Field-consistency – toward a science of constrained multi-strain-field finite element formulations

G PRATHAP

Structural Sciences Division, National Aeronautical Laboratory, Bangalore 560 017, India

Abstract. The finite element method has progressed rapidly from simple engineering principles through rigorous mathematical formalisations to become a body of engineering knowledge incorporating well-defined conceptual schemes and operational procedures. The current practice however leads to a series of difficulties which are actually inter-related. In this paper, we describe an additional conceptual scheme called 'field-consistency', and demonstrate new methodology and terminology that helps to clear up this area of the finite element method that had remained difficult to understand for some time. The new procedure also enhances the predictive capabilities of error analyses of the method, giving it some of the self-criticality needed for a 'falsifiable' scientific basis. This paper is also in a limited sense, an essay on how the processes of the method of science are seen to work in this area of engineering know-how.

Keywords. Constrained strain-field problems; finite element; shear locking; membrane locking; parasitic shear; field-consistency; functional reconstitution; additional stiffening parameter; the scientific method; falsifiability.

1. The field-consistency paradigm – introduction

1.1 *History of the finite element method*

The finite element method, as we know it today, was invented by practising aeronautical engineers and its early developments proceeded from simple engineering judgement. Its initial acceptance as a valid and reliable method of structural analysis was based on an *a posteriori* understanding of the results obtained by it in terms of known analytical results of familiar and tractable structural problems (Turner *et al* 1956; Argyris 1960).

The method soon began to accept more rigorous mathematical formalisations of its procedures. This involved the identification of variational and weighted residual principles such as those of minimum total potential, complementary energy, mixed

theorems of Hellinger-Reissner and Hu-Washizu and virtual work theorems. These now offered rigorous operational procedures that not only accommodated the physics of the problems but also the mathematics of the discretisation process itself. To this extent, the finite element method was becoming a formal science.

However, the weakest link in the chain remained that of the actual choice of the discretisation domain (in terms of shape and size and orientation with respect to coordinate axes) and the shape function definitions over that domain that would produce the weighted residual discretised approximations of the continuum behaviour in that domain. Here, the method was practised more in the form of a skilled 'art' rather than as a precise 'science'. We shall presently see what we mean by this.

The strategy of discretisation was based on the quick acceptance of a broad conceptual scheme which crystallised into a conventional wisdom that was readily backed by the authority of teacher and text. Around 1977 (Zienkiewicz 1977), the cardinal principles that informed this wisdom were summarised thus – the piecewise approximations used over element domains should be based on complete polynomials: these should be compatible, i.e., satisfy the continuity of these functions or their derivatives, where required, across inter-element boundaries; these functions should be able to represent states of constant strain in the limit; and finally, that these functions should be able to recover strain-free rigid body motions.

1.2 *Introduction of field-consistency concepts*

At this stage of its development, the conceptual scheme outlined above still suffered from many inner contradictions. It was a matter of great embarrassment that elements derived rigorously from the principles outlined above could still behave in dramatically erratic ways when applied to certain important practical applications, even though these same elements were in many other respects, excellent. Some familiar examples of these were known under the names 'shear locking', 'membrane locking', 'parasitic shear' etc., some of which we shall examine later. The terms of the conceptual schemes embodied in the conventional wisdom so far did not therefore provide all the sufficient conditions. Conversely, it was possible to design elements which deliberately violated one or more of these conditions and which could behave extremely well in these critical applications – an indication that some of these conditions were not even necessary ones in all instances.

Some of these problems were recorded in the literature quite early as 'locking', 'parasitic shear' etc. and quite often, it is believed now, some went unrecorded as there was no adequate framework or terminology with which to classify it (Zienkiewicz 1977). Nor was it known whether these separate problems were related in a special way or whether the problem of violently oscillatory stresses within elements was related to these phenomena. Initial attempts to explain away these deviations were made in such terms as the rank, the singularity or the lack of it, or of the spectral condition numbers of the matrices involved. Other arguments included counting of constraints and free degrees of freedom etc. (Zienkiewicz 1977). However, it was soon realised that these heuristic arguments could at best be symptoms of the general syndrome rather than the origin of the actual trouble.

The need for a new conceptual scheme, not so much to supercede the old wisdom as to complement it, was apparent. The early work of the author and his colleagues (Prathap & Viswanath 1982; Prathap & Bhashyam 1982; Somashekar 1983; Viswanath & Prathap 1983) focussed on some of the strategies or 'tricks' that were used to salvage elements that suffered from deficiencies such as shear locking, membrane locking, parasitic shear, locking at incompressible limits etc. This was performed within the context of the prevailing paradigms and its inadequacies, especially with respect to its inability to provide a 'falsifiability' capability of error analysis, became quite obvious. In the terminology of Kuhn (1962), a 'paradigm change' was called for. Subsequent work on curved beams (Prathap 1985a), plane stress modelling of flexure (Prathap 1985b), plates and shells (Prathap 1985c,d) and bricks (Prathap 1984a,b) led to the formulation of the new paradigm, called for brevity, the field-consistency principle, and definition of new terminology (Prathap 1984c,d), a new error norm (Prathap 1985e) and its use to bring all these phenomena under a single classification.

1.3 *Field-consistency as a 'falsifiable' scientific basis*

Before passing on to the main body of the new theory, we shall briefly consider here how the identification of the problem area, the founding of a conceptual basis to examine the problems and tackle them and see why they are effective, can be examined in the context of the processes of the method of science, and in particular, the ideas of Popper (1972) and Kuhn (1962) on the scientific method.

The test of any body of knowledge to succeed as a science is established from its ability to proceed from its conceptual schemes to predict or anticipate *a priori*, a wide range of observed, observable or verifiable phenomena, using specially designed operational procedures. In this paraphrasing of Popper's falsifiability theorem (Popper 1972), we shall consider the body of knowledge known as the 'finite element method' to have become a true science only when it incorporates within itself specific procedures which constitute the elements that can contribute to its own self-criticism.

We therefore seek to associate a certain paradigm or conceptual scheme that guides the selection of the discretisation process involved in the finite element method so that predictable results will always emerge. The critical aspects of this are seen to be the choice of shape and size of the elements and the form of the interpolation functions chosen. It is further necessary to incorporate auxiliary procedures that can anticipate the type and degree of error that will result from the discretisation procedure alone.

Having done this, it is logical to see if these *a priori* anticipations are really true by carrying out actual digital computations with the finite element. Thus, in the finite element method, we are in a position to perform numerical experiments, which could confirm or falsify the predictions we make, starting from the postulates of the conceptual schemes that inform the critical stage of this discretisation process. These could be viewed as Kantian experiments, in the terminology of Medawar (1979, pp. 69–77).

The work initiated by the writer and his colleagues has been directed towards founding an additional conceptual scheme called field-consistency, defining the new terminology to give it a proper vocabulary, devising an operational procedure

called the functional re-constitution technique to give it a self-critical error analysis capability and using it to examine a series of problems in the finite element method that had defied a proper 'scientific' analysis earlier.

In this paper, we review the work done in this area and its possible applications to the design and development of finite element software.

2. A brief overview of the field-consistency approach

We describe very briefly in this section, the basic features of the principle. In subsequent sections, we shall take up a more comprehensive description of the conceptual scheme and procedures. We shall then consider the error analyses aspects of the new methodology and show how the present definitions allow a greater degree of self-criticality of the procedures involved. We shall finally review the work done to develop specific elements based on this theory and their potential applications in general purpose software.

The field-consistency approach emphasizes the need to recognize that problems in structural mechanics (or more generally, in continuum mechanics), which need a description in terms of several strain-fields derived from one or more field variables, may require that one or more of these strain-fields must vanish in certain constraining limits. Conventional displacement method approaches based on independent low order functions that satisfied only the principles outlined earlier under the conventional wisdom could not ensure that these strain-fields would vanish in what, in our new terminology (see below), was called a 'consistent' way. These field-inconsistencies led to the enforcement of spurious constraints in the limit.

It remained now to devise an operational procedure that could perform an error analysis, keeping track of how field-inconsistencies emerged through the mathematics of discretisation to cause the degraded performance known as 'locking', or poor or delayed convergence even where locking was absent, and violent spurious stress oscillations. This operational procedure, called the functional re-constitution technique, was successfully applied to a wide range of problems (Prathap & Bhashyam 1982; Prathap & Viswanath 1982, 1983; Prathap 1984a–d, 1985a–e). The conceptual scheme offered new strategies to devise field-consistent elements free of these errors and the field re-distribution strategies (Ramesh Babu & Prathap 1985a,b; Prathap & Ramesh Babu 1985a,b) appear to be the most elegant method of designing elements free of locking and spurious stress oscillations. An elaboration of these ideas follows.

3. Definition of the field-consistency terminology

We begin by defining the terminology and vocabulary of the new paradigm in as consistent a way as the present state of the art makes it possible.

3.1 *Constrained multi-strain-field problems*

We first recognize a single classification scheme for those problems which show these common characteristics – locking, poor convergence, stress oscillations etc.

when a finite element discretisation is made. These are problems in structural mechanics which need a description of its continuum behaviour by more than one strain-field. In many practical situations, the engineering dimensions of the problem (e.g. a very thin Timoshenko beam, a very thin Mindlin plate) or the physical response due to the nature of the loading and the supports (e.g. inextensional bending of an arch) or the problem statement (e.g. an exterior penalty function approach) will require that one or more of these strain-fields must be constrained in particular ways. The discretised functionals for the strain energy of a finite element corresponding to the constrained strain-fields should be able to vanish in a way that ensures that only true constraints emerge in the limit. Where it fails to do so, we have the problems associated with field-inconsistency.

Table 1 gives a representative list of exercises which come under this class. In this review, we shall examine one of these exercises to show that the inability of the chosen field interpolations of a finite element to represent consistently the constrained strain-fields, can result in the very poor performance reflected as locking, poor convergence, stress oscillations etc.

3.2 *Field-consistency*

As we have stated earlier, one of the critical steps in the discretisation process is the selection of interpolation functions to describe the field variables over the element domain. These functions are usually chosen in the form of polynomials of a certain order. The number of constants used in each interpolation will depend on the number of nodes at which the field variable appears as a degree of freedom, and any other additional nodeless degrees of freedom corresponding to that field variable that may be introduced. From these definitions, one can compute the strain-fields also as interpolations associated with these constants by obtaining the correct derivatives of the field variables. In a multi-strain-field problem, these strain-fields will have as coefficients, terms from more than one field variable. Depending on the order of derivatives of each field variable appearing in the definition of that strain-field and on the order of the interpolation functions used for each contributing field variable, the coefficients of the strain-field interpolations may have constants from all contributing field variable interpolations or from only one or some of these. In some limiting cases of physical behaviour, it will become

Table 1. Some constrained multi-strain-field problems.

Name of exercise	Strain-fields	
	Unconstrained	Constrained
Plane stress, plane strain, 2D; 3D brick modes of flexure	Normal	Shear
Plane strain, 2D; 3D bricks modes of nearly incompressible elasticity	Distortional	Dilatational
Shear flexible beams (Timoshenko) and plates (Mindlin)	Bending	Transverse shear
Inextensional bending of curved beams; thin shells	Bending	Membrane
Shear flexible thick curved beams; shells	Bending	Membrane and shear

necessary for these strain-fields to be constrained to zero values. These can arise naturally, but indirectly, from classical variational statements (e.g. the minimum energy principle in displacement type elements) on which the finite element models are based, in certain geometrical situations, e.g., vanishing thickness of shear flexible beams, plates and shells, or they can arise directly in exterioₗ penalty type formulations, e.g., the discrete Kirchhoff hypothesis enforcement in thin plates and shells by a large penalty factor. In the former case, the constraints are embedded internally in the formulation and emerge as the geometrical or material limits are reached at which these constrained states are the natural condition for the problem. In the latter case, the constraints are introduced externally and reflect the limiting constrained state directly.

Where the discretised strain-field is such that all the terms in it (i.e. the constant, linear, quadratic etc.) have, associated with it, coefficients containing contributions from all the independent interpolations of the field variables that appear in the definition of that strain-field, the constraint that appears in the limit can be correctly enforced. Such a representation is said to be field consistent. The constraints thus enforced are called 'true constraints'. Where the discretised strain-field has coefficients in which the contributions from some of the field variables are absent, the constraints may incorrectly constrain the contributions from the field variables present. These are called 'spurious constraints'.

3.3 *Errors of the first and second kinds*

We now try to draw a distinction between the special errors that arise due to the 'spurious constraints' and the more familiar errors of discretisation. In a continuum problem with unconstrained strain-fields and only a single field variable, the errors of discretisation are of simple form and these usually vanish rapidly as the mesh size is reduced. However, in a multi-strain-field problem in which at limiting cases, constrained strain-fields must be enforced, the 'spurious constraints' can give rise to a form of errors that vanish very slowly with reduction in mesh size, and whose slowness of convergence and magnitude of error is greatly exaggerated by changes in the structural parameters that emphasise the appearance of the limiting physical situations.

By way of illustration of these concepts, we shall examine the results from the use of linear two-noded Timoshenko beam elements to predict the vertical deflection at the tip of a cantilever beam loaded by a tip shear force (Prathap & Bhashyam 1982). Figure 1 shows the rate of convergence obtained when the conventional element FI (for field-inconsistent) is based on an exact integration of all matrices. It is clear that convergence is extremely slow, making the use of such elements impractical. Obviously, there are errors of a form which are not those recognized as the usual errors of discretisation. To see this distinction clearly, we shall also show by the line marked by FC (for field-consistent), the convergence obtained by the use of elements which have been modified according to the field-consistency requirements. The convergence is rapid and is similar to that obtained in unconstrained problems.

Walz and coworkers (Walz *et al* 1970) recognized the existence of these two kinds of errors. This interpretation in terms of 'errors of the first kind' and 'errors of the second kind' was found useful when Prathap & Bhashyam (1982) separated
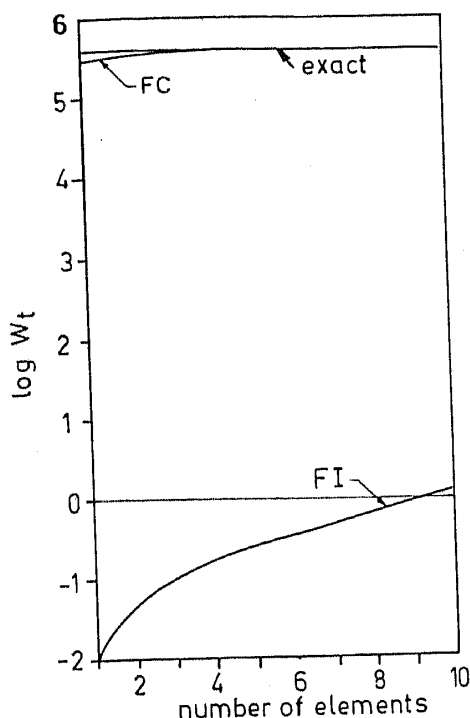
Figure 1. Convergence plot of tip deflection of a cantilever beam under tip shear load for $L/t = 10^4$.

the spurious constraints from the true constraints for an exactly integrated shear flexible element. It was seen that a field consistent element obtained by reduced integration, which had only the true Kirchhoff constraints, had errors which vanished rapidly as the mesh size was reduced. These errors were independent of the related structural parameter – the thinness of the beam, (l/t) – and these were identified as errors of the first kind. On the other hand, the field-inconsistent element which would result from an exact integration and which will have the additional spurious constraints, will have errors due to these constraints which are exaggerated in a $(l/t)^2$ fashion (Prathap 1985e). In the very thin limits, these errors are so large as to make the results virtually meaningless. This phenomenon is known as 'locking', and these locking errors are called 'errors of the second kind'.

3.4 *Functional reconstitution after discretisation*

We examine now how we can provide a procedure that will give this class of problems in the finite element method, a 'falsifiability' capability – essentially an error analysis procedure interpreted after Popper's falsifiability criterion (Popper 1972). Starting from the basic premises of the paradigm, its definitions and its operational procedures, the finite element model will produce a numerical result to a physical problem. This answer may be in error because the physical model is wrongly stated, or because the mathematical modelling due to the finite element discretisation process was faulty. Here, we are primarily concerned with the laws of numerical analysis alone – we must be able to satisfy ourselves that the discretisation process has introduced a certain predictable degree of error and that it converges at a predictable rate, i.e., it is removed in a predictable manner as the discretisation is improved in terms of the mesh refinement.

To make this sort of error analysis possible, we try to devise auxiliary procedures that can anticipate the type and degree of error that will result from the

discretisation process alone and see that this is really the case when an actual digital computation with that element is made.

The field-consistency paradigm allows the incorporation of such a procedure to make accurate *a priori* estimates of errors of the second kind in constrained multi-strain-field problems which are 'inconsistently' formulated. We call this procedure, the functional re-constitution technique (Prathap & Bhashyam 1982; Prathap 1985a,b).

The method of functional reconstitution after discretisation (Prathap & Bhashyam 1982) showed how shear locking emerges in a simple linear shear flexible beam element in which the normal displacement w and face rotation θ are independently interpolated. We shall elaborate on this when we consider the shear flexible beam element in greater detail later. Here, we shall briefly describe this technique.

The problem of locking arises because of the need to work with a finite size element. Thus, in a finite element idealisation, the mathematical operations of defining a small but non-infinitesimal element, prescribing interpolations for the field variables over this domain to a certain order of polynomial, and of subsequently integrating the functional for the strain energy of the element will result in a discretised estimate for the strain energy of that region in terms of the values of the field variables at the nodes and element sizes. In many practical situations, it is required that in limiting cases, some of the strain-fields must vanish and the strain energies associated with it must also vanish. This will be possible if the structural parameters multiplying the strain energy terms take large values in a penalty-limiting sense and the energy terms associated with it reflect true constraints. However, because the original interpolations are not 'field consistent', we have terms in which a structural parameter involving an element dimension multiplies a 'spurious constraint'. In an ideal infinitesimal case of vanishing size of the element, this structural parameter vanishes, leaving the 'spurious constraint' unenforced. However, in a practical analysis, we need to work with elements of reasonably finite size. In this case, the structural parameter enforces the 'spurious constraints' and adds an additional 'spurious energy' and therefore an 'additional stiffening' effect. These constraints are physically equivalent to an altered system in the non-penalty regime, which we can obtain by reconstituting a functional for the effective strain energy of an element based on a finite size discretisation. This is done, by carrying out the discretisation operations for a finite size element, re-grouping the true energy terms, the energy terms that should vanish due to the true constraints and the energy terms that should vanish but do not because of its appearance in the form of the spurious constraints. The strain energy density of such a discretised element is obtained by dividing this strain energy by the element volume. This strain energy density represents a physical system that contains the 'locking' effect due to field inconsistency. This procedure of obtaining a strain energy functional for the field inconsistent system will be called functional reconstitution.

3.5 *The additional stiffening parameter*

Conventional error analysis norms in the finite element method are based on the percentage error or equivalent in some computed value as compared to the

theoretically predicted value. In multi-strain-field problems which are inconsistently modelled, errors of the second kind can be exaggerated without limit as the structural parameter that acts as a penalty multiplier becomes indefinitely large. The percentage error norms therefore saturate quickly to a value approaching 100% and do not sensibly reflect the relationship between error and structural parameter even on a logarithmic plot.

Prathap (1985e) introduced a new error norm called the additional stiffening parameter, e. This helps to recognise the manner in which the errors of the second kind can be blown out of proportion by a large variation in the structural parameter. Essentially, this takes into account the fact that the spurious constraints give rise to a spurious energy term, and consequently alter the rigidity of the system being modelled. In many examples, it was seen that the rigidity, I, of the field consistent system and the rigidity, I', of the inconsistent system, were related to the structural parameters in the form,

$$I'/I - 1 = \alpha(l/t)^2 ,$$

where l is an element dimension and t is the element thickness. Thus, if w is the deflection of a reference point as predicted by an analytical solution to the theoretical description of the problem and $w(\text{fem})$ is the deflection predicted by a field inconsistent finite element model, we would expect

$$e = w/w(\text{fem}) - 1 = \alpha(l/t)^2 \tag{1}$$

A logarithmic plot of the new error norm against the parameter (l/t) will show a quadratic relationship that will continue indefinitely as (l/t) is increased. This was found to be true of the many multi-strain-field problems reviewed by Prathap (1985e) and this will be further described in the subsequent sections.

By way of illustration of the distinction made by this definition, we shall anticipate again the results from § 4. If we represent the conventional error norm in the form

$$E = [w - w(\text{fem})]/w, \tag{2}$$

and plot both E and the new error norm e from the results for the same problem illustrated in figure 1 using 4 FI elements against the penalty multiplier $(l/t)^2$ on a logarithmic scale, the dependence is as shown in figure 3. It can be seen that E saturates quickly to a value approaching 100% and cannot meaningfully show how the error propogates as the penalty multiplier increases indefinitely. On the other hand, e captures this relationship, if it exists, very accurately.

3.6 *Spurious stress oscillations*

The fact that finite element models can give rise to violent stress oscillations have been known for some time (Barlow 1976). It is also known that in such problems, there may be optimal locations where the stresses can be correctly sampled in spite of the violent stress oscillations. We now know that these oscillations arise uniquely in constrained multi-strain-field problems (Prathap & Ramesh Babu 1985c) and the stresses that oscillate thus are those that correspond to 'inconsistently' represented constrained strain-fields. These oscillations represent spurious self-equilibrating stress-fields generating the spurious energy terms leading to 'locking' or poor

convergence. The stress oscillations can therefore be thought of as the mechanism by which the spurious stiffening energy of locking is distributed as a self-equilibrating stress variation in the element. It is possible using the functional re-constitution procedure to devise estimates for these oscillations starting from the field-consistency interpolation (Prathap & Ramesh Babu 1985c). Since the inconsistent part of the constrained strain-fields are self-equilibrating, zero-points emerge from such terms and consequently sampling at these points give correct stresses. It was possible in all such cases (Prathap & Ramesh Babu 1985c) to determine these sampling points from a consideration of self-equilibrated re-distribution of the inconsistent part of the constrained strain-fields alone, without reference to orders of numerical integration used as was done earlier (Hinton & Campbell 1974). An example of this will be seen in the next section.

4. An illustrative example. The poor bending response of a shear flexible beam element

This is perhaps the simplest and also the earliest example in which the problems of field-inconsistency were identified and modelled mathematically through definitions and operations of the new conceptual scheme. We shall now use it as a very brief illustration of the concepts introduced in § 3 of this review.

4.1 Historical introduction

A brief historical introduction will be in order. The classical two-noded beam element with two degrees of freedom at each node, the transverse deflection w and slope dw/dx based on elementary beam theory is very well-known. This is the simplest element which can be constructed in terms of the C^1 continuity requirement of this theory. It was soon very tempting to formulate a similar two-noded beam element based on the shear flexible Timoshenko beam theory – the attractive aspect about this being that the two degrees of freedom at each node follow from a transverse displacement w and a section rotation θ which are independent field variables. Such an element will need only C^0 continuity and can be based on simple linear interpolations.

It is apparent that early experiments with such an element proved to be disastrous. With the advantage of hindsight, we can now understand why no record of such an element appeared until 1977 (Hughes et al 1977), when a deceptively elementary 'trick', that of using a reduced integration of the shear strain energy, produced a remarkably accurate element. The initial response was to dismiss this as a 'useful trick'. Implicit in this was the assumption that reduced integration introduced errors that compensated somehow for the other constraining errors. These early interpretations were based on an understanding in terms of the rank and singularity of the matrices corresponding to the penalty-linked, i.e., the shear energy terms. These interpretations argued that the exact integration of the shear energy terms introduced too many constraints among the limited number of variables (degrees of freedom) available per element and that this was reflected in the very degraded 'locking' behaviour of the element.

Figure 2 shows a typical illustration of what we mean by the very poor behaviour of the original unmodified element (FI – for field-inconsistent) as compared to the

dramatic improvement in efficiency obtained by making it field-consistent (FC). A map of percentage error E can be made for N-element models of a cantilever beam under tip shear load for various ratios of length of beam L to beam depth t. Figure 2 shows the relationship on logarithmic scales. It is evident that over the practical range in which the Timoshenko beam theory is appropriate (i.e. say $L/t = 5$ to $L/t = 1000$), the FI models are virtually impractical to use – needing as many as $100(L/t)$ elements to achieve the same accuracy as that obtained with 10 FC elements!

A reworking of the same shear flexible linear beam element (Prathap & Bhashyam 1982) afforded a better understanding of the principles involved. Although the improvement was brought about by reduced integration, this was rationalised in terms of the types of constraints seen at the penalty limits. The foundation of the functional re-constitution procedure was laid in the paper and it was seen to give very accurate predictions of how an exactly integrated element would behave in static and dynamic problems. The elaboration in terms of field-consistency interpretations came later and in this section, we reinterpret all this from this view point.

4.2 *Description of element*

The strain energy of a Timoshenko beam of length $2l$, taking into account the effect of transverse shear deformation, can be written as the sum of its bending and shear energy terms,

$$U = U_B + U_S, \tag{3}$$

$$U_B = \frac{1}{2} \int_{-1}^{1} EI\theta,_x^2 \, dx, \tag{4a}$$

$$U_S = \frac{1}{2} \int_{-1}^{1} kGA(\theta - w,_x)^2 \, dx. \tag{4b}$$

In (4a) and (4b), w is the transverse displacement and θ, the section rotation. E and

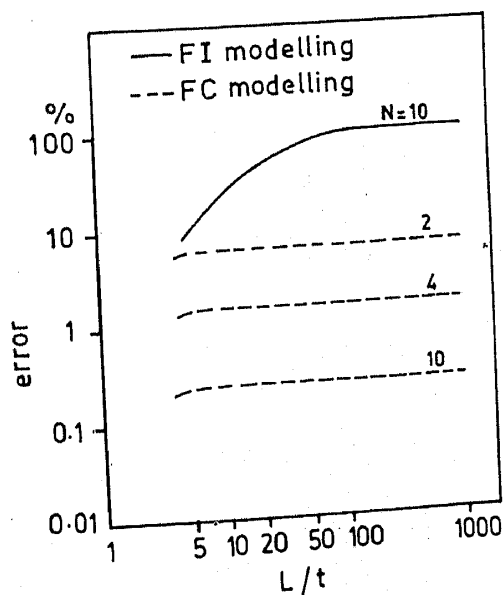Figure 2. Error map for N-element model of a Timoshenko cantilever beam with tip shear load.

G are the Young's and Shear moduli, respectively, and k, the shear correction factor. I and A are the moment of inertia and the area of cross-section. The strain energies correspond to flexural and shear strains of the form

$$\chi = \theta_{,x},$$ (5a)

$$\gamma = \theta - w_{,x}.$$ (5b)

4.3 *The field-inconsistent element*

The conventional procedure is to start with a linear isoparametric representation of the two field variables w and θ. These functions are

$$N_1 = (1-\xi)/2,$$

$$N_2 = (1+\xi)/2,$$ (6)

where the dimensionless coordinate $\xi = x/l$ varies from -1 to $+1$ for an element of length $2l$. The strain energies in (4a) and (4b) are then directly computed, in an analytically or numerically exact way, using these interpolation functions in the expressions for the strain-fields.

We can now examine, from the field-consistency point of view, the implications of this approach. If we start from the linear interpolations seen in (6), we can associate two constants with each of the field variable interpolations in the following manner,

$$w = a_0 + a_1 (x/l),$$ (7a)

$$\theta = b_0 + b_1 (x/l).$$ (7b)

These lead to the following interpolations for the strain-fields

$$\chi = (b_1/l),$$ (8a)

$$\gamma = (b_0 - a_1/l) + b_1 (x/l).$$ (8b)

An exact evaluation of the strain energies of the element will now give

$$U_B = \tfrac{1}{2} (EI)(2l)\{(b_1/l)\}^2,$$ (9a)

$$U_S = \tfrac{1}{2} (kGA)(2l)\{(b_0 - a_1/l)^2 + 1/3\ b_1^2\}.$$ (9b)

It is possible to see from this that in the constraining physical limit of a very thin beam modelled by element of length $2l$ and thickness t, the shear strain energy in (9b) must vanish. An examination of the conditions produced by this requirement show that the following constraints would emerge in such a limit

$$b_0 - a_1/l \to 0,$$ (10a)

$$b_1 \to 0.$$ (10b)

In our terminology, constraint (10a) is field-consistent as it contains constants from both of the two contributing field-variables relevant to the shear strain-field. These constraints can then accommodate the true Kirchhoff constraint in a physically meaningful way. In direct contrast, constraint (10b) contains only a term

from θ. A constraint imposed on this will lead to an undesired restriction of θ – this is the 'spurious constraint' that leads to 'shear locking' and violent disturbances in the shear force prediction over the element, as we shall see presently.

4.4 *The field-consistent element*

A field-consistent element can be formulated free of this deficiency in many ways. The 'trick' is to evaluate the shear strain energy so that only the consistent term will contribute to the shear strain energy. This can be done by simply dropping the undesirable term from the shear strain interpolation in (9b). This same result can be achieved by a reduced integration rule – i.e. a Gaussian one-point integration of shear energy. Other techniques like addition of bubble modes, hybrid and mixed methods can produce the same result if the manipulations are judiciously done to ensure field consistency.

A field-consistent redistribution strategy is offered here as the most elegant procedure. In this, substitute-shape functions,

$$\overline{N}_1 = \overline{N}_2 = 1/2,$$

are used for the θ field variable in the shear-strain definition alone [equation (5b)]. This substitute-shape function can be thought of as a least squares smoothed equivalent of the original linear interpolation so that the shear-strain field is now 'consistently' interpolated in terms of the original nodal degrees of freedom. This technique is versatile in that it restores field-consistency without disturbing the number of the nodal degrees of freedom, and in freeing the order of integration from the field-consistency requirement. The shear-forces are computed directly from these smoothed functions and are free of the violent fluctuations seen in an inconsistent formulation.

4.5 *Functional reconstitution*

It is simple to derive the shear locking present through the functional re-constitution procedure. The strain energy of a uniform beam of length L and thickness t can be written as

$$\pi = \frac{1}{2} \int_0^L EI\theta,_x^2 \, dx + \frac{1}{2} \int_0^L kGA (\theta - w,_x)^2 \, dx. \tag{11}$$

If an element of length $2l$ is isolated, the discretisation process produces energy for the element of the form [see (9)],

$$\pi_e = \frac{1}{2} (EI)(2l)(\theta,_x)^2 + \frac{1}{2}(kGA)(2l)(\theta - w,_x)^2$$
$$+ \frac{1}{2} (kGAl^2)(2l)(\theta,_x^2/3), \tag{12}$$

where the constants in (9) which appear in the discretisation process have been re-constituted into their original continuum form as contributing field variables. From this re-constituted functional, we can consider that an idealisation of a beam region of length $2l$ into a linear finite element would produce a modified strain energy density within that region of

$$\pi'_e = \tfrac{1}{2}(EI + kGAl^2/3)\,\theta,_x^2 + \tfrac{1}{2}kGA\,(\theta - w,_x)^2. \tag{13}$$

This strain energy density reflects the altered physical system that contains the shear locking effect. The re-constituted functional for the whole beam is then,

$$\pi' = \frac{1}{2}\int_0^L EI(1 + kGAl^2/3EI)\theta,_x^2\,dx + \frac{1}{2}\int_0^L kGA\,(\theta - w,_x)^2\,dx. \tag{14}$$

We now have a representation that is stiffer than the undiscretised beam by the facto. $kGAl^2/3EI$. For a very thin beam, this can be very large, and produces the additional stiffening effect known as 'shear locking'.

4.6 *Additional stiffening parameter*

An exactly-integrated or field-inconsistent finite element model tends to behave as a shear-flexible beam with stiffened rigidity I'. This can be related to the rigidity I of the original system as

$$I'/I = 1 + kGAl^2/3EI, \tag{15}$$

It is possible to show that this *a priori* estimate for the error due to field-inconsistency anticipates very accurately the behaviour of a field-inconsistent linearly interpolated shear flexible beam element in an actual digital computation. Exact solutions are available for the static deflection w of a Timoshenko cantilever beam of length L and thickness t under a vertical tip load. If $w(\text{fem})$ is the result from a finite-element digital computation using elements of length $2l$, and thickness t, the additional stiffening parameter for such a model is

$$e(\text{fem}) = w/w(\text{fem}) - 1. \tag{16}$$

From (15), we can postulate the error model prediction to be

$$e = I'/I - 1 = kGAl^2/3EI. \tag{17}$$

Figure 3 shows the variation of e with the structural parameter that denotes the penalty multiplier in this case, namely kGL^2/Et^2, for the case presented in figure 1. The crosses indicate the additional stiffening parameter computed for the finite
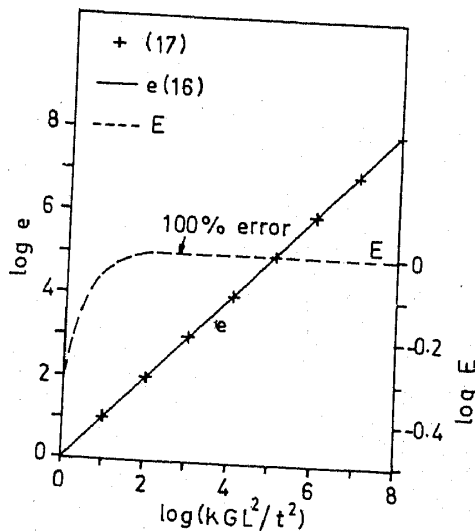

Figure 3. Error norms, e, E as functions of penalty multiplier for a cantilever beam under tip force.

element computation [(16)] and the solid line shows the variation as predicted by the error model [(17)].

In the example, therefore, we have been able to demonstrate that the field consistency approach has been given a powerful 'falsifiability' capability by the functional re-constitution procedure. We are able to predict that an element, designed in a certain way and containing certain measurable deficiencies, should behave in a precisely quantifiable erratic way. Prathap & Bhashyam (1982) also give an example where it is shown that the error prediction in (14) is equally accurate for a dynamic case – the free vibration of a simply-supported Timoshenko beam. Figure 4 shows how an estimate of the additional stiffening parameter based on the square of the frequencies of the first mode, i.e.

$$e = \omega^2(\text{fem})/\omega^2 - 1,$$

varies with the additional stiffening parameter for a simply-supported beam.

4.7 *Shear force oscillations*

The inconsistent constraint will further appear as a violent spurious shear force oscillation. If $\overline{V}$ is the shear force predicted by a field-consistent shear strain field and if V is the shear force obtained from the original shear strain field, we can write from (8b),

$$\overline{V} = kGA\,(b_0 - a_1/l), \tag{18a}$$

$$V = \overline{V} + kGA\,(b_1)\,(x/l). \tag{18b}$$

We see that V has a linear term that relates directly to the constant that appeared in the spurious constraint (10b). Thus, in a field-inconsistent formulation, this constant will activate a violent linear shear force variation when the shear forces are evaluated directly from the shear strain field given in (8b). Using an extension of the functional re-constitution procedure, it is possible to make *a priori* estimates for the magnitude of these oscillations. This establishes that field-inconsistency is the
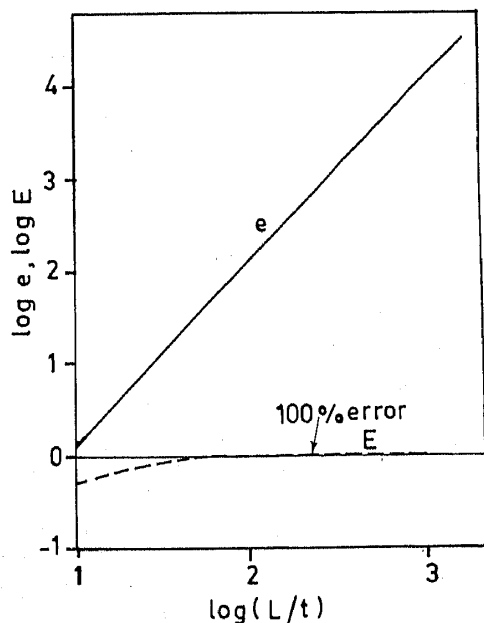

Figure 4. E norms e, E as functions of slenderness ratio for the first frequency of a cantilever beam.
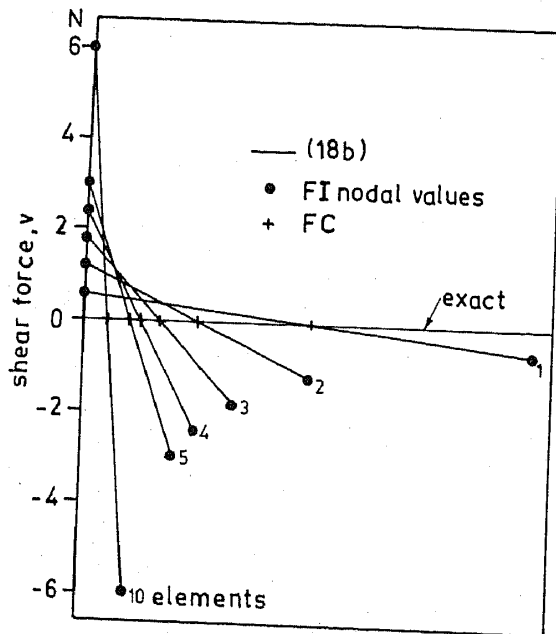
Figure 5. Shear force oscillations for a cantilever beam with tip moment.

source of these oscillations. Figure 5 shows the shear force oscillations in a typical problem – a straight cantilever beam with a concentrated moment at the tip. One to ten equal length field-inconsistent elements were used and shear forces were computed at the nodes of each element. In each case, only the variation of the element at the fixed end is shown, as the pattern repeats itself over all other elements. At element mid-nodes, the correct shear force, i.e. $\overline{V} = 0$, is correctly reproduced. Over the length of the element, the oscillations are seen to be linear functions corresponding to the $kGA\,(b_1)\,(x/l)$ term. Also indicated by the solid lines is the prediction made by the functional re-constitution exercise (Ramesh Babu & Prathap 1985b).

5. Field-consistency and shape, size and orientation of elements

The preceding analysis has focussed attention on one aspect of the discretisation strategy, namely, the selection of field-interpolations over an element domain and the subsequent derivation of strain-fields and strain energies from it, and how this can have disastrous effects if consistency of constrained strain-fields is not assured in problems that require it. Equally critical to the task is the division of the continuum region into element domains – it is now known that shape, size and orientation of the element grids are very important factors in assuring that the true constraints are met.

One of the earliest features noticed about the locking phenomenon was that the errors propagated indefinitely as the element aspect ratios increased. We have already seen this aspect earlier, e.g., the dependence of the additional stiffening factor on $(l/t)^2$ for the beam element. In fact, a simple test to see if locking has been removed is to monitor the additional stiffening parameter as the geometrical ratios to which it is sensitive are increased indefinitely.

It was also well-known that the errors of locking become larger if elements are distorted – in terms of location of nodes, curved edges, deviation from rectangular forms etc. We shall consider this now.

In one-dimensional problems, as in shear flexible beams and in curved beams, the choice is unambiguous as long as the description is linear – i.e. two nodes at the ends of the element. However, where more than two nodes are to be distributed over an element and isoparametric transformations are used, the field-consistency requirements can be accurately met only if the disposition of the nodes ensures that the transformation from the local-coordinate axes to the global coordinate system is linear so that no new spurious constraints are brought in.

In two-dimensional problems, e.g. plane stress modelling of flexure, shear flexible plates and shells etc., it is now known that division into low order triangular elements often leads to locking, even with reduced integration. It is also known that quadrilateral elements can be used safely only if they are rectangular and well-aligned with the global axes directions. If these requirements are not met, it may not be possible to devise simple methods of field-redistribution of the strain fields that will ensure that the constrained strain-fields yield only and all the true constraints.

A good example to demonstrate this is the finite element modelling of a shear flexible theory such as the Mindlin plate theory (Mindlin 1951). In the thin plate limit, the shear strains must vanish to produce the Kirchhoff constraints i.e. $\gamma_{xz} = \theta_x - w,_x \to 0$ and $\gamma_{yz} = \theta_y - w,_y \to 0$. The field-redistribution strategies can work efficiently only if the isoparametric transformations ensure that these constraints remain true even after the transformations. In fact, the emerging wisdom on the use of low-order C° continuous Mindlin plate elements is as follows–

(i) Linear triangles behave very poorly unless they are right-angled, have the sides containing the right-angle parallel to global x and y axes, and have separate integration points for the shear strain energies associated with γ_{xz} and γ_{yz}, respectively (Prathap 1985c).

(ii) Bi-linear quadrilaterals can be used most efficiently only in rectangular form with edges parallel to global x, y axes, and with consistently re-derived shear-strain fields based on the original nodal degrees of freedom (Prathap 1984e).

These restrictions are not unexpected as the formulation of the shear strain energy in the Mindlin plate theory, and the subsequent identification of the Kirchhoff constraints in the penalty limit of the thin plate theory, depends intrinsically on the definition of an x and y orthogonal cartesian system. Therefore, the retention and enforcement of all the true constraints is possible only in a rectangular form of the element. Many elements free of this restriction exist (Hughes & Tezduyar 1981; Crisfield 1984), but are made possible by manipulating the shear energies so that true constraints which are non-invariant are omitted and any zero-energy modes introduced thereby may have to be suppressed in some way, e.g., by addition of a stabilisation matrix (Belytschko *et al* 1981).

The behaviour of 3-D elements is very similar. Hexahedral elements behave much more accurately than tetrahedral elements, especially in problems where high-aspect ratio elements are used to model regions of flexure or nearly-incompressible elasticity. Again, expectedly, the accuracy is best if rectangular elements are used.

6. Some constrained multi-strain-field problems

In table 1, we have seen a list of problems in structural mechanics which need a description by multiple strain-fields and in which one or more of these strain-fields

are constrained in certain physical regimes. We shall now examine, briefly, the issues involved in trying to design for field-consistency, and also survey some of the elements that have been designed taking this into consideration. It is seen that design based on field-consistency interpretations provide better and more reliable performance.

6.1 *Plane-stress flexure – 4-noded and 8-noded elements*

Although many plane stress finite elements can exactly represent rigid-body and constant-strain modes, they are too stiff in their response to the simple flexural action of a beam. The problem is due to the inability of the plane stress elements to yield shear-strains that are consistently represented. In the case of the 4-noded element (Prathap 1985b), the errors result in 'locking' and these errors progress indefinitely as the element aspect ratio increases. The 8-noded element does not lock, but performs with a reduced efficiency. Both elements show violent shear-stress oscillations in modelling flexure – linear in the 4-noded case and quadratic in the 8-noded case. All these difficulties are overcome when field-consistent shear-strain definitions are used to design the improved 4-noded and 8-noded plane stress elements (Prathap *et al* 1985).

6.2 *Two-noded and three-noded curved beam elements*

Early attempts to derive curved beam and shell elements in a curvilinear system were dramatically unsuccessful. This was often wrongly attributed to the failure of these elements to recover strain-free rigid-body displacement modes in a curvilinear coordinate description. Recent evidence (Prathap 1985a) has pointed to a 'membrane locking' phenomena that arises when the constrained strain-fields corresponding to inextensional bending are not 'consistently' recovered.

The field-consistency concepts have allowed two-noded and three-noded thick curved beam elements based on curvilinear deep shell formulations to be derived free of shear and membrane locking (Ramesh Babu & Prathap 1985b; Prathap & Ramesh Babu 1985a). These will be the most efficient elements of their class.

6.3 *Two-noded and three-noded axisymmetric shell elements*

Field-consistent representations have allowed very efficient two-noded linear and three-noded quadratic curved axisymmetric shell elements to be designed. These are free of shear and membrane locking (Ramesh Babu & Prathap 1985a; Prathap & Ramesh Babu 1985b). These elements have been tested over a wide range of problems, some of which have special consistency requirements that can be met by the new theory.

6.4 *8-Node and 20-node brick elements*

Brick elements in flexure suffer from locking due to parasitic shear in the same way that plane stress elements do. Brick elements can be made field-consistent in flexure by using a field re-distribution strategy to define consistent shear strain-fields. An 8-node element has been successfully tested (Prathap 1984a). Brick elements also perform very poorly in near-incompressible elastic limits. This can be corrected if the volumetric strains are designed to be field-consistent and the

modifications for this were successfully made for a 8-node brick (Prathap 1984b). Work is now underway to design a field-consistent 20-noded brick element.

6.5 *A C° continuous 4-node plate/shell element*

A C° continuous doubly curved shell element was considered very tricky to achieve because of the combined problems of shear and membrane locking. We have been able to design a simple 4-noded element using field-redistribution to eliminate membrane and shear locking. The element has been successfully tested in a laminated anisotropic doubly curved shell version.

7. Field-consistency and 'variational tricks and crimes'

In the early years of the formal rationalisation of the finite element discretisation process in terms of variational calculus, it was believed that the functionals describing the total potential or complementary energy, or weighted residual equations, or virtual work equations, must be integrated exactly over the element domain for the process to be mathematically complete and correct. The subsequent experience with constrained multi-strain-field problems was to indicate that such an approach may be too unrealistic. Thus, the exact integration of functionals derived from independent low order polynomials will tend to impose a set of strain-fields which are uncharacteristic of the actual physical conditions in the constraining limit at practical levels of idealisation.

The first response to such an impasse was to experiment with short-cuts such as reduced integration, selective integration, addition of bubble-modes etc. The types of techniques varied greatly and their number has been growing rapidly. Very often, the successful use of the technique preceded the actual understanding of the mechanics of its operation, and not surprisingly, these techniques could and did fail when not properly used. In this section, it will be instructive to review briefly, some of these 'variational crimes' from the point of view of field-consistency. The list given below is by no means exhaustive, as many ingenious variations continue to be invented, but it is believed that it is a good representative sample of some of the procedures available.

7.1 *Unequal order interpolation*

This is perhaps the simplest to understand, as no 'trick' such as reduced integration is involved. This proceeds from an understanding of the fact that field-inconsistency arises when equal order interpolations are used for field variables which appear in different orders of its derivatives in the strain-field that has to be constrained. Thus, if one ensures that the strain-field is consistently represented by a proper *a priori* choice of unequal order interpolations for the contributing field variables, there would be no 'locking'.

A simple illustration of this is to consider a mixed linear/quadratic interpolation version of the beam element considered in §4. We can associate the following constants with each of the field variable interpolations thus:

$$w = a_0 + a_1(x/l) + a_2(x/l)^2, \tag{19a}$$

$$\theta = b_0 + b_1(x/l). \tag{19b}$$

The shear strain-field now becomes

$$\gamma = (b_0 - a_1/l) + (b_1 - 2a_2/l)(x/l).$$

(20)

It is clear that in the penalty limit, the two constraints that appear will enforce the Kirchhoff constraints in a 'consistent' manner.

7.2 Reduced integration/selective integration

This is perhaps the first of the 'variational crimes' to have been discovered. It proved to be very effective in many cases and was at first quite difficult to justify. Early attempts to explain its effectiveness were based on the number of constraints introduced at integration points and on its relation to the total number of degrees of freedom available in the model. Often, the orders of integration, although lower than that needed for an exact integration of the constrained strain energy functional, were still too high to remove locking, or so low that singularities or zero-energy mechanisms were introduced. This led to considerable experimentation and confusion.

With the field-consistency paradigm, we have clearer rules as to the optimal integration strategy for constrained multi-strain-field problems. What is important is that the strain energy functional comprising the constrained strain-fields must be 'consistently' represented. An optimal order of integration is one that will integrate all the contributions from the 'consistent' terms and remove all the contributions from the 'inconsistent' terms in the strain-field interpolation. If the order of integration is not high enough to cover all the 'consistent' terms, one or more of the true constraints will vanish, leading to a rank-deficiency and to zero-energy mechanisms that can degrade the performance of the element in certain applications. Again, an order of integration that is high enough to retain only one spurious constraint is enough to cause 'locking'.

It is also clear that the optimal rule for the constrained part of the strain energy functional need not be the correct rule for the unconstrained part of the energies. This suggests that a selective integration strategy is often required. Again, it is possible that, in some cases, an optimal integration strategy cannot be found for a particular problem and some other technique may be required to achieve field-consistency.

7.3 Reduced interpolation – smoothed function approach

We have considered an aspect of this and its equivalence to reduced integration when we examined the linear beam element earlier in this paper. In this technique, we re-derive interpolations for the field variables contributing to the inconsistent terms of the constrained strain-field such that the new interpolations are 'consistent' with the other field definitions. This can be obtained by a simple least-squares smoothing operation on the original shape function. The smoothed shape function is now a 'consistent' version built up from the same degrees of freedom as the original function – thus the number of degrees of freedom or disposition of the nodes in the element is not disturbed. For the unconstrained strain-fields, the original interpolations are retained.

In most cases, the use of smoothed interpolations is identical to the use of an optimal integration strategy. However, one can envisage cases where an optimal

reduced integration strategy may not exist, but consistent field interpolations can be derived using a least-squares technique to reduce the original field to the new field descriptions for these variables. Moreover, once a set of smoothed substitute functions are available, these can be easily incorporated in a shape function sub-routine that can be called up as desired from the field-consistency requirements of the problem. A uniform order of integration can now be used throughout – a desirable feature as it releases the restriction on the order of integration from the field-consistency requirement and allows it to be determined from other considerations, e.g., tapered or varying properties over an element.

7.4 Addition of incompatible modes or bubble functions

This is one of the earlier techniques used to restore field-consistency in a multi-strain-field problem. We can show, through the simple illustrative example of the linear beam, how this works.

We introduce what is called a bubble function (often, this is an incompatible mode in the sense that it does not satisfy the continuity requirement across some of the element boundaries) through a nodeless variable. Thus, for the linear beam this becomes:

$$w = a_0 + a_1(x/l) + A[1 - (x/l)^2], \tag{21a}$$

$$\theta = b_0 + b_1(x/l). \tag{21b}$$

The shear strain-field now becomes

$$\gamma = (b_0 - a_1/l) + (b_1 + 2A/l)(x/l). \tag{22}$$

Thus, in the penalty limit, the two constraints that appear will enforce the Kirchhoff constraints in a 'consistent' manner. The similarity to the method of using unequal order interpolations (§7.1) is obvious. However, A is no longer associated with a node and is best condensed out of the stiffness matrix. The element that results will be identical to each of the element versions considered in §§7.1 to 7.3 so far.

It is important to recognise that the addition of bubble functions must be done carefully, keeping in mind that it is the field-consistency requirement that is paramount. In fact, it is this over-riding considertion that dictates that in many problems, the functions so added must be chosen in a way that makes it non-conforming or incompatible.

7.5 Field-redistribution

This method is one of the most recently introduced (Oleson 1983). This recognises that the directly derived constrained strain-field which is inconsistent is also of a higher order than the consistent version of this strain-field. This higher order field is therefore redistributed in a special way that makes it very similar to the smoothed function approach (§7.3) or the hybrid/mixed method (§7.6). Again, the lower order field to which the redistribution strategy is directed must be based on field-consistency criteria for the method to be effective. -

7.6 *Hybrid and mixed methods*

Since the invention of the assumed stress strategy of deriving stiffness matrices (Pian 1964), hybrid or mixed elements, based on complementary energy, Hellinger-Reissner or Hu-Washizu theorems, have been very powerful in many problems where conventional displacement type elements had failed. With the advantage of hindsight, it is possible to reconcile this. The power of these methods resides in their ability to control the stiffness matrix by varying the number of terms in each stress-field. It is becoming obvious to us now that this can be construed as another method to achieve field-consistency by a smoothing type operation. If the assumed stress fields correspond to the inconsistent strain-fields derived from the displacement fields, then locking and stress oscillations will still remain in the hybrid/mixed element. However, if the assumed stress-fields are chosen carefully so as to cover only the consistent parts of the constrained strain-fields, the element so obtained will be identical to a field-consistent displacement model derived by any of the other techniques illustrated in the various sub-sections of §7.

Thus a hybrid/mixed element that starts with linear interpolations for displacements and rotations, and constants for the moment and shear force fields, will produce a beam element that is identical to one obtained by any other technique such as reduced integration, addition of bubble functions etc. If, a linear shear stress field had been used, the element would have locked. This has been observed when continuous force field mixed elements were derived (i.e. with force quantities also as nodal degrees of freedom) and the practice now is to use only discontinuous force field mixed elements. The field-consistency paradigm, therefore, offers a comfortable position from which the success of hybrid and mixed formulations can be easily evaluated.

Since an element with the same efficiency can be derived through the displacement type process with any of the techniques described here, it would seem that the additional mathematical operations and complexity that follow from a hybrid/mixed formulation do not justify its use for achieving field-consistency only, in such problems.

7.7 *Energy compensation*

From our field-consistency point of view, this falls into the realm of inexcusable 'variational crimes'. The idea of compensation can be traced to the fact that field-inconsistency leads to spurious energies emerging from the constrained strain-fields, which are then grossly exaggerated as the structural parameter that enforces the penalty limit becomes very large. This can be artificially removed by using the inverse of this parameter, together with an arbitrary constant C, to multiply the constrained part of the strain energy so that the spurious energy is pegged at an artificially low value. An early form of this energy compensation was the residual energy balancing technique (Fried 1974). Many variations of this exist, but in all, the prevailing idea is to compensate for spurious energies by manipulating the structural parameters that enforce the locking.

A very interesting version of this is the artificial softening technique (Cook 1977) as applied to the poor plane stress modelling capability of the 4-node element. In this, a predetermined factor is used to multiply the stiffness matrix derived in a conventional way (and therefore, with field-inconsistencies present). Not surpri-

singly, in this example, this factor is exactly the same as the additional stiffening parameter for this problem (Prathap 1985b).

The constant C is problem-dependent and needs to be manipulated from problem to problem and even from mesh to mesh or support condition to support condition in order to achieve good results. Again, quite expectedly, stress predictions, especially those pertaining to the constrained strain-fields from such methods can be very unreliable. The point to note here is that the energy compensation procedure is an artificial one in that it deliberately introduces an error to compensate for another error, without seeking to remove the error at its very source (i.e. without trying to establish field-consistency at all).

7.8 *Reduced interaction*

This is another interesting variation that has appeared recently (Mohr 1980). This method seeks to trace the inconsistency in a constrained strain-field to the field-variable that leads to the inconsistent terms and replaces the original interpolation for it (i.e. the one used for that field-variable in the unconstrained strain-fields) with a lower order interpolation so that consistency is achieved. This is very similar to the smoothed function approach, except that it differs only in the strategy for choosing the lower order interpolation. In the reduced interaction approach, this is often done by simple inspection, whereas in the smoothed function approach, it is done by ensuring a least squares smoothed relationship between the inconsistent and consistent strain-fields. The method works well, especially when the choice of the lower order interpolation is the same or closely resembles the smoothed function equivalent.

8. Concluding remarks

In this paper, we have surveyed an area that has generated considerable interest and controversy in recent years. We understand, in the terminology of the philosophy of science (Kuhn 1972), that this uncertainty resulted from the incompleteness of the current paradigms that informed the state of the art in the finite element method. It was also clear that whereas the conventional wisdom succeeded very well in most situations so as to sustain an applications industry of a billion dollars a year, there were still large areas where they failed, and a re-definition of paradigms was needed.

The first task was therefore to classify the problem area. The emerging wisdom on the subject at present seems to indicate that these are continuum field problems that need description by several strain-fields and in which one or more of these strain-fields are to be constrained. The next appropriate task is to give this class of problems a name for quick reference of the principles and issues involved – it is hoped that the requirement of 'field-consistency' will join the other more well-known principles used in constructing finite elements.

The last but equally important task for ensuring that the present body of knowledge acquires the semblance of a hard science is to provide it auxiliary terms of reference and procedures that can give it an error-analysis capability. The concept of 'errors of the second kind' to delineate a special form of discretisation errors and the 'additional stiffening parameter' were found necessary to make *a*

priori error models and *a posteriori* evaluations of elements used in this class of problems.

The extension of these concepts to other similar field problems will be of further interest and the invention of new techniques or 'tricks' that can restore field-consistency will be interesting academic exercises. Of more immediate practical relevance will be the re-design of the existing element libraries of major general purpose software packages so that many capabilities which did not exist can now be included. The library of field-consistent elements offered here (see §6) is a small beginning in this direction.

References

Argyris J H 1960 *Energy theorems and structural analysis* (London: Butterworths) (reprinted from *Aircr. Eng.* 1954–55)

Barlow J 1976 *Int. J. Numer. Methods Eng.* 10: 243–251

Belytschko T, Tsay C S, Liu W K 1981 *Comput. Struct.* 29: 313–327

Cook R D 1977 *Int. J. Numer. Methods Eng.* 11: 1334–1339

Crisfield M A 1984 *Comput. Struct.* 18: 833–852

Fried I 1974 *Comput. Struct.* 8: 771–778

Hinton E, Campbell J S 1974 *Int. J. Numer. Methods Eng.* 8: 461–480

Hughes T J R, Taylor R L, Kanoknukulchal W 1977 *Int. J. Numer. Methods Eng.* 11: 1529–1543

Hughes T J R, Tezduyar T E 1981 *J. Appl. Mech.* 48: 587–596

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

Medawar P B 1979 *Advice to a young scientist* (New York: Harper and Row)

Mindlin R D 1951 *J. Appl. Mech.* 18: 31–38

Mohr G A 1980 *Comput. Struct.* 11: 565–571

Oleson J F 1983 *Comput. Struct.* 17: 157–159

Pian T H H 1964 *AIAA J.* 2: 1333–1336

Popper K R 1972 *Conjectures and refutations – the growth of scientific knowledge* 4th edn (London & Henley: Routledge and Kegan Paul)

Prathap G 1984a A field consistent 8-node brick element, DFVLR IB 131-84/34, Institute of Structural Mechanics, Braunschweig, W Germany

Prathap G 1984b Field-consistent representation of incompressible strain states in a brick element, DFVLR IB 131-84/35, Institute of Structural Mechanics, Braunschweig, W Germany

Prathap G 1984c Field consistent finite element formulations, DFVLR IB 131-84/33, Institute of Structural Mechanics, Braunschweig, W Germany

Prathap G 1984d *Analysis of structures* (eds) K A V Pandalai and B R Somashekar (Bangalore: Natl. Aeronaut. Lab.)

Prathap G 1984e *Comput. Struct.* 18: 789–794

Prathap G 1985a *Int. J. Numer. Methods Eng.* 21: 389–407

Prathap G 1985b *Int. J. Numer. Methods Eng.* 21: 825–835

Prathap G 1985c *Int. J. Numer. Methods. Eng.* 21: 1149–1156