PERGAMON

Computers and Structures 79 (2001) 1275–1285

Computers & Structures

www.elsevier.com/locate/compstruc

On higher-order-accuracy points in isoparametric finite element analysis and an application to error assessment

Jean-François Hiller, Klaus-Jürgen Bathe *

Department of Mechanical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

Received 14 November 2000; accepted 31 January 2001

Abstract

We consider the systematic calculation of higher-order-accuracy points in finite element analysis. The cases of varying material properties and distortion of elements are included. The results obtained are used to establish an error estimator that is element-based and applicable to varying material properties. Some example problems are solved and show the applicability of the error estimator. While one-dimensional problems are studied in detail, the extension to two- and three-dimensional analysis is indicated. © 2001 Elsevier Science Ltd. All rights reserved.

1. Introduction

Accurate assessment of the error is a crucial issue in finite element analysis. In particular, error estimates are an indispensable part of any mesh refinement process. Numerous contributions have been made to the development of error estimates; see the developments of implicit and explicit estimators, the truth mesh approach and recovery-based procedures (see Refs. [1,2]).

Much attention has been given to the Zienkiewicz–Zhu algorithm (see Refs. [3,4]), a recovery-based procedure which employs a least squares fit technique to extrapolate the finite element values of stresses at appropriately chosen sampling points to nodes. However, the Zienkiewicz–Zhu algorithm is not ideal with respect to some issues. Recovery-based techniques, because they are patch-based, smoothen out the effects of material discontinuities on the solution unless special procedures are used. Also, the program data structures required by patch-based algorithms when employed in engineering practice are complicated. These two disadvantages would not be present if an element-based error estimator were available. Further, recovery-based algorithms require the calculation of stresses at certain sampling points that are not the points used for the evaluation of the stiffness matrix (assuming that "full" numerical integration is employed [5]). This is a significant shortcoming in general non-linear analysis.

The choice of appropriate sampling points is a major issue in the development of a recovery-based error estimator. For the smoothing algorithm to yield a recovered stress field that is significantly closer to the exact solution than the finite element stress field, sampling points need be chosen that provide higher-order convergence [6]. Frequently, the algorithm developed by Barlow [7] is used to select these points, but these points have been the subject of a debate (see Refs. [8–10]) mainly due to the fact that Barlow's algorithm may only yield approximate locations for the points where higher-order convergence is observed. MacKinnon and Carey have proposed an expression for the location of the higher-order-accuracy points when solving the Laplace equation (see Ref. [11]). However, the results given are rather restrictive, since even for the one-dimensional scalar problem of a bar in tension with varying material properties and/or cross-sectional area, an expression for the location of the higher-order-accuracy points was not given.

With the present paper, we want to confirm some of the known results, clear up some of the controversies on earlier results published by other authors, and give

* Corresponding author. Tel.: +1-617-253-6645; fax: +1-617-253-2275.

E-mail address: kjb@mit.edu (K.J. Bathe).

0045-7949/01/$ - see front matter © 2001 Elsevier Science Ltd. All rights reserved.
PII: S0045-7949(01)00017-7

insight and applications that should be useful in the further development of error measures.

The first objective of this paper is to show how the existence and the exact location of higher-order-accuracy points can be systematically established. We also identify under what conditions the higher-order-accuracy points coincide with the lower-order Gauss points (MacKinnon and Carey proved this result for the Laplace equation) and thereby clear up the controversy on these points. The second objective is to show how these results on higher-order-accuracy points can be used to develop an error estimator that is *element-based* (that is, not patch-based), that does not require the evaluation of the finite element stresses at sampling points, and that is applicable to varying material properties. This error estimator extends the work of Babuška and Szabó using hierarchical shape functions (i.e. the p method, see e.g. [12]). We give detailed results for a one-dimensional scalar problem solved with isoparametric elements and indicate how these results can be extended to more challenging problems, such as the Poisson equation solved over two- or three-dimensional domains.

2. Higher-order-accuracy points

The objective in this section is to derive the higher-order-accuracy points. We consider one-dimensional problems.

2.1. Earlier work

The notion that the finite element method yields strains that are more accurate at certain points than at others is the foundation of recovery-based error estimators (see Ref. [1]).

We shall from here on use the following definition:

In an N-node element, the higher-order-accuracy points are defined as those points, constant in the natural coordinate variable, at which the finite element strains are equal to the exact strains whenever the exact displacements are any polynomial of order N or lower in the local coordinate.

We shall observe that such points do not always exist.

Since the finite element strain energy converges to the exact strain energy with order $2(N-1)$, the finite element strains converge to the exact strains with order $(N-1)$ on average; hence it is said that a higher order of convergence is observed at the higher-order-accuracy points.

The expression "Barlow points" (as well as "superconvergent points") has been used since Barlow's first paper on the subject in 1976 [7] to denote two different items. On the one hand, the term has been used to denote the locations of points of higher-order accuracy *in general*, and on the other hand the term has been used to denote the locations of the points predicted by a method proposed by Barlow in his original paper.

In this work, we call "Barlow points" the points determined by the method proposed by Barlow, whereas we call "higher-order-accuracy points" the higher-order-accuracy points defined above.

This distinction is necessary because the method proposed by Barlow is in general only approximate in giving the higher-order-accuracy points because of a key assumption. This assumption is that the exact solution corresponding to a displacement pattern one order higher than what the finite element can represent and the finite element solution have the same nodal displacements (see Ref. [7]). A notable fact is that the local coordinates of the Barlow points are independent of the material properties and element distortions.

MacKinnon and Carey showed that in an undistorted element the higher-order-accuracy points for the Laplace equation are located at the lower-order Gauss points [11]. They also showed that this is true for the Poisson equation in two and three dimensions. However, the Barlow points are located at the lower order Gauss points only when certain elements are used. For instance, for the Laplace equation in one dimension, the Barlow points are not located at the Gauss points for elements with more than four nodes.

Using numerical studies, some authors [4] have concentrated on lower order elements and confirmed that higher-order-accuracy is obtained at the lower order Gauss points. But because the term "Barlow points" was used to denote both the higher-order-accuracy points and the points calculated by Barlow's method, there has been some confusion as to where the higher-order-accuracy points are located in the general case (varying material properties, etc). Even in the case when constant material properties are used, it is still a debated question [8–10] whether the higher-order-accuracy points coincide for elements of any order with the lower-order Gauss points despite the proof by MacKinnon and Carey [11] of this result for this specific simple case. In the more complex case of non-constant material properties and cross-sectional area, even less is known, and some authors [10] have questioned the very existence of predictable higher-order-accuracy points in this case.

We present in this section a proof of the existence of higher-order-accuracy points in the general case of varying material property and varying cross-sectional area and distortion of elements for a bar subjected to a distributed load. This proof is based on element orthogonal displacement patterns. In the past, element orthogonal displacement patterns were used by various researchers in the development of finite elements (see Refs. [5,12,13]).

2.2. Bar with distributed load

Let us consider a one-dimensional problem involving a bar of varying Young's modulus $E(x)$ and cross-sectional area $A(x)$ subjected to a loading force per unit length $f(x)$, and modeled using P (possibly distorted) N-node elements ($N > 1$). The exact solution to the problem u satisfies the equation

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(E(x)A(x)\frac{\mathrm{d}u}{\mathrm{d}x}\right) = -f(x) \tag{1}$$

in the domain $\Omega =]L, R[$, plus boundary conditions at the two ends of the domain. Possible boundary conditions are:

1. Essential boundary condition: the boundary displacement is known, the boundary force is unknown.
2. Natural boundary condition: the boundary displacement is unknown, the boundary force is known.

For well-posedness, at least one of the two boundary conditions has to be essential. We will always assume here that the boundary condition at $x = L$ is essential. The other boundary condition (at $x = R$) can be either natural or essential.

For any element M, we can define the element energy scalar product for the derivatives of the displacement field [5]:

$$a^M\left(\frac{\mathrm{d}u}{\mathrm{d}r}, \frac{\mathrm{d}v}{\mathrm{d}r}\right) = \int_M EA\frac{\mathrm{d}u}{\mathrm{d}r}\frac{\mathrm{d}v}{\mathrm{d}r}\frac{\mathrm{d}r}{\mathrm{d}x}\,\mathrm{d}r \tag{2}$$

where isoparametric elements are used, with r being the natural coordinate, and the global scalar product for the problem:

$$a\left(\frac{\mathrm{d}u}{\mathrm{d}x}, \frac{\mathrm{d}v}{\mathrm{d}x}\right) = \int_\Omega EA\frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}v}{\mathrm{d}x}\,\mathrm{d}x \tag{3}$$

We assume that in every element the usual polynomial shape functions (denoted by h_1 to h_N) for isoparametric displacement-based finite elements are employed so that

$$\mathrm{Span}(h_1(r), \ldots, h_N(r)) = \mathrm{Span}(1, r, \ldots, r^{N-1}) \tag{4}$$

Let us define $T_0^M(r) = 1$ and $T_1^M(r) = r$. Then the following result can be shown to hold:

For any element M, there exists a unique set T^M of polynomials T_k^M ($T_k^M = a_{kk}^M r^k + \cdots + a_{k1}^M r + a_{k0}^M$) of order k, $k \geqslant 2$, in the local coordinates, such that these polynomials satisfy:

1. The leading coefficient a_{kk}^M of T_k^M is 1.
2. The trailing coefficient a_{k0}^M of T_k^M is 0.
3. $a^M\left(\frac{\mathrm{d}T_k^M}{\mathrm{d}r}, \frac{\mathrm{d}T_j^M}{\mathrm{d}r}\right) = 0 \quad \text{for } j = 0, \ldots, k-1$ (5)

These polynomials can readily be calculated using the Gram–Schmidt procedure.

We can now determine the location of higher-order-accuracy points in a specific element M for this problem.

2.2.1. Undistorted elements, constant Young's modulus and constant cross-sectional area

In this section we consider the case $EA\,(\mathrm{d}r/\mathrm{d}x)$ constant over every element M in the mesh. This is clearly the case when we consider undistorted elements and constant Young's modulus and cross-sectional area.In this case, the polynomials T_k^M (where $k \geqslant 2$) in the set T^M take on a particular form:

$$T_k^M(r) = \frac{k2^{(k-1)}((k-1)!)^2}{(2(k-1))!}\int_0^r P_{k-1}(\xi)\,\mathrm{d}\xi \tag{6}$$

where P_k is the kth Legendre polynomial.

Proof. This result is obtained from the following two standard properties of the Legendre polynomials [14]:

$$1.\ \int_{-1}^{1} P_i(r)P_j(r)\,\mathrm{d}r = \frac{2\delta_{ij}}{2i+1} \tag{7}$$

$$2.\ P_i(r) = \frac{(2i)!}{2^i(i!)^2}r^i + c_i(r) \tag{8}$$

where δ_{ij} denotes the Kronecker delta and $c_i(r)$ is a polynomial of order lower than i. Uniqueness of the set T^M allows us to conclude that T_k^M is given by Eq. (6).

In this case the polynomials in the set T^M are independent of the element M. We have, considering $k = 1$ to 7:

$$\begin{aligned}
T_0^M(r) &= 1\\
T_1^M(r) &= r\\
T_2^M(r) &= r^2\\
T_3^M(r) &= r^3 - r\\
T_4^M(r) &= r^4 - \frac{6}{5}r^2\\
T_5^M(r) &= r^5 - \frac{10}{7}r^3 + \frac{3}{7}r\\
T_6^M(r) &= r^6 - \frac{5}{3}r^4 + \frac{5}{7}r^2\\
T_7^M(r) &= r^7 - \frac{63}{33}r^5 + \frac{35}{33}r^3 - \frac{5}{33}r
\end{aligned} \tag{9}$$

For $j \geqslant 2$, $T_j^M(r)$ is orthogonal to r in the element scalar product defined above, which immediately implies that $T_j^M(-1) = T_j^M(1)$. Let us define the set S^M of polynomials S_j^M defined over the whole computational domain Ω with $S_j^M(r) = T_j^M(r) - T_j^M(1) = T_j^M(r) - T_j^M(-1)$, $j \geqslant 2$, over element M and S_j^M equal zero everywhere else.

We note that in this case (with $EA(\mathrm{d}r/\mathrm{d}x)$ constant over element M) the S_j^M's ($j \geqslant 2$) are the hierarchical functions used as shape functions in the design of finite elements (see Ref. [12]).

Also, we define $S_1^M = r + 1$ over element M and $S_1^M = 0$ elsewhere. We also define $\tilde{S}_1^M = r + 1$ over element M, constant elsewhere and continuous. Finally, We define S_0^M to be equal to 1 over element M and 0 elsewhere.

It is immediately seen that if $1 < j < N$ (where N still denotes the number of nodes per element), S_j^M is an element of the finite element space, whereas if $j \geqslant N$, S_j^M is orthogonal to the finite element space in the global scalar product $a(\cdot,\cdot)$ (see Eq. (3)).

Let us consider that the boundary condition at $x = R$ is natural, then $\tilde{S}_1^M$ is also in the finite element space.

Let us assume that the exact solution (i.e. the exact solution to Eq. (1) over element M is analytic in the local coordinate (i.e. it can be developed in the form of a Taylor expansion of r over element M). Then, because the set S^M is a basis of the space of polynomials over element M, there is a unique set a^M of coefficients a_k^M ($k \geqslant 0$) such that, over element M,

$$u(r) = \sum_{i=0}^{\infty} a_i^M S_i^M(r) \tag{10}$$

Let us develop the finite element solution obtained from the solution of this problem with N-noded elements on the same basis over element M:

$$u_h(r) = \sum_{i=0}^{N-1} \alpha_i^M S_i^M(r) \tag{11}$$

and define $b_i^M = a_i^M - \alpha_i^M$ for any $1 \leqslant M \leqslant P$ (where P is the number of elements) and for any $i \geqslant 0$.

We have that, for any virtual displacement v in the finite element space,

$$a\left(\frac{\mathrm{d}u_h}{\mathrm{d}r}, \frac{\mathrm{d}v}{\mathrm{d}r}\right) = \int_\Omega fv = a\left(\frac{\mathrm{d}u}{\mathrm{d}r}, \frac{\mathrm{d}v}{\mathrm{d}r}\right) \tag{12}$$

With v equal to, in turn $S_2^M, S_3^M, \ldots, S_{N-1}^M$, we obtain $\alpha_i^M = a_i^M$, $i = 2, \ldots, N-1$. If one of the boundary conditions is natural then $\tilde{S}_1^M$ is in the finite element space and we also obtain that $a_1^M = \alpha_1^M$.

Hence

$$u_h(r) = \alpha_0^M + \sum_{i=1}^{N-1} a_i^M S_i^M(r) \tag{13}$$

Finally, continuity of the error in the displacements immediately implies that for any $1 \leqslant M \leqslant P$ we have $\alpha_0^M = a_0^M$. We conclude that in any element M the error can be decomposed into the form

$$e_h(r) = \sum_{i=N}^{\infty} b_i^M S_i^M(r) \tag{14}$$

Over element M, the error in the derivative $\mathrm{d}e_h/\mathrm{d}r$ then satisfies

$$\frac{\mathrm{d}e_h}{\mathrm{d}r} = \sum_{i=N}^{\infty} b_i^M \frac{\mathrm{d}S_i^M}{\mathrm{d}r} = \sum_{i=N}^{\infty} b_i^M \frac{\mathrm{d}T_i^M}{\mathrm{d}r} \tag{15}$$

and the accuracy of the derivative of the finite element solution is one order higher at the zeros of $\mathrm{d}T_N^M/\mathrm{d}r$.

Let us now consider the case when both boundary conditions are essential. In this case, $\tilde{S}_1^M$ is not an element of the finite element space, so that we cannot use it as our test function in equation (12). However, we still have, over element M

$$e_h = b_0^M S_0^M + b_1^M S_1^M + \sum_{i=N}^{\infty} b_i^M S_i^M \tag{16}$$

For $1 \leqslant M < P$ (where P is the number of elements), we can construct a test function v_M that has a slope of $+1$ in the local coordinate over element M, a slope of -1 in the local coordinate over element $M+1$, is continuous, and satisfies the boundary conditions (see Fig. 1).

Using the orthogonality equation (12) with v_M as our test function and continuity, we obtain (with some algebra) as in the earlier case

$$b_i^M = 0 \quad \text{for } i = 0, 1 \text{ and } M = 1, \ldots, P \tag{17}$$

Conclusions: For an undistorted N-node element M used to model a bar of constant Young's modulus and constant cross-sectional area, higher-order-accuracy points are located at the zeros of $\mathrm{d}T_N^M/\mathrm{d}r$, i.e. at the lower-order Gauss points of the element, because of Eq. (6). Since it was shown that T_N^M is unique up to a multiplicative and an additive constant, at no point other than the zeros of $\mathrm{d}T_N^M/\mathrm{d}r$ is higher-order accuracy achieved. As a side result, we note that the error in the displacements is zero at the end points of all elements, a classic result [5].

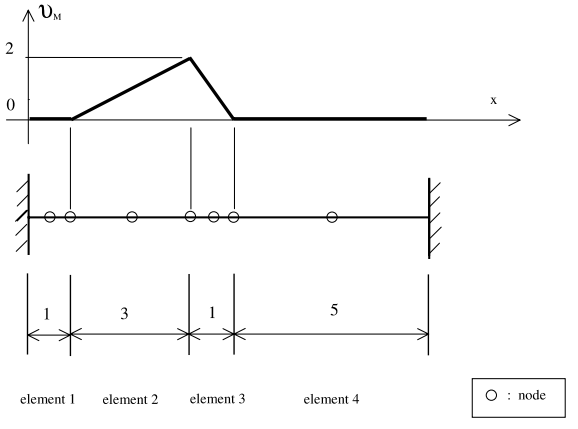

Fig. 1. v_M function. Elements M and $M+1$ are not distorted ($M = 2$).

Physical interpretation: Of all functions that are supported by element M, S_N^M is the polynomial of the lowest order that is orthogonal to the finite element space. S_N^M is therefore the first "hidden" displacement pattern: it is the lowest order polynomial displacement pattern that cannot be picked up by the N-node element. However, at the points where $\mathrm{d}S_N^M/\mathrm{d}r = 0$, this displacement pattern does not create any strain, and therefore we obtain the same accuracy in the strain at these points whether we include this hidden displacement pattern or not.

Hence, whereas implicit algorithms explore the functional space orthogonal to the finite element space, recovery-based algorithms are based on the use of the zeros of the derivative of the lowest order polynomial in this orthogonal space as sampling points.

2.2.2. Distorted elements, varying Young's modulus and varying cross-sectional area

In the general case (i.e. when we no longer make the assumption that the product of the Young's modulus times the cross-sectional area and the determinant of the mapping function is a constant), we need to modify the above proof. Note that in this case the functions T_J^M need to be calculated for each element M depending on the conditions of the element.

In the case of a non-constant product $EA\,(\mathrm{d}r/\mathrm{d}x)$, the element orthogonal polynomials T_j^M in the set T^M in general no longer satisfy $T_j^M(-1) = T_j^M(1)$ so that it is not possible to construct a set S^M of polynomials S_j^M that are globally orthogonal to the finite element space by the above method. As a result, it is not immediately obvious that the higher-order-accuracy points are located at the zeros of the derivative of T_N^M, like in the case of a constant product $EA\,(\mathrm{d}r/\mathrm{d}x)$. Let us first consider the case of one natural boundary condition.

Again, we assume that the exact solution (i.e. the mathematical solution to Eq. (1)) over element M is analytic in the local coordinates. Thus we can find a set b^M of coefficients b_i^M such that over element M

$$u(r) = u_h(r) + \sum_{i=0}^{\infty} b_i^M T_i^M(r) \tag{18}$$

In the case $EA\,(\mathrm{d}r/\mathrm{d}x)$ being constant, we could argue that for $0 \leqslant i < N$ we had $b_i^M = 0$ because of the global orthogonality of the polynomials S_i^M. We can no longer use this argument. However, we can still write for any test function v

$$a^M\left(\frac{\mathrm{d}u}{\mathrm{d}r}, \frac{\mathrm{d}v}{\mathrm{d}r}\right) = \int_M fv + \left[EA\frac{\mathrm{d}u}{\mathrm{d}x}v\right]_{M-}^{M+} \tag{19}$$

where $M+$ and $M-$ are the coordinates of the right and left ends of element M.

For every polynomial T_k^M $(k > 1)$, we define as g_k^M the function of the local coordinate r that is linear and that equals T_k^M at 1 and at -1:

$$g_k^M(r) \overset{\text{def}}{=} \frac{T_k^M(1) + T_k^M(-1)}{2} + \frac{T_k^M(1) - T_k^M(-1)}{2} r \tag{20}$$

Using $T_j^M - g_j^M$ (with $j > 1$) as our test function v in Eq. (19), and replacing u by expression (18), we obtain

$$a^M\left(\frac{\mathrm{d}}{\mathrm{d}r}\left(u_h(r) + \sum_{i=0}^{\infty} b_i^M T_i^M(r)\right), \frac{\mathrm{d}(T_j^M(r) - g_j^M(r))}{\mathrm{d}r}\right) = \int_M f(T_j^M - g_j^M) \tag{21}$$

because the boundary terms vanish due to our choice of v. Using the bilinearity of the element scalar product, we obtain

$$\sum_{i=0}^{\infty} b_i^M \left(a^M\left(\frac{\mathrm{d}T_i^M}{\mathrm{d}r}, \frac{\mathrm{d}T_j^M}{\mathrm{d}r}\right) - a^M\left(\frac{\mathrm{d}T_i^M}{\mathrm{d}r}, \frac{\mathrm{d}g_j^M}{\mathrm{d}r}\right)\right) = \int_M f(T_j^M - g_j^M) - a^M\left(\frac{\mathrm{d}u_h}{\mathrm{d}r}, \frac{\mathrm{d}(T_j^M - g_j^M)}{\mathrm{d}r}\right) \tag{22}$$

Finally, exploiting the orthogonality properties of the set T^M we are left with

$$b_j^M a^M\left(\frac{\mathrm{d}T_j^M}{\mathrm{d}r}, \frac{\mathrm{d}T_j^M}{\mathrm{d}r}\right) - b_0^M a^M\left(\frac{\mathrm{d}T_0^M}{\mathrm{d}r}, \frac{\mathrm{d}g_j^M}{\mathrm{d}r}\right) - b_1^M a^M\left(\frac{\mathrm{d}T_1}{\mathrm{d}r}, \frac{\mathrm{d}g_j^M}{\mathrm{d}r}\right) = \int_M f(T_j^M - g_j^M) - a^M\left(\frac{\mathrm{d}u_h}{\mathrm{d}r}, \frac{\mathrm{d}(T_j^M - g_j^M)}{\mathrm{d}r}\right) \tag{23}$$

If we choose $1 < j < N$, the right-hand side of Eq. (23) vanishes, hence

$$b_j^M a^M\left(\frac{\mathrm{d}T_j^M}{\mathrm{d}r}, \frac{\mathrm{d}T_j^M}{\mathrm{d}r}\right) = b_0^M a^M\left(\frac{\mathrm{d}T_0^M}{\mathrm{d}r}, \frac{\mathrm{d}g_j^M}{\mathrm{d}r}\right) + b_1^M a^M\left(\frac{\mathrm{d}T_1^M}{\mathrm{d}r}, \frac{\mathrm{d}g_j^M}{\mathrm{d}r}\right) \quad \text{for } 1 < j < N \tag{24}$$

or, using the explicit form of the element scalar product $a^M(\cdot,\cdot)$ with T_0^M,

$$b_j^M a^M\left(\frac{\mathrm{d}T_j^M}{\mathrm{d}r}, \frac{\mathrm{d}T_j^M}{\mathrm{d}r}\right) = b_1^M a^M\left(\frac{\mathrm{d}T_1^M}{\mathrm{d}r}, \frac{\mathrm{d}g_j^M}{\mathrm{d}r}\right) \quad \text{for } 1 < j < N \tag{25}$$

Now we make use of the orthogonality property of the error (see Ref. [5]). With V_h the finite element interpolation space, we have

$$\forall v \in V_h,\ a(u - u_h, v) = 0 \tag{26}$$

with v chosen to be the $\tilde{S}_1^M$ function. It should be noted that this choice of v is a valid one for as long as one of the boundary conditions is natural. This gives

$$b_1^M = 0 \tag{27}$$

From Eqs. (27) and (25), we conclude that

$$\text{if } 1 \leqslant j < N,\ b_j^M = 0 \tag{28}$$

We can therefore write

$$\frac{\mathrm{d}u}{\mathrm{d}r} = \frac{\mathrm{d}u_h}{\mathrm{d}r} + \sum_{i=N}^{\infty} b_i^M T_i^M(r) \tag{29}$$

Conclusion: For a distorted N-node element used to model a bar of varying Young's modulus and varying cross-sectional area, higher-order-accuracy points are located at the zeros of $\mathrm{d}T_N^M/\mathrm{d}r$ if one of the boundary conditions is natural.

We now consider the case when both boundary conditions are essential: in this case, we show by a counter-example that higher-order-accuracy points do not exist.

Counterexample: We choose $L = 0,\ R = 4,\ A = 1,\ E = 1 + x$, and two identical three-node non-distorted elements, so that we have a three degree of freedom problem, both ends being constrained. We solve this finite element problem exactly using a symbolic mathematics program.

First, we impose forces such that the exact solution is

$$u_1(r) = r^3 - \frac{12}{55}r^2 - \frac{1}{2}r + \frac{79}{110} \tag{30}$$

over the first element ($r = x - 1$ being the local coordinate pertaining to that element) and

$$u_1(r) = r^3 - \frac{24}{235}r^2 - \frac{3}{2}r + \frac{283}{470} \tag{31}$$

over the second element ($r = x - 3$ being the local coordinate pertaining to that element). The error corresponding to this loading is, over the first element,

$$e_{h_1}(r) = r^3 - \frac{14}{69}r^2 - \frac{341}{345}r + \frac{74}{345} \tag{32}$$

and over the second element:

$$e_{h_1}(r) = r^3 - \frac{34}{345}r^2 - \frac{349}{345}r + \frac{38}{345} \tag{33}$$

Then, we impose forces such that the exact solution is

$$u_2(r) = 2r^3 - \frac{122}{55}r^2 - \frac{3}{2}r + \frac{299}{110} \tag{34}$$

over the first element ($r = x - 1$ being the local coordinate pertaining to that element) and

$$u_2(r) = r^3 - \frac{24}{235}r^2 - \frac{3}{2}r + \frac{283}{470} \tag{35}$$

over the second element ($r = x - 3$ being the local coordinate pertaining to that element). The error corresponding to this loading is, over the first element,

$$e_{h_2}(r) = 2r^3 - \frac{47}{115}r^2 - \frac{226}{115}r + \frac{51}{115} \tag{36}$$

and over the second element:

$$e_{h_2}(r) = r^3 - \frac{11}{115}r^2 - \frac{119}{115}r + \frac{3}{23} \tag{37}$$

In the second element, the error in the strain ($\mathrm{d}e_{h_1}/\mathrm{d}r$) corresponding to the first loading is zero at the points with coordinates $r = -0.548765\ldots$ and $r = 0.614466\ldots$, whereas the error in the strain ($\mathrm{d}e_{h_2}/\mathrm{d}r$) corresponding to the second loading is zero at the points with coordinates $r = -0.556286\ldots$ and $r = 0.620054\ldots$ Similarly, the zeros of the error in the strain depend also in the first element on the exact solution. This proves that in this case higher-order-accuracy points do not exist.

2.3. Other one-dimensional scalar products: the beam in flexure

The approach used in Section 2.2 can be more generally employed for other one-dimensional scalar problems that can be written in the form:

$$\text{Find } u \in V / \forall v \in V \quad a(u, v) = (f, v) \tag{38}$$

where $a(\cdot,\cdot)$ is a scalar product, $(f,\cdot)$ is a linear form, and V includes the space of polynomials. In this more general setting, we can still derive a unique set T^M of polynomials T_k^M of order k orthogonal to each other for the element scalar product $a^M(\cdot,\cdot)$. Again the set T^M can be constructed by recurrence using the Gram–Schmidt procedure. In general, only the polynomial T_0^M can be chosen arbitrarily. In the previous section we could also choose T_1^M arbitrarily because the displacements only appeared through their derivatives in the scalar product. Generally, if the lowest order derivative of the displacements that appears in the scalar product is the mth derivative, we can choose the first $(m+1)$ functions arbitrarily.

As an example, we consider the case of a beam in flexure, for which the bilinear form is (see Ref. [5])

$$a(u, v) = \int EI \frac{\mathrm{d}^2 u}{\mathrm{d}x^2} \frac{\mathrm{d}^2 v}{\mathrm{d}x^2}\, \mathrm{d}x \tag{39}$$

For this problem, we can impose $T_0^M(x) = 1$, $T_1^M(x) = x$, $T_2^M(x) = x^2 - 1$.

For $j > 2$, the polynomials T_j^M are determined by the Gram–Schmidt procedure up to a linear part which is arbitrary (because this part does not enter the bilinear form (39)). We can therefore choose this linear part such

that the polynomials T_j^M are zero at $r = 1$ and $r = -1$. A proof similar to the one used in the case of constant $EA\,(\mathrm{d}r/\mathrm{d}x)$ for the problem of the bar in tension still applies here and we obtain that for the beam in flexure problem

$$\frac{\mathrm{d}^2(u-u_h)}{\mathrm{d}r^2} = \sum_{i=N}^{\infty} b_i^M \frac{\mathrm{d}^2 S_i^M}{\mathrm{d}r^2} = \sum_{i=N}^{\infty} b_i^M \frac{\mathrm{d}^2 T_i^M}{\mathrm{d}r^2} \tag{40}$$

Hence the points where curvature is assessed with higher-order-accuracy in an N-degree of freedom finite element are located at the zeros of the second derivative of the polynomial T_N^M (which are the two order lower Gauss points in the case of a constant product $EI\,(\mathrm{d}r/\mathrm{d}x)$, because in this case the polynomials T_j^M are obtained by integrating the Legendre polynomials twice and then fixing their linear part as indicated above).

It should be realized that the set T^M is dependent on the order of the problem (as can be noticed by comparing the definitions for the sets T^M for the bar in tension problem and the beam problem) and also on the constants involved in the problem, as shown by the following example: consider the problem $-(\mathrm{d}^2u/\mathrm{d}x^2) + cu = f$ over $(-1, 1)$. For this problem, we can show that, for any non-distorted element M, the set T^M starts with the following polynomials:

$$\begin{aligned}
T_0^M(x) &= 1\\
T_1^M(x) &= x\\
T_2^M(x) &= x^2 - \frac{1}{3}\\
T_3^M(x) &= x^3 - \frac{2+\frac{2c}{5}}{2+\frac{2c}{3}}x\\
T_4^M(x) &= x^4 - \frac{6(c+21)}{7(c+15)}x^2 + \frac{3(35+c)}{35(15+c)}
\end{aligned} \tag{41}$$

We note that, for this new problem, only when j is even (and therefore T_j^M is even) do we have $T_j^M(1) = T_j^M(-1)$, unlike in the bar in tension problem. Clearly, the polynomials in the set T^M for this problem differ from those we obtain for the bar in tension and depend on the parameter c.

3. The Poisson problem

Let us now show how the same ideas can be extended to two-dimensional problems, such as the Poisson problem:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -f(x,y) \tag{42}$$

The bilinear form for this problem is

$$\int_\Omega \frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y}\,\mathrm{d}\Omega \tag{43}$$

For clarity, let us assume that this problem is solved using non-distorted nine-node quadrilateral isoparametric elements, and that the usual polynomial shape functions h_1 to h_9 are used so that

$$\begin{aligned}
&\mathrm{Span}(h_1(x,y), \ldots, h_9(x,y))\\
&\quad = \mathrm{Span}(1, x, y, x^2, xy, y^2, x^2y, xy^2, x^2y^2)
\end{aligned} \tag{44}$$

Since the order of convergence of the finite element strain energy is dictated by the degree of completeness of the polynomial approximation, we see that all that is needed to increase the order of convergence of this element is to include one polynomial with x^3 and one with y^3.

We can show that the polynomial which includes x^3 and is orthogonal to the finite element space is $x^3 - x$ and the polynomial that includes y^3 and is orthogonal to the finite element space is $y^3 - y$. The points where the derivatives of these two polynomials are zero are the lower order Gauss points for the element of interest, $x = \pm 1/\sqrt{3}$ and $y = \pm 1/\sqrt{3}$.

Obviously, the same approach is valid for elements other than the nine-node element. For elements used to solve the Poisson problem, the zeros of the derivatives of the lowest order polynomials orthogonal to the finite element space functions are always the lower order Gauss points, as was shown by MacKinnon and Carey.

4. Application to error assessment

In this section we show how the results of Section 2 can be used to develop a class of error estimators.

4.1. Error estimator for the bar problem

Again, let us consider the deformation of the bar subjected to the loading f. For the problem solution we use three-noded isoparametric finite elements.

We have shown in Section 2 that in any element, distorted or not, it is possible to develop a set of polynomials T_k^M $(k \geqslant 0)$ that are energy orthogonal to each other.

The error in element M can then be written as, using Eq. (18),

$$e_h = u - u_h = \sum_{i=0}^{\infty} b_i^M T_i^M(r) \tag{45}$$

For every polynomial T_k^M $(k \geqslant 2)$, we define as before $g_k^M(r) = ((T_k^M(1) + T_k^M(-1))/2) + ((T_k^M(1) - T_k^M(-1))/2)r$.

Starting from the differential equation for the problem, applying the Galerkin procedure to the test function $T_l^M - g_l^M$ $(l > 1)$, and using the orthogonality properties of the set T^M, we obtain:

$$b_l^M = \frac{\int_M f(T_l^M - g_l^M)\,\mathrm{d}x - \int_M EA\frac{\mathrm{d}(T_l^M - g_l^M)}{\mathrm{d}x}\frac{\mathrm{d}u_h}{\mathrm{d}x}\,\mathrm{d}x + b_1^M \int_M EA\frac{\mathrm{d}T_1^M}{\mathrm{d}x}\frac{\mathrm{d}g_l^M}{\mathrm{d}x}\,\mathrm{d}x}{\int_M EA(\frac{\mathrm{d}T_l^M}{\mathrm{d}x})^2\,\mathrm{d}x} \tag{46}$$

We have shown in Section 2.2 that under certain circumstances the constant $b_1^M = 0$, so that all the terms on the right-hand side are known or can be evaluated, and therefore for any l, b_l^M can be calculated. In case we do not have $b_1^M = 0$, we are still able to use Eq. (46) to obtain an estimate for b_l^M by neglecting the $b_1^M a^M((\mathrm{d}T_1^M/\mathrm{d}r), (\mathrm{d}g_l^M/\mathrm{d}r))$ term in the numerator.

In the general case, we use

$$b_l^{M*} = \frac{\int_M f(T_l^M - g_l^M)\,\mathrm{d}x - \int_M EA\frac{\mathrm{d}(T_l^M - g_l^M)}{\mathrm{d}x}\frac{\mathrm{d}u_h}{\mathrm{d}x}\,\mathrm{d}x}{\int_M EA(\frac{\mathrm{d}T_l^M}{\mathrm{d}x})^2\,\mathrm{d}x} \tag{47}$$

as an estimate for b_l^M and

$$e_h^*(r) = \sum_{i=N}^{i=N'} b_i^{M*} T_i^M(r) \tag{48}$$

as an estimate of the error. N' is an integer which determines the desired accuracy of our error estimator. We indicate in Section 4.2 that $N' = N + 1$ might be a good choice.

In the particular case when the product $EA\,(\mathrm{d}r/\mathrm{d}x)$ is constant, the second integral in the numerator of Eq. (46) equals zero. Since the set T^M is known, the b_l^M's can be calculated prior to the finite element solution. We have in this case an a priori error estimator, in essence by applying p-refinement. When the product $EA\,(\mathrm{d}r/\mathrm{d}x)$ is not constant, the second integral in the numerator must be evaluated and we have an a posteriori error estimator.

4.2. Examples

In this section, we test the error estimator presented above on some sample problems to assess its efficiency. We choose to include only two terms in the error estimate of Eq. (48), i.e. we take $N' = N + 1$. All problems are solved on the non-uniform mesh consisting of three-node undistorted elements given in Fig. 2. The Young's modulus varies as $E(x) = 1 + x$. In Examples 1 and 2, one end ($x = 0$) is clamped and one end ($x = 10$) is free. In Example 3, both ends are clamped to enforce zero displacements.

Example 1. We impose a loading per unit length such that the exact displacement is a polynomial of order 4:

$$u(x) = (x+1)^2(x-10)^2 - 100 \tag{49}$$

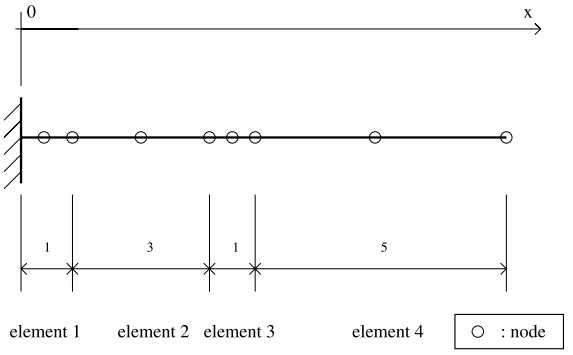

Fig. 2. Mesh: the four-element mesh employed to solve the example problems 1 and 2 of Section 4.2. In problem 3, the displacement at $x = 10$ is imposed to be zero.

Because $N' = N + 1 = 4$ and the exact solution is a polynomial of order 4, we expect our error estimator to give the exact error for this problem. The recovered strain $\mathrm{d}(u_h + e^*)/\mathrm{d}x$, the finite element strain $\mathrm{d}u_h/\mathrm{d}x$ and the exact strain are plotted in Fig. 3. We see that the recovered strain equals the exact strain, as expected.

Example 2. We next impose a loading per unit length such that the exact displacement cannot be developed as a finite Taylor series:

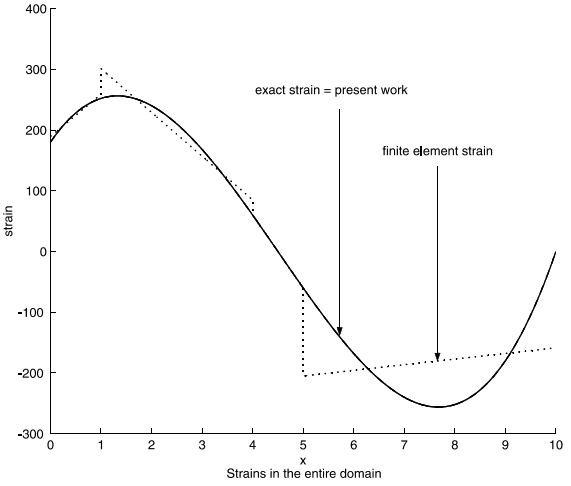

Fig. 3. Strains for Example 1. (—) $\mathrm{d}u/\mathrm{d}x = \mathrm{d}(u_h + e_h^*)/\mathrm{d}x$; $(\cdots)$ $\mathrm{d}u_h/\mathrm{d}x$.

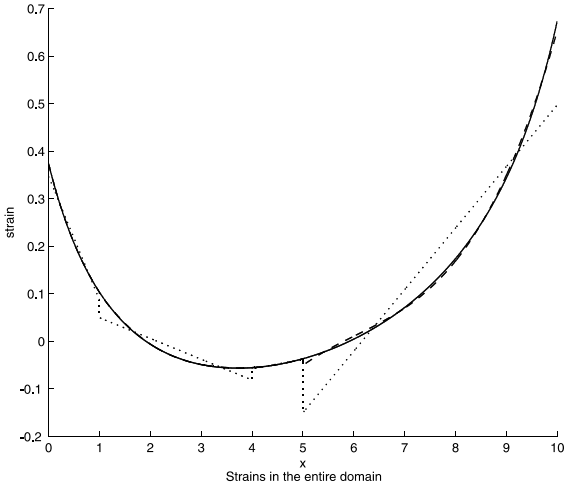

Fig. 4. Strains for Example 2. (—) $\mathrm{d}u/\mathrm{d}x$ (···) $\mathrm{d}u_h/\mathrm{d}x$; (- - -) $\mathrm{d}(u_h + e_h^*)/\mathrm{d}x$.

$$u(x) = \tan\left(\frac{x-5}{5}\right) - \frac{x-5}{5}\tan(1) + 0.0001x^3 \qquad (50)$$

Figs. 4 and 5 show that the recovered strain obtained with the algorithm introduced in this paper is much closer to the exact strain than is the finite element strain. The recovered strain is not equal to the exact strain because only two terms have been included in Eq. (48). Fig. 6 illustrates the performance of our algorithm compared to the Zienkiewicz–Zhu algorithm.

Example 3. In this problem, we apply forces corresponding to the exact displacement:

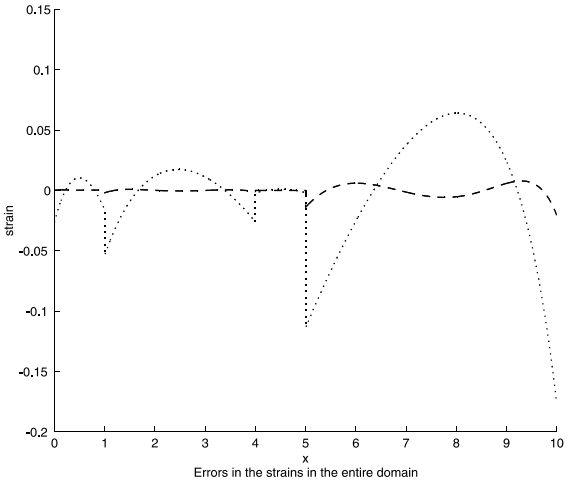

Fig. 5. Errors in the strain for Example 2. (···) $\mathrm{d}u_h/\mathrm{d}x - \mathrm{d}u/\mathrm{d}x$; (- - -) $\mathrm{d}(u_h + e_h^*)/\mathrm{d}x - \mathrm{d}u/\mathrm{d}x$.

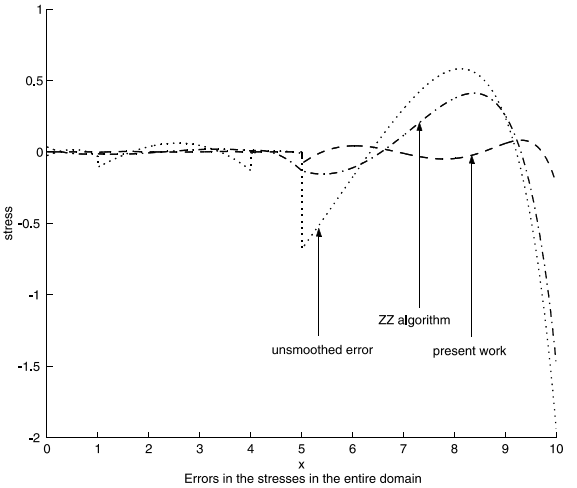

Fig. 6. Errors in the stress for Example 2. (···) $E(x)\mathrm{d}u_h/\mathrm{d}x - \sigma$; (- - -) $E(x)\mathrm{d}(u_h + e_h^*)/\mathrm{d}x - \sigma$; (-·-·-) $\sigma_{ZZ} - \sigma$, where σ and σ_{ZZ} denote the exact stress and the stress recovered by the Zienkiewicz–Zhu algorithm, respectively.

$$u(x) = \tan\left(\frac{x-5}{5}\right) - \frac{x-5}{5}\tan(1) \qquad (51)$$

Both boundary conditions are essential, and $EA\,(\mathrm{d}r/\mathrm{d}x)$ is not constant over the elements which implies that the calculated coefficients given by Eq. (47) are not equal to the exact coefficients in Eq. (46). Besides, the exact displacement can not be written in the form of a finite Taylor series, but again we choose to include only two terms in Eq. (48).

The strains presented in Fig. 7 show that despite these two approximations, the recovered strain is much closer to the exact strain than is the finite element strain (see also Fig. 8). Comparing the results of Example 3 to those of Example 2, we see that the accuracy of our error estimator is hardly affected by the fact that the boundary condition at $x = 10$ is now essential.

Fig. 9 illustrates the performance of our algorithm compared to the Zienkiewicz–Zhu algorithm.

5. Concluding remarks

In this paper, we have shown how the existence of higher-order-accuracy points can be related to hidden displacement patterns. In fact, we have shown that higher-order-accuracy points are the points where the lowest order hidden displacement pattern results in no strain. This interpretation enables us to calculate higher-order-accuracy points in cases that were not considered before. In particular, the case of distorted elements and varying material properties can be treated. Also, this

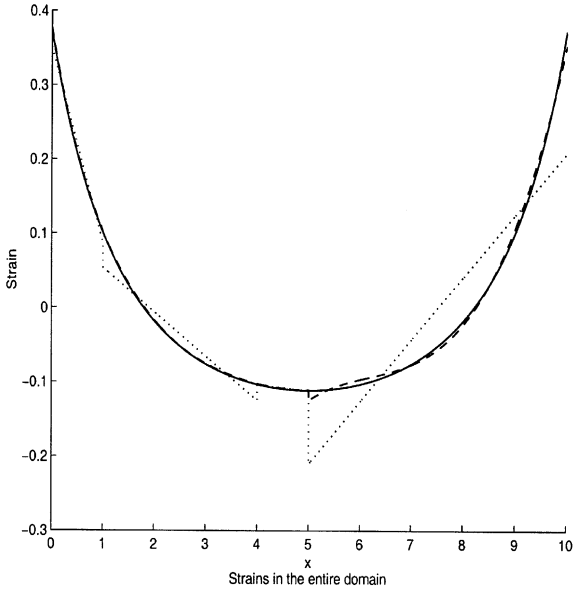

Fig. 7. Strains for Example 3. (—) $\mathrm{d}u/\mathrm{d}x$; (···) $\mathrm{d}u_h/\mathrm{d}x$; (- - -) $\mathrm{d}(u_h + e_h^*)/\mathrm{d}x$.

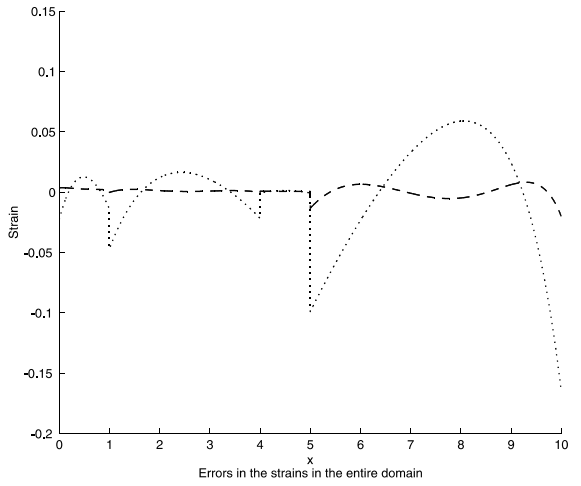

Fig. 8. Errors in the strain for Example 3. (···) $\mathrm{d}u_h/\mathrm{d}x - \mathrm{d}u/\mathrm{d}x$; (- - -) $\mathrm{d}(u_h + e_h^*)/\mathrm{d}x - \mathrm{d}u/\mathrm{d}x$.

interpretation allows to find similarities between implicit error-estimators and recovery-based error-estimators.

We have shown how these hidden displacement patterns can be used to design an a posteriori error estimator which is element-based and does not require the calculation of the strains at lower order Gauss points. Since the error estimator is element based, the estimator performs efficiently even in the case of non-uniform meshes. Also, since the hidden patterns are dependent on the material properties, the estimator performs well

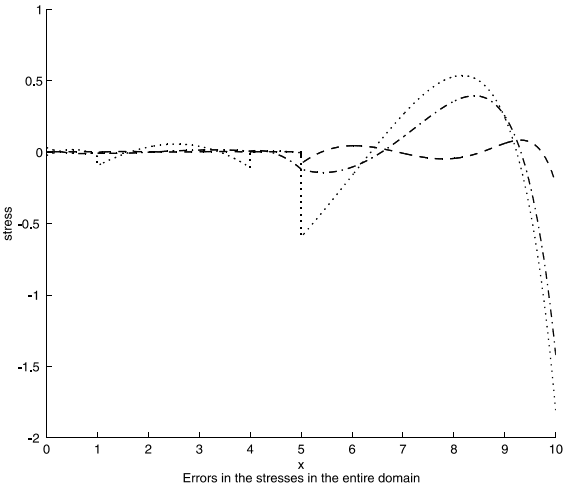

Fig. 9. Errors in the stress for example 3. (···) $E(x)\mathrm{d}(u_h)/\mathrm{d}x - \sigma$; (- - -) $E(x)\mathrm{d}(u_h + e_h^*)/\mathrm{d}x - \sigma$; (-·-·-) $\sigma_{ZZ} - \sigma$, where σ and σ_{ZZ} denote the exact stress and the stress recovered by the Zienkiewicz–Zhu algorithm, respectively.

even when the material properties vary not only from element to element but also inside an element. This last property is essential for an error estimator used in the solution of plasticity problems.

However, we considered in this paper primarily one-dimensional problems. For two- and three-dimensional problems, significant further work is necessary.

Acknowledgement

We thank I. Babuška of the University of Texas, Austin, for his valuable comments on this paper.

References

[1] Ainsworth M, Oden JT. A posteriori error estimation in finite element analysis. Comput Meth Appl Mech Engng 1997;142:1–88.

[2] Paraschivoiu M, Peraire J, Patera AT. A posteriori finite element bounds for linear-functional outputs of elliptic partial differential equations. Comput Meth Appl Mech Engng 1997;150:289–312.

[3] Zienkiewicz OC, Zhu JZ. The superconvergent patch recovery and a posteriori error estimates. Part 1: the recovery technique. Int J Numer Meth Engng 1992;33:1331–64.

[4] Zienkiewicz OC, Zhu JZ. The superconvergent patch recovery and a posteriori error estimates. Part 2: error estimates and adaptivity. Int J Numer Meth Engng 1992;33:1365–82.

[5] Bathe KJ. Finite element procedures. Englewood Cliffs, NJ: Prentice Hall; 1996.
[6] Krizek M, Neittaanmaki P, Stenberg R. Finite element methods, superconvergence, post-processing, and a posteriori estimates. New York: Marcel Dekker; 1998.
[7] Barlow J. Optimal stress locations in finite element models. Int J Numer Meth Engng 1976;10:243–51.
[8] MacNeal R. Comments on Barlow points and Gauss points and the aliasing and best fit paradigms. Comput Struct 1997;63:185.
[9] Prathap G. A variational basis for Barlow points. Comput Struct 1993;49:381–3.
[10] Prathap G. Barlow points and Gauss points and the aliasing and best fit paradigms. Comput Struct 1996;58:321–5.
[11] MacKinnon RG, Carey F. Superconvergent derivatives: a Taylor series analysis. Int J Numer Meth Engng 1989;28(3):489–509.
[12] Szabó B, Babuška I. Finite element analysis. New York: Wiley; 1991.
[13] Bergan PG. Finite elements based on energy orthogonal functions. Int J Numer Meth Engng 1980;15: 1541–55.
[14] Kreyszig E. Adv Engng Math. New York: Wiley; 1979.