SIAM J. NUMER. ANAL.
Vol. 19, No. 4, August 1982

0036-1429/82/1904-0006 $01.00/0

AN INTERIOR PENALTY FINITE ELEMENT METHOD WITH DISCONTINUOUS ELEMENTS*

DOUGLAS N. ARNOLD†

Abstract. A new semidiscrete finite element method for the solution of second order nonlinear parabolic boundary value problems is formulated and analyzed. The test and trial spaces consist of discontinuous piecewise polynomial functions over quite general meshes with interelement continuity enforced approximately by means of penalties. Optimal order error estimates in energy and L^2-norms are stated in terms of locally expressed quantities. They are proved first for a model problem and then in general.

1. Introduction. In this paper we define a semidiscrete finite element procedure for the numerical solution of a second order parabolic initial-boundary value problem. The piecewise polynomial trial functions employed are, in general, discontinuous. Approximate continuity is imposed by including penalty terms in the form which defines the method. These are weighted L^2 inner products of the jumps in the function values across element edges. In the case of Dirichlet boundary conditions, the penalty terms on the boundary of the domain penalize the deviation of the approximate solution from the specified value of the true solution, exactly as in a well-known method of Nitsche [9].

The primary motivation for the interior penalty method is the enhanced flexibility afforded by discontinuous elements. This allows meshes which are more general in their construction and degree of nonuniformity than is permitted by more conventional finite element methods. Moreover, the local nature of the trial space and the capability to regulate the degree of smoothness of the approximate solution by local variation of the penalty weighting function should enable closer approximation of solutions which vary in character from one part of the domain to another and should allow the incorporation of partial knowledge of the solution into the scheme. An important particular class of difficult equations is that of parabolic equations with dominant transport terms for which the solution varies rapidly on a small moving part of the domain.

The inclusion of penalty terms in the variational form defining a finite element method is not new. The method of Nitsche referred to above and the penalty method of Babuska [3] both employ this technique in order to impose essential boundary conditions weakly. Zienkiewicz [13] discussed the use of penalties in the formulation of nonconforming methods for fourth order problems for which the trial functions, though continuous, are not contained in H^2. Babuska and Zlámal [4] have presented a scheme implementing this idea using interior penalties analogous to the boundary penalties of Babuska's method to solve the biharmonic equation. More recently, Douglas and Dupont [7] have analyzed a method analogous to ours which uses interior penalties to enforce behavior between C^0 and C^1 on conforming elements for linear elliptic and parabolic problems. Numerical experiments with that method have clearly demonstrated the value of penalties for solving certain problems which have proved intractable to more conventional methods (see, e.g., [7]). Closest to the present method are an interior penalty method which Wheeler [12] has presented and analyzed for second order linear elliptic equations, and a similar procedure due to Baker [5] for the biharmonic equation.

* Received by the editors July 23, 1979, and in final revised form June 15, 1981.

† Department of Mathematics and Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742.

Since we wish to allow meshes which are relatively coarse in some parts of the domain and fine in others, we do not assume quasi-uniformity, and we have stated the error estimates in a manner which relates the size of each finite element to the smoothness of the solution on that element. Thus, if h_T is the diameter of the element T, we have bounded the discretization error by quantities of the form

$$\left(\sum_T h_T^{2j}\|w\|_{H^k(T)}^2\right)^{1/2}$$

rather than the more usual

$$(\max_T h_T)^j\left(\sum_T \|w\|_{H^k(T)}^2\right)^{1/2}.$$

(In fact, we shall even allow j and k to depend on T.) This provides motivation and some justification for schemes incorporating adaptive mesh refinement, in which a new mesh is selected from time to time using partial knowledge of the solution to equalize $h_T^j\|w\|_{H^k(T)}$. We also feel that finite element methods based on discontinuous elements will prove more amenable to such adaptive schemes than do conforming methods.

The problem considered for most of the paper is the initial-boundary value problem

$$\begin{aligned} &w_t(x,t)-\nabla\cdot[a(x,t,w(x,t))\nabla w(x,t)+b(x,t,w(x,t))]\\ &\qquad\qquad\qquad\qquad =f(x,t,w(x,t)),\qquad (x,t)\in\Omega\times I,\\ &w(x,t)=g(x,t),\qquad (x,t)\in\partial\Omega\times I,\\ &w(x,0)=w_0(x),\qquad x\in\Omega. \end{aligned}\tag{1.1}$$

Here Ω is a convex polygon in the plane, $I=[0,t^*]\subset\mathbb{R}$, $a\in C_b^2(\bar\Omega\times I\times\mathbb{R})$, $b\in C_b^1(\bar\Omega\times I\times\mathbb{R})\times C_b^1(\bar\Omega\times I\times\mathbb{R})$, $f\in C_b^1(\bar\Omega\times I\times\mathbb{R})$. ($C_b^n$ is the space of functions with continuous, bounded partial derivatives of order up to n.) It is assumed that $\underline{a}\leqq a(x,t,\rho)\leqq\bar a$ where $\underline{a}$ and $\bar a$ are positive constants and that w and w_t are in $C(I;C^1(\bar\Omega))$ and $w\in L^\infty(I;H^2(\Omega))$. (For the definition of this latter space see § 2.) In fact many of the results proved below are valid so long as Ω is a bounded domain with Lipschitz boundary (cf. [2]), but more involved proofs are required.

The paper is divided into six sections. In the following section some notations are collected and preliminary results relating to the mesh and finite element space are presented. The procedure is discussed in a simplified context in § 3, and is defined and analyzed in general in the following section. In § 5 some generalizations and extensions are considered. Various observations concerning the penalty functions are collected in the final section.

The work presented here is based on the author's thesis [2], written under the guidance of Professor Jim Douglas, Jr., for whose generous help and skillful supervision the author is most grateful.

2. Preliminaries. We shall use the usual L^2-based Sobolev spaces $H^k(S)$ with $\|\cdot\|_{k,S}$ and seminorm $|\cdot|_{k,S}$ and the L^∞-based Sobolev spaces with norm $\|\cdot\|_{W^k_\infty(S)}$. $H_0^1(S)$ denotes the subspace of $H^1(S)$ consisting of functions which vanish on ∂S.

If $S\subseteq\mathbb{R}^2$, $(\cdot,\cdot)_S$ [respectively $\langle\cdot,\cdot\rangle_S$] will denote the inner product in $L^2(S)$ where S is measured by the Lebesgue [respectively, the one-dimensional Hausdorff] measure.

By default, $(\cdot,\cdot)=(\cdot,\cdot)_\Omega$, $\langle\cdot,\cdot\rangle=\langle\cdot,\cdot\rangle_{\partial\Omega}$, $H^k=H^k(\Omega)$, and $\|\cdot\|=\|\cdot\|_{0,\Omega}=\|\cdot\|_{L^2(\Omega)}$.

If K is an interval, X is one of the function spaces introduced above, and ϕ is a function on $\Omega\times K$ then $\|\phi\|_{L^p(K;X)}$ denotes the norm in $L^p(K)$ of the function $t\to\|\phi(\cdot,t)\|_X$. $L^p(X)$ is short for $L^p(I,X)$.

By a (triangular) *mesh* on Ω we mean a finite set $\mathcal{T}$ of closed uniformly nondegenerate triangles with disjoint interiors such that $\bigcup\mathcal{T}=\bar{\Omega}$. By uniformly nondegenerate we mean that there is a positive lower bound K_0, called an *angle bound*, for all angles of all triangles of all meshes under consideration. For a given mesh $\mathcal{T}$ we define

$$\mathcal{E}_0=\{T_1\cap T_2 \mid T_1, T_2\in\mathcal{T} \text{ are distinct, } T_1\cap T_2 \text{ contains at least two points}\}$$

$$\mathcal{E}_\partial=\{T\cap\partial\Omega \mid T\in\mathcal{T},\ T\cap\partial\Omega \text{ contains at least two points}\},$$

$$\mathcal{E}=\mathcal{E}_0\cup\mathcal{E}_\partial,$$

$$\mathcal{E}_T=\{e\in\mathcal{E}\mid e\subset T\},\qquad T\in\mathcal{T}.$$

For $T\in\mathcal{T}$, we set $h_T=\operatorname{diam}(T)$. For $e\in\mathcal{E}$, $l_e=\operatorname{diam}(e)$. Let $h=\max\{h_T\mid T\in\mathcal{T}\}$.

Next we introduce a property relating adjacent triangles. One of the advantages of the present method is that it is not necessary to assume that the mesh be *edge-to-edge*, i.e., that distinct intersecting triangles meet in either a common vertex or common edge. Instead we shall assume only that all meshes considered are uniformly graded. A mesh $\mathcal{T}$ is said to be *graded with grade constant* K_1 if for all $T\in\mathcal{T}$ and all $e\in\mathcal{E}_T$, $K_1 l_e\geqq h_T$. Note that for such a mesh the cardinality of the sets $\mathcal{E}_T$ are uniformly bounded by $3K_1$. An edge-to-edge mesh is graded with grade constant $(\sin K_0)^{-1}$ where K_0 is an angle bound.

Since our finite element space will consist of discontinuous elements, it will not lie in $H^1(\Omega)$ but rather in the piecewise Sobolev space defined by

$$H^l(\mathcal{T})=\{\phi\in L^2(\Omega)\,\big|\,\phi|_T\in H^l(\mathcal{T}) \text{ for all } T\in\mathcal{T}\}.$$

Differential operators will be understood to act on such spaces piecewise and *not* in the sense of distributions. Thus, for example, if $\phi\in H^1(\mathcal{T})$, we view $\nabla\phi$ as a function in $L^2(\Omega)\times L^2(\Omega)$.

For $\phi\in H^1(\mathcal{T})$, we define the jump and average of ϕ—denoted $[\phi]$ and $\{\phi\}$, respectively—as functions on $\bigcup\mathcal{E}$ as follows. For each $e\in\mathcal{E}$, $[\phi], \{\phi\}\in L^2(e)$. If $e\in\mathcal{E}_0$, then $e=T_1\cap T_2$ with n_e exterior to T_1 for some pair of elements (T_1, T_2). Set

$$[\phi]=(\phi|_{T_1})|_e-(\phi|_{T_2})|_e,\qquad \{\phi\}=[(\phi|_{T_1})|_e+(\phi|_{T_2})|_e]/2.$$

If $e\in\mathcal{E}_\partial$, then $[\phi]=\{\phi\}=\phi|_e$.

For $e\in\mathcal{E}_0$, we select one of the two unit normals to e and denote it n_e or simply n. If $e\in\mathcal{E}_\partial$, we choose n_e to point exterior to Ω. In this notation we can state the basic integration by parts formula

$$(2.1)\qquad (\nabla\cdot\phi,\psi)=-(\phi,\nabla\psi)+\langle\phi\cdot n,\psi\rangle+\sum_{e\in\mathcal{E}_0}(\langle\{\phi\}\cdot n,[\psi]\rangle_e+\langle[\phi]\cdot n,\{\psi\}\rangle_e),$$

valid for $\phi\in H^1(\mathcal{T})\times H^1(\mathcal{T})$ and $\psi\in H^1(\mathcal{T})$.

Our first result is an inequality of Poincaré–Friedrichs type valid for $\phi\in H^1(\mathcal{T})$.

LEMMA 2.1. *Let $\mathcal{T}$ be a mesh on Ω. There exists a constant C depending only on Ω and an angle bound and grade constant for $\mathcal{T}$ such that*

$$(2.2)\qquad \|\phi\|^2\leqq C\Big(\|\nabla\phi\|^2+\sum_{e\in\mathcal{E}} l_e^{-1}\|[\phi]\|_{0,e}^2\Big),\qquad \phi\in H^1(\mathcal{T}).$$

Proof. Define $\psi\in H^2(\Omega)\cap\mathring{H}^1(\Omega)$ by $-\Delta\psi=\phi$. Then [8] there exists a constant

C_1 depending only on Ω such that $\|\phi\|_{2,\Omega} \leqq C_1\|\phi\|$. By (2.1),

$$\|\phi\|^2 = (\phi, -\Delta\psi) = (\nabla\phi, \nabla\psi) - \sum_{e\in\mathscr{E}} \left\langle [\phi], \frac{\partial\psi}{\partial n}\right\rangle_e$$

$$\leqq \|\nabla\phi\|^2 + \sum_{e\in\mathscr{E}} l_e^{-1}\|[\phi]\|_{0,e}^2)^{1/2}\left(\|\nabla\psi\|^2 + \sum_{e\in\mathscr{E}} l_e\left\|\frac{\partial\psi}{\partial n}\right\|_{0,e}^2\right)^{1/2}.$$

Now, as is easily verified by employing an affine transformation onto an isoceles right triangle (see (2.5) below), for each triangle T there is a constant C_2 depending only on an angle bound for T such that

$$l_e\left\|\frac{\partial\psi}{\partial n}\right\|_{0,e}^2 \leqq C_2\|\psi\|_{2,T}^2, \qquad e \in \mathscr{E}_T.$$

Hence $\|\nabla\psi\|^2 + \sum_{e\in\mathscr{E}} l_e\|\partial\psi/\partial n\|_{0,e}^2 \leqq C\|\phi\|^2$ where C has the stated dependence. Then (2.2) follows. $\square$

For the remainder of the paper we concentrate—for simplicity of notation—on a fixed mesh $\mathscr{T}$ with angle bound K_0 and grade constant K_1. Let r be a positive integer. For each $T \in \mathscr{T}$ let $\mathscr{P}_r(T)$ be the set of restrictions to T of polynomials of degree at most r. There exists a constant C depending only on r and K_0 such that the local inverse inequalities

$$\|\phi\|_{1,T}^2 \leqq Ch_T^{-2}\|\phi\|_{0,T}^2, \qquad \|\phi\|_{0,\partial T}^2 \leqq Ch_T^{-1}\|\phi\|_{0,T}^2,$$

$$\left\|\frac{\partial\phi}{\partial n}\right\|_{0,T}^2 \leqq Ch_T^{-1}\|\nabla\phi\|_{0,T}^2, \qquad \|\phi\|_{W_\infty^1(T)}^2 \leqq Ch_T^{-1}\|\phi\|_{1,T}^2$$

are valid for $\phi \in \mathscr{P}_r(T)$.

If ϕ is continuous on the triangle T, define $\mathscr{I}_T\phi$ to be the unique function in $\mathscr{P}_r(T)$ interpolating ϕ at the $(r+1)(r+2)/2$ points of T with barycentric coordinates in $\{0, 1/r, 2/r, \cdots, 1\}$. Then

$$\|\phi - \mathscr{I}_T\phi\|_{i,T} \leqq Ch_T^{j-i}\|\phi\|_{j,T}, \qquad 0 \leqq i \leqq j \leqq r+1, \quad j \geqq 2, \tag{2.3}$$

where C depends only on r and K_0.

The finite element space we shall employ is $\mathscr{M} = \prod_{T\in\mathscr{T}} \mathscr{P}_r(T)$. We shall identify an element $(\chi_{T\in T})_{\mathscr{T}}$ of this space with the element of $L^2(\Omega)$ defined unambiguously on $\Omega\backslash\cup\mathscr{E}_0$ by $\chi|_T = \chi_T$, $T \in \mathscr{T}$. Define $\mathscr{I}: H^2(\mathscr{T}) \to \mathscr{M}$ by

$$(\mathscr{I}\phi)|_T = \mathscr{I}_T(\phi|_T), \qquad T \in \mathscr{T}.$$

Now for $e \in \mathscr{E}_T$ the trace inequalities

$$\|\phi\|_{0,e}^2 \leqq C(l_e^{-1}\|\phi\|_{0,T}^2 + l_e|\phi|_{1,T}^2), \qquad \phi \in H^1(T), \tag{2.4}$$

$$\left\|\frac{\partial\phi}{\partial n}\right\|_{0,e}^2 \leqq C(l_e^{-1}|\phi|_{1,T}^2 + l_e|\phi|_{2,T}^2), \qquad \phi \in H^2(T), \tag{2.5}$$

are valid [1, Thm. 3.10]. Combining with (2.3) we get for all $\phi \in H^j(\mathscr{T})$

$$\begin{aligned}\|\phi - \mathscr{I}\phi\|_{1,T}^2 + \sum_{e\in\mathscr{E}_T}\left[l_e^{-1}\|\phi - \mathscr{I}\phi\|_{0,e}^2 + l_e\left\|\frac{\partial}{\partial n}(\phi - \mathscr{I}\phi)\right\|_{0,e}^2\right]\\ \leqq Ch_T^{2(j-1)}\|\phi\|_{j,T}^2, \qquad 2 \leqq j \leqq r+1,\end{aligned} \tag{2.6}$$

where C depends on r, K_0, and K_1.

In order to derive a weak formulation of (1.1) we note that (2.1) implies

$$-(\nabla\cdot[a(w)\nabla\phi],\psi)=(a(w)\nabla\phi,\nabla\psi)-\sum_{e\in\mathscr{E}_0}\left\langle a(w)\frac{\partial\phi}{\partial n},[\psi]\right\rangle-\left\langle a(g)\frac{\partial\phi}{\partial n},\psi\right\rangle$$

for $\phi\in H^2(\Omega)$ and $\psi\in H^1(\mathscr{T})$. If we symmetrize the form appearing on the right hand side and replace the unknown solution w by a function ρ, we arrive at the form

$$A(\rho;\phi,\psi)=(a(\rho)\nabla\phi,\nabla\psi)-\sum_{e\in\mathscr{E}_0}\left[\left\langle a(\{\rho\})\left\{\frac{\partial\phi}{\partial n}\right\},[\psi]\right\rangle_e+\left\langle a(\{\rho\})[\phi],\left\{\frac{\partial\psi}{\partial n}\right\}\right\rangle_e\right]$$
$$-\left[\left\langle a(g)\frac{\partial\phi}{\partial n},\psi\right\rangle+\left\langle a(g)\phi,\frac{\partial\psi}{\partial n}\right\rangle\right]$$

defined for $\rho,\phi,\psi\in H^2(\mathscr{T})$.

Penalties will be introduced via the form

$$J(\phi,\psi)=\sum_{e\in\mathscr{E}}l_e^{-1}\langle\sigma[\phi],\psi\rangle_e,\qquad \phi,\psi\in H^1(\mathscr{T}),$$

where $\sigma:\bigcup\mathscr{E}\times I\to[\gamma_0,\infty)\subset(0,\infty)$ is a measurable function, differentiable in t when viewed as a function into $L^\infty(\bigcup\mathscr{E})$. J depends on t through σ. Note that the definitions of A and J are independent of the choice of the interior normals n_e.

We also set

$$B(\rho;\cdot,\cdot)=A(\rho;\cdot,\cdot)+J(\cdot,\cdot),\qquad B(\cdot,\cdot)=B(w;\cdot,\cdot).$$

It follows from (2.1) that the solution w satisfies

$$(2.7)\qquad \begin{aligned}&(w_t,\chi)+B(w,\chi)+(b(w),\nabla\chi)-\sum_{e\in\mathscr{E}_0}\langle b(w)\cdot n,[\chi]\rangle_e-\langle b(g)\cdot n,\chi\rangle\\&\quad=(f(w),\chi)-\left\langle a(g)g,\frac{\partial}{\partial n}\right\rangle+\sum_{e\in\mathscr{E}_\partial}l_e^{-1}\langle\sigma g,\chi\rangle\end{aligned}$$

for all $\chi\in H^2(\mathscr{T})$. Here and elsewhere we suppress some of the arguments of the coefficients in the notation.

On the space $H^l(\mathscr{T})$ we place the obvious norm:

$$\|\phi\|_{l,\mathscr{T}}=\left(\sum_{T\in\mathscr{T}}\|\phi\|_{l,\mathscr{T}}^2\right)^{1/2}.$$

The following norm, which incorporates a measure of discontinuity into the $H^1(\mathscr{T})$ norm, is naturally associated with the form A. Define

$$|||\phi|||^2=\|\phi\|_{1,\mathscr{T}}^2+\sum_{e\in\mathscr{E}}\left(l_e^{-1}\|[\phi]\|_{0,e}^2+l_e\left\|\left\{\frac{\partial\phi}{\partial n}\right\}\right\|_{0,e}^2\right),$$

for $\phi\in H^2(\mathscr{T})$. We have immediately the inequality

$$|A(\rho;\phi,\psi)|\leqq\bar{a}|||\phi|||\,|||\psi|||,\qquad \rho,\phi,\psi\in H^2(\mathscr{T}).$$

The following lemma shows that restricted to $\mathscr{M}$, $|||\cdot|||$ is equivalent to a simpler norm.

LEMMA 2.2. *There exists a constant C depending only on K_0 such that*

$$|||\phi|||^2\leqq C\left(\|\phi\|_{1,\mathscr{T}}^2+\sum_{e\in\mathscr{E}}l_e^{-1}\|[\phi]\|_{0,e}^2\right),\qquad \phi\in\mathscr{M}.$$

The proof follows directly from the inverse inequality

$$\sum_{e\in\mathscr{E}} l_e \left\| \left\{ \frac{\partial\phi}{\partial n} \right\} \right\|_{0,e}^2 \leqq C\|\phi\|_{1,T}^2, \qquad \phi \in \mathscr{M}. \tag{2.8}$$

From (2.6) it follows that for integer $j(T)$ with $2 \leqq j(T) \leqq r+1$, $T \in \mathscr{T}$,

$$|||\phi - \mathscr{I}\phi||| \leqq C\Big(\sum_{T\in\mathscr{T}} h_T^{2[j(T)-1]} \|\phi\|_{j(T),T}^2\Big)^{1/2}, \tag{2.9}$$

with C depending only on r, K_0 and K_1.

In the course of the analysis we shall impose restrictions on the penalty function σ. These restrictions will refer to various quantities which are collected here for reference.

$$\gamma_0 = \text{a positive lower bound for } \sigma,$$
$$\gamma_1 = \sup\{\sigma(x,t) \big| x \in \cup\mathscr{E}, t \in I\},$$
$$\gamma_2 = \sup\{|\sigma_t(x,t)| \big| x \in \cup\mathscr{E}, t \in I\}.$$

When the statement of a result refers to some γ_i, it is tacitly assumed that γ_i exists and is finite.

As has already been seen, the letter C will denote a generic constant, changing from appearance to appearance. Occasionally its dependence will be noted explicitly. Other times the dependence will be indicated implicitly; e.g., $C(r, K_0, K_1)$. However, dependence on r, K_0, K_1, $\bar{a}$, $\underline{a}$, Ω, t^*, γ_0, $\|a\|_{W^2_\infty}$, $\|b\|_{W^1_\infty \times W^1_\infty}$, $\|f\|_{W^1_\infty}$, $\|w\|_{L^\infty(W^1_\infty)}$ and $\|w_t\|_{L^\infty(W^1_\infty)}$ will not necessarily be noted. Similar remarks apply to ε, which will be used to denote a generic small positive constant.

3. A model problem. In this section we consider an interior penalty method for the heat equation. The results we obtain here are special cases of sharper ones which will be proven in the following sections, and a number of dispensable assumptions are made in the interest of simplicity.

Let w be a smooth function satisfying the heat equation

$$\begin{aligned} w_t - \Delta w &= 0 \quad \text{on } \Omega \times I, \\ w &= 0 \quad \text{on } \partial\Omega \times I, \\ w(\cdot, 0) &= w_0 \quad \text{on } \Omega. \end{aligned}$$

We assume that the mesh $\mathscr{T}$ is edge-to-edge. It need not be quasi-uniform, but the estimates will not be stated in a manner which reflects the advantage of local refinement.

We consider only a constant penalty function

$$\sigma(x,t) = \gamma_0, \qquad x \in \Omega, \quad t \in I.$$

Thus,

$$A(\phi,\psi) = (\nabla\phi, \nabla\psi) - \sum_{e\in\mathscr{E}} \left[\left\langle \left\{ \frac{\partial\phi}{\partial n} \right\}, [\psi] \right\rangle_e + \left\langle [\phi], \left\{ \frac{\partial\psi}{\partial n} \right\} \right\rangle_e \right],$$

and

$$J(\phi,\psi) = \gamma_0 \sum_{e\in\mathscr{E}} l_e^{-1} \langle [\phi], [\psi] \rangle_e.$$

Clearly,

$$|A(\phi, \psi)| \leqq |||\phi||| \, |||\psi|||,$$
$$|J(\phi, \psi)| \leqq \gamma_0 |||\phi||| \, |||\psi|||,$$

for $\phi, \psi \in H^2(\mathcal{T})$.

Let C_1 and C_2 be the constants appearing in (2.9) and (2.2), respectively. Then

$$\begin{aligned} A(\phi, \phi) &\geqq \|\nabla \phi\|^2 - 2\Big(\sum l_e \Big\| \Big\{\frac{\partial \phi}{\partial n}\Big\} \Big\|_{0,e}^2 \Big)^{1/2} \big(\sum l_e^{-1} \|[\phi]\|_{0,e}^2\big)^{1/2} \\ &\geqq \tfrac{1}{2}\|\nabla\phi\|^2 - 2C_1 \sum l_e^{-1} \|[\phi]\|_{0,e}^2 \\ &\geqq \tfrac{1}{4}\|\nabla\phi\|^2 + \frac{1}{4C_1}\sum l_e \Big\| \Big\{\frac{\partial^{\varphi}}{\partial n}\Big\}\Big\|_{0,e}^2 - 2C_1 \sum l_e^{-1}\|[\phi]\|_{0,e}^2 \\ &\geqq \tfrac{1}{8}\|\nabla\phi\|^2 + \frac{1}{8C_2}\|\phi\|^2 + \frac{1}{4C_1}\sum l_e \Big\| \Big\{\frac{\partial \phi}{\partial n}\Big\}\Big\|_{0,e}^2 - (2C_1 + \tfrac{1}{8}) \sum l_e^{-1} \|[\phi]\|_{0,e}^2, \end{aligned}$$

for all $\phi \in \mathcal{M}$. Now assume that $\gamma_0 \geqq 4C_1 + \frac{1}{2}$. Then

$$(3.1) \qquad B(\phi, \phi) \geqq \varepsilon_1 |||\phi|||^2 + \tfrac{1}{2} J(\phi, \phi), \qquad \phi \in \mathcal{M},$$

where $\varepsilon_1 = \min(1/8, 1/4C_1, 1/8C_2) > 0$.

Now w satisfies

$$(3.2) \qquad (w_t, \chi) + B(w, \chi) = 0, \qquad \chi \in H^2(\mathcal{T}).$$

The interior penalty finite element approximation to w is defined analogously as the unique function $W: I \to \mathcal{M}$ such that

$$(3.3) \qquad (W_t, \chi) + B(W, \chi) = 0, \quad \chi \in \mathcal{M}, \qquad W(0) = \mathcal{I} w_0.$$

Upon choice of a basis for $\mathcal{M}$, (3.3) determines W as the solution to an initial value problem for a linear system of ordinary differential equations. If the basis is chosen in the obvious way as the union of bases for each $\mathcal{P}_r(T)$ (with all functions extended to Ω by zero), the linear system is sparse. The primary disadvantage of the interior penalty method compared to a standard finite element method with continuous piecewise polynomial elements of the same degree is that the linear system is larger. In fact the dimension of $\mathcal{M}$ is $(r+1)(r+2)\,\text{card}\,(\mathcal{T})/2$, while the dimension of the continuous subspace of $\mathcal{M}$ is only slightly greater than $r^2\,\text{card}\,(\mathcal{T})/2$.

We now analyze the proposed procedure by the method of energy estimates. Let $\zeta = W - w$. Then from (3.2) and (3.3)

$$(\zeta_t, \chi) + B(\zeta, \chi) = 0, \qquad \chi \in \mathcal{M}.$$

Decompose ζ as $\mu - \nu$ where $\mu = \mathcal{I} w - w$, $\nu = \mathcal{I} w - W$. Note that $[\mu] \equiv 0$ on $\bigcup \mathcal{E} \times I$; thus,

$$(3.4) \qquad (\nu_t, \chi) + B(\nu, \chi) = (\mu_t, \chi) + A(\mu, \chi), \qquad \chi \in \mathcal{M}.$$

Since $\nu(t) \in \mathcal{M}$ we can set $\chi = \nu(t)$, obtaining

$$\frac{1}{2}\frac{d}{dt}\|\nu\|^2 + B(\nu, \chi) = (\mu_t, \nu) + A(\mu, \nu) \leqq C(\|\mu_t\|^2 + |||\mu|||^2) + \frac{\varepsilon_1}{2}|||\nu|||.$$

Therefore, we can apply the coercivity result (3.1) to get

$$\frac{d}{dt}\|\nu\|^2 + \varepsilon_1 |||\nu|||^2 + J(\nu, \nu) \leqq C(\|\mu_t\|^2 + |||\mu|||^2).$$

Since $\nu(0)=0$, integration in time yields

$$\|\nu\|^2_{L^\infty(L^2)}+\int_I |||\nu|||^2\,dt+\int_I J(\nu,\nu)\,dt \leqq C\Big(\|\mu_t\|^2_{L^2(L^2)}+\int_I |||\mu|||\,dt\Big), \tag{3.5}$$

Since $\zeta=\mu-\nu$ and $[\mu]=0$,

$$\|\zeta\|^2_{L^\infty(L^2)}+\int_I |||\zeta|||^2\,dt+\int_I J(\zeta,\zeta)\,dt \leqq C\Big(\|\mu\|^2_{L^\infty(L^2)}+\int_I |||\mu|||^2\,dt+\|\mu_t\|^2_{L^2(L^2)}\Big).$$

Thus error bounds for the finite element approximation to the true solution reduce to error bounds for the piecewise polynomial interpolant. These latter bounds have already been noted in (2.3) and (2.9) and hence we have obtained the following theorem.

THEOREM 3.1. *The error ζ in the interior penalty finite element method for the heat equation satisfies the inequality*

$$\|\zeta\|_{L^\infty(L^2)}+\Big(\int_I |||\zeta|||^2\,dt\Big)^{1/2}+\Big(\int_I J(\zeta,\zeta)\,dt\Big)^{1/2}$$
$$\leqq Ch^r(\|w\|_{L^\infty(H^r)}+\|w\|_{L^2(H^{r+1})}+\|w_t\|_{L^2(H^r)}).$$

The constant C depends only on r, K_0, and Ω.

Remark 3.2. This theorem does not supply an optimal order estimate on $\|\zeta\|_{L^\infty(L^2)}$. In § 4 we use the technique of comparing W to an interior penalty elliptic projection to derive an $O(h^{r+1})$ bound on ζ in $L^\infty(L^2)$.

The choice $\chi=\nu_t$ in (3.4) leads to the second energy estimate. Then,

$$\|\nu_t\|^2+\frac{1}{2}\frac{d}{dt}B(\nu,\nu)\leqq\frac{1}{2}\|\mu_t\|^2+\frac{1}{2}\|\nu_t\|^2+A(\mu,\nu_t),$$

so

$$\int_0^t \|\nu_t\|^2\,dt+B(\nu(t),\nu(t))\leqq\|\mu_t\|^2_{L^2(L^2)}+2\Big|\int_0^t A(\mu,\nu_t)\,dt\Big|. \tag{3.6}$$

The final term may be integrated by parts in time. Hence,

$$2\Big|\int_0^t A(\mu,\nu_t)\,dt\Big|\leqq 2|A(\mu(t),\nu(t))|+2\Big|\int_0^t A(\mu_t,\nu)\,dt\Big|$$
$$\leqq\frac{\varepsilon_1}{2}|||\nu(t)|||^2+\int_I |||\nu|||^2\,dt+C\Big(\sup_I |||\mu|||^2+\int_I |||\mu_t|||^2\,dt\Big).$$

Moreover, by (3.5), $\int |||\nu|||^2\,dt$ can be absorbed into the last term. If (3.1) is applied to (3.6), it follows that

$$\int_0^t \|\nu_t\|^2\,dt+|||\nu(t)|||+J(\nu(t),\nu(t))\leqq C\Big(\sup_I |||\mu|||^2+\int_I |||\mu_t|||^2\,dt\Big).$$

Since this result obviously remains true if we replace ν by μ, it holds also for ζ, giving the following theorem.

THEOREM 3.3. *There exists a constant C depending only on r, K_0, and Ω such that*

$$\|\zeta_t\|_{L^2(L^2)}+\sup_I |||\zeta|||+\sup_I\,[J(\zeta,\zeta)]^{1/2}\leqq Ch^r(\|w\|_{L^\infty(H^{r+1})}+\|w_t\|_{L^2(H^{r+1})}).$$

4. Definition and analysis of the interior penalty method. In light of (2.7) we define the approximate solution $W: I \to \mathcal{M}$ by the equations

$$(W_t, \chi) + B(W; W, \chi) + (b(W), \nabla\chi) - \sum_{e \in \mathscr{E}_0} \langle b(\{W\}) \cdot n, [\chi] \rangle_e$$
(4.1)
$$= (f(W), \chi) - \left\langle a(g)g, \frac{\partial \chi}{\partial n} \right\rangle + \sum_{e \in \mathscr{E}_\partial} l_e^{-1} \langle \sigma g, \chi \rangle_e + \langle b(g) \cdot n, \chi \rangle, \qquad \chi \in \mathcal{M}, \quad t \in I.$$

Once an initial condition is imposed, it follows that, for small t, W is determined uniquely and is computable from f and g as the solution to an initial value problem in ordinary differential equations. It follows from the estimates made in the proof of Theorem 4.3 that the solution persists for all $t \in I$. (These estimates require that γ_0 be sufficiently large.) We assume that the initial value $W(0) \in \mathcal{M}$ satisfies

$$\|W(0) - w_0\|^2 \leqq Ch^2 \sum_{T \in \mathcal{T}} h_T^{2[j(T)-1]} \|w_0\|_{j(T),T}^2, \qquad 2 \leqq j(T) \leqq r+1. \tag{4.2}$$

(That is, (4.2) is supposed to hold for all integer valued functions $T \to j(T) \in [2, r+1]$.) Acceptable choices of $W(0)$ are, for example, the interpolant $\mathcal{I} w_0$ of w_0, the L^2 projection of w_0 into $\mathcal{M}$, or the elliptic projection of w_0 defined by the linear system

$$B(W(0), \chi) = B(w_0, \chi), \qquad \chi \in \mathcal{M}.$$

In the first two cases,

$$\|W(0) - w_0\|^2 \leqq C \sum h_T^{2j(T)} \|w_0\|_{j(T),T}^2, \qquad 1 \leqq j(T) \leqq r+1,$$

while the error estimate (4.2) will be shown for the elliptic projection in Theorem 4.5 below.

Let $\zeta = W - w$. Below we derive estimates on ζ which are extensions of those stated in Theorem 3.1, Remark 3.2, and Theorem 3.3, respectively.

We begin with a coercivity result for the form B.

THEOREM 4.1. *There exists a positive constant ε such that if γ_0 is sufficiently large, then*

$$B(\rho; \phi, \phi) \geqq \varepsilon |||\phi|||^2 + \tfrac{1}{2} J(\phi, \phi)$$

for all $\phi \in \mathcal{M}$ and $\rho \in H^2(\mathcal{T})$.

Proof. For arbitrary $\delta > 0$,

$$A(\rho; \phi, \phi) \geqq \underline{a} \|\nabla\phi\|^2 - 2\bar{a} \sum_{e \in \mathscr{E}} \left\| \left\{ \frac{\partial \phi}{\partial n} \right\} \right\|_{0,e}^2 \|[\phi]\|_{0,e}^2$$
$$\geqq \underline{a} \|\nabla\phi\|^2 - \delta \sum_{e \in \mathscr{E}} l_e \left\| \left\{ \frac{\partial \phi}{\partial n} \right\} \right\|_{0,e}^2 - C\delta^{-1} \sum_{e \in \mathscr{E}} l_e^{-1} \|[\phi]\|_{0,e}^2.$$

Using (2.8) and Lemmas 2.1 and 2.2 we see that the theorem results from a sufficiently small choice of δ. □

Hereafter it is assumed that γ_0 is sufficiently large in the sense of Theorem 4.1. The following lemma will be used repeatedly in the estimates.

LEMMA 4.2. *Let α, β_1, β_2, and γ be real-valued functions on $\bar{\Omega} \times \mathbb{R}$ each of which satisfies a Lipschitz condition with respect to its second argument uniformly over $\bar{\Omega}$, with Lipschitz constant M. Let $\phi \in C^1(T)$ for each $T \in \mathcal{T}$ and set* $\|\phi\|_{W^1_\infty(\mathcal{T})} =$

$\sup_T \|\phi\|_{W^1_\infty(T)}$. *Then there is a constant* $C = C(M, \|\phi\|_{W^1_\infty(\mathcal{T})})$ *such that, for all* ρ_1, ρ_2, $\psi \in H^2(\mathcal{T})$,

$$(4.3)\quad \begin{aligned} &|([\alpha(\rho_1)-\alpha(\rho_2)]\nabla\phi, \nabla\psi)| + |(\beta(\rho_1)-\beta(\rho_2), \nabla\psi)| \\ &\qquad + |(\gamma(\rho_1)-\gamma(\rho_2), \psi)| \leqq C\|\rho_1-\rho_2\| \, \|\psi\|_{1,\mathcal{T}};\end{aligned}$$

$$(4.4)\quad \begin{aligned} &\sum_{e\in\mathscr{E}_0} \left(\left| \left\langle [\alpha(\{\rho_1\})-\alpha(\{\rho_2\})]\left\{\frac{\partial\phi}{\partial n}\right\}, [\psi]\right\rangle_e \right| + |\langle[\beta(\{\rho_1\})-\beta(\{\rho_2\}]\cdot n, [\psi]\rangle_e| \right) \\ &\qquad \leqq C\|\rho_1-\rho_2\| + \left(\sum_{T\in\mathscr{E}} h_T^2\|\rho_1-\rho_2\|_{1,T}^2\right)^{1/2} \left(\sum_{e\in\mathscr{E}_0} l_e^{-1}\|[\psi]\|_{0,e}^2\right)^{1/2};\end{aligned}$$

$$(4.5)\quad \begin{aligned} &\sum_{e\in\mathscr{E}_0} \left| \left\langle [(\{\rho_1\})-(\{\rho_2\})]\left\{\frac{\partial\psi}{\partial n}\right\}, [\phi]\right\rangle_e \right| \\ &\qquad \leqq C \sup_{e\in\mathscr{E}_0} (l_e^{-1}\|[\phi]\|_{L^\infty(e)}) \left[\|\rho_1-\rho_2\| + \left(\sum_{T\in\mathcal{T}} h_T^2\|\rho_1-\rho_2\|_{1,T}^2\right)^{1/2}\right] \\ &\qquad \cdot \left(\sum_{e\in\mathscr{E}_0} l_e \left\|\left\{\frac{\partial\psi}{\partial n}\right\}\right\|_{0,e}^2\right)^{1/2}.\end{aligned}$$

Proof. The inequality (4.3) is clear with $C = M(\|\phi\|_{W^1_\infty(\mathcal{T})}+2)$. For (4.4), note that

$$(4.6)\quad \begin{aligned} &\left| \left\langle [\alpha(\{\rho_1\})-\alpha(\{\rho_2\})]\left\{\frac{\partial\phi}{\partial n}\right\}, [\Psi]\right\rangle_e \right| + |\langle[\beta(\{\rho_1\})-(\{\rho_2\})]\cdot n, [\psi]\rangle_e| \\ &\qquad \leqq M(\|\phi\|_{W^1_\infty(\mathcal{T})}+1) l_e^{1/2}\|\{\rho_1-\rho_2\}\|_{0,e} l_e^{-1/2}\|[\Psi]\|_{0,e}.\end{aligned}$$

Now, by (2.4)

$$\begin{aligned} \sum_{e\in\mathscr{E}} l_e\|\{\rho_1-\rho_2\}\|_{0,e}^2 &\leqq \sum_{T\in\mathcal{T}} \sum_{e\in\mathscr{E}_T} \|(\rho_1-\rho_2)|_T\|_{0,e}^2 \\ &\leqq C \sum_{T\in\mathcal{T}} \|\rho_1-\rho_2\|_{0,T}^2 + h_T^2\|\rho_1-\rho_2\|_{1,T}^2.\end{aligned}$$

Thus (4.4) results from summing (4.6) over $e \in \mathscr{E}_0$. Similarly one proves (4.5). □

We are now ready to prove the first energy estimate.

THEOREM 4.3. *Assume that* (4.2) *holds for a selection of* $j(T)$, $T \in \mathcal{T}$, *satisfying* $2 \leqq j(T) \leqq r+1$. *Then there exists a constant* C *depending on* γ_1 *such that the error satisfies the inequality*

$$(4.7)\quad \begin{aligned} &\|\zeta\|_{L^\infty(L^2)}^2 + \int_I [\|\!|\zeta|\!\|^2 + J(\zeta,\zeta)]\,dt \\ &\qquad \leqq C \sum_{T\in\mathcal{T}} h_T^{2[j(T)-1]}(\|w_t\|_{L^1(H^{j(T)-1}(T))}^2 + \|w\|_{L^2(H^{j(T)}(T))}^2 + h^2\|w_0\|_{j(T),T}^2).\end{aligned}$$

Proof. Subtracting (2.7) from (4.1) yields

$$(4.8)\quad \begin{aligned} (\zeta_t, \chi) + B(W; \zeta, \chi) = &-(b(W)-b(w), \nabla\chi) \\ &+ \sum_{e\in\mathscr{E}_0} \langle[b(\{W\})-b(w)]\cdot n, [\chi]\rangle_e + (f(W)-f(w), \chi) \\ &+ A(w; w, \chi) - A(W; w, \chi), \qquad \chi \in \mathcal{M}.\end{aligned}$$

Set $\mu = \mathcal{I}w - w$ and $\nu = \mathcal{I}w - W$, and substitute

$$(4.9)\quad \mu - \nu = \zeta$$

into (4.8) to obtain

$$(\nu_t, \chi)+B(W; \nu, \chi)=(\mu_t, \chi)+B(W; \mu, \chi)$$
$$+(b(W)-b(W), \nabla\chi)-\sum_{e\in\mathscr{E}_0}\langle[b(\{W\})-b(w)]\cdot n, [\chi]\rangle_e$$
$$-(f(W)-f(w), \chi)+[A(W; w, \chi)-A(w; w, \chi)], \qquad \chi\in\mathscr{M}.$$

Applying Lemma 4.2 and noting that $[w]\equiv 0$ on $\cup\mathscr{E}_0$, we may bound the last four of the six terms in this equation by

$$C(\|\zeta\|^2+\sum_{T\in\mathscr{T}} h_T^2\|\zeta\|_{1,T}^2)+\frac{\varepsilon}{2}|||\chi|||^2$$

where ε is the value furnished by Theorem 4.1.

From the triangle inequality and an inverse inequality, we see that

$$\sum_{T\in\mathscr{T}} h_T^2\|\zeta\|_{1,T}^2 \leqq 2\Big(\sum_{T\in\mathscr{T}} h_T^2\|\nu\|_{1,T}^2+\sum_{T\in\mathscr{T}} h_T^2\|\mu\|_{1,T}^2\Big)\leqq C\Big(\|\nu\|^2+\sum_{T\in\mathscr{T}} h_T^2\|\mu\|_{1,T}^2\Big), \tag{4.10}$$

so

$$(\nu_t, \chi)+B(W; \nu, \chi)\leqq(\mu_t, \chi)+B(W; \mu, \chi)+C(\|\mu\|^2+\|\nu\|^2+h_T^2\|\mu\|_{1,T}^2)+\frac{\varepsilon}{2}|||\chi|||^2$$

for all $\chi\in\mathscr{M}$ and $t\in I$.

We now set $\chi=\nu(t)\in\mathscr{M}$ and apply Theorem 4.1:

$$\begin{aligned}&\frac{d}{dt}\|\nu\|^2+\varepsilon|||\nu|||^2+J(\nu, \nu)\\ &\qquad\leqq 2(\mu_t, \nu)+2B(W; \mu, \nu)+C(\|\mu\|^2+\|\nu\|^2+\sum h_T^2\|\mu\|_{1,T}^2).\end{aligned} \tag{4.11}$$

Dominating $B(W; \mu, \nu)$ by $C(\gamma_1)|||\mu|||^2+(\varepsilon/2)|||\nu|||^2$ and integrating (4.11) over $t\in[0, t_0]\subseteq I$, we get

$$\begin{aligned}&\|\nu(t_0)\|^2+\int_0^{t_0}\Big[\frac{\varepsilon}{2}|||\nu|||^2+J(\nu, \nu)\Big]\,dt\\ &\qquad\leqq\|\nu(0)\|^2+\tfrac{1}{2}\|\nu\|_{L^\infty(L^2)}^2+C\Big(\|\mu_t\|_{L^1(L^2)}^2+\int_0^{t_0}\|\nu\|^2\,dt+\int_I|||\mu|||^2\,dt\Big).\end{aligned}$$

As this holds for all $t_0\in I$, Gronwall's lemma implies that

$$\begin{aligned}&\|\nu\|_{L^\infty(L^2)}^2+\int_I(|||\nu|||^2+J(\nu, \nu))\,dt\\ &\qquad\leqq C\Big(\|\nu(0)\|^2+\|\mu_t\|_{L^1(L^2)}^2+\int_I|||\mu|||^2\,dt\Big).\end{aligned} \tag{4.12}$$

By (4.2),

$$\|\nu(0)\|^2\leqq 2(\|\mu(0)\|^2+\|\zeta(0)\|^2\leqq Ch^2\sum_{T\in\mathscr{T}} h_T^{2[j(T)-1]}\|w_0\|_{J(T),T}^2.$$

Since $\mu_t=\mathscr{I}w_t-w_t$,

$$\|\mu_t\|^2\leqq C\sum h_T^{2[j(T)-1]}\|w_t\|_{j(T)-1,T}^2.$$

Finally, by (2.9),

$$|||\mu|||^2 \leqq C \sum h_T^{2[j(T)-1]} \|w\|_{j(T),T}^2.$$

Hence (4.12) implies the assertion of Theorem 4.3 with ζ replaced by ν. Since the assertion is clear when ζ is replaced by μ, Theorem 4.3 is proven. □

The bound for $\|\zeta\|_{L^\infty(L^2)}$ provided by Theorem 4.3 is not of optimal order in h. To achieve an optimal order bound we use the technique introduced by Wheeler [11] of comparing the approximate solution to an elliptic projection of the true solution. This approach could also be used to produce optimal order estimates of $|||\zeta|||$ and $J(\zeta, \zeta)$, but these results would not be as satisfactory as those of the last section for two reasons. First, we shall have to impose restrictions on the growth of the penalty function σ as a function of time; second, the bounds derived through the projection are not expressed entirely locally and hence are weaker in the case of a family of meshes which is not quasi-uniform.

We begin with a lemma based on a duality argument.

LEMMA 4.4. *Let $t \in I$ be fixed and suppose that $\Phi \in H^2(\mathcal{T})$ satisfies*

$$B(\Phi, \chi) = F(\chi), \qquad \chi \in \mathcal{M},$$

where $F: H^2(\mathcal{T}) \to \mathbb{R}$ is a linear map. Let M_1 and M_2 be constants for which

$$|F(\rho)| \leqq M_1 |||\rho|||, \qquad \rho \in H^2(\mathcal{T}),$$

and

$$|F(\Psi)| \leqq M_2 \|\Psi\|_{2,\Omega}, \qquad \Psi \in H^2 \cap H_0^1.$$

Then

$$\|\Phi\| \leqq C(|||\Phi||| + M_1)h + M_2,$$

where C depends on γ_1.

Proof. Define $\Psi \in H^2 \cap H_0^1$ by the relation

$$-\nabla \cdot (a(w)\nabla\Psi) = \Phi.$$

Then, by regularity results found in [8],

$$\|\Psi\|_{2,\Omega} \leqq C\|\Phi\|, \tag{4.13}$$

the constant depending on $\underline{a}$, $\|a(t, \cdot, \cdot)\|_{W_\infty^1}$ and $\|w(t, \cdot)\|_{W_\infty^1}$. Now, by (2.9) and (4.13),

$$\begin{aligned}
\|\Phi\|^2 &= (\Phi, -\nabla \cdot (a(w)\nabla\Psi)) = (a(w)\nabla\Phi, \nabla\Psi) - \sum_{e\in\mathcal{E}} \left\langle a(w)[\Phi], \frac{\partial\Psi}{\partial n}\right\rangle_e \\
&= B(\Phi, \Psi) = B(\Phi, \Psi - \mathcal{I}\Psi) - F(\Psi - \mathcal{I}\Psi) + F(\Psi) \\
&\leqq |B(\Phi, \Psi - \mathcal{I}\Psi)| + M_1|||\Psi - \mathcal{I}\Psi||| + M_2\|\Psi\|_{2,\Omega} \\
&\leqq C(\gamma_1)|||\Phi|||\,|||\Psi - \mathcal{I}\Psi||| + M_1|||\Psi - \mathcal{I}\Psi||| + M_2\|\Psi\|_{2,\Omega} \\
&\leqq [C(\gamma_1)(|||\Phi||| + M_1)h + M_2]\|\Phi\|.
\end{aligned}$$

□

The next theorem introduces the elliptic projection and contains the analysis of the interior penalty method for an associated elliptic problem. It generalizes to our situation [12, Thm. 1].

THEOREM 4.5. *There exists a unique function $Z: I \to \mathcal{M}$ satisfying*

$$B(Z, \chi) = B(w, \chi), \qquad \chi \in \mathcal{M}.$$

The error $\eta = Z - w$ *satisfies at each* $t \in I$

$$\||\eta\||^2 + J(\eta, \eta) \leqq C \sum_{T \in \mathcal{T}} h_T^{2[j(T)-1]} \|w\|_{j(T),T}^2, \tag{4.14}$$

$$\|\eta\|^2 \leqq Ch^2 \sum_{T \in \mathcal{T}} h_T^{2[j(T)-1]} \|w\|_{j(T),T}^2, \tag{4.15}$$

$$\||\eta_t\||^2 \leqq C\Big[\sum_{T \in \mathcal{T}} h_T^{2[j(T)-1]} (\|w\|_{j(T),T}^2 + \|w_t\|_{j(T),T}^2)\Big], \tag{4.16}$$

$$\|\eta_t\|^2 \leqq Ch^2\Big[\sum_{T \in \mathcal{T}} h_T^{2[j(T)-1]} (\|w\|_{j(T),T}^2 + \|w_t\|_{j(T),T}^2)\Big] \tag{4.17}$$

for $2 \leqq j(T) \leqq r+1$. *The constants depend on* γ_1 *and* γ_2.

Proof. The uniqueness of Z and therefore its existence follow from the positivity of the form B, which was established in Theorem 4.1. Moreover, since

$$B(\eta, \chi) = 0, \qquad \chi \in \mathcal{M},$$

(4.15) is a consequence of (4.14) and Lemma 4.4. To prove (4.14) we apply Theorem 4.1 to $\theta = Z - \mathcal{I}w$. It follows that

$$\||\theta\||^2 + J(\theta, \theta) \leqq CB(\theta, \theta) = CB(w - \mathcal{I}w, \theta) \leqq C(\gamma_1)\||w - \mathcal{I}w\|| \, \||\theta\||.$$

Thus, by (2.9)

$$\||\eta\|| \leqq \||\theta\|| + \||w - \mathcal{I}w\|| \leqq C\||w - \mathcal{I}w\|| \leqq C\Big(\sum_{T \in \mathcal{T}} h_T^{2[(j(T)-1])} \|w\|_{j(T),T}^2\Big)^{1/2}. \tag{4.18}$$

Since $J(\eta, \eta) \leqq \gamma_1 \||\eta\||_1^2$, (4.14) follows from (4.18).

To estimate η_t, differentiate the defining equation for Z to obtain

$$B(\eta_t, \chi) + B'(\eta, \chi) = 0, \qquad \chi \in \mathcal{M},$$

where

$$B'(\phi, \Psi) = \Big(\frac{d}{dt} a(w) \nabla \phi, \nabla \Psi\Big) - \sum_{e \in \mathcal{E}} \Big[\Big\langle \frac{d}{dt} a(w)[\phi], \Big\{\frac{\partial \Psi}{\partial n}\Big\}\Big\rangle_e + \Big\langle \frac{d}{dt} a(w) \Big\{\frac{\partial \phi}{\partial n}\Big\}, [\Psi]\Big\rangle_e\Big]$$
$$+ \sum_{e \in \mathcal{E}} l_e^{-1} \langle \sigma_t[\phi], [\Psi] \rangle_e.$$

Note that

$$|B'(\eta, \rho)| \leqq C(\gamma_2)\||\eta\|| \, \||\rho\||, \qquad \rho \in H^2(\mathcal{T}).$$

Moreover, for $\psi \in H^2 \cap H_0^1$ an integration by parts using (2.1) shows that

$$|B'(\eta, \psi)| = \Big|\Big(-\eta, \nabla \cdot \Big[\frac{d}{dt} a(w) \nabla \psi\Big]\Big)\Big| \leqq C\|\eta\| \, \|\psi\|_{2,\Omega}.$$

Thus, Lemma 4.4 applies, and

$$\|\eta_t\| \leqq C[(\||\eta_t\|| + \||\eta\||)h + \|\eta\|].$$

Consequently, (4.14)–(4.16) imply (4.17), and it remains only to demonstrate (4.16).

Recall that

$$\Big\|\Big|\frac{\partial}{\partial t}(w - \mathcal{I}w)\Big|\Big\|^2 \leqq C \sum_{T \in \mathcal{T}} h_T^{2[j(T)-1]} \|w_t\|_{j(T),T}^2, \qquad 2 \leqq j(T) \leqq r+1. \tag{4.19}$$

Also, by Theorem 4.1,

$$
\begin{aligned}
|||\theta_t|||^2 + J(\theta_t, \theta_t) &\leqq CB(\theta_t, \theta_t) \\
&= C\left[B\left(\left[\frac{\partial}{\partial t}(w - \mathcal{I}w), \theta_t\right) - B'(\eta, \theta_t)\right] \\
&\leqq C(\gamma_1)\left(\left|\left|\left|\frac{\partial}{\partial t}(w - \mathcal{I}w)\right|\right|\right| + |||\eta|||\right)|||\theta_t||| + \sum_{e\in\mathscr{E}} l_e^{-1}|\langle\sigma_t[\eta], [\theta_t]\rangle_e| \\
&\leqq \tfrac{1}{2}\left(|||\theta_t|||^2 + \sum_{e\in\mathscr{E}} l_e^{-1}\|\sigma^{1/2}[\theta_t]\|_{0,e}^2\right) \\
&\quad + C\left(\left|\left|\left|\frac{\partial}{\partial t}(w - \mathcal{I}w)\right|\right|\right|^2 + |||\eta|||^2 + \sum_{e\in\mathscr{E}} l_e^{-1}\|\sigma_t\sigma^{-1/2}[\eta]\|_{0,e}^2\right).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
|||\eta_t|||^2 &\leqq 2\left(|||\theta_t|||^2 + \left|\left|\left|\frac{\partial}{\partial t}(w - \mathcal{I}w)\right|\right|\right|^2\right) \\
&\leqq C\left[\left|\left|\left|\frac{\partial}{\partial t}(w - \mathcal{I}w)\right|\right|\right|^2 + |||\eta|||^2 + \gamma_2^2\gamma_0^{-2}J(\eta, \eta)\right],
\end{aligned}
$$

and (4.16) follows from (4.19) and (4.14). □

In analogy with (4.9) we shall use the decomposition $\eta - \xi = \zeta$ where $\xi = Z - W \in \mathcal{M}$. Substituting this into (4.8) leads to the relation

$$
\begin{aligned}
(\eta_t, \chi) + B(W; \eta, \chi) = (\xi_t, \chi) &+ B(W; \xi, \chi) - (b(W) - b(w), \nabla\chi) \\
&+ \sum_{e\in\mathscr{E}} \langle[b(\{W\} - b(w)]\cdot n, [\chi]\rangle_e + (f(W) - f(w), \chi) \\
&+ A(w; w, \chi) - A(W; w, \chi).
\end{aligned}
$$

Now,

$$
\begin{aligned}
B(W; \eta, \chi) - A(w; w, \chi) + A(W; w, \chi) &= B(W; \eta, \chi) - B(w; w, \chi) + B(W; w, \chi) \\
&= B(W; Z, \chi) - B(w; Z, \chi) \\
&= A(W; Z, \chi) - A(w; Z, \chi).
\end{aligned}
$$

Thus

$$
\begin{aligned}
(\xi_t, \chi) + B(W; \xi, \chi) = (\eta_t, \chi) &+ [A(W; Z, \chi) + A(w; Z, \chi)] + (b(W) - b(w), \nabla\chi) \\
&- \sum_{e\in\mathscr{E}_0} \langle[b(\{W\}) - b(w)]\cdot n, [\chi]\rangle_e - (f(W) - f(w), \chi).
\end{aligned}
\tag{4.20}
$$

The last four of the five terms on the right-hand side of (4.20) can be estimated by Lemma 4.2, and

$$
(\xi_t, \chi) + B(W; \xi, \chi) \leqq (\eta_t, \chi) + C\left(\|\eta\|^2 + \|\xi\|^2 + \sum_{T\in\mathcal{T}} h_T^2\|\eta\|_{1,T}^2\right) + \frac{\varepsilon}{4}|||\chi|||^2,
$$

where C depends on $\|Z\|_{W^1_\infty(\mathcal{T})}$ and $\sup_{e\in\mathscr{E}_0} l_e^{-1}\|[Z]\|_{L^\infty(e)}$ and ε is derived from Theorem 4.1. Note also that the triangle inequality and an inverse inequality have been used in the same manner as in (4.10). If the choice $\chi = \xi$ is made and the argument by which (4.12) was derived from (4.11) is adapted, then the inequality

$$\begin{aligned}\|\xi\|^2_{L^\infty(L^2)} &+ \int_I [|||\xi|||^2 + J(\xi, \xi)]\,dt \\ &\leqq C(\|\xi(0)\|^2 + \|\eta_t\|^2_{L^1(L^2)} + \|\eta\|^2_{L^2(L^2)} + \sum h_T^2 \|\eta\|^2_{L^2(H^1(T))}) \\ &\leqq Ch^2 \sum h_T^{2[j(T)-1]}(\|w_0\|^2_{j(T),T} + \|w\|^2_{L^2(H^{j(T)}(T))} + \|w_t\|^2_{L^1(H^{j(T)}(T))})\end{aligned}$$

results from (4.2) and Theorem 4.5. This estimate together with (4.15) implies the following theorem.

THEOREM 4.6. *There exists a constant C depending on γ_1, γ_2, $\|Z\|_{L^\infty(W^1_\infty(\mathcal{T}))}$ and $\sup_{e\in\mathscr{E}_0} l_e^{-1}\|[Z]\|_{L^\infty(e)}$ such that the error ζ satisfies the inequality*

$$\|\zeta\|_{L^\infty(L^2)} \leqq Ch\left[\sum_{T\in\mathcal{T}} h_T^{2[j(T)-1]}(\|w_0\|^2_{j(T),T} + \|w\|^2_{L^\infty(H^{j(T)}(T))} + \|w_t\|^2_{L^1(H^{j(T)}(T))})\right]$$

for $2 \leqq j(T) \leqq r+1$.

In certain cases the dependence on Z of the constant in this theorem can be suppressed. This dependence was introduced in bounding $|A(W; Z, \chi) - A(w; Z, \chi)|$ via Lemma 4.2. Hence if the coefficient a is independent of w, then the constant can be taken independent of Z. In particular, this is the case if the differential equation is linear or even semilinear.

Also, in the case of a quasi-uniform family of edge-to-edge meshes the dependence of the constant C of Theorem 4.6 on Z reduces to dependence on the solution.

THEOREM 4.7. *Suppose that $\mathcal{I}w$ is continuous. Let $M = \sup_{T\in\mathcal{T}} h/h_T$. Then there exists a constant $C = C(M, \|w\|_{2,\Omega})$ such that*

$$\|Z\|_{W^1_\infty(\mathcal{T})} + \sup_{e\in\mathscr{E}_0} l_e^{-1}\|[Z]\|_{L^\infty(e)} \leqq C, \qquad t \in I.$$

Proof. Set $\theta = Z - \mathcal{I}w$. From an inverse inequality, (4.14), and (2.3), we obtain

$$\begin{aligned}\|\theta\|_{W^1_\infty(\mathcal{T})} = \sup_{T\in\mathcal{T}} \|\theta\|_{W^1_\infty(T)} &\leqq C \sup_{T\in\mathcal{T}} h_T^{-1}\|\theta\|_{1,T} \\ &\leqq CMh^{-1}\|\theta\|_{1,\mathcal{T}} \leqq CMh^{-1}(\|\eta\|_{1,\mathcal{T}} + \|w - \mathcal{I}w\|_{1,\mathcal{T}}) \\ &\leqq CM\|w\|_{2,\Omega}.\end{aligned}$$

Since $\|\mathcal{I}w\|_{W^1_\infty(\Omega)} \leqq C\|w\|_{W^1_\infty(\mathcal{T})}$, $\|Z\|_{W^1_\infty(\mathcal{T})} \leqq C$.

Finally, for $e \in \mathscr{E}_0$ a one-dimensional inverse inequality and (4.14) imply that

$$l_e^{-1}\|[Z]\|_{L^\infty(e)} = l_e^{-1}\|[\theta]\|_{L^\infty(e)} \leqq Cl_e^{-3/2}\|[\theta]\|_{0,e} = Cl_e^{-3/2}\|[\eta]\|_{0,e} \leqq C\|w\|_{2,\Omega}. \qquad \square$$

It is also possible to prove a second energy estimate analogous to that proved in Theorem 3.3 for the heat equation. In the nonlinear case the required estimates are quite lengthy, and so we only state the result. Details may be found in [2].

It is necessary to assume that $W(0)$ is chosen to satisfy

$$|||W(0) - w_0||| \leqq C \sum_{T\in\mathcal{T}} h^{2[j(T)-1]}\|w_0\|^2_{j(T),T}, \qquad 2 \leqq j(T) \leqq r+1.$$

For example, the elliptic projection, the L^2 projection, and the interpolant of w_0 are all satisfactory choices of $W(0)$ [2].

THEOREM 4.8. *There exists $C = C(\gamma_1, \gamma_2, \|\zeta\|_{L^\infty(W^1_\infty(\mathcal{T}))})$ so that for h sufficiently* *small and any selection of integers* $j(T)\in[2, r+1]$, *the error* ζ *satisfies the inequality*

$$\|\zeta_t\|^2_{L^2(L^2)}+\sup_I |||\zeta|||^2+\sup_I J(\zeta,\zeta)$$
$$\leqq C\sum h_T^{2[j(T)-1]}(\|w_t\|^2_{L^2(H^{j(T)-1}(T))}+\|w\|^2_{L^2(H^{j(T)}(T))}+\|w_0\|^2_{j(T),T}).$$

Moreover, if $a(x, t, w)$ *is independent of* w, *then the constant may be chosen independent of* ζ.

5. Extensions. The interior penalty method can be applied more generally than indicated above. For example the analysis can be extended to handle an equation of the form

$$c(x,t,w)\frac{\partial w}{\partial t}-\sum_{p,q=1}^{2}\frac{\partial}{\partial x_p}a_{pq}(x,t,w)\frac{\partial w}{\partial x_q}=f(x,t,w,\nabla w).$$

Neumann boundary conditions are also permissible. Rectangular elements may be used in place of triangular ones, and elements with curved sides may be used along the boundary of a nonpolygonal domain. Most of the analysis applies also to three-dimensional domains with brick elements. Details concerning all these extensions may be found in [2].

We also mention two extensions of the method which exploit its flexibility. The first simple extension allows the degree r to vary from element to element. In the context of discontinuous elements this is an easy matter. This fact was exploited by Percell and Wheeler [10] in their local residual finite element procedure, and they proposed the strategy of using polynomials of low degree subordinate to a fine mesh in regions where the solution is relatively rough and higher degree polynomials subordinate to a coarse mesh in regions of smoothness of the solution.

There is no difficulty in adapting our analysis to allow for this possibility. Given an integer-valued function $T\in\mathcal{T}\mapsto r(T)\geqq 1$, set

$$\mathcal{M}=\{\chi\in L^2(\Omega)\,\big|\,\chi|_T\in\mathcal{M}_{r(T)},\ T\in\mathcal{T}\}.$$

The usual range $2\leqq j(T)\leqq r+1$ should then be replaced with $2\leqq j(T)\leqq r(T)+1$. All the results previously stated remain valid.

Finally we discuss a multipenalty method. If the mesh $\mathcal{T}$ is to be changed from time to time as the character of the solution w changes, it is necessary to interpolate (or project, etc.) the approximate solution from one mesh to another, inevitably introducing interpolation errors. Let us sketch briefly and heuristically how interior penalties can be used to minimize such errors.

Let

$$J_1(\phi,\psi)=\sum_{e\in\mathcal{E}_0} l_e\left\langle\sigma_1\left[\frac{\partial\phi}{\partial n}\right],\left[\frac{\partial\psi}{\partial n}\right]\right\rangle_e,$$

where $\sigma_1\in L^\infty(\bigcup\mathcal{E}_0\times I)$ is a nonnegative function, and set $B_1=B+J_1$. Define $W^1: I\to\mathcal{M}$ by the equations derived from (4.1) by replacing B with B_1. Then it is easy to show, as is indicated below, that $J(W^1, W^1)+J_1(W^1, W^1)$ is bounded by the right-hand side of (4.7). If σ and σ_1 are large on some edge $e\in\mathcal{E}_0$, this estimate tells us that the discontinuities of W^1 and its normal derivative across e_0 are small and decrease with h. Suppose now that $r=1$, so that W^1 is a piecewise linear function. Such a function is determined by its values and those of its normal derivative along a line segment. Hence, when σ and σ_1 are large on e_0, W^1 is essentially the same function on both sides of e. Therefore, if we interpolate W^1 into the mesh derived from $\mathcal{T}$ by removing the edge e, the error should be small.

Conversely, to introduce a new edge into the mesh, we can begin with the penalties at that edge large and reduce them to pass smoothly from the old mesh to the refinement.

For $r \geqq 1$, the same heuristic considerations apply if we use the form $B_r = B + J_1 + J_2 + \cdots + J_r$, where

$$J_k(\phi, \psi) = \sum_{e \in \mathscr{E}_0} l_e^{2k-1} \left\langle \sigma_k \left[\frac{\partial^k \phi}{\partial n^k}\right], \left[\frac{\partial^k \psi}{\partial n^k}\right] \right\rangle_e.$$

Now, for all $\phi \in H^k(T)$ and $e \in \mathscr{E}_T$,

$$\left\| \frac{\partial^k \phi}{\partial n^k} \right\|_{0,e}^2 \leqq C(l_e^{-1} |\phi|_{k,T}^2 + l_e |\phi|_{k+1,T}^2).$$

It follows that, if $\phi \in H^j(T)$ with $k+1 \leqq j \leqq r+1$, then

$$l_e^{2k-1} \left\| \frac{\partial^k}{\partial n^k}(\phi - \mathscr{I}\phi) \right\|_{0,e}^2 \leqq C(l_e^{2k-2} |\phi - \mathscr{I}\phi|_{k,T}^2 + l_e^{2k} |\phi - \mathscr{I}\phi|_{k+1,T}^2).$$

Thus, for $\phi \in H^{r+1}(\mathscr{T})$,

$$(5.1) \quad \sum_{e \in \mathscr{E}} l_e^{2k-1} \left\| \left[\frac{\partial^k}{\partial n^k}(\phi - \mathscr{I}\phi)\right] \right\|_{0,e}^2 \leqq \sum_{T \in \mathscr{T}} h_T^{2[j(T)-1]} \|\phi\|_{j(T),T}^2, \qquad k+1 \leqq j(T) \leqq r+1.$$

In the multipenalty method we define W^k via the form B_k and let $\zeta_k = W^k - w$.

THEOREM 5.1. *There exists a constant depending on* γ_1 *and* $\sup\{|\sigma_i(x,t)| \mid x \in \bigcup \mathscr{E}_0, t \in I, 1 \leqq i \leqq k\}$ *such that*

$$\|\zeta_k\|_{L^\infty(L^2)}^2 + \int_I [|||\zeta_k|||^2 + J(\zeta_k, \zeta_k) + J_1(\zeta_k, \zeta_k) + \cdots + J_k(\zeta_k, \zeta_k)]\, dt$$

$$\leqq C \sum_{T \in \mathscr{T}} h_T^{2[j(T)-1]} (\|w_t\|_{L^1(H^{j(T)-1}(T))}^2 + \|w\|_{L^2(H^{j(T)}(T))}^2 + h^2 \|w_0\|_{j(T),T}^2)$$

for $k+1 \leqq J(T) \leqq r+1$.

The proof of Theorem 4.3, almost unchanged, gives Theorem 5.1. Since $J_i(w, \chi) = 0$ for all $i \geqq 1$ and $\chi \in \mathscr{M}$, the error equation (4.8) holds with B_k and ζ_k replacing B and ζ. Moreover, it is clear from Theorem 4.1 that

$$B_k(\rho; \phi, \phi) \geqq \varepsilon |||\phi|||^2 + \tfrac{1}{2} J_1(\phi, \phi) + J_1(\phi, \phi) + \cdots + J_k(\phi, \phi).$$

Thus, the claimed bounds reduce to bounds on μ, which hold by (5.1).

In a similar manner analogues of the $L^\infty(L^2)$ and second energy estimates can be shown for the multipenalty method.

We note that the form J_1 is exactly the one used by Douglas and Dupont [7] in creating their conforming interior penalty method mentioned in the introduction.

6. The penalty function. In the previous section we suggested an application of the interior penalty method to mesh refinement for which it is clearly valuable to be able to choose the penalty functions with some degree of flexibility. In addition, one of the initial motivations of this study was the possibility of using interior penalties to adjust the smoothness of the approximation to the behavior of the solution. For these reasons we have avoided placing undue restrictions on the penalty function σ, even when this would have simplified the analysis. Let us recall what restrictions have been made. First we have assumed throughout that γ_0, a lower bound for σ, is sufficiently large. This is necessary for the coercivity result of Theorem 4.1 and is entirely to be expected. Our estimates also depend on γ_1, an upper bound for σ, and the $L^\infty(L^2)$ estimate of Theorem 4.6 and the second energy estimate of Theorem 4.8 depend on γ_2, the least upper bound of σ_t, as well.

Reasoning heuristically, we can understand the dependence of the estimates on an upper bound for σ as follows. If the constant C in (4.7) were to remain bounded as we let σ tend to infinity, then W would tend to a *continuous* optimal order approximation of w in the subspace $\mathcal{M}$. But if the subspaces are constructed over a general family of meshes, the best approximation in the continuous subspace, $\mathcal{M} \cap H^1(\Omega)$, may not be of optimal order. Of course if $\mathcal{T}$ is an edge-to-edge mesh, then $\mathcal{I}w$ is a continuous optimal order approximation to w. In this case, or more generally if the mesh permits continuous optimal order approximation, the restrictions on the penalty function may be eased considerably.

More precisely, suppose that there is a linear operator $\mathcal{J}: H^2(\mathcal{T}) \to \mathcal{M}$ such that

$$\|\phi - \mathcal{J}\phi\|_{i,T} \leqq C(r, K_0) h_T^{j-i} \|\phi\|_{j,T}, \qquad 0 \leqq i \leqq j \leqq r+1, \quad j \geqq 2, \quad T \in \mathcal{T},$$

and for which $\mathcal{J}\phi$ is continuous if ϕ is continuous. We then say that the subspace *approximates smoothly*. In this case it can be shown that the estimates holds with γ_1 replaced by

$$\gamma_3 = \sup\{\sigma(x, t) | x \in \partial\Omega, t \in I\}.$$

This leads to a very mild restriction on σ since our main interest is in adjusting the interior penalties, and it is not unreasonable to fix $\sigma|_{\partial\Omega \times I}$ at some sufficiently large constant value.

In the case of Theorem 4.8, which bounds $\|\zeta_t\|_{L^2(L^2)}$, some dependence of the estimate on σ_t is to be expected since a change in σ causes a change in W. However the dependence of the constant on $\gamma_2 = \sup|\sigma_t|$ can be considerably weakened to dependence on $\gamma_4 = \sup|\sigma^{-1}\sigma_t|$ in case the subspace approximates smoothly. Thus exponential growth of the penalty function is permitted. The same substitution of γ_4 for γ_2 can be made in Theorem 4.6, although it is not clear whether even the weakened restriction is necessary. For proofs of the above claims see [2].

Finally, let us note that σ need *not* be furnished as an *explicit* function of x and t. For example, in the favorable case of smooth approximation the basic energy estimate remains valid if $\sigma|_{\cup \mathscr{E}_0 \times I}$ is *any* function which is bounded below by γ_0. In particular, σ can depend on the approximate solution at an earlier time.

Note added in proof. Finite element computations incorporating interior penalties as described here have been used to solve equations of multiphase flow through porous media. Results of such computations have been reported in the following papers: *Self-adaptive finite element simulation of miscible displacement in porous media*, by J. Douglas, Jr., M. Wheeler, B. Darlow and R. Kendall, to appear in SIAM J. Sci. Stat. Comput.; *Finite elements with characteristics for two component incompressible miscible displacement*, by T. Russell, Soc. Pet. Eng. report SPE 10500, Dallas, 1982; *Mixed finite element methods for miscible displacement problems in porous media*, by B. Darlow, R. Ewing and M. Wheeler, Soc. Pet. Eng. report SPE 10501, Dallas, 1982. The former paper includes experimental results on an adaptive grid refinement scheme using penalties as discussed in § 5 above.

REFERENCES

[1] S. AGMON, *Lectures on Elliptic Boundary Value Problems*, Van Nostrand, Princeton, NJ, 1965.

[2] D. ARNOLD, *An interior penalty finite element method with discontinuous elements*, Ph.D. thesis, University of Chicago, Chicago, IL, 1979.

[3] I. BABUSKA, *The finite element method with penalty*, Math. Comp., 27 (1973), pp. 221–228.

[4] I. BABUSKA AND M. ZLÁMAL, *Nonconforming elements in the finite element method with penalty*, this Journal, 10 (1973), pp. 863–875.

[5] G. BAKER, *Finite element methods for elliptic equations using nonconforming elements*, Math. Comp., 31 (1977), pp. 45–59.

[6] J. DOUGLAS, JR., *The numerical simulation of miscible displacement in porous media*, Proc. Second International Conference on Computational Methods in Nonlinear Mechanics, Univ. of Texas, Austin, March 26–29, 1979.

[7] J. DOUGLAS, JR. AND T. DUPONT, *Interior Penalty Procedures for Elliptic and Parabolic Galerkin Methods*, Lecture Notes in Physics 58, Springer–Verlag, Berlin, 1976.

[8] J. KADLEC, *On the regularity of the solution of the Poisson problem on a domain with boundary locally similar to the boundary of a convex set*. Czech. Math. J., 14 (1964), pp. 386–393.

[9] J. A. NITSCHE, *Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind*, Abh. Math. Sem. Univ. Hamburg, 36 (1971), pp. 9–15.

[10] P. PERCELL AND M. WHEELER, *A local residual finite element procedure for elliptic equations*, this Journal, 15 (1978), pp. 705–714.

[11] M. WHEELER, *A priori L_2 error estimates for Galerkin approximations to parabolic partial differential equations*, this Journal, 10 (1973), pp. 723–759.

[12] M. WHEELER, *An elliptic collocation-finite element method with interior penalties*, this Journal, 15 (1978), pp. 152–161.

[13] O. ZIENKIEWICZ, *Constrained variational principles and penalty function methods in finite element analysis*, Conference on the Numerical Solution of Differential Equations, Lecture Notes in Mathematics 363, Springer, New York, 1973.