Triangular Elements in the Finite Element Method

By James H. Bramble and Miloš Zlámal

Abstract. For a plane polygonal domain Ω and a corresponding (general) triangulation we define classes of functions $p_m(x, y)$ which are polynomials on each triangle and which are in $C^{(m)}(\Omega)$ and also belong to the Sobolev space $W_2^{(m+1)}(\Omega)$. Approximation theoretic properties are proved concerning these functions. These results are then applied to the approximate solution of arbitrary-order elliptic boundary value problems by the Galerkin method. Estimates for the error are given. The case of second-order problems is discussed in conjunction with special choices of approximating polynomials.

1. Introduction. The classical Ritz and Galerkin method has several advantages over the finite-difference method. Nevertheless two things have prevented its more extensive use: 1. The practical construction of the basic functions in more dimensions was possible only for some simple domains. 2. Even for these domains the procedures can be highly unstable.

The finite element method is nothing else than the Ritz or Galerkin method using special trial functions. The first idea goes back to Courant [9] who suggested triangulating the given domain and using functions which are linear on each triangle as trial functions for solving boundary value problems of the second order. This idea was rediscovered by the engineers and developed, originally as a concept of structural analysis, into a method called the finite element method (see Turner, Clough, Martin and Topp [15] and the references in Zienkiewicz [18]). Practical experience, the large amount of numerical results and the first theoretical results show that the finite element method removes the above mentioned shortcomings of the classical Ritz and Galerkin method.

One feature of the procedures described by the engineers consists in introducing higher degree polynomials for interpolation of the solution on the given element. Some procedures of this kind for triangular elements were proposed and justified by the second of the authors [19]. For fourth-order equations the trial functions used are polynomials of the fifth degree.[1] The results and the method of [19] were generalized by Ženíšek [17]. He proposed to use polynomials of the degree $4m + 1$ introduced later in this paper[2] and he justified the method for $m = 2, 3$ (the case $m = 1$ being justified in [19]).

The method of this paper differs completely from the method of [19]. A lemma about linear functionals on $W_p^{(k)}$ by Bramble and Hilbert [7] allows us to get general results for any m. We prove a general interpolation theorem and apply it to V-elliptic

---

[1] Almost simultaneously this procedure was described and applied to bending of plates by Bell [3] and [4], Visser [16], Bosshard [6] and Argyris, Fried, Scharpf [2].

[2] As a matter of fact, he also introduces polynomials of the degree $4m + 2$, $4m + 3$ and $4(m + 1)$. We restrict ourselves to the case of polynomials of the degree $4m + 1$. The others are easy to deal with in the same way.

boundary value problems of arbitrary order. The seminorm used in this paper for the discretization error is more appropriate than that used in [19].

2. **Interpolation Polynomials on Triangles.** To define the interpolation polynomials introduced by Ženíšek [17] we denote by P_j ($j = 1, 2, 3$) the vertices of a triangle T,[3] by (x_j, y_j) the coordinates of P_j, by P_0 the center of gravity of T, by l_j the sides of T, by ν_j the normals to l_j. We divide every side l_j in $r + 1$ equal parts ($r = 1, 2, \cdots$) by the points $Q_j^{(\rho, r)}$ ($j = 1, 2, 3, \rho = 1, \cdots, r$).

Now a polynomial $p_m(x, y)$ in two variables of the degree $4m + 1$ ($m = 0, 1, \cdots$) has $(2m + 1)(4m + 3)$ coefficients. Hence we cannot prescribe more than $(2m + 1) \cdot (4m + 3)$ conditions for such a polynomial. Let us prescribe the following values:

(1) $\qquad D^i p_m(P_j),$[4] $\qquad j = 1, 2, 3, \quad |i| \leqq 2m,$

(2) $\qquad \dfrac{\partial^r p_m(Q_j^{(\rho, r)})}{\partial \nu_j^r}, \qquad j = 1, 2, 3, \quad \rho = 1, \cdots, r, \quad r = 1, \cdots, m,$

(3) $\qquad D^i p_m(P_0), \qquad |i| \leqq m - 2.$

We must add that we leave out the values (2) and (3) if $m = 0$ and $m = 0, 1$, respectively. Thus, $p_0(x, y)$ is a linear polynomial determined by the values of $u(x, y)$ at the vertices of T and $p_1(x, y)$ is the polynomial introduced in [19, p. 404] and in the papers quoted in footnote 1.

The importance of the polynomials $p_m(x, y)$ follows from the property proved in [17] which we formulate in this way: Suppose the values of the form (1), (2), (3) determine uniquely a polynomial $p_m(x, y)$ of the degree not greater than $4m + 1$. Let Ω be a polygonal domain triangulated by triangles $\{T_k\}_{k=1}^M$ and let values of the form (1), (2), (3) be prescribed at every vertex of the triangulation, at every point $Q_j^{(\rho, r)}$ and at every center of gravity. Then the function $v(x, y)$ which on every T_k is equal to a polynomial $p_m^k(x, y)$ defined in the way just described belongs to $C^{(m)}(\bar{\Omega})$. Later we shall construct trial functions for the Galerkin method by means of the polynomials $p_m(x, y)$. First, we must, of course, prove the existence and uniqueness of $p_m(x, y)$.

THEOREM 1. *There exists exactly one polynomial $p_m(x, y)$ of the degree not greater than $4m + 1$ assuming the values* (1), (2), (3).

Proof. The assertion is trivial for $m = 0$, hence we consider $m \geqq 1$. It is sufficient to prove that if

(4) $\qquad D^i p_m(P_j) = 0, \qquad j = 1, 2, 3, \quad |i| \leqq 2m,$

(5) $\qquad \dfrac{\partial^r p_m(Q_j^{(\rho, r)})}{\partial \nu_j^r} = 0, \qquad j = 1, 2, 3, \quad \rho = 1, \cdots, r, \quad r = 1, \cdots, m,$

(6) $\qquad D^i p_m(P_0) = 0, \qquad |i| \leqq m - 2,$

and $p_m(x, y)$ is a polynomial of a degree not greater than $4m + 1$ then $p_m(x, y) \equiv 0$. (That is, the linearity of (1), (2) and (3) permits the uniqueness of their solution to imply existence of a solution.)

---

[3] At the same time T means the interior of T; it will always be clear what meaning of T is necessary to be taken.

[4] Here $i = (i_1, i_2)$, $|i| = i_1 + i_2$, $D^i u = \partial^{|i|} u / \partial x^{i_1} \partial y^{i_2}$.

The derivatives $\partial^r p_m / \partial v_j^r$ $(r = 0, \cdots, m, j = 1, 2, 3)$ are Hermite polynomials (see, for instance, [5]) in one variable on the corresponding sides of the triangle T which, with respect to (4) and (5), assume homogeneous boundary values. Therefore they are identically equal to zero on the sides of T. Using the reasoning of the proof of Theorem 1 in [17] we find out that

$$(7) \qquad D^i p_m(x, y)\,|_{\delta T} = 0, \qquad |i| \leq m.$$

Now let us consider the transformation

$$(8) \qquad \begin{aligned} x &= x(\xi, \eta) \equiv x_1 + (x_2 - x_1)\xi + (x_3 - x_1)\eta, \\ y &= y(\xi, \eta) \equiv y_1 + (y_2 - y_1)\xi + (y_3 - y_1)\eta \end{aligned}$$

and the polynomial $\tilde{p}_m(\xi, \eta) = p_m[x(\xi, \eta), y(\xi, \eta)]$. The equations (8) map T onto the triangle T_1 with vertices $\tilde{P}_1(0, 0)$, $\tilde{P}_2(1, 0)$, $\tilde{P}_3(0, 1)$. The points $Q_j^{(\rho, r)}$ are mapped on the points $\tilde{Q}_j^{(\rho, r)}$ which again divide the new sides $\tilde{l}_i$ into $r + 1$ equal parts and P_0 is mapped on the center of gravity $\tilde{P}_0(\frac{1}{3}, \frac{1}{3})$ of the triangle T_1. From (7) and (6) it follows that

$$(9) \qquad D^i \tilde{p}_m(\xi, \eta)\,|_{\delta T_1} = 0, \qquad |i| \leq m,$$

$$(10) \qquad D^i \tilde{p}(P_0) = 0, \qquad |i| \leq m - 2,$$

(if we use the symbol D applied to functions of ξ and η we always mean a derivative with respect to ξ and η; thus $D^i \tilde{p}_m(\xi, \eta) = \partial^{|i|} \tilde{p}_m(\xi, \eta)/(\partial \xi^{i_1} \partial \eta^{i_2})$). A consequence of (9) is that $\partial^r \tilde{p}_m(\xi, 0)/\partial \eta^r = 0$ for $0 \leq \xi \leq 1, r = 0, \cdots, m$. Therefore $\tilde{p}_m(\xi, \eta)$ is divisible by η^{m+1}. Similarly, one can show that $\tilde{p}_m(\xi, \eta)$ is divisible by $(1 - \xi - \eta)^{m+1}$ and by ξ^{m+1}. Hence, if $m = 1$ it must be that $\tilde{p}_1(\xi, \eta) \equiv 0$, and consequently $p_1(x, y) \equiv 0$, and if $m \geq 2$

$$\tilde{p}_m(\xi, \eta) = [\xi \eta (1 - \xi - \eta)]^{m+1} Q(\xi, \eta),$$

where $Q(\xi, \eta)$ is a polynomial of the degree not greater than $m - 2$. Now it is sufficient to use (10). Since $[\xi \eta (1 - \xi - \eta)]_{\xi = \eta = 1/3} \neq 0$ we get

$$D^i Q(\tilde{P}_0) = 0, \qquad |i| \leq m - 2,$$

and since $Q(\xi, \eta)$ is a polynomial of the degree not greater than $m - 2$ it follows that $Q(\xi, \eta) \equiv 0$, hence $\tilde{p}_m(\xi, \eta) \equiv 0$ and $p_m(x, y) \equiv 0$.

Next what we need is some estimate of the error arising when we approximate a function $u(x, y) \in C^{(2m)}(\bar{T})$ by a polynomial $p_m(x, y)$. We will say that $p_m(x, y)$ is the interpolation polynomial corresponding to $u(x, y)$ if

$$(11) \qquad D^i p_m(P_j) = D^i u(P_j), \qquad j = 1, 2, 3, \quad |i| \leq 2m,$$

$$(12) \qquad \partial^r p_m(Q_j^{(\rho, r)})/\partial v_j^r = \partial^r u(Q_j^{(\rho, r)})/\partial v_j^r,$$

$$j = 1, 2, 3, \quad \rho = 1, \cdots, r, \quad r = 1, \cdots, m,$$

$$(13) \qquad D^i p_m(P_0) = D^i u(P_0), \qquad |i| \leq m - 2.$$

To get the estimate we make use of a lemma by Bramble and Hilbert [7]. First we introduce some notation. By $W_2^{(k)}(\Omega)$ we denote the Hilbert space of all functions which together with their generalized derivatives up to the kth order belong to $L_2(\Omega)$.

The norm is given by

$$||u||^2_{k,\Omega} = \sum_{j=1}^{k} |u|^2_{j,\Omega}, \quad \text{where} \quad |u|^2_{j,\Omega} = \sum_{|i|=j} ||D^i u||^2_{L_2(\Omega)}.$$

LEMMA.[5] *Let Ω be a bounded domain in E_N with* diam $(\Omega) = 1$. *Assume that Ω satisfies the ordinary cone condition (see [1]). Let $F(u)$ be a bounded linear functional on $W_2^{(k)}(\Omega)$,*

$$|F(u)| \leqq C_1 ||u||_{k,\Omega},$$

and let $F(q) = 0$ for every polynomial q of the degree less than k. Then there exists a constant C_2 depending on the cone condition only such that

(14) $$|F(u)| \leqq C_1 C_2 |u|_{k,\Omega}$$

for all $u \in W_2^{(k)}(\Omega)$.

THEOREM 2. *Let $u(x, y) \in W_2^{(k)}(T)$ where $2m + 2 \leqq k \leqq 4m + 2$. Let $p_m(x, y)$ be the interpolation polynomial corresponding to $u(x, y)$. Then, for $0 \leqq n \leqq k$,*

(15) $$||u - p_m||_{n,T} \leqq \frac{K}{(\sin \alpha)^{m+n}} c^{k-n} |u|_{k,T},$$

where the constant K does not depend on the triangle T and the function u and where α is the smallest angle and c is the length of the greatest side of T.

Proof. We denote by $\alpha \leqq \beta \leqq \gamma$ the angles of the triangle T and we choose the notation of the vertices such that α lies at P_1, β at P_2 and γ at P_3. The lengths of the sides are denoted by a, b, c, a being the smallest and c the greatest. As $a + b > c$ we have $b > \frac{1}{2}c$. The area of T is equal to one half of $|J|$ where J is the Jacobian of the transformation (8) so that

$$|J|^{-1} = \frac{1}{bc \sin \alpha} < \frac{2}{c^2 \sin \alpha}.$$

For the inverse transformation to (8) we easily find out that

(16) $$\left|\frac{\partial \xi}{\partial x}\right|, \; \left|\frac{\partial \xi}{\partial y}\right|, \; \left|\frac{\partial \eta}{\partial x}\right|, \; \left|\frac{\partial \eta}{\partial y}\right| \leqq \frac{2}{c \sin \alpha}.$$

Let us denote $w(x, y) = u(x, y) - p_m(x, y)$ and consider the function $\tilde{w}(\xi, \eta) = w[x(\xi, \eta), y(\xi, \eta)]$. The derivatives $D^i \tilde{w}(\xi, \eta)$ are linear combinations of the derivatives $D^i w(x, y)$ and using (16) we easily obtain

(17) $$||w||_{n,T} \leqq \frac{K_1}{(c \sin \alpha)^n} |J|^{1/2} ||\tilde{w}||_{n,T_1}.$$

Here K_1 is a constant which does not depend on T and the functions considered (in the sequel we shall denote such constants by $K_1, K_2, \cdots$).

Now to get an estimate for $||\tilde{w}||_{n,T_1}$ we apply the Lemma. Let us consider the linear functional $F(\bar{u}) = (\bar{u} - \tilde{p}_m, v)_{n,T_1}$ on $W_2^{(k)}(T_1)$ where $(\tilde{w}, v)_{n,T_1}$ means the scalar product in $W_2^{(n)}(T_1)$ and v is an arbitrary function from $W_2^{(n)}(T_1)$. If $\bar{u}(\xi, \eta)$ is a polynomial of the degree less than k then $u(x, y)$ is also a polynomial of the degree less

_____

[5] Actually, it is true for more general spaces $W_p^{(k)}(\Omega)$ and the formulation introduced here differs a little from the formulation introduced in [7].

than k. For $k \leq 4m + 2$ it follows by Theorem 1 that $u(x, y) - p_m(x, y) \equiv 0$, hence $\tilde{u}(\xi, \eta) - \tilde{p}_m(\xi, \eta) \equiv 0$ and $F(\tilde{u}) = 0$. Further,

$$|F(\tilde{u})| \leqq ||v||_{n, T_1} ||\tilde{u} - \tilde{p}_m||_{n, T_1} \leqq ||v||_{n, T_1} \{||\tilde{u}||_{k, T_1} + ||\tilde{p}_m||_{k, T_1}\}.$$

Assume we succeed in proving

(18)
$$||\tilde{p}_m||_{k, T_1} \leqq \frac{K_2}{(\sin \alpha)^m} ||\tilde{u}||_{k, T_1}.$$

Then

$$|F(\tilde{u})| \leqq \frac{K_3}{(\sin \alpha)^m} ||v||_{n, T_1} ||\tilde{u}||_{k, T_1}$$

and applying the Lemma (actually in our case diam (Ω) = diam (T_1) = $\sqrt{2}$; however obviously (14) is also true with C_2 being an absolute constant) we have

$$|F(\tilde{u})| \leqq \frac{K_4}{(\sin \alpha)^m} ||v||_{n, T_1} |\tilde{u}|_{k, T_1}.$$

Choosing $v = \tilde{u} - \tilde{p}_m$ we get

$$||\tilde{u} - \tilde{p}_m||_{n, T_1} \leqq \frac{K_4}{(\sin \alpha)^m} |\tilde{u}|_{k, T_1}.$$

From (17) it follows

$$||u - p_m||_{n, T} \leqq \frac{K_5}{(\sin \alpha)^{m+n}} c^{-n} |J|^{1/2} |\tilde{u}|_{k, T_1}$$

and since

$$|\tilde{u}|_{k, T_1} \leqq K_6 c^k |J|^{-1/2} |u|_{k, T}$$

the final result is the estimate (15).

To prove (18) we remark that the polynomial $\tilde{p}_m(\xi, \eta)$ is, according to Theorem 1, uniquely determined by the values

$$D^i \tilde{p}_m(\tilde{P}_j), \qquad j = 1, 2, 3, \quad |i| \leqq 2m,$$

$$\partial^r \tilde{p}_m(\tilde{Q}_j^{(\rho, r)})/\partial \tilde{v}_j^r, \qquad j = 1, 2, 3, \quad \rho = 1, \cdots, r, \quad r = 1, \cdots, m,$$

$$D^i(\tilde{P}_0), \qquad |i| \leqq m - 2.$$

If we order these values in some way and denote by a_j $(j = 1, \cdots, N_0 = (2m + 1)$ $\cdot(4m + 3))$, it obviously holds that $\tilde{p}_m(\xi, \eta) = \sum_{i=1}^{N_0} a_j r_i(\xi, \eta)$, where $r_i(\xi, \eta)$ are polynomials such that from the above-mentioned N_0 values one of their values is equal to 1 and the others are zero. Hence, the polynomials $r_i(\xi, \eta)$ as well as their derivatives of an arbitrary order are bounded by absolute constants and it is sufficient to prove that

(19)
$$|a_j| \leqq \frac{K_7}{(\sin \alpha)^m} ||\tilde{u}||_{k, T_1}.$$

Now from (11) and (13) it follows immediately

(20)
$$D^i \tilde{p}_m(\tilde{P}_j) = D^i \tilde{u}(\tilde{P}_j), \qquad j = 1, 2, 3, \quad |i| \leqq 2m,$$

$$D^i \tilde{p}_m(\tilde{P}_0) = D^i \tilde{u}(\tilde{P}_0), \qquad |i| \leqq m - 2,$$

so that for these values we get (19) by means of the Sobolev lemma (actually, (19) is true even without the factor $1/(\sin \alpha)^m$).

To prove (19) for the remaining values we first notice the following formula. Let the direction l make an angle ϕ with the positive ξ-axis and let ν be the direction perpendicular to l. Let τ be a direction making an angle ω with the positive ξ-axis and let $\partial f(P)/\partial \tau = 0$. Then

$$(21) \qquad \partial f(P)/\partial \nu = -\sigma \partial f(P)/\partial l, \qquad \sigma = \cot g(\omega - \phi)$$

(the Eq. (21) follows from the formula $\partial f/\partial \tau = \cos(\omega - \phi)\,\partial f/\partial l + \sin(\omega - \phi)\partial f/\partial \nu$). Further we notice that the condition (12) is just the condition

$$\partial^r \bar{p}_m(\tilde{Q}_i^{(\rho,r)})/\partial \tau_i^r = \partial^r \bar{u}(\tilde{Q}_i^{(\rho,r)})/\partial \tau_i^r,$$

where τ_i $(j = 1, 2, 3)$ are certain directions which are easy to find. Apply now (21) to $f = \bar{p}_m - \bar{u}$ and to $l = \bar{l}_i, \tau = \tau_i$. By elementary calculations, which we leave out, we get

$$(22) \qquad \frac{\partial \bar{p}_m(\tilde{Q}_i^{(1,1)})}{\partial \bar{\nu}_i} = \frac{\partial \bar{u}(\tilde{Q}_i^{(1,1)})}{\partial \bar{\nu}_i} + \sigma_i \frac{\partial \bar{u}(\tilde{Q}_i^{(1,1)})}{\partial \bar{l}_i} - \sigma_i \frac{\partial \bar{p}_m(\tilde{Q}_i^{(1,1)})}{\partial \bar{l}_i},$$

where

$$|\sigma_1| = \frac{|(x_2 - x_1)(x_3 - x_1) + (y_2 - y_1)(y_3 - y_1)|}{c^2}, \qquad |\sigma_2| = \frac{c^2 - b^2}{a^2},$$

$$|\sigma_3| = \frac{|(x_2 - x_1)(x_3 - x_1) + (y_2 - y_1)(y_3 - y_1)|}{b^2}.$$

σ_1 and σ_3 are bounded by absolute constants:

$$|\sigma_1| \leq bc/c^2 \leq 1, \qquad |\sigma_3| \leq bc/b^2 < 2,$$

whereas σ_2 is not bounded by an absolute constant.[6] However,

$$|\sigma_2| = (c - b)(c + b)/a^2 < 2ac/a^2 \leq 2/\sin \alpha.$$

Since $\bar{p}_m$ is a Hermite interpolation polynomial in one variable on the sides of T_1 which is determined by the values (20) it follows from (22) by means of the Sobolev lemma that

$$\left|\frac{\partial \bar{p}_m(\tilde{Q}_i^{(1,1)})}{\partial \bar{\nu}_i}\right| \leq \frac{K_8}{\sin \alpha} ||\bar{u}||_{k,T_1}.$$

In general, we get

$$\left|\frac{\partial^r \bar{p}_m(\tilde{Q}^{(\rho,r)})}{\partial \bar{\nu}_i^r}\right| \leq \frac{K_9}{(\sin \alpha)^r} ||\bar{u}||_{k,T_1}, \qquad r = 1, \cdots, m,$$

if we proceed by induction and use the formula

$$\frac{\partial^r f(P)}{\partial \nu^r} = -\sum_{i=1}^r \binom{r}{j}\sigma^i \frac{\partial^r f(P)}{\partial l^i \, \partial \nu^{r-i}}$$

which holds if $\partial^r f(P)/\partial \tau^r = 0$.

_____

[6] There are triangles for which $(c^2 - b^2)/a^2 \geq 1/2\sin \alpha$.

Remark. In a similar way one can prove

$$\max_{T} |D^i(u - p_m)| \leqq \frac{K}{(\sin \alpha)^{m+|i|}} |J|^{-1/2} c^{k-|i|} |u|_{k, T}$$

if

$$2m + 2 \leqq k \leqq 4m + 2, \qquad |i| \leqq k - 2.$$

3. Application to *V*-Elliptic Boundary Value Problems. Let Ω be a bounded simply or multiply connected domain in E_2 with a boundary Γ consisting of a finite number of polygons Γ_j $(j = 0, 1, \cdots, s)$; $\Gamma_1, \cdots, \Gamma_s$ lie inside of Γ_0 and do not intersect. This assumption enables one to triangulate Ω. Let V be a Hilbert space such that

$$\overset{\circ}{W}_2^{(n)}(\Omega) \subset V \subset W_2^{(n)}(\Omega),$$

with the norm induced by $W_2^{(n)}(\Omega)$. Here, $\overset{\circ}{W}_2^{(n)}(\Omega)$ is the completion with respect to the norm $||\cdot||_n$[7] of functions from $C^{(\infty)}(\Omega)$ with compact support in Ω. Let $a(u, v)$ be a bilinear form continuous on $V \times V$ and *V*-elliptic, i.e., a mapping $(u, v) \rightarrow a(u, v)$ from $V \times V$ into the field of complex numbers which is linear in u, antilinear in v and satisfies the conditions of boundedness and coerciveness

(23) $\qquad |a(u, v)| \leqq M ||u||_n ||v||_n , \quad \forall u, v \in V, \quad M = \text{const} > 0,$

(24) $\qquad \text{Re } a(v, v) \geqq \alpha ||v||_n^2 , \quad \forall v \in V, \quad \alpha = \text{const} > 0.$

Finally, let $L(v)$ be an antilinear functional continuous on V. Under these conditions there exists just one $u \in V$ such that

(25) $\qquad\qquad a(u, v) = L(v), \quad \forall v \in V,$

(see Lions and Magenes [13]).

We shall approximate the problem (25) by the Galerkin method (see Céa [8]) using the following finite-dimensional subspaces V_h^m of V. We triangulate Ω, i.e., we cover $\bar{\Omega}$ by a finite number of arbitrary triangles such that any two triangles are either disjoint or have a common vertex or a common side. To every triangulation we associate two parameters: h, ϑ. h is the largest side and ϑ the smallest angle of all triangles of the given triangulation. In the sequel we assume that as $h \rightarrow 0$, ϑ remains bounded away from zero,

(26) $\qquad\qquad\qquad \vartheta \geqq \vartheta_0 > 0.$

Now V_h^m is the finite-dimensional subspace of V consisting of all functions which on the triangles of the given triangulation are equal to polynomials $p_m(x, y)$ introduced in the preceding section. Every function from V_h^m belongs to $C^{(m)}(\bar{\Omega})$ and, at the same time, to $W_2^{(m+1)}(\Omega)$.

Let us consider the problem of finding u_h^m such that

(27) $\qquad\qquad a(u_h^m, v) = L(v), \quad \forall v \in V_h^m.$

THEOREM 3. *Let* $n \leqq m + 1$. *Under the assumptions* (23), (24) *and* (26) *there exists*

---

[7] In this and the last section we write $||\cdot||_n$ instead of $||\cdot||_{n,\Omega}$ and $|\cdot|_n$ instead of $|\cdot|_{n,\Omega}$.

just one $u_h^m \in V_h^m$ satisfying (27) *and*

(28)
$$||u - u_h^m||_n \to 0 \quad as \quad h \to 0.$$

Proof. It is an immediate consequence of the theorem by Céa about the Galerkin method (see [8, p. 363, Théorème 3.1]) and of Theorem 2 proved in Section 2. We must show that the subspaces V_h^m have the following property of density: there is a subspace $\mathcal{V} \subset V$ which is dense in V and a family of linear operators r_h^m from $\mathcal{V}$ into V_h^m such that

(29)
$$||v - r_h^m v||_n \to 0, \quad \forall v \in \mathcal{V} \quad as \quad h \to 0.$$

For $\mathcal{V}$ we choose functions from V belonging to $W_2^{(k)}(\Omega)$ with $2m + 2 \leq k \leq 4m + 2$. As $k \geq 2m + 2$ it follows by Sobolev's lemma that $\mathcal{V} \subset C^{(2m)}(\bar{\Omega})$. $r_h^m v$ is then the function which on every triangle of the corresponding triangulation is equal to the interpolation polynomial $p_m(x, y)$ corresponding to $v(x, y)$. According to (15) and (26) we have

$$||v - r_h^m v||_{n,T}^2 \leq K_{10} h^{2(k-n)} |v|_{k,T}^2.$$

Hence,

(30)
$$||v - r_h^m v||_n \leq K_{11} |v|_k h^{k-n}$$

and (29) follows.

Theorem 3 proves only the convergence of the finite element method. Of course, we did not ask more than that the solution u of the boundary value problem (25) of the 2nth order belongs to $W_2^{(n)}(\Omega)$. If we suppose more about the smoothness of u we get an asymptotic estimate of the rate of convergence:

THEOREM 4. *Suppose that the form $a(u, v)$ is Hermitian. Let the assumptions of Theorem 3 hold and let*

$$u(x, y) \in W_2^{(k)}(\Omega), \quad 2m + 2 \leq k \leq 4m + 2.$$

Then

(31)
$$||u - u_h^m||_n \leq K |u|_k h^{k-n},$$

where the constant K does not depend on the triangulation and on the solution u.

Proof. We use a lemma by Céa [8, p. 365, Proposition 3.1]. According to the inequality 3.14 of this lemma

$$||u - u_h^m||_n \leq (M/\alpha)^{1/2} ||u - r_h^m u||_n$$

holds. As $u \in W_2^{(k)}(\Omega)$ we can set $v = u$ in (30) and the proof is finished.

In case $n = 2$, $m = 1$, (31) gives

$$||u - u_h^1||_2 \leq K |u|_k h^{k-2}, \quad 4 \leq k \leq 6,$$

for $u \in W_2^{(k)}(\Omega)$. The highest order of accuracy is attained for $k = 6$,

$$||u - u_h^1||_2 \leq K |u|_6 h^4.$$

This result is a generalization of the result of [19] where instead of $|u|_6$ the seminorm $M_6 = \sup_\Omega |D^i u|$, $|i| = 6$, is used. In the same way we get for $n = 3$, $m = 2$ and $n = 4$, $m = 3$ the generalization of the results of [17].

4. Some Special Cases. 1. To get the asymptotic estimate (31) we had to assume a greater smoothness of the solution $u(x, y)$ than that guaranteed by the conditions (23) and (24) which, on the other hand, are sufficient for the uniqueness and existence of $u(x, y)$. In one case we need not impose any additional condition on the smoothness of the solution and yet we obtain an asymptotic error estimate, even in terms of data only. Consider, namely, the Dirichlet problem

$$(32) \qquad Lu \equiv - \sum_{i,k=1}^{2} \frac{\partial}{\partial x_i} \left(a_{ik} \frac{\partial u}{\partial x_k} \right) + au = f$$

on a convex polygon Ω. Let us assume that

$$(33) \qquad a_{ik}(x, y) \in C^{(0,1)}(\Omega), \qquad a(x, y), f(x, y) \in L_2(\Omega) \qquad (j, k = 1, 2),$$

that the operator Lu is uniformly elliptic,

$$\sum_{i,k=1}^{2} a_{ik}(x, y)\xi_i \xi_k \geqq \alpha_0 \sum_{j=1}^{2} \xi_j^2, \qquad \alpha_0 > \text{const} > 0,$$

and that $a(x, y) \geqq 0$. Then the form $a(u, v)$ corresponding to the above Dirichlet problem,

$$a(u, v) = \int_{\Omega} \left[\sum_{i,k=1}^{2} a_{ik} \frac{\partial u}{\partial x_k} \frac{\partial v}{\partial x_i} + auv \right] dx \, dy,$$

is $\mathring{W}_2^{(1)}(\Omega)$-elliptic. According to a theorem of Kadlec [11] the solution $u(x, y)$ belongs to $W_2^{(2)}(\Omega)$ and

$$(34) \qquad ||u||_2 \leqq C \, ||f||_{L_2(\Omega)},$$

where the constant C depends only on the coefficients of the operator Lu and on the domain Ω. Actually, the result is stated in [11] for the equation

$$L_0 u \equiv - \sum_{i,k=1}^{2} \frac{\partial}{\partial x_i} \left(a_{ik} \frac{\partial u}{\partial x_k} \right) = f.$$

However, if we write (32) in the form $L_0 u = -au + f$ we see that the right-hand side belongs to $L_2(\Omega)$. By the theorem of Kadlec

$$||u||_2 \leqq C \, ||-au + f||_{L_2(\Omega)} \leqq C(K_{12} \, ||u||_1 + ||f||_{L_2(\Omega)}).$$

As $a(u, u) = (Lu, u)_L$ for $u \in W_2^{(2)}(\Omega) \cap \mathring{W}_2^{(1)}(\Omega)$ it follows from the $\mathring{W}_2^{(1)}(\Omega)$-ellipticity of $a(u, v)$ that $||u||_1 \leqq (1/\alpha)||f||_{L_2(\Omega)}$. Hence (34) is true. Now the assumptions of Theorem 4 are satisfied ($n = 1$, $m = 0$, $k = 2$) and we have the following.

THEOREM 5. *Let Ω be a convex polygon and suppose that the real coefficients and the right-hand side of Eq. (32) satisfy (33). Further, let Lu be uniformly elliptic and let $a(x, y) \geqq 0$. Then*

$$(35) \qquad ||u - u_h^0||_1 \leqq C \, ||f||_{L_2(\Omega)} h,$$

where the constant C depends on the coefficients of Lu and on the domain Ω only.

Using an argument similar to that of Nitsche [14] we can obtain the additional result

$$(36) \qquad ||u - u_h^0||_{L_2(\Omega)} \leqq Ch^2 \, ||f||_{L_2(\Omega)}.$$

Remark. The estimate of the form (36) is also given in the paper by L. A. Oganesjan, P. A. Ruchovec: "Investigation of the convergence rate of variational-difference schemes for elliptic second order equations in a two-dimensional domain with a smooth boundary," *Ž. Vyčisl. Mat. i Mat. Fiz.*, v. 9, 1969, 1102–1120. (Russian)

The proof is as follows: Write

$$||u - u_h^0||_0 = \sup_{\psi \in L_2(\Omega)} \frac{|(u - u_h^0, \psi)_0|}{||\psi||_0}.$$

Now let ϕ satisfy

$$a(v, \phi) = (v, \psi)_0, \qquad \forall v \in \mathring{W}_2^{(1)}(\Omega).$$

Then, as in (34), we have $||\phi||_2 \leq C||\psi||_0$. Hence,

(37) $$||u - u_h^0||_0 \leq C \sup |a(u - u_h^0, \phi)|/||\phi||_2.$$

But

$$a(u - u_h^0, \phi) = a(u - u_h^0, \phi - \bar\phi), \qquad \forall \bar\phi \in V_h^0.$$

Hence,

(38) $$|a(u - u_h^0, \phi)| \leq M ||u - u_h^0||_1 ||\phi - \bar\phi||_1.$$

Choose $\bar\phi$ such that

(39) $$||\phi - \bar\phi||_1 \leq K_{13} ||\phi||_2 h.$$

Then (37), (38) and (39) imply

$$||u - u_h^0||_0 \leq K_{14}h ||u - u_h^0||_1.$$

This together with Theorem 5 yields the result.

2. In [19] there was also introduced a cubic polynomial $p(x, y)$ determined by ten values

$$p(P_i), \quad \partial p(P_i)/\partial x, \quad \partial p(P_i)/\partial y, \quad p(P_0), \qquad j = 1, 2, 3.$$

This element can be used for solving second-order boundary value problems. It is easy to show, in the same way as Theorem 2 was proved, that

$$||u - p||_{n, T} \leq \frac{K}{(\sin \alpha)^n} c^{k-n} |u|_{k, T}, \qquad k = 3, 4, \quad n \leq k,$$

if $u \in W_2^{(k)}(T)$. For the corresponding finite element procedure (again under the assumptions (23), (24) and (26)) it follows first that it converges in the $||\cdot||_1$ norm, and secondly that

$$||u - u_h||_1 \leq K |u|_k h^{k-1}, \qquad k = 3, 4,$$

if $u \in W_2^{(k)}(\Omega)$. For $k = 4$ this result is a generalization of the estimate (13) in [19].

3. The polynomial $p_1(x, y)$ is a 21-degree-of-freedom element. However, the values $\partial p_1(Q_j^{(1,1)})/\partial \nu_j$ ($j = 1, 2, 3$) are not necessary in applications. Bell proposed in [3] (also Goël in [10]) an 18-degree-of-freedom element and applied it to bending of thin plates. We get it from $p_1(x, y)$ if we eliminate the three above mentioned values by imposing on $p_1(x, y)$ the condition that $\partial p_1/\partial \nu_j$ ($j = 1, 2, 3$) be cubic

polynomials on the corresponding sides of T. In general, $\partial p_1/\partial \nu_i$ is a polynomial of the fourth degree in one variable on the side l_i of T and it is easy to see that the above condition determines uniquely the values $\partial p_1(Q_i^{(1,1)})/\partial \nu_i$ as linear combinations of the remaining 18 values

$$D^i p_1(P_j), \qquad j = 1, 2, 3, \quad |i| \leq 2.$$

We denote this 18-degree-of-freedom element by $q(x, y)$. If we inspect the proof of Theorem 2, we easily find out that an estimate corresponding to (18) is again true in case of the element $q(x, y)$ and that the only change is that the functional F vanishes for polynomials of the degree less than 5, whereas, in case of $p_1(x, y)$ it vanishes for polynomials of the degree less than 6. We have

$$\|u - q\|_{n,T} \leq \frac{K}{(\sin \alpha)^n} c^{k-n} |u|_{k,T}, \qquad n = 1, 2, \quad k = 4, 5,$$

if $u \in W_2^{(k)}(T)$. For the corresponding finite element procedure (again under the assumptions (23), (24) and (26)), it follows first that it converges in the norm $\|\cdot\|_1$ and $\|\cdot\|_2$, respectively, and secondly that

$$\|u - u_h\|_n \leq K |u|_k h^{k-n}, \qquad n = 1, 2, \quad k = 4, 5,$$

if $u \in W_2^{(k)}(\Omega)$. Thus, for bending of thin plates the highest order of accuracy is the third order.

Similarly one can generalize the results of [20] where, by eliminating the value $p(P_0)$ from the cubic element $p(x, y)$, there was constructed a 9-degree-of-freedom element.

4. For practical applications it is desirable (see [20, p. 395]) that as many parameters determining the polynomials as possible are prescribed at the vertices only. In [12] it is remarked that in the case of polynomials of degree $4m + 1$ and $4m + 3$ (see footnore 2) the parameters prescribed on the sides of the triangle can be eliminated by imposing on the polynomials the condition that the normal derivatives of the kth order be polynomials of degree $n - 2k$ along the sides of the triangle. For the corresponding finite element procedure one can easily prove that

$$\|u - {}_i u_h\|_n \leq K |u|_k h^{k-n}$$

for $2m + 2 \leq k \leq 3m + 2$ and $2m + 3 \leq k \leq 3m + 4$, respectively, if $n \leq m + 1$ and $u \in W_2^{(k)}(\Omega)$.

It is also possible to eliminate the parameters prescribed at the center of gravity by imposing some restrictions on the polynomials. However, in this case a better practical way is to retain them and to use the method of condensation of internal parameters (see [21] or [22]).

Department of Mathematics
Cornell University
Ithaca, New York 14850

Technical University
Brno
Czechoslovakia

1. S. AGMON, *Lectures on Elliptic Boundary Value Problems,* Van Nostrand Math. Studies, no. 2, Van Nostrand, Princeton, N. J., 1965. MR 31 #2504.

2. J. H. ARGYRIS, I. FRIED & D. W. SCHARPF, "The tuba family of plate elements for the matrix displacement method," *Aeronautical J. Roy. Aeronautical Soc.,* v. 72, 1968, pp. 618–623.

3. K. BELL, *Analysis of Thin Plates in Bending Using Triangular Finite Elements,* The Technical University of Norway, Trondheim, 1968.

4. K. BELL, "A refined triangular plate bending finite element," *Internat. J. Numer. Methods in Engrg.,* v. 1, 1969, pp. 101–122.

5. I. S. BEREZIN & N. P. ŽIDKOV, *Computing Methods.* Vol. I, 2nd ed., Fizmatgiz, Moscow, 1962; English transl., of 1st ed., Pergamon Press, New York, 1965. MR 30 #4372; MR 31 #1756.

6. W. BOSSHARD, "Ein neues, vollverträgliches endliches Element für Plattenbiegung," *Abh. Int. Verein. Brückenbau and Hochbau, Zürich,* v. 28, 1968, pp. 27–40.

7. J. H. BRAMBLE & S. R. HILBERT, "Estimation of linear functionals on Sobolev spaces with application to Fourier transforms and spline interpolation," *Siam. J. Numer. Anal.,* v. 7, 1970, pp. 113–124.

8. J. CÉA, "Approximation variationnelle des problèmes aux limites," *Ann. Inst. Fourier (Grenoble),* v. 14, 1964, pp. 345–444, MR 30 #5037.

9. R. COURANT, "Variational methods for the solution of problems of equilibrium and vibrations," *Bull. Amer. Math. Soc.,* v. 49, 1943, pp. 1–23. MR 4, 200.

10. J. J. GOËL, *List of Basic Functions for Numerical Utilisation of Ritz's Method. Application to the Problem of the Plate,* École Polytechnique Féderale, Lausanne, 1969.

11. J. KADLEC, "The regularity of the solution of the Poisson problem in a domain whose boundary is similar to that of a convex domain, *Czechoslovak Math. J.,* v. 14(89), 1964, pp. 386–393. (Russian) MR 30 #329.

12. J. KRATOCHVÍL, A. ŽENÍŠEK & M. ZLÁMAL, "A simple algorithm for the stiffness matrix of triangular plate bending finite elements," *Numer. Methods in Engineering.* (To appear.)

13. J. L. LIONS & E. MAGENES, *Problèmes aux Limites non Homogènes et Applications.* Vol. 1, Travaux Recherches Math., no. 17, Dunod, Paris, 1968.

14. J. NITSCHE, "Lineare Spline-Funktionen und die Methode von Ritz für elliptische Randwertprobleme." (To appear.)

15. M. J. TURNER, R. W. CLOUGH, H. C. MARTIN & L. J. TOPP, "Stiffness and deflection analysis of complex structures," *J. Aeronautical Sci.,* v. 23, 1956, pp. 805–823.

16. M. VISSER, *The Finite Element Method in Deformation and Heat Conduction Problems,* Delft, 1968.

17. A. ŽENÍŠEK, "Interpolation polynomials on the triangle," *Numer. Math.,* v. 15, 1970, pp. 283–296.

18. O. C. ZIENKIEWICZ, *The Finite Element Method in Structural and Continuum Mechanics,* McGraw-Hill, New York, 1967.

19. M. ZLÁMAL, "On the finite element method," *Numer. Math.,* v. 12, 1968, pp. 394–409. MR 39 #5074.

20. M. ZLÁMAL, "A finite element procedure of the second order of accuracy," *Numer. Math.,* v. 16, 1970, pp. 394–402.

21. A. C. FELIPPA, *Refined Finite Element Analysis of Linear and Nonlinear Two-Dimensional Structures,* SESM Report No. 66–22, University of California, Berkeley, Calif., 1967.

22. E. ANDERHEGGEN, *Programme zur Methode der finiten Elemente,* Institut für Baustatik, Eidgenössische Technische Hochschule, Zürich, 1969.

23. G. R. COWPER, E. KOSKO, G. M. LINDBERG & M. D. OLSON, "Formulation of a new triangular plate bending element," *C.A.S.I. Trans.,* v. 1, 1968, pp. 86–90.

24. G. R. COWPER, E. KOSKO, G. M. LINDBERG & M. D. OLSON, "Static and dynamic applications of a high-precision triangular plate bending element," *AIAA J.,* v. 7, 1969, pp. 1957–1965.