Numer. Math. 65, 23–50 (1993)

Numerische Mathematik

A unified approach to a posteriori error estimation using element residual methods

Mark Ainsworth* and J. Tinsley Oden

Texas Institute for Computational Mechanics, The University of Texas at Austin, Austin, TX 78712, USA

Received September 11, 1991 / Revised version received June 24, 1992

Summary. This paper deals with the problem of obtaining numerical estimates of the accuracy of approximations to solutions of elliptic partial differential equations. It is shown that, by solving appropriate local residual type problems, one can obtain upper bounds on the error in the energy norm. Moreover, in the special case of adaptive *h-p* finite element analysis, the estimator will also give a realistic estimate of the error. A key feature of this is the development of a systematic approach to the determination of boundary conditions for the local problems. The work extends and combines several existing methods to the case of full *h-p* finite element approximation on possibly irregular meshes with elements of non-uniform degree. As a special case, the analysis proves a conjecture made by Bank and Weiser [Some A Posteriori Error Estimators for Elliptic Partial Differential Equations, Math. Comput. **44**, 283–301 (1985)].

Mathematics Subject Classification (1991): 65N30

1. Introduction

In this work we address the problem of computing *a posteriori* error estimates for approximations to elliptic boundary value problems. Although we have in mind adaptive *h-p* finite element computations, the analysis also includes various other types of approximation.

The error estimates are based on local residual problems similar in type to those discussed in [5, 6, 11, 12]. There are, however, significant differences in the approach.

Although the boundary value problem which we are approximating may be associated with an operator of the form

$$Lu \equiv -\nabla \cdot (a \nabla u) + cu$$

* On leave from Mathematics Department, Leicester University, Leicester LE1 7RU, UK.
Correspondence to: J.T. Oden

the local residual problem will always be Poisson's equation. This means that the error estimation analysis can be done independently of the analysis of the particular form of the operator L. More importantly, error analysis routines can be developed which exploit properties of the Laplace operator to develop a very efficient routine for solving the local problems. The first main result is to show that the error estimator always overestimates the true error. Essentially, we do not need to make any further regularity assumptions on the true solution other than $u \in H^1(\Omega)$.

In order to show that the estimator does not give an unduly pessimistic estimate, further assumptions are made on the source of the approximation u_h. In particular, we assume u_h is obtained by an *h-p* adaptive finite element computation, possibly on irregular meshes with elements of differing shapes and non-uniform polynomial degree. A preliminary analysis reveals that the estimate can become very pessimistic unless the boundary conditions for the local problems are chosen carefully.

The question of boundary conditions is examined in detail with the result that a scheme proposed by Bank and Weiser [5] for piecewise affine approximation on triangular elements is extended to the case of full *h-p* approximation on irregular meshes. It is of interest to notè that in determining the boundary conditions one works on the same element patches which occur in the related works of Babuška et al. [3, 4]. As special cases of our results, we obtain the result conjectured in Bank and Weiser [5] that a certain error estimator always overestimates the true error, and provide theoretical support for the heuristic results of Kelly [9].

2. Notations and preliminaries

Let Ω denote an open bounded Lipschitzian domain in $\mathbb{R}^2$ with a piecewise smooth boundary $\partial\Omega$. The boundary consists of a finite number of smooth arcs meeting with internal angle $\theta \in (0, 2\pi)$.

The Sobolev space $H^m(\Omega)$, $m \in \mathbb{Z}^+$, is a Hilbert space defined as the completion of $C^\infty(\Omega)$ in the Sobolev norm

$$\|u\|_{m,\Omega} = \left\{ \sum_{|\alpha| \leq m} \int_\Omega |D^\alpha u|^2 \, dx \right\}^{1/2} \tag{2.1}$$

where $\alpha = (\alpha_1, \alpha_2)$, $\alpha_i \in \mathbb{Z}^+$, $|\alpha| = \alpha_1 + \alpha_2$ and

$$D^\alpha u = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2}} \tag{2.2}$$

is the distributional derivative. $H^m(\Omega)$ is equipped with the inner product

$$(u, v)_{m,\Omega} = \sum_{|\alpha| \leq m} \int_\Omega D^\alpha u \cdot D^\alpha v \, dx \tag{2.3}$$

We use the notation $H^0(\Omega) = L_2(\Omega)$ in the case $m = 0$. Let $\mathscr{P}$ be a partition of Ω into a collection of $N = N(\mathscr{P})$ subdomains Ω_K with boundaries $\partial\Omega_K$, $1 \leq K \leq N$, such that

(i) $N(\mathscr{P}) < \infty$

(ii) $\bar{\Omega} = \bigcup_{K=1}^{N} \bar{\Omega}_K$, $\Omega_K \cap \Omega_L = \emptyset$, $K \neq L$

(iii) Ω_K are Lipschitzian with piecewise smooth boundaries $\partial\Omega_K$
(iv) $\Gamma_{KL} = \partial\Omega_K \cap \partial\Omega_L$, $1 \leqq K, L \leqq N$ are sets consisting of a finite number $\rho(K, L)$ of components such that

$$\Gamma_{KL} = \bigcup_{M=1}^{\rho(K,L)} \Gamma_{KL}^M \cup S_{KL}$$

where Γ_{KL}^M are smooth arcs and S_{KL} is a set of isolated points such that

$$\bigcup_{M=1}^{\rho(K,L)} \Gamma_{KL}^M \cap S_{KL} = \varnothing \quad 0 \leqq K, L \leqq N$$

We set $\Gamma_{0K} = \partial\Omega_K \cap \partial\Omega$. Notice that

$$\partial\Omega_K = \bigcup_{L,M} \bar{\Gamma}_{KL}^M$$

With these notations, it is possible to unambiguously characterize the boundary segments of the partition as

$$E = E(\mathscr{P}) = \bigcup_{\substack{K,L=0 \\ K>L \\ 1 \leqq M \leqq \rho(K,L)}}^{N} \Gamma_{KL}^M \tag{2.4}$$

The boundary segments lying on the interior of Ω are denoted by

$$E_I = E_I(\mathscr{P}) = \bigcup_{\substack{K,L=1 \\ K>L \\ 1 \leqq M \leqq \rho(K,L)}}^{N} \Gamma_{KL}^M \tag{2.5}$$

The outward pointing unit normal vector on Ω_K is denoted by $\boldsymbol{n}_K$. Let

$$\sigma_{KL} = -\sigma_{LK} = \begin{cases} +1, & K > L \\ -1, & K < L \end{cases}$$

and define $\boldsymbol{n}(s) = \sigma_{KL}\boldsymbol{n}_K(s) = \sigma_{LK}\boldsymbol{n}_L(s)$, $s \in \Gamma_{KL}^M$. That is, $\boldsymbol{n}$ points outward from the subdomain with the largest index.

Throughout, if v is some function defined on Ω, then its restriction to Ω_K is denoted by

$$v_K \equiv v|_{\Omega_K}, \quad 1 \leqq K \leqq N \tag{2.6}$$

In addition to the global Sobolev spaces and norms above, we introduce the broken Sobolev spaces $H^m(\mathscr{P})$:

$$H^m(\mathscr{P}) = \{v \in L_2(\Omega)\colon v_K \in H^m(\Omega_K);\ 1 \leqq K \leqq N\} \tag{2.7}$$

equipped with the norm

$$\| v \|_{m,\mathscr{P}} = \left\{ \sum_{K=1}^{N} \| v_K \|_{m,\Omega_K}^2 \right\}^{1/2} \tag{2.8}$$

Evidently $H^m(\Omega) \subset H^m(\mathscr{P})$ for $m \in \mathbb{N}$ and

$$H^0(\mathscr{P}) = L_2(\mathscr{P}) = L_2(\Omega) = H^0(\Omega) \tag{2.9}$$

Let $L_2(E)$ denote the space of classes of square integrable functions defined on E with the norm

$$\|\varphi\|_{0,E} = \left\{ \sum_{\substack{K,L=0 \\ K>L \\ 1\leq M\leq \rho(K,L)}}^{N} \int_{\Gamma_{KL}^{M}} |\varphi_{KL}^{M}|^2 \, ds \right\}^{1/2} \tag{2.10}$$

and inner product

$$(\varphi, \chi)_{0,E} = \sum_{\substack{K,L=0 \\ K>L \\ 1\leq M\leq \rho(K,L)}}^{N} \int_{\Gamma_{KL}^{M}} \varphi_{KL}^{M} \cdot \chi_{KL}^{M} \, ds \tag{2.11}$$

where $\varphi_{KL}^{M} = \varphi|_{\Gamma_{KL}^{M}}$ is the restriction of φ to Γ_{KL}^{M}. Analogous inner product and norm are defined on $L_2(E_I)$.

Throughout let γ denote the various trace operators associated with mappings to the boundaries or segments of boundaries [1]. For any curve Γ, the space $H^{1/2}(\Gamma)$ is the completion of $C^{\infty}(\Gamma)$ in the norm on $H^{\frac{1}{2}}(\Gamma)$ given by

$$|w|_{1/2,\Gamma}^2 = \|w\|_{1/2,\Gamma}^2 + \|w\|_{0,\Gamma}^2 \tag{2.12}$$

where

$$|w|_{1/2,\Gamma}^2 = \int_\Gamma \int_\Gamma \frac{|w(\boldsymbol{x}(s)) - w(\boldsymbol{x}(t))|^2}{|\boldsymbol{x}(s) - \boldsymbol{x}(t)|^2} \, ds \, dt \tag{2.13}$$

It is well known that $\gamma \in \mathscr{L}(H^1(\Omega_K), H^{\frac{1}{2}}(\partial\Omega_K))$ and is surjective. Finally, we denote the space of continuous linear functionals on $H^{1/2}(\Gamma)$ by $H^{-1/2}(\Gamma)$.

3. Model problem

Consider the following boundary-value problem for given data $f \in L_2(\Omega)$ and $g \in L_2(\Gamma_N)$:

$$\left.\begin{aligned} Lu &= f \quad \text{in } \Omega \\ \gamma u &= 0 \quad \text{on } \Gamma_D \\ a\frac{\partial u}{\partial n} &= g \quad \text{on } \Gamma_N \end{aligned}\right\} \tag{3.1}$$

where

$$Lu \equiv -\nabla \cdot (a \nabla u) + cu \tag{3.2}$$

and

$$\Gamma_N \cap \Gamma_D = \varnothing, \quad \bar{\Gamma}_N \cup \bar{\Gamma}_D = \partial\Omega \tag{3.3}$$

We shall assume there exist constants $\underline{a}, \bar{a}, \underline{c}$ and $\bar{c}$ such that the coefficients $a \in C^1(\bar{\Omega})$ and $c \in C(\bar{\Omega})$ satisfy

$$0 < \underline{a} \leq a(\boldsymbol{x}) \leq \bar{a}, \quad 0 < \underline{c} \leq c(\boldsymbol{x}) \leq \bar{c} \quad \text{for } \boldsymbol{x} \in \bar{\Omega}$$

The important case of $c \equiv 0$ is therefore excluded but will be discussed in the concluding remarks.

Let

$$V(\Omega) = \{v \in H^1(\Omega)\colon \gamma v = 0 \text{ on } \Gamma_D\}$$

Then $u \in V(\Omega)$ is the weak solution to (3.1) if

$$a(u, v) = l(v) \quad \forall v \in V(\Omega) \tag{3.4}$$

where

$$a(u, v) = \int_{\Omega} (a \nabla u \cdot \nabla v + cuv)\, dx \tag{3.5}$$

and

$$l(v) = \int_{\Omega} fv\, dx + \int_{\Gamma_N} g\gamma v\, ds \tag{3.6}$$

Under the above hypotheses $a\colon V(\Omega) \times V(\Omega) \to \mathbb{R}$ is a continuous and coercive bilinear form and $l\colon V(\Omega) \to \mathbb{R}$ is a continuous linear form. The Lax-Milgram theorem guarantees the existence of a unique solution to (3.4).

It is convenient to break the global forms $l(\cdot)$ and $a(\cdot, \cdot)$ into sums of contributions from each subdomain Ω_K of the partition $\mathscr{P}$:

$$a_K(u, v) = \int_{\Omega_K} (a \nabla u \cdot \nabla v + cuv)\, dx \tag{3.7}$$

$$l_K(v) = \int_{\Omega_K} fv\, dx + \int_{\partial\Omega_K \cap \Gamma_N} g\gamma v\, ds \tag{3.8}$$

so that

$$a(u, v) = \sum_{K=1}^{N} a_K(u, v) \tag{3.9}$$

and

$$l(v) = \sum_{K=1}^{N} l_K(v) \tag{3.10}$$

We shall use the notation $\|\cdot\|_E$ to denote the *energy norm*

$$\|v\|_E = \sqrt{a(v, v)} \tag{3.11}$$

and

$$\|v\|_{E,K} = \sqrt{a_K(v, v)} \tag{3.12}$$

so that

$$\|v\|_E^2 = \sum_{K=1}^{N} \|v\|_{E,K}^2 \tag{3.13}$$

Finally, we introduce the space V_K, $1 \leqq K \leqq N$

$$V_K = \{v \in H^1(\Omega_K)\colon \gamma v = 0 \text{ on } \partial\Omega_K \cap \Gamma_D\} \tag{3.14}$$

and

$$V(\mathscr{P}) = \{v \in H^1(\mathscr{P}) \colon \gamma v = 0 \text{ on } \Gamma_D\} \tag{3.15}$$

so that, equivalently,

$$V(\mathscr{P}) = \prod_{K=1}^{N} V_K \tag{3.16}$$

Let $H(\mathrm{div}, \Omega)$ denote the space

$$H(\mathrm{div}, \Omega) = \{\boldsymbol{q} \in [L^2(\Omega)]^2 : \mathrm{div}\, \boldsymbol{q} \in L^2(\Omega)\} \tag{3.17}$$

which is equipped with the norm

$$\|\boldsymbol{q}\|^2_{H(\mathrm{div},\Omega)} = \|\mathrm{div}\,\boldsymbol{q}\|^2_{L^2(\Omega)} + \|\boldsymbol{q}\|^2_{L^2(\Omega)} \tag{3.18}$$

Further, define the subspace $\mathscr{Q} \subset H(\mathrm{div}, \Omega)$ to be

$$\mathscr{Q} = \left\{\boldsymbol{q} \in H(\mathrm{div}, \Omega) : \oint_{\partial\Omega} v\boldsymbol{n} \cdot \boldsymbol{q}\, ds = 0 \ \forall v \in V(\Omega)\right\} \tag{3.19}$$

Finally, denote the space of continuous linear functionals on $V(\mathscr{P})$ and which vanish on $V(\Omega)$ by $\mathscr{M}$.

Lemma 3.1. *A continuous linear functional μ on $V(\mathscr{P})$ vanishes on $V(\Omega)$ (i.e. $\mu \in \mathscr{M}$) if and only if there exists $\boldsymbol{q} \in \mathscr{Q}$ such that*

$$\mu(v) = \sum_{K=1}^{N} \oint_{\partial\Omega_K} v_K \boldsymbol{n}_K \cdot \boldsymbol{q}\, ds \tag{3.20}$$

Proof. The result follows in a similar way to Lemma 1 in [14]. □

Let $\alpha^M_{KL} \colon \Gamma^M_{KL} \to \mathbb{R}$ be such that for $s \in \Gamma^M_{KL}$

$$\alpha^M_{KL}(s) + \alpha^M_{LK}(s) = 1, \quad 0 \leqq K, L \leqq N,\ 1 \leqq M \leqq \rho(K, L) \tag{3.21}$$

For any $v \in H^2(\mathscr{P})$, $[\![v]\!]$, $\langle v \rangle_\alpha$, $[\![\partial v/\partial n]\!]$ and $\langle \partial v/\partial n \rangle_\alpha \in L_2(E)$ are defined as follows (with the convention $v_0 \equiv 0$) for $s \in \Gamma^M_{KL}$ and for all $\Gamma^M_{KL} \in E$:

$$\left.\begin{aligned}
[\![v]\!]_{\Gamma^M_{KL}} &= \sigma_{KL}\gamma v_K + \sigma_{LK}\gamma v_L \\
\left[\!\!\left[\frac{\partial v}{\partial n}\right]\!\!\right]\Bigg|_{\Gamma^M_{KL}} &= \sigma_{KL}\gamma_1 v_K + \sigma_{LK}\gamma_1 v_L \\
\langle v \rangle_\alpha\big|_{\Gamma^M_{KL}} &= \alpha^M_{KL}\gamma v_K + \alpha^M_{LK}\gamma v_L \\
\left\langle \frac{\partial v}{\partial n} \right\rangle_\alpha\Bigg|_{\Gamma^M_{KL}} &= \alpha^M_{KL}\gamma_1 v_K + \alpha^M_{LK}\gamma_1 v_L
\end{aligned}\right\}. \tag{3.22}$$

In other words, $[\![v]\!]$ is the difference or *jump* in v between neighboring subdomains whilst $\langle v \rangle_\alpha$ is a linear combination or *weighted average* of v between neighboring subdomains.

A key result in the development is the following generalization of a result (Percell-Wheeler) found in [13]:

Lemma 3.2. *Let $v \in H^2(\mathscr{P})$ and $w \in H^1(\mathscr{P})$. Then*

$$\sum_{K=1}^{N} \oint_{\partial\Omega_K} \frac{\partial v_K}{\partial n_K} \gamma w_K \, ds = \left(\left[\!\!\left[\frac{\partial v}{\partial n} \right]\!\!\right], \langle w \rangle_\alpha \right)_{0,E} + \left(\left\langle \frac{\partial v}{\partial n} \right\rangle_{1-\alpha}, [\![w]\!] \right)_{0,E} \tag{3.23}$$

Proof. Owing to the regularity $v \in H^2(\mathscr{P})$, it follows that $\boldsymbol{n} \cdot \nabla v \in L^2(\partial\Omega_K)$ (see [8, p. 9, Remark 1.1]) and therefore the duality pairings may be treated as Lebesgue integrals over the boundaries of the subdomains. We have

$$\begin{aligned}
\sum_{K=1}^{N} \oint_{\partial\Omega_K} \frac{\partial v_K}{\partial n_K} \gamma w_K \, ds &= \sum_{K=1}^{N} \sum_{L=0}^{N} \sum_{M=1}^{\rho(K,L)} \int_{\Gamma_{KL}^M} \frac{\partial v_K}{\partial n_K} \gamma w_K \, ds \\
&= \sum_{\substack{L,K=0 \\ K>L}}^{N} \sum_{M=1}^{\rho(K,L)} \int_{\Gamma_{KL}^M} \frac{\partial v_K}{\partial n_K} \gamma w_K \, ds + \sum_{\substack{L,K=0 \\ K<L}}^{N} \sum_{M=1}^{\rho(K,L)} \int_{\Gamma_{KL}^M} \frac{\partial v_K}{\partial n_K} \gamma w_K \, ds \\
&= \sum_{\substack{L,K=0 \\ K>L}}^{N} \sum_{M=1}^{\rho(K,L)} \int_{\Gamma_{KL}^M} \frac{\partial v_K}{\partial n_K} \gamma w_K \, ds + \sum_{\substack{K,L=0 \\ L<K}}^{N} \sum_{M=1}^{\rho(L,K)} \int_{\Gamma_{LK}^M} \frac{\partial v_L}{\partial n_L} \gamma w_L \, ds \\
&= \sum_{\substack{L,K=0 \\ K>L}}^{N} \sum_{M=1}^{\rho(K,L)} \int_{\Gamma_{KL}^M} \left(\frac{\partial v_K}{\partial n_K} \gamma w_K + \frac{\partial v_L}{\partial n_L} \gamma w_L \right) ds \\
&= \sum_{\substack{L,K=0 \\ K>L}}^{N} \sum_{M=1}^{\rho(K,L)} \int_{\Gamma_{KL}^M} \left(\sigma_{KL} \frac{\partial v_K}{\partial n} \gamma w_K + \sigma_{LK} \frac{\partial v_L}{\partial n} \gamma w_L \right) ds
\end{aligned}$$

Now

$$\sigma_{KL} \frac{\partial v_K}{\partial n} = \left[\!\!\left[\frac{\partial v}{\partial n} \right]\!\!\right] \alpha_{KL} + \sigma_{KL} \left\langle \frac{\partial v}{\partial n} \right\rangle_{1-\alpha}$$

and

$$\gamma w_K = \alpha_{LK} \sigma_{KL} [\![w]\!] + \langle w \rangle_\alpha$$

so since $\sigma_{KL} + \sigma_{LK} \equiv 0$ we obtain

$$\sigma_{KL} \frac{\partial v_K}{\partial n} \gamma w_K + \sigma_{LK} \frac{\partial v_L}{\partial n} \gamma w_L = (\alpha_{KL} + \alpha_{KL}) \left[\!\!\left[\frac{\partial v}{\partial n} \right]\!\!\right] \langle w \rangle_\alpha + (\alpha_{LK} + \alpha_{KL}) [\![w]\!] \left\langle \frac{\partial v}{\partial n} \right\rangle_{1-\alpha}$$

and the result follows as claimed since $\alpha_{KL} + \alpha_{LK} \equiv 1$. □

4. A posteriori error bounds

Let u_h be *any* approximation to the true solution of the model problem such that

(4.1) $$u_h \in V(\Omega) \cap H^2(\mathscr{P})$$

Such an approximation might, for example, be obtained using a finite element discretization with Ω_K as the elements. Conversely, u_h might simply be some guess at the true solution.

The problem of interest is that of numerically estimating the accuracy of the approximation. The error $e(\boldsymbol{x})$ in the approximation at $\boldsymbol{x}$ is defined as $u(\boldsymbol{x}) - u_h(\boldsymbol{x})$. We shall be specifically interested in obtaining bounds on the error measured in the energy norm, $\| e \|_E$.

Lemma 4.1. *Let $J : V(\Omega) \to \mathbb{R}$ be defined as*

(4.2) $$J(v) = \tfrac{1}{2} a(v, v) - l(v) + a(u_h, v)$$

Then the error $e = u - u_h$ is the unique minimizer of J over $V(\Omega)$. Moreover,

(4.3) $$-\tfrac{1}{2} \| e \|_E^2 = J(e) = \inf_{v \in V(\Omega)} J(v)$$

Proof. First, notice that $e \in V(\Omega)$ since $u_h \in V(\Omega)$. Now

$$J(v) = \tfrac{1}{2} a(v, v) - l(v) + a(u_h, v) = \tfrac{1}{2} a(v, v) - a(u, v) + a(u_h, v)$$
$$= \tfrac{1}{2} a(v, v) - a(e, v)$$

and hence
$$J(e) = -\tfrac{1}{2} \| e \|_E^2$$

Let $\lambda \in \mathbb{R}$ and $\omega \in V(\Omega)$. Then

$$J(e) - J(e + \lambda\omega) = -\tfrac{1}{2} \lambda^2 a(\omega, \omega) \leqq 0$$

and equality holds iff $\| \lambda\omega \|_E^2 = 0 \Leftrightarrow \lambda\omega = 0$. Thus $J(e) \leqq J(v)\ \forall v \in V(\Omega)$. □

Let $J_{\mathscr{P}} : V(\mathscr{P}) \to \mathbb{R}$ be defined as

(4.4) $$J_{\mathscr{P}}(v) = \sum_{K=1}^{N} \left\{ \frac{1}{2} a_K(v, v) - l_K(v) + a_K(u_h, v) \right\}$$

That is, $J_{\mathscr{P}}$ is an extension of J to $V(\mathscr{P}) \supset V(\Omega)$. Let $\mathscr{L}_{\mathscr{P}} : V(\mathscr{P}) \times \mathscr{M} \to \mathbb{R}$ denote the Lagrangian functional

(4.5) $$\mathscr{L}_{\mathscr{P}}(v, \mu) = J_{\mathscr{P}}(v) - \mu(v)$$

associated with the constraint $v \in V(\Omega)$, μ representing the Lagrange multiplier. It is instructive to compare the Lagrangian $\mathscr{L}_{\mathscr{P}}$ with analogous functionals used in the analysis of *primal-hybrid* finite element methods [14].

Lemma 4.2. *With the above assumptions and notations there follows*

$$-\tfrac{1}{2}\|e\|_E^2 = \inf_{v\in V(\mathscr{P})} \sup_{\mu\in\mathscr{M}} \mathscr{L}_{\mathscr{P}}(v,\mu) \tag{4.6}$$

Proof. Let $\Phi\colon V(\mathscr{P}) \to \mathbb{R}$ be the functional

$$\Phi(v) = \sup_{\mu\in\mathscr{M}} \mathscr{L}_{\mathscr{P}}(v,\mu)$$

and observe that

$$\Phi(v) = \begin{cases} J(v), & v\in V(\Omega) \\ +\infty, & \text{otherwise} \end{cases}$$

Therefore

$$\inf_{v\in V(\mathscr{P})} \sup_{\mu\in\mathscr{M}} \mathscr{L}_{\mathscr{P}}(v,\mu) = \inf_{v\in V(\mathscr{P})} \Phi(v) = \inf_{v\in V(\Omega)} \Phi(v)$$

$$= \inf_{v\in V(\mathscr{P})} J(v) = -\tfrac{1}{2}\|e\|_E^2$$

where the final step follows from Lemma 4.1. □

Let $v\in V(\mathscr{P})$. Then applying Green's identity on each subdomain, we have

$$a(u_h,v) = \sum_{K=1}^{N} a_K(u_h,v) = \sum_{K=1}^{N} \left\{ \oint_{\partial\Omega_K} a\frac{\partial u_h}{\partial n_K}\gamma v\,ds + \oint_{\Omega_K} vLu_h\,dx \right\} \tag{4.7}$$

Applying Lemma 3.2 to the first term gives

$$a(u_h,v) = \sum_{K=1}^{N}\int_{\Omega_K} vLu_h\,dx + \sum_{\substack{K,L=0\\K>L\\1\le M\le\rho(K,L)}}^{N} \int_{\Gamma_{KL}^M} \left\{ \left[\!\left[a\frac{\partial u_h}{\partial n} \right]\!\right] \langle v\rangle_\alpha \right. \tag{4.8}$$

$$\left. + \left\langle a\frac{\partial u_h}{\partial n} \right\rangle_{1-\alpha} [\![v]\!] \right\} ds$$

$$= \sum_{K=1}^{N}\int_{\Omega_K} vLu_h\,dx + \sum_{\substack{K,L=1\\K>L\\1\le M\le\rho(K,L)}}^{N} \int_{\Gamma_{KL}^M} \left\{ \left[\!\left[a\frac{\partial u_h}{\partial n} \right]\!\right] \langle v\rangle_\alpha \right.$$

$$\left. + \left\langle a\frac{\partial u_h}{\partial n} \right\rangle_{1-\alpha} [\![v]\!] \right\} ds + \sum_{\substack{K=1\\1\le M\le\rho(K,0)}}^{N} \int_{\Gamma_{K0}^M} a\frac{\partial u_h}{\partial n}\gamma v\,ds$$

since $\alpha_{K0} \equiv 1$, from (3.22) we obtain

$$\left\langle a\frac{\partial u_h}{\partial n} \right\rangle_{1-\alpha} = 0, \qquad \gamma v_0 = 0, \qquad \langle v\rangle_\alpha = v_K \quad \text{and}$$

$$\left[\!\left[a\frac{\partial u_h}{\partial n} \right]\!\right] = a\frac{\partial u_h}{\partial n} \quad \text{on } \Gamma_{K0}^M$$

Now $\gamma v = 0$ on $\Gamma^M_{K0} \cap \Gamma_D$ and so

$$(4.9)\quad J_{\mathscr{P}}(v) = \sum_{K=1}^{N} \left\{ \frac{1}{2} a_K(v,v) - (f,v)_{0,K} + (Lu_h, v)_{0,K} \right\}$$

$$+ \sum_{\substack{K,L=1\\ K>L\\ 1 \leqq M \leqq \rho(K,L)}}^{N} \int_{\Gamma^M_{KL}} \left\{ \left[\!\!\left[a\frac{\partial u_h}{\partial n} \right]\!\!\right] \langle v \rangle_\alpha + \left\langle a \frac{\partial u_h}{\partial n} \right\rangle_{1-\alpha} [\![v]\!] \right\} ds$$

$$+ \sum_{\substack{K=1\\ 1 \leqq M \leqq \rho(K,0)}}^{N} \int_{\Gamma^M_{K0}} \left(a\frac{\partial u_h}{\partial n} - g \right) \gamma v \, ds$$

$$= \sum_{K=1}^{N} \left\{ \frac{1}{2} a_K(v,v) + (Lu_h - f, v)_{0,K} + \sum_{1 \leqq M \leqq \rho(K,0)} \int_{\Gamma^M_{K0}} \left(a\frac{\partial u_h}{\partial n} - g \right) \gamma v_K \, ds \right.$$

$$\left. + \sum_{\substack{L=1\\ 1 \leqq M \leqq \rho(K,L)}}^{N} \int_{\Gamma^M_{KL}} \alpha^M_{KL} \left[\!\!\left[a\frac{\partial u_h}{\partial n} \right]\!\!\right] \gamma v_K \, ds \right\}$$

$$+ \sum_{\substack{K,L=1\\ K>L\\ 1 \leqq M \leqq \rho(K,L)}}^{N} \int_{\Gamma^M_{KL}} \left\langle a\frac{\partial u_h}{\partial n} \right\rangle_{1-\alpha} [\![v]\!] \, ds$$

since

$$\sum_{\substack{K,L=1\\ K>L\\ 1 \leqq M \leqq \rho(K,L)}}^{N} \int_{\Gamma^M_{KL}} \left[\!\!\left[a\frac{\partial u_h}{\partial n} \right]\!\!\right] \langle v \rangle_\alpha \, ds = \sum_{K=1}^{N} \sum_{L=1}^{N} \sum_{M=1}^{\rho(K,L)} \int_{\Gamma^M_{KL}} \alpha^M_{KL} \left[\!\!\left[a\frac{\partial u_h}{\partial n} \right]\!\!\right] \gamma v_K \, ds$$

Let

$$(4.10)\qquad r_K(x) = f(x) - Lu_h(x), \quad x \in \Omega_K$$

denote the elementwise residual and define

$$(4.11)\quad R_K = \begin{cases} g - a\dfrac{\partial u_h}{\partial n} & \text{on } \Gamma_N \cap \Gamma^M_{K0},\ 1 \leqq K \leqq N,\ 1 \leqq M \leqq \rho(K,0) \\[2ex] -\alpha_{KL} \left[\!\!\left[a\dfrac{\partial u_h}{\partial n} \right]\!\!\right] & \text{on } \Gamma^M_{KL},\ 1 \leqq K, L \leqq N,\ 1 \leqq M \leqq \rho(K,L) \end{cases}$$

Then:

$$(4.12)\qquad J_{\mathscr{P}}(v) = \sum_{K=1}^{N} J_{\mathscr{P},K}(v) + \left(\left\langle a\frac{\partial u_h}{\partial n} \right\rangle_{1-\alpha}, [\![v]\!] \right)_{0,E}$$

where $J_{\mathscr{P},K}\colon V_K \to \mathbb{R}$ is given by

$$(4.13)\qquad J_{\mathscr{P},K}(v_K) = \tfrac{1}{2} a_K(v_K, v_K) - (r_K, v_K)_{0,K} - \oint_{\partial\Omega_K} R_K \gamma v_K \, ds$$

Notice that it is unnecessary to define R_K on Γ_D, since $\gamma v_K = 0$ on Γ_D.

Roughly speaking, we may interpret this result as showing that $J_{\mathscr{P}}$ may be decomposed into a sum of local contributions from each of the subdomains Ω_K, plus an extra term coupling the local contributions though the regularity $v \in V(\Omega)$. The significant feature, from our viewpoint, is the possibility of choosing the Lagrange multiplier $\mu \in \mathscr{M}$ in such a way that the coupling is severed. If this step could be achieved, then it would result in a *sequence of local problems* rather than a *single global problem.* This situation is highly desirable in practice since the computational cost of dealing with a sequence of local problems is much smaller than dealing with a single global problem.

In the following, we focus on the problem of constructing a suitable choice of μ. Making use of the definitions (3.22) we find

$$(4.14)\qquad \left(\left\langle a\frac{\partial u_h}{\partial n}\right\rangle_{1-\alpha}, [\![v]\!]\right)_{0,E} = \sum_{K=1}^{N} \left\{ \sum_{\substack{L=0\\ 1\le M\le \rho(K,L)}}^{N} \int_{\Gamma_{KL}^{M}} \left\langle a\frac{\partial u_h}{\partial n}\right\rangle_{1-\alpha} v\,ds \right\}$$

Noting that

$$(4.15)\qquad \partial\Omega_K = \bigcup_{\substack{L=0\\ 1\le M\le \rho(K,L)}}^{N} \Gamma_{KL}^{M}$$

we may formally define the linear functional ζ on V_K by

$$(4.16)\qquad \zeta(v) = \sum_{\substack{L=0\\ 1\le M\le \rho(K,L)}}^{N} \int_{\Gamma_{KL}^{M}} \left\langle a\frac{\partial u_h}{\partial n}\right\rangle_{1-\alpha} v_K\,ds$$

Examining this expression one readily concludes that, in fact, ζ is a bounded linear functional on $L^2(\partial\Omega_K)$ and hence a fortiori $\zeta \in H^{-1/2}(\partial\Omega_K)$. Recalling that the Trace Operator

$$H(\operatorname{div}, \Omega_K) \ni \boldsymbol{q} \mapsto \boldsymbol{n}_K \cdot \boldsymbol{q} \in H^{-1/2}(\partial\Omega_K)$$

is surjective, we may construct $\boldsymbol{q} \in H(\operatorname{div}, \Omega_K)$ such that

$$(4.17)\qquad \int_{\partial\Omega_K} v_K \boldsymbol{n}_K \cdot \boldsymbol{q}\,ds = \sum_{\substack{L=0\\ 1\le M\le \rho(K,L)}}^{N} \int_{\Gamma_{KL}^{M}} \left\langle a\frac{\partial u_h}{\partial n}\right\rangle_{1-\alpha} v_K\,ds$$

Performing this procedure over each subdomain results in our having constructed $\boldsymbol{q} \in H(\operatorname{div}, \Omega)$ satisfying

$$(4.18)\qquad \sum_{K=1}^{N} \int_{\partial\Omega_K} v_K \boldsymbol{n}_K \cdot \boldsymbol{q}\,ds = \sum_{K=1}^{N} \sum_{\substack{L=0\\ 1\le M\le \rho(K,L)}}^{N} \int_{\Gamma_{KL}^{M}} \left\langle a\frac{\partial u_h}{\partial n}\right\rangle_{1-\alpha} v_K\,ds$$

$$(4.19)\qquad = \left(\left\langle a\frac{\partial u_h}{\partial n}\right\rangle_{1-\alpha}, [\![v]\!]\right)_{0,E}$$

Moreover, suppose $v \in V(\Omega)$, then

$$(4.20)\qquad \sum_{K=1}^{N} \int_{\partial\Omega_K} v_K \boldsymbol{n}_K \cdot \boldsymbol{q}\,ds = \oint_{\partial\Omega} v\boldsymbol{n} \cdot \boldsymbol{q}\,ds$$

and

$$\left(\left\langle a\frac{\partial u_h}{\partial n}\right\rangle_{1-\alpha}, [\![v]\!]\right)_{0,E} = \oint_{\partial\Omega}\left\langle a\frac{\partial u_h}{\partial n}\right\rangle_{1-\alpha} v\,ds = 0 \tag{4.21}$$

where the final quantity vanishes since $\gamma v = 0$ on Γ_D and $\langle a\,(\partial u_h/\partial n)\rangle_{1-\alpha} = 0$ on Γ_N by definition. In view of the above developments, it follows that $\boldsymbol{q} \in \mathcal{Q}$. Applying Lemma 3.1 then shows that there exists $\hat{\mu} \in \mathcal{M}$ such that

$$\hat{\mu}(v) = \left(\left\langle a\frac{\partial u_h}{\partial n}\right\rangle_{1-\alpha}, [\![v]\!]\right)_{0,E} \tag{4.22}$$

Hence, returning to $J_{\mathscr{P}}$, we find we may write

$$J_{\mathscr{P}}(v) = \sum_{K=1}^{N} J_{\mathscr{P},K}(v) = \hat{\mu}(v) \tag{4.23}$$

and, equally well,

$$\mathscr{L}_{\mathscr{P}}(v, \mu) = \sum_{K=1}^{N} J_{\mathscr{P},K}(v) + \hat{\mu}(v) - \mu(v) \tag{4.24}$$

Therefore, we can obtain the desired decoupling in the case $\mu = \hat{\mu}$.

Lemma 4.3. *Let $\hat{\mu} \in \mathcal{M}$ be constructed as above. Then*

$$\|e\|_E^2 \leqq \sup_{v \in V(\mathscr{P})} -2\mathscr{L}_{\mathscr{P}}(v, \hat{\mu}) \tag{4.25}$$

where

$$\mathscr{L}_{\mathscr{P}}(v, \hat{\mu}) = \sum_{K=1}^{N} J_{\mathscr{P},K}(v) \tag{4.26}$$

Proof. From Lemma 4.2, we have

$$-\tfrac{1}{2}\|e\|_E^2 = \inf_{v \in V(\mathscr{P})} \sup_{\mu \in \mathcal{M}} \mathscr{L}_{\mathscr{P}}(v, \mu) \geqq \sup_{\mu \in \mathcal{M}} \inf_{v \in V(\mathscr{P})} \mathscr{L}_{\mathscr{P}}(v, \mu)$$

$$\geqq \inf_{v \in V(\mathscr{P})} \mathscr{L}_{\mathscr{P}}(v, \hat{\mu})$$

so $\|e\|_E^2 \leqq \sup_{v \in V(\mathscr{P})} -2\mathscr{L}_{\mathscr{P}}(v, \hat{\mu})$. □

One serious drawback of this result as a means of estimating the error is the presence of the supremum term which either will not be attained by our computed choice of $v \in V(\mathscr{P})$ or the cost involved in calculating a suitable v will prove to be disproportionately expensive compared to the computational effort expended in obtaining u_h itself. This type of difficulty was successfully resolved in [2] using *dual variational principles* (see [7]).

Lemma 4.4. *Let*

(4.27) $$W_K = \left\{ p \in H(\mathrm{div}, \Omega_K): \oint_{\partial\Omega_K} n_K \cdot p v \, ds = \oint_{\partial\Omega_K} R_K v \, ds \ \forall v \in V_K \right\}$$

In addition, let $\mathscr{G}_K: W_K \to \mathbb{R}$ *be the functional*

(4.28) $$\mathscr{G}_K(p) = -\frac{1}{2} \oint_{\Omega_K} \frac{1}{a} p \cdot p \, dx - \frac{1}{2} \oint_{\Omega_K} \frac{1}{c} (\nabla \cdot p + r_K)^2 \, dx \qquad (4.28)$$

where r_K *is the element residual of* (4.10). *Further, let* $v_K^* \in V_K$ *be the solution of the local problem*

(4.29) $$a_K(v_K^*, \omega) = (r_K, \omega)_{0,K} + \oint_{\partial\Omega_K} R_K \gamma \omega \, ds \quad \forall \omega \in V_K$$

Then

(4.30) $$\text{(i)} \quad p_K^* \stackrel{\text{def}}{=} a \nabla v_K^* \in W_K$$

and

(4.31) $$\text{(ii)} \quad \inf_{v_K \in V_K} J_{\mathscr{P},K}(v_K) = J_{\mathscr{P},K}(v_K^*) = \mathscr{G}_K(p_K^*) = \sup_{p_K \in W_K} \mathscr{G}_K(p_K)$$

Proof. (i) The existence and uniqueness of v_K^* is a consequence of the Lax-Milgram Theorem. The strong form of (4.29) is

$$-\nabla \cdot (a \nabla v_K^*) + c v_K^* = r_K \quad \text{in } \Omega_K$$

$$a \frac{\partial v_K^*}{\partial n_K} = R_K \quad \text{on } \partial\Omega_K \backslash \Gamma_D$$

$$\gamma v_K^* = 0 \quad \text{on } \partial\Omega_K \cap \Gamma_D$$

Let $p_K^* = a \nabla v_K^*$. Then

$$\nabla \cdot p_K^* = c v_K^* - r_K \in L_2(\Omega_K)$$

whilst

$$n_K \cdot p_K^* = a \frac{\partial v_K^*}{\partial n_K} = R_K \quad \text{on } \partial\Omega_K \backslash \Gamma_D$$

Thus $p_K^* \in W_K$.

(ii) That v_K^* is the minimizer of $J_{\mathscr{P},K}$ follows in a similar manner to Lemma 4.1. Moreover $v_K^* \in V_K$ so

$$J_{\mathscr{P},K}(v_K^*) = \frac{1}{2} a_K(v_K^*, v_K^*) - (r_K, v_K^*)_{0,K} - \oint_{\partial\Omega_K} R_K \gamma v_K^* \, ds$$

$$= \frac{1}{2} a_K(v_K^*, v_K^*) - a_K(v_K^*, v_K^*) = -\frac{1}{2} a_K(v_K^*, v_K^*)$$

Furthermore,

$$\mathscr{G}_K(p_K^*) = -\frac{1}{2} \int_{\Omega_K} a |\nabla v_K^*|^2 \, dx - \frac{1}{2} \int_{\Omega_K} \frac{1}{c} (c v_K^*)^2 \, dx = -\frac{1}{2} a_K(v_K^*, v_K^*)$$

Since $\mathcal{G}_K$ is strictly concave and quadratic, it suffices to show $\mathcal{G}_K$ is stationary at $\boldsymbol{p}_K^*$.
Let

$$q \in \left\{ q \in H(\mathrm{div}, \Omega_K) \colon \oint_{\partial\Omega_K} \gamma v \boldsymbol{n} \cdot \boldsymbol{q}\, ds = 0 \ \forall v \in V_K \right\} \tag{4.32}$$

and let $\lambda \in \mathbb{R}$. Then $\boldsymbol{p}_K^* + \lambda \boldsymbol{q} \in W_K$ and

$$\begin{aligned}\frac{d}{d\lambda}\mathcal{G}_K(\boldsymbol{p}_K^* + \lambda \boldsymbol{q})|_{\lambda=0} &= -\int_{\Omega_K} \frac{1}{a}\boldsymbol{q}\cdot\boldsymbol{p}_K^*\, dx - \int_{\Omega_K} \frac{1}{c}(\nabla\cdot\boldsymbol{q})(\nabla\cdot\boldsymbol{p}_K^* + r_K)\, dx \\ &= -\int_{\Omega_K} \boldsymbol{q}\cdot\nabla v_K^*\, dx - \int_{\Omega_K} (\nabla\cdot\boldsymbol{q}) v_K^*\, dx \\ &= -\oint_{\partial\Omega_K} \gamma v_K^*(\boldsymbol{n}_K\cdot\boldsymbol{q})\, ds\end{aligned}$$

This now vanishes owing to (4.32). □

Theorem 4.5. *Let W_K be as in* (4.27). *Then*

$$\|e\|_E^2 \leqq -2\sum_{K=1}^{N} \mathcal{G}_K(\boldsymbol{p}) \quad \forall \boldsymbol{p} \in \prod_{K=1}^{N} W_K \tag{4.33}$$

Proof. From Lemmas 4.3 and 4.4 we have

$$\begin{aligned}\|e\|_E^2 &\leqq \sup_{v\in V(\mathcal{P})} \sum_{K=1}^{N} -2J_{\mathcal{P},K}(v) = -2\sum_{K=1}^{N} \inf_{v_K\in V_K} J_{\mathcal{P},K}(v_K) = -2\sum_{K=1}^{N} \sup_{\boldsymbol{p}_K\in W_K} \mathcal{G}_K(\boldsymbol{p}_K) \\ &\leqq -2\sum_{K=1}^{N} \mathcal{G}_K(\boldsymbol{p}_K)\ \forall\, \boldsymbol{p}_K \in W_K = -2\sum_{K=1}^{N} \mathcal{G}_K(\boldsymbol{p}) \quad \forall \boldsymbol{p} \in \prod_{K=1}^{N} W_K\end{aligned}$$ □

The main result, Theorem 4.5, shows that computable error bounds can be obtained merely by constructing elements of the linear manifolds W_K. Obviously to obtain realistic estimates it is necessary to choose $\boldsymbol{p}$ with some care.

5. Local element residual error estimator

In this section we propose a strategy for constructing $\boldsymbol{p}$. Let $\varepsilon_K \colon W_K \to \mathbb{R}$ and $\Lambda_K \colon W_K \to \mathbb{R}$ be defined as

$$\varepsilon_K^2(\boldsymbol{p}) = \int_{\Omega_K} \frac{1}{a}\boldsymbol{p}\cdot\boldsymbol{p}\, dx \tag{5.1}$$

and

$$\Lambda_K^2(\boldsymbol{p}) = \int_{\Omega_K} \frac{1}{c}(\nabla\cdot\boldsymbol{p} + r_K)^2\, dx \tag{5.2}$$

Furthermore, let

(5.3) $$\eta_K^2(\boldsymbol{p}) = \varepsilon_K^2(\boldsymbol{p}) + \Lambda_K^2(\boldsymbol{p})$$

Then Theorem 4.5 may be rewritten as

(5.4) $$\|e\|_K^2 \leqq \sum_{K=1}^{N} \eta_K^2(\boldsymbol{p}) \quad \forall \boldsymbol{p} \in \prod_{K=1}^{N} W_K$$

This result suggests that $\boldsymbol{p}$ should be chosen to minimize the local functionals $\eta_K^2(\boldsymbol{p})$. Owing to the fact that the functionals and spaces are *local* this is a viable practical method. However, an alternative approach is pursued for reasons which will become clear. Rather than minimizing $\varepsilon_K^2 + \Lambda_K^2$ we shall concentrate all our efforts on the term Λ_K^2 alone. The strategy is therefore to choose $\boldsymbol{p}_K \in W_K$ such that

(5.5) $$\Lambda_K(\boldsymbol{p}_K) \leqq \Lambda_K(\boldsymbol{q}_K) \quad \forall \boldsymbol{q}_K \in W_K$$

On subdomains $\Omega_K: \partial\Omega_K \cap \Gamma_D \neq \varnothing$ it is always possible to choose $\boldsymbol{p}_K$ to give $\Lambda_K(\boldsymbol{p}_K) \equiv 0$. Specifically we choose $\boldsymbol{p}_K = \nabla\varphi_K$ where $\varphi_K \in H^1(\Omega_K)$ satisfies

(5.6) $$\left.\begin{aligned} -\Delta\varphi_K &= r_K \quad \text{in } \Omega_K \\ \frac{\partial\varphi_K}{\partial n} &= R_K \quad \text{on } \partial\Omega_K \backslash \Gamma_D \\ \gamma\varphi_K &= 0 \quad \text{on } \partial\Omega_K \cap \Gamma_D \end{aligned}\right\}$$

The existence of a unique φ_K is again guaranteed by the Lax-Milgram Theorem.

For subdomains $\Omega_K: \partial\Omega_K \cap \Gamma_D = \varnothing$, it is in general not possible to choose $\boldsymbol{p}_K$ such that $\Lambda_K(\boldsymbol{p}_K) = 0$. The following result quantifies this statement.

Lemma 5.1. *Suppose $\partial\Omega_K \cap \Gamma_D$ is empty. Let $\varphi_K \in H^1(\Omega_K)$ be such that*

(5.7) $$\left.\begin{aligned} -\Delta\varphi_K &= r_K - c\delta_K \quad \text{in } \Omega_K \\ \frac{\partial\varphi_K}{\partial n_K} &= R_K \quad \text{on } \partial\Omega_K \end{aligned}\right\}$$

where

(5.8) $$\delta_K = \left\{ \int_{\Omega_K} r_K\, dx + \oint_{\partial\Omega_K} R_K\, ds \right\} \Big/ \left\{ \oint_{\Omega_K} c\, dx \right\}$$

then

(5.9) $$(i) \quad \nabla\varphi_K \in W_K$$

and

(5.10) $$(ii) \quad \Lambda_K(\nabla\varphi_K) \leqq \Lambda_K(\boldsymbol{q}_K) \quad \forall \boldsymbol{q}_K \in W_K$$

Proof. (i) First, observe that

$$\int_{\Omega_K} (r_K - c\delta_K)\, dx + \oint_{\partial\Omega_K} R_K\, ds = 0$$

It is well known that this *compatibility condition* is necessary and sufficient for the existence of a solution $\varphi_K \in H^1(\Omega_K)$ to (5.7), (φ_K being unique up to the addition of an arbitrary constant). Now

$$\nabla \cdot (\nabla \varphi_K) = -r_K + c\delta_K \in L_2(\Omega_K)$$

and

$$\boldsymbol{n}_K \cdot \nabla \varphi_K = \frac{\partial \varphi_K}{\partial n_K} = R_K \quad \text{on } \partial\Omega_K$$

so $\nabla \varphi_K \in W_K$.

(ii) Let $\boldsymbol{q}_K \in W_K$ be given and define

$$\xi_K = \nabla \cdot \boldsymbol{q}_K + r_K \in L_2(\Omega_K)$$

then

$$\Lambda_K^2(\boldsymbol{q}_K) = \int_{\Omega_K} \frac{1}{c} \xi_K^2 \, dx$$

Moreover, since $\boldsymbol{q}_K \in W_K$,

$$\int_{\Omega_K} \xi_K \, dx = \oint_{\partial\Omega_K} \boldsymbol{n}_K \cdot \boldsymbol{q}_K \, ds + \int_{\Omega_K} r_K \, dx = \oint_{\partial\Omega_K} R_K \, ds + \int_{\Omega_K} r_K \, dx = \delta_K \oint_{\Omega_K} c \, dx$$

By the Cauchy-Schwarz Inequality,

$$\left(\delta_K \int_{\Omega_K} c \, dx \right)^2 = \left(\int_{\Omega_K} \xi \, dx \right)^2 \leq \int_{\Omega_K} c \, dx \cdot \int_{\Omega_K} \frac{1}{c} \xi^2 \, dx$$

and, hence,

$$\Lambda_K^2(\nabla \varphi_K) = \delta_K^2 \int_{\Omega_K} c \, dx \leq \int_{\Omega_K} \frac{1}{c} \xi^2 \, dx = \Lambda_K^2(\boldsymbol{q}_K) \qquad \square$$

In view of this result, we choose $\boldsymbol{p}_K = \nabla \varphi_K$ where φ_K is any solution of the local problem (5.7). The local error estimator on subdomain Ω_K is taken as

$$\eta_K^2(\nabla \varphi_K) = \varepsilon_K^2(\nabla \varphi_K) + \Lambda_K^2(\nabla \varphi_K)$$

where φ_K is the solution of (5.6) or (5.7) depending on whether or not Γ_D intersects $\partial\Omega_K$.

Theorem 5.2. *Let φ_K be the solution of* (5.6) *or* (5.7). *Then*

$$\|e\|_E^2 \leq \sum_{K=1}^{N} \eta_K^2(\nabla \varphi_K) \tag{5.11}$$

Proof. Follows immediately from foregoing arguments and (5.4). $\square$

The importance of this result is that *no matter how we choose* $\alpha(s)$ subject to (3.21), the resulting error estimator always gives an upper bound on the true error.

At this stage it is worthwhile to compare the result with other types of element residual methods. Almost all existing methods are associated with the specific choice $\alpha(s) \equiv \frac{1}{2}$.

A more fundamental difference between the method proposed here and existing methods is that the local problem involves only the Laplacian operator whilst other methods have local problems based on the operator L. While it might seem more advantageous to have local problems based on L this actually does not appear to be the case.

One requirement of error estimators which is of paramount importance is that they must be *cheap* to compute. The local problems are often solved using finite elements in an h, p, or h-p mode. The main cost in solving by finite elements is the assembly of stiffness matrices and solution of the resulting matrix equation. By basing the local problems on the Laplacian one can assemble the stiffness matrix a priori on the reference element and keep this as data within the code. Secondly, one can easily construct an orthogonal basis for the p-version finite element approximation in the case of the Laplacian operator. This not only makes the solution of the linear system extremely cheap but allows one to increase the accuracy of the approximation of the local problem very simply. As a consequence, one can assume that the local problems have been solved exactly as indeed we shall do throughout our analysis.

Lemma 5.3. *Let φ_K be a solution of the local residual problem* (5.7). *Then*

$$\Lambda_K(\nabla\varphi_K) \leq \left(\int_{\Omega_K} ce^2\,dx\right)^{1/2} + \left(\int_{\Omega_K} c\,dx\right)^{-1/2} \left|\oint_{\partial\Omega_K} \left\langle \frac{\partial e}{\partial n_K}\right\rangle_{1-\alpha} ds\right| \tag{5.12}$$

Proof. From Lemma 5.1 we have

$$\Lambda_K^2(\nabla\varphi_K) = \delta_K^2 \int_{\Omega_K} c\,dx$$

where

$$\delta_K \int_{\Omega_K} c\,dx = \int_{\Omega_K} r_K\,dx + \oint_{\partial\Omega_K} R_K\,ds$$

Now $r_K = f - Lu_h = Le_K$, so

$$\begin{aligned}\delta_K \int_{\Omega_K} c\,dx &= -\int_{\Omega_K} \nabla\cdot(a\nabla e_K)\,dx + \int_{\Omega_K} ce_K\,dx + \oint_{\partial\Omega_K} R_K\,ds \\ &= \oint_{\partial\Omega_K} \left\{-a\frac{\partial e_K}{\partial n_K} + R_K\right\} ds + \int_{\Omega_K} ce_K\,dx\end{aligned}$$

On $\partial\Omega_K\backslash\partial\Omega$ we have

$$\begin{aligned}-a\frac{\partial e_K}{\partial n_K} + R_K &= -a\frac{\partial u}{\partial n_K} + a\frac{\partial u_h}{\partial n_K}\bigg|_K - \alpha_{KL}\left[\!\!\left[a\frac{\partial u_h}{\partial n}\right]\!\!\right] \\ &= -a\frac{\partial u}{\partial n_K} + a\frac{\partial u_h}{\partial n_K}\bigg|_K - \alpha_{KL}\left(a\frac{\partial u_h}{\partial n}\bigg|_K \sigma_{KL} + a\frac{\partial u_h}{\partial n}\bigg|_L \sigma_{LK}\right) \\ &= \sigma_{KL}\left(\left\langle a\frac{\partial u_h}{\partial n}\right\rangle_{1-\alpha} - a\frac{\partial u}{\partial n}\right) \\ &= -\sigma_{KL}\left\langle a\frac{\partial e}{\partial n}\right\rangle_{1-\alpha}\end{aligned} \tag{5.13}$$

By the Triangle and Cauchy-Schwarz Inequalities,

$$\begin{aligned}\left|\delta_K \int_{\Omega_K} c\,dx\right| &\leqq \left|\int_{\Omega_K} ce\,dx\right| + \left|\oint_{\partial\Omega_K} \sigma_{KL}\left\langle a\frac{\partial e}{\partial n}\right\rangle_{1-\alpha} ds\right| \\ &\leqq \left(\int_{\Omega_K} c\,dx \cdot \int_{\Omega_K} ce^2\,dx\right)^{1/2} + \left|\oint_{\partial\Omega_K} \left\langle a\frac{\partial e}{\partial n}\right\rangle_{1-\alpha} ds\right|\end{aligned}$$

and hence

$$\Lambda_K(\nabla\varphi_K) \leqq \left(\int_{\Omega_K} ce^2\,dx\right)^{1/2} + \left(\int_{\Omega_K} c\,dx\right)^{-1/2} \left|\oint_{\partial\Omega_K} \left\langle a\frac{\partial e}{\partial n_K}\right\rangle_{1-\alpha} ds\right| \qquad \square$$

Lemma 5.4. *Let $\psi \in H^1(\Omega_K)$ and let*

$$\psi_c = \int_{\Omega_K} c\psi\,dx \Big/ \int_{\Omega_K} c\,dx \tag{5.14}$$

Then there exists a constant $C > 0$ such that

$$\|\psi - \psi_c\|_{0,\Omega_K} \leqq Ch_K |\psi|_{1,\Omega_K} \tag{5.15}$$

where $h_K = \mathrm{diam}(\Omega_K)$.

Proof. If ψ is constant then $\psi_c \equiv \psi$. The result follows from a standard application of the Bramble-Hilbert Lemma. $\square$

Lemma 5.5. *There exists a constant $C > 0$ such that*

$$\begin{aligned}\varepsilon_K(\nabla\varphi_K) \leqq &\left(\int_{\Omega_K} a|\nabla e|^2\,dx\right)^{1/2} + Ch_K\left(\int_{\Omega_K} ce^2\,dx\right)^{1/2} \\ &+ Ch_K^{1/2}\left(\oint_{\partial\Omega_K} \left\langle\frac{\partial e}{\partial n}\right\rangle_{1-\alpha}^2 ds\right)^{1/2}\end{aligned} \tag{5.16}$$

Proof. By the Triangle Inequality

$$\begin{aligned}\varepsilon_K(\nabla\varphi_K) &= \left(\int_{\Omega_K} \frac{1}{a}|\nabla\varphi_K|^2\,dx\right)^{1/2} \\ &\leqq \left(\int_{\Omega_K} a|\nabla e|^2\,dx\right)^{1/2} + \left(\int_{\Omega_K} \frac{1}{a}|\nabla\varphi_K - a\nabla e|^2\,dx\right)^{1/2}\end{aligned}$$

Let $a\nabla\psi = \nabla\varphi_K - a\nabla e$ then

$$\begin{aligned}-\nabla\cdot(a\nabla\psi) &= -\nabla\varphi_K + \nabla\cdot(a\nabla e) = r_K - c\delta_K + \nabla\cdot(a\nabla e) \\ &= Le - c\delta_K + \nabla\cdot(a\nabla e) = c(e - \delta_K)\end{aligned}$$

While on $\partial\Omega_K \backslash \Gamma_D$, using (5.13)

$$a\frac{\partial\psi}{\partial n_K} = \frac{\partial\varphi_K}{\partial n_K} - a\frac{\partial e_K}{\partial n_K} = R_K - a\frac{\partial e_K}{\partial n_K} = -\sigma_{KL}\left\langle a\frac{\partial e}{\partial n}\right\rangle_{1-\alpha}$$

Therefore

$$\oint_{\Omega_K} \frac{1}{a}|a\nabla\psi|^2\,dx = \oint_{\Omega_K} a\frac{\partial\psi}{\partial n_K}\psi\,ds - \oint_{\Omega_K} \psi\,\nabla\cdot(a\nabla\psi)\,dx$$

$$= -\oint_{\partial\Omega_K} \sigma_{KL}\left\langle a\frac{\partial e}{\partial n}\right\rangle_{1-\alpha}\psi\,ds + \int_{\Omega_K} \psi c(e-\delta_K)\,dx$$

$$= -\oint_{\partial\Omega_K} \sigma_{KL}\left\langle a\frac{\partial e}{\partial n}\right\rangle_{1-\alpha}\psi\,ds + \int_{\Omega_K} c\psi e\,dx - \psi_c\delta_K\int_{\Omega_K} c\,dx$$

Examining the proof of Lemma 5.3 we find

$$\delta_K\int_{\Omega_K} c\,dx = -\oint_{\partial\Omega_K} \sigma_{KL}\left\langle a\frac{\partial e}{\partial n}\right\rangle_{1-\alpha} ds + \int_{\Omega_K} ce\,dx$$

and so

$$\int_{\Omega_K} \frac{1}{a}|a\nabla\psi|^2\,dx = -\oint_{\partial\Omega_K} \sigma_{KL}\left\langle a\frac{\partial e}{\partial n}\right\rangle_{1-\alpha}(\psi-\psi_c)\,ds + \int_{\Omega_K} ce(\psi-\psi_c)\,dx$$

$$\leqq \left(\oint_{\partial\Omega_K}\left\langle a\frac{\partial e}{\partial n}\right\rangle_{1-\alpha}^2 ds\right)^{1/2}\|\psi-\psi_c\|_{0,\partial\Omega_K}$$

$$+ C\left(\int_{\Omega_K} ce^2\,dx\right)^{1/2}\|\psi-\psi_c\|_{0,\Omega_K}$$

By Lemma 5.4 and standard Trace Inequalities

$$\|\psi-\psi_c\|_{0,\partial\Omega_K} \leqq Ch_K^{-1/2}\|\psi-\psi_c\|_{0,\Omega_K} \leqq Ch_K^{1/2}|\psi|_{1,\Omega_K}$$

and since a is bounded below by $\underline{a} > 0$,

$$|\psi|_{1,\Omega_K}^2 \leqq \frac{1}{\underline{a}}\int_{\Omega_K}\frac{1}{a}|a\nabla\psi|^2\,dx$$

Therefore

$$\left(\int_{\Omega_K}\frac{1}{a}|a\nabla\psi|^2\,dx\right)^{1/2} \leqq Ch_K^{1/2}\left(\oint_{\partial\Omega_K}\left\langle a\frac{\partial e}{\partial n}\right\rangle_{1-\alpha}^2 ds\right)^{1/2} + Ch_K\left(\int_{\Omega_K} ce^2\,dx\right)^{1/2}$$

Collecting these results gives (5.15). □

The following result complements Theorem 5.2.

Theorem 5.6. *For any $\mu_K > 0$ there exists $C > 0$ such that*

(5.17)
$$\eta_K^2(\nabla\varphi_K) \leqq (1+\mu_K)\|e\|_{E,\Omega_K}^2 + C(1+\mu_K^{-1})\left\{\left\|h_K\left\langle\frac{\partial e}{\partial n}\right\rangle_{1-\alpha}\right\|_{0,\partial\Omega_K}^2 + Ch_K^2\|e\|_{0,\Omega_K}^2 + \left|\oint_{\partial\Omega_K}\sigma_{KL}\left\langle\frac{\partial e}{\partial n}\right\rangle_{1-\alpha} ds\right|^2 \Big/ \int_{\Omega_K} c\,dx\right\}$$

Proof. Follows from previous lemmas and the elementary inequality

$$(a+b)^2 \leqq (1+\mu_K)a^2 + (1+\mu_K^{-1})b^2 \quad \forall \mu_K > 0 \qquad \square$$

This result shows that a necessary condition in order for $\sum_{K=1}^{N} \eta_K^2$ to realistically estimate the true error is that $\langle \partial e/\partial n \rangle_{1-\alpha}$ be sufficiently small. The controlling term in the right hand bound is

$$\left| \oint_{\partial\Omega_K} \sigma_{KL} \left\langle \frac{\partial e}{\partial n} \right\rangle_{1-\alpha} ds \right|^2 \bigg/ \int_{\Omega_K} c\,dx \tag{5.18}$$

Notice that the denominator behaves like means $(\Omega_K)^{-1}$ as $h_K \to 0$. In particular there is a danger that the right hand side will blow up as $h_K \to 0$, making the bound meaningless.

6. Flux splitting for finite element approximations

In this section we suppose u_h is an approximation obtained using the finite element method with Ω_K corresponding to the elements. A related approach has been developed in [10]; but the types considered here of finite element schemes may be *h*, *p*, or *h-p* versions, including *k*-irregular meshes [3].

Let $\mathscr{F}(\mathscr{P})$ denote the set of *unconstrained or proper nodes* in the partition $\mathscr{P}$, see [3] for details. Let $A \in \mathscr{F}(\mathscr{P})$ and let L_A be the degree one basis function associated with node A. We suppose L_A to be scaled so that it takes the value 1 at the node A. Let S_A denote the patch of elements forming the support of L_A

$$S_A = \{\Omega_K : \Omega_K \cap \text{supp}(L_A) \text{ is non-empty}\} \tag{6.1}$$

and let Q_A denote the boundary segments lying on the interior of the support of L_A ($\text{supp}^0 L_A$)

$$Q_A = \{\Gamma_{KL} : \Gamma_{KL} \subset \text{supp}^0(L_A) \text{ or } \Gamma_{KL} \subset \Gamma_N \cap \text{supp}(L_A)\} \tag{6.2}$$

In the following, we shall show how α_{KL} may be constructed such that the scalar,

$$\Delta_K = \delta_K \int_{\Omega_K} c\,dx = \int_{\Omega_K} r_K\,dx + \oint_{\partial\Omega_K} R_K\,ds \tag{6.3}$$

vanishes on all elements $\Omega_K : \partial\Omega_K \cap \Gamma_D = \varnothing$.

Let $A \in \mathscr{F}(\mathscr{P})$. For each $\Omega_K \in S_A$ define

$$R_K^A = \int_{\Omega_K} r_K L_A\,dx \tag{6.4}$$

and for each $\Gamma_{KL} \in Q_A$ define

$$\rho_{KL}^A = \begin{cases} \displaystyle\int_{\Gamma_{KL}} \left[\!\left[a\frac{\partial u_h}{\partial n} \right]\!\right] \gamma L_A\,ds & K, L \neq 0 \\ \displaystyle\int_{\Gamma_{K0}} \left(a\frac{\partial u_h}{\partial n} - g \right) \gamma L_A\,ds & L = 0 \end{cases} \tag{6.5}$$

With each boundary segment $\Gamma_{KL}\in Q_A$ we associate a constant λ^A_{KL}. Notice that both Γ_{KL} and $\Gamma_{LK}\in Q_A$ and it will usually be found that $\lambda^A_{KL} \neq \lambda^A_{LK}$.

Theorem 6.1. *For each $A\in\mathscr{F}(\mathscr{P})$ there exist $\{\lambda^A_{JL}\}_{\Gamma_{JL}\in Q_A}$ such that*

(6.6) $$\text{(i)}\quad \lambda^A_{JL} + \lambda^A_{LJ} \equiv 1 \quad \forall\, \Gamma_{JL}\in Q_A$$

(6.7) $$\text{(ii)}\quad \lambda^A_{J0} \equiv 1 \quad \forall\, \Gamma_{J0}\in Q_A$$

(6.8) $$\text{(iii)}\quad \sum_{L:\Gamma_{KL}\in Q_A} \lambda^A_{KL}\rho^A_{KL} \equiv R^A_K \quad \forall\, \Omega_K\in S_A$$

Proof. Let $\mu^A_{JL} = \lambda^A_{JL}\rho^A_{JL}$, $\Gamma_{JL}\in Q_A$. With this notation, (i)–(iii) become (since $\rho^A_{JL} = \rho^A_{LJ}$)

$$\text{(i)}\quad \mu^A_{JL} + \mu^A_{LJ} \equiv \rho^A_{LJ}$$

$$\text{(ii)}\quad \mu^A_{J0} \equiv \rho^A_{J0}$$

$$\text{(iii)}\quad \sum_{L:\Gamma_{KL}\in Q_A} \mu^A_{KL} = R^A_K$$

We may eliminate the unknowns μ_{LJ}, $L < J$ by using the first two conditions to give, for each $\Omega_K\in S_A$,

(6.9) $$\sum_{\substack{L:\Gamma_{KL}\in Q_A\\ 0<L<K}} \mu^A_{KL} - \sum_{\substack{L:\Gamma_{KL}\in Q_A\\ L>K>0}} \mu^A_{LK} = R^A_K - \sum_{\substack{L:\Gamma_{KL}\in Q_A\\ L>K>0}} \rho^A_{LK} - \sum_{\Gamma_{K0}\in Q_A} \rho^A_{K0}$$

This represents a linear system of $|S_A|$ equations in the unknowns μ^A_{LJ}, $L > J > 0$, $\Gamma_{LJ}\in Q_A$. Let M_A denote the underlying matrix for this system.

We examine the null space $\ker(M^*_A)$ of M^*_A. Suppose $\xi\in\mathbb{R}^{|S_A|}$ is such that $M^*_A\xi \equiv \mathbf{0}$. First, notice that each column of M_A (and therefore each row of M^*_A) has precisely two non-zero entries corresponding to each unknown μ^A_{JL}, $J > L > 0$ occurring once in the equations when $\Omega_K = \Omega_J$ and $\Omega_K = \Omega_L$. Moreover, these non-zero entries are $+1$ and -1. Therefore

$$M^*_A\xi = \mathbf{0} \Leftrightarrow \xi_K = \xi_J \quad \forall\, \Gamma_{KJ}\in Q_A,\ K > J > 0$$

Since $\operatorname{supp}(L_A)$ is connected this is now equivalent to all of the components of ξ being the same constant. That is

$$\ker(M^*_A) = \operatorname{span}\{\lambda\}$$

where $\lambda = (1, 1, \ldots, 1)\in\mathbb{R}^{|S_A|}$.

By the Fredholm Alternative, a necessary and sufficient condition for the existence of solutions to (6.9) is that the data be orthogonal to $\ker(M^*_A)$. That is

(6.10) $$\sum_{\Omega_K\in S_A} R^A_K = \sum_{\Omega_K\in S_A}\ \sum_{\substack{L:\Gamma_{KL}\in Q_A\\ L>K}} \rho^A_{LK} + \sum_{\Omega_K\in S_A}\ \sum_{\Gamma_{K0}\in Q_A} \rho^A_{K0}$$

Now

$$R^A_K = \int_{\Omega_K} r_K L_A\,dx$$

$$= \int_{\Omega_K} f L_A\,dx - \int_{\Omega_K} (a\nabla u_h\cdot\nabla L_A + c u_h L_A)\,dx + \oint_{\partial\Omega_K} a\frac{\partial u_h}{\partial n_K}\gamma L_A\,ds$$

and, hence, since $\bigcup_{\Omega_K \in S_A} \Omega_K = \text{supp}(L_A)$, we have

$$\sum_{\Omega_K \in S_A} R_K^A = (f, L_A) - a(u_h, L_A) + \sum_{\Omega_K \in S_A} \oint_{\partial\Omega_K} a \frac{\partial u_h}{\partial n_K} \gamma L_A \, ds$$

Applying Lemma 3.2 and noting $[\![L_A]\!] = 0$, $\langle L_A \rangle = L_A$ gives

$$\sum_{\Omega_K \in S_A} \oint_{\partial\Omega_K} a \frac{\partial u_h}{\partial n_K} \gamma L_A \, ds = \sum_{\Omega_K \in S_A} \sum_{\Gamma_{KL} \in Q_A} \int_{\Gamma_{KL}} \left[\!\left[a \frac{\partial u_h}{\partial n} \right]\!\right] \gamma L_A \, ds$$

$$= \sum_{\Omega_K \in S_A} \sum_{\substack{L: \Gamma_{KL} \in Q_A \\ L > 0}} \rho_{KL}^A + \sum_{\Omega_K \in S_A} \sum_{\Gamma_{K0} \in Q_A} \int_{\Gamma_{K0}} a \frac{\partial u_h}{\partial n} \gamma L_A(S) \, ds$$

Moreover, by the definition of the Galerkin finite element solution, there holds

$$a(u_h, L_A) = (f, L_A) + \sum_{\Omega_K \in S_A} \sum_{\Gamma_{K0} \in Q_A} \int_{\Gamma_{K0}} g \gamma L_A \, ds$$

Combining these results gives

$$\sum_{\Omega_K \in S_A} R_K^A = \sum_{\Omega_K \in S_A} \sum_{\substack{L: \Gamma_{KL} \in Q_A \\ L > 0}} \rho_{KL}^A + \sum_{\Omega_K \in S_A} \sum_{\Gamma_{K0} \in Q_A} \int_{\Gamma_{K0}} \left(a \frac{\partial u_h}{\partial n} - g \right) \gamma L_A \, ds$$

which in view of (6.5) gives (6.10). It therefore follows that there exists a solution of (i)–(iii). □

Applying the Fredholm Alternative once again, we deduce that there will be an infinite number of solutions in the case of $\Gamma_N \cap \text{supp}^0(L_A)$ being empty. Otherwise the solution will be unique.

The constants λ_{KJ}^A are used to construct the α_{KJ}'s used in the average $\langle \cdot \rangle_\alpha$ and $\langle \cdot \rangle_{1-\alpha}$. Specifically, for each element $\Omega_K \in \mathscr{P}$ and for each $\Gamma_{KJ} \in Q_A$, we define

(6.11)
$$\alpha_{KJ}(s) = \sum_{\substack{A \in \mathscr{F}(\mathscr{P}): \\ \Gamma_{KJ} \in Q_A}} \lambda_{KJ}^A L_A(s)$$

Some special cases of (6.11) for 0- and 1-irregular meshes in two dimensions are shown in Figs. 1 and 2.

Theorem 6.2. *Let α_{KJ} be constructed as in (6.11). Then*

(i) $\alpha_{KJ}(s) + \alpha_{JK}(s) \equiv 1, \quad s \in \Gamma_{KJ} : \Gamma_{KJ} \cap \Gamma_D = \varnothing$

(ii) $\int_{\Omega_K} r_K \, dx + \oint_{\partial\Omega_K} R_K \, ds = 0, \quad \forall \Omega_K : \bar{\Omega}_K \cap \Gamma_D = \varnothing$

Proof. Notice that $\sum_{A \in \mathscr{F}(\mathscr{P})} L_A(x) \equiv 1$, $x \in \Omega_K : \bar{\Omega}_K \cap \Gamma_D = \varnothing$.

1. Let $s \in \Gamma_{KJ}$; then

$$\alpha_{KJ}(s) + \alpha_{JK}(s) = \sum_{\substack{A \in \mathscr{F}(\mathscr{P}) \\ \Gamma_{KJ} \in Q_A}} \lambda_{KJ}^A L_A(s) + \sum_{\substack{A \in \mathscr{F}(\mathscr{P}) \\ \Gamma_{JK} \in Q_A}} \lambda_{JK}^A L_A(s)$$

$$= \sum_{A \in \mathscr{F}(\mathscr{P})} (\lambda_{KJ}^A + \lambda_{JK}^A) L_A(s) \equiv 1$$

since $\lambda_{KJ}^A + \lambda_{JK}^A \equiv 1$.

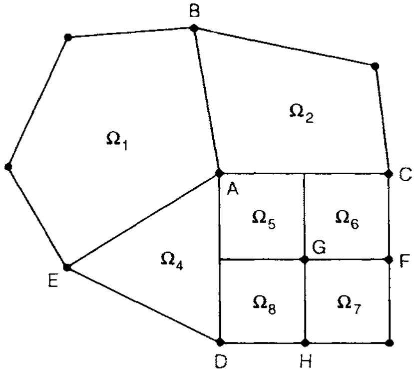

Fig. 1. Flux splitting for 0-irregular mesh (λ^A_{KJ} are as in Theorem 6.1).

$\alpha_{12}(s) = \lambda^A_{12} L_A(s) + \lambda^B_{12} L_B(s)$

$\alpha_{23}(s) = \lambda^A_{23} L_A(s) + \lambda^C_{23} L_C(s)$

$\alpha_{34}(s) = \lambda^A_{34} L_A(s) + \lambda^D_{34} L_D(s)$

$\alpha_{41}(s) = \lambda^A_{41} L_A(s) + \lambda^E_{41} L_E(s)$

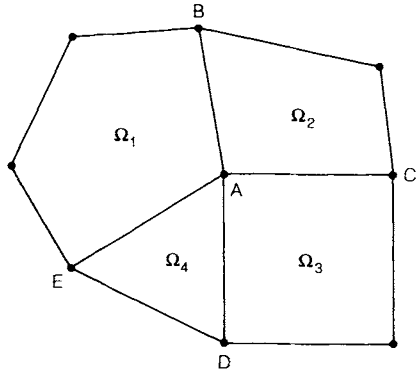

Fig. 2. Flux splitting for 1-irregular mesh (λ^A_{KJ} are as in Theorem 6.1).

$\alpha_{12}(s) = \lambda^A_{12} L_A(s) + \lambda^B_{12} L_B(s)$

$\alpha_{25}(s) = \lambda^A_{25} L_A(s) + \lambda^C_{25} L_C(s)$

$\alpha_{26}(s) = \lambda^A_{26} L_A(s) + \lambda^C_{26} L_C(s)$

$\alpha_{56}(s) = \lambda^A_{56} L_A(s) + \lambda^C_{56} L_C(s) + \lambda^G_{56} L_G(s)$

2.

$$\oint_{\partial\Omega_K} R_K\, ds = -\sum_{\substack{J:\Gamma_{KJ}\in E\\ J>0}} \int_{\Gamma_{KJ}} \alpha_{KJ}\left[\!\left[a\frac{\partial u_h}{\partial n} \right]\!\right] ds + \sum_{\Gamma_{K0}\in E} \int_{\Gamma_{K0}} \left(g - a\frac{\partial u_h}{\partial n} \right) ds$$

$$= -\sum_{\substack{J:\Gamma_{KJ}\in E\\ J>0}} \int_{\Gamma_{KJ}} \sum_{\substack{A\in\mathscr{F}(\mathscr{P})\\ \Gamma_{KJ}\in Q_A}} \lambda^A_{KJ} L_A \left[\!\left[a\frac{\partial u_h}{\partial n} \right]\!\right] ds - \sum_{\Gamma_{K0}\in E} \sum_{\substack{A\in\mathscr{F}(\mathscr{P})\\ \Gamma_{K0}\in Q_A}} \rho^A_{K0}$$

$$= -\sum_{\substack{A\in\mathscr{F}(\mathscr{P})\\ \Gamma_{KJ}\in Q_A}} \lambda^A_{KJ} \rho^A_{KJ}$$

$$= -\sum_{A\in\mathscr{F}(\mathscr{P})} R^A_K = -\sum_{A\in\mathscr{F}(\mathscr{P})} \int_{\Omega_K} r_K L_A\, dx = -\int_{\Omega_K} r_K\, ds$$

and, hence, $\Delta_K = \int_{\Omega_K} r_K\, dx + \oint_{\partial\Omega_K} R_K\, ds = 0.$ □

7. Equivalence of estimator

While Theorem 5.2 assures that our estimate will always bound the error, it is of importance to examine whether the estimator provides an equivalent measure of the error. Theorem 5.6 suggests that unless care is taken to control the quantity $\langle \partial e/\partial n \rangle_{1-\alpha}$, the bound can be very poor. The aim of this section is to show that the construction for α in Sect. 6 does control this quantity effectively.

In addition to the previous assumptions we shall further assume that Ω_K are convex subdomains. Let

$$\mathcal{N}_K = \{\Omega_L \in \mathcal{P}: \partial\Omega_L \cap \partial\Omega_K = \Gamma_{LK} \text{ is non-empty}\} \tag{7.1}$$

Then $\mathcal{N}_K$ is the set of elements neighboring Ω_K. Let

$$l_{KJ} = |\Gamma_{KJ}| = \text{length of } \Gamma_{KJ} \tag{7.2}$$

and suppose that there is a fixed constant $\kappa > 0$ such that for all $\Omega_K \in \mathcal{P}$:

(i)

$$\frac{1}{\kappa} \leqq \frac{l_{KJ}}{h_K} \leqq \kappa \quad \forall J \in \mathcal{N}_K \tag{7.3}$$

This means that the meshes should be *locally* quasi-uniform.

Furthermore, we assume there is a fixed constant M such that

(ii)

$$\operatorname{card}(\mathcal{N}_K) \leqq M \quad \forall \Omega_K \in \mathcal{P} \tag{7.4}$$

where $\operatorname{card}(\mathcal{N}_K)$ indicates the number of elements in the set $\mathcal{N}_K$. This assumption does not exclude the case of k-irregular meshes [3], but it does force the degree of irregularity (k) to be finite.

(iii) Finally, we assume that the number of edges meeting at any given node is uniformly bounded independently of $\mathcal{P}$.

The finite element approximation is assumed to be a piecewise polynomial on each element Ω_K, but we shall not assume the polynomial degree is constant. However, we suppose the maximum degree $\bar{p}$ to be bounded above, independently of $\mathcal{P}$. This assumption excludes the p or h-p versions proper (but is satisfied for essentially every practical implementation of these versions).

These assumptions are placed on the regularity of the mesh. The following represents an assumption on the regularity of the true solution u:

There exists a piecewise polynomial π on $\mathcal{P}$ of degree at most $\bar{p}+1$ such that

$$\|u-\pi\|_E^2 + \left\| h_K^{1/2} \left\langle a\frac{\partial}{\partial n}(u-\pi) \right\rangle_{1-\alpha} \right\|_{0,E_I}^2 \leqq C\|e\|_E^2 \tag{7.5}$$

for some constant C independent of $\mathcal{P}$. This is similar to the *saturation assumption* made in [5], but the present version is weaker in that we do not assume $C \to 0$ as $\mathcal{P}$ is refined. Finally, due to (7.3), we assume that the following *local inverse estimates* for $v \in \mathcal{P}_{\bar{p}}(\Omega_K)$ are valid

$$\left.\begin{aligned} |v|_{0,\Gamma_{JK}} &\leqq C l_{JK}^{1/2} \|v\|_{0,\Omega_K} \\ \left|\frac{\partial v}{\partial n_K}\right|_{0,\Gamma_{JK}} &\leqq C l_{JK}^{-1/2} \|v\|_{1,\Omega_K} \\ |v|_{2,\Omega_K} &\leqq C h_K^{-1/2} \|v\|_{1,\Omega_K} \end{aligned}\right\} \tag{7.6}$$

and that the *trace inequality* holds,

$$\left|\frac{\partial v}{\partial n}\right|^2_{0,\Gamma_{KJ}} \leqq C\left(\frac{1}{l_{KJ}}|v|^2_{1,\Omega_K} + h_K|v|^2_{2,\Omega_K}\right). \tag{7.7}$$

Lemma 7.2. *Under the previous assumptions*

$$\left\| h_K^{1/2}\left\langle a\frac{\partial e}{\partial n}\right\rangle_{1-\alpha}\right\|_{0,E_I} \leqq C\|e\|_E$$

for some constant $C > 0$.

Proof. By the Triangle Inequality

$$\left\| h_K^{1/2}\left\langle a\frac{\partial e}{\partial n}\right\rangle_{1-\alpha}\right\|_{0,E_I} \leqq \left\| h_K^{1/2}\left\langle a\frac{\partial}{\partial n}(u-\pi)\right\rangle_{1-\alpha}\right\|_{0,E_I} + \left\| h_K^{1/2}\left\langle a\frac{\partial}{\partial n}(\pi-u_h)\right\rangle_{1-\alpha}\right\|_{0,E_I}$$

Now

$$\left\| h_K^{1/2}\left\langle a\frac{\partial}{\partial n}(\pi-u_h)\right\rangle_{1-\alpha}\right\|^2_{0,\Gamma_{KJ}} \leqq Ch_K\left(\left\| a\frac{\partial}{\partial n_K}(\pi-u_h)\right\|^2_{0,\Gamma_{KJ}} + \left\| a\frac{\partial}{\partial n_J}(\pi-u_h)\right\|^2_{0,\Gamma_{KJ}}\right)$$

Using the Trace Inequality gives

$$\begin{aligned} h_K\left\| a\frac{\partial}{\partial n_K}(\pi-u_h)\right\|^2_{0,\Gamma_{KJ}} &\leqq Ch_K(l_{KJ}^{-1}|\pi-u_h|^2_{1,\Omega_K} + h_K|\pi-u_h|^2_{2,\Omega_K}) \\ &\leqq Ch_K(l_{KJ}^{-1}|\pi-u_h|^2_{1,\Omega_K} + h_K^{-1}|\pi-u_h|^2_{1,\Omega_K}) \\ &\leqq C\|\pi-u_h\|^2_{E,\Omega_K} \end{aligned}$$

Summing over all edges gives

$$\left\| h_K^{1/2}\left\langle a\frac{\partial}{\partial n}(\pi-u_h)\right\rangle_{1-\alpha}\right\|_{0,E_I} \leqq C\|\pi-u_h\|_E$$

Moreover

$$\|\pi-u_h\|_E \leqq \|\pi-u\|_E + \|e\|_E$$

Therefore using (7.5)

$$\begin{aligned} \left\| h_K^{1/2}\left\langle a\frac{\partial e}{\partial n}\right\rangle_{1-\alpha}\right\|_{0,E_I} &\leqq C\left(\left\| h_K^{1/2}\left\langle a\frac{\partial e}{\partial n}(u-\pi)\right\rangle_{1-\alpha}\right\|_{0,E_I} + \|\pi-u\|_E\right) + C\|e\|_E \\ &\leqq C\|e\|_E \end{aligned}$$

□

Theorem 7.3. *Let* φ_K *be a solution of the problem*

$$-\Delta\varphi_K = r_K \quad \text{in } \Omega_K$$

subject to

$$\gamma\varphi_K = 0 \quad \text{on } \partial\Omega_K \cap \Gamma_D$$

$$\frac{\partial\varphi_K}{\partial n_K} = \begin{cases} -\alpha_{KJ}\left[\!\left[a\dfrac{\partial u_h}{\partial n}\right]\!\right] & \text{on } \partial\Omega_K\backslash\Gamma_N \\ g - a\dfrac{\partial u_h}{\partial n} & \text{on } \partial\Omega_K \cap \Gamma_N \end{cases}$$

where α_{KJ} is constructed as in Sect. 6. Then under the above assumptions there exists a constant C, independent of $\mathscr{P}$, such that

(7.8)
$$\|e\|_E^2 \leqq \sum_{K=1}^{N} \varepsilon_K^2(\nabla\varphi_K) \leqq C\|e\|_E^2$$

where $\bar{h} = \max h_K$ *and*

$$\varepsilon_K^2(\nabla\varphi_K) = \int_{\Omega_K} \frac{1}{a}|\nabla\varphi_K|^2\,dx.$$

Proof. From Theorem 6.2 we have $\delta_K = 0$ on Ω_K: $\bar{\Omega}_K \cap \Gamma_D = \emptyset$ and so

$$\eta_K^2(\nabla\varphi_K) = \varepsilon_K^2(\nabla\varphi_K) + \delta_K^2 \int_{\Omega_K} c(x)\,dx = \varepsilon_K^2(\nabla\varphi_K)$$

Applying Theorem 5.2 gives

$$\|e\|_E^2 \leqq \sum_{K=1}^{N} \eta_K^2(\nabla\varphi_K) = \sum_{K=1}^{N} \varepsilon_K^2(\nabla\varphi_K)$$

By Lemma 5.5 we have

$$\varepsilon_K^2(\nabla\varphi_K) \leqq (1 + Ch_K^2)\|e\|_{E,K}^2 + C\left\| h_K^{1/2}\left\langle \frac{\partial e}{\partial n}\right\rangle_{1-\alpha}\right\|_{0,\partial\Omega_K}^2$$

and by Lemma 7.2 we then obtain

$$\varepsilon_K^2(\nabla\varphi_K) \leqq (1 + Ch_K^2)\|e\|_{E,K}^2 + C\|e\|_{E,K}^2$$

Summing over all elements gives

$$\sum_{K=1}^{N} \varepsilon_K^2(\nabla\varphi_K) \leqq C\|e\|_E^2.$$ □

8. Summary and examples

The foregoing analysis can be regarded as consisting of two main sections.

The culmination of the first part is Theorem 5.2 which states that the error estimator generated using the local element residual method should always provide an upper bound on the true error, so long as the boundary conditions do not entail any loss in flux, that is to say, condition (3.21) holds. The most common type of element residual method is to choose the symmetrical splitting factor $\frac{1}{2}$. However, further analysis reveals that the estimator, while bounding the error, can be very pessimistic unless the boundary conditions for the local problem are chosen carefully (see comments which follow (5.18)). The upper bound property was conjectured by Bank and Weiser [5], on the basis of numerical experiment for the case of piecewise linear approximation on triangles.

The second part of the work then focuses on the determination of boundary conditions used in the local problems, and, in particular, on the choice of splitting which determines the boundary conditions. It is shown that there exists splittings which mean that the term previously leading to gross overestimation will now vanish. One by-product of this work is that the "equilibration" used by Kelly [9] in one dimension is extended to higher dimensions, more general operators and irregular meshes.

Throughout we have assumed that the operator is of the form

$$Lu \equiv -\nabla\cdot(a(x)\nabla u) + c(x)u$$

with $c > 0$. However, if the boundary conditions are chosen as suggested so that Theorem 6.2 is valid, then the restriction $c > 0$ can be relaxed, meaning that both the theory and the method extend to problems with no absolute terms such as Poisson problems.

In the one-dimensional case, the splitting factors were given explicitly by Kelly [9]. Suppose that the element I_K is the interval (x_K, x_{K+1}). In this case we need only choose the splitting at each node. Let

$$J_K = [\![a u_h'(x_K)]\!]$$

and define

$$\alpha_{K,K-1} = \frac{R_K^K}{J_K}.$$

Using the standard orthogonal projection property of finite element approximation one can easily show that

$$\alpha_{K-1,K} = 1 - \alpha_{K,K-1} = \frac{R_{K-1}^K}{J_K}.$$

This choice of splitting then leads to the satisfaction of the condition of the condition $\Delta_K = 0$. In order to illustrate the necessity of employing the equilibration procedure, we consider the simple problem of finding u:

$$-u'' + u = f \quad \text{on } (0, 1)$$

subject to

$$u(0) = u(1) = 0$$

The function f is chosen so that the true solution is of the form

$$u(x) = x^7 + 10(1-x)^8 - x - 10(1-x)$$

We present results of approximating this problem on uniform meshes with elements of uniform degree. The results in Table 1 show the effectivity indices (ratio of estimated to true error) in the case of symmetrical splitting ($\alpha = \frac{1}{2}$). The results for the cases $p = 2$ and $p = 4$ are seen to be unsatisfactory owing to the poor approximation to the boundary flux obtained using a simple averaging between neighboring elements. In Table 2, we give the corresponding effectivity indices for the splitting described above.

Table 1. Effectivity indices for splitting $\alpha = 1/2$

Degree (p)	Uniform mesh spacing				
	1/4	1/8	1/16	1/32	1/64
1	1.529	1.272	1.088	1.017	1.000
2	5.536	6.238	7.264	7.828	8.127
3	1.497	1.211	1.074	1.026	1.007
4	4.494	5.395	5.816	5.995	5.670

Table 2. Effectivity indices for splitting $\Delta_K = 0$

Degree (p)	Uniform mesh spacing		
	1/4	1/8	1/16
1	1.003	1.0007	1.000
2	1.001	1.0002	1.000
3	1.024	1.0006	1.000
4	1.167	1.166	1.165

For the case of Poisson's equation, Kelly attempted to satisfy the condition $\Delta_K = 0$ by means of a global minimization of the functional $\sum \Delta_K^2$ over the splittings α subject to the condition (3.21). It was found that the objective functional could be driven to zero to machine accuracy in each case. This comes as no surprise in view of Theorem 6.2 above. Numerical results given by Kelly [9] show that the constant appearing in (7.8) is close to unity.

Acknowledgements. The support of DARPA and ONR under Grants N00014-89-J-1451 and N00014-89-J-3109 is gratefully acknowledged.

References

1. Adams, R.A. (1975): Sobolev spaces, Academic Press, New York
2. Ainsworth, M., Craig, A.W. (1992): A posteriori error estimators in the finite element method. Numer. Math. **60,** 429–463
3. Babuška, I., Miller, A.D. (1987): A feedback finite element method with a posteriori error estimation: Part 1. Comput. Methods Appl. Mech. Eng. **61**, pp. 1–40
4. Babuška, I., Rheinboldt, W.C. (1978): A posteriori error estimates for adaptive finite element computations. SIAM J. Numer. Anal. **15**, pp. 736–754
5. Bank, R.E., Weiser, A. (1985): Some a posteriori error estimators for elliptic partial differential equations, Math. Comput. **44**, 283–301
6. Demkowicz, L., Oden, J.T., Strouboulis, T. (1985): An adaptive *p*-version finite element method for transient flow problems with moving boundaries. In: R.H. Gallagher, G. Carey, J.T. Oden, O.C. Zienkiewicz, eds., Finite elements in fluids, Vol. 6. Wiley, Chichester, pp. 291–306
7. Fraeijs de Veubeke, B. (1974): Variational principles and the patch test, Int. J. Numer. Methods Eng. **8**, 783–801
8. Girault, V., Raviart, P.A. (1986): Finite element methods for Navier Stokes equations: Theory and algorithms. Springer Series in Computational Mathematics **5**. Springer, Berlin Heidelberg New York
9. Kelly, D.W. (1984): The self equilibration of residuals and complementary a posteriori error estimates in the finite element method. Int. J. Numer. Methods Eng. **20**, 1491–1506
10. Ladeveze, D., Leguillon (1983): Error estimate procedure in the finite element method and applications. SIAM J. Numer. Anal. **20**, 485–509
11. Oden, J.T., Demkowicz, L., Rachowicz, W., Westermann, T.A. (1989): Towards a universal *h-p* finite element strategy. Part II. A posteriori error estimation, Comput. Methods Appl. Mech. Eng. **77**, 113–180
12. Oden, J.T., Demkowicz, L., Strouboulis, T., Devloo, P. (1986): Adaptive methods for problems in solid and fluid mechanics. In: I. Babuška, O.C. Zienkiewicz, J. Gago, E.R. de A. Oliveira, eds., Accuracy estimates and adaptive refinements in finite element computations, Wiley Chichester, pp. 249–280.
13. Percell, P., Wheeler, M.F. (1978): A local finite element procedure for elliptic equation. SIAM J. Numer. Anal. **15**, pp. 705–714
14. Raviart, P.A., Thomas, J.M. (1977): Primal hybrid finite element methods for 2nd order elliptic equations. Math. Comput. **31**, 391–413