RAIRO. ANALYSE NUMÉRIQUE

JIM JR. DOUGLAS
TODD DUPONT
PETER PERCELL
RIDGWAY SCOTT

A family of C^1 finite elements with optimal approximation properties for various Galerkin methods for 2nd and 4th order problems

RAIRO. Analyse numérique, tome 13, nº 3 (1979), p. 227-255
<http://www.numdam.org/item?id=M2AN_1979__13_3_227_0>

R.A.I.R.O. Analyse numérique/Numerical Analysis
(vol. 13, n° 3, 1979, p. 227 à 255)

A FAMILY OF C^1 FINITE ELEMENTS WITH OPTIMAL APPROXIMATION PROPERTIES FOR VARIOUS GALERKIN METHODS FOR 2ND AND 4TH ORDER PROBLEMS (*) (1)

by Jim DOUGLAS, Jr., Todd DUPONT (2)
Peter PERCELL (3) and Ridgway SCOTT (4)

Abstract. — *Families of C^1 piecewise polynomial spaces of degree $r \geqq 3$ on triangles and quadrilaterals in two dimensions are constructed, and approximation properties of the families are studied. Examples of the use of the families in Galerkin methods for 2nd and 4th order elliptic boundary value problems on arbitrarily shaped domains are given. The approximation properties on the boundary are such that the rate of convergence of the Galerkin methods is the optimal rate determined by the degree r of the piecewise polynomial space.*

Résumé. — *En dimension deux, on construit des familles d'espaces de classe C^1, formés de polynômes de degré $r \geqq 3$ par morceaux, sur des triangles et des quadrilatères, et on étudie les propriétés d'approximation de ces familles. On en donne des exemples d'application à des méthodes de Galerkin pour les problèmes aux limites elliptiques du 2e et du 4e ordre posés sur des domaines de forme arbitraire. Les propriétés d'approximation de la frontière sont telles que le taux de convergence des méthodes de Galerkin est le taux optimal, déterminé par le degré r de l'espace des polynômes par morceaux.*

INTRODUCTION

In finite element approximation of fourth order elliptic boundary value problems and in some Galerkin methods for second order problems, C^1 piecewise polynomial spaces are required in order to satisfy conformity. In addition, if the boundary is curved, essential boundary conditions must be closely approximated in a negative Sobolev boundary norm [2]. These two requirements rule out the use of standard triangular and quadrilateral elements, and the purpose of this note is to present families of macroelements that do possess the necessary smoothness and boundary condition approximation

(*) Reçu juin 1978.
(1) Work supported by NSF Grant #MPS 74-12461 AØ1.
(2) Department of Mathematics, University of Chicago, Chicago, Illinois.
(3) Department of Mathematics, University of Houston.
(4) The submitted manuscript has been authored under contract EY-79-C-02-0016 with the U.S. Department of Energy. Applied Mathematics Department, Brookhaven National Laboratory, Upton, L.I., New York.

properties. An additional bonus of the families is that they have minimal smoothness: e.g., the standard 21 degree of freedom C^1 quintic [34] is C^2 at vertices, whereas the elements described here comprise the full space of C^1 piecewise polynomials, with no additional constraints.

A basic idea of the Galerkin methods treated here is to interpolate the boundary conditions rather than attempt to satisfy them exactly, without imposing a penalty on the nonsatisfaction of boundary conditions. This idea was first studied in [2] and then amplified in [4] and [30]; in these papers, only second order Dirichlet problems using C^0 piecewise polynomial spaces on triangles were considered. Here, we treat fourth order problems, as well as general second order problems using Galerkin techniques [13], [15], [16] that require C^1 spaces. To achieve C^1 continuity while retaining enough degrees of freedom at the boundary to have the necessary approximation properties, macroelements based on the Clough-Tocher [12] and Fraeijs de Veubeke-Sander [18], [27] macrocubics are used.

The families of macroelements studied here were first considered in [32], and further developed in [21], [25], although accurate approximation of essential boundary conditions was not treated. In these works, the convergence parameter is the degree r of the piecewise polynomials. The point of view in the present paper is that r is fixed and the convergence parameter is the mesh size h.

1. FAMILIES OF MACROELEMENTS

In this section we present a family of C^1 triangular macroelements (also *see* [32]) for which we shall prove useful approximation properties. The family contains an element of degree n for each $n \geqq 3$ beginning with the well known cubic Clough-Tocher element [12], [8]. We shall discuss several modifications of this basic family and also briefly discuss a very similar family of quadrilateral macroelements which begins with the cubic Fraeijs de Veubeke-Sander macroelement [18], [27], [11].

First we give some general definitions. Following Ciarlet [9], we define a *finite element* to be a triple (K, F, Σ) such that

(*a*) $K \subset \mathbf{R}^2$ is a compact region having the restricted cone property,

(*b*) F is a finite dimensional vector space of real-valued functions on K, and

(*c*) Σ is a finite set of linear functionals φ_i, $1 \leqq i \leqq N$, called the *degrees of freedom* of the finite element, which are defined on a vector space of functions containing F and which have the property that for any real numbers α_i, $1 \leqq i \leqq N$, there exists a unique function $f \in F$ which satisfies

$$\varphi_i(f) = \alpha_i, \qquad 1 \leqq i \leqq N.$$

A *nodal finite element* is a finite element (K, F, Σ) for which each degree of freedom is a functional which picks out the value or the value of some derivative at a point, called a *node*, in K. For $n \geqq 0$ and $U \subset \mathbf{R}^2$, let $P_n(U)$, denote the space of restrictions to U of polynomials (in the coordinates of $\mathbf{R}^2$) of total degree not greater than n. The *degree* of a finite element (K, F, Σ) is the largest integer n such that $P_n(K) \subset F$ or $-\infty$ if F does not contain $P_n(K)$ for any $n \geqq 0$.

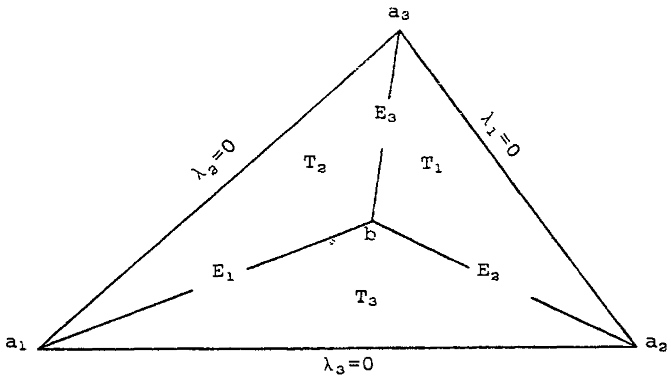

Figure 1. — A macrotriangle T.

Now we define the family of triangular macroelements. For $n \geqq 3$, an element of degree n in the family of triangular macroelements is a nodal finite element $(T, S_n(T), \Sigma_n)$ defined as follows:

(*a*) $T \subset \mathbf{R}^2$ is a macrotriangle; i.e., a triangle T triangulated by three subtriangles T_1, T_2 and T_3 (see *fig.* 1). Note that the point b is allowed to be anywhere in the interior of T;

(*b*) $S_n(T) = \{f \in C^1(T) : f \mid T_i \in P_n(T_i),\ 1 \leqq i \leqq 3\}$, where the vertical bar denotes restriction of a function;

(*c*) The degrees of freedom Σ_n are

1. the value and gradient (i.e. $\partial/\partial x$ and $\partial/\partial y$) at the exterior vertices,
2. the value at $n-3$ distinct points in the interior of each exterior edge of T,
3. the normal derivative at $n-2$ distinct points in the interior of each exterior edge of T,

and if $n \geqq 4$,

4. the value and gradient at the interior vertex,
5. the value at $n-4$ distinct points in the interior of each interior edge of T,
6. the normal derivative at $n-4$ distinct points in the interior of each interior edge of T, and

7. the value at (1/2) $(n-4)$ $(n-5)$ distinct points in the interior of each T_i chosen so that if a polynomial of degree $n-6$ vanishes at those points, then it vanishes identically.

For an example of one of these macroelements, see *fig.* 2 in which a dot denotes a value, a circle denotes a gradient and a dash denotes a normal derivative.

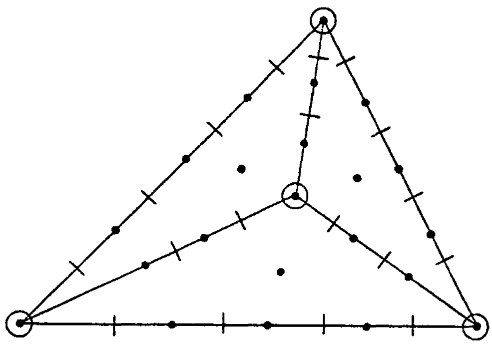

Figure 2. — The degrees of freedom Σ_6

For use in the proof of the following theorem, let λ_i, $1 \leqq i \leqq 3$, be the (unique) affine function on $\mathbf{R}^2$ such that $\lambda_i(b)=1$ and λ_i vanishes at the exterior vertices of T_i. Furthermore, let λ be the function on T defined by

$$\lambda \mid T_i = \lambda_i \mid T_i, \qquad 1 \leqq i \leqq 3.$$

Note that λ is well defined and continuous on T because λ_i and λ_{i+1} agree on the edge $E_{i+2}=T_i \cap T_{i+1}$ since both are zero at a_{i+2} and one at b. (Here and through the following proof, subscripts referring to triangles are counted modulo 3).

THEOREM 1: *$(T, S_n(T), \Sigma_n)$, $n \geqq 3$, is a well defined nodal finite element of degree n.*

Proof: It is easy to see by simply counting that, for all $n \geqq 3$, the number of degrees of freedom, denoted by $\# \Sigma_n$, is $(3/2)(n^2-n+2)$. On the other hand, the following argument due to Strang (*see* [29], [31] for example) shows that

$$\dim S_n(T) \geqq \frac{3}{2}(n^2-n+2).$$

Let $R_n(T)$, $n \geqq 3$, denote the space of all (discontinuous) piecewise polynomials of degree n on the macrotriangle T. Note that $R_n(T)$ is isomorphic to $\prod_{j=1}^{3} P_n(T_j)$, so dim $R_n(T)=(3/2)(n+1)(n+2)$. Consider $S_n(T)$ to be a subspace of $R_n(T)$ consisting of functions satisfying certain constraints imposed across the interior edges of T. In general, it takes $2n+1$ constraints ($n+1$ on values and n on normal derivatives) to make two polynomials of degree n agree in a C^1 fashion along an edge. However, if a function in $R_n(T)$ satisfies the $2n+1$ constraints along two of the interior edges of T, then it automatically has a uniquely defined value and gradient at the interior vertex of T. Thus it takes the satisfaction of no more than $3(2n+1)-3=6n$ constraints to ensure that a function in $R_n(T)$ is actually in $S_n(T)$. Therefore,

$$\dim S_n(T) \geqq \dim R_n(T)-6n=\frac{3}{2}(n^2-n+2).$$

To show that $(T, S_n(T), \Sigma_n)$ is a well defined finite element we must show that Σ_n is a basis for the dual space of $S_n(T)$, which we denote $(S_n(T))'$. But we have just shown that

$$\#\Sigma_n \leqq \dim S_n(T)=\dim (S_n(T))',$$

so it is enough to show that Σ_n spans $(S_n(T))'$. This is equivalent to showing that if all the degrees of freedom for $f \in S_n(T)$ are zero, then f vanishes identically.

When the degrees of freedom are all zero, in particular those of types 1, 2 and 3, f and ∇f vanish on the boundary of T. Thus $f=p\lambda^2$, where $p_i=p \mid T_i$ is a polynomial of degree $n-2$ and p is continuous (because f and λ are continuous and because λ does not vanish in the interior of T). Since $f \in C^1(T)$, ∇f is well defined on E_{i-1} and hence can be expressed there by both

$$2p\lambda\nabla\lambda_i+\lambda^2\nabla p_i \qquad \text{and} \qquad 2p\lambda\nabla\lambda_{i+1}+\lambda^2\nabla p_{i+1}.$$

Therefore

$$2p\nabla(\lambda_{i+1}-\lambda_i)+\lambda\nabla(p_{i+1}-p_i)=0 \quad \text{on } E_{i-1}. \tag{1}$$

Since $\lambda(a_{i-1})=0$ and $\nabla(\lambda_{i+1}-\lambda_i)\neq 0$ (because the lines $\lambda_i=0$ and $\lambda_{i+1}=0$ cannot be parallel) it follows that $p(a_{i-1})=0$. Now, suppose for a moment that $n=3$. (*See* also [8] and [26] for this case.) Then each p_i is a linear polynomial which vanishes at the exterior vertices a_{i-1} and a_{i+1} of the triangle T_i, so p_i vanishes on the entire exterior edge of T_i. Hence $p_i=c_i\lambda_i$ for some $c_i \in \mathbf{R}$, so $f_i=c_i\lambda_i^3$. Thus $\nabla f_i(b)=3c_i\nabla\lambda_i(b)$. Since the vectors $\nabla\lambda_i(b)$, $1 \leqq i \leqq 3$, are linearly independent and f is differentiable at b [i.e., $\nabla f_i(b)=\nabla f_{i+1}(b)$], it follows that all the c_i's are zero; i.e., f vanishes identically. This finishes the case $n=3$. From now on we suppose that $n \geqq 4$. Then $p(b)=0$, since $f(b)=0$ and $\lambda(b)=1$, and $\nabla p_i(b)=0$ since

$$0=\nabla f(b)=(2p\lambda\nabla\lambda_i+\lambda^2\nabla p_i)(b)=\nabla p_i(b).$$

Thus $p \mid E_{i-1}$ is a polynomial of degree $n-2$ which vanishes along with its derivative at b and is zero at a_{i-1} and the $n-4$ nodes of type 5 on E_{i-1}; consequently, $p \mid E_{i-1}$ is identically zero. Therefore,

$$p_i = q_i (\lambda_{i+1} - \lambda_i)(\lambda_i - \lambda_{i-1})$$

and

$$f_i = q_i \lambda_i^2 (\lambda_{i+1} - \lambda_i)(\lambda_i - \lambda_{i-1}),$$

where q_i is a polynomial of degree $n-4$, since $p_i = 0$ on the two interior edges of T_i and since $\lambda_{i+1} - \lambda_i$ vanishes on one of these edges while $\lambda_i - \lambda_{i-1}$ vanishes on the other. Also, using (1) and the fact that $p=0$ on E_{i-1}, we see that $\nabla p_i = \nabla p_{i+1}$ on E_{i-1}. But on E_{i-1},

$$\nabla p_i = q_i (\lambda_i - \lambda_{i-1}) \nabla (\lambda_{i+1} - \lambda_i)$$

and

$$\nabla p_{i+1} = q_{i+1} (\lambda_{i-1} - \lambda_{i+1}) \nabla (\lambda_{i+1} - \lambda_i);$$

so since $\lambda_i = \lambda_{i+1}$ on E_{i-1}, we find that $q_i = -q_{i+1}$ on E_{i-1}. In particular, $q_i(b) = -q_{i+1}(b)$, and

$$q_i(b) = -q_{i+1}(b) = q_{i+2}(b) = -q_i(b),$$

which implies that $q_i(b) = 0$. At this point we are finished if $n=4$ because then q_i is a constant polynomial (*cf.* [26]). If $n \geqq 5$, then $\nabla f_i = 0$ at the $n-4$ nodes of type 6 on $E_{i\pm 1}$ because $f = p\lambda^2 = 0$ on $E_{i\pm 1}$. Since

$$\nabla f_i = q_i \lambda_i^2 (\lambda_i - \lambda_{i\pm 1}) \nabla (\lambda_{i\mp 1} - \lambda_i) \quad \text{on } E_{i\pm 1},$$

it follows that q_i vanishes at the same $n-4$ nodes on E_{i+1} and E_{i-1}. As it also vanishes at b and is a polynomial of degree $n-4$, it vanishes identically on E_{i+1} and E_{i-1}. If $n=5$, this means that $q_i \equiv 0$, since it is linear. For $n \geqq 6$, this means that

$$q_i = r_i (\lambda_{i+1} - \lambda_i)(\lambda_i - \lambda_{i-1})$$

and

$$f_i = r_i \lambda_i^2 (\lambda_{i+1} - \lambda_i)^2 (\lambda_i - \lambda_{i-1})^2,$$

where r_i is a polynomial of degree $n-6$. Finally, since f_i is zero at the nodes of type 7, it follows that r_i is also zero at these nodes and hence r_i must vanish identically because it has degree $n-6$. //

REMARK 1: Let α_i: $[0, L_i] \to T_i$, $1 \leqq i \leqq 3$, be a parametrization by arclength of the exterior edge of T_i and let $\partial / \partial n_i$ denote differentiation normal to this edge. If for some i, $1 \leqq i \leqq 3$, the degrees of freedom along the exterior edge of T_i of types 2 and 3 in the set Σ_n are replaced by

$$\tilde{2}. \qquad f \to \int_0^{L_i} p_j(s) f(\alpha_i(s))\, ds, \qquad 0 \leqq j \leqq n-4,$$

and

$$\tilde{3}. \qquad f \to \int_0^{L_i} q_j(s) \frac{\partial f}{\partial n_i}(\alpha_i(s))\, ds, \qquad 0 \leqq j \leqq n-3,$$

where $\{p_j\}$ (resp. $\{q_j\}$) is a basis for the space of polynomials in one variable of degree $n-4$ (resp. $n-3$), then the result is a new (non-nodal) finite element of degree n, which we denote $(T, S_n(T), \tilde{\Sigma}_n)$. This follows from the proof of theorem 1 because the number of degrees of freedom is unchanged and because the degrees of freedom of types 1, $\tilde{2}$ and $\tilde{3}$ again uniquely determine f and ∇f on an exterior edge of T when $f \in S_n(T)$. Finite elements using degrees of freedom of type $\tilde{2}$ were studied by Blair [4] in the context of second order problems. If the sets $\{p_j\}$ and $\{q_j\}$ are chosen to be orthogonal polynomials, a hierarchical structure may be achieved [25]. //

REMARK 2: In both the original element $(T, S_n(T), \Sigma_n)$ and the modified element $(T, S_n(T), \tilde{\Sigma}_n)$, the exterior edges of T need not be straight but can be smooth curves which are just C^1 close to being straight. Using the notation introduced above, we say that the exterior edge of T_i is C^1 close to being straight if

$$\sup_{0 \leqq s \leqq L_i} \left\{ |\lambda_i(\alpha_i(s))|, \left| \frac{d}{ds}[\lambda_i(\alpha_i(s))] \right| \right\}$$

is small. It follows immediately from this definition that a C^1 small perturbation of a straight edge moves points on the edge and normals to the edge only slightly. Thus if T has exterior edges which are sufficiently C^1 close to being straight, then the degrees of freedom Σ_n or $\tilde{\Sigma}_n$ are close enough to those for the corresponding element for the macrotriangle having straight edges and the same vertices so that the degrees of freedom still determine uniquely a function $f \in S_n(T)$. (The distance between two degrees of freedom is measured in the dual space of C_b^1, the Banach space of bounded C^1 functions with bounded first derivatives.) //

REMARK 3: The elements in theorem 1 have the virtue that they may be pieced together to form C^1 functions. For this reason, they may be referred to as "C^1 finite elements". To see why this is so, let T^1 and T^2 be two macro-triangles that share (only) a common edge E (the vertices of E are required to be vertices of both T^1 and T^2). Let Σ_n^i be degrees of freedom defined on $S_n(T^i)$, $i = 1, 2$, that are *consistent*, i.e., the nodal points of type 2 and 3 on E for Σ_n^1 and Σ_n^2 pairwise coincide. Let $f_i \in S_n(T^i)$, $i=1, 2$, be such that the degrees of freedom of types $1-3$ of f_1 and f_2 pairwise coincide, and let f be defined on $T^1 \cup T^2$ by $f \mid T^i = f_i$, $i=1, 2$. Then $f \in C^1(T^1 \cup T^2)$. (*Proof:* Let $\tau_i \subset T^i$, $i=1, 2$, be the subtriangle having E as an edge. The polynomial $P = f_1 \mid \tau_1 - f_2 \mid \tau_2$ vanishes to second order at the vertices of E and at $n-3$ other points on E, and since the degree of P is n, $P \mid E \equiv 0$. Similar reasoning shows that the normal derivative of P vanishes on E. Thus P vanishes to second order on E, and this means that f is C^1.)

The elements in remark 1 can also be pieced together to give C^1 functions, requiring only a matching of the orientations of E. Elements with curved edges as in remark 2 will arise only when the curved edge lies on the boundary of a domain; thus the problem of piecing together elements across a curved edge is avoided. Elements of all three types can be attached to each other across straight edges by matching the types of degrees of freedom on the shared edges. In fact, while the degrees of freedom of types 2 and 3, or $\tilde{2}$ and $\tilde{3}$, will be required in section 2 on the boundary of the domain, other finite elements may be used in the interior, with the appropriate matching. For $n \geq 5$, there is a well known (*cf.* [34] for the case $n=5$) C^1 finite element of degree n, which we shall denote by $(\tau, P_n(\tau), \overline{\Sigma}_n)$, such that τ is an ordinary triangle and the degrees of freedom $\overline{\Sigma}_n$ are

1. the value plus all first and second derivatives at each vertex,
2. the value at $n-5$ distinct points on each edge,
3. the normal derivative at $n-4$ distinct points on each edge, and
4. the value at $(1/2)(n-4)(n-5)$ distinct points in the interior of τ chosen so that if a polynomial of degree $n-6$ vanishes at those points, then it vanishes identically.

The transition from this element to the one in theorem 1 is *via* two unsymmetric macroelements, denoted $(T, S_n'(T), \Sigma_n')$ and $(T, S_n''(T), \Sigma_n'')$, for which $S_n'(T)$ and $S_n''(T)$ are proper subspaces of $S_n(T)$ consisting of functions which have second derivatives at certain exterior vertices of T, and whose degrees of freedom when $n=6$ are presented in figure 3. In this figure a second circle around a node means that all the second derivatives at that point are degrees of freedom. We shall not write out the degrees of freedom Σ_n' and Σ_n'' in detail because the pattern should be clear. Proofs that these elements are well defined can be given along the same lines as in the proof of theorem 1. Note that remarks 1 and 2 hold for these elements. //

We now quickly present the family of C^1 quadrilateral macroelements, discussing just the differences between this family and the family of triangular macroelements. For $n \geqq 3$, an element of degree n in the family of quadrilateral macroelements is a nodal finite element $(Q, S_n(Q), \Lambda_n)$ such that

(*a*) $Q \subset \mathbf{R}^2$ is a convex quadrilateral triangulated by the four triangles Q_i, $1 \leqq i \leqq 4$, obtained by drawing in the diagonals of Q,

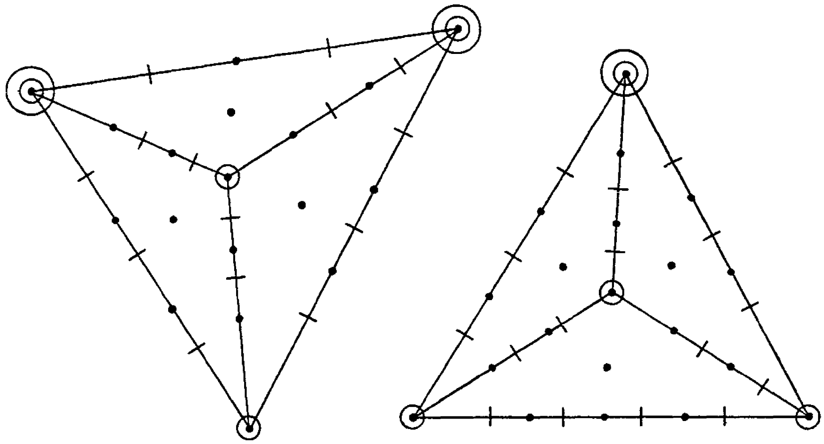

Figure 3. — Triangular transition elements of degree 6.

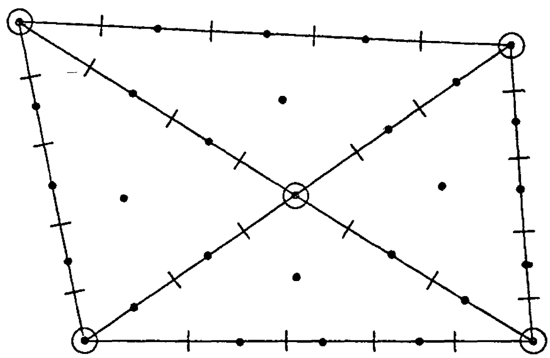

Figure 4. — The degrees of freedom Λ_6

(*b*) $S_n(Q) = \{ f \in C^1(Q) : f \,|\, Q_i \in P_n(Q_i),\ 1 \leqq i \leqq 4 \}$, and

(*c*) after the obvious change of T' s to Q' s, the description of Λ_n is the same as the description of Σ_n, except that (when $n \geqq 4$) an extra normal derivative at a point along just one of the interior edges of Q is added to the degrees of freedom of type 6.

The proof that $(Q, S_n(Q), \Lambda_n)$ is a well defined nodal finite element of degree n is essentially the same as the proof of theorem 1 with only two changes of any significance needed. The first change is necessary because the interior vertex in the quadrilateral macroelement is *singular* in the sense of [29]. Counting shows that $\# \Lambda_n = 2(n^2 - n + 2)$, while the methods used in the proof of theorem 1 only show that

$$\dim S_n(Q) \geqq 2(n^2 - n + 2) - 1.$$

However, in order to be able to proceed as before we need to know that $\# \Lambda_n \leqq \dim S_n(Q)$. In [29], it is shown that this is true since one of the constraints on normal derivatives one expects to be necessary is actually redundant in the

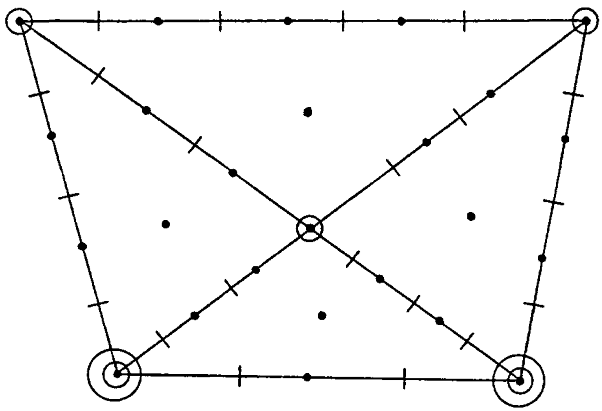

Figure 5. — Quadrilateral transition element of degree 6.

presence of the singular vertex. The only other change needed in the proof comes when one is trying to show that the q_i's are zero at the interior vertex. The alternating sign approach used before breaks down for the quadrilateral macroelements because an even number of edges abut the interior vertex. However, it is easy to see that the proof is rescued by the extra degree of freedom of type 6 in Λ_n available when $n \geqq 4$.

Remarks 1 and 2 clearly apply to this family of quadrilateral macroelements, and the ideas discussed in remark 3 may be extended to the present situation as follows. First note that triangular and quadrilateral macroelements of the same degree are compatible and can be used together in the same domain Ω because the degrees of freedom associated with exterior edges are the same for both types of element. One may thus use the standard C^1 triangular element $(\tau, P_n(\tau), \overline{\Sigma}_n)$, $n \geqq 5$, as before, in the interior of Ω and a layer of quadrilateral macroelements one element thick along $\partial\Omega$. Only one type of boundary layer *quadrilateral* macroelement is required for the transition, and this is depicted in figure 5 for $n=6$. Transition elements of other degrees may be easily constructed using the ideas discussed earlier. Note that the transition element in figure 5 may be used as a boundary element directly without having an additional layer of elements of the type depicted in figure 4, if desired.

2. APPROXIMATION PROPERTIES

In this section we prove approximation properties for the finite elements introduced in section 1. We start by fixing some notation to be used throughout the remainder of this section.

For any domain $\Omega \subset \mathbf{R}^2$ having the restricted cone property and for any smooth arc $\Gamma \subset \mathbf{R}^2$ parametrized by arclength, denoted by s, let

$$(v, w)_\Omega = \int_\Omega v(x)\, w(x)\, dx$$

be the (real) inner product in $L^2(\Omega)$ and let

$$\langle v, w \rangle_\Gamma = \int_\Gamma v(s)\, w(s)\, ds$$

be the inner product in $L^2(\Gamma)$. Furthermore, if m is a non-negative integer, let $[\cdot]_{m,\Omega}$ denote the seminorm defined by

$$[v]^2_{m,\Omega} = \sum_{|\alpha|=m} (D^\alpha v, D^\alpha v)_\Omega \,;$$

let $\|\cdot\|_{m,\Omega}$ be the norm for the Sobolev space $H^m(\Omega)$ defined by

$$\| v \|^2_{m,\Omega} = \sum_{j=0}^{m} [v]^2_{j,\Omega} \,.$$

When m is a nonintegral positive real number, let $\bar{m}$ denote the integral part of m and define the semi-norm $[\cdot]_{m,\Omega}$ by

$$[v]^2_{m,\Omega} = \sum_{|\alpha|=\bar{m}} \iint_{\Omega\times\Omega} \frac{|D^\alpha v(x) - D^\alpha v(y)|^2}{|x-y|^{2+2(m-\bar{m})}}\, dx\, dy;$$

define the norm $\| \quad \|_{m,\Omega}$ for $H^m(\Omega)$ by

$$\| v \|^2_{m,\Omega} = \| v \|^2_{\bar{m},\Omega} + [v]^2_{m,\Omega} \,.$$

Let $\|\cdot\|_{-m,\Omega}$ denote the dual norm for the space $H^{-m}(\Omega)=(H^m(\Omega))'$ defined by

$$\|v\|_{-m,\Omega}=\sup\left\{\frac{|(v,\varphi)_\Omega|}{\|\varphi\|_{m,\Omega}}:\varphi\in H^m(\Omega),\ \varphi\neq 0\right\}.$$

Similarly, for m a non-negative integer, let $|.|_{m,\Gamma}$ denote the norm for $H^m(\Gamma)$ defined by

$$|v|_{m,\Gamma}^2=\sum_{j=0}^{m}\left\langle\frac{\partial^j v}{\partial s^j},\frac{\partial^j v}{\partial s^j}\right\rangle_\Gamma.$$

For nonintegral positive real m, define the norm $|.|_{m,\Gamma}$ for $H^m(\Gamma)$ by

$$|v|_{m,\Gamma}^2=|v|_{\bar{m},\Gamma}^2+\iint_{\Gamma\times\Gamma}\frac{|v^{(m)}(s)-v^{(m)}(t)|^2}{d(s,t)^{1+2(m-\bar{m})}}\,ds\,dt$$

where $v^{(m)}(s)=\partial^{\bar{m}}v/\partial s^{\bar{m}}(x(s))$ with $x(s)$ parametrizing Γ and $d(s,t)$ is the minimum arc length on Γ between $x(s)$ and $x(t)$. Let

$$|v|_{-m,\Gamma}=\sup\left\{\frac{|\langle v,\varphi\rangle_\Gamma|}{|\varphi|_{m,\Gamma}}:\varphi\in H^m(\Gamma),\ \varphi\neq 0\right\}$$

be the norm for $H^{-m}(\Gamma)=(H^m(\Gamma))'$. Finally, for m a non-negative integer, let $\{.\}_{m,\Omega}$ denote the $C^m(\Omega)$ norm defined by

$$\{v\}_{m,\Omega}=\sup\{|D^\alpha v(x)|:x\in\Omega,\ 0\leqq|\alpha|\leqq m\}.$$

Let Ω be a fixed, bounded domain in $\mathbf{R}^2$ with C^∞ boundary $\partial\Omega$ parametrized by arclength. A *triangulation* of Ω will mean a collection $\mathscr{T}=\{T_1,\ldots,T_N\}$ of polygons each of which is a triangle, macrotriangle, or convex quadrilateral such that $\bar{\Omega}=\bigcup_{j=1}^{N}T_j$ and such that distinct polygons intersect at most in a common vertex or a common edge. By a *simple triangle* in such a triangulation, we mean either a triangle in $\mathscr{T}$ (but not a macrotriangle), a subtriangle of a macrotriangle, or a traingle obtained by subdividing a convex quadrilateral by drawing its diagonals. Simple triangles are assumed to have straight edges except that an edge between two vertices on $\partial\Omega$ is assumed to be contained in $\partial\Omega$.

Now suppose that Ω is provided with a family of triangulations $\mathscr{T}_h$, $0<h\leqq h_0\leqq 1$, as described above, such that

(*a*) if $T\in\mathscr{T}_h$, then diam $(T)\leqq h$, and

(*b*) the family is *non-degenerate* in the sense that there exists a number $\rho>0$ such that if T is a polygon in the family and $\tau\subset T$ is a simple triangle, then τ is starlike with respect to a disk whose diameter is ρ diam (T).

Consider for a moment an arbitrary finite element (K, F, Σ) with degrees of freedom $\varphi_1, \ldots, \varphi_N$ that are continuous on $H^m(K)$. We associate with (K, F, Σ) an *interpolation operator*

$$\Pi : \quad H^m(K) \to F$$

by letting Πv be the unique function in F such that

$$\varphi_i(\Pi v) = \varphi_i(v), \qquad i=1, \ldots, N.$$

Let $\delta = \operatorname{diam}(K)$ and let $\hat{x} = x/\delta$ and $\tilde{K} = \{ \hat{x} : x \in K \}$.

For any function f defined on K, set $\hat{f}(\hat{x}) = f(x)$, and define an operator $\hat{\Pi}$ on $H^m(\hat{K})$ by

$$\hat{\Pi} \hat{v} = \Pi v, \qquad v \in \mathrm{H}^m(K)$$

Note that $\hat{\Pi}$ is the interpolation operator associated with the finite element $(\hat{K}, \hat{F}, \hat{\Sigma})$, and that, for any of the finite elements discussed in section 1, including that of remark 1, $\hat{\Sigma}$ has the same description as Σ (because dilations – as opposed to arbitrary linear maps – preserve orthogonality).

Now, suppose h_0 is small enough (see remark 2 of section 1) so that each polygon $T_{h,j} \in \mathscr{I}_h$, $1 \leqq j \leqq N_h$, can be given the structure $(T_{h,j}, F_{h,j}, \Sigma_{h,j})$ of one of the finite elements of section 1 of some fixed degree $r \geqq 3$ for $0 < h \leqq h_0$. Let $\Pi_{r,h,j}$ be the interpolation operator associated with $(T_{h,j}, F_{h,j}, \Sigma_{h,j})$ and let m_0 be a real number such that each $\varphi \in \Sigma_{h,j}$ is continuous on $H^{m_0}(T_{h,j})$ for all $j=1, \ldots, N_h$ and $0 < h \leqq h_0$. We assume that the following *uniformity condition* holds:

$$(c) \qquad \| \hat{\Pi}_{r,h,j} \hat{v} \|_{2, \hat{T}_{h,j}} \leqq C \| \hat{v} \|_{m_0, \hat{T}_{h,j}}$$

for $1 \leqq j \leqq N_h$, $0 < h \leqq h_0$, and $\hat{v} \in H^{m_0}(\hat{T}_{h,j})$. (Here and throughout the rest of this section, unless otherwise stated, C denotes a generic constant which may be different in different places and which depends only on Ω, r, h_0, and the families of triangulations and interpolation operators.) The existence of such a constant C follows from Sobolev's inequality (*cf.* Grisvard [19]), with m_0 any real number greater than one plus the highest order of derivative occurring as a degree of freedom (hence note that $m_0 > 2$ always). The condition can be satisfied with C independent of h because the family of triangulations is non-degenerate: there is room in the normalized (hatted) polygons for degrees of freedom that are bounded away from degenerate configurations. Examples of degeneracy are coalescence of two value nodes or colinearity of three value nodes in the interior of a simple triangle associated with an element of degree 7. Now define a global interpolation operator $\Pi_{r,h}$ on $H^{m_0}(\Omega)$ by

$$(\Pi_{r,h} v) \mid T_{h,j} = \Pi_{r,h,j}(v \mid T_{h,j}), \qquad j=1, \ldots, N_h.$$

Let $S_{r,h}$ be the image of $H^{m_0}(\Omega)$ *via* the mapping $\Pi_{r,h}$. We assume that the family of elements is *consistent* (see remark 3 of section 1), namely, we assume that

(d) $S_{r,h} \subset C^1(\bar{\Omega})$.

Since each function in $S_{r,h}$ is a polynomial in each simple triangle in $\mathcal{T}_h$, condition (d) implies that $S_{r,h} \subset H^2(\Omega)$.

REMARK 4: When each element in $\mathcal{T}_h$ is one of the macroelements of degree r of section 1 (except for the transition macroelements described in figures 3 and 5), then $S_{r,h}$ is the space of all C^1 piecewise polynomials of degree r on the triangulation $\bar{\mathcal{T}}_h$ of Ω obtained by considering all simple triangles in $\mathcal{T}_h$ separately. As shown in [23], when $r \geqq 5$, this space has a nodal basis, but of a more complicated variety than considered in section 1: the macroelements provide a simpler nodal parametrization of $S_{r,h}$. In the case $r \geqq 5$, the theorems to follow may be generalized to hold for $S_{r,h}(\bar{\mathcal{T}}_h)$ for an arbitrary triangulation $\bar{\mathcal{T}}_h$ of Ω, i. e., one not coming from a macro-triangulation, by changing the nodal parameters in [23] at the boundary appropriately. The essential change required is algebraic, as the analysis of this section involves only the simple triangles. //

For the remainder of this section, we make the following

ASSUMPTION: For $0 < h \leqq h_0$, the family of triangulations $\mathcal{T}_h$ of Ω satisfies conditions (a) and (b) and the associated family of interpolation operators $\Pi_{r,h}$, $r \geqq 3$, satisfies conditions (c) and (d).

THEOREM 2: *Let q and m be real numbers such that $0 \leqq q \leqq 2$ and $m_0 \leqq m \leqq r+1$, and let $v \in H^m(\Omega)$. Then*

$$\| v - \Pi_{r,h} v \|_{q,\Omega} \leqq C h^{m-q} [v]_{m,\Omega}.$$

Proof: Let $T_{h,j} \in \mathcal{T}_h$. From the fact that $\hat{\Pi}_{r,h,j}$ preserves polynomials of degree r and the uniformity condition (c), it follows that, for any $\varphi \in P_r(\hat{T}_{h,j})$,

$$\| \hat{v} - \hat{\Pi}_{r,h,j} \hat{v} \|_{q,\hat{T}_{h,j}} = \| (\hat{v} - \varphi) - \hat{\Pi}_{r,h,j} (\hat{v} - \varphi) \|_{q,\hat{T}_{h,j}} \leqq C \| \hat{v} - \varphi \|_{m,\hat{T}_{h,j}}.$$

The following version (proved in [17] and [20]) of the standard Bramble-Hilbert lemma (see [5] and [10]) is now required:

LEMMA 1: *Given $0 < \rho < 1$ and a positive real number m, there exists a constant $C = C(\rho, m)$ such that if $K \subset \mathbf{R}^2$ is a domain having diameter at most one that is* *starlike with respect to a disc of diameter* ρ *and if* $v \in H^m(K)$, *then*

$$\inf_{\varphi \in P_{\overline{m}}(K)} \| v - \varphi \|_{m,K} \leqq C [v]_{m,K}$$

where $\overline{m}$ *is the greatest integer less than m.* //

Applying the lemma,

$$\| \hat{v} - \hat{\Pi}_{r,h,j} \hat{v} \|_{q,T_{h,j}} \leqq C [\hat{v}]_{m,T_{h,j}}.$$

Since diam $(T_{h,j}) \leqq h$, the homogeneity of the norms implies that

$$\| v - \Pi_{r,h} v \|_{q,\hat{T}_{h,j}} \leqq C h^{m-q} [v]_{m,\hat{T}_{h,j}}. \quad //$$

From now on, we shall make frequent use of the following easily proved inequalities: if $K \subset \mathbf{R}^2$ is a region such that diam $(K) \leqq \delta \leqq 1$ and K contains a disc of diameter $\rho\delta$, $\rho > 0$, then for $\varphi \in P_r(K)$ and any m,

$$\{ \varphi \}_{m,K} \leqq C \delta^{-1} \| \varphi \|_{m,K} \quad (*) \tag{2}$$

and

$$\| \varphi \|_{m,K} \leqq C \delta^{j-m} \| \varphi \|_{j,K}, \qquad 0 \leqq j \leqq m, \tag{3}$$

where C depends only on ρ and r. The following lemma will be used in the proofs of the remaining theorems of this section (for a proof, see the appendix of [30], or use the trace theorem [22] with a little care).

LEMMA 2: *Let* $B(\mathscr{I}_h)$ *be the collection of all simple triangles in* $\mathscr{I}_h$ *having an edge on* $\partial\Omega$ *and let* $v \in H^m(\Omega)$ *with* $m_0 \leqq m \leqq r+1$. *Then*

$$\Big(\sum_{\tau \in B(\mathscr{I}_h)} \| \Pi_{r,h} v \|_{m-1,\tau}^2 \Big)^{1/2} \leqq C h^{1/2} \| v \|_{m,\Omega}. \quad //$$

THEOREM 3: *Let* $v \in H^m(\Omega)$ *for some m satisfying* $m_0 \leqq m \leqq r+1$. *Then if* $v \in H^m(\partial\Omega)$,

(*a*)
$$| v - \Pi_{r,h} v |_{q,\partial\Omega} \leqq C h^{m-q} (| v |_{m,\partial\Omega} + \| v \|_{m,\Omega})$$

when $0 \leqq q \leqq 2$. *If* $\partial v / \partial n \in H^{m-1}(\partial\Omega)$, *then*

(*b*)
$$\left| \frac{\partial v}{\partial n} - \frac{\partial}{\partial n} (\Pi_{r,h} v) \right|_{q,\partial\Omega} \leqq C h^{m-q-1} \left(\left| \frac{\partial v}{\partial n} \right|_{m-1,\partial\Omega} + \| v \|_{m,\Omega} \right)$$

when $0 \leqq q \leqq 1$, *where* $\partial/\partial n$ *denotes differentiation normal to* $\partial\Omega$.

(*) Here m is assumed to be a non-negative integer for $\{\ \}_{m,k}$ to be defined.

Proof: The proofs of part (*a*) and (*b*) are similar, so we just prove (*a*). Let $T \in \mathscr{I}_h$ be an element with an edge e on $\partial\Omega$, let $\delta = \operatorname{diam} T$ and let $\tau \subset T$ be the simple triangle having e as an edge. Note that since $\partial\Omega$ is smooth, the length of e, $l(e)$, may be bounded by $C\delta$. Let φ be the polynomial in arclength on e of degree r defined using the degrees of freedom associated with values, tangential derivatives, or orthogonalities on e in the definition of $\Pi_{r,h}$ which interpolates both $v|e$ and $\Pi_{r,h}v|e$. Then

$$|v-\Pi_{r,h}v|_{q,e} \leqq |v-\varphi|_{q,e} + |\varphi-\Pi_{r,h}v|_{q,e}$$
$$\leqq C\left(\delta^{m-q}|v|_{m,e} + \delta^{r-q+1}\left|\frac{\partial^{r+1}}{\partial s^{r+1}}(\Pi_{r,h}v)\right|_{0,e}\right)$$

for $0 \leqq q \leqq m$. Furthermore,

$$\left|\frac{\partial^{r+1}}{\partial s^{r+1}}(\Pi_{r,h}v)\right|_{0,e} \leqq C\delta^{1/2}\sup_e\left|\frac{\partial^{r+1}}{\partial s^{r+1}}(\Pi_{r,h}v)\right|$$
$$\leqq C\delta^{1/2}\{\Pi_{r,h}v\}_{r+1,\tau} = C\delta^{1/2}\{\Pi_{r,h}v\}_{r,\tau}$$
$$\leqq C\delta^{-1/2}\|\Pi_{r,h}v\|_{r,\tau} \leqq C\delta^{m-r-(3/2)}\|\Pi_{r,h}v\|_{m-1,\tau},$$

for $1 \leqq m \leqq r+1$, where we have used the fact that $\Pi_{r,h}v$ is a polynomial of degree r on τ and inequalities (2) and (3). Thus

$$|v-\Pi_{r,h}v|_{q,e} \leqq C\delta^{m-q}(|v|_{m,e} + \delta^{-1/2}\|\Pi_{r,h}v\|_{m-1,\tau})$$
$$\leqq Ch^{m-q}(|v|_{m,e} + h^{-1/2}\|\Pi_{r,h}v\|_{m-1,\tau}).$$

Part (*a*) of the theorem now follows from lemma 2. //

Up to this point we have made no essential use of the macroelements of degree $r \geqq 5$ introduced in section 1: the results proved so far hold for the standard C^1 element $(\tau, P_r(\tau), \overline{\Sigma}_r)$, $r \geqq 5$, of remark 3, section 1. Now, however, we demonstrate that the use of the macroelements of section 1 along $\partial\Omega$ results in a reduced interpolation error for boundary values and normal derivatives when measured in negative Sobolev norms. We deal first with the case of non-nodal degrees of freedom (see remark 1, section 1) on $\partial\Omega$ because the proof is easier and the result is better in the sense that less smoothness is required of the function being interpolated. For Γ a smooth arc in $\mathbf{R}^2$, let $\tilde{P}_m(\Gamma)$ denote the space of functions on Γ which are polynomials in arclength of degree not greater than m. Adopt the convention that $\tilde{P}_{-1}(\Gamma)$ is the set consisting of the null function.

THEOREM 4: *Suppose that each element in $\mathscr{I}_h$ with an edge e in $\partial\Omega$ is such that for $v \in H^{m_0}(\Omega)$*

(*a*) $$\langle v-\Pi_{r,h}v, \psi\rangle_e = 0 \quad \text{when} \quad \psi \in \tilde{P}_{r-4}(e)$$

and/or

$$(b) \qquad \left\langle \frac{\partial}{\partial n}(v-\Pi_{r,h}v),\ \psi \right\rangle e=0 \quad \textit{when } \psi\in \tilde{P}_{r-3}(e).$$

Suppose that $m_0 \leqq m \leqq r+1$. *Then there exists a* C *such that*

(*a*) *for* $v\in H^m(\Omega)\cap H^m(\partial\Omega)$,

$$|v-\Pi_{r,h}v|_{-p,\partial\Omega} \leqq Ch^{m+p}(|v|_{m,\partial\Omega}+\|v\|_{m,\Omega}), \qquad 0\leqq p\leqq r-3,$$

and/or

(*b*) *for* $v\in H^m(\Omega)$ *such that* $\partial v/\partial n\in H^{m-1}(\partial\Omega)$,

$$\left|\frac{\partial}{\partial n}(v-\Pi_{r,h}v)\right|_{-p,\partial\Omega} \leqq Ch^{m+p-1}(|v|_{m-1,\partial\Omega}+\|v\|_{m,\Omega}), \qquad 0\leqq p\leqq r-2.$$

Proof: Let $\varphi\in H^p(\partial\Omega)$ and let $\psi\in L^2(\partial\Omega)$ be such that $\psi|_e\in \tilde{P}_{\bar{p}}(e)$ for all boundary edges e in $\mathscr{I}_h$ and (*cf.* lemma 1)

$$|\varphi-\psi|_{0,\partial\Omega}\leqq Ch^q|\varphi|_{q,\partial\Omega}, \qquad 0\leqq q\leqq p.$$

Thus for $m_0\leqq m\leqq r+1$ and $v\in H^m(\Omega)\cap H^m(\partial\Omega)$,

$$\langle v-\Pi_{r,h}v,\ \varphi\rangle_{\partial\Omega}=\langle v-\Pi_{r,h}v,\ \varphi-\psi\rangle_{\partial\Omega}$$
$$\leqq |v-\Pi_{r,h}v|_{0,\partial\Omega}|\varphi-\psi|_{0,\partial\Omega}\leqq Ch^{m+p}(|v|_{m,\partial\Omega}+\|v\|_{m,\Omega})\,|\varphi|_{p,\partial\Omega},$$

proving part (*a*). Part (*b*) is similar. //

Note that in part (*a*) the upper limit on p is just the number of value degrees of freedom associated with the interior of an edge lying on $\partial\Omega$ and in part (*b*) the upper limit on p is the number of normal derivative degrees of freedom associated with the interior of an edge lying on $\partial\Omega$. It is clear that if the standard C^1 element $(\tau,\ P_r(\tau),\ \overline{\Sigma}_r)$ were used along $\partial\Omega$ (after modification as in remark 1), then a result similar to theorem 4 would hold except that the upper limit on p would be reduced by two in both parts of the theorem.

When nodal degrees of freedom are used along $\partial\Omega$, the orthogonality conditions in theorem 4 can no longer be satisfied exactly. Nevertheless (see [30]), it is possible to place the nodes on $\partial\Omega$ in such a way that the integrals involved in those orthogonality conditions are sufficiently small for the rates of convergence of theorem 4 to be retained, as will now be shown.

Let e be an edge of $\mathscr{I}_h$ on $\partial\Omega$, let $l(e)$ be the length of e and suppose e is parametrized by $s\in[0,\ l(e)]$. Let

$$A_{r,e}(s)=\frac{d^{r-3}}{ds^{r-3}}(s^{r-1}(l(e)-s)^{r-1})$$

and let

$$B_{r,e}(s)=\frac{d^{r-2}}{ds^{r-2}}\left(s^{r-1}(l(e)-s)^{r-1}\right).$$

Integration by parts shows that

$$\langle A_{r,e}, \psi\rangle_e=0, \qquad \psi\in\tilde{P}_{r-4}(e)$$

and

$$\langle B_{r,e}, \psi\rangle_e=0, \qquad \psi\in\tilde{P}_{r-3}(e).$$

It follows that $A_{r,e}$, whose degree is $r+1$, has $r-3$ distinct zeroes in the interior of $[0, l(e)]$ in addition to the second order zeroes at the endpoints and that $B_{r,e}$, whose degree is r, has $r-2$ distinct zeroes in the interior of $[0, l(e)]$ in addition to the zeroes at the endpoints. The zeroes of $B_{r,e}$ are the Lobatto quadrature points for e. Quadrature rules based on both $A_{r,e}$ and $B_{r,e}$ were extensively used in [14].

THEOREM 5: (*a*) *Suppose that each element in* $\mathcal{T}_h$ *with an edge e on* $\partial\Omega$ *is a macroelement with nodes on e placed so that* $v-\Pi_{r,h}v$ *has the same zeroes on e (each with the same order) as* $A_{r,e}$. *If* $m_0\leqq m\leqq r+1$, $0\leqq p\leqq r-3$, *and* $v\in H^m(\Omega)\cap H^{m+p}(\partial\Omega)$, *then*

$$|v-\Pi_{r,h}v|_{-p,\partial\Omega}\leqq Ch^{m+p}\left(|v|_{m+p,\partial\Omega}+\|v\|_{m,\Omega}\right).$$

(*b*) *Suppose that each element in* $\mathcal{T}_h$ *with an edge e on* $\partial\Omega$ *is a macroelement with nodes on e placed so that* $(\partial/\partial n)(v-\Pi_{r,h}v)$ *has the same zeroes on e as* $B_{r,e}$. *If* $m_0\leqq m\leqq r+1$, $0\leqq p\leqq r-2$, $v\in H^m(\Omega)$, *and* $\partial v/\partial n\in H^{m+p-1}(\partial\Omega)$, *then*

$$\left|\frac{\partial v}{\partial n}-\frac{\partial}{\partial n}(\Pi_{r,h}v)\right|_{-p,\partial\Omega}\leqq Ch^{m+p-1}\left(\left|\frac{\partial v}{\partial n}\right|_{m+p-1,\partial\Omega}+\|v\|_{m,\Omega}\right).$$

Proof: We give the proof of part (*a*); part (*b*) is similar.

Let T, τ, e, δ and $l(e)$ be as in the proof of theorem 3. Let $\varphi\in H^p(\partial\Omega)$. Choose ψ as in theorem 4 so that

$$h^{-q}|\varphi-\psi|_{0,e}+|\psi|_{q,e}\leqq C|\varphi|_{q,e}, \qquad 0\leqq q\leqq p.$$

Note that $\rho\delta\leqq l(e)$, where ρ is the constant associated with the non-degeneracy of the family of triangulations. Since

$$\langle v-\Pi_{r,h}v, \varphi\rangle_{\partial\Omega}=\langle v-\Pi_{r,h}v, \varphi-\psi\rangle_{\partial\Omega}+\langle v-\Pi_{r,h}v, \psi\rangle_{\partial\Omega}$$

and

$$\langle v-\Pi_{r,h}v, \varphi-\psi\rangle_{\partial\Omega}\leqq Ch^{m+p}\left(|v|_{m,\partial\Omega}+\|v\|_{m,\Omega}\right)|\varphi|_{p,\partial\Omega}$$

as in the proof of theorem 4, we just need to estimate $\langle v-\Pi_{r,h}v, \psi\rangle_{\partial\Omega}$.

Consider the linear functional E on $C^1([0, l(e)])$ given by

$$E(f)=\int_0^{l(e)} (f(s)-f_I(s))\,ds,$$

where f_I is the unique polynomial of degree r such that the set of zeroes $f-f_I$ include the zeroes (counting multiplicity) of $A_{r,e}$. Note that $E(f)=0$ if f is a polynomial of degree $2r-3$ since $f-f_I$ factors into $A_{r,e}$ times a polynomial of degree $r-4$. Thus, by the Peano Kernel theorem, there exists C such that for $f\in C^k([0, l(e)])$

$$|E(f)| \leqq Cl(e)^k \left\| \left(\frac{d}{ds}\right)^k f \right\|_{L^1([0,\, l(e)])}, \qquad 2\leqq k\leqq 2r-2.$$

Let $\psi\in \tilde{P}_{\bar{p}}(e)$. Then by Schwarz's inequality

$$|E(f\psi)| \leqq Cl(e)^k |f|_{k,e} |\psi|_{p,e}$$

for k an integer in the range $2\leqq k\leqq 2r-2$, with C independent of ψ. Viewing $f\to E(f\psi)$ as a linear functional, it follows by interpolation [22] that the above holds for all real k in that range.

Let $\theta\in\tilde{P}_r(e)$ be such that $\theta-v$ (as a function of arc length) vanishes at the zeroes of $A_{r,e}$ (counting multiplicity). Then

$$\begin{aligned}\langle v-\Pi_{r,h}v, \psi\rangle_e &= \langle v-\theta, \psi\rangle_e + \langle \theta-\Pi_{r,h}v, \psi\rangle_e \\ &\leqq C(\delta^{m+p}|v-\theta|_{m+p,e}+\delta^{2r-2}|\theta-\Pi_{r,h}v|_{2r-2,e})\,|\psi|_{p,e}.\end{aligned}$$

Using an argument similar to that in the proof of theorem 3, we see that

$$|\Pi_{r,h}v|_{2r-2,e}\leqq C\,\delta^{m-r-3/2}\,\|\Pi_{r,h}v\|_{m-1,\tau}.$$

Since θ is a polynomial, we see that

$$\delta^{2r-2}|\theta|_{2r-2,e}=\delta^{2r-2}|\theta|_{r,e}\leqq C\,\delta^{r+m-3}\,|\theta|_{m-1,e}\leqq C\,\delta^{r+m-3}\,|v|_{m+p,e}.$$

Thus from the above estimates and the fact that $p\leqq r-3$, we see that

$$\langle v-\Pi_{r,h}v, \psi\rangle_e\leqq C(\delta^{m+p}\,|v|_{m+p,e}+\delta^{m+r-7/2}\,\|\Pi_{r,h}v\|_{m-1,\tau})\,|\psi|_{p,e}.$$

Summing over e and applying lemma 2, we get

$$\langle v-\Pi_{r,h}v, \psi\rangle_{\partial\Omega}\leqq Ch^{m+p}(|v|_{m+p,\partial\Omega}+\|v\|_{m,\Omega})\,|\varphi|_{p,\partial\Omega},$$

which finishes the proof. //

The last result in this section is that we can extract from the proofs of theorems 3, 4 and 5 the fact that $S_{r,h}$ contains subspaces consisting of functions with "nearly zero" boundary values and normal derivatives (see [24] and [30]).

THEOREM 6: *Let $S^0_{r,h}$ be any subspace of $S_{r,h}$ such that if $\chi \in S^0_{r,h}$ and e is an edge of $\mathscr{I}_h$ on $\partial\Omega$, then*

(*a*) *$\chi=0$ and $\partial\chi/\partial s=0$ at the endpoints of e, and either $\langle \chi, \psi \rangle_e=0$ for $\psi \in \tilde{P}_{r-4}(e)$ or χ vanishes at the zeroes of $A_{r,e}$ in the interior of e, and/or*

(*b*) *$\partial\chi/\partial n=0$ at the endpoints of e and either $\langle \partial\chi/\partial n, \psi \rangle_e=0$ for $\psi \in \tilde{P}_{r-3}(e)$ or $\partial\chi/\partial n$ vanishes at the zeroes of $B_{r,e}$ in the interior of e,*

Then, for $\chi \in S^0_{r,h}$ and $0 \leqq m \leqq 2$,

(*a*) $$|\chi|_{-p,\partial\Omega} \leqq C h^{m+p+(1/2)} \|\chi\|_{m,\Omega}$$

when $-2 \leqq p \leqq r-3$ and $m+p+1/2 \geqq 0$, and/or

(*b*) $$\left|\frac{\partial\chi}{\partial n}\right|_{-p,\partial\Omega} \leqq C h^{m+p-(1/2)} \|\chi\|_{m,\Omega}$$

when $-1 \leqq p \leqq r-2$ and $m+p-1/2 \geqq 0$

Proof: As usual, we only prove part (*a*).

Let e be an edge of $\mathscr{I}_h$ on $\partial\Omega$, let $l(e)$ be the length of e, let τ be the simple triangle associated with $\mathscr{I}_h$ which has e as an edge and let δ be the diameter of τ. Since the family of triangulations is non-degenerate,

$$C^{-1}\delta \leqq l(e) \leqq C\delta.$$

No matter which version of (*a*) is satisfied on e, if $\chi \in S^0_{r,h}$, then $\chi|e$ has $r+1$ zeros, counting multiplicity, so by the Poincaré inequality,

$$|\chi|_{q,e} \leqq C\delta^{r-q+1} |\chi|_{r+1,e}$$

for $0 \leqq q \leqq r+1$. As in the proof of theorem 3, the inequalities (2) and (3) imply that, for $0 \leqq m \leqq r \leqq j$,

$$|\chi|_{j,e} \leqq C\delta^{m-r-(1/2)} \|\chi\|_{m,\tau}. \tag{4}$$

Thus, for $0 \leqq q \leqq r+1$ and $0 \leqq m \leqq r$,

$$|\chi|_{q,e} \leqq C\delta^{m-q+(1/2)} \|\chi\|_{m,\tau}.$$

Summing over all e and using the fact that $\delta \leqq h$, we obtain

$$|\chi|_{q, \partial\Omega} \leqq C h^{m-q+(1/2)} \|\chi\|_{m, \Omega},$$

for $0 \leqq q \leqq 2$, $0 \leqq m \leqq 2$, provided that $m-q+(1/2) \geqq 0$. Letting $q=-p$, we see that part (*a*) of the theorem is now proved for $-2 \leqq p \leqq 0$.

To prove part (*a*) in the range $0 \leqq p \leqq r-3$, let φ and ψ be as in the proof of theorem 5. Then

$$\langle \chi, \varphi-\psi \rangle_e \leqq |\chi|_{0, e} |\varphi-\psi|_{0, e} \leqq C \delta^{m+p+(1/2)} \|\chi\|_{m, \tau} |\varphi|_{p, e}$$

for $0 \leqq m \leqq r$. Furthermore, either $\langle \chi, \psi \rangle_e$ is automatically zero by the definition of $S^0_{r, h}$ in terms of orthogonalities or $\chi\psi$ vanishes at the zeros of $A_{r, e}$. Therefore, as in the proof of theorem 5,

$$\langle \chi, \psi \rangle_e \leqq C \delta^{2r-2} |\chi|_{2r-2, e} |\psi|_{p, e}.$$

Using (4) and the definition of ψ,

$$\langle \chi, \psi \rangle_e \leqq C \delta^{r+m-(5/2)} \|\chi\|_{m, \tau} |\psi|_{p, e}.$$

For $0 \leqq m \leqq r$ and $0 \leqq p \leqq r-3$, we therefore have

$$\langle \chi, \varphi \rangle_e = \langle \chi, \varphi-\psi \rangle_e + \langle \chi, \psi \rangle_e \leqq C \delta^{m+p+(1/2)} \|\chi\|_{m, \tau} |\varphi|_{p, e}.$$

Thus, for $0 \leqq m \leqq 2$, $0 \leqq p \leqq r-3$, and $m+p+1/2 \geqq 0$,

$$\langle \chi, \varphi \rangle_{\partial\Omega} \leqq C h^{m+p+(1/2)} \|\chi\|_{m, \Omega} |\varphi|_{p, \partial\Omega},$$

which completes the proof. //

REMARK 5: Since the techniques of proof in Theorems 2-6 are purely local, it is not necessary to segregate the methods of orthogonality and interpolation. Indeed, when the boundary data is singular, it would be wise to use orthogonality to impose the boundary conditions near the singularity, while interpolation could be used away from the singularity (*cf.* the different smoothness requirements on the data in theorem 4 versus theorem 5). Also, it is not necessary to require that $\partial\Omega$ be smooth globally; if $\partial\Omega$ is piecewise smooth and if a boundary vertex is placed at every point of $\partial\Omega$ where it is not C^∞, the same results follow. //

REMARK 6: It may be desirable to impose the orthogonalities in theorem 4 by evaluating the integral over e using a numerical quadrature rule:

$$\int_e f \sim \sum_{i=1}^{k} w_i f(z_i).$$

In the case that the weights w_i are all nonzero and $\{z_i\}$ corresponds to the set of zeroes of $A_{r,e}$ (resp. $B_{r,e}$) in the interior of $[0, l(e)]$ then orthogonality of $v-\Pi_{r,h}v$ to $\Psi \in \tilde{P}_{r-4}(e)$ [resp. $(\partial/\partial n)(v-\Pi_{r,h}v)$ to $\Psi \in \tilde{P}_{r-3}(e)$] with respect to the quadrature on e is equivalent to vanishing at the quadrature points, i.e., the situation covered by theorem 5. The error induced by using other quadrature rules may be studied using the techniques in the proof of theorem 5. /

3. APPLICATION TO THE PLATE BENDING PROBLEM

Consider the bilinear form $a(.,.)$ on $H^2(\Omega)$ defined by

$$a(u, v) = \frac{D}{2} \int_{\Omega} \{ \Delta u \, \Delta v - (1-\nu)(u_{xx} v_{yy} + u_{yy} v_{xx} - 2 u_{xy} v_{xy})\} \, dx \, dy,$$

where D and ν are constants such that $D > 0$ and $0 \leqq \nu \leqq 1/2$.

Let V be the subspace of $H^2(\Omega)$ consisting of functions that vanish on $\partial\Omega$. Since $a(v, v) \geqq (1/2) D (1-\nu) [v]^2_{2,\Omega}$, it follows from Rellich's lemma [1, chapt. 10] that for some $\gamma < \infty$,

$$a(v, v) \geqq \frac{1}{\gamma} \| v \|^2_{2,\Omega} \quad \text{for all } v \in V. \tag{5}$$

Given $F \in H^{-2}(\Omega)$ and $g \in H^2(\Omega)$, there is a unique $u \in H^2(\Omega)$ such that

$$u - g \in V (\text{i.e., } u = g \text{ on } \partial\Omega)$$

and

$$a(u, v) = F(v) \quad \text{for all } v \in V$$

(the Lax-Milgram theorem). Suppose that the inner product with F is given by

$$F(v) = \int_{\Omega} f v \, dx + \int_{\partial\Omega} M \frac{\partial v}{\partial n} \, ds,$$

where $f \in L^2(\Omega)$ and $M \in L^2(\partial\Omega)$. Then u is the solution to the simply supported plate bending problem corresponding to a loading f, an edge displacement g, and a moment M applied to the edge. The constant ν in the definition of $a(.,.)$ is Poisson's ratio and D is the flexural rigidity [3]. When $f \in H^{s-4}(\Omega)$, $g \in H^{s-(1/2)}(\partial\Omega)$, and $M \in H^{s-(5/2)}(\partial\Omega)$, then $u \in H^s(\Omega)$, with the obvious norm inequality ($s \geqq 2$), since u is related to (f, g, M) by a properly elliptic boundary value problem [22], [1]. We now consider a Galerkin approximation to u.

Let $\Pi_{r,h}$, $0 < h \leqq h_0$, be a family of interpolation operators as studied in the previous section, and let $S^0_{r,h}$ be the image of $V \cap H^{m_0}(\Omega)$ *via* the mapping $\Pi_{r,h}$.

Suppose that

$$a(\chi, \chi) \geqq \frac{1}{2\gamma} \|\chi\|_{2,\Omega}^2 \quad \text{for all } \chi \in S_{r,h}^0. \tag{6}$$

(As can be seen from [30], (6) follows from (5) for h_0 sufficiently small.) Then there is a uniquely determined $u^h \in S_{r,h}$ such that

$$u^h - \Pi_{r,h} g \in S_{r,h}^0 \text{(i.e., } u^h \text{ interpolates } g \text{ on } \partial\Omega)$$

and

$$a(u^h, \chi) = F(\chi) \quad \text{for all } \chi \in S_{r,h}^0.$$

Note that u and u^h depend only on the values of g on $\partial\Omega$.

The Galerkin methods above are direct generalizations to a fourth order elliptic boundary value problem of the methods studied by Blair [4] and in the papers [2] and [30].

THEOREM 7: *Let $\Pi_{r,h}$ be as in either theorem 4 or 5. Suppose that (6) holds for $0 < h \leqq h_0$ and that $u \in H^m(\Omega)$ for m in the range $7/2 < m_0 \leqq m \leqq r+1$. Then*

$$\|u - u^h\|_{2,\Omega} \leqq ch^{m-2} \|u\|_{m,\Omega}.$$

Suppose further that $g \in H^m(\partial\Omega)$ and that $\Pi_{r,h}$ is as in theorem 4. Then

$$\|u - u^h\|_{s,\Omega} \leqq ch^{m-s}(\|u\|_{m,\Omega} + |g|_{m,\partial\Omega})$$

for $3-r \leqq s \leqq 2$. When $\Pi_{r,h}$ is as in theorem 5, suppose that $g \in H^{q+m}(\partial\Omega)$, $0 \leqq q \leqq r-3$. Then

$$\|u - u^h\|_{s,\Omega} \leqq ch^{m-s}(\|u\|_{m,\Omega} + |g|_{m+q,\partial\Omega})$$

for $-q \leqq s \leqq 2$.

Proof: As in [30, section A.1] (6), Green's theorem and the trace theorem [22] imply that, for $m > 7/2$,

$$\|u - u^h\|_{2,\Omega} \leqq C_1 \|u - \Pi_{r,h} u\|_{2,\Omega} + c_2 \|u\|_{m,\Omega} \sup_{\chi \in S_{r,h}^0} \frac{|\chi|_{7/2-m,\partial\Omega}}{\|\chi\|_{2,\Omega}}.$$

Thus theorems 2 and 6 yield the first conclusion. Now let $\varphi \in H^k(\Omega)$, $0 \leqq k \leqq r-3$, and let $\Phi \in V$ solve (*), $a(\Phi), v) = (\varphi, v)$ for all $v \in V$. By elliptic regularity theory (see above), $\|\Phi\|_{k+4,\Omega} \leqq c \|\varphi\|_{k,\Omega}$. Integration by parts (Green's theorem) and the trace theorem yield

$$(u - u^h, \varphi) \leqq |a(u - u^h, \Phi - \Pi_{r,h}\Phi)| \\ + c\|\Phi\|_{k+4,\Omega} |g - u^h|_{-(1/2)-k,\partial\Omega} + c\|u\|_{m,\Omega} |\Pi_{r,h}\Phi|_{7/2-m,\partial\Omega}.$$

(*) Hence forth $(.,.)$ and $\langle .,. \rangle$ will denote $(.,.)_\Omega$ and $\langle .,. \rangle_{\partial\Omega}$.

Using the first part of this theorem, theorem 2, and either theorem 4 or 5,

$$(u-u^h, \varphi) \leqq c\,(h^{m+k}\|u\|_{m,\Omega}+|g-u^h|_{-(1/2)-k,\partial\Omega})\|\Phi\|_{k+4,\Omega}$$

(recall that $\Phi \in V$, i.e., $\Phi \equiv 0$ on $\partial\Omega$, so that the restrictions of theorems 4 and 5 are the same). Thus regardless of the type of interpolation,

$$\|u-u^h\|_{-k,\Omega} \leqq c\,(h^{m+k}\|u\|_{m,\Omega}+|g-u^h|_{-(1/2)-k,\partial\Omega}).$$

Now consider u^h. Theorem 6 implies that

$$\begin{aligned}|g-u^h|_{-(1/2)-k,\partial\Omega} &\leqq |u^h-\Pi_{r,h}u|_{-(1/2)-k,\partial\Omega}+|g-\Pi_{r,h}u|_{-(1/2)-k,\partial\Omega}\\ &\leqq ch^{k+2}\|u^h-\Pi_{r,h}u\|_{2,\Omega}+|u-\Pi_{r,h}u|_{-(1/2)-k,\partial\Omega}.\end{aligned}$$

Inserting $(u-u)$ in the first term and using the triangle inequality, the first part of this theorem and theorem 2 imply that

$$\|u-u^h\|_{-k,\Omega} \leqq c\,(h^{m+k}\|u\|_{m,\Omega}+|u-\Pi_{r,h}u|_{-(1/2)-k,\partial\Omega}).$$

Apply either theorem 4 or 5 to estimate the second term on the right hand side, and the result follows. //

The above theorem can clearly be extended to allow more general coefficients in $a(.,.)$. The clamped plate problem may be treated by simply changing V in the above to be the subspace of $H^2(\Omega)$ consisting of functions that vanish to second order on $\partial\Omega$ and using part (*b*) of theorems 4-6.

4. THE H^1-GALERKIN METHOD FOR SECOND ORDER ELLIPTIC PROBLEMS

Consider the Dirichlet problem

$$\begin{aligned} Lu=\nabla\cdot(a\nabla u)=f \quad &\text{in } \Omega,\\ u=g \quad &\text{on } \partial\Omega, \end{aligned} \tag{7}$$

where $a=a(x)$ is a smooth, positive function on $\overline{\Omega}$. Again let $V=\{v\in H^2(\Omega) : v=0 \text{ on } \partial\Omega\}$, and let $S^0_{r,h}$ be the image of $V\cap H^{m_0}(\Omega)$ under the mapping $\Pi_{r,h}$, where $\Pi_{r,h}$ is determined by either the interpolation procedure associated with theorem 5 or the orthogonality conditions associated with theorem 4. Assume that g has a (theoretical) extension to $H^{m_0}(\Omega)$, so that $\Pi_{r,h}\,g$ is definable; practically, this involves only the values of g on $\partial\Omega$. Then the H^1-Galerkin method for approximating the solution of (7) consists of finding $u^h\in S_{r,h}$ such that

$$\begin{aligned} (Lu^h, \Delta v)=(f, \Delta v), \qquad &v\in S^0_{r,h},\\ u^h-\Pi_{r,h}g\in S^0_{r,h}, \end{aligned} \tag{8}$$

H^1-Galerkin methods have been proposed earlier in [16] for both elliptic and parabolic problems for the special case $L=\Delta$, and the nonlinear Dirichlet problem based on (7) with $a=a(x, u)$, plus a linear problem with lower order terms included, have been treated in [13]. In both these papers the Dirichlet boundary condition was imposed weakly through penalty-like terms on the boundary. The method (8) corresponding to interpolation was mentioned briefly in [13], but no analysis was given in that case. The orthogonality method (8) presents an H^1 analogue of the method of Blair [4].

Note that the algebraic equations arising from (8) do not, in general, generate a symmetric matrix, and care must be taken to show that a solution of (8) exists. There are significant practical advantages of (8) over least squares methods [6, 7], particularly for nonlinear problems and in applications to transient problems, since the algebraic equations become simpler. The analysis of (8) below is similar to that given in [13], but the details are noticeably different. Both rely on ideas discussed by Schatz [28] earlier.

LEMMA 3 (Gårding inequality): *There exist constants* $h_0 > 0$, $\rho > 0$, *and* C *such that*

$$(L v, \Delta v) \geqq \rho \| v \|_{2,\Omega}^2 - C \| v \|_{0,\Omega}^2$$

for $0 < h \leqq h_0$ *and* $v \in \overset{\circ}{S}_{r,h}$.

Proof: Since a is bounded below positively and ∇a is bounded, it is trivial to see that

$$(L v, \Delta v) \geqq \rho_1 \| \Delta v \|_{0,\Omega}^2 - C_1 \| v \|_{1,\Omega}^2, \qquad v \in H^2(\Omega).$$

For $v \in \overset{\circ}{S}_{r,h}$, theorem 6 implies that $| v |_{3/2,\partial\Omega} \leqq C h \| v \|_{2,\Omega}$. Since $\| v \|_{2,\Omega}$ and $\| \Delta v \|_{0,\Omega} + | v |_{3/2,\partial\Omega}$ are equivalent [22], a simple version [1] of interpolation of Sobolev norms implies that

$$(L v, \Delta v) \geqq \rho_2 \| v \|_{2,\Omega}^2 - C_1 (\varepsilon \| v \|_{2,\Omega}^2 + \varepsilon^{-1} \| v \|_{0,\Omega}^2) \geqq \rho \| v \|_{2,\Omega}^2 - C \| v \|_{0,\Omega}^2$$

for $v \in \overset{\circ}{S}_{r,h}$ and h sufficiently small.

LEMMA 4: *If* $\zeta = u - u^h$ *and* $s \in [-2, r-3]$, *then*

$$\| \zeta \|_{-s,\Omega} \leqq C \{ h^{s+2} \| \zeta \|_{2,\Omega} + | \zeta |_{-s-(1/2),\partial\Omega} \}.$$

Proof: Let $w \in H^s(\Omega)$ and determine $\chi \in H^{s+2}(\Omega)$ and $\psi \in H^{s+(1/2)}(\partial\Omega)$ such that

$$(v, w) = (L v, \chi) + \langle v, \psi \rangle$$

for all $v \in H^2(\Omega)$. The existence of χ and ψ follows from lemma 6.1 of Douglas-Dupont [15]; also the following bound holds:

$$\|\chi\|_{s+2,\Omega}+|\psi|_{s+(1/2),\partial\Omega} \leqq C\|w\|_{s,\Omega}$$

($\psi = a(\partial\chi/\partial n)|_{\partial\Omega}$ whenever the latter makes sense). Let φ solve Poisson's equations:

$$\Delta\varphi=\chi \quad \text{in } \Omega,$$

$$\varphi=0 \quad \text{on } \partial\Omega.$$

Then $\varphi \in H^{s+4}(\Omega)$ and

$$(\zeta, w)=(L\zeta, \Delta\varphi)+\langle\zeta, \psi\rangle.$$

Since $(L\zeta, \Delta v)=0$ for $v \in S^0_{r,h}$, an appropriate choice of $v \in S^0_{r,h}$ yields

$$\begin{aligned}(\zeta, w)&=(L\zeta, \Delta(\varphi-v))+\langle\zeta,\psi\rangle\\ &\leqq C\|\zeta\|_{2,\Omega}h^{s+2}\|\varphi\|_{s+4,\Omega}+|\zeta|_{-s-(1/2),\partial\Omega}|\psi|_{s+(1/2),\partial\Omega}\\ &\leqq C(\|\zeta\|_{2,\Omega}h^{s+2}+|\zeta|_{-s-(1/2),\partial\Omega})\|w\|_{s,\Omega}\end{aligned}$$

and the proof is finished. //

LEMMA 5: *Let $\alpha=0$ if orthogonality determines $\Pi_{r,h}$, and let $\alpha=s+(1/2)$ if interpolation determines $\Pi_{r,h}$. If $-1/2 \leqq s \leqq r-(7/2)$, then*

$$|\zeta|_{-s-(1/2),\partial\Omega} \leqq C h^{s+3}\|\zeta\|_{2,\Omega}+C h^{m+s+(1/2)}(\|u\|_{m,\Omega}+|g|_{m+\alpha,\partial\Omega})$$

for $m_0 \leqq m \leqq r+1$.

Proof: Let $-1/2 \leqq s \leqq r-(7/2)$. Then

$$|\zeta|_{-s-(1/2),\partial\Omega} \leqq |u-\Pi_{r,h}u|_{-s-(1/2),\partial\Omega}+|\Pi_{r,h}u-u^h|_{-s-(1/2),\partial\Omega}.$$

Apply either theorem 4 or 5 to the first term on the right hand side and theorem 6 to the second. After a trivial simplification, the desired inequality results. //

Logically it would perhaps have been better to show existence and uniqueness of a solution of (8) before treating lemmas 4 and 5, but it would have induced an unnecessary duplication of argument.

LEMMA 6: *For h sufficiently small there exists a unique solution $u^h \in S_{r,h}$ of* (8).

Proof: It is clear that uniqueness implies existence and that the difference z of two solutions of (8) is an element of $S^0_{r,h}$ satisfying $(Lz, \Delta v)=0$ for $v \in S^0_{r,h}$. Thus, z corresponds to ζ in the case $u\equiv 0$, and lemma 3, 4, and 5 with $s=0$ imply that

$$\|z\|_{2,\Omega} \leqq C h^2\|z\|_{2,\Omega}.$$

Hence $z=0$ for small h. //

THEOREM 8: *Let $\zeta = u - u^h$, where u is the solution of* (7) *and u^h is the unique solution of* (8) *for small h. Let $m_0 \leqq k \leqq r+1$ and*

$$m = m(p) = \max\{m_0,\ k - p - (1/2)\},$$

where m_0 appears in condition (c) *of section* 2. *Then, if $r \geqq 4$,*

$$\|\zeta\|_{p,\Omega} \leqq C(h^{k-p}\|u\|_{k,\Omega} + h^{m+(1/2)}|g|_{m+\alpha,\partial\Omega}), \qquad 0 \leqq p \leqq 2,$$

where $\alpha = 0$ for $\Pi_{r,h}$ determined by orthogonality and $\alpha = 1/2$ for $\Pi_{r,h}$ determined by interpolation. If $r = 3$, $|g|_{m+\alpha,\partial\Omega}$ should be replaced by $|g|_{m+\alpha+(1/2),\partial\Omega}$ in the inequality. If $0 \leqq s \leqq r - (7/2)$, then

$$\|\zeta\|_{-s,\Omega} \leqq C(h^{k+s}\|u\|_{k,\Omega} + h^{n+s+(1/2)}|g|_{n+\beta,\partial\Omega}),$$

where $n = \max\{m_0,\ k - (1/2)\}$ and $\beta = 0$ or $s + (1/2)$ if $\Pi_{r,h}$ is determined by orthogonality or interpolation, respectively. If $r - (7/2) \leqq s \leqq r - 3$, then

$$\|\zeta\|_{-s,\Omega} \leqq C(h^{k+s}\|u\|_{k,\Omega} + h^{t+r-3}|g|_{t+\gamma,\partial\Omega}),$$

where $t = \max(m_0,\ k + s - r + 3)$ and $\gamma = 0$ or $r - 3$ for $\Pi_{r,h}$ as above.

Proof: We have seen that $\|\zeta\|_{0,\Omega} \leqq C h^2 \|\zeta\|_{2,\Omega} + |\zeta|_{-1/2,\partial\Omega}$. Let

$$\xi = \Pi_{r,h}\, u - u^h \in S^0_{r,h},$$

and apply the Gårding inequality to ξ. Then

$$\begin{aligned}\rho\|\xi\|^2_{2,\Omega} &\leqq (L\xi,\ \Delta\xi) + C\|\xi\|^2_{0,\Omega}\\ &= (L(\Pi_{r,h}\, u - u),\ \Delta\xi) + C\|\xi\|^2_{0,\Omega}\\ &\leqq C\|\Pi_{r,h}\, u - u\|_{2,\Omega}\|\xi\|_{2,\Omega} + C\|\xi\|^2_{0,\Omega},\end{aligned}$$

and

$$\|\xi\|_{2,\Omega} \leqq C(\|u\|_{k,\Omega}\, h^{k-2} + \|\xi\|_{0,\Omega}), \qquad m_0 \leqq k \leqq r+1.$$

Thus,

$$\|\zeta\|_{2,\Omega} \leqq C(\|u\|_{k,\Omega}\, h^{k-2} + \|\zeta\|_{0,\Omega}) \leqq C(\|u\|_{k,\Omega}\, h^{k-2} + \|\zeta\|_{2,\Omega}\, h^2 + |\zeta|_{-1/2,\partial\Omega}).$$

For small h and by lemma 5, if $r \geqq 4$,

$$\begin{aligned}\|\zeta\|_{2,\Omega} &\leqq C(\|u\|_{k,\Omega}\, h^{k-2} + |\zeta|_{-1/2,\partial\Omega})\\ &\leqq C\{\|u\|_{k,\Omega}\, h^{k-2} + h^{m+(1/2)}(\|u\|_{m,\Omega} + |g|_{m+\alpha,\partial\Omega})\}.\end{aligned}$$

Choose $m = \max\{m_0, k-5/2\}$ as called for in the statement of theorem 8 to obtain the desired bound for $p=2$ and $r \geqq 4$. The remainder of the proof consists of a careful application of lemmas 4 and 5. //

The two H^1 – Galerkin methods for the Dirichlet problem can be used to motivate H^1 – treatments of parabolic problems. See [16] for a simple case (however, with penalty-set boundary values) and [15] for a somewhat analogous development. See also [33] for another related concept.

REFERENCES

1. S. AGMON, *Elliptic Boundary Values Problems*, D. van Nostrand, 1965.
2. A. BERGER, R. SCOTT and G. STRANG, *Approximate Boundary Conditions in the Finite Element Method*, Symposia Mathematica, X, Academic Press, 1972, pp. 295-313.
3. S. BERGMAN and M. SCHIFFER, *Kernel Functions and Elliptic Differential Equations in Mathematical Physics*, Academic Press, 1953.
4. J. J. BLAIR, *Higher Order Approximations to the Boundary Conditions for the Finite Element Method*, Math. Comp., Vol. 30, 1976, pp. 250-262.
5. J. H. BRAMBLE and S. R. HILBERT, *Estimation of Linear Functionals on Sobolev Spaces with Applications to Fourier Transforms and Spline Interpolation*, S.I.A.M. J. Numer. Anal., Vol. 7, 1970, pp. 112-124.
6. J. H. BRAMBLE and A. H. SCHATZ, *Rayleigh-Ritz-Galerkin Methods for Dirichlet's Problem Using Subspaces Without Boundary Conditions*, Comm. Pure App. Math., Vol. 23, 1970, pp. 653-674.
7. J. H. BRAMBLE and A. H. SCHATZ, *Least Squares Methods for 2m-th Order Elliptic Boundary-Value Problems*, Math. Comp., Vol. 25, 1971, pp. 1-32.
8. P. G. CIARLET, *Sur l'élément de Clough et Tocher, R.A.I.R.O.*, Analyse numérique, Vol. 2, 1974, pp. 19-27.
9. P. G. CIARLET, *Numerical Analysis of the Finite Element Method*, Séminaire de Mathématiques supérieures, Université de Montréal, 1975.
10. P. G. CIARLET and P.-A. RAVIART, *General Lagrange and Hermite Interpolation in* $\mathbf{R}^n$ *with Applications to Finite Element Methods*, Arch. Rational Mech. Anal., Vol. 46, 1972, pp. 177-199.
11. J. F. CIAVALDINI and J. C. NEDELEC, *Sur l'élément de Fraeijs de Veubeke et Sander, R.A.I.R.O.*, Analyse numérique, Vol. 2, 1974, pp. 29-46.
12. R. W. CLOUGH and J. L. TOCHER, *Finite Element Stiffness Matrices and Analysis of Plates in Bending*, Proceedings of Conference on Matrix Methods in Structural Mechanics, Wright-Patterson AFB, 1965.
13. J. DOUGLAS, Jr., *H^1-Galerkin Methods for a Nonlinear Dirichlet Problem*, Mathematical Aspects of Finite Element Methods, Rome, 1975, Lecture Notes in Mathematics, n° 606, Springer-Verlag, 1977, pp. 64-86.
14. J. DOUGLAS, Jr. and T. DUPONT, *Collocation Methods for Parabolic Equations in a Single Space Variable*, Lecture Notes in Mathematics, n° 385, Springler-Verlag, 1974.
15. J. DOUGLAS, Jr. and T. DUPONT, *H^{-1}-Galerkin Methods for Problems Involving Several Space Variables*, Topics in Numerical Analysis, III, John J. H. MILLER, ed., Academic Press, 1977, pp. 125-141.

16. J. DOUGLAS, Jr., T. DUPONT and M. F. WHEELER, *H^1-Galerkin Methods for the Laplace and Heat Equations*, Mathematical Aspects of Finite Elements in Partial Differential Equations, C. DE BOOR, ed., Academic Press, 1974, pp. 383-416.

17. T. DUPONT and R. SCOTT, *Polynomial Approximation of Functions in Sobolev Spaces*, submitted Math. Comp.

18. B. FRAEIJS DE VEUBEKE, *Bending and Stretching of Plates*, Proceedings of Conference on Matrix Methods in Structural Mechanics, Wright-Patterson AFB, 1965.

19. P. GRISVARD, *Behavior of the Solutions of an Elliptic Boundary Value Problem in Polygonal or Polyhedral Domain*, Numerical Solution of Partial Differential Equations, III (Synspade, 1975), Bert HUBBARD, ed., Academic Press, 1976, pp. 207-274.

20. P. JAMET, *Estimation de l'erreur d'interpolation dans un domaine variable et application aux éléments finis quadrilatéraux dégénérés*, in *Méthodes numériques en mathématiques appliquées* (Séminaire de Mathématiques supérieures, été 1975), Presses de l'Université de Montréal, Vol. 60, 1977.

21. I. N. KATZ, A. G. PEANO and B. A. SZABO, *Nodal Variables for Arbitrary Order Conforming Finite Elements*, U. S. Dept. of Transportation Tech. Rep. DOT-OS-30108-5, Washington Univ., June, 1975.

22. J. L. LIONS and E. MAGENES, *Problèmes aux limites non homogènes et applications*, Vol. 1, Dunod, Paris, 1968.

23. J. MORGAN and R. SCOTT, *A Nodal Basis for C^1 Piecewise Polynomials of Degree $n \geqq 5$*, Math. Comp., Vol. 29, 1975, pp. 736-740.

24. J. NITSCHE, *On Dirichlet Problems Using Subspaces with Nearly Zero Boundary Conditions*, The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, A. K. AZIZ, ed., Academic Press, 1972, pp. 603-628.

25. A. G. PEANO, *Hierarchies of Conforming Finite Elements for Plane Elasticity and Plate Bending*, Comp. and Maths. with Appls., Vol. 2, 1976, pp. 211-224.

26. P. PERCELL, *On Cubic and Quartic Clough-Tocher Finite Elements*, S.I.A.M. J. Numer. Anal., Vol. 13, 1976. pp. 100-103.

27. G. SANDER, *Bornes supérieures et inférieures dans l'analyse matricielle des plaques en flexion-torsion*, Bull. Soc. Royale des Sc. de Liège, Vol. 33, 1964, pp. 456-494.

28. A. H. SCHATZ, *An Observation Concerning Ritz-Galerkin Methods with Indefinite Bilinear Forms*, Math. Comp., Vol. 28, 1974, pp. 959-962.

29. R. SCOTT, *C^1 Continuity via Constraints for 4th Order Problems*, Mathematical Aspects of Finite Elements in Partial Differential Equations, C. DE BOOR, ed., Academic Press, 1974, pp. 171-193.

30. R. SCOTT, *Interpolated Boundary Conditions in the Finite Element Method*, S.I.A.M. J. Numer. Anal., Vol. 12, 1975, pp. 404-427.

31. G. STRANG, *Piecewise Polynomials and the Finite Element Method*, Bull. A.M.S., Vol. 79, 1973, pp. 1128-1137.

32. B. A. SZABO *et al.*, *Advanced Design Technology for Rail Transportation Vehicles*, U.S. Dept. of Transportation Tech. Rep. DOT-OS-30108-2, Washington Univ., June, 1974.

33. V. THOMÉE and L. WAHLBIN, *On Galerkin Methods in Semi-Linear Parabolic Problems*, S.I.A.M. J. Num. Anal., Vol. 12, 1975, pp. 378-389.

34. O. C. ZIENKIEWICZ, *The Finite Element Method in Engineering Science*, McGraw-Hill, 1971.