SOME RECENT DEVELOPMENTS IN
FINITE ELEMENT ANALYSIS

B. A. SZABO
Greensfelder Professor of Civil Engineering, Washington University, St. Louis, Missouri, MO 63130,
U.S.A.

Communicated by I. N. Katz
(Received October 1978)

Abstract. A survey of some recent results in the development of adaptive finite element software systems
is presented. The points of view and terminology are those of an engineer-analyst and emphasis is on
applications to structural stress analysis.

1. INTRODUCTION
Finite element analysis is the most widely used numerical method for solving certain types of
differential equations of importance in a variety of engineering disciplines. The method is
especially important in structural stress analysis. Although it enjoys general acceptance, it still
has a large potential for further growth. There are substantial incentives for making the method
more efficient and more responsive to the technological requirements of current design practice.
In this paper some recent developments are surveyed which, in the opinion of the writer, will
influence, perhaps to a significant degree, the course of future development of the finite element
method. In order to clearly delineate the objective and scope of this paper, some definitions are
necessary.

First, we must distinguish between two convergence processes in finite element analysis. In
the conventional finite element convergence process, which we shall call \textit{h-convergence}, the
number and type of the interpolating (basis) functions over each element are fixed and the finite
element mesh is refined in such a way that the maximum diameter of the elements, \( h \),
approaches zero. In the second convergence process, to which we shall refer as \textit{p-convergence},
the number and distribution of finite elements is fixed and the number of basis functions, which
are necessarily locally complete polynomials of order \( p \), is progressively increased.

A finite element software system is said to be \textit{adaptive} when it possesses some local a
posteriori error estimation capability and a capability to increase, either automatically, or with
minimal user-interaction, the number of degrees of freedom in regions where an error tolerance
has been exceeded. Adaptive software systems should be designed to utilize the available
solutions in order to obtain improved solutions as efficiently as possible.

Adaptive finite element software systems can be based on \textit{h-convergence}, \textit{p-convergence},
or a combination of both, and on any type of finite element formulation such as displacement,
hybrid or mixed, for which convergence is guaranteed. Important considerations in the design
of adaptive systems are the rate of convergence and the efficiency of the procedure for obtaining
improved solutions.

Adaptiveness based on \textit{h} convergence has been examined by Babuska and Rheinboldt\cite{1-4}.\newpage
This paper is concerned with adaptive finite element software systems based on \textit{p-convergence}
and the displacement formulation. The development of such software systems has been made
possible by three recent developments.

1. Hierarchic families of finite elements, based on the displacement formulation, have been
developed.
2. It has been established that whenever \textit{h-convergence} occurs, \textit{p-convergence} will also
occur, (and vice versa) and the rate of \textit{p-convergence} cannot be slower than the rate of
\textit{h-convergence}, when \textit{h} convergence is based uniform or quasi-uniform mesh refinement.
3. Substantial computational experience has been accumulated which indicates that uniform
\textit{p-convergent} finite element approximations are superior in efficiency to uniform \textit{h-convergent

approximations. (In this context “efficiency” is meant to be measured in terms of both the computational and data preparation efforts involved in finite element analysis tasks).

Chronologically these developments did not occur in the sequence listed above. $p$-Convergent approximations were first studied numerically. Solutions were obtained for typical plate problems\cite{5,6}, plane elastic problems\cite{7}, stability problems\cite{8} and shell problems\cite{9}. It was found in all cases that $p$-convergent approximations required substantially fewer degrees of freedom to achieve a given level of precision than $h$-convergent approximations based on uniform mesh refinement did. Thus it seemed promising to pursue investigation of $p$-convergence in finite element analysis.

The first serious obstacle to computer implementation of $p$-convergent procedures for plate and shell problems was encountered in attempts to properly enforce $C^1$ continuity requirements. In order to achieve generality and preserve uniformity of convergence, it was considered necessary to enforce the appropriate continuity requirements exactly and minimally and to preserve the completeness of the polynomial approximating sequences. Thus overconformity, nonconformity and incompleteness, common undesirable characteristics of plate and shell finite elements, were to be avoided. The use of superelement techniques was considered but rejected mainly because they require cumbersome operations at the element level and make it difficult to efficiently control the number and distribution of the degrees of freedom. In work reported in Refs. \cite{5-9}, standard numerical and analysis procedures were employed to identify and remove redundant constraints which arise when minimal $C^1$ continuity is enforced. Subsequently linear programming techniques were developed and programmed\cite{10-13}. The name “constraint method” was used to refer to these procedures.

Drawing on the work of Dunne\cite{14} and Irons\cite{15}, Peano clarified the reasons for occurrence of redundancies in $C^1$ continuity constraints and proposed methods for eliminating them a priori\cite{16,34}. Incidental to this work, Peano determined the dimensionality of $C^1$ continuous polynomial approximation spaces for the class of triangulations normally used in engineering practice. The same problem was solved independently by Morgan and Scott\cite{17}.

Another very important development is also due to Peano\cite{16}. He generalized the idea of hierarchic approximations, first proposed by Zienkiewicz et al.\cite{18} in connection with joining finite elements of different polynomial order in $h$-convergent approximations, and developed shape functions for both $C^0$ and $C^1$ continuous finite elements\cite{16,19}. Hierarchic elements have the property that the shape functions corresponding to an element of order $p$ constitute a subset of the shape functions of all higher order hierarchic elements of the same kind. Thus the stiffness matrix of each hierarchic element is embedded in the stiffness matrices of all higher order hierarchic elements of the same kind. In $p$-convergent approximations this property is very important, because the computational labor spent on triangularizing the (initial) stiffness matrix can be saved in subsequent computations.

Numerical studies of typical problems in linear elastic fracture mechanics revealed that the strain energy converges strongly and monotonically as the polynomial order is increased on a fixed mesh even when the approximated function contains severe stress singularities. It was also shown that in such cases pointwise convergence of the stresses can be achieved by averaging\cite{20,21} and the rate of energy convergence, with respect to the number of degrees of freedom in the elements surrounding the crack tip, was established empirically\cite{22,23}. Petruska investigated the important question of whether $p$-convergence can be guaranteed in all cases in which $h$-convergence occurs. He concluded that, at least for problems requiring $C^0$ approximations, the question can be answered in the affirmative\cite{24}. More recently, Babuska has shown that the asymptotic rate of $p$-convergence in energy, with respect to the number of degrees of freedom, cannot be slower than the rate of $h$-convergence based on uniform or locally uniform mesh refinement\cite{25}. This confirms and generalizes the results presented in Refs. \cite{22-24}.

Two finite element computer codes, with adaptive features based on $p$-convergence, are now in existence. COMET-X, developed at Washington University, was designed as an experimental prototype for an advanced finite element software system\cite{26}. Another computer code, developed at the Istituto Sperimentale Modelli e Strutture (ISMES) in Bergamo, Italy has a self-adaptive capability based on $p$-convergence which was designed as a postprocessor for
Some recent developments in finite element analysis

An existing conventional finite element computer program [27, 28]. Some of the numerical results obtained by means of COMET-X are discussed in this paper.

A substantial amount of work has been completed on making the computation of elemental stiffness matrices and load vectors for arbitrary order polynomial approximations as efficient as possible. This can be accomplished with the use of precomputed arrays, i.e. arrays which are computed once, stored on permanent file, and then re-used in all subsequent applications. References [29, 30] present the development of precomputed arrays for two dimensional $C^0$ finite elements of arbitrary polynomial order and examples in which the relative efficiency of $h$ and $p$-convergent approximations are compared. It is concluded that $p$-convergent approximations are consistently superior in efficiency to $h$-convergent approximations based on uniform or quasi-uniform mesh refinement, even when the approximated function is not smooth. Other efforts to make the computation of elemental stiffness matrices more efficient were concerned with the integration of polynomials over elements with curved boundaries [31] and efficient integration of products of polynomials and rational functions over triangular domains [32].

In the following the key developments which make $p$-convergence not only a feasible but an attractive alternative to conventional finite element approximation procedures are presented and an outline for an adaptive finite element software system, based on $p$-convergence and the displacement formulation, is presented.

2. HIERARCHIC FINITE ELEMENTS

$p$-Convergent finite element approximations can be most efficiently implemented through the use of hierarchic families of finite elements. As mentioned in the introduction, hierarchic finite elements were first considered by Zienkiewicz et al. [18] in conjunction with joining finite elements of different polynomial order. The hierarchic elements described in Ref. [18] did not have nodal variables in the conventional sense. Rather, continuity was achieved by equating the amplitudes of corresponding quadratic, cubic and higher polynomial modes at interelement boundaries. In addition to this possibility, there is a unique set of nodal variables which lead to hierarchic $C^0$ and $C^1$ finite elements [16]. We shall briefly describe these families in the following.

2.1 The $C^0$ family

The first three members of the family are listed in Table 1. The first member is the well known “constant strain triangle”, the nodal variables of which are the displacements at the vertices and the corresponding shape functions, in triangular coordinates, are $L_1$, $L_2$, $L_3$. (References [16] and [19] present a clear interpretation of triangular coordinates.) The next member of the family is the “linear strain triangle” or the quadratic $C^0$ finite element. The nodal variables for this element are the displacements at the vertices and the second tangential derivatives at the midside points. Since all shape functions of the linear $C^0$ element have zero second derivatives, they can be retained as shape functions for the quadratic $C^0$ element. It is necessary only that the shape functions corresponding to the second derivatives of the quadratic element vanish at the vertices. This is indeed the case for the shape functions given in Table 1.

The third member of the family is the cubic $C^0$ finite element. The additional nodal variables are the third tangential derivatives at the midside nodes. There is also an internal mode, $N_{10}$, which vanishes along the edges of the triangle. We note that the shape functions corresponding to the cubic $C^0$ finite element are not unique; all shape functions can be modified by adding $N_{10} = L_1 L_2 L_3$ multiplied by an arbitrary constant. The constants can be chosen such that $N_{10}$ is energy-orthogonal to all other shape functions. The number and type of operations involved are precisely those of static condensation.

It is shown in Refs. [16, 19] that the nodal variables just described constitute a nodal basis for $C^0$ piecewise polynomials. Other choices of shape functions are possible also, those however do not correspond to a nodal basis. The question of which set of shape functions is optimal in terms of numerical stability has not been investigated, but the shape functions presented here have been numerically stable within the range of $p$ investigated (1 through 8). Also, the shape functions comprise the minimum number of monomial terms in area coordinates. This is advantageous from the point of view of computational economy.

Given the hierarchic nodal basis and the corresponding shape functions for triangles, a straightforward extension for tetrahedra is possible: For linear approximation the first four...
Table 1. Nodal variables and shape functions for the first three hierarchic triangular $C^0$ elements

1. Linear

3 terms: Values of the approximating polynomial at the vertices.

Normalizing factor: 1

\[ N_1 = L_1, \quad N_2 = L_2, \quad N_3 = L_3 \]

2. Quadratic

3 additional terms: Second derivatives of the approximating polynomial at the vertices.

Normalizing factor: $-1/2$

\[ N_4 = L_1 L_2, \quad N_5 = L_2 L_3, \quad N_6 = L_3 L_1 \]

3. Cubic

4 additional terms: Three third derivatives of the approximating polynomial at the midpoints of edges.

Normalizing factor: $1/12$

\[ N_7 = L_1^3 L_2 - L_1 L_2^3, \quad N_8 = L_2^3 L_3 - L_2 L_3^3, \quad N_9 = L_3^3 L_1 - L_3 L_1^3 \]

One internal mode:

\[ N_{10} = L_1 L_2 L_3 \]

shape functions are simply $I_1, I_2, I_3, I_4$, where $I_4$ now denote tetrahedral coordinates. For quadratic approximation we again use the second derivatives along the edges as nodal variables. The corresponding (normalized) shape functions are: $-1/2L_1 L_2, -1/2L_2 L_3, \ldots$, a total of six terms (Table 2). For the cubic tetrahedral elements we add the edgewise third derivatives. The corresponding shape functions are: $1/12(L_i^3 L_j - L_i L_j^3) \ldots$ etc. In addition, four face modes appear: $L_1 L_2 L_3, L_1 L_2 L_4, L_1 L_3 L_4, L_2 L_3 L_4$. The face modes need not be associated with specific nodal variables, however care must be taken in assembling the elemental stiffness matrices to ensure that the corresponding modes are scaled correctly.

The quartic tetrahedral element has a total of 35 displacement modes. 20 of these are identical with the displacement modes of the cubic tetrahedral element, 6 correspond to the edgewise fourth derivatives. For example, for edge 1–2 the shape function corresponding to the fourth derivative is: $-1/48(L_i^3 L_j^2 + L_i L_j^3)$. There are 8 additional face modes: $L_i^2 L_j L_k, L_i L_j^2 L_k, L_i^2 L_j L_k, L_i L_j^2 L_k, L_i L_j L_k^2, L_i L_j L_k^2, L_i^2 L_j L_k, L_i L_j^2 L_k, L_i^2 L_j L_k$, and there is an internal node: $L_1 L_2 L_3 L_4$. These results are summarized in Table 2.

The proof that these shape functions constitute a three-dimensional $C^0$ piecewise polynomial basis is similar to the proof given in[16, 19] for two-dimensional $C^0$ piecewise polynomials on the triangle. In particular, we note that the polynomial basis of order $p$ contains $n = (p + 1)(p + 2)(p + 3)/6$ terms. The same number of monomial terms are given in Table 1 for the hierarchic tetrahedral elements of order $p$. The order of polynomials representing the shape functions is always less than or equal to $p$. Finally, we observe that the shape functions are linearly independent: For example, letting $L_4 = 0$, we have the same shape functions as for the two-dimensional elements which are known to be linearly independent[16, 19]. The same observations can be made by letting $L_3 = 0$, however we note that the shape functions pertaining to edge 1–2 are contained in both sets and therefore must not be counted again. The rest of the shape functions vanish over the face $L_4 = 0$ and are therefore linearly independent from the first set. With similar arguments it can be shown that all shape functions pertaining to the vertices, edges and faces of the tetrahedron are linearly independent. The internal mode $L_1 L_2 L_3 L_4$ vanishes over the vertices, edges and faces and is therefore linearly independent from the other shape functions. Similarly, hierarchic shape functions for prismatic $C^0$ elements with triangular, rectangular and other polygonal cross-sections can be developed also.
Table 2. Nodal variables and shape functions for the first four hierarchic tetrahedral $C^0$ elements

1. Linear

<table>
<thead>
<tr>
<th>Terms</th>
<th>Description</th>
<th>Shape Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Values of the approximating function at the vertices.</td>
<td>$N_1 = L_1$, $N_2 = L_2$, $N_3 = L_3$, $N_4 = L_4$</td>
</tr>
</tbody>
</table>

2. Quadratic

<table>
<thead>
<tr>
<th>Terms</th>
<th>Description</th>
<th>Shape Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Second derivatives at the midpoints of edges.</td>
<td>$N_5 = L_1 L_2$, $N_6 = L_3 L_4$, ..., $N_{10} = L_1 L_4$</td>
</tr>
</tbody>
</table>

3. Cubic

<table>
<thead>
<tr>
<th>Terms</th>
<th>Description</th>
<th>Shape Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Six third derivatives at the midpoints of edges.</td>
<td>$N_{11} = L_1^3 L_2 - L_1 L_2^2$, $N_{12} = L_2^3 L_3 - L_2 L_3^2$, ..., $N_{16} = L_1^2 L_4 - L_1 L_4^2$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Face Modes</th>
<th>Shape Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>$N_{17} = L_1 L_2 L_3$, $N_{18} = L_1 L_2 L_4$, $N_{19} = L_1 L_3 L_4$, $N_{20} = L_2 L_3 L_4$</td>
</tr>
</tbody>
</table>

4. Quartic

<table>
<thead>
<tr>
<th>Terms</th>
<th>Description</th>
<th>Shape Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>Six fourth derivatives at the midpoints of edges.</td>
<td>$N_{21} = L_1^4 L_2 + L_1 L_2^3$, $N_{22} = L_2^4 L_3 + L_2 L_3^3$, ..., $N_{26} = L_1^3 L_4 + L_1 L_4^3$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Face Modes</th>
<th>Shape Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eight</td>
<td>$N_{27} = L_1^2 L_2 L_3$, $N_{28} = L_1 L_2 L_4$, $N_{29} = L_2^2 L_3$, $N_{30} = L_1 L_2 L_4$, $N_{31} = L_2 L_3 L_4$, $N_{32} = L_1 L_3 L_4$, $N_{33} = L_2 L_3 L_4$, $N_{34} = L_2 L_3 L_4$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Internal Mode</th>
<th>Shape Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>$N_{35} = L_1 L_2 L_3 L_4$</td>
</tr>
</tbody>
</table>

2.2 The $C^1$ family

The subject of hierarchic $C^1$ elements is best approached by first examining two complete, exactly and minimally conforming non-hierarchic finite elements which preceded the development of the hierarchic $C^1$ finite elements. Such developments were clearly motivated by the desire to avoid problems caused by overconforming elements in approximating certain boundary and connectivity conditions such as those that occur at reentrant corners, sudden changes in plate thickness, shell intersections, etc.

A very important exactly conforming finite element based on complete polynomials was developed by Irons[15]. A schematic representation of the nodal variables is given in Fig. 1(a). Although proof is not provided in Ref. [15], it is not difficult to show that the shape functions proposed by Irons are indeed equivalent to complete fourth order polynomial approximation supplemented by three rational functions. Rational functions were first employed by Baseley et al.[33], who named these functions $e_1$, $e_2$, $e_3$. Irons[15] introduced three additional rational functions named $\delta_1$, $\delta_2$, $\delta_3$. Peano[16, 19] had shown that the functions can be obtained from
simple combinations of a new set of rational functions, called $\eta_1, \eta_2, \ldots, \eta_6$:

$$
\eta_1 = \frac{L_1^2 L_2 L_3}{L_1 + L_3}, \quad \eta_2 = \frac{L_1^2 L_2 L_3^2}{L_2 + L_3}, \quad \eta_3 = \frac{L_1 L_2^2 L_3}{L_1 + L_3},
$$

$$
\eta_4 = \frac{L_1^2 L_2^2 L_3}{L_1 + L_3}, \quad \eta_5 = \frac{L_1 L_2^2 L_3^2}{L_2 + L_3}, \quad \eta_6 = \frac{L_1^2 L_2 L_3^2}{L_2 + L_3}.
$$

The sums $(\eta_1 + \eta_2), (\eta_3 + \eta_4), (\eta_5 + \eta_6)$ yield three quartic monomial terms. The $\epsilon$'s are given by:

$$
\epsilon_1 = \eta_3 + \eta_5, \quad \epsilon_2 = \eta_2 + \eta_6, \quad \epsilon_3 = \eta_1 + \eta_4
$$

and the $\delta$'s are given by [32]:

$$
\delta_1 = \eta_3 - \eta_5, \quad \delta_2 = \eta_2 - \eta_6, \quad \delta_3 = \eta_1 - \eta_4.
$$

Thus the three independent rational functions which supplement the complete quartic polynomial in Irons' quartic element are: $(\eta_1 - \eta_2), (\eta_3 - \eta_4), (\eta_5 - \eta_6)$. The $\eta$ functions have further significance: they are the shape functions associated with the mixed second partial derivatives, taken along the sides and normal to the sides of a triangle and evaluated at the vertices [16, 19, 34].

The purpose in introducing rational functions is to destroy the analyticity of the (polynomial) approximant at the vertices and thereby to make possible proper enforcement of the continuity of the first derivatives. To clarify this point, we note that $C^1$ continuity of a function $w(x, y)$ implies continuity of the second derivatives: $\partial^2 w/\partial s_1^2$, $\partial^2 w/\partial s_2^2$, $\partial^2 w/\partial s_3^2$, $\partial^2 w/\partial s_1 \partial n_1$, $\partial^2 w/\partial s_2 \partial n_2$ (Fig. 2). These are not, however, linearly independent when the second derivatives of $w$ are single valued. In particular, the following relationship exists:

$$
\frac{\partial^2 w}{\partial s_1^2} \cos \phi + \frac{\partial^2 w}{\partial s_1 \partial n_1} \sin \phi = \frac{\partial^2 w}{\partial s_2^2} \cos \phi - \frac{\partial^2 w}{\partial s_2 \partial n_1} \sin \phi . \tag{1}
$$

Because of this relationship, any set of nodal variables which enforce $C^1$ continuity of polynomial functions over polygonal domains are linearly dependent. The number of linear dependencies is the same as the number of vertices in the polygon. On the other hand, when the approximant contains functions whose second derivatives are multivalued at the vertices, such as the $\eta, \epsilon$ and $\delta$ functions, the relationship given in eqn (1) no longer holds, and the nodal variables are independent. A somewhat undesirable consequence of this is that certain physical

Fig. 1. Nodal bases for complete and conforming $C^1$ finite elements. (a), Complete quartic, supplemented by three rational functions; (b), Complete quintic. Constraints on the second derivatives must be enforced in the assembly process; (c), Complete quintic, supplemented by three rational functions.
quantities of interest, such as shear forces and bending moments, are in general multi-valued at vertices, even when the approximated function is analytic, and must therefore be evaluated by averaging. Simple explicit formulas are now available for calculating the average values of second derivatives [35].

The second precursor to hierarchic $C^1$ finite elements is illustrated in Fig. 1(b). This element, proposed by Morgan and Scott [36], is based on the complete polynomial of order 5 and is in fact the first member of a family of higher order $C^1$ elements. There are 18 degrees of freedom. Only three second derivatives are specified at the vertices, which are evaluated along the sides of the triangle. It is proven in [36] that the shape functions associated with the nodal variables determine a $C^1$ polynomial basis over arbitrary triangulation. Computer implementation of this element poses an unusual problem in assembling the structural stiffness matrix, however: the constraints representing relationships among second derivative nodal variables at vertices must be enforced in the assembly process. This can be done without incurring redundant equations only if the number of independent polynomial coefficients (i.e. the dimension) of $C^1$ polynomial spaces over those types of triangulation which are used in finite element analysis can be determined. As recently as in 1973, this was an open question. For this reason, computational studies of exactly and minimally conforming $C^1$ elements necessarily involved the identification of redundant constraints by numerical rank analysis [5, 6, 8–11, 37]. The problem of dimensionality of piecewise polynomial spaces over triangular subdivisions was first posed formally by Strang [38], who offered a conjecture for the dimension of $C^1$ piecewise polynomial cubics over arbitrary triangulation. The problem was solved independently for $p \geq 5$ (where $p$ is the polynomial order) by Morgan and Scott [36, 37] for all triangulations of practical importance. A very significant contribution has been the identification of singular vertices. A brief summary follows.

Let us consider a vertex at which two or more triangles meet (Fig. 3). The relationships among second derivatives is:

$$\frac{\partial^2 w}{\partial s_i^2} (\cot \phi_{i-1} + \cot \gamma_i) = \frac{1}{\sin \phi_{i-1}} \frac{\partial^2 w}{\partial s_{i-1} \partial s_i} + \frac{1}{\sin \phi_i \sin \phi_{i-1}} \frac{\partial^2 w}{\partial s_i \partial s_{i+1}}$$

(2)
This constraint must be enforced in the process of assembling the structural stiffness matrix when the nodal basis proposed by Morgan and Scott is employed. A detailed discussion of this point can be found in [16, 34]. A singular vertex occurs when four edges meet at a vertex and $\phi_1 = \phi_3, \phi_2 = \phi_4$ (Fig. 5). In this case $\cot \phi_{1-1} + \cot \phi_1 = 0$ and the relationship, given in eqn (2), enforced around the vertex, results in four homogeneous algebraic equations among the mixed second derivatives, one of which is redundant. In order to avoid such redundancies, one of the four constraints must not be enforced when a singular vertex is encountered.

The first member of the hierarchic family of $C^1$ elements is illustrated in Fig. 1(c). The element has 24 degrees of freedom, and its basis is equivalent to the complete fifth-order polynomial supplemented by the same three rational functions as Irons’ quartic element[15]. For full description of this element, the reader is referred to the original expositions[16, 19, 34]. Shape functions are available for higher-order hierarchic $C^1$ elements as well.

As a practical matter, the importance of $C^1$ plate and shell elements has decreased considerably with the introduction of numerically stable and efficient $C^0$ plate bending elements. Hierarchic $C^0$ plate bending elements have been implemented in the experimental computer program, COMET-X[26]. Computational studies have shown the elements to be stable at least over the range of polynomial orders 1–8 permitted by COMET-X.

3. RATE OF $p$-CONVERGENCE

Important advances have been made in the understanding of the $p$-convergence process in finite element approximation of elliptical boundary value problems. The results, briefly summarized in the following, have practical significance in that they establish $p$-convergence as a feasible and attractive method for achieving improved accuracy in adaptive finite element computations and provide an independent method for computing certain quantities of interest.

The rapid convergence in energy as well as pointwise convergence of finite element approximations employing high order finite elements, when the approximated function is smooth, has been recognized by several investigators[5, 7, 9, 40–41]. In practical problems, however, the approximated function is very seldom smooth. In the large majority of cases, the lack of smoothness is due to angular corners at external boundaries. In two dimensional elasticity, for example, the displacement vector components at corners are of the form:

$$u_i = r^n g_i(\theta) + G_i(r, \theta)$$

in which $i = 1, 2$; $r$ and $\theta$ are polar coordinates with the origin at the corner; $n > 1/4$, its value depends on the corner angle and the boundary conditions imposed at the corner; $g_i$ and $G_i$ are smoother functions than $u_i$ in the neighborhood of $r = 0$. The values of $n$ for plane elastic and plate problems are given in [42].

Singularities, such as a corner singularity, perturb $u_n$, the polynomial approximation to $u_n$, and the computed stresses, which are linear combinations of the derivatives of $u_n$, oscillate in the elements containing the source of perturbation. It has been observed that this oscillation is...
similar to the oscillation of Fourier series in the vicinity of jump discontinuities, in the sense that the average of successive partial sums (the first Cesaro sums) of the computed stresses approximate the true stress distribution with increasing precision as the order of polynomial approximation is increased\[21,23,43\]. This phenomenon has not been studied in detail and it certainly deserves further attention. However, obtaining close pointwise stress approximation in the vicinity of singularities is not necessary in stress analysis. Of primary importance is that the true energy density distribution be closely approximated in the finite element solution and the singularity should not significantly perturb the displacement and stress approximations in elements away from the crack tip. It has been demonstrated that rapid convergence in energy occurs when the polynomial displacement approximation is progressively increased on a fixed finite element mesh\[22,23,43\] and the stress oscillations caused by singularities are confined to the elements in the immediate vicinity of the singularities\[23,43\].

From the point of view of designing an adaptive finite element software system, it is important to consider the rate of (energy) convergence. When the number of degrees of freedom, $N$, is increased through uniform (or locally uniform) mesh refinement then the energy error is bounded by:

$$U - \hat{U} \leq \frac{c}{N^{\nu^p}}$$  \hspace{1cm} (4)

in which $U$ is the exact strain energy; $\hat{U}$ is the computed strain energy; $c$ is a constant which depends on the order $p$ of the polynomial approximation, the value of the elastic constants and the geometry of the mesh, $\nu^p = \min(\nu, p)$\[44\].

In contrast, when the number of degrees of freedom is increased through increasing the order of polynomial approximation, the error in strain energy is bounded by:

$$U - \hat{U} \leq \frac{k(\epsilon)}{N^{2\nu - \epsilon}}, \hspace{1cm} \epsilon > 0 \text{ arbitrary}$$  \hspace{1cm} (5)

in which $k$ is a constant which depends only on the elastic constants, the geometry of the mesh and the distribution of polynomial orders over the mesh. The rate of convergence relationship expressed in eqn (5) was discovered in numerical experiments\[22\] and subsequently verified and generalized theoretically\[25\]. Thus, it has been established that the rate of $p$-convergence is twice the rate of $h$-convergence when corner singularities are present and the number of degrees of freedom are increased by uniform or quasi-uniform mesh refinement in the $h$-convergence process.

Numerical studies have indicated that the energy error is closely approximated by the bound given in eqn (5), even when the number of degrees of freedom is small. This is illustrated by an example, involving the centrally cracked square panel shown in Fig. 5(a). Due to symmetry, only one quarter of the panel had to be analyzed. Two finite element meshes were used; the eight-element mesh (Fig. 5b) and the three-element mesh (Fig. 5c) which comprises the minimum number of elements needed for describing the geometry for this problem. In the eight element mesh the polynomial orders of the displacement approximation were distributed in two ways: uniformly and non-uniformly. In the non-uniform or graded distribution the polynomial orders were increased over $p = 3$ only in the crack tip elements (numbered 1, 3, 4); the polynomial order was constant at $p = 3$ in the remote elements (numbered 5, 6, 8) and the transition elements (numbered 2, 7). However, in order to preserve completeness of the polynomial approximating functions in the crack tip elements, as well as $C^0$ continuity, it was necessary to add higher order shape functions to the transition elements corresponding to the higher order hierarchic nodal variables at their interelement boundaries with the crack tip elements.

For this problem $\nu = 1/2$. Therefore, the computed strain energy was plotted against $N^{-1}$ in Fig. 6. The convergence paths are seen to be very nearly linear for all $p \geq 4$. For the eight-element mesh, the pre-asymptotic rate of convergence is slightly slower than that predicted by eqn (5), when uniform $p$-distribution is used and slightly faster when graded $p$ distribution is used.
Fig. 6. Centrally cracked panel. Computed strain energy (quarter panel only) for varying polynomial orders (p) vs reciprocal of the net degrees of freedom. \(E\), modulus of elasticity; \(t\), Panel thickness.

The bound given by eqn (5) is sufficiently close within the range of mesh divisions and polynomial orders that can be employed in practical computations to permit estimation of the exact strain energy by extrapolation to well within one per cent of accuracy. For example, let us define \(\bar{U}_p\) as an estimate of \(U\) based on eqn (5):

\[
\bar{U}_p = \frac{N_p^{2p} \hat{U}_p - N_{p-1}^{2p} \hat{U}_{p-1}}{N_p^{2p} - N_{p-1}^{2p}}
\]

in which \(N_p\) represents the (net) number of degrees of freedom corresponding to \(p\)th order polynomial approximation in the elements at the crack tip. \(\bar{U}_p\) is plotted against \(p\) in Fig. 7 for the centrally cracked panel problem. It is seen that \(\bar{U}_p\) tends to 0.7702\(\sigma^2t^4/E\) and the error is probably not more than 0.0002\(\sigma^2t^4/E\) or 0.03%. The exact solution for this problem is not known.

This rapid and regular energy convergence can be utilized for finding the stress intensity factor in linear elastic fracture mechanics very efficiently [22, 23, 43].

4. ADAPTIVE CONTROL

One of the key problems in developing adaptive finite element software systems is to find suitable estimators of error to indicate whether and where the distribution of degrees of freedom should be modified over the solution domain. Some estimators have been proposed already:

(a) Babuska and Rheinboldt developed a local, asymptotic, \textit{a posteriori} error estimator for
Some recent developments in finite element analysis

109

The estimator requires element level computations only and
measures the error in strain energy associated with an element. An optimal distribution of the
degrees of freedom is obtained when this error measure is the same for all elements.

(b) Melosh and Marcal[45] proposed to measure the specific energy difference (SED),
defined as the largest difference over the element domain between the computed strain energy
density function and the same function evaluated at the origin of the elemental coordinate
system. This measures the effect of modes, higher than those associated with the (generalized)
constant strain states, on the distribution of strain energy density within finite elements. The
criterion for mesh refinement is that SED be approximately the same for each finite element. In
practical computations SED is approximated by the largest of the strain energy density
differences evaluated at quadrature points only.

(c) Peano et al.[28, 46] proposed a criterion for p-convergent approximations. This criterion
is based on the rate of change of the total potential energy with respect to higher order
displacement modes, evaluated before the higher displacement modes are actually introduced
When the rate of change of the potential energy exceeds a prescribed tolerance, the stiffness
terms corresponding to the higher modes are assembled and the new system or equations is
solved. The hierarchic structure of the elemental stiffness matrices permits efficient use of
block relaxation procedures in obtaining improved solutions.

(d) In Refs. [47, 48], investigation of error estimators, similar to the estimators proposed by
Babuška and Rheinboldt[4] but modified for p-convergent approximations, was reported. The
results of these investigations are reviewed in the following:

Two estimators were defined. One, called the r-estimator, is a function of the unbalanced
body forces within individual elements:

\[ R_k(\alpha) = \frac{A_k}{p_k^2 E} \int_{A_k} \delta_{ij} r_i W_k (t \, dA) \]  \hspace{1cm} (7)

in which: \( A_k \) is the area of the kth element; \( p_k \) is the polynomial order of the displacement
approximation over the kth element; \( \alpha \) is a constant, determined by numerical experiments; \( E \)
is the modulus of elasticity; \( \delta_{ij} \) is the Kronecker delta; \( r_i \) is the unbalanced body force vector
component: \( r_i = G \hat{u}_{ij} + (\lambda + G) \hat{u}_{ij} + X_i \). \( G \) and \( \lambda \) are Lamé's parameters; \( \hat{u}_i \) is the ith com-
ponent of the displacement vector computed by the finite element method; \( X_i \) is the imposed
body force; \( W_k \) is a weight function; and \( t \) is the panel thickness.

The other estimator, called the t-estimator, is a function of the unbalanced tractions at
interelement boundaries:

\[ T_k(\beta) = \frac{\sqrt{A_k}}{p_k^2 E} \int_{t_{ij}} \delta_{ij} t_i \omega_k (t \, ds) \]  \hspace{1cm} (8)

in which: \( \beta \) is a constant, determined by numerical experiments; \( \Gamma_k \) denotes the kth element
boundary; \( t_i \) is the vector of unbalanced tractions: \( t_i = [\hat{\sigma}_{ij}^{(a)}(s) - \hat{\sigma}_{ij}^{(b)}(s)]n_j \) \( (s \) is the coordinate
along the interelement boundary. \( \hat{\sigma}_{ij}^{(a)} \) is the finite element solution for the stress tensor for the
a\textsuperscript{th} element, \( n_j \) is the unit normal to the interelement boundary between elements \( a \) and \( b \); \( \omega_k \)
is a weight function.

The typical variation of \( t_i \) with increasing polynomial orders along interelement boundaries
is illustrated in Figs. 8 and 9. In Fig. 8 the variation of \( t_2 \) at \( x_2 = l/2 \) is shown for the centrally
cracked square panel discussed in the preceding section and shown in Fig. 5. Uniform
\emph{p}-distribution was employed. \( t_2 \) attenuates rather slowly, however the signs alternate for even
and odd \( p \) values indicating that the rate of attenuation of \( t_2 \) can be increased by averaging.

The variation of \( t_2 \) along \( x_1 = l/2 \) is shown in Fig. 9. The attenuation of \( t_2 \) between elements 5
and 6 is so rapid that only the value for \( p = 3 \) can be shown at the scale of this diagram. There
is no appreciable attenuation on the other hand between the crack tip elements. Successive
averaging smoothens \( t_2 \) and shifts the peak progressively closer to the crack tip.

The computed dimensionless strain energy values, the corresponding errors and the values
of the two error estimators for parameter values \( \alpha = 3, \beta = 2 \) and weight function values
\( W_k = \omega_k = 1 \) are given in Table 3.
The per cent changes in the estimators as $p$ is increased tends to bound the corresponding per cent change in the strain energy error with increasing precision such that the change in the $r$-estimator is smaller, the change in the $t$-estimator is larger than that of the strain energy error. Furthermore, the values of these percentage changes are monotonically decreasing for $p > 2$. This suggests the possibility that two constants, $c_r$ and $c_t$, could be found such that a relationship:

$$c_r \sum R_k(3) \leq U - \hat{U}_p \leq c_t \sum T_k(2)$$

remains valid for all $p$-values and that the upper and lower bounds become progressively closer
Some recent developments in finite element analysis

Table 3. Estimators for the error in strain energy. Centrally cracked panel, 8-element mesh, uniform p-distribution.

<table>
<thead>
<tr>
<th>p</th>
<th>( \hat{U}_p )</th>
<th>( U - \hat{U}_p )</th>
<th>( \sum R_k(3) )</th>
<th>( \sum T_k(2) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.6936</td>
<td>0.0766</td>
<td>0.01080</td>
<td>0.0305</td>
</tr>
<tr>
<td>3</td>
<td>0.7305</td>
<td>0.0397</td>
<td>0.00668</td>
<td>0.0189</td>
</tr>
<tr>
<td>4</td>
<td>0.7461</td>
<td>0.0241</td>
<td>0.00401</td>
<td>0.0113</td>
</tr>
<tr>
<td>5</td>
<td>0.7538</td>
<td>0.0164</td>
<td>0.00273</td>
<td>0.0077</td>
</tr>
<tr>
<td>6</td>
<td>0.7584</td>
<td>0.0118</td>
<td>0.00201</td>
<td>0.0057</td>
</tr>
<tr>
<td>7</td>
<td>0.7613</td>
<td>0.0089</td>
<td>0.00158</td>
<td>0.0045</td>
</tr>
</tbody>
</table>

*To convert entries into dimensioned values, multiply by \( \sigma t^4 / E \). Poisson’s ratio: 0.3.*

to the strain energy error as \( p \) is increased. This is, of course, highly speculative at the present because no theoretical justification exists, but consistent with the observations of this numerical experiment. For example, if we choose \( c_r = 5.633 \) and \( c_t = 0.748 \), the two estimators will give the value of the strain energy error at \( p = 7 \). The resulting relationship between the strain energy error and the two estimators is illustrated in Fig. 10.

Similar results were obtained in other computational experiments in which the strength of corner singularity, as characterized by \( \nu \), was varied[48]. Thus, the indications are that the \( r \) and \( t \) estimators give respective lower and upper bounds for the strain energy error regardless of the strength of corner singularity.

The question naturally arises whether the same estimators would bound the energy error at the element level as well. Clearly, this approach will be useful only if the indicators tend to zero with increasing \( p \) at about the same rate as the error in strain energy does, not only for the entire solution domain but for individual finite elements as well. Study of the local error in the energy norm is complicated by the fact that the local error can be computed only if the exact solution is known. Studies concerned with the local behavior of the estimators are in progress. The available results indicate that the element to element variation of the \( r \) and \( t \) estimators conforms closely with the probable error variation in strain energy, indicating that the \( r \) and \( t \)

Fig. 10. Centrally cracked square panel. Energy error and energy error estimators vs reciprocal of the net number of degrees of freedom. The estimators were multiplied by suitable scaling constants.
estimators or similar estimators may provide a suitable basis for adaptive redistribution of the
degrees of freedom in finite element analysis.

5. ADAPTIVE FINITE ELEMENT SOFTWARE SYSTEM BASED ON
p-CONVERGENCE

The contributions described in this paper, particularly the development of hierarchic finite
elements, the proofs concerning the occurrence and rate of p convergence and the indications
that readily computable local a-posteriori error estimators can be developed in the near future,
make it possible to develop adaptive finite element computer programs based on p-conver-
genence. Let us briefly consider the architecture and advantages of such a software system. A
schematic representation is shown in Fig. 11.

From the user's point of view, the system represented by the upper part of the diagram
differs from conventional finite element capabilities only in two respects: the number of finite
elements generated in the preprocessor does not have to be greater than the number of
elements needed to define the geometry of the structure and the orders of polynomial
approximation (as well as optional use of a number of non-polynomial basis functions) may be
controlled over each finite element. Internally, this system differs from conventional systems in
that it contains hierarchic elements; the elemental stiffness matrices and load vectors are
generated through extensive use of precomputed arrays; the data needed for successive
solution steps are saved on the SAVE FILE and it computes and displays estimators indicating
the error distribution over the finite element mesh.

The system represented by the lower part of the diagram in Fig. 11 implements the
p-convergence procedure. Having obtained a solution and an indication of the error distri-
bution, the user may choose to increase the orders of polynomial approximation either
uniformly or selectively over the finite elements and/or introduce non-polynomial basis func-
tions. The secondary solver utilizes results from the previous solution step and computes the
error distribution using the procedures employed by the primary solver. In addition to providing
the capability for interactive looping through the secondary solver, it is also possible to develop
an automated capability, called the "self-adaptive module" which would serve to update the
input files to the secondary solver until some prespecified conditions, such as uniformity of
error distribution or expended CPU time are met.

Optimally, the preprocessor, postprocessor and update modules would be installed in a
minicomputer and operated in the interactive mode whereas the solvers and the self-adaptive
module would be installed on a large computer and operated in the batch mode. It may become
possible to develop secondary solvers which will execute efficiently in a mini-computer
environment. This would further reduce the cost of analysis and improve turn-around time,
which are important considerations in practical analysis.

Fig. 11. Architecture of an adaptive finite element software system based on p-convergence.
Development of an adaptive finite element software system based on \( p \)-convergence yields the following significant advantages:

(a) A very large finite element library is available. Elements of any polynomial order can be used and the polynomials can be supplemented with non-polynomial basis functions.

(b) The number of finite elements is determined by the geometry rather than the considerations of accuracy.

(c) The distribution and number of degrees of freedom can be modified with minimal interactive and computational effort. Mesh refinement is not involved and the effort expended in triangularizing the stiffness matrix is saved for successive analyses (Hierarchical property).

(d) Very rapid energy convergence is obtained when the orders of polynomial approximation are increased on a fixed mesh even when the approximated function is not smooth, provided that discontinuities in loading, angular corners in boundaries and, in general, the sources of perturbation are located at interelement boundaries or vertices. This restriction is an acceptable one in most practical analyses.

(e) \( p \)-Convergence provides an efficient new method for computing stress intensity factors in linear elastic fracture mechanics.

(f) Based on the available information, it is estimated that in comparison with state-of-the-art methods, an order of magnitude reduction can be realized in the overall cost and time required for performing stress analysis tasks.

6. SOME UNSOLVED PROBLEMS

Until quite recently, convergence in finite element approximation was generally understood to mean \( h \)-convergence only. In the three fundamental aspects of the method: conditions for convergence to occur; asymptotic estimates for the rate of convergence and local, a-posteriori estimates of error distribution in various norms, a penetrating and elaborate mathematical support exists for \( h \)-convergent approximations. Adequate mathematical support now exists for \( p \)-convergent approximations in the first two areas although some open questions still remain. In the third area, initial studies have indicated that readily computable \textit{a posteriori} error bounds that approach zero at about the same rate as the total error in strain energy exist but theoretical support is still lacking. The relationship between the local error and local error indicators is yet to be explored. This is perhaps the most important uncharted area in the theory of the finite element method.

There are other challenging problems as well: one concerns obtaining pointwise stress estimates in the neighborhood of singularities. It is known that averaging the stress solutions that result from progressively increasing orders of polynomial approximation on a fixed finite element mesh can smoothen the stress oscillations caused by singularities but techniques for optimal stress smoothing have not yet been developed.

Development of efficient secondary solvers is yet another unsolved problem. Block relaxation procedures have been employed and the results were found to be quite satisfactory: only very few iterations were needed\([27, 28]\). This subject deserves a systematic study, however, with particular reference to developing secondary solvers designed to operate on minicomputers or in small partitions of large timesharing systems.

Extension of the convergence rate studies to three-dimensional problems in elasticity, particularly for studying problems posed by linear elastic fracture mechanics, is yet another promising area for research. Computational as well as theoretical studies are needed. Work in this area is currently in progress. For example, hierarchic shape functions of prismatic finite elements have been developed and implemented\([49, 50]\).

Application of adaptive finite element approximation methods based on \( p \)-convergence to non-linear problems is a completely unexplored area. The advantages of having a capability to adjust the distribution of degrees of freedom without mesh refinement and dealing with substantially fewer variables that require global operations, are likely to outweigh the disadvantage of having to employ high order quadrature rules in generating elemental stiffness matrices. A particularly important computational problem in non-linear mechanics is posed by recent developments in the understanding of ductile fracture processes, which require evaluation of two energy related quantities, known as the "\( J \)-integral" and the "tearing modulus", \( installation of a new method for computing stress intensity factors in linear elastic fracture mechanics.

(f) Based on the available information, it is estimated that in comparison with state-of-the-art methods, an order of magnitude reduction can be realized in the overall cost and time required for performing stress analysis tasks.

6. SOME UNSOLVED PROBLEMS

Until quite recently, convergence in finite element approximation was generally understood to mean \( h \)-convergence only. In the three fundamental aspects of the method: conditions for convergence to occur; asymptotic estimates for the rate of convergence and local, a-posteriori estimates of error distribution in various norms, a penetrating and elaborate mathematical support exists for \( h \)-convergent approximations. Adequate mathematical support now exists for \( p \)-convergent approximations in the first two areas although some open questions still remain. In the third area, initial studies have indicated that readily computable \textit{a posteriori} error bounds that approach zero at about the same rate as the total error in strain energy exist but theoretical support is still lacking. The relationship between the local error and local error indicators is yet to be explored. This is perhaps the most important uncharted area in the theory of the finite element method.

There are other challenging problems as well: one concerns obtaining pointwise stress estimates in the neighborhood of singularities. It is known that averaging the stress solutions that result from progressively increasing orders of polynomial approximation on a fixed finite element mesh can smoothen the stress oscillations caused by singularities but techniques for optimal stress smoothing have not yet been developed.

Development of efficient secondary solvers is yet another unsolved problem. Block relaxation procedures have been employed and the results were found to be quite satisfactory: only very few iterations were needed\([27, 28]\). This subject deserves a systematic study, however, with particular reference to developing secondary solvers designed to operate on minicomputers or in small partitions of large timesharing systems.

Extension of the convergence rate studies to three-dimensional problems in elasticity, particularly for studying problems posed by linear elastic fracture mechanics, is yet another promising area for research. Computational as well as theoretical studies are needed. Work in this area is currently in progress. For example, hierarchic shape functions of prismatic finite elements have been developed and implemented\([49, 50]\).

Application of adaptive finite element approximation methods based on \( p \)-convergence to non-linear problems is a completely unexplored area. The advantages of having a capability to adjust the distribution of degrees of freedom without mesh refinement and dealing with substantially fewer variables that require global operations, are likely to outweigh the disadvantage of having to employ high order quadrature rules in generating elemental stiffness matrices. A particularly important computational problem in non-linear mechanics is posed by recent developments in the understanding of ductile fracture processes, which require evaluation of two energy related quantities, known as the "\( J \)-integral" and the "tearing modulus", \( installation of a new method for computing stress intensity factors in linear elastic fracture mechanics.
which measure the effect of crack length increments on the potential energy of non-linearly elastic solids \[51, 52\].

7. SUMMARY AND CONCLUSIONS

The present state of development of a new finite element approximation procedure has been reviewed. In this procedure the finite element mesh is fixed and the number or type of basis functions is varied over the mesh, either uniformly or selectively, until some desired level of precision is reached. The basis functions are complete polynomials, optionally supplemented by rational or other non-polynomial basis functions. They are hierarchic, i.e. the set of basis functions associated with an element is a subset of the basis functions associated with each higher order element of the same kind. Consequently, the stiffness matrix of each element is embedded in the stiffness matrices of all higher order elements of the same kind. There are several computational advantages, the most important of which are that the rate of convergence is much faster than the rate of convergence achieved through mesh refinement and the number and distribution of degrees of freedom can be modified without sacrificing the effort spent on triangularizing the initial stiffness matrix. Thus the approximation error can be efficiently controlled.

It is concluded that development of adaptive finite element software systems, in which redistribution of the degrees of freedom is achieved through controlling the number and type of basis functions, is feasible. Such systems are likely to be more efficient, by an order of magnitude or more, than conventional finite element software systems.

Acknowledgements—The material surveyed here is the result of several years of effort of many investigators. The writer is particularly indebted to Professor Ivo Babuska of the University of Maryland, Dr. Alberto Peano of the Polytechnic University of Milan and his colleagues at Washington University, Drs. I. N. Katz, M. P. Rossow and P. K. Basu for their contributions. Various phases of work on p-convergent approximations have been supported by the National Science Foundation, the U.S. Department of Transportation, the Association of American Railroads, AMCAR Division of ACF Industries, the Air Force Office of Scientific Research and the Electric Power Research Institute through grants and contracts awarded to Washington University.

REFERENCES

17. J. Morgan and R. Scott, The Dimension of the Space of C^1 Piecewise Polynomials. manuscript.
Some recent developments in finite element analysis

25. I. Babuska, Private Communication.
35. I. N. Katz, Private Communication.