AD-A053 028 MARYLAND UNIV COLLEGE PARK COMPUTER SCIENCE CENTER F/G 12/1
ON THE RELIABILITY AND OPTIMALITY OF THE FINITE ELEMENT METHOD.(U)
APR 78 I BABUSKA, W C RHEINBOLDT N00014-77-C-0623
UNCLASSIFIED TR-643 NL

END
DATE FILMED
5-78
DDC

COMPUTER SCIENCE TECHNICAL REPORT SERIES

UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND
20742

12

Technical Report TR-643
ONR-N00014-77-C-0623-643

April 1978

ON THE RELIABILITY AND OPTIMALITY OF THE FINITE ELEMENT METHOD[1)]

by

I. Babuška[2)] and W. C. Rheinboldt[3)]

To be presented as an invited address at the "Symposium on Future Trends in Computerized Structural Analysis and Synthesis", Washington, D.C., Oct. 30-Nov. 1, 1978.

1)This work was in part supported under DOE Contract E(401)3443, NSF Grant MCS-72-03721A06, and ONR Grant N00014-77-C-0623.

2)Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742.

3)Computer Science Center, University of Maryland, College Park, MD 20742.

Abstract

An overview is presented of the authors' recent theoretical and experimental results on reliable and computationally efficient a-posteriori error bounds for finite element solutions. These estimates are composed from error indicators evaluated on the individual elements, and these indicators in turn allow for a very effective approach to the effective construction of optimal meshes. Finally, some views are presented about possible future trends in the development of finite element software and an outline is given of the design of an experimental finite element system currently under development which incorporates many of these ideas and results.

1. Introduction

In recent years considerable progress has been made in the theoretical analysis and the variety and depth of applications of the finite element method (see, e.g., [1]). The pace of these developments does not appear to be diminishing (see, e.g., [2]). It is surprising, however, that a subject of considerable practical importance, namely, the assessment of the reliability of the results of the finite element computations, has not been studied more intensively.

Many of the current reliability studies consist of various comparisons for different benchmark problems. Theoretical error estimations are almost always of an a-priori type and, as important as such estimates are for the theory of the method, they are rarely of much use for practical error determinations. In particular, since in practical applications an accuracy of 5-10% may be entirely acceptable, any error bound that typically represents an overestimate by a factor of five is clearly worthless.

It appears that for any effective computation of reliable error bounds for finite element solutions we should use a-posteriori estimates that are based on information obtained during the solution process itself. Moreover, such estimates may also provide a tool for the adaptive design of optimal finite element meshes.

This is the topic of this paper. More specifically, we present here an overview of recent theoretical and experimental results ([3],[4],[5],[6]) on reliable and computationally efficient a-posteriori error bounds for the finite element method. These estimates are composed from "error-indicators"

evaluated on the elements used in the solution, and these error indicators in turn allow for a very effective approach to the adaptive construction of optimal meshes. After a brief discussion in Section 2 of typical a-priori estimates and their shortcomings for practical purposes, we present in Section 3 some of the principal results about the a-posteriori estimates. This includes also a few illustrative numerical experiments. Then Section 4 addresses the design of optimal finite element meshes using these estimates. Finally, in Section 5 we sketch some views about possible future trends in the development of finite element software and outline an experimental finite element system currently under development which incorporates many of these ideas and results.

2. Rates of Convergence

In this section we present some typical examples of theorems about the rate of convergence of the finite element solutions to the exact solution. For reasons of simplicity we restrict the discussion to the simple model problem in R^2

$$-\Delta u + u = f, \text{ on } \Omega \subset R^2$$

$$\frac{\partial u}{\partial n} = 0, \text{ on } \partial\Omega \tag{2.1}$$

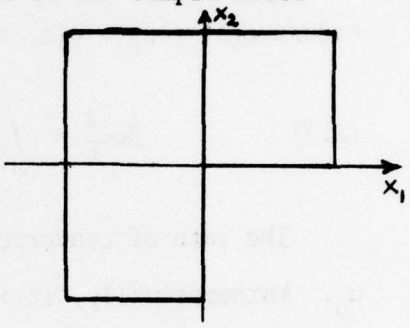


Figure 1

where Ω is an L-shaped domain as shown in Figure 1, n represents the outward normal of $\partial\Omega$, and f is a suitable given function on $\bar{\Omega}$.

On $\bar{\Omega}$ we consider a family P of quasi-uniform triangular meshes Δ; that is, the ratio of the largest triangle-side to the smallest triangle-side occurring in Δ is bounded by a constant depending only on P. For any $\Delta \in P$ let $S(\Delta)$ be the set of all continuous functions on $\bar{\Omega}$ which are linear on each triangle of Δ. The dimension of the space $S(\Delta)$ shall be denoted by $N(\Delta)$. In addition, let Q_p be the set of all polynomials of degree at most p on $\bar{\Omega}$ and N_p the dimension of Q_p.

For any $\Delta \in P$ the finite element solution of (2.1) is now the function $u = u(\Delta) \in S(\Delta)$ such that for all $v \in S(\Delta)$

$$(2.2) \qquad \int_{\Omega} \left[\frac{\partial u}{\partial x_1} \frac{\partial v}{\partial x_1} + \frac{\partial u}{\partial x_2} \frac{\partial v}{\partial x_2} + uv \right] dx = \int_{\Omega} fv dx \ .$$

Correspondingly $u = u(p) \in Q_p$ is the function for which (2.2) holds for all $v \in Q_p$. The exact solution of (2.1) is denoted by u_0, and we measure the errors $e(\Delta) = u_0 - u(\Delta)$ and $e(p) = u_0 - u(p)$ in the energy norm

$$(2.3) \qquad \|e\|_E^2 = \int_{\Omega} \left[\left(\frac{\partial e}{\partial x_1}\right)^2 + \left(\frac{\partial e}{\partial x_2}\right)^2 + e^2 \right] dx \ .$$

The rate of convergence depends on the smoothness of the exact solution u_0. Mathematically, it is advantageous to introduce the so-called Besov space $B^{\alpha}_{2,\infty}(\Omega)$ (see, e.g., [7]) and for $u_0 \in B^{\alpha}_{2,\infty}(\Omega)$ to express the error in terms of the norm of that space. We shall not enter into details here, but note that $H^{\alpha}(\Omega) \subset B^{\alpha}_{2,\infty} \subset H^{\alpha+\varepsilon}(\Omega)$ where $H^{\alpha}(\Omega)$ is the usual Sobolev space of fractional order α on Ω. We mention also that the function $r^{2/3}\cos 2\theta/3$--representing the natural singularity of our problem on the L-shaped domain Ω--belongs to $B^{5/3}_{2,\infty}(\Omega)$ but not to $H^{5/3}(\Omega)$.

For the following rate of convergence result we consider sequences of meshes $\Delta_i \in P$ for which $S(\Delta_i) \subset S(\Delta_{i+1})$ and $N(\Delta_i) \to \infty$, as $i \to \infty$. We call this a sequence of proper refinements.

Theorem 2.1: Suppose that the exact solution u_0 of (2.1) belongs to $B^{\alpha}_{2,\infty}(\Omega)$, then

$$(2.4) \qquad \|e(p)\|_E \le C\, N_p^{\frac{1-\alpha}{2}} \|u_0\|_{B^{\alpha}_{2,\infty}} \ , \text{ as } N_p \to \infty$$

where the constant C is independent of p and u_0 but depends on α and Ω, etc. If $1 < \alpha < 2$, then for any sequence of proper refinements

$$\|e(\Delta)\|_E \leq C\, N(\Delta)^{\frac{1-\alpha}{2}} \|u_0\|_{B^{\alpha}_{2,\infty}}, \text{ as } N(\Delta) \to \infty, \tag{2.5}$$

where C is independent of u_0 and the meshes used, but once again depends on α, Ω, etc.

For the finite element solution the case $\alpha = 2$ requires special attention. But we note that when linear finite elements are used, the rate cannot be better than $N^{-1/2}$.

It is important that conversely the rate of convergence completely characterizes the regularity of the solution u_0:

Theorem 2.2: Suppose that either

$$\|e(p)\|_E \leq K\, N_p^{\frac{1-\alpha}{2}}, \text{ as } N_p \to \infty, \tag{2.6}$$

or for $1 < \alpha < 3/2$ and for any sequence of proper refinements,

$$\|e(\Delta)\|_E \leq K\, N(\Delta)^{\frac{1-\alpha}{2}}, \text{ as } N(\Delta) \to \infty, \tag{2.7}$$

then $u_0 \in B^{\alpha}_{2,\infty}(\Omega)$ and

$$\|u_0\|_{B^{\alpha}_{2,\infty}} \leq C\,[K + (\int_{\Omega} u_0^2 dx)^{1/2}]\,. \tag{2.8}$$

In particular, since $r^{2/3} \cos 2\theta/3 \in B^{5/3}_{2,\infty}(\Omega)$, the natural singularity slows down the convergence rate to $N^{-1/3}$ [*] as compared to the value

*Usually only the value $N^{-1/3+\varepsilon}$ is proved.

$N^{-1/2}$ for smooth solutions. In [8]--and, with more details in [9] --it was shown that for properly refined sequences of meshes the rate $N^{-1/2}$ can be obtained even in the presence of the natural singularity. In that case it is natural to use weighted Besov spaces with weights that depend on the refinement. This type of result reflects the importance of the use of properly refined meshes.

The comparison between the behavior of the polynomial approximations u_p and the finite element solutions $u(\Delta)$ shows that for the same degree of freedom $N = N_p = N(\Delta)$ the polynomials give the same rate of convergence if the solution is badly behaved. Of course, for smooth u_0 the convergence rate for the polynomial solutions will be better or at least never worse than that of the $u(\Delta)$. This is the basis of the p-convergence approach developed in [10],[11].

For practical purposes the estimates of Theorem 2.1 are not useful. The constants are computationally unavailable, and so is the norm of the exact solution. The estimates (2.4),(2.5) are asymptotic in nature and hence characterize the convergence rate for large N. For smaller N, the decrease of the error-norm with growing dimension is often faster than the asymptotic rate. This effect has been observed in many practical situations. It is intuitively understandable since changes of u_0 have to be compared to the mesh size.

As stated at the outset, the discussion in this section is only meant to illustrate a typical situation. Theorems of the form given here apply also in many more general cases (see, e.g., [12]).

3. A-Posteriori Error Estimates

The discussion of the previous section shows that a-priori estimates, such as (2.4),(2.5), are not directly usable for practical error assessments. In order to obtain computable and reliable bounds, we need to go to a-posteriori estimates which use information obtained during the solution process itself. The theory of these estimates differs somewhat for problems in R^1 and those in R^d with $d > 1$, and accordingly we discuss these cases separately.

3.1 One-Dimensional Problems

Once again--for simplicity--we restrict the presentation to a simple model problem. The results also extend to more general cases, as, for instance, the beam equation. Consider the two-point boundary problem

$$
\begin{aligned}
&-\frac{d}{dx}\, a(x)\, \frac{du}{dx} + b(x)u = f(x), \quad 0 < x < 1 \\
&u(0) = u(1) = 0
\end{aligned}
\tag{3.1}
$$

with sufficiently smooth coefficients a,b,f on $[0,1]$ such that $a(x) \geq a > 0$, and $b(x) \geq 0$ for $0 \leq x \leq 1$.

Let P be a family of partitions

$$
\Delta:\ 0 = x_0^\Delta < x_1^\Delta < \ldots < x_{m-1}^\Delta < x_m^\Delta = 1, \quad m = m(\Delta) \geq 1\ , \tag{3.2}
$$

on the interval $[0,1]$, and introduce the notations

$$h_j(\Delta) = x_j^\Delta - x_{j-1}^\Delta, \quad j = 1,\ldots,m$$

(3.3)

$$\bar{h}(\Delta) = \max_{j=1,\ldots,m} h_j(\Delta), \quad \underline{h}(\Delta) = \min_{j=1,\ldots,m} h_j(\Delta) .$$

The family $\mathcal{P}$ will be called κ-regular if there are constants $\lambda > 0$, $\kappa \geq 1$ such that

(3.4) $$\underline{h}(\Delta) \geq \lambda \bar{h}(\Delta)^\kappa, \ \forall\ \Delta \in \mathcal{P} .$$

For this presentation we shall use linear elements on the intervals $I_j(\Delta) = [x_{j-1}^\Delta, x_j^\Delta]$, $j = 1,\ldots,m$, to solve (3.1). As before, we denote by $u(\Delta)$ the resulting finite element solution and by u_0 the true solution of our problem. On every separate interval $I_j(\Delta)$ the function $u(\Delta)$ is linear and hence the residuals

(3.5) $$r_j(x) = -\frac{d}{dx} a(x) \frac{d}{dx} u(\Delta)(x) + b(x)u(\Delta)(x) - f(x), \ x_{j-1} \leq x \leq x_j, \quad j = 1,\ldots,m,$$

and their integrals

(3.6) $$\nu_j(\Delta)^2 = \int_{x_{j-1}}^{x_j} r_j(x)^2 dx, \quad j = 1,\ldots,m$$

are easily computable. We introduce the error indicators

(3.7) $$\varepsilon_j(\Delta)^2 = \frac{1}{12} \frac{h_j^2 \nu_j(\Delta)^2}{a(\frac{1}{2}(x_{j-1}+x_j))}, \quad j = 1,\ldots,m$$

and the error estimator

$$\varepsilon(\Delta)^2 = \sum_{j=1}^{m} \varepsilon_j(\Delta)^2 \ . \tag{3.8}$$

Then the following result holds; see [5],[6]:

Theorem 3.1: Let u_0 be the true solution of (3.1) and $u(\Delta)$ the finite element solution for any $\Delta \in P$. Then without any regularity assumption about P the error $e(\Delta) = u_0 - u(\Delta)$ satisfies

$$\|e(\Delta)\|_E \le \frac{\sqrt{12}}{\pi}\, \varepsilon(\Delta)(1+O(\bar{h}(\Delta))), \text{ as } \bar{h}(\Delta) \to 0 \ . \tag{3.9}$$

The bound of the O-term depends on a,b but not on u_0. If P is κ-regular with $1 \le \kappa < 2$ and $u_0''(x)$ has only simple roots in $[0,1]$, then

$$\|e(\Delta)\|_E = \varepsilon(\Delta)(1+O(\bar{h}(\Delta)^{\bar{\kappa}})), \text{ as } \bar{h}(\Delta) \to 0 \ , \tag{3.10}$$

where $\bar{\kappa} = 1 - \kappa/2$.

In the present case the energy norm used in (3.9) and (3.10), of course, has the form

$$\|e\|_E^2 = \int_0^1 (a(x)(\frac{de}{dx})^2 + b(x)e^2)dx \ . \tag{3.11}$$

Note that in the case of (3.10) the error indicator (3.8) provides the exact error for $\bar{h} \to 0$. In other words, the effectivity constant

$$\theta(\Delta) = \frac{\|e(\Delta)\|_E}{\varepsilon(\Delta)} \tag{3.12}$$

tends to one for $\bar{h} \to 0$. Even the general estimate (3.9) ensures that asymptotically we have $\theta(\Delta) \le \sqrt{12}/\pi < 1.103$.

In applications we may readily accept, say, $1/2 \le \theta(\Delta) \le 2$, that is a 100% error in $\theta(\Delta)$ with respect to the limiting value $\theta(\Delta) = 1$. On the other hand, a corresponding error in the solution relative to u_0, that is, $1/2 \le \|e(\Delta)\|_E / \|u_0\|_E \le 2$, would be completely unacceptable. This points out once again the essential difference in the estimates in this section and those of Section 2.

As an illustration of the effectivity of Theorem 3.1, we consider the following example:

$$- \frac{d}{dx} (x+\alpha)^p \frac{du}{dx} + (x+\alpha)^q u = f, \quad 0 < x < 1, \quad \alpha > 0,$$

$$u(0) = u(1) = 0 . \tag{3.12}$$

Here f was chosen such that the exact solution of (3.12) is

$$u_0(x) = (x+\alpha)^r - [\alpha^r(1-x)+(1+\alpha)^r x] . \tag{3.13}$$

Note that for small α and negative r severe near singularities may be created. In the numerical examples given below we chose

$$p = 0, \quad q = 1, \quad r = -\frac{1}{4}, \quad \alpha = \frac{1}{100}, \tag{3.14}$$

in which case $\|u_0\|_E = 6.09811$. Two different types of meshes are used, namely, uniform meshes with $h_j = 1/m$ and asymptotically optimal meshes in the sense of Section 4 below. Table 1 below shows the coordinates of the nodal points of the asymptotically optimal mesh with $m = 10$.

Table 1

Asymptotically Optimal Mesh for m = 10

j	x_j	j	x_j	j	x_j
0	0	4	.01443	8	.11781
1	.00207	5	.02398	9	.26831
2	.00487	6	.03732	10	1
3	.00877	7	.06318		

In Table 2 we show the relative error $E = 100\ \|e\|_E/\|u_0\|_E$ in percent of the energy norm of the exact solution and the effectivity constant θ of (3.12) for the two types of meshes. In addition, two quantities ω and E_0 are included which relate to the optimality of the mesh as discussed in Section 4 below. More specifically, ω is the ratio of the maximal and minimal error indicators. As we shall see in Section 4, a mesh is asymptotically optimal if all its error indicators are equal. Thus ω is a measure of optimality. The value E_0 is the value of E for the asymptotically optimal partition.

The reliability and effectivity of the estimates are evident. A 26% error relative to the norm of the exact solution corresponds only to a 15% error in estimating the actual error by means of the estimator (3.12). For the "more optimal" meshes the results are even better.

Table 2

Uniform mesh

m	E	E_0	θ	ω
5	85.301	22.613	.1706	8.84(+6)
10	73.768	11.306	.2950	2.34(+7)
20	58.784	5.653	.4702	5.31(+7)
40	41.933	2.827	.6708	1.11(+8)
80	26.316	1.413	.8419	2.19(+8)

Asymptotically optimal mesh

m	E	E_0	θ	ω
5	22.243	22.613	.6524	5.854
10	11.289	11.306	.9025	2.274
20	5.652	5.653	.9757	1.372
40	2.826	2.827	.9940	1.111
80	1.413	1.413	.9984	1.031

3.2 Higher-Dimensional Problems

Once more we restrict ourselves to a special model problem, namely, the two-dimensional boundary value problem

$$
\begin{aligned}
&-\sum_{i,j=1}^{2} \frac{\partial}{\partial x_i}\, a_{ij}\, \frac{\partial u}{\partial x_j} + bu = f(x), \quad \forall\, x \in \Omega \\
&u(x) = g(x), \quad \forall\, x \in \partial\Omega\ .
\end{aligned}
\tag{3.15}
$$

Here the coefficients a_{ij}, b are constants with

$$
\begin{aligned}
&\bar{a}(x_1^2+x_2^2) \geq \sum_{i,j=1}^{2} a_{ij}x_ix_j \geq \underline{a}(x_1^2+x_2^2), \quad \forall\, x \in R^2, \quad a_{12} = a_{21} \\
&b \geq 0\ ,
\end{aligned}
\tag{3.16}
$$

f is continuous on $\bar{\Omega}$ and continuously differentiable on Ω, and g is continuous and piecewise continuously differentiable on $\partial\Omega$.

We restrict the domain Ω to sets of R^2 bounded by lines that are parallel to the coordinate axis. For example, Ω may be the L-shaped domain of Figure 1 in Section 2.

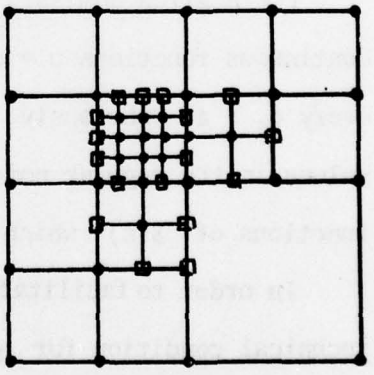


Figure 2

The admissible meshes Δ on Ω shall be sets $\Delta = \{q_j\}$ of squares q_j of the form

$$
q_j = \{x \in R^2 \;\big|\; \alpha < x_i < \alpha + h,\ i = 1,2\}, \quad h = h(q_j) > 0\ ,
\tag{3.17}
$$

such that

(i) for any $q_j \in \Delta$ the closure $\bar{q}_j$ belongs to $\bar{\Omega}$;

(3.18) (ii) $\bigcup_{q_j \in \Delta} \bar{q}_j = \bar{\Omega}$;

(iii) $q_i \cap q_j = \emptyset$ for any $q_i, q_j \in \Delta$, $i \neq j$; that is, $\bar{q}_i \cap \bar{q}_j$ consists only of boundary points of q_i and q_j.

A typical mesh on a square Ω is given in Figure 2. The corner points of the squares $q_j \in \Delta$ are the nodal points of the mesh. A nodal point $\nu \in \bar{\Omega}$ is <u>regular</u> if it is a cornerpoint of all the squares $q_j \in \Delta$ for which $\nu \in \bar{q}_j$; otherwise ν is <u>irregular</u>. Note that then any nodal point on $\partial\Omega$ is necessarily regular. In Figure 1 the regular and irregular nodes are marked by black dots and open squares, respectively.

A collection P of admissible meshes Δ on Ω is N-regular if for any $\Delta \in P$ the number of irregular nodes on any side of a square $q_j \in \Delta$ is at most equal to N.

For a given admissible mesh Δ on Ω let $S(\Delta)$ be the set of all continuous functions $u = u(x_1,x_2)$ on $\bar{\Omega}$ which are bilinear in x_1,x_2 on every $q_j \in \Delta$. Obviously, any $u \in S(\Delta)$ is then uniquely determined by its values in the regular nodes of Δ. By $\overset{\circ}{S}(\Delta)$ we denote the subset of all functions of $S(\Delta)$ which are equal to zero on $\partial\Omega$.

In order to facilitate the error estimation on the boundary we add the technical condition for g that for any $q_j \in \Delta$ which has a side σ on $\partial\Omega$ we have with a suitable constant $c > 0$

$$(3.19) \qquad \int_\sigma (g'-u(\Delta)')^2 dx \le \frac{c}{h(q_j)} \int_\sigma (g-u(\Delta))^2 dx, \ \sigma = \partial q_j \cap \partial\Omega \ .$$

This holds, for instance, for quadratic polynomials g.

For a given admissible mesh, the finite element solution of (3.15) is now the (unique) function $u(\Delta) \in S(\Delta)$ for which

$$(3.20) \qquad \int_\Omega \Big(\sum_{i,j=1}^{2} a_{ij} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} + buv \Big) dx = \int_\Omega fv dx, \ \forall\, v \in \mathring{S}(\Delta)$$

and $u(\Delta) = g$ holds at all nodal points of Δ on $\partial\Omega$. As before u_0 denotes the exact solution of (3.15) and we are interested in estimating the error $e = u_0 - u(\Delta)$ in the energy norm

$$\|e\|_E^2 = \int_\Omega \Big(\sum_{i,j=1}^{2} a_{ij} \frac{\partial e}{\partial x_i} \frac{\partial e}{\partial x_j} + be^2 \Big) dx \ .$$

As in the one-dimensional case we may compute on each $q_j \in \Delta$ the residual

$$(3.21) \qquad r_j(x) = - \sum_{i,j=1}^{2} \frac{\partial}{\partial x_i} a_{ij} \frac{\partial}{\partial x_j} u(\Delta) + bu(\Delta) - f(x), \ \forall\, x \in \bar{q}_j \ .$$

If, for instance, $a_{11} = a_{22} = 1$, $a_{12} = 0$, $b = 0$ then it is easily seen that $r_j(x) \equiv 0$ in any q_j for which $f(x) \equiv 0$, $x \in \bar{q}_j$. This prevents us from proceeding analogously to the one-dimensional case. Instead we use an approach which is equivalent to that presented in [3].

Let $q_j \in \Delta$ be any square of the mesh and as indicated in Figure 3 let $\sigma_1, \bar{\sigma}_1, \sigma_2, \bar{\sigma}_2$ be the four sides of q_j. For a square q_j in the interior of the domain, we denote by J_{σ_1} and $J_{\bar{\sigma}_1}$ the jumps of $\frac{\partial}{\partial x_1} u(\Delta)$ on σ_1 and $\bar{\sigma}_1$, respectively, and correspondingly,

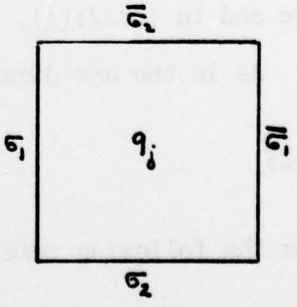


Figure 3

by J_{σ_2} and $J_{\bar{\sigma}_2}$ those of $\frac{\partial}{\partial x_2} u(\Delta)$ on σ_2 and $\bar{\sigma}_2$, respectively. Now we compute the following quantities:

If $q_j \subset \Omega$ then

$$\text{(i)} \quad \eta(q_j)^2 = \frac{1}{24} \frac{h(q_j)^2}{\underline{a}} \int_{q_j} r_j(x)^2 dx$$

$$\text{(ii)} \quad \rho_\sigma(q_j)^2 = \begin{cases} \frac{1}{24} a_{11} h(q_j) \int_\sigma J_\sigma(x)^2 dx, & \sigma = \sigma_1, \bar{\sigma}_1 \\ \\ \frac{1}{24} a_{22} h(q_j) \int_\sigma J_\sigma(x)^2 dx, & \sigma = \sigma_2, \bar{\sigma}_2 \end{cases} \tag{3.22}$$

and the error indicator of q_j is defined by

$$\varepsilon(q_j)^2 = \eta(q_j)^2 + \rho_{\sigma_1}(q_j)^2 + \rho_{\bar{\sigma}_1}(q_j)^2 + \rho_{\sigma_2}(q_j)^2 + \rho_{\bar{\sigma}_2}(q_j)^2 \tag{3.23}$$

On the other hand, if $\bar{q}_j$ has sides in common with $\partial\Omega$ then for any such side $\sigma \in \bar{q}_j \cap \partial\Omega$ the corresponding ρ-term in (3.23) is replaced by

$$\tilde{\rho}_\sigma^2(q_j) = \rho_{\bar{\sigma}}^2(q_j) + \frac{3\bar{a}}{h(q_j)} \int_\sigma (g - u(\Delta))^2 dx \ . \tag{3.24}$$

Here and in (3.22)(i), $\bar{a}$ and $\underline{a}$ are the constants of (3.16).

As in the one-dimensional case we introduce now the error estimator

$$\varepsilon(\Delta)^2 = \sum_{q_j \in \Delta} \varepsilon(q_j)^2 \ . \tag{3.25}$$

Then the following result holds [3]:

Theorem 3.2: The error $e = u_0 - u(\Delta)$ between the exact solution and the finite element solution of (3.15) satisfies

$$\frac{1}{K}\,\varepsilon(\Delta) \le \|e\|_E \le K\,\varepsilon(\Delta) \tag{3.26}$$

with some $K \ge 1$ which is independent of Δ but depends on a, N, etc.

In contrast to Theorem 3.1 we cannot claim here that $K \to 1$ when $\bar{h}(\Delta) \to 0$. But all experimental experience so far indicates that K never exceeds a value of two for a large class of problems.

Theorem 3.2 holds no matter which constant factors are used in the definition of the quantities (3.22),(3.24). But the powers of $h(q_j)$ occurring in these quantities are, of course, essential. The particular constants chosen in (3.22),(3.24) arose from an asymptotic analysis of the case of a uniform mesh Δ.

We illustrate Theorem 3.2 on the following numerical example:

$$\frac{\partial^2 u}{\partial x_1^2} + \alpha \frac{\partial^2 u}{\partial x_1 \partial x_2} + \frac{\partial^2 u}{\partial x_2^2} = 0, \quad x \in \Omega = (0,1) \times (0,1) \subset R^2$$

$$u = g, \quad x \in \Omega \tag{3.27}$$

The boundary condition is chosen to provide for a specific exact solution u_0. In particular, we consider the following two cases:

$$\text{(I)} \quad \alpha = 0, \quad u_0 = \mathrm{Re}(z - z_0)^{1/2}, \quad z = x_1 + ix_2, \quad z_0 = 1.03,$$

$$\text{(II)} \quad \alpha = 1, \quad u_0 = \mathrm{Re}(w - w_0)^{1/2}, \quad w = y_1 + iy_2, \quad w_0 = y_1^0 + iy_2^0$$

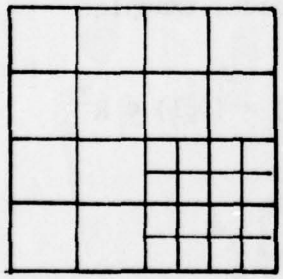
where

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} , \quad \begin{pmatrix} y_1^0 \\ y_2^0 \end{pmatrix} = T \begin{pmatrix} 1.03 \\ 0 \end{pmatrix}$$

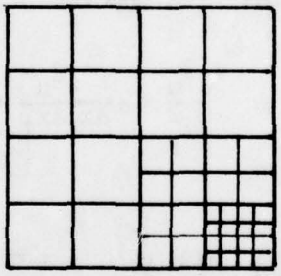
and

$$T = \frac{1}{\sqrt{2}} \begin{pmatrix} (1 + \frac{\alpha}{2})^{-1/2} & (1 + \frac{\alpha}{2})^{-1/2} \\ (1 - \frac{\alpha}{2})^{-1/2} & -(1 - \frac{\alpha}{2})^{-1/2} \end{pmatrix}$$

Five different meshes are used, denoted by (a), (b), (c), (d), (e). The meshes (a), (b), (c) are regular partitions of the unit square with step-size $h = \frac{1}{4}, \frac{1}{8}, \frac{1}{16}$, respectively. The other two meshes (d) and (e) are shown in Figure 4.



Mesh (d)



Mesh (e)

Figure 4

Conforming bilinear elements were used. Table 3 gives the corresponding results. Here "active nodes" means the number of degrees of freedom, that is, the size of the corresponding system of linear equations; the other notation is the same as that used in Table 2 for the one-dimensional case. Once again, the a-posteriori estimates are clearly very accurate and reliable.

Table 3

Case	Number of active nodes	Number of irreg. nodes	elem.	Energy error abs. $\|e\|_E$	Energy error in % E	Effect. Const. θ	Ratio of error ind. ω
I(a)	9	0	16	.0867	13.51	.917	1.17(3)
I(b)	49	0	64	.0492	7.68	.914	2.54(4)
I(c)	225	0	256	.0253	3.95	.921	3.43(5)
I(d)	21	4	28	.0534	8.35	.922	1.31(3)
I(e)	33	8	40	.0271	5.79	.952	6.22(3)
II(a)	9	0	16	.0677	11.28	.683	7.09(2)
II(b)	49	0	64	.0384	6.40	.711	3.98(3)
II(c)	225	0	256	.0201	3.35	.736	1.57(4)
II(d)	21	4	28	.0418	6.97	.731	3.03(2)
II(e)	33	8	40	.0294	4.90	.805	2.14(2)

4. Optimal Meshes

4.1 One-Dimensional Problems

Once again we consider the model problem (3.1). The theory of a-posteriori estimates sketched in Section 3.1 provides a basis for the construction of optimal meshes (3.2) in this case.

For this, we call a partition (3.2) a (ξ,m)-partition if there exists a continuous, increasing function ξ on $[0,1]$ which satisfies

$$\xi(x_j^\Delta) = \frac{j}{m}, \quad j = 0,1,\ldots,m(\Delta) \tag{4.1}$$

and is twice continuously differentiable on each subinterval $(x_{j-1}^\Delta, x_j^\Delta)$, $j = 1,\ldots,m$. Any partition Δ is a (ξ,m)-partition for the piecewise linear function

$$\xi(x) = \frac{j-1}{j} + \frac{1}{mh_j}\,(x-x_{j-1}), \quad x_{j-1}^\Delta \le x \le x_j^\Delta, \quad j = 1,\ldots,m\,. \tag{4.2}$$

The error corresponding to the (ξ,m)-partition Δ shall be denoted by $e(\xi,m)$.

It now turns out that the (ξ_0,m)-partition Δ_0 defined by

$$\xi_0(x) = \gamma_0 \int_0^x [a(t)u_0''(t)^2]^{1/3}dt, \quad \gamma_0^{-1} = \int_0^1 [a(t)u_0''(t)^2]^{1/3}dt \tag{4.3}$$

is asymptotically optimal in the following sense ([6]):

Theorem 4.1: Suppose that $u_0''(x)$ has only simple roots in $[0,1]$, then the following statements hold:

(a) The (ξ_0,m)-partition Δ_0 of (4.3) satisfies the κ-regularity condition (3.4) with $\kappa = 5/3$ for all m.

(b) The error for the partition Δ_0 is given by

$$\|e(\xi_0,m)\|_E = \frac{1}{12m^2\gamma_0^3}\,(1+O(\bar{h}(\Delta_0)),\ \text{as}\ \bar{h}(\Delta_0)\to 0, \tag{4.4}$$

where the constant in the bound of the O-term depends on ξ_0 but not on m.

(c) Let P be a family of 5/3-regular (ξ,m)-partitions. Then

$$\limsup_{m\to\infty}\ \frac{\|e(\xi_0,m)\|_E}{\|e(\xi,m)\|_E} < 1 \tag{4.5}$$

for any partition in P with $\xi \neq \xi_0$.

Part (c) of this theorem shows that the partition Δ_0 of (4.3) is asymptotically optimal in the sense that for any other partition in P the error is larger for sufficiently large m. Accordingly, we call Δ_0 the asymptotically optimal mesh.

The error indicators $\varepsilon_j(\Delta)$ of (3.7) constitute a simple tool for the construction of meshes with errors that are asymptotically equal to $\|e(\xi_0,m)\|_E$. This is the content of the following result (see again [6]):

Theorem 4.2: Suppose that $u_0''(x)$ has only simple roots in $[0,1]$. Let Δ be any partition (3.2) for which there is a constant μ such that

$$\varepsilon_j(\Delta) = \mu(1+O(h(\Delta)^{\bar{\kappa}})),\quad \bar{\kappa} = 1/12,\ \text{as}\ \bar{h}(\Delta)\to 0\ . \tag{4.6}$$

Then Δ satisfies (3.4) with $\kappa = 5/3$ and

$$\|e(\Delta)\|_E \le \|e(\Delta_0)\|_E(1+O(\bar{h}(\Delta))^{\bar{\kappa}}),\ \text{as}\ \bar{h}(\Delta)\to 0\ . \tag{4.7}$$

The condition (4.6) means that the error indicators should be asymptotically equal. This provides the basis for the construction of partitions Δ satisfying (4.7). A natural approach is here the adaptive mesh-refinement algorithm discussed in the next section.

It can also be shown (see [6]) that the error for the (ξ_0,m)-partition is very stable. More precisely, if Δ is any (ξ,m)-partition with $\xi = \xi_0 + \lambda\eta$ then

$$\|e(\xi,m)\|_E = \|e(\xi_0,m)\|_E(1+O(\lambda^2)) \text{ as } \lambda \to 0 \ . \tag{4.8}$$

On the other hand, the distribution of the nodal points of the optimal partition is not particularly stable. In other words, we should concentrate on making the error indicators nearly equal and not concern ourselves with an accurate computation of Δ_0 itself.

Numerical results illustrating these results were already included in Tables 1 and 2 of Section 3.1. In particular, in the second part of Table 2 the ω-values are nearing one and as expected the values of E and E_0 are very close.

4.2 Higher-Dimensional Problems

For problems in two- or higher dimensions there exists as yet no theory of the extent sketched in Section 4.1 for the one-dimensional case. In [3] it is shown that here again asymptotic optimality of the mesh calls for the (asymptotic) equality of the error indicators. This provides the basis for the following type of adaptive mesh-refinement algorithm.

Suppose that the error indicators ε_i have an asymptotic behavior of the form

$$\varepsilon_i = c\, h_i^{\lambda}, \quad \text{as } h_i \to 0 \,. \tag{4.9}$$

If any element τ_i with corresponding value ε_i was generated by subdividing an element in a previous mesh with value ε_i^{old} then (4.9) suggests that the (worst) indicator after dividing τ_i will be approximately

$$\varepsilon_i^{new} = (\varepsilon_i)^2/\varepsilon_i^{old} \,. \tag{4.10}$$

Practical experience has shown that, in general, this prediction is very satisfactory.

Clearly, we should refine only those elements which have an indicator above the largest predicted new indicator in the new mesh. In order to start this process, all elements in the basic mesh are refined in the first step. Hence we have the following refinement algorithm:

1. cut := 0
2. If "current mesh Δ is the basic mesh" then go to 4
3. For "each element τ in Δ" do
   - 3.1 Compute error indicator ε of element τ
   - 3.2 $\varepsilon^{new} = \varepsilon^2/\varepsilon^{old}$
   - 3.3 If $\varepsilon^{new} >$ cut then cut := ε^{new}
4. For "each element τ in Δ" do
   - 4.1 If $\varepsilon >$ cut then "subdivide τ and for each new element set $\varepsilon^{old} := \varepsilon$

Implementation details will be discussed elsewhere.

5. Future Trends

5.1 Computational Goals

In the development of typical mathematical library programs, the aim is to achieve--whenever possible--a prescribed and typically high accuracy. Here "accuracy" refers usually to the size of the errors introduced by the particular numerical method, and the program is reliable if its overall failure rate is low.

For the general-purpose, finite-element software used, say, in structural engineering the requirements are rather different. The user needs to obtain results which predict, with an acceptable degree of reliability and accuracy, the actual behavior of the mechanical structure at hand. In other words, the terms "accuracy" and "reliability" are used in a much broader context than before. It is clearly inadequate to assess the results solely on the basis of the errors introduced by the numerical calculations. For example, an earlier decision to use a plate model rather than a full three-dimensional formulation may result in final stress values that differ by 10-20% from each other. In contrast to this, the numerical errors are entirely negligible. In other words, the accuracy of the numerical results should be a measure of their deviation from the solution of a "higher" mathematical model. Moreover, these results can only be considered reliable if changes in the reference model and, more generally, in the entire sequence of steps from the problem formulation to the final results have relatively small effects.

Generally, the computations involved in finite-element applications are very extensive and highly demanding on the available computer resources, not to mention the corresponding demands on the user personnel. In most cases the user is forced to set a maximum allowable cost for any one computational analysis. Ideally then the goal is to achieve optimal, acceptable accuracy (in the above sense) within the prescribed cost range. Here it is essential that often relatively low accuracies--often in the range of 5-10%--are entirely satisfactory.

Without doubt the design goal of producing reliable results of optimal accuracy within some prescribed cost range represents a serious challenge and is not realized in today's finite-element software. However, our own experience has shown that this is not an unrealistic goal (see, e.g., [3], [4], [13], and it appears also that it corresponds well with the trends that are now beginning to emerge in the design of the next generation of such software.

The principal tool for the realization of this design goal is the availability of relatively easily computable error estimates. On some simple model problems we have illustrated in the previous two sections that a-posteriori error estimates are feasible and not very costly to compute. Moreover, these estimates are very reliable. The approaches presented here extend readily to much more general cases including, for instance, the typical elasticity problems in structural analysis and even nonlinear problems. Moreover, the estimates can also be developed for norms other than the energy norm. Of course, the formulas for the error

indicators are then different, but the structure of the estimates remains the same.

The a-posteriori estimates are developed on the basis of a general definition of the finite element method in terms of a weak mathematical formulation involving bivariate forms on suitable function spaces (see, e.g., [13]). This suggests that finite element software should no longer be designed for specific classes of applications but instead should constitute more general "finite element solvers" based on entire classes of such weak formulations. This will provide the needed closer interaction with the theoretical foundations and, at the same time, it will open up a much wider applicability of the software.

In order to simplify the use of the software, the user should not have to understand the internal operation or to make difficult a-priori decisions about the desired computations. For this the user interfaces should be strictly separated into flexible pre- and post-processors which adapt the system to various classes of applications.

The need for reducing the range of decisions asked of the user in today's finite element software requires the introduction of extensive adaptive techniques into the computational process. These techniques are also needed to meet the goal of achieving an optimal solution within a prescribed cost range. The adaptive control of the computation in turn requires the availability of reliable error estimates at all stages of the calculation. In other words, here is a further reason for the introduction of the a-posteriori estimates discussed in the previous sections.

In connection with the optimality requirement we have to distinguish between the design of a (nearly) optimal mesh and the achievement of a (nearly) optimal a-posteriori error estimate for the solution. As we saw before, the optimal error (for a given degree of freedom) is rather stable while this is not true for the optimal mesh. Thus we should certainly concentrate on achieving a (nearly) optimal error within our given cost range and not overplay the search for the optimal mesh. The discussions of Section 4 show that we have a rather simple criterion for the (near) optimality of the error in the near-equality of the error indicators. It provides the basis of simple and yet highly effective adaptive mesh-refinement algorithms (see, e.g., [3],[13]). Since these algorithms aim to keep all error indicators as close together as possible, all resulting meshes--after some initial phase--have a fairly optimal error for the degrees of freedom they incorporate. Thus we may stop the process whenever either a prescribed error tolerance has been achieved or the given computational cost range has been exceeded (see [3],[4]).

5.2 A Prototype Adaptive Finite Element Solver

As stated before, the general design goals for future finite element software still represent serious challenges and are not yet incorporated into any operational programs. However, as we said, they are within the realm of practicality.

In substantiation of this claim an experimental finite element system is being designed by several of us at the University of Maryland that meets

the following four design criteria:

(a) The system constitutes an applications-independent finite element solver for a certain class of two-dimensional, linear, elliptic problems based on a weak mathematical formulation;

(b) it incorporates extensive adaptive approaches to minimize the critical decisions demanded of the user;

(c) it uses a-posteriori error estimates of the type discussed above to control the adaptive approach and to provide a solution with a (near) optimal error within a prescribed cost range;

(d) it takes advantage of modern advances in the design of data structures and software systems, and, in particular, it is designed to use the natural parallelism and modularity of the finite element method to increase the size of the practically solvable problems.

The general design of the system has been described in [13]. It is being implemented not as a production system but as an experimental system for the evaluation of the various new ideas in it. Accordingly, extensive provisions for evaluating the performance are incorporated in it.

Beyond the new approaches reflected in the criteria (a), (b) and (c) already discussed in the previous section, the introduction of parallelism appears to be a particularly novel aspect of the design. This parallelism is on the procedural level rather than the instructional level because there the expected payoff is much greater. Thus, parallelism is specified in terms of processes which are autonomous units with their own programs and data.

The processes run in parallel and communicate asynchronously in a limited and highly structured manner.

The natural parallel process structure for the finite element system derives from the familiar substructure analysis in engineering design. In principle the substructure segmentation of the data and processing is the same as that used in large industrial applications except that ours is essentially automatic and more flexible, while current finite element systems demand that their users create a rigid segmentation at the manual or even managerial level.

In the design the local data of a process contain almost all the information needed for the computations of that process. Thus the process structure induces on the finite element data a segmentation which is natural to the problem and provides a basis for intelligent storage management in the environment of a single large computer. At the same time the use of parallel processes makes it possible to apply multiple processors effectively--hopefully for significant gains in speed.

References

1. D. H. Norrie, G. de Vries, A Finite Element Bibliography, Plenum Press, New York, 1976.

2. The Finite Element Method, Smithsonian Science Information Exchange, Publ. KSO 2-51, Washington, D.C., 1977.

3. I. Babuška, W. Rheinboldt, Error estimates for adaptive finite element computations, University of Maryland, Institute for Physical Science and Technology, Tech. Note BN-854, May 1977; SIAM J. Num. Anal. 15, 4, August 1978 (in press).

4. I. Babuška, W. Rheinboldt, Computational aspects of finite element analysis, in Mathematical Software III, ed. J. R. Rice, Academic Press, New York, 1977, 225-255.

5. I. Babuška, W. Rheinboldt, A-posteriori error estimates for the finite element method, University of Maryland, Computer Science Technical Report TR-581, Sept. 1977; Int. J. for Numer. Meth. in Eng., 1978, (in press).

6. I. Babuška, W. Rheinboldt, Analysis of optimal finite element meshes in R^1, University of Maryland, Institute for Physical Science and Technology, Tech. Note BN-869, March 1978.

7. J. Bergh, J. Löfström, Interpolation Spaces. An Introduction, Springer Verlag, Heidelberg, New York, 1976.

8. I. Babuška, Finite element method for domains with corners, Computing 6, 1970, 264-273.

9. I. Babuška, R. Kellogg, J. Pitkaranta, Direct and inverse error estimates for finite elements with mesh refinement (in preparation).

10. A. Peano, B. Szabo, A. Mehta, Self-adaptive finite elements in fracture mechanics, presented at Symposium on Innovative Numerical Analysis in Applied Engineering Science, Versailles, France, May 23-27, 1977.

11. B. Szabo, Recent developments in finite element analysis, Comp. and Math. with Appl. (to appear).

12. I. Babuška, A. Aziz, Survey lectures on the mathematical foundations of the finite element methods, in The Mathematical Foundations of the Finite Element Method, ed. A. K. Aziz, Academic Press, New York, 1972, 3-363.

13. P. Zave, W. Rheinboldt, Design of an adaptive parallel finite element system, University of Maryland, Computer Science Technical Report TR-593, Nov. 1977.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: ONR-N00014-77-C-0623-643	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE *(and Subtitle)*: ON THE RELIABILITY AND OPTIMALITY OF THE FINITE ELEMENT METHOD.	5. TYPE OF REPORT & PERIOD COVERED: Technical Report.
6. PERFORMING ORG. REPORT NUMBER: TR-643	
7. AUTHOR(s): Ivo Babuska and Werner C. Rheinboldt	8. CONTRACT OR GRANT NUMBER(s): N00014-77-C-0623, EC(AL)-3443
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Computer Science Center, University of Maryland, College Park, MD 20742	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Mathematics Branch, Office of Naval Research, Arlington, VA 22217	12. REPORT DATE: April 1978
	13. NUMBER OF PAGES: 33
14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)*: 34p.	15. SECURITY CLASS. *(of this report)*: UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

20. Abstract (continued)

which incorporates many of these ideas and results.

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

finite element method
error estimates
mesh refinement
optimal mesh
adaptivity
finite element software

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

An overview is presented of the authors' recent theoretical and experimental results on reliable and computationally efficient a-posteriori error bounds for finite element solutions. These estimates are composed from error indicators evaluated on the individual elements and these indicators in turn allow for a very effective approach to the effective construction of optimal meshes. Finally, some views are presented about possible future trends in the development of finite element software and an outline is given of the design of an experimental finite element system currently under development...

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

403 018