overline{\epsilon} - \epsilon) dV = 0,$	(14b)
$\overline{\epsilon}$	Orthogonality (Equilibrium)	$\int \delta \overline{\epsilon}^T (\overline{\overline{\sigma}} - \overline{\sigma}) dV = 0.$	(14c)

Equation (14a) shows that the variation on the displacement field u requires that the independent stress field $\overline{\overline{\sigma}}$ must satisfy the equilibrium equations (∇ signifies the operators that describe the equilibrium condition). Equation (14c) is a variational condition to restore the equilibrium imbalance between $\overline{\sigma}$ and $\overline{\overline{\sigma}}$. In the displacement type formulation, we choose $\overline{\overline{\sigma}} = \overline{\sigma}$. This satisfies the orthogonality condition seen in (14c) identically, and leaves us with the orthogonality condition in (14b). We can now argue that this tries to restore the compatibility imbalance between the exact strain field ϵ and the discretized strain field $\overline{\epsilon}$. In the displacement type formulation this can be stated as

$$\int \delta \overline{\sigma}^T (\overline{\epsilon} - \epsilon) dV = 0. \tag{15}$$

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. There is therefore a uniquely defined correspondence between the approximate stress and the true stress in such finite element computations.

3. Numerical experiments

So far, our knowledge has been based on theoretical derivations from fundamental principles and paradigms which originated from intelligent or intuitive conjecture or guess. The deductions we made quantitatively in § 2 from the stress correspondence paradigm belonged to this category. Bertrand Russell had pointed out that knowledge based only on universal principles is sterile; it is the world of Platonic ideas. Science needs proof, in the form of empiricism – only then does it become complete. In our present discipline, which is finite element modelling of problems in structural engineering, these experiments would

Table 4. Case a – Where are the optimal stress points? $(-1 < \xi < -1)$.

Element	Observed/ computed	Expected/predicted	
		SC	DC
2-node	0	0	0
3-node	$\pm 1/\sqrt{3}$	$\pm 1/\sqrt{3}$	$\pm 1/\sqrt{3}$
4-node	$0, \pm\sqrt{3}/5$	$0, \pm\sqrt{3}/5$	$0, \pm\sqrt{5}/3$

not be actual physical ones but would be numerical, computational or digital in nature. We must therefore adopt the following course of action to ensure that our understanding of the stress correspondence paradigm is scientifically complete and coherent: Starting with a guess (that the stress correspondence paradigm gives the correct description of the finite element analysis procedure), we predict quantitatively the various consequences of the guess (for example, where the optimal stress points should be), and then design and conduct numerical experiments that would verify (falsify) that the real, observed situation (as obtained from a routine 'first order' tradition of finite element computation) agree (disagree) with these predicted or expected values (from the 'second order' traditional exercise).

In § 2.5a we had elaborated on a simple bar element model under various loading conditions and we had made predictions for the location of the optimal points according to the DC and SC paradigms. This will now be experimentally verified as case a below. Three other examples we choose are: Case b – the rate of convergence of a beam element; Case c – the transverse shear stress for the fundamental thickness shear mode of a hinged-hinged beam; and case d – the shear force resultant in an axisymmetric circular plate. Case b is based on Timoshenko theory, and cases c and d use elements based on a higher order shear deformation theory.

3.1 Case a – Location of optimal stress points

In § 2.5a we worked out analytically what the location of the optimal points according to the DC and SC paradigms are. We shall now perform numerical experiments using 2-node, 3-node and 4-node bar elements and in each case apply distributed axial loads whose intensity will vary to produce linear, quadratic and cubic variations of axial strain along the length. Table 4 compares the results observed/computed from experiment with those deduced analytically from the competing paradigms and listed under the column expected/predicted. Note that for the two simpler elements, the DC paradigm proved to be deceptively accurate; it is only in the cubic element that it is seen that only the SC paradigm makes the correct prediction.

3.2 Case b – Rate of convergence for a tip-loaded cantilever idealized with linear Timoshenko beam elements

Figure 4 shows a cantilever beam under tip load. The dimensions are chosen such that the tip deflection under the load will be $w = 4.0$. The example chosen represents a thin beam
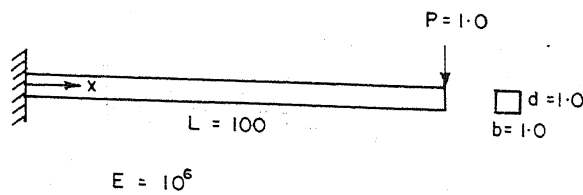
Figure 4. Cantilever beam under tip load.

so that the influence of shear deformation and shear strain energy is negligible. The finite element idealization is performed using equal length two-node C^0 beam elements based on independent linear interpolations of the transverse displacement w and normal rotation θ. This element permits constant bending and shear strain accuracy within each element – the simplest representation possible under the circumstances and therefore an advantage in seeing how it works in this problem.

The second column of table 5 shows the tip deflections obtained from the finite element digital computation. This was the actual epistemological sequence in which the understanding was obtained – these results were known to this writer in 1980, much before any explanatory paradigm was offered. What was noticed was that if w is the true (i.e. analytical) solution to the exact problem and $\overline{w}_0$ the deflection observed from the finite element experiment, then the quantity $\{(w - \overline{w}_0)\}/w$, turned out to be exactly $1/4N^2$, where N is the number of elements used. The predictions based on the SC paradigm were made much later, around 1988.

The challenge now is to see if this relationship describing the rate of convergence can be established by arguing that it emerges from the fact that FEA operates according to the SC paradigm and that within this paradigm, strains are sought in the 'best-fit' manner. We pay attention to the bending moment variation observed from the finite element model and its relation to the actual bending moment variation in this problem. Figure 5 shows the bending moments obtained from 1, 2 and 4-element idealizations of the present problem. The true bending moment (shown by the solid line) varies linearly. The computed bending moments are distributed in a piecewise constant manner as shown by the broken lines. In each case, the elements pick up the bending moment at the centroid correctly i.e., in a 'best-fit' manner. We shall now attempt to relate this to the accuracy (and convergence) of results.

Consider the case where the beam is modelled by equal length beam elements, so that any beam region of length L is replaced by an element of length $2l$. Let the moment and

Table 5. Tip deflections of a thin cantilever beam, $L/t = 100$.

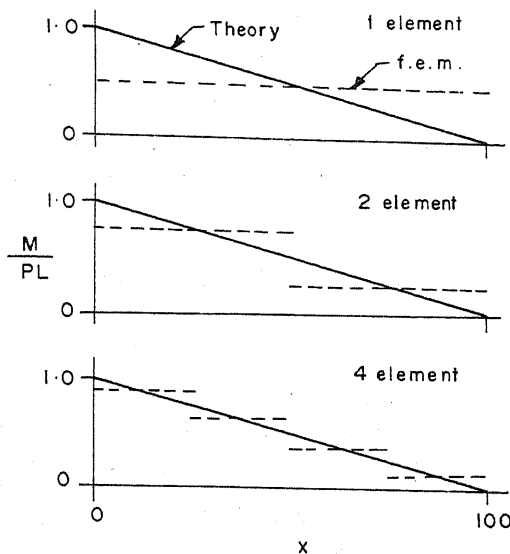| No. of elements	Observed/ computed $\overline{w}_0$	Expected/ predicted $\overline{w}_e$
1	3.0000	3.0000
2	3.7500	3.7500
4	3.9375	3.9375

Figure 5. Bending moment diagrams for 1-, 2- and 4-element idealizations of a cantilever beam under tip load.

shear force at the centroid be M and V. Thus the true bending moment over the element region for our problem can be taken to vary as $M + Vx$ (this follows from the simple fact that equilibrium requires the rate of change of bending moment to be equal to the shear force). The discretized bending moment sensed by our linear element would therefore be M – it cannot do any better. We shall now compute the actual bending energy in the element region (i.e., from a continuum analysis) and that given by the finite element model. We can show that

$$\text{Energy in continuum model} = (l/EI)(M^2 + V^2l^2/3), \qquad (16a)$$

$$\text{Energy in discretized model} = (l/EI)(M^2). \qquad (16b)$$

Thus, as a result of the discretization process involved in replacing each continuum segment of length $2l$ by a linear Timoshenko beam element which can give only a constant value M for the bending moment, there is a reduction (error) in energy in each element equal to $(l/EI)(V^2l^2/3)$. It is simple now to show from this that for the cantilever beam of length L with a tip load P, the total reduction in strain energy of the discretized model for the beam is $U/4N^2$ where $U = P^2L^3/6EI$ is the energy of the beam under tip load.

We can now show that this error in strain energy translates into an error in the deflection under load P. From (16a) and (16b) and Castigliano's second theorem, it can be deduced that the tip deflection, w, of the continuum, and that expected (or predicted) from the SC paradigm, $\overline{w}_e$, will differ as $w - \overline{w}_e/w = 1/4N^2$. The third column in table 5 shows this expected or predicted rate of convergence. This follows from the fact that if any linear variation is approximated in a piecewise manner by constant values as seen in figure 5, this is the manner in which the square of the error in the stresses/strains (or, equivalently, the difference in work or energy) will converge with idealization.

The agreement between observed and expected results is exact. Thus from the SC paradigm and using a simple example, we can deduce quantitatively that the rate of convergence observed in the numerical experiment could be exactly predicted. This reinforces again our conviction that the SC paradigm is the correct description of the finite element process.
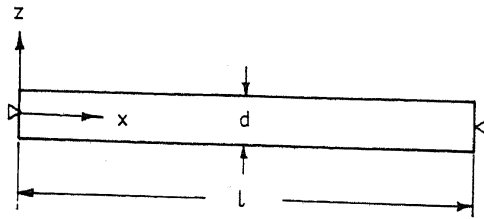
Figure 6. Hinged-hinged beam undergoing fundamental thickness shear vibration.

3.3 Case c – the transverse shear stress distribution for the fundamental thickness shear mode of a hinged-hinged beam

Figure 6 shows a hinged-hinged beam of depth d, length l and rectangular cross-section and for simplicity an assumed shear modulus $G = 1$. It is possible for the beam to vibrate in what is called a fundamental thickness-shear mode – without transverse deflection (i.e., $w = 0$), the displacement being entirely parallel to the neutral axis. The mode-shape describing the cross-sectional distortion through the depth is

$$u = \sin(\pi z/d), \tag{17}$$

where z is measured from the neutral axis of the beam. This implies that the shear strain/stress distribution varies as

$$\tau_{xz} = (\pi/d)\cos(\pi z/d), \tag{18}$$

through the thickness. Note that this distribution satisfies the traction-free conditions specified at the top and bottom surfaces. The nature of motion is such that it produces a pure shear stress state that varies only through the thickness and not along the length.

We shall now model this problem using finite elements based on what is called a higher-order shear deformation theory. The displacement field chosen for such a problem is quasi-two-dimensional (stresses are now functions of the x- and z-axes), as can be seen below:

$$u = u_0 + z\theta + z^2 u_0^* + z^3\theta^*, \tag{19a}$$

$$w = w_0 + z\psi + z^2 w_0^*. \tag{19b}$$

If a three-noded beam element is the basis for the finite element formulation, then the degrees of freedom are interpolated in the x-direction using quadratic shape functions. Also, the variations in the z-direction are as depicted in (19). The discretization process is now two-dimensional – in addition to the approximation along the beam axis (x-axis) represented by interpolations of variables defined at nodes spaced along the length using conventional shape functions, there is also an approximation in the thickness direction (z-axis) represented as a Taylor series expansion in terms of variables available at the node at that location. This allows us to examine the nature of the computed stress variation in the thickness direction when elements permitting this are used to see if the best-fit stress correspondence paradigm covers this problem as well.

Computations for hinged-hinged beams of various slenderness ratios (d/l ranging from 0.1 to 3) showed that the lowest thickness-shear frequency was picked up accurately. The cross-sectional distortion pattern obtained from the FEM model for the thickness-shear mode showed that the vibration had no transverse or symmetric axial deflection (w_0, ψ,
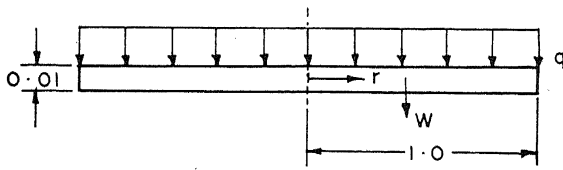
Figure 7. Axisymmetric circular plate under uniformly distributed transverse load.

w_0^*, u_0, $u_0^* \leq 10^{-14}$) and only a predominantly antisymmetric axial deflection pattern characterized by $\theta^* d^2/\theta = -1.399$ for all the values of d/l considered and at all sections along the length of the hinged-hinged beam.

We shall now try to rationalize the observation that the mode-shape obtained from the finite element computations was characterized by the value $\theta^* d^2/\theta = -1.399$ for all d/l considered. We can simplify the analysis by noting that for this mode, the following description suffices:

$$\bar{u} = z\theta + z^3\theta^*, \tag{20a}$$

$$\bar{\tau}_{xz} = \theta + 3z^3\theta^*. \tag{20b}$$

We shall derive predictions using both the DC and SC paradigms; in the former we shall argue that $\bar{u}$ replaces u in a best-fit manner and that in the latter, $\bar{\tau}_{xz}$ replaces τ_{xz} in a similar manner. The variation δ is carried out over the generalized degrees of freedom θ and θ^* leading to two simultaneous equations in each case and finally to a value of $\theta^* d^2/\theta$ for each case.

$$\text{(DC)} \int \delta\bar{u}^T (\bar{u} - u)\mathrm{d}z = 0 \quad \text{gives} \quad \theta^* d^2/\theta = -1.448,$$

$$\text{(SC)} \int \delta\bar{u}_{xz}^T (\bar{\tau}_{xz} - \tau_{xz})\mathrm{d}z = 0 \quad \text{gives} \quad \theta^* d^2/\theta = -1.402.$$

Only the SC prediction agrees very closely with the computed value of -1.399. The DC paradigm is therefore not successful in predicting this factor.

3.4   *Case d – the shear force resultant in an axisymmetric circular plate*

Figure 7 shows a simply-supported circular plate loaded by a uniformly distributed transverse loading of intensity $q = 1.0$. It is modelled using axisymmetric plate elements based on the same higher-order theory that was used in case c above (see § 3.3). Again, three-node elements are used – these elements are capable of computing the shear stress resultant exactly to a linear variation along the length of the element. Interestingly, for this problem, the shear stress resultant $Q = qr/2$ varies linearly along the radius of the plate. Thus, according to the SC paradigm, the finite element model should be able to pick up the shear stress resultant exactly.

Table 6 shows the results from a four-element model of the circular plate. We see from the fourth and fifth columns that the observed values (from the FE computation) and the expected values (from the SC paradigm predictions) agree to seven decimal figures. On the other hand, the observed and expected transverse deflections agree only to two decimal places! This is of course a specially chosen example to highlight the SC paradigm. In most general computation, the computed displacements are, on the average, of greater reliability than computed stresses. However, the SC paradigm allows us to take advantage

Table 6. Case d – Transverse displacement w and shear force resultant Q for a simply-supported axisymmetric circular plate under uniformly distributed load from a 4-element model using 3-node higher-order axisymmetric plate elements.

r	w		Q	
	Observed	Expected	Observed	Expected
0.00	0.1902	0.1875	0	0
0.25	0.1658	0.1648	0.1250000	0.1250000
0.50	0.1064	0.1055	0.2500000	0.2500000
0.75	0.0357	0.0359	0.3750000	0.3750000
1.00	0	0	0.5000000	0.5000000

of the fact that stresses at optimal points are of comparable or greater accuracy than the displacements.

4. What does the finite element method do?

The persistence of the DC paradigm for a very long time can be attributed to the fact that it is a very natural or obvious or common-sensical interpretation of what seems to be happening in a finite element computation. After all, at the end of the computation, it is the global degrees of freedom that are usually the nodal displacements, which are presented first and the strains/stresses are seemingly processed from these displacements. It was therefore 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 fields. 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 centres – this, it is thought, is a consequence of the mean value theorem for derivatives.

However, we have demonstrated theoretically and by using numerical results that, in actual fact, the Ritz approximation process and the displacement type FEM, which can be interpreted as a piecewise Ritz procedure, do exactly the opposite and more unnatural or less common-sensical thing – 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, but the present writer feels that the most intellectually satisfying proof can be arrived at by starting with the Hu–Washizu theorem. 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 globally assembled stiffness matrix

equation actually reflecting this integration process and the continuity of fields across element boundaries, and the suppression of the field values at domain edges reflecting the imposition of boundary conditions. It must therefore be argued that displacements are on an average, more accurate than strains because integrals of smooth functions appear generally more accurate than the original data. We have thus turned the whole argument on its head.

5. Conclusions

In this paper, we postulated a few models to explain how displacement type FEM works. We worked out a series of simple problems of increasing complexity to establish whether 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 σ and σ_f were actual or true physical states that were sought to be modelled. The stress terms $\overline{\sigma}_r$ and $\overline{\sigma}_f$ were the quantities that emerged in what we can call the 'first-order' tradition analysis in the language of Sir Karl Popper – where the RR or FEM operations are mechanically carried out using functional approximation and finite element stiffness equations respectively. We noticed certain features which seemed to relate these computed stresses to the true system they were modelling 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 quantities 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' and stress correspondence paradigms had successfully survived a falsification test.

Another important step we took was to prove that the 'best fit' nature of the stress correspondence 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, which is a variation of the least action principle, was found valuable to determine that the best-fit paradigm had a rational basis.

One important conclusion we can derive from the best-fit nature of the stress correspondence paradigm is that an interpolation field for the stresses $\overline{\sigma}$ (or stress resultants as the case may be) which is of higher order than the strain fields ϵ 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, 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.

The scientist-philosopher Lewis Wolpert, argued in his influential *The unnatural nature of science* that

"... the world is not constructed on a common-sensical basis. ...'natural' thinking – ordinary, day-to-day common sense – will never give an understanding about the nature of science. Scientific ideas are, with rare exceptions, counter-intuitive ... common sense is prone to error when applied to problems requiring rigorous and quantitative thinking; lay theories are highly unreliable."

It is very easy to see that the DC paradigm had direct commonsensical appeal. The SC paradigm is counter-intuitive and requires rigorous quantitative analysis to establish its validity.

References

Barlow J 1976 Optimal stress locations in finite element models. *Int. J. Numer. Methods Eng.* 10: 243–251

Dasgupta S 1994 Understanding design: Artificial intelligence as an explanatory paradigm. *Sādhanā* 19: 5–21

Hu H C 1955 On some variational principles in the theory of elasticity and the theory of plasticity. *Sci. Sin.* 4: 33–54

Kuhn T S 1962 *The structure of scientific revolution* (Chicago: University Press)

Mackerle J 1995 Some remarks on progress with finite elements. *Comput. Struct.* 55: 1101–1106

MacNeal R H 1994 *Finite elements: Their design and performance* (New York: Marcel Dekker)

Magee B 1988 *Popper* (London: Fontana)

Prathap G 1993 *The finite element method in structural mechanics* (Dordrecht: Kluwer Academic Press)