SOME A POSTERIORI ERROR ESTIMATORS FOR ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS

RANDOLPH E. BANK AND ALAN WEISER

Abstract. We present three new a posteriori error estimators in the energy norm for finite element solutions to elliptic partial differential equations. The estimators are based on solving local Neumann problems in each element. The estimators differ in how they enforce consistency of the Neumann problems. We prove that as the mesh size decreases, under suitable assumptions, two of the estimators approach upper bounds on the norm of the true error, and all three estimators are within multiplicative constants of the norm of the true error. We present numerical results in which one of the estimators appears to converge to the norm of the true error.

1. Introduction

In this work, we will describe several methods for computing a posteriori error estimates for finite element calculations. That is, given some piecewise polynomial approximation U to $u_{\mathcal{H}}$, the solution of an elliptic partial differential equation, we seek some practical method for computing an estimate of $\|\!|u_{\mathcal{H}} - U|\!\|$ for an appropriate norm $\|\!|\cdot|\!\|$. A priori estimates can give asymptotic rates of convergence as the mesh parameter h tends to zero, but often cannot provide much practical information about the actual errors encountered on a given mesh with a fixed h. A posteriori estimates, on the other hand, attempt to provide the user of a finite element package with such information, enhancing the robustness of the package, and the reliability of the approximations it produces.

There has been a great deal of work by Babuška and his coworkers on local mesh refinement strategies and the a posteriori error indicators necessary for their success [4, 10, 6, 7, 5, 8, 9]. The indicators in [5], for example, is based on solving local Dirichlet problems in the patch of elements surrounding each vertex in the finite element mesh. In this scheme, the Dirichlet boundary conditions insure well-posedness of the local problems. Error indicators can also be based on the computation of the norm of the local residual of the elliptic equation and the jump in the normal derivative of the computed solution at interelement boundaries (e.g., [8] [13]). Such schemes as a rule require less computation than the ones involving the solution of local problems. They also appear to give good results when used in local mesh refinement algorithms. However, with highly nonuniform triangular meshes, as arise in the finite element code $PLTMG$ [11], it is sometimes difficult to weight the residual and boundary terms properly.

Date: Received November 23, 1982; revised March 15, 1984.
1980 *Mathematics Subject Classification.* Primary 65N15, 65N30, 65N50.
This work was supported in part by the Office Of Naval research under grants N00014-80-C-0645 and N00014-76-C-0277.

The error indicators described in this paper are based on solving local Neumann problems in each finite element. Such an approach leads directly to error indicators defined element by element, and at first seems a simple and natural approach. However, some care must be taken to insure the Neumann problems are well posed. As the element diameter tends to zero, the lower order terms in the elliptic operator lose significance and the problems tend to singularity. To be well posed the data (right-hand side) for these problems must be consistent in the limit $h \to 0$. The three procedures we analyze here differ in the method by which this consistency condition is satisfied. In some respects, our schemes are similar to the local residual method of Percell and Wheeler [14], but without the penalty terms.

When the approximation U is the finite element approximation, our error indicators $\bar{e} = \hat{e}$, $\tilde{e}$, and $\check{e}$ yield error estimators $|\!|\!|\hat{e}|\!|\!|$, $|\!|\!|\tilde{e}|\!|\!|$, and $|\!|\!|\check{e}|\!|\!|$ which satisfy the inequalities

$$(1-\epsilon_1)|\!|\!|u_{\mathcal{H}} - U|\!|\!| \le |\!|\!|\bar{e}|\!|\!| \le (1+\epsilon_2)|\!|\!|u_{\mathcal{H}} - U|\!|\!| \tag{1.1}$$

where $\epsilon_1 < 1$ and ϵ_2 is bounded. When U is an arbitrary function in the finite element subspace, extra terms in (1.1) are introduced which measure the difference between U and the finite element solution. Such terms are of significance if the variational crimes are committed in assembling the finite element stiffness matrix and right-hand side, or if the resulting linear system is only approximately solved, say, by an iterative method.

Our theoretical results are for linear, elliptic, selfadjoint, positive-definite problems. The algorithms and many of our results extend readily to some nonselfadjoint, indefinite, and quasilinear elliptic problems. Essentially, the linear highest-order term dominates the process as the element size shrinks, so the lower-order nonlinear terms contribute only perturbations to the basic error bounds of Sections 4-6. Our results also extend with only small modification to the situation where homogeneous Dirichlet boundary conditions are specified on all or part of $\partial\Omega$. For elements which intersect the Dirichlet boundary, a local problem with homogeneous Dirichlet boundary conditions on the Dirichlet portion of the element boundary and Neumann conditions on the remaining edges is solved. The analysis of these local problems is easier than the pure Neumann case because the consistency condition need not be satisfied.

Our error indicators are based on local computations for efficiency reasons. That is, they are calculated using computations which involve only one or a few neighboring elements at a time. Assembling a global stiffness matrix and right-hand side, and solving the resulting linear system, would generally be more expensive than computing the original solution U. For example, if the finite element solution U is a continuous piecewise polynomial of degree r, a reasonable choice for an error indicator might be a discontinuous piecewise polynomial of degree $r+1$. Although the assumptions of our analysis, in particular, inequality (2.8), require the error indicators to be of higher degree than U, we have had practical success using continuous piecewise linear triangular basis functions for U and discontinuous piecewise linear basis functions for the error indicators.

We can prove a local analogue of the right-hand inequality of (1.1) under some circumstances, if we replace $|\!|\!|u_{\mathcal{H}} - U|\!|\!|_\tau$ with $|\!|\!|u_{\mathcal{H}} - U|\!|\!|_{N_\tau}$, where N_τ is a small set of elements in the neighborhood of triangle τ.

The remainder of the paper is organized as follows: Section 2 gives definitions and establishes notation. In Section 3, we give some preliminary results to be used

in the analyses of our error indicators $\hat{e}$, $\tilde{e}$, and $\check{e}$, which are analyzed in Sections 4, 5, and 6, respectively. In Section 7, we present numerical results indicating the behavior of our error indicators on several model problems.

For the error indicator $\hat{e}$ presented in Section 4, we show that ϵ_1 in (1.1) tends to zero for arbitrary U in the finite element space. However, we can bound ϵ_2 only when $u_{\mathcal{H}} - U$ exactly satisfies a subset (2.9) of the orthogonalities satisfied by the the finite element error, and in this case, our method for insuring a well-defined error indicator in an element requires some computations involving quantities from neighboring elements.

For the error indicator $\tilde{e}$ presented in Section 5, we can only prove that ϵ_1 in (1.1) tends to zero when $u_{\mathcal{H}} - U$ satisfies the orthogonality conditions (2.9). We can always bound ϵ_2, with somewhat stronger bounds if $u_{\mathcal{H}} - U$ satisfies (2.9). The computation of $\tilde{e}$ in an element requires only quantities from within that element.

For the error indicator $\check{e}$ presented in Section 6, we bound ϵ_1 and ϵ_2. The computation of $\check{e}$ requires only quantities from within that element, and requires even less work than the computation of $\tilde{e}$.

In Section 7, we find that, except on very coarse grids, on several model problems our error estimators are upper bounds on the norm of the true error, and that our estimators are always within a factor of three of the norm of the true error. The simplest and cheapest error indicator, $\check{e}$, also performs the best for fine grids, with its norm appearing to converge to the norm of the true error as the mesh size decreases.

2. Notation

Consider the linear, selfadjoint, positive definite Neumann problem

$$\begin{aligned} L(u) &= -\nabla a \nabla u + bu = f \quad && \text{in } \Omega \subset I\!R^2, \\ a\frac{\partial u}{\partial n} &= g && \text{on } \partial\Omega, \end{aligned} \tag{2.1}$$

with Ω a bounded domain, $a \in \mathcal{C}^1(\bar{\Omega})$, $b \in \mathcal{C}^0(\bar{\Omega})$. We assume there exist constants $\underline{a}$, $\bar{a}$, $\underline{b}$, $\bar{b}$ such that

$$\begin{aligned} 0 < \underline{a} \le a(x) \le \bar{a} \\ \text{for } x \in \bar{\Omega}. \\ 0 \le \underline{b} \le b(x) \le \bar{b} \end{aligned}$$

The boundary of Ω is assumed piecewise smooth. The weak form of (2.1) is: find $u \in \mathcal{H}^1(\Omega)$ such that

$$a(u,v) = (f,v) + \langle g, v \rangle \qquad \text{for all } v \in \mathcal{H}^1(\Omega), \tag{2.2}$$

where

$$a(u,v) = \int\!\!\int_\Omega (a\nabla u \nabla v + buv)\, dx, \qquad (f,v) = \int\!\!\int_\Omega fv\, dx, \qquad \langle g, v\rangle = \int_{\partial\Omega} gv\, ds.$$

$\mathcal{H}^k(\Omega)$ for nonnegative integer k will denote the usual Sobolev space equipped with the norm

$$\|u\|_k^2 = \sum_{|\beta| \le k} \|D^\beta u\|_0^2 = \sum_{|\beta| \le k} (D^\beta u, D^\beta u).$$

Negative and nonintegral spaces will not be used. We will set $\mathcal{H}^1(\Omega) = \mathcal{H}$ and $\|\cdot\|_0 = \|\cdot\|$. We will also use the energy norm associated with the bilinear form $a(\cdot,\cdot)$

$$|||u|||^2 = a(u,u).$$

We will consider the solution of (2.2) using a standard Rayleigh-Ritz-Galerkin procedure based on $\mathcal{C}^0$ triangular finite elements. Specifically, let $\mathcal{F} = \{\mathcal{T}\}$ denote a family of triangulations of Ω, where $\mathcal{T} \in \mathcal{F}$ is a collection of closed triangles such that, for distinct $\tau_1, \tau_2 \in \mathcal{T}$, $\tau_1 \cap \tau_2$ is either empty, a single vertex or a common edge. Triangles will normally have straight edges, although triangles with edges coincident with some part of $\partial\Omega$ will be allowed one curved boundary edge. Such an edge must be a smooth arc.

For $\tau \in \mathcal{T}$, let h_τ denote the diameter of τ and let

$$h = \max_{\tau\in\mathcal{T}} h_\tau.$$

Let $\mathcal{E}$ be the collection of curves which form an edge of a triangle in $\mathcal{T}$. The set of edges may be decomposed as the union of two disjoint sets $\mathcal{E}_B \cup \mathcal{E}_I$ where $\mathcal{E}_B$ is the set of boundary edges and $\mathcal{E}_I$ the set of (straight) interior edges. We denote by $\mathcal{E}_\tau$ the set of three edges of the triangle $\tau \in \mathcal{T}$. We define the neighbors of τ, $\mathcal{N}_\tau$ by

$$\mathcal{N}_\tau = \{\tau' \in \mathcal{T} \,|\, \tau' \cap \tau \neq \emptyset\}.$$

For $\varepsilon \in \mathcal{E}$, let h_ε denote the length of ε.

Let δ_0 be a fixed positive constant independent of $\mathcal{T}$ and $\mathcal{F}$. We require each triangle $\tau \in \mathcal{T}$ to be star-shaped with respect to a circle of diameter $\delta_0 h_\tau$ contained in τ. The constant δ_0 is a measure of the shape regularity of the triangles in $\mathcal{T}$. Shape regularity does not require the triangulations to be globally quasi-uniform, although it does imply a small angle condition and a local quasi-uniformity of the mesh. In particular, there exists a positive constant $\delta_1 = \delta_1(\delta_0)$ independent of $\mathcal{T}$ such that for $\varepsilon \in \mathcal{E}_\tau$, $\tau' \in \mathcal{N}_\tau$, and $\tau \in \mathcal{T}$,

$$\delta_1^{-1} h_\tau \leq h_\varepsilon \leq \delta_1 h_\tau.$$

For $\mathcal{T} \in \mathcal{F}$, let $\underline{\mathcal{S}} \subseteq \mathcal{S} \subset \bar{\mathcal{S}}$ denote three finite-dimensional spaces of $\mathcal{C}^0$ piecewise polynomials associated with $\mathcal{T}$. The space $\mathcal{S}$ is the space in which we will seek an approximate solution of (2.2), while $\bar{\mathcal{S}}$ will be associated with a larger space in which we seek an approximation to the error. The space $\underline{\mathcal{S}}$ will be a (possibly) smaller space, which contains at least the $\mathcal{C}^0$ piecewise linear functions on $\mathcal{T}$; we allow $\underline{\mathcal{S}} = \mathcal{S}$.

For $\mathcal{Q} = \mathcal{H}$, $\bar{\mathcal{S}}$, $\mathcal{S}$, or $\underline{\mathcal{S}}$, we define $u_\mathcal{Q} \in \mathcal{Q}$ by

$$a(u_\mathcal{Q}, v) = (f,v) + \langle g, v\rangle \qquad \text{for all } v \in \mathcal{Q}. \tag{2.3}$$

If $b \equiv 0$, we impose the additional requirement

$$(u_\mathcal{Q}, 1) = 0. \tag{2.4}$$

$u_\mathcal{H}$ is therefore the weak solution of (2.2), while $u_\mathcal{S}$, $u_{\bar{\mathcal{S}}}$, and $u_{\underline{\mathcal{S}}}$ are finite element approximations of $u_\mathcal{H}$. Let $U \in \mathcal{S}$ denote the computed approximation of $u_\mathcal{H}$, and define $e_\mathcal{Q} = u_\mathcal{Q} - U$, $\mathcal{Q} = \mathcal{H}$, $\mathcal{S}$, $\bar{\mathcal{S}}$. Normally one has $U - u_\mathcal{S}$ so that $e_\mathcal{S} = 0$, but our analysis does not require it. Thus our error estimates will allow for the inclusion of effects due to roundoff errors, variational crimes, or the approximate solution of the linear equations by an iterative method.

To facilitate the introduction of the local function spaces and inner products required for our analysis, let ω be an open set in $\mathbb{R}^2$ and γ a simple piecewise-smooth curve in $\mathbb{R}^2$. Let $(\cdot,\cdot)_\omega$ and $\|\cdot\|_\omega$ denote the $\mathcal{L}^2(\omega)$ inner product and norm, $a(\cdot,\cdot)_\omega$ and $|||\cdot|||_\omega$ denote the energy inner product and norm restricted to ω, and $\langle\cdot,\cdot\rangle_\gamma$ and $|\cdot|_\gamma$ denote the inner product and norm on $\mathcal{L}^2(\gamma)$.

Let $\mathcal{T}\in\mathcal{F}$ be a fixed triangulation and let

$$\mathcal{H}_\mathcal{T}=\prod_{\tau\in\mathcal{T}}\mathcal{H}^1(\tau)=\left\{\phi\,|\,\phi|_\tau\in\mathcal{H}^1(\tau),\,\tau\in\mathcal{T}\right\}$$

denote the space of piecewise $\mathcal{H}^1$ functions. For $v,w\in\mathcal{H}_\mathcal{T}$ we define broken $\mathcal{L}^2$ and energy inner products and norms by

$$(v,w)=\sum_{\tau\in\mathcal{T}}(v,w)_\tau,\qquad\qquad \|v\|_0^2=(v,v),$$
$$a(v,w)=\sum_{\tau\in\mathcal{T}}a(v,w)_\tau,\qquad\qquad |||v|||^2=a(v,v).$$

Note that $\mathcal{H}_\mathcal{T}\subset\mathcal{L}^2(\Omega)$, so the broken $\mathcal{L}^2$ inner product is just the usual $\mathcal{L}^2$ inner product. Also, $\mathcal{H}=\mathcal{H}^1(\Omega)\subset\mathcal{H}_\mathcal{T}$, and the above definitions reduce to the usual ones whenever $v,w,\in\mathcal{H}$. Similarly, for $\Gamma=\mathcal{E}$, $\mathcal{E}_I$, $\mathcal{E}_B$, $\mathcal{E}_\tau$, or some other subset of $\mathcal{E}$, let

$$\langle v,w\rangle_\Gamma=\sum_{\varepsilon\in\Gamma}\langle v,w\rangle_\varepsilon,\qquad\qquad |v|_\Gamma^2=\langle v,v\rangle_\Gamma.$$

Let $\mathcal{S}_\tau$ denote the restriction of $\mathcal{S}$ to $\tau\in\mathcal{T}$ and let

$$\mathcal{S}_\mathcal{T}=\prod_{\tau\in\mathcal{T}}\mathcal{S}_\tau.$$

Similarly, define $\bar{\mathcal{S}}_\tau$, $\underline{\mathcal{S}}_\tau$, $\bar{\mathcal{S}}_\mathcal{T}$, and $\underline{\mathcal{S}}_\mathcal{T}$. The space $\mathcal{S}_\mathcal{T}$ (and $\bar{\mathcal{S}}_\mathcal{T}$, $\underline{\mathcal{S}}_\mathcal{T}$) is a space of discontinuous piecewise polynomials locally defined in each element $\tau\in\mathcal{T}$. Note the inclusions $\underline{\mathcal{S}}_\mathcal{T}\subseteq\mathcal{S}_\mathcal{T}\subset\bar{\mathcal{S}}_\mathcal{T}\subset\mathcal{H}_\mathcal{T}$.

For each edge $\varepsilon\in\mathcal{E}$, we define a normal direction $n=n_\varepsilon$. If $\varepsilon\in\mathcal{E}_B$, this will be the usual outward normal. If $\varepsilon\in\mathcal{E}_I$, the choice is arbitrary. Since we are dealing with discontinuous spaces, it is useful to have notation describing the jump and average of functions along edges. We denote two triangles sharing an edge $\varepsilon\in\mathcal{E}_I$ as τ_{in} and τ_{out}, where the normal is outward from τ_{in} (see Figure 2.1). Then for x on ε,

$$[v]_J(x)=v(x)|_{\text{out}}-v(x)|_{\text{in}}$$

is the jump of v across ε, and

$$[v]_A(x)=\frac{1}{2}\left\{v(x)|_{\text{out}}+v(x)|_{\text{in}}\right\}$$

is the average of v on ε. Note that the quantity $[\partial v/\partial n]_J$ is independent of the direction of n.

Let $\mathcal{I}:\bar{\mathcal{S}}_\mathcal{T}\to\underline{\mathcal{S}}_\mathcal{T}$ denote a local polynomial interpolation operator. For example, for each $\tau\in\mathcal{T}$, we can let $\mathcal{I}$ be the usual Lagrange interpolant for triangular elements. We assume that $\mathcal{I}$ satisfies

(i) if $v\in\underline{\mathcal{S}}_\mathcal{T}$, $\mathcal{I}v=v$ ($\mathcal{I}$ behaves like the identity on $\underline{\mathcal{S}}_\mathcal{T}$: in particular, on piecewise constant functions);

(ii) if $v\in\bar{\mathcal{S}}$, $\mathcal{I}v\in\underline{\mathcal{S}}$ ($\mathcal{I}$ preserves continuity);

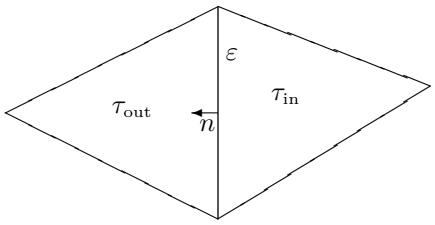

FIGURE 2.1

(iii) there exists a constant C_0, depending on δ_0, δ_1, a, b, the maximum degree of polynomials in $\mathcal{S}$, and the particular interpolation operator, but independent of h, such that

$$\sup_{v\in\bar{\mathcal{S}}_\tau, v\neq 0} \frac{|\!|\!| \mathcal{I}v |\!|\!|}{|\!|\!| v |\!|\!|} \leq C_0.$$

In Lemma 4 and Section 6, we will use the space

$$\check{\mathcal{S}}_\tau = \left\{ v \,|\, v \in \bar{\mathcal{S}}_\tau \text{ and } \mathcal{I}v = 0 \right\}$$

and the product space

$$\check{\mathcal{S}}_{\mathcal{T}} = \prod_{\tau\in\mathcal{T}} \check{\mathcal{S}}_\tau.$$

Note that any function $\bar{v} \in \bar{\mathcal{S}}_{\mathcal{T}}$ can be written uniquely as $\bar{v} = \underline{v} + \check{v}$, where $\underline{v} \in \underline{\mathcal{S}}_{\mathcal{T}}$ and $\check{v} \in \check{\mathcal{S}}_{\mathcal{T}}$.

Let $\tau \in \mathcal{T}$ and $\varepsilon \in \mathcal{E}_\tau$. In our analyses, we will use the following inequalities

$$|v|^2_{\mathcal{E}_\tau} \leq C_1^2 \left\{ h_\tau^{-1} \|v\|^2_\tau + h_\tau \|\nabla v\|^2_\tau \right\}, \qquad v \in \mathcal{H}^1(\tau), \tag{2.5}$$

$$\left| \frac{\partial v}{\partial n} \right|^2_{\mathcal{E}_\tau} \leq C_1^2 \left\{ h_\tau^{-1} \|\nabla v\|^2_\tau + h_\tau \|\nabla^2 v\|^2_\tau \right\}, \qquad v \in \mathcal{H}^2(\tau), \tag{2.6}$$

$$\|\nabla^p v\|_\tau \leq C_2 h_\tau^{q-p} \|\nabla^q v\|_\tau, \qquad v \in \bar{\mathcal{S}}_{\mathcal{T}},\ 0 \leq q \leq p \leq 2. \tag{2.7}$$

The trace inequalities (2.5)-(2.6) follow from results in Agmon [1]; the constant $C_1 = C_1(\delta_0, \delta_1)$. Inequality (2.7) is a local inverse inequality; the constant C_2 depends on δ_0 and on the maximum degree of polynomials contained in $\bar{\mathcal{S}}_{\mathcal{T}}$. Both C_1 and C_2 are independent of h.

We require some notion of convergence of the finite element solutions $u_{\mathcal{S}}$ and $u_{\bar{\mathcal{S}}}$ to the weak solution $u_{\mathcal{H}}$ of (2.2) as a function of h. In particular, we make the saturation assumption

$$|\!|\!| u_{\mathcal{H}} - u_{\bar{\mathcal{S}}} |\!|\!|^2 + \left| h_\varepsilon^{1/2} \left[a \frac{\partial (u_{\mathcal{H}} - u_{\bar{\mathcal{S}}})}{\partial n} \right]_A \right|^2_{\mathcal{E}_I} \leq \beta^2 |\!|\!| u_{\mathcal{H}} - u_{\mathcal{S}} |\!|\!|^2, \tag{2.8}$$

where $\beta = \beta(h)$ and $\lim_{h\to 0} \beta = 0$ (Babuška [2], Babuška and Luskin [3]). This assumption is reasonable, since $\bar{\mathcal{S}}$ contains polynomials of higher degree than $\mathcal{S}$.

Finally, some of our results will hold only when U satisfies

$$a(e_{\mathcal{S}}, \chi) = a(u_{\mathcal{S}} - U, \chi) = 0 \qquad \text{for all } \chi \in \underline{\mathcal{S}}, \tag{2.9}$$

i.e., the elliptic projection of the error on $\underline{\mathcal{S}}$ is zero. Note that (2.9) holds if U is the finite element solution $u_{\mathcal{S}}$ or if $U = u_{\underline{\mathcal{S}}}$. It may also hold when U is computed using certain two-level iterative schemes, where the coarse level corresponds to $\underline{\mathcal{S}}$. If the iteration terminates on a coarse subspace correction, then the approximate solution U will automatically satisfy (2.9).

3. Preliminary Results

In this section we present some results which are useful in analyzing the error estimation procedures to be discussed in Sections 4-6. Let $e_{\mathcal{Q}} = u_{\mathcal{Q}} - U$ for $\mathcal{Q} = \mathcal{H}$, $\underline{\mathcal{S}}$, $\mathcal{S}$, $\bar{\mathcal{S}}$. Then for $v \in \mathcal{Q}$, using (2.3) and integration by parts in each element,

$$\begin{aligned} a(e_{\mathcal{Q}}, v) &= (f, v) + \langle g, v\rangle - \sum_{\tau\in\mathcal{T}} \left\langle a\frac{\partial U}{\partial n_\tau}, v \right\rangle_{\mathcal{E}_\tau} - (L(U), v) \\ &= (r, v) + \langle r_B, v\rangle - \sum_{\tau\in\mathcal{T}} \left\langle a\frac{\partial U}{\partial n_\tau}, v \right\rangle_{\mathcal{E}_\tau\cap\mathcal{E}_I}. \end{aligned} \tag{3.1}$$

Here n_τ is the outward normal for τ, $r = f - L(U)$ is defined elementwise with possible discontinuities on $\mathcal{E}_I$, and $r_B = g - a\,\partial U/\partial n$ is defined edgewise on $\mathcal{E}_B$.

Noting that the outward normal for τ is the inward normal for a neighbor sharing a common edge, we may write the last term in (3.1) as

$$-\sum_{\tau\in\mathcal{T}} \left\langle a\frac{\partial U}{\partial n_\tau}, v \right\rangle_{\mathcal{E}_\tau\cap\mathcal{E}_I} = \left\langle \left[a\frac{\partial U}{\partial n}\right]_J, v \right\rangle_{\mathcal{E}_I}.$$

We define the local right-hand side of (3.1) by

$$F_\tau(v) = (r, v)_\tau + \langle r_B, v\rangle_{\mathcal{E}_\tau\cap\mathcal{E}_B} + \frac{1}{2}\left\langle \left[a\frac{\partial U}{\partial n}\right]_J, v \right\rangle_{\mathcal{E}_\tau\cap\mathcal{E}_I} \tag{3.2}$$

for $v \in \mathcal{Q}_\tau$. Summing over triangles, we obtain the linear functional $F(v)$, where

$$F(v) = \sum_{\tau\in\mathcal{T}} F_\tau(v) = (r, v) + \langle r_B, v\rangle + \left\langle \left[a\frac{\partial U}{\partial n}\right]_J, [v]_A \right\rangle_{\mathcal{E}_I} \tag{3.3}$$

for $v \in \mathcal{Q}_{\mathcal{T}}$.

In terms of this linear functional, (3.1) can be written as

$$a(e_{\mathcal{Q}}, v) = F(v) \qquad \text{for all } v \in \mathcal{Q}. \tag{3.4}$$

Equation(3.4) gives a useful characterization of the error. Noting $e_{\mathcal{H}} - e_{\mathcal{Q}} = u_{\mathcal{H}} - u_{\mathcal{Q}}$ ($\mathcal{Q} = \underline{\mathcal{S}}$, $\mathcal{S}$, $\bar{\mathcal{S}}$), (3.4) implies the orthogonality relations

$$a(e_{\mathcal{H}} - e_{\mathcal{Q}}, v) = 0 \qquad v \in \mathcal{Q} = \underline{\mathcal{S}},\, \mathcal{S},\, \bar{\mathcal{S}}, \tag{3.5}$$

normally associated with the error in the finite element method, as well as

$$a(e_{\bar{\mathcal{S}}} - e_{\mathcal{Q}}, v) = 0 \qquad v \in \mathcal{Q} = \underline{\mathcal{S}},\, \mathcal{S}, \tag{3.6}$$

$$a(e_{\mathcal{S}} - e_{\underline{\mathcal{S}}}, v) = 0 \qquad v \in \underline{\mathcal{S}}. \tag{3.7}$$

The orthogonality relations (3.5)-(3.7) imply

$$\|e_{\mathcal{H}}\|^2 = \|e_{\mathcal{Q}}\|^2 + \|e_{\mathcal{H}} - e_{\mathcal{Q}}\|^2, \qquad \mathcal{Q} = \underline{\mathcal{S}},\, \mathcal{S},\, \bar{\mathcal{S}}; \tag{3.8}$$

$$\|e_{\bar{\mathcal{S}}}\|^2 = \|e_{\mathcal{Q}}\|^2 + \|e_{\bar{\mathcal{S}}} - e_{\mathcal{Q}}\|^2, \qquad \mathcal{Q} = \underline{\mathcal{S}},\, \mathcal{S}; \tag{3.9}$$

$$\|e_{\mathcal{S}}\|^2 = \|e_{\underline{\mathcal{S}}}\|^2 + \|e_{\mathcal{S}} - e_{\underline{\mathcal{S}}}\|^2. \tag{3.10}$$

Equation (3.4) characterizes the inner product $a(e_-\mathcal{Q}, v)$ when $v \in \mathcal{Q}$. For $e_{\mathcal{H}}$, we also need to characterize the inner product $a(e_{\mathcal{H}}, v)$ for $v \in \mathcal{H}_{\mathcal{T}}$. To do so, we start from the elementwise version of (3.1). Suppose $e_{\mathcal{H}} \in \mathcal{H}^2(\tau)$; then, integrating by parts,

$$a(e_{\mathcal{H}}, v)_\tau = (r, v)_\tau + \langle r_B, v\rangle_{\mathcal{E}_\tau \cap \mathcal{E}_B} + \left\langle a\frac{\partial e_{\mathcal{H}}}{\partial n_\tau}, v \right\rangle_{\mathcal{E}_\tau \cap \mathcal{E}_I} \tag{3.11}$$

for all $v \in \mathcal{H}^1(\tau)$. In summing over elements, the third term is somewhat more complicated than in (3.1), since $[v]_J \neq 0$. In particular,

$$\sum_{\tau\in\mathcal{T}} \left\langle a\frac{\partial e_{\mathcal{H}}}{\partial n_\tau}, v \right\rangle_{\mathcal{E}_\tau \cap \mathcal{E}_I} = \left\langle \left[a\frac{\partial U}{\partial n}\right]_J, [v]_A \right\rangle_{\mathcal{E}_I} + \left\langle \left[a\frac{\partial e_{\mathcal{H}}}{\partial n}\right]_A, [v]_J \right\rangle_{\mathcal{E}_I}$$

This gives the global equation

$$a(e_{\mathcal{H}}, v) = F(v) + \left\langle \left[a\frac{\partial e_{\mathcal{H}}}{\partial n}\right]_A, [v]_J \right\rangle_{\mathcal{E}_I} \qquad \text{for all } v \in \mathcal{H}_{\mathcal{T}}. \tag{3.12}$$

Note that (3.12) reduces to (3.4) whenever $v \in \mathcal{H}$.

In the course of analysis in Sections 4-6, there are several basic estimates that will be used often. We summarize them in Lemmas 1-3.

Lemma 1. *Assume (2.6), (2.7), and (2.8) hold. Then*

$$\left| h_\varepsilon^{1/2} \left[a\frac{\partial e_{\mathcal{H}}}{\partial n}\right]_A \right|_{\mathcal{E}_I} \leq C_3 |\!|\!| e_{\mathcal{H}} |\!|\!|,$$

where $C_3 = C_3(\underline{a}, \bar{a}, C_1, C_2, \delta_1, \beta)$.

Proof. First note

$$\left| h_\varepsilon^{1/2} \left[a\frac{\partial e_{\mathcal{H}}}{\partial n}\right]_A \right|_{\mathcal{E}_I} \leq \left| h_\varepsilon^{1/2} \left[a\frac{\partial e_{\bar{\mathcal{S}}}}{\partial n}\right]_A \right|_{\mathcal{E}_I} + \left| h_\varepsilon^{1/2} \left[a\frac{\partial (u_{\mathcal{H}} - u_{\bar{\mathcal{S}}})}{\partial n}\right]_A \right|_{\mathcal{E}_I}.$$

The second term can be bounded using (2.8)

$$\left| h_\varepsilon^{1/2} \left[a\frac{\partial (u_{\mathcal{H}} - u_{\bar{\mathcal{S}}})}{\partial n}\right]_A \right|_{\mathcal{E}_I} \leq \beta |\!|\!| u_{\mathcal{H}} - u_{\mathcal{S}} |\!|\!| \leq \beta |\!|\!| e_{\mathcal{H}} |\!|\!|. \tag{3.13}$$

To bound the first term, let $\varepsilon \in \mathcal{E}_I$ and $\tau_{\text{in}} \cap \tau_{\text{out}} = \varepsilon$. Then

$$\left| h_\varepsilon^{1/2} \left[a\frac{\partial e_{\bar{\mathcal{S}}}}{\partial n}\right]_A \right|_\varepsilon^2 \leq \frac{h_\varepsilon}{2} \left\{ \left| a\frac{\partial e_{\bar{\mathcal{S}}}}{\partial n}|_{\tau_{\text{in}}} \right|_\varepsilon^2 + \left| a\frac{\partial e_{\bar{\mathcal{S}}}}{\partial n}|_{\tau_{\text{out}}} \right|_\varepsilon^2 \right\}.$$

Using (2.6)-(2.7) we obtain

$$\begin{aligned} h_\varepsilon \left| a\frac{\partial e_{\bar{\mathcal{S}}}}{\partial n}|_\tau \right|_\varepsilon^2 &\leq \bar{a}^2 C_1^2 h_\varepsilon \left\{ h_\tau^{-1} \|\nabla e_{\bar{\mathcal{S}}}\|_\tau^2 + h_\tau \|\nabla^2 e_{\bar{\mathcal{S}}}\|_\tau^2 \right\} \\ &\leq \bar{a}^2 C_1^2 \delta_1 (1 + C_2^2) \|\nabla e_{\mathcal{S}}\|_\tau^2 \\ &\leq \bar{a}^2 \underline{a}^{-1} C_1^2 \delta_1 (1 + C_2^2) |\!|\!| e_{\mathcal{S}} |\!|\!|_\tau^2 \end{aligned}$$

Summing over edges, and using the fact that each triangle has at most 3 edges, we obtain

$$\left| h_\varepsilon^{1/2} \left[a\frac{\partial e_{\bar{\mathcal{S}}}}{\partial n}\right]_A \right|_{\mathcal{E}_I}^2 \leq \frac{3}{2} \bar{a}^2 \underline{a}^{-1} C_1^2 \delta_1 (1 + C_2^2) |\!|\!| e_{\mathcal{S}} |\!|\!|^2. \tag{3.14}$$

Combining (3.13), (3.14), and (3.8) proves the lemma. □

Lemma 2. *Let $v \in \mathcal{H}_{\mathcal{T}}$, and suppose*

$$\|v\|_\tau \le C_4 h_\tau |\!|\!| v |\!|\!|_\tau \qquad \text{for all } \tau \in \mathcal{T},$$

where C_4 is independent of τ and v. Then

$$\left| h_\varepsilon^{1/2} [v]_J \right|_{\mathcal{E}_I} \le C_5 |\!|\!| v |\!|\!|,$$

where $C_5 = C_5(C_1, C_4, \underline{a}, \delta_1)$.

Proof. As in the proof of Lemma 1,

$$\left| \, [v]_J \right|_\varepsilon^2 \le 2 \left(\left| \, v|_{\tau_{\rm in}} \right|_\varepsilon^2 + \left| \, v|_{\tau_{\rm out}} \right|_\varepsilon^2 \right).$$

Using (2.5) and the hypothesis of the lemma, we obtain

$$\left| \, v|_\tau \right|_\varepsilon^2 \le C_1^2 \left(h_\tau^{-1} \|v\|_\tau^2 + h_\tau \|\nabla v\|_\tau^2 \right) \le C_1^2 (1 + C_4^2) \underline{a}^{-1} \delta_1 h_\varepsilon |\!|\!| v |\!|\!|_\tau^2.$$

Thus, summing over the edges,

$$\left| h_\varepsilon^{1/2} [v]_J \right|_{\mathcal{E}_I}^2 \le 6 C_1^2 (1 + C_4^2) \underline{a}^{-1} \delta_1 |\!|\!| v |\!|\!|^2,$$

and the lemma follows. □

Lemma 3. *Assume (2.8) holds and $\beta < 1$. Then*

$$(1 - \beta^2)^{1/2} |\!|\!| e_{\mathcal{H}} |\!|\!| \le |\!|\!| e_{\bar{\mathcal{S}}} |\!|\!|.$$

Proof. Using (3.8) for the cases $\mathcal{Q} = \mathcal{S}$ and $\mathcal{Q} = \bar{\mathcal{S}}$, and (2.8), we have

$$|\!|\!| e_{\mathcal{H}} |\!|\!|^2 \le \beta^2 |\!|\!| e_{\mathcal{H}} - e_{\mathcal{S}} |\!|\!|^2 + |\!|\!| e_{\bar{\mathcal{S}}} |\!|\!|^2 \le \beta^2 |\!|\!| e_{\mathcal{H}} |\!|\!|^2 + |\!|\!| e_{\bar{\mathcal{S}}} |\!|\!|^2,$$

and the result follows. □

In Section 6, we will use the following strengthened Cauchy-Schwarz inequality, proved in Bank and Dupont [12].

Lemma 4. *There exists a constant $\gamma < 1$, depending on δ_0, δ_1, a, b, the maximum degree of polynomials in $\bar{\mathcal{S}}$, and the particular choice of interpolation operator $\mathcal{I}$, but independent of h, such that*

$$|a(\underline{v}, \check{v})| \le \gamma \, |\!|\!| \underline{v} |\!|\!| \, |\!|\!| \check{v} |\!|\!|$$

for any $\underline{v} \in \underline{\mathcal{S}}$ and $\check{v} \in \check{\mathcal{S}}$.

4. An Error Estimator

Let $\theta \in \mathcal{L}^2(\mathcal{E})$ be a function defined on each edge ε with $\theta|_{\mathcal{E}_B} = 0$. If $\varepsilon = \tau_{\rm in} \cap \tau_{\rm out}$, define

$$\theta_{\tau_{\rm out}} = \theta, \qquad \theta_{\tau_{\rm in}} = -\theta.$$

so that

$$\sum_{\tau \in \mathcal{T}} \langle \theta_\tau, v \rangle_{\mathcal{E}_\tau} = \langle \theta, [v]_J \rangle_{\mathcal{E}_I}$$

for all $v \in \bar{\mathcal{S}}_{\mathcal{T}}$.

Given θ, we can formulate equations in each τ for the error indicator $\hat{e} \in \bar{\mathcal{S}}_{\mathcal{T}}$ as follows:

$$a(\hat{e}, v)_\tau = F_\tau(v) + \langle \theta_\tau, v \rangle_{\mathcal{E}_\tau \cap \mathcal{E}_I} \tag{4.1}$$

for all $v \in \bar{\mathcal{S}}_\tau$. For the moment assume θ is chosen so that (4.1) makes $\hat{e}$ well defined. Appropriate choices for θ will be discussed shortly. Summing over triangles, $\hat{e}$ satisfies the global equations

$$a(\hat{e}, v) = F(v) + \langle \theta, [v]_J \rangle_{\mathcal{E}_I} \tag{4.2}$$

for all $v \in \bar{\mathcal{S}}_{\mathcal{T}}$.

We immediately have the following:

Theorem 1. *If $\hat{e}$ exists,*

$$(1 - \beta^2)^{1/2} |\!|\!| e_{\mathcal{H}} |\!|\!| \le |\!|\!| \hat{e} |\!|\!|.$$

Proof. Since $e_{\bar{\mathcal{S}}} \in \bar{\mathcal{S}}$, $[e_{\bar{\mathcal{S}}}]_J \equiv 0$, and by (3.4) and (4.2),

$$a(e_{\bar{\mathcal{S}}}, e_{\bar{\mathcal{S}}}) = F(e_{\bar{\mathcal{S}}}) = a(\hat{e}, e_{\bar{\mathcal{S}}})$$

so $|\!|\!| e_{\bar{\mathcal{S}}} |\!|\!| \le |\!|\!| \hat{e} |\!|\!|$. The theorem now follows from Lemma 3. □

There are several ways to choose θ such that $\hat{e}$ is well defined. If $\underline{b} > 0$, we can simply choose $\theta \equiv 0$. Unfortunately, this may allow $|\!|\!| \hat{e} |\!|\!|$ to be much larger than $|\!|\!| e_{\mathcal{H}} |\!|\!|$. Another possibility is to let θ be an approximation to $[a\, \partial e_{\mathcal{H}}/\partial n]_A$ (see equation (4.8)).

If $b \equiv 0$, since $a(v, 1)_\tau = 0$ for any $v \in \bar{\mathcal{S}}_\tau$, we will require θ_τ to satisfy

$$\langle \theta_\tau, 1 \rangle_{\mathcal{E}_\tau \cap \mathcal{E}_I} = -F_\tau(1). \tag{4.3}$$

Then (4.1) will be consistent and $\hat{e}$ will be well-defined. Summing over an arbitrary set of triangles $\mathcal{T}^0$, (4.3) implies

$$\sum_{\tau \in \mathcal{T}^0} \langle \theta_\tau, 1 \rangle_{\mathcal{E}_\tau \cap \mathcal{E}_I} = - \sum_{\tau \in \mathcal{T}^0} F_\tau(1). \tag{4.4}$$

By (2.4) and (3.4), $F(1) = 0$, so (4.4) holds holds for $\mathcal{T}^0 = \mathcal{T}$. Now suppose $\mathcal{T}^1 = \mathcal{T}^2 \cup \mathcal{T}^3$ is a set of triangles, $\mathcal{T}^2 \cap \mathcal{T}^3 = \emptyset$, and we have chosen θ on $\partial\mathcal{T}^1$ such that (4.4) holds for $\mathcal{T}^0 = \mathcal{T}^1$. Then on $\partial\mathcal{T}^2 - \partial\mathcal{T}^1$, we can choose θ such that (4.4) holds for $\mathcal{T}^0 = \mathcal{T}^2$. For example, we could choose θ to be the constant

$$\theta_\tau|_{\partial\mathcal{T}^2 - \partial\mathcal{T}^1} = \frac{\sum_{\tau \in \mathcal{T}^2} F_\tau(1) - \langle \theta_\tau, 1 \rangle_{\partial\mathcal{T}^2 - \partial\mathcal{T}^1}}{\sum_{\varepsilon \in \partial\mathcal{T}^2 - \partial\mathcal{T}^1} h_\varepsilon}.$$

The best choice for $\theta_\tau|_{\partial\mathcal{T}^2 - \partial\mathcal{T}^1}$ is an open question. Then (4.4) automatically holds for $\mathcal{T}^0 = \mathcal{T}^3$. Continuing in the manner insures (4.4) holds for all triangles; always choosing $\mathcal{T}^2$ to be a single triangle allows local computation of θ. Unfortunately, this manner of choosing θ may also allow $|\!|\!| \hat{e} |\!|\!|$ to be much larger than $|\!|\!| e_{\mathcal{H}} |\!|\!|$.

We now show a way of choosing θ when (2.9) holds, which insures both that $\hat{e}$ is well-defined, and that $|\!|\!| \hat{e} |\!|\!|$ is not too much larger than $|\!|\!| e_{\mathcal{H}} |\!|\!|$. Let $\{x_i\}$ be the vertices of the triangulation $\mathcal{T}$ and let $\{b_i\}$ be the usual Lagrange basis for the $\mathcal{C}^0$ piecewise linear functions on $\mathcal{T}$, with $b_i(x_j) = \delta_{ij}$.

Recall $\underline{\mathcal{S}}$ contains $\{b_i\}$. We will choose θ to be a linear function on each edge ε, such that

$$\langle \theta_\tau, b_i \rangle_{\mathcal{E}_\tau \cap \mathcal{E}_I} = -F_\tau(b_i) \tag{4.5}$$

for all b_i and τ. In particular, since $\sum_i b_i = 1$, (4.3) will hold, and $\hat{e}$ will always be well-defined.

Let $\nu(\varepsilon, j)$ be an indexing function such that edge ε connects vertex $x_{\nu(\varepsilon,0)}$ and $x_{\nu(\varepsilon,1)}$, and define

$$\sigma_{\varepsilon,j} = \frac{4b_{\nu(\varepsilon,j)}(x) - 2b_{\nu(\varepsilon,1-j)}(x)}{h_\varepsilon}, \qquad j = 0, 1,$$

for $x \in \varepsilon$. By construction

$$\langle \sigma_{\varepsilon,j}, b_{\nu(\varepsilon,k)} \rangle_\varepsilon = \delta_{kj} \qquad j, k = 0, 1.$$

Expressing θ as a linear combination of the $\sigma_{\varepsilon,j}$'s,

$$\theta = \sum_{\varepsilon \in \mathcal{E}} \sum_{j=0,1} \theta_{\varepsilon,j} \sigma_{\varepsilon,j}.$$

Equations (4.5) decouple into sets of equations for each b_i,

$$\sum_{\nu(\varepsilon,j)=i} R_{\tau,\varepsilon} \theta_{\varepsilon,j} = -F_\tau(b_i)$$

for all $\tau \in \mathrm{supp}(b_i)$, where

$$R_{\tau,\varepsilon} = \begin{cases} 1 & \text{if } \tau = \tau_{\mathrm{out}}(e), \\ -1 & \text{if } \tau = \tau_{\mathrm{in}}(e). \end{cases}$$

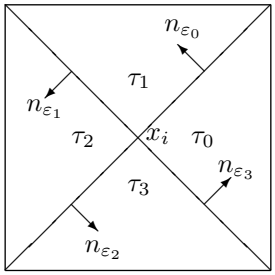

FIGURE 4.1

For example, for the vertex x_i depicted in Figure 4.1, the resulting system of linear equations is

$$\begin{bmatrix} 1 & 0 & 0 & -1 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} \theta_{\varepsilon_0,j_0} \\ \theta_{\varepsilon_1,j_1} \\ \theta_{\varepsilon_2,j_1} \\ \theta_{\varepsilon_3,j_3} \end{bmatrix} = \begin{bmatrix} -F_{\tau_0}(b_i) \\ -F_{\tau_1}(b_i) \\ -F_{\tau_2}(b_i) \\ -F_{\tau_3}(b_i) \end{bmatrix},$$

where $i = \nu(\varepsilon_0, j_0) = \nu(\varepsilon_1, j_1) = \nu(\varepsilon_2, j_2) = \nu(\varepsilon_3, j_3)$. The matrix has exactly one eigenvector $[1, 1, 1, 1]^t$ with eigenvalue zero. However since $b_i \in \underline{\mathcal{S}}$, $a(e_{\mathcal{H}}, b_i) = F(b_i) = 0$, and thus

$$-\sum_{k=0}^{3} F_{\tau_k}(b_i) = 0,$$

so the system is consistent. One solution of the system is

$$\theta_{\varepsilon_m,j_m} = -\sum_{k=0}^{m} F_{\tau_k}(b_i),$$

so that we can choose θ such that

$$|\theta_{\varepsilon,j}| \le \sum_{\tau_k \in \mathrm{supp}(b_i)} |F_{\tau_k}(b_i)| \tag{4.6}$$

where $i = \nu(\varepsilon, j)$.

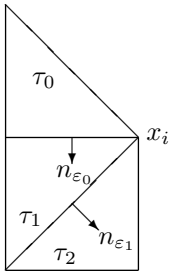

FIGURE 4.2

Inequality (4.6) holds in a like manner for all other coefficients, including the ones associated with boundary vertices. For example, for the vertex x_i depicted in Figure 4.2, the resulting system of linear equations is

$$\begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \theta_{\varepsilon_0,j_0} \\ \theta_{\varepsilon_1,j_1} \end{bmatrix} = \begin{bmatrix} -F_{\tau_0}(b_i) \\ -F_{\tau_1}(b_i) \\ -F_{\tau_2}(b_i) \end{bmatrix}.$$

Again, the system is consistent because $F(b_i) = 0$, and this time the unique solution is

$$\theta_{\varepsilon_m,j_m} = -\sum_{k=0}^{m} F_{\tau_k}(b_i).$$

With these choices for the coefficients of θ, we have

Theorem 2. *If $u_{\mathcal{H}} \in \mathcal{H}^2(\tau)$ for all τ,*

$$\left| h_{\varepsilon}^{1/2} \right|_{\mathcal{E}} \le C_8 |\!|\!| e_{\mathcal{H}} |\!|\!|,$$

where $C_8 = C_8(\bar{a}, \bar{b}, \delta_0, C_3)$.

Proof. On any edge ε, since $|\sigma_{\varepsilon,j}|_{\varepsilon} = 2h_{\epsilon}^{-1/2}$, by the definition of θ and (4.6)

$$\left| h_{\epsilon}^{1/2} \theta \right|_{\mathcal{E}_\tau} \le 4 \sum_{x_i \in \tau} \sum_{\tau' \in \mathrm{supp}(b_i)} |F_{\tau'}(b_i)|. \tag{4.7}$$

By (3.2) and (3.11), since

$$\begin{aligned} a \frac{\partial e_{\mathcal{H}}}{\partial n_\tau} &= \left[a \frac{\partial e_{\mathcal{H}}}{\partial n_\tau} \right]_A + \frac{1}{2} \left[a \frac{\partial U}{\partial n_\tau} \right]_J, \\ F_\tau(b_i) &= a(e_{\mathcal{H}}, b_i)_\tau - \left\langle \left[a \frac{\partial e_{\mathcal{H}}}{\partial n_\tau} \right]_A, b_i \right\rangle_{\mathcal{E}_\tau \cap \mathcal{E}_I}. \end{aligned}$$

Since

$$\left| b_i \right|_{\varepsilon} = \left(\frac{h_\varepsilon}{3} \right)^{1/2} \quad \text{and} \quad |\!|\!| b_i |\!|\!|_\tau \le C_6 \equiv \left(\frac{2\bar{a}}{\delta_0^2} + \bar{b} h^2 \right)^{1/2},$$

then

$$|F_\tau(b_i)| \le C_6 \|\!| e_{\mathcal{H}} |\!\|_\tau + 3^{-1/2} \left| h_\varepsilon^{1/2} \left[a \frac{\partial e_{\mathcal{H}}}{\partial n_\tau} \right]_A \right|_{\mathcal{E}_\tau \cap \mathcal{E}_I}.$$

By (4.7) and Lemma 1

$$\begin{aligned} \left| h_\varepsilon^{1/2} \theta \right|_{\mathcal{E}_I}^2 &= \frac{1}{2} \sum \tau \in \mathcal{T} \left| h_\varepsilon^{1/2} \theta \right|_{\mathcal{E}_\tau \cap \mathcal{E}_I}^2 \\ &\le \frac{1}{2} (12 C_7)^2 \sum_{\tau \in \mathcal{T}} \sup_{x_i \in \tau} |F_\tau(b_i)|^2 \\ &\le (12 C_7)^2 \left\{ C_6^2 \|\!| e_{\mathcal{H}} |\!\|^2 + \frac{1}{3} \left| h_\varepsilon^{1/2} \left[a \frac{\partial e_{\mathcal{H}}}{\partial n} \right]_A \right|_{\mathcal{E}_I}^2 \right\} \\ &\le (12 C_7)^2 (C_6^2 + C_3^2/3) \|\!| e_{\mathcal{H}} |\!\|^2, \end{aligned}$$

where $C_7 = C_7(\delta_1)$ is the maximum number of triangles sharing a vertex. □

We are now ready to bound $\|\!| \hat{e} |\!\|$.

Theorem 3. *If $u_{\mathcal{H}} \in \mathcal{H}^2(\tau)$ for all $\tau \in \mathcal{T}$, and θ is chosen such that (4.5) and Theorem 2 hold, then*

$$\|\!| \hat{e} |\!\| \le (1 + C_{\hat{e}}) \|\!| e_{\mathcal{H}} |\!\|,$$

where $C_{\hat{e}} \le (C_3 + C_8) |h_\varepsilon^{-1/2} [\hat{e}]_J|_{\mathcal{E}_I} / \|\!| \hat{e} |\!\| \le C_9$.

Proof. By (3.12) and (4.2), we have

$$a(\hat{e}, v) = a(e_{\mathcal{H}}, v) + \left\langle \theta - \left[a \frac{\partial e_{\mathcal{H}}}{\partial n} \right]_A [v]_J \right\rangle_{\mathcal{E}_I} \tag{4.8}$$

for all $v \in \mathcal{S}_{\mathcal{T}}$. Taking $v = \hat{e}$, we obtain, using Lemma 1 and Theorem 2,

$$\|\!| \hat{e} |\!\|^2 \le \|\!| e_{\mathcal{H}} |\!\| \, \|\!| \hat{e} |\!\| + C_{\hat{e}} \|\!| e_{\mathcal{H}} |\!\| \, \|\!| \hat{e} |\!\|.$$

To see that $C_{\hat{e}}$ is bounded, note that by (4.5), in each element τ, $a(\hat{e}, v)_\tau = 0$ for any linear polynomial v. A standard duality argument the implies

$$\|\hat{e}\|_\tau \le C_4(a, b, \delta_1) h_\tau \|\!| \hat{e} |\!\|_\tau.$$

Thus, by Lemma 2

$$C_{\hat{e}} \le (C_3 + C_8) C_5 \equiv C_9.$$

□

The term $C_{\hat{e}}$ measures a discontinuity of $\hat{e}$; in particular, if $\hat{e}$ is continuous, $C_{\hat{e}} = 0$ (i.e., $\hat{e} = e_{\bar{\mathcal{S}}}$). We could *force* $\hat{e}$ to converge to $e_{\bar{\mathcal{S}}}$ by penalizing its jumps in value across the interelement boundaries $\mathcal{E}_I$, but the penalties would necessarily destroy the local nature of the problems (4.1), and thus make the computation of $\hat{e}$ much more expensive.

5. A Second Error Estimator

One of the prime considerations in choosing the function θ in Section 4 was to make the problems (4.1) consistent. In this section, we present an alternative algorithm for computing error indicators $\tilde{e} \in \mathcal{S}_{\mathcal{T}}$ which avoids the computation of θ. The function $\tilde{e}$ is defined on $\tau \in \mathcal{T}$ as the solution of the local problem

$$a(\tilde{e}, v)_\tau = F_\tau(v - \mathcal{I}v) \qquad \text{for all } v \in \bar{\mathcal{S}}_\tau. \tag{5.1}$$

Since $1 = \mathcal{I}(1)$, Equations (5.1) are consistent. If $b \equiv 0$ on τ, we require

$$\iint_\tau \tilde{e}\, dx = 0$$

to insure a unique solution. Summing over τ, we obtain a global definition of $\tilde{e} \in \bar{\mathcal{S}}_{\mathcal{T}}$:

$$a(\tilde{e}, v) = F(v - \mathcal{I}v) \qquad \text{for all } v \in \bar{\mathcal{S}}_{\mathcal{T}}. \tag{5.2}$$

Theorem 1.

$$(1 - \beta^2)^{1/2} |\!|\!| e_{\mathcal{H}} |\!|\!| \le |\!|\!| \tilde{e} |\!|\!| + C_0 |\!|\!| e_{\mathcal{S}} |\!|\!|.$$

Proof. From (5.2) and (3.4) with $\mathcal{Q} = \bar{\mathcal{S}}$, $v = e_{\bar{\mathcal{S}}}$,

$$a(\tilde{e}, e_{\bar{\mathcal{S}}}) = a(e_{\bar{\mathcal{S}}}, e_{\bar{\mathcal{S}}} - \mathcal{I}e_{\bar{\mathcal{S}}}).$$

Using (3.6) and the fact that $\mathcal{I}e_{\bar{\mathcal{S}}} \in \underline{\mathcal{S}} \subseteq \mathcal{S}$,

$$\begin{aligned} |\!|\!| e_{\bar{\mathcal{S}}} |\!|\!|^2 &= a(e_{\bar{\mathcal{S}}}, e_{\bar{\mathcal{S}}}) \\ &= a(\tilde{e}, e_{\bar{\mathcal{S}}}) + a(e_{\bar{\mathcal{S}}}, \mathcal{I}e_{\bar{\mathcal{S}}}) \\ &\le |\!|\!| \tilde{e} |\!|\!|\, |\!|\!| e_{\bar{\mathcal{S}}} |\!|\!| + C_0 |\!|\!| e_{\mathcal{S}} |\!|\!|\, |\!|\!| e_{\bar{\mathcal{S}}} |\!|\!| \end{aligned} \tag{5.3}$$

Theorem 1 follows by dividing by $|\!|\!| e_{\bar{\mathcal{S}}} |\!|\!|$ and using Lemma 3. □

Theorem 2. *If $u_{\mathcal{H}} \in \mathcal{H}^2(\tau)$ for all $\tau \in \mathcal{T}$,*

$$|\!|\!| \tilde{e} |\!|\!| \le (1 + C_{\tilde{e}}) |\!|\!| e_{\mathcal{H}} |\!|\!| + C_0 |\!|\!| e_{\mathcal{S}} |\!|\!|,$$

where

$$C_{\tilde{e}} = [C_0 + C_3 C_5 (1 + C_0)] \inf_{\chi \in \bar{\mathcal{S}}} \frac{|\!|\!| \tilde{e} - \chi |\!|\!|}{|\!|\!| \tilde{e} |\!|\!|}.$$

Proof. From (3.12) and (5.2)

$$a(\tilde{e}, v) = a(e_{\mathcal{H}}, v - \mathcal{I}v) - \left\langle \left[a \frac{\partial e_{\mathcal{H}}}{\partial n} \right]_A, [v - \mathcal{I}v]_J \right\rangle_{\mathcal{E}_I}$$

for all $v \in \bar{\mathcal{S}}_{\mathcal{T}}$. Take $v = \tilde{e} - \phi$, where ϕ is the elliptic projection of $\tilde{e}$ onto $\bar{\mathcal{S}}$, so that

$$|\!|\!| \tilde{e} - \phi |\!|\!| = \inf_{\chi \in \bar{\mathcal{S}}} |\!|\!| \tilde{e} - \chi |\!|\!| \quad \text{and} \quad |\!|\!| \phi |\!|\!| \le |\!|\!| \tilde{e} |\!|\!|.$$

Noting that $a(\tilde{e}, \phi) = a(e_{\mathcal{H}}, \phi - \mathcal{I}\phi)$, we obtain

$$\begin{aligned} a(\tilde{e}, \tilde{e}) &= a(e_{\mathcal{H}}, \tilde{e}) - a(e_{\mathcal{H}}, \mathcal{I}(\tilde{e} - \phi)) - a(e_{\mathcal{H}}, \mathcal{I}\phi) \\ &\quad - \left\langle \left[a \frac{\partial e_{\mathcal{H}}}{\partial n} \right]_A, [\tilde{e} - \phi - \mathcal{I}(\tilde{e} - \phi)]_J \right\rangle_{\mathcal{E}_I}. \end{aligned} \tag{5.4}$$

Ww now bound the four terms on the right-hand side of (5.4). The first two are trivial:

$$|a(e_{\mathcal{H}}, \tilde{e})| \leq \|\!| e_{\mathcal{H}} |\!\| \, |\!\| \tilde{e} |\!\|; \tag{5.5}$$

$$|a(e_{\mathcal{H}}, \mathcal{I}(\tilde{e} - \phi))| \leq C_0 \|\!| e_{\mathcal{H}} |\!\| \, |\!\| \tilde{e} - \phi |\!\|. \tag{5.6}$$

For the third, note $\mathcal{I}\phi \in \underline{\mathcal{S}}$. Using (3.5),

$$|a(e_{\mathcal{H}}, \mathcal{I}\phi)| = |a(e_{\mathcal{S}}, \mathcal{I}\phi)| \leq C_0 \|\!| e_{\mathcal{S}} |\!\| \, |\!\| \tilde{e} |\!\|. \tag{5.7}$$

The fourth term uses Lemmas 1-2. Note that $\tilde{e} - \phi - \mathcal{I}(\tilde{e} - \phi)$ is zero at the interpolation points, so that the hypothesis of Lemma 2 is satisfied. Thus we obtain

$$\left| \left\langle \left[a \frac{\partial e_{\mathcal{H}}}{\partial n} \right]_A , [\tilde{e} - \phi - \mathcal{I}(\tilde{e} - \phi)]_J \right\rangle_{\mathcal{E}_I} \right| \leq C_3 C_5 (1 + C_0) \|\!| e_{\mathcal{H}} |\!\| \, |\!\| \tilde{e} - \phi |\!\|. \tag{5.8}$$

Theorem 2 follows from (5.4)-(5.8). □

If $e_{\mathcal{S}}$ satisfies (2.9), the bounds are somewhat stronger.

Theorem 3. *If (2.9) is satisfied and $u_{\mathcal{H}} \in \mathcal{H}^2(\tau)$ for all $\tau \in \mathcal{T}$, then*

$$(1 - \beta^2)^{1/2} \|\!| e_{\mathcal{H}} |\!\| \leq \|\!| \tilde{e} |\!\| \leq (1 + C_{\tilde{e}}) \|\!| e_{\mathcal{H}} |\!\|.$$

Proof. In this case, the last term in (5.3) and the term bounded in (5.7) are both zero. □

In analogy with Theorem 3, $C_{\tilde{e}}$ measures the discontinuity of $\tilde{e}$; if $\tilde{e}$ is continuous, $C_{\tilde{e}} = 0$. Except for the global use of (2.8), we could prove a local analogue of Theorem 2 for each element.

6. A Third Error Estimator

Equations (5.1) define an error indicator $\tilde{e}$ on τ as the solution of a linear system with size equal to $\dim(\bar{\mathcal{S}}_\tau)$, the dimension of $\bar{\mathcal{S}}_\tau$. In this section, we present an alternative algorithm for computing an error indicator $\check{e}$ on τ which is the solution of a smaller linear system which is automatically positive definite.

The error indicator $\check{e} \in \check{\mathcal{S}}$, defined on $\tau \in \mathcal{T}$ as the solution of the local problem

$$a(\check{e}, v)_\tau = F_\tau(v) \qquad \text{for all } v \in \check{\mathcal{S}}_\tau. \tag{6.1}$$

Since the constant function $1 \notin \check{\mathcal{S}}_\tau$, the linear system (6.1) is positive definite. Note that $\dim(\check{\mathcal{S}}_\tau) = \dim(\bar{\mathcal{S}}_\tau) - \dim(\underline{\mathcal{S}}_\tau)$. By the definition of $\check{\mathcal{S}}_\tau$ and (5.1)

$$a(\check{e}, v)_\tau = F_\tau(v - \mathcal{I}v) = a(\tilde{e}, v) \qquad \text{for all } v \in \check{\mathcal{S}}_\tau. \tag{6.2}$$

Equation (6.2) shows that $\check{e}$ is the elliptic projection of $\tilde{e}$ into $\check{\mathcal{S}}_\tau$.

Theorem 1.

$$(1 - \gamma^2)^{1/2} \|\!| \check{e} |\!\| \leq \|\!| \tilde{e} |\!\|.$$

Proof. Let $\tilde{e} = \tilde{e}_1 + \tilde{e}_2$, where $\tilde{e}_1 \in \underline{\mathcal{S}}_{\mathcal{T}}$ and $\tilde{e}_2 \in \check{\mathcal{S}}_{\mathcal{T}}$. By Lemma 4

$$\|\!| \tilde{e} |\!\|^2 = a(\tilde{e}_1, \tilde{e}_1) + 2a(\tilde{e}_1, \tilde{e}_2) + a(\tilde{e}_2, \tilde{e}_2) \geq (1 - \gamma^2) \|\!| \tilde{e}_2 |\!\|^2.$$

Since $\tilde{e}_1 = \mathcal{I}\tilde{e}_1$, $a(\tilde{e}, \tilde{e}_1) = 0$. Then by (6.2), with $v = \tilde{e}_2$,

$$\|\!| \tilde{e} |\!\|^2 = a(\tilde{e}, \tilde{e}_1) + a(\tilde{e}, \tilde{e}_2) = a(\tilde{e}, \tilde{e}_2) \leq \|\!| \check{e} |\!\| \, |\!\| \tilde{e_2} |\!\|,$$

and the theorem follows. □

By (6.2) with $v = \check{e}$, we immediately have

Theorem 2.

$$|||\check{e}||| \leq |||\tilde{e}|||.$$

Using Theorems 1-3 and 1-2, we obtain

Theorem 3.

$$(1-\beta^2)^{1/2}|||e_{\mathcal{H}}||| \leq (1-\gamma^2)^{-1/2}|||\check{e}||| + C_0|||e_{\mathcal{S}}|||.$$

If $u_{\mathcal{H}} \in \mathcal{H}^2(\tau)$ for all $\tau \in \mathcal{T}$,

$$|||\check{e}||| \leq (1+C_{\check{e}})|||e_{\mathcal{H}}||| + C_0|||e_{\mathcal{S}}|||.$$

If in addition (2.9) is satisfied, then

$$\{(1-\beta^2)(1-\gamma^2)\}^{1/2}|||e_{\mathcal{H}}||| \leq |||\check{e}||| \leq (1+C_{\check{e}})|||e_{\mathcal{H}}|||.$$

7. Numerical Results

In this section, we present some example calculations comparing the error estimators $|||\hat{e}|||$, $|||\tilde{e}|||$, and $|||\check{e}|||$, described in Section 4, 5, and 6, respectively. In these calculations, $\underline{\mathcal{S}} = \mathcal{S}$ is the space of $\mathcal{C}^0$ piecewise linear triangular finite elements, and $\bar{\mathcal{S}}$ is the space of $\mathcal{C}^0$ piecewise quadratic elements. The test problems are of the form

$$\begin{aligned} -\Delta u &= 0 && \text{in } \Omega, \\ u &= 0 && \text{on } \partial\Omega_1, \\ \frac{\partial u}{\partial n} &= \frac{2}{k}\cos\left(\frac{2\theta}{k}\right) && \text{on } \partial\Omega_2, \\ \frac{\partial u}{\partial n} &= 0 && \text{on } \partial\Omega_3, \end{aligned} \tag{7.1}$$

where $\Omega = \{(r,\theta) : 0 < r < 1, 0 < \theta < k\pi/4\}$; $\partial\Omega_1$ is the line $\theta = 0$, $0 \leq r \leq 1$; $\partial\Omega_2$ is the are $0 \leq \theta \leq k\pi/4$, $r = 1$; $\partial\Omega_3$ is the line $\theta = k\pi/4$, $0 \leq r \leq 1$; (r,θ) are polar coordinates; and $k = 1, 3, 4$ and 8. The solution of (7.1) is

$$u_{\mathcal{H}} = r^{2/k}\sin\left(\frac{2\theta}{k}\right), \tag{7.2}$$

the leading term of the point singularity for a corner with interior angle $k\pi/4$, so that $u_{\mathcal{H}} \in \mathcal{H}^{1+2/k-\epsilon}(\Omega)$ for any $\epsilon > 0$, but not for $\epsilon = 0$, when $k = 3, 4$, or 8 [15]. When $k = 3, 4$, or 8, the analogues to Theorems 2, 3, 2, 3, and 3 must use weaker bounds on $|||\bar{e}|||_\tau$ in elements touching the origin, and count on adaptive mesh refinement to make those elements very small.

The problems were solved using the Fortran package *PLTMG* [11]. This code used local adaptive mesh refinement and created for each problem a sequence of meshes of varying degrees of nonuniformity. The meshes for the case $k = 8$ are illustrated in Figure 7.1. *PLTMG* uses a multilevel iterative method for solving the linear systems and typically generates solutions $U \in \mathcal{S}$ such that $|||e_{\mathcal{S}}||| = |||U - u_{\mathcal{S}}|||$ is somewhat less than the discretization error.

For each problem and each mesh, we computed the quantities $|||e_{\mathcal{H}}|||$, $|||\hat{e}|||$, $|||\tilde{e}|||$, and $|||\check{e}|||$. For $|||e_{\mathcal{H}}|||$, a numerical quadrature rule using six quadrature points per element was used. $\hat{e}$, $\tilde{e}$, and $\check{e}$ were all piecewise quadratic polynomials and their norms were computed exactly, except for small errors for elements with curved boundary edges.

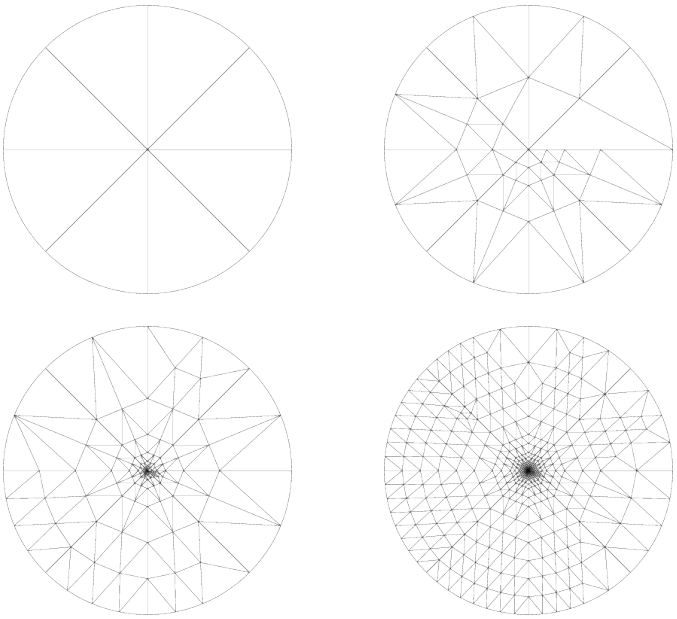

FIGURE 7.1

The effectivity index [8] or efficiency index [9] for an error estimator $\bar{e}$, $\bar{e} = \hat{e}, \tilde{e}$, or $\check{e}$ is defined as

$$\text{eff}(\bar{e}) = \frac{\|\bar{e}\|}{\|e_{\mathcal{H}}\|}. \tag{7.3}$$

We chose to measure the relative error

$$\rho(\bar{e}) = \text{eff}(\bar{e}) - 1. \tag{7.4}$$

Note that if $\bar{e}$ satisfies (1.1), we have $-\epsilon_1 \leq \rho(\bar{e}) \leq \epsilon_2$.

Having $\rho(\bar{e})$ near or converging to zero is clearly the most desirable situation. Positive values of ρ indicate an overestimate of the true error and are acceptable as long as ρ is not too much larger than one. Negative values of ρ mean the error estimator has given an erroneously optimistic values of $\|e_{\mathcal{H}}\|$. In Table 7.1, we tabulate the values of $\rho(\bar{e})$ for each error indicator on each mesh for each test case. NV is the number of vertices in the mesh.

On the basis of Theorems 1, 1, and 3, we expect $\rho(\hat{e}) \geq 0$, $\rho(\tilde{e}) \geq 0$, and $\rho(\check{e}) \geq (1-\gamma^2)^{1/2} - 1$, since once the initial mesh is appreciable refined, the effects of β and $\|e_{\mathcal{S}}\|$ are small. For all error indicators, we have $\rho(\bar{e}) \leq \epsilon_2$, where the data suggest that $\epsilon_2 \leq 3$ for this problem class.

The apparent convergence of $\rho(\check{e})$ to zero represents a particularly nice state of affairs, not only because it means $\|\check{e}\|$ is quite accurate, but also because $\check{e}$ is the least costly of the three indicators to compute.

The calculation of $\hat{e}$ is the most expensive since it involves solving a linear system like (4.5) for each vertex in order to obtain θ. We must then assemble and solve a symmetric positive semidefinite 6×6 linear system (of rank 5) in each element. The computation of $\tilde{e}$ also involves assembling and solving a 6×6 linear system in each element, but the calculation of θ is avoided. For $\check{e}$, only a 3×3 symmetric,

k	NV	$\|\|\|e_{\mathcal{H}}\|\|\|$	$\rho(\hat{e})$	$\rho(\tilde{e})$	$\rho(\check{e})$
1	3	.428	.247	.150	.121
	15	.129	.508	.385	.0117
	45	.0653	.681	.430	.00366
	153	.0328	.974	.452	.00131
	561	.0164	1.44	.463	.000480
3	5	.241	.408	.152	-.0960
	17	.123	.520	.335	.0591
	75	.0522	.998	.625	.196
	137	.0379	1.24	.697	.255
	480	.0200	1.50	.589	.154
	1908	.00946	2.23	.579	.110
4	6	.368	.618	.0166	-.199
	27	.181	.579	.335	.0280
	119	.0889	.695	.522	.135
	501	.0389	.842	.588	.148
	1363	.0225	.959	.541	.0963
8	10	.499	1.46	-.230	-.3903
	41	.368	1.87	.0545	-.197
	161	.208	1.22	.407	.0645
	681	.120	1.32	.378	.0263

TABLE 7.1

positive definite linear system (corresponding to the three edge midpoint Lagrange basis function) is assembled and solved.

From other tests using square elements [16], we have observed apparent convergence of $\rho(\check{e})$ and $\rho(\tilde{e})$ to zero for some test problems. Under suitable assumptions, Babuška and Miller [4] have recently proved convergence of the error estimators used in their two-dimensional code (with piecewise bilinear basis functions) to the norm of the true error.

For problems in one space dimension, when $\mathcal{S}$ consists of piecewise linear functions, the formulation leading to $\hat{e}$ and $\tilde{e}$ are equivalent. When, in addition, $\bar{\mathcal{S}}$ consists of piecewise quadratic functions, the analogue of $\check{e}$ is only slightly different from an error indicator of Babuška and Rheinboldt ($\|\|\check{e}\|\|_\tau$ is analogous to $\hat{\eta}_j(\Delta)$ in [9]), When in addition $a(x) = 1$ and $b(x) = 0$, all three indicators are equivalent. Under suitable assumptions, it can be shown that the corresponding one-dimensional error estimators converge to the norm of the true error.

Acknowledgments. This work was initiated while the first author was a summer professor at Exxon Production Research Company. We are grateful to the management of EPR for permission to publish this paper, and to Linda Scott and Marie Mason of EPR for help in preparation of the manuscript.

This research was supported in part by the Office of Naval Research through contracts N00014-82-K-0197 (University of California at San Diego) and N00014-76-C-0277 (Yale University).

Department of Mathematics
University of California at San Diego

La Jolla, California 92093

Exxon Production Research Company
Houston, Texas 77001

References

[1] S. Agmon, *Lectures on Elliptic Boundary Value Problems*, Van Nostrand, New York, 1965.
[2] I. Babuška. private communication.
[3] I. Babuška and M. Luskin, *An adaptive time discretization procedure for parabolic problems*, in Proc. Fourth IMACS Inter. Sympos. on Computer Methods for Partial Differential Equations, Lehiegh University, Bethlehem, Pennsylvania, 1981, pp. 5–8.
[4] I. Babuška and A. Miller, *A posteriori error esitmates and adaptive techniques for the finite element method*, Tech. Rep. BN-968, Institute for Physical Science and Technology, University of Maryland, 1981.
[5] I. Babuška and W. C. Rheinboldt, *Error estimates for aadptive finite element computations*, SIAM J. Numer. Anal., 15 (1978), pp. 736–754.
[6] ———, *A posteriori error error estimates for the finite element method*, Internat. J. Numer. Methods Engrg., 12 (1978), pp. 1597–1615.
[7] ———, *Analysis of optimal finite element meshes in* $I\!R^1$, Math. Comp., 33 (1979), pp. 435–463.
[8] ———, *On the reliability and optimality of the finite element method*, Comput. & Structures, 10 (1979), pp. 87–94.
[9] ———, *Reliable error estimation and mesh adaptation for the finite element method*, in Computational Methods in Nonlinear Mechanics, North-Holland, New York, 1980, pp. 67–108.
[10] ———, *A posteriori error analysis of finite element solutions for one-dimensional problems*, SIAM J. Numer. Anal., 18 (1981), pp. 565–589.
[11] R. E. Bank, *PLTMG user's guide, June 1981 version*, tech. rep., Department of Mathematics, University of California at San Diego, 1982.
[12] R. E. Bank and T. F. Dupont, *Analysis of a two level scheme for solving finite element equations*, Tech. Rep. CNA-159, Center for Numerical Analysis, University of Texas at Austin, 1980.
[13] R. E. Bank and A. H. Sherman, *A multilevel iterative method for solving finite element equations*, in Proc. Fifth Sympos. on Reservoir Engineering, Society of Petroleum Engineers of AIME, Dallas, 1979, pp. 117–126.
[14] P. Percell and M. F. Wheeler, *A local residual finite element procedure for elliptic equations*, SIAM J. Numer. Anal., 15 (1978), pp. 705–714.
[15] G. Strang and G. Fix, *An Analysis of the Finite Element Method*, Prentice-Hall, Englewood Cliffs, New Jersey, 1973.
[16] A. Weiser, *Local-mesh, local-order, adaptive finite element methods with a posteriori error estimators*, Tech. Rep. 213, Computer Science Department, Yale University, 1981.