

## Examining the best-fit paradigm for FEM at element level

HIMANSHU MISHRA<sup>1</sup> and SOMENATH MUKHERJEE<sup>2</sup>

<sup>1</sup>Mechanical Engineering Department, Punjab Engineering College, Chandigarh, India

<sup>2</sup>Structures Division, National Aerospace Laboratories, Bangalore 560 017, India  
e-mail: himanshmishra@yahoo.com; somu@css.cmmacs.ernet.in

MS received 11 June 2004; revised 17 September 2004

**Abstract.** Interestingly, an esoteric branch of pure mathematics called “functional analysis”, more general and profound than variational calculus and originally developed by mathematicians, can be employed to explain clearly how the finite element machinery works. In a very abstract way, finite element results can be portrayed as “shadows” or orthogonal projections on predetermined function subspaces of the analytical results, known or unknown. Herein lies the philosophy of the finite element method. In the present work, an effort has been made to validate this important aspect of the finite element method at the element level through a particular differential equation representing a specific case of equilibrium.

**Keywords.** Strain projections; function subspaces; projection theorem; function space; orthogonal projection.

### 1. Introduction

Finite element analysis (FEA) has been a powerful tool for solving differential equations of mathematical physics that are not always amenable to analytical methods. In the conventional finite element method computation the algorithm below is followed. The domain of interest is first discretized into a number of elements, and then equilibrium equations of the elements, derived from approximate interpolation functions, are assembled in the form of matrices and solved. In essence, the method yields a matrix form for differential equations that represent domain characteristics. Solutions of these matrices with specified boundary conditions yield the desired results. Convergence is achieved by finer meshing. These solutions are easily recognized as being a best-fit approximation at the global level. In this paper, we examine critically what happens at the element level, which is of course the level at which the engineer is most interested in.

The function space interpretation of the finite element method has been presented by Strang & Fix (1973). Variational calculus based finite element analysis (Burnett 1987; Prathap 1993; Reddy 2003) is about half a century old. The fact that in FEA, the individual elements with approximate solutions respond in a best-fit manner to the analytical solution has been observed

---

A list of symbols is given at the end of the paper

(Prathap 1996, Prathap & Hattangadi 1998) through the orthogonality conditions that result from Hu–Washizu’s principle of approximate computation, that apply, in general, to the usual Rayleigh–Ritz method. The function space approach, with the best-fit rule has been directly used by Mukherjee & Prathap (2001,2002) for explaining shear locking in Timoshenko beam elements. It has been shown that violation of this best-fit rule may occur when variational incorrectness is introduced through reduced integration methods. Prathap & Mukherjee (2003) have reexamined critically Strang & Fix’s (1973) theorems in the light of function spaces and have identified the factors that determine the validity or violation of the best-fit paradigm at the element level. Cases where the best-fit rule is violated due to spurious forces at element connectivity (from nodal reaction errors) have been studied by Sangeeta *et al* (2003).

This paper deals with the subtle picture of the finite element method in terms of function spaces. The motivation behind this work is to expose what actually happens when the finite element computation is made, even for an atypical type of differential equation considered here. The paper begins with an introduction to the *projection theorem* for the finite element method followed by its validation using a problem called an *equilibrium equation* which is representative of many natural phenomena. The conditions of validity of the paradigm are confirmed by the exactness of the computed nodal reaction vectors for the elements.

## 2. Projection theorem for finite element analysis

### 2.1 At the global level

Consider the linear differential equation for a conservative system of the following form,

$$A(u) - f = 0, \quad (1)$$

where  $u$  is the analytical solution. At a global level, boundary terms in the weak form vanish due to either zero forces or vanishing of the test function from Dirichlet conditions. Therefore they do not appear in the following weak forms (Strang & Fix 1973) for the global system

$$a(u, u) = (f, u), \quad (2)$$

where the operator  $A$  is such that  $a(u, v) = a(v, u)$  is the bilinear symmetric functional and  $a(u, u)$  is positive definite, representing physically twice the strain energy of the system for the differential equation (1). Invoking the virtual work principle, one can use an approximate but admissible solution  $u^h$  and have

$$a(u, u^h) = (f, u^h), \quad (3)$$

If the exact solution  $u$  in (2) is replaced by the finite element test function  $u^h$  the global equilibrium equation becomes,

$$a(u^h, u^h) = (f, u^h), \quad (4)$$

(2) to (4) are valid at the global level with zero displacements or vanishing force conditions of the boundary of the global domain. Subtracting (3) from (4), we have

$$a(u^h, u^h) - a(u, u^h) = 0, \quad (5)$$

which implies,

$$a(u - u^h, u^h) = 0. \quad (6)$$

From the fact that the energy inner product is bilinear, we can argue that

$$\begin{aligned} a(u - u^h, u - u^h) &= a(u, u) + a(u^h, u^h) - 2a(u, u^h) \\ &= a(u, u) - a(u^h, u^h) - 2[a(u, u^h) - a(u^h, u^h)] \\ &= a(u, u) - a(u^h, u^h) - 2[a(u - u^h, u^h)]. \end{aligned}$$

Introducing the result from (6), we get an energy error theorem from the above expression as

$$a(u - u^h, u - u^h) = a(u, u) - a(u^h, u^h), \quad (7)$$

i.e.

$$\text{energy of the error} = \text{error of the energy}.$$

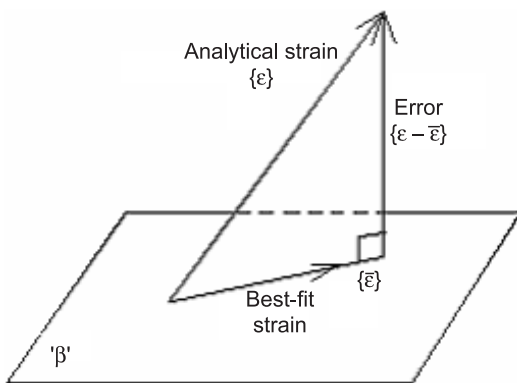
This leads to a useful statement that the left hand side of (7) is always positive definite. Hence,

$$a(u^h, u^h) < a(u, u). \quad (8)$$

Thus, in a variationally correct approach, the energy inner product of the approximate (Ritz or finite element) solution is always a lower bound of the exact energy (Strang & Fix 1973; Prathap & Mukherjee 2003). Equation (6) is the statement that shows that the approximate solution  $u^h$  is actually a best-fit to the analytical solution  $u$ . As will be shown in the following sections, the best-fit is actually with respect to the strain (or more generally, the strain vector) of the domain. This fact is shown geometrically in figure 1. The computed approximate strain vector  $\{\bar{\epsilon}\}$  is actually the orthogonal projection (best-fit) to the analytical strain vector  $\{\epsilon\}$  onto a predetermined function subspace.

## 2.2 At the element level

At the element level, however, the boundary terms do not necessarily vanish, since the vanishing Dirichlet conditions do not necessarily apply at element nodes, and there can be nodal



**Figure 1.** Geometric representation of the best-fit strain vector as an orthogonal projection of the analytical strain vector onto a function sub-space ' $\beta$ ' generated by strain-displacement relation.

reactions from adjacent elements/supports. Including the work done due to the displacements at the element boundary, one can express the weak forms at an element level as

$$a(u, u)^e = (f, u)^e + [uR^e]_b, \quad (9)$$

and

$$a(u, u^h)^e = (f, u^h)^e + [u^h R^e]_b, \quad (10)$$

where  $R^e$  represents the analytical nodal reaction forces acting on the element from adjacent element(s) or support(s). The terms  $[uR^e]_b$  and  $[u^h R^e]_b$  represent respectively, the work done by these nodal forces on the exact and the test functions  $u$  and  $u^h$  at the element boundary.

If the exact solution  $u$  in (10) is replaced by the finite element test function  $u^h$  in an element,  $e$ , of the discretized system, the equilibrium equation is given by,

$$a(u^h, u^h)^e = (f, u^h)^e + [u^h R^{h,e}]_b, \quad (11)$$

where  $R^{h,e}$  is the nodal force at the element level due to the approximation which is necessary for the equilibrium of an element having an approximation test function  $u^h$  as its displacement.

Subtracting (10) from (11)

$$a(u^h, u^h)^e - a(u, u^h)^e = [u^h (R^{h,e} - R^e)]_b. \quad (12)$$

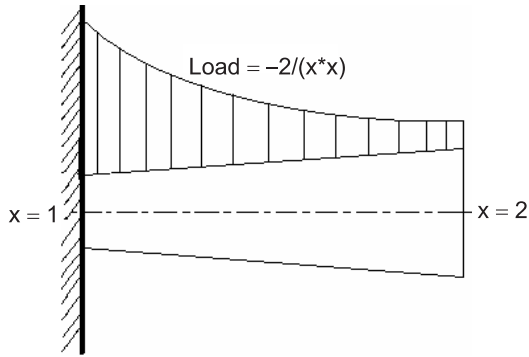
Thus, it is evident that only for the problems in which the discretization process in the finite element analysis conserves the nodal reactions (i.e.  $R^{h,e} = R^e$ ), so that the right hand side of (12) vanishes, the finite element method solution is the best-fit to the exact solution even at the element level (figure 1). Under such situations, (5) to (8), derived originally for the global domain, also apply at the element level. The best-fit rule, when satisfied, effectively guarantees that the strain energy of the approximate (Ritz or finite element) solution is always a lower bound of the exact strain energy, even at the element level.

Examples of some atypical problems that violate the best fit rule have been discussed by Prathap & Mukherjee (2003) and studied in full details by Sangeeta *et al* (2003). Even with variationally correct formulations, violation of the best-fit rule can be attributed to the disagreement between the analytical and finite element computed reactions ( $R^{h,e} \neq R^e$ ). The right hand side of (12) does not vanish, and hence the violation of the best-fit rule occurs. Physically, the deviation from the best-fit rule can be interpreted as the additional response of the element from spurious nodal forces ( $R^{h,e} - R^e$ ). A typical example where such a violation of best-fit rule occurs is the analysis of a bar, embedded in an elastically distributed spring foundation and axially loaded with distributed load intensity.

### 3. A specific boundary value problem

The one-dimensional boundary value problem taken here is an interesting problem discussed in the book '*Finite element analysis: From concepts to applications*' by David S Burnett. The governing equation is the following ordinary differential equation:

$$\frac{d}{dx} \left( x \frac{du}{dx} \right) = \frac{2}{x^2}, \quad (13a)$$



**Figure 2.** A structural analogy of the present problem. A bar of varying rigidity  $x$  with distributed axial loading  $q(x) = -2/x^2$ .

or expressed in the form of (1), it is

$$-\frac{d}{dx} \left( x \frac{du}{dx} \right) - \left( -\frac{2}{x^2} \right) = 0. \quad (13b)$$

The self-adjoint operator is given as

$$A = -\frac{d}{dx} \left( x \frac{d}{dx} \right)$$

and the distributed 'loading' function  $f$  is given by  $f = -2/x^2$ . The domain is the interval  $1 \leq x \leq 2$  along the  $x$ -axis and subjected to boundary conditions:

$$u(1) = 2. \quad (14)$$

At the boundary  $x = 2$ , flux condition is

$$R = \left( +x \frac{du}{dx} \right)_{x=2} = -\frac{1}{2}. \quad (15)$$

A closer examination of (13) reveals that it can be interpreted as that of an axially loaded bar of linear variation of section rigidity  $D = x$ , and loaded by a distributed load intensity  $f = -2/x^2$ . This is shown in figure 2.

#### 4. Analytical solution

The analytical solution of the differential equation (13) is given by

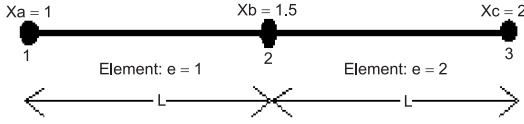
$$u = (2/x) + c_1 \ln(x) + c_2, \quad (16)$$

and the analytical strain is

$$du/dx = -(2/x^2) + (c_1/x).$$

From the boundary conditions (14) and (15) the particular solution is

$$u = (2/x) + (1/2) \ln(x). \quad (17)$$



**Figure 3.** Discretization details for the problem. The domain is divided into two elements, each of length  $L$ .

Thus the analytical strain is:

$$\varepsilon = (du/dx) = -(2/x^2) + (1/2x). \quad (18)$$

Analytical reactions from flux across the boundaries at  $x = x_1$  and  $x = x_2$  of a sub-domain, identified by the element 'e' can be easily obtained from the weak forms (9) and (10) for the present problem as

$$R_1^e = (-x(du/dx))_{x=x_1}, \quad R_2^e = (x(du/dx))_{x=x_2}. \quad (19)$$

By convention, a positive flux is directed along the left to the right direction (along positive  $x$ -axis). If the full domain is discretized into two elements, as shown in figure 3, the analytical nodal reactions for the elements for the problem can be thus be expressed as

$$R_1^{e=1} = -(-2/x_a) + (1/2), \quad (20a)$$

$$R_2^{e=1} = +((-2/x_b) + (1/2)), \quad (20b)$$

$$R_1^{e=2} = -((-2/x_b) + (1/2)), \quad (20c)$$

$$R_2^{e=2} = ((-2/x_c) + (1/2)) = -\frac{1}{2} = R. \quad (20d)$$

Note that  $R_2^{e=1} + R_1^{e=2} = 0$ , a condition for nodal junction of two elements.

## 5. The conventional finite element computation

### 5.1 Formulation

The domain has been discretized into two equal regions as shown in figure 3 and represented by

$$\text{element 1:}(x_a \leq x \leq x_b), \text{ element 2:}(x_b \leq x \leq x_c).$$

From the given boundary points,  $x_a = 1$  for the first node and length of each element is  $L = 0.5$ . Coordinates for the second and third nodes are respectively  $x_b = x_a + L$  and  $x_c = x_a + 2L$ . In each element the approximate solution is  $u^h$  given by

$$u^h = \sum_{i=1}^2 N_i \delta_i^e = [N_1 N_2] \{\delta^e\} = [N] \{\delta^e\}, \quad (21)$$

where  $\{\delta^e\}$  is the nodal displacement vector for the element  $e$ , given by  $\{\delta^e\} = [u_1 u_2]^T$  and  $N_1$  and  $N_2$  are the linear Lagrange interpolation shape functions. For first element the shape functions are given by  $N_1^{e=1} = (x_b - x)/L^e$  and  $N_2^{e=1} = (x - x_a)/L^e$  for first element and

$N_1^{e=2} = (x_c - x)/L^e$  and  $N_2^{e=2} = (x - x_b)/L^e$  for second and  $L^e$  is the length of the element  $e$ . Following the usual procedure, the element strain can be approximated by

$$\{\varepsilon^{h,e}\} = du^h/dx = [B]\{\delta^e\}, \quad (22)$$

where,

$$[B] = \left[ \frac{dN_1}{dx} \frac{dN_2}{dx} \right], \quad \text{or} \quad [B] = \left[ -\frac{1}{L^e} \frac{1}{L^e} \right]. \quad (23)$$

Since the displacement function is approximated as a linear one, the approximate strain should be a constant in each element, as is obvious by (22) and (23). Using the conventional Rayleigh–Ritz method, one can get the following form of the element equation of equilibrium in the finite element method,

$$[K^e]\{\delta^e\} = \{F^e\} + \{R^{h,e}\}, \quad (24)$$

The element stiffness and equivalent nodal load vectors (from the distributed load) are respectively given as

$$[K^e] = \int_e [B]^T [D] [B] dx \quad \{F^e\} = \int_e [N]^T q(x) dx. \quad (25)$$

For our problem in (23) we have the ‘elastic’ rigidity matrix  $[D] = x$  and the distributed loading  $q(x) = -(2/x^2)$ . After element assembly, the element nodal reactions at the node joining two elements cancel each other, and the global equation is obtained. With the incorporation of the boundary condition  $u_1 = 2$ , one can solve this equation to determine the nodal displacements. After solving for these nodal displacements (from global analysis), the element reaction vectors through the finite element method can be obtained from

$$\begin{aligned} \{R^{h,e}\} &= [K^e]\{\delta^e\} - \{F^e\} \\ &= [R_1^{h,e}, R_2^{h,e}]^T. \end{aligned} \quad (26)$$

It must be emphasized here that the analytical element nodal reactions  $\{R^e\}$ , as expressed in (19), are obtained through the analytical solution  $u$ . However, with any approximate solution  $u^h$ , these expressions cannot be extended for evaluating element reactions  $\{R^{h,e}\}$  by the finite element method. For computation of element nodal reactions  $\{R^{h,e}\}$  through a finite element formulation, (26) should be used.

It may also be noted that while the displacement gradient  $du/dx$  in the analytical solution is continuous across the node joining two elements, its counterpart ( $du^h/dx$ ) in the finite element approximation is not necessarily continuous across nodal connectivity. Yet, just like the analytical reaction components, the nodal reaction components obtained by finite element computation at the connecting node of two element  $e$  and  $e + 1$  automatically cancel each other during element assembly. This fact can easily be verified for the problem discussed in this paper.

For the purpose of our analysis, the domain has been discretized into two equal elements, each of length  $L$ , i.e.  $L^{e=1} = L^{e=2} = L$  (figure 3). With numerical values,  $x_a = 1$ ,  $L = 0.5$ ,  $x_b = x_a + L = 1.5$ ,  $x_c = x_a + 2L = 2$ .

For the first element then, the element stiffness matrix can be obtained as

$$\begin{aligned} [K^{e=1}] &= \int_{x_a}^{x_b} [B]^T x [B] dx = \frac{x_a + x_b}{2} \frac{1}{L^e} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \\ &= \frac{1}{L} \left( x_a + \frac{L}{2} \right) \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \end{aligned}$$

or

$$[K^{e=1}] = \begin{bmatrix} A_1 & A_2 \\ A_2 & A_1 \end{bmatrix}, \text{ where } A_1 = \frac{L + 2x_a}{2L} \text{ and } A_1 = -A_2. \quad (27a)$$

Similarly the stiffness matrix for the second element is given as

$$[K^{e=2}] = \int_{x_b}^{x_c} [B]^T x [B] dx = \begin{bmatrix} A_3 & A_4 \\ A_4 & A_3 \end{bmatrix}, \text{ where } A_3 = \frac{3L + 2x_a}{2L} = -A_4 \quad (27b)$$

The applied force vector for the first element from distributed loading is

$$\begin{aligned} \{F^{e=1}\} &= \int_{x_a}^{x_b} [N^{e=1}]^T \left( -\frac{2}{x^2} \right) dx = \frac{2}{L} \begin{Bmatrix} \ln \frac{x_b}{x_a} + \left( 1 - \frac{x_b}{x_a} \right) \\ -\ln \frac{x_b}{x_a} + \left( 1 - \frac{x_a}{x_b} \right) \end{Bmatrix} \\ &= \frac{2}{L} \begin{Bmatrix} \ln \left( \frac{x_b}{x_a} \right) - \frac{L}{x_a} \\ -\ln \left( \frac{x_b}{x_a} \right) + \frac{L}{x_b} \end{Bmatrix} \end{aligned} \quad (28)$$

and the applied force vector for the second element from distributed loading is

$$\begin{aligned} \{F^{e=2}\} &= \int_{x_b}^{x_c} [N^{e=2}]^T \left( -\frac{2}{x^2} \right) dx = \frac{2}{L} \begin{Bmatrix} \ln \frac{x_c}{x_b} + \left( 1 - \frac{x_c}{x_b} \right) \\ -\ln \frac{x_c}{x_b} + \left( 1 - \frac{x_b}{x_c} \right) \end{Bmatrix} \\ &= \frac{2}{L} \begin{Bmatrix} \ln \left( \frac{x_c}{x_b} \right) - \frac{L}{x_b} \\ -\ln \left( \frac{x_c}{x_b} \right) + \frac{L}{x_c} \end{Bmatrix}. \end{aligned} \quad (29)$$

## 5.2 Assembly of elements and solution

The global force-displacement equation after assembly can be expressed as

$$\begin{bmatrix} A_1 & A_2 & 0 \\ A_2 & A_1 + A_3 & A_4 \\ 0 & A_4 & A_3 \end{bmatrix} \begin{Bmatrix} u_1 \\ u_2 \\ u_3 \end{Bmatrix} = \begin{Bmatrix} F_1^{e=1} \\ F_2^{e=1} + F_1^{e=2} \\ F_2^{e=2} \end{Bmatrix} + \begin{Bmatrix} R_1^{h,e=1} \\ 0 \\ R = -0.5 \end{Bmatrix}, \quad (30)$$



This is of the form  $[K]\{\delta\} = \{F\} + \{R\}$ . Here  $[K]$ ,  $\{F\}$  and  $\{R\}$  represent respectively the global stiffness, applied nodal force vector (equivalent to distributed loading) and the nodal reaction vector. Note that the elements interact mutually at the second node, so that the net reaction force there vanishes as a connectivity condition, as is obvious from the vanishing of the second row of  $\{R\}$ . Thus  $R_2^{h,e=1} + R_1^{h,e=2} = 0$  is a condition for nodal junction 2 of the elements.

The right hand side of (30) represents the global force vector, which can be expanded to yield the following form

$$\begin{Bmatrix} F_1^{e=1} + R_1^{h,e=1} \\ F_2^{e=1} + F_1^{e=2} \\ F_2^{e=2} + (R = -0.5) \end{Bmatrix} = \frac{2}{L} \begin{Bmatrix} -L/x_a + \ln(x_b/x_a) + \frac{L}{2} (R_1^{h,e=1}) \\ \ln(x_a x_c/x_b^2) \\ L/x_c - \ln(x_c/x_b) + (-0.5)\frac{L}{2} \end{Bmatrix} \quad (31)$$

In (30), the displacement boundary condition of (14) is  $u_1 = 2$  at node 1 of coordinate  $x_a = 1$ . By the force boundary condition of (15), the externally applied flux  $R = -0.5$  (at the terminal node 3, of coordinate  $x_c = x_a + 2L = 2$ ) the unknowns namely the nodal displacements at nodes 2 and 3 ( $u_2$  and  $u_3$ ) can be obtained by solving (30) as:

$$\begin{Bmatrix} u_1 \\ u_2 \\ u_3 \end{Bmatrix} = \begin{Bmatrix} 2 \\ \frac{1}{A_1} \left[ R - \frac{2}{L} \left( \left( \ln \frac{x_a + L}{x_a} \right) - \frac{L}{x_a + 2L} \right) - A_2 u_1 \right] \\ \frac{1}{A_3} \left[ \frac{2}{x_c} - \frac{2}{L} \ln \left( \frac{x_c}{x_b} \right) + R \right] + \frac{1}{A_1} \left[ R - \frac{2}{L} \left( \ln \left( \frac{x_b}{x_a} \right) - \frac{L}{x_c} \right) - A_2 u_1 \right] \end{Bmatrix}. \quad (32)$$

The finite element computed strains in the two elements can be obtained from (22) as

$$\varepsilon^{h,e=1} = (u_2 - u_1)/L, \quad \varepsilon^{h,e=2} = (u_3 - u_2)/L. \quad (33a,b)$$

With numerical substitution,  $x_a = 1$ ,  $L = 0.5$ ,  $R = -0.5$ , (32) and (33) give the global nodal displacements and element strains as

$$\begin{Bmatrix} u_1 \\ u_2 \\ u_3 \end{Bmatrix} = \begin{Bmatrix} 2 \\ 1.551 \\ 1.365 \end{Bmatrix}, \quad \varepsilon^{h,e=1} = -0.897488, \quad \varepsilon^{h,e=2} = -0.371844. \quad (34)$$

## 6. Computed strain and nodal reactions in the elements

We present here the algebraic expressions and numerical values for the approximate strain and nodal reactions for the different elements as computed by the conventional finite element method.

For first element ( $x_a \leq x \leq x_b$ )

$$\varepsilon^{h,e=1} = \frac{1}{A_1 L} \left[ R - \frac{2}{L} \left( \ln \left( \frac{x_b}{x_a} \right) - \frac{L}{x_c} \right) \right]. \quad (35a)$$

With  $R = -0.5$  we have from (35a) the strain in element 1

$$\varepsilon^{h,e=1} = \left( \frac{-2 \ln(x_b/x_a) + 1/2(x_b - x_a)}{(1/2)((x_b)^2 - (x_a)^2)} \right) = \left( \frac{L - 4 \ln(x_b/x_a)}{L(2x_a + L)} \right). \quad (35b)$$

Now substituting  $x_a = 1$ ,  $L = 0.5$  we have from (35b) the strain in element 1,

$$\varepsilon^{h,e=1} = \frac{2}{2.5} \left[ \frac{1}{2} - 4 \ln \left( \frac{1.5}{1} \right) \right] = -0.897488. \quad (35c)$$

For the second element ( $x_b \leq x \leq x_c$ )

$$\varepsilon^{h,e=2} = \frac{1}{A_3 L} \left[ R - \frac{2}{L} \left[ \ln \left( \frac{x_c}{x_b} \right) - \frac{2}{x_c} \right] \right]. \quad (35d)$$

With  $R = -0.5$  we have from (35d) the strain in element 2

$$\varepsilon^{h,e=2} = \left( \frac{-2 \ln(x_c/x_b) + (1/2)(x_c - x_b)}{(1/2)((x_c)^2 - (x_b)^2)} \right) = \left( \frac{L - 4 \ln(x_c/x_b)}{L(2x_b + L)} \right). \quad (35e)$$

Now substituting  $x_b = x_a + L = 1.5$ ,  $x_c = 2$ ,  $L = 0.5$  we have from (35e) the strain in element 2 as

$$\varepsilon^{h,e=2} = \frac{2}{3.5} \left[ \frac{1}{2} - 4 \ln \left( \frac{4}{3} \right) \right] = -0.371844. \quad (35f)$$

From the analytical solution the elemental nodal reactions for the first element are, from (20a,b)

$$(R_1^{e=1}) = - \left( -\frac{2}{x_a} + \frac{1}{2} \right), \quad (R_2^{e=1}) = \left( -\frac{2}{x_b} + \frac{1}{2} \right). \quad (36a,b)$$

Using (26) and (28), the finite element computed nodal reactions for the first element are

$$R_1^{h,e=1} = - \left( R + \frac{2}{x_a + 2L} - \frac{2}{x_a} \right), \quad R_2^{h,e=1} = \left( R + \frac{2}{x_a + 2L} - \frac{2}{x_a + L} \right). \quad (37a,b)$$

Substituting  $x_a = 1$ ,  $L = 0.5$  and  $R = -0.5$  these expressions yield the finite element computed nodal reactions for the first element that match with the analytical reactions (given in (20a,b)) exactly,

$$R_1^{h,e=1} = R_1^{e=1} = (3/2) \text{ and } R_2^{h,e=1} = R_2^{e=1} = (-5/6) \quad (38a,b)$$

Thus the finite element computed reactions at the element level exactly matches the analytical reactions for the first element. The same may be easily shown for the second element in the following way:

From (20), the analytical elemental nodal reactions for the second element are

$$(R_1^{e=2}) = -\left(-\frac{2}{x_b} + \frac{1}{2}\right), \quad (R_2^{e=2}) = \left(-\frac{2}{x_c} + \frac{1}{2}\right) = R. \quad (39a,b)$$

Using (26) and (29), the finite element computed nodal reactions for the second element are

$$R_1^{h,e=2} = -R + \frac{2}{L} \left[ \frac{L}{x_a + L} - \frac{L}{x_a + 2L} \right] \text{ and } R_2^{h,e=2} = R. \quad (40a,b)$$

Substituting  $x_a = 1$ ,  $L = 0.5$  and  $R = -0.5$  these expressions yield the finite element computed nodal reactions for the second element, that match with the analytical reactions (given in (20c,d)) exactly,

$$R_1^{h,e=2} = R_1^{e=2} = (5/6) \text{ and } R_2^{h,e=2} = R_2^{e=2} = (-1/2). \quad (41a,b)$$

## 7. Prediction of finite element results from projection theorem

It is observed from (38a,b) and (41a,b) that, for the case considered, the discretization process conserves the nodal reactions. Thus, for these elements, according to the projection theorem, the finite element solutions behave in a best-fit sense to the analytical solutions. To demonstrate this, the best-fit is computed independently and compared with finite element results. The best-fit rule may be stated in the following form.

*When the element reactions, obtained analytically, are conserved by finite element discretization, ( $R^{h,e} = R^e$ ) (12), the element strains as computed by the finite element method are actually the best-fits to, or the orthogonal projections onto a pre-determined subspace 'β', of the element analytical strains (Prathap & Mukherjee 2003).*

From (12) where,

$$a(u^h, u^h)^e = \langle \varepsilon^{h,e}, \varepsilon^{h,e} \rangle = \{\delta^e\}^T \int_e [B]^T [D][B] dv^e \{\delta^e\}$$

$$\text{and } a(u, u^h)^e = \langle \varepsilon^{h,e}, \varepsilon^e \rangle = \{\delta^e\}^T \int_e [B]^T [D] \{\varepsilon^e\} dv^e,$$

we get the following form

$$\begin{aligned} & \{\delta^e\}^T \int_e [B]^T [D][B] dv^e \{\delta^e\} - \{\delta^e\}^T \int_e [B]^T [D] \{\varepsilon^e\} dv^e \\ & = \{\delta^e\}^T [\{R^{h,e}\} - \{R^e\}]. \end{aligned} \quad (42a)$$

When analytical reactions in an element are conserved, the right hand side of (12) vanishes, giving rise to the following *normal equation* for arbitrary, non-trivial nodal displacements in the element

$$\int_e [B]^T [D][B] dv^e \{\delta^e\} = \int_e [B]^T [D] \{\varepsilon^e\} dv^e, \quad (42b)$$

where  $[D]$  is the positive definite element rigidity matrix. The best-fit strain in an element,  $\bar{\varepsilon}^e$ , can then be obtained from by the projection formula (Mukherjee & Prathap 2001, 2002), given as

$$\bar{\varepsilon}^e = \sum_{j=1}^m \frac{\langle \varepsilon^e, v_j \rangle}{\langle v_j, v_j \rangle} \{v_j\}, \quad \langle v_i, v_j \rangle = 0 (i \neq j), \quad (43)$$

where,  $\varepsilon^e$  is the exact strain and  $v_j$  are the  $m$  orthogonal basis vectors spanning the finite  $m$ -dimensional vector subspace ' $\beta$ '. These orthogonal basis vectors can be generated by applying the *Gram-Schmidt* procedure (Edward & Penny 1988) to the column vectors of the strain-displacement matrix  $[B]$ . The inner product of two vectors  $\{a\}$  and  $\{b\}$  each of size  $r$  rows, with respect to a weighting symmetric and positive definite element rigidity matrix  $[D]$  of size  $r \times r$  is defined as,

$$\langle a, b \rangle = \int_e \{a\}^T [D] \{b\} dx. \quad (44)$$

For our problem the element rigidity matrix is given by  $[D] = x$  and the generic coordinate is  $v^e = x$ . From (33) the strain-displacement matrix for our problem is,

$$[B] = \begin{bmatrix} -\frac{1}{L^e} & \frac{1}{L^e} \end{bmatrix}. \quad (45a)$$

For the present problem, it can be shown, using the Gram-Schmidt procedure upon the column vectors of the  $[B]$  matrix for an element (45), that a one-dimensional space ' $\beta$ ' (spanned by a single basis vector) is generated which contains the column vectors of  $[B]$ . On this subspace, the projection of the analytical strain is captured as the best-fit strain (figure 1). The basis vector spanning space  $\beta$  is

$$\{v_1\} = 1. \quad (45b)$$

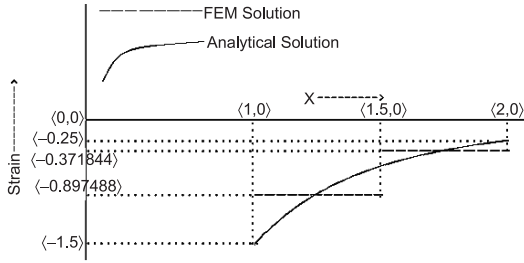
### 7.1 Best-fit and finite element strains for the elements

Using the projection formula (43) for strain projection, we have the best-fit strain for the first element as the orthogonal projection of the corresponding analytical strain onto the one-dimensional subspace ' $\beta$ ' ( $m = 1$ )

$$\bar{\varepsilon}^{e=1} = \frac{\langle \varepsilon^{e=1}, v_1 \rangle}{\langle v_1, v_1 \rangle} \{v_1\} \quad (46a)$$

$$= \left( \frac{\int_{x_a}^{x_b} (\varepsilon^{e=1}) \cdot x \cdot 1 \cdot dx}{\int_{x_a}^{x_b} 1 \cdot x \cdot 1 \cdot dx} \right) = \left( \frac{\int_{x_a}^{x_b} [-(2/x^2) + (1/2x)] \cdot x \cdot 1 \cdot dx}{\int_{x_a}^{x_b} 1 \cdot x \cdot 1 \cdot dx} \right),$$

$$\bar{\varepsilon}^{e=1} = \left( \frac{-2 \ln(x_b/x_a) + (1/2)(x_b - x_a)}{(1/2)((x_b)^2 - (x_a)^2)} \right). \quad (46b)$$



**Figure 4.** Comparison of the values of analytical strains and the finite element computed strains in the elements of the domain.

Substituting  $x_b = x_a + L = 1.5$ ,  $x_a = 1$  and  $L = 1/2$ , the best-fit strain in the first element is given as

$$\bar{\varepsilon}^{e=1} = \left( \frac{L - 4 \ln(x_b/x_a)}{L(2x_a + L)} \right) = \frac{2}{2.5} \left[ \frac{1}{2} - 4 \ln \left( \frac{1.5}{1} \right) \right] = -0.897488. \quad (46c)$$

Clearly, (46b,c) representing the best fit strain for element 1 exactly matches with (35b,c) which represents the element strain obtained by the conventional finite element method, i.e.  $\varepsilon^{h,e=1} = \bar{\varepsilon}^{e=1}$ . Likewise for the second element, the best-fit strain is given as

$$\begin{aligned} \bar{\varepsilon}^{e=2} &= \frac{\langle \varepsilon^{e=2}, v_1 \rangle}{\langle v_1, v_1 \rangle} \{v_1\} \\ &= \left( \frac{-2 \ln(x_c/x_b) + (1/2)(x_c - x_b)}{(1/2)((x_c)^2 - (x_b)^2)} \right) = \left( \frac{L - 4 \ln(x_c/x_b)}{L(2x_b + L)} \right) \end{aligned} \quad (46d)$$

Substituting  $x_b = 1.5$ ,  $x_c = 2$  and  $L = 1/2$ , the best-fit strain in the first element is given as

$$\bar{\varepsilon}^{e=2} = \frac{2}{3.5} \left[ \frac{1}{2} - 4 \ln \left( \frac{2}{1.5} \right) \right] = -0.371844. \quad (46e)$$

Again, (46d,e) representing best-fit strain for element 2 exactly matches with (35e,f) that represents the element strain obtained by the conventional finite element method, i.e.  $\varepsilon^{h,e=2} = \bar{\varepsilon}^{e=2}$ .

These results have been confirmed by actual results from conventional finite element codes. These strains are shown in figure 4. The finite element strains are actually the best-fits to (or the orthogonal projections of) the analytical strains in each element, as shown earlier in figure 1. This validates the postulates of the best-fit paradigm.

## 8. A critical examination of weak forms at the element level

We present here the physical significance of the element nodal reactions through a critical examination of the weighted residual statement of the present problem at the element level. For the differential equation (13), virtual work principle leads to the following statement for the element ( $x_1 \leq x \leq x_2$ ),

$$\int_e \left[ -\frac{d}{dx} \left( x \frac{du}{dx} \right) - \left( -\frac{2}{x^2} \right) \right] \cdot u^h \cdot dx = 0 \quad (47)$$

where  $u$  is the analytical solution and  $u^h$  is the approximate solution within the element. Expanding the above equation through integration by parts, we have

$$\int_e \frac{du^h}{dx} \left( x \frac{du}{dx} \right) dx - \int_e u^h \left( -\frac{2}{x^2} \right) dx - \left[ u_1 \left( -x \frac{du}{dx} \right)_{x=x_1} + u_2 \left( x \frac{du}{dx} \right)_{x=x_2} \right] = 0$$

or

$$a(u, u^h)^e - (f, u^h)^e - [u_1 R_1^e + u_2 R_2^e] = 0$$

or

$$a(u, u^h)^e - (f, u^h)^e - [u^h R^e]_b = 0, \quad (48)$$

where

$$R_1^e = (-x(du/dx))_{x=x_1}, \quad R_2^e = (x(du/dx))_{x=x_2} \quad (49a,b)$$

are the analytical values of the nodal reactions, or constraints, on the element. Thus (48) is directly reduced to form (10). This equation is satisfied since  $u$  is the exact solution of the differential equation, and therefore equilibrium is satisfied at the element level. Note that it is true also when the approximate function  $u^h$  is replaced by the analytical function  $u$ , as a special case. The analytical derivative  $(du/dx)$  is continuous across nodal connectivity.

Now we examine virtual work with the approximate function  $u^h$  substituted in the actual differential equation for  $u$ . This leads to, using integration by parts,

$$\int_e \left[ -\frac{d}{dx} \left( x \frac{du^h}{dx} \right) - \left( -\frac{2}{x^2} \right) \right] \cdot u^h \cdot dx = \int_e \frac{du^h}{dx} \left( x \frac{du^h}{dx} \right) dx - \int_e u^h \left( -\frac{2}{x^2} \right) dx - \left\{ u_1 \left( -x \frac{du^h}{dx} \right)_{x=x_1} + u_2 \left( x \frac{du^h}{dx} \right)_{x=x_2} \right\}.$$

Thus

$$\int_e \left[ -\frac{d}{dx} \left( x \frac{du^h}{dx} \right) - \left( -\frac{2}{x^2} \right) \right] \cdot u^h \cdot dx = a(u^h, u^h)^e - (f, u^h)^e - \left\{ u_1 \left( -x \frac{du^h}{dx} \right)_{x=x_1} + u_2 \left( x \frac{du^h}{dx} \right)_{x=x_2} \right\}. \quad (50)$$

In principle, the right hand side of (50), with arbitrary nodal displacements for the element, does not really vanish to zero, since  $u^h$  does not exactly satisfy the differential equation and therefore equilibrium is not satisfied at the element level. We can *force* this weak form (50)

to vanish, provided we replace the boundary force terms (acting through nodal connectivity) by appropriate nodal reactions for element equilibrium, so that

$$a(u^h, u^h)^e - (f, u^h)^e - [u_1 R_1^{h,e} + u_2 R_2^{h,e}] = 0$$

or

$$a(u^h, u^h)^e - (f, u^h)^e - [u^h R^{h,e}]_b = 0. \quad (51)$$

Here, unlike the analytical case, the computed nodal reactions have the following properties (inequalities, denoted by  $\neq$ ),

$$R_1^{h,e} \neq \left( -x \frac{du^h}{dx} \right)_{x=x_1} \quad \text{and} \quad R_2^{h,e} \neq \left( x \frac{du^h}{dx} \right)_{x=x_2} \quad (52a,b)$$

Comparing (49) with (52), we appreciate the penalty introduced for using an approximate function  $u^h$  in the actual differential equation. Note that while the gradient ( $du^h/dx$ ) is constant within the element with linear displacement approximation, it is not continuous across the nodal connectivity. The weak form (51) leads to (11). In practice, element nodal reactions  $\{R^{h,e}\}$  with approximate function  $u^h$  in the differential equation can be computed only from (26) after the solution for the nodal displacements is available.

In (51), it can be observed that with single element solution of the full domain, however, *either* vanishing kinematic *or* vanishing force conditions apply at the element boundaries. For such cases, kinematic constraints again impose nodal reactions in the form of Lagrange multipliers. With vanishing of the boundary terms, the weak form (4) for the global domain is obtained.

It has been demonstrated in this paper that for variationally correct formulations, the best-fit paradigm is valid when the finite element computed reactions  $\{R^{h,e}\}$  agree exactly with the analytical reactions  $\{R^e\}$ . Cases of violation of the best-fit rule by a precise amount have been examined by Sangeeta *et al* (2003). Such *prima facie* violations result from nodal reaction errors. It can be shown that the finite element strains of such systems are best-fits to the analytical strains of *modified* systems stiffened by spurious forces arising from nodal reaction errors of the original systems.

## 9. Conclusions

This paper confirms that in finite element modelling the best-fit paradigm (Prathap 1993) gives an accurate picture of how strains and stresses are computed at element and global level. The present study indicates that for any system discretized such that it conserves the analytical element nodal reactions, the finite element computed element strain always behaves as the best-fit to the analytical strain (figure 1). For the present work, a constant strain element has been taken for the analysis yielding a constant basis vector for strain projection. However, in general, with arbitrary variation of the elastic rigidity and strain vector within the element, appropriate element rigidity matrix  $[D]$  should be used in the definition of the inner product, (44), of each element for evaluation of the orthogonal basis vectors spanning subspace ' $\beta$ ' using the Gram-Schmidt method and for the determination of the strain projections onto  $\beta$ .

One of the authors (HM) is particularly thankful to Dr Gangan Prathap who introduced him to the subject and also to Dr K Sangeeta for her kind help and support.

### List of symbols

$[B]$	strain-displacement matrix;
$[D]$	rigidity matrix;
$\{F^e\}$	force vectors;
$[K], [K^e]$	stiffness matrices;
$L$	element length;
$v_j$	orthogonal basis vector of ‘ $\beta$ ’;
$N_i, [N]$	shape function and shape function matrix;
$\{R^e\}$	nodal reaction vector;
$u$	analytical solution;
$\beta$	function sub-space;
$\{\delta^e\}$	nodal displacement vector;
$\varepsilon^e$	elemental analytical strain;
$\varepsilon^{h,e}$	finite element method strain;
$\bar{\varepsilon}^e$	best-fit strain.

### References

- Burnett D S 1987 *Finite element analysis – From concepts to applications* (California: Addison-Wesley)
- Edwards L H, Penny D E 1988 *Elementary linear algebra* (Englewood Cliffs, NJ: Prentice Hall)
- Mukherjee S, Prathap G 2001 Analysis of shear locking in Timoshenko beam elements using function space approach. *Commun. Numer. Methods Engng.* 17: 385–393
- Mukherjee S, Prathap G 2002 Analysis of delayed convergence in the three-noded Timoshenko beam element using the function approach. *Sādhanā* 27: 507–526
- Prathap G 1993 *The finite element method in structural mechanics* (Dordrecht: Kluwer Academic)
- Prathap G 1996 Finite element analysis and the stress correspondence paradigm. *Sādhanā* 21: 525–546
- Prathap G, Hattangadi A 1998 Stress without strain-finite element analysis as least action process. NAL Project Document PD ST 9803
- Prathap G, Mukherjee S 2003 The engineer grapples with theorem 1·1 and lemma 6·3 of Strang and Fix. *Curr. Sci.* 85: 989–994
- Reddy J N 2003 *An introduction to the finite element method* (New Delhi: Tata McGraw-Hill)
- Sangeeta K, Mukherjee S, Prathap G 2003 Generalization of the projection theorem for finite element analysis. *Research Report C-MMACS*: CM 0305
- Strang G, Fix G J 1973 *An analysis of the finite element method* (Englewood Cliffs, NJ: Prentice Hall)