MATHEMATICS OF COMPUTATION, VOLUME 32, NUMBER 143
JULY 1978, PAGES 663–685

Superconvergence and Reduced Integration in the Finite Element Method

By Miloš Zlámal

Abstract. The finite elements considered in this paper are those of the Serendipity family of curved isoparametric elements. There is given a detailed analysis of a superconvergence phenomenon for the gradient of approximate solutions to second order elliptic boundary value problems. An approach is proposed how to use the superconvergence in practical computations.

1. Introduction. Among finite elements the curved isoparametric elements of the Serendipity family (see Zienkiewicz [8]) are mostly used in the finite element codes prepared for engineering computations. It has been observed (see, e.g., Veryard [7], Irons and Razzaque [5], Barlow [1]) that applying quadratic members of this family a considerable improvement in accuracy of stresses is achieved if a reduced numerical integration—Gauss' 2×2 or $2 \times 2 \times 2$ product formulas—is used and the stresses are computed at Gaussian points, i.e. at points of these formulas. Here we want to analyze and justify this phenomenon. The results proved in the paper constitute a substantial extension of earlier results of the author [9].

We consider first the Dirichlet problem in two dimensions for a selfadjoint second order elliptic equation with variable coefficients as a model problem. We assume that the finite element partitions of the given domain are 2-strongly regular (see definition in the next section). In Section 4 we prove superconvergence of the gradient of the approximate solution at Gaussian points if Gauss' 2×2 formula for the two-dimensional cube C_2: $-1 \leqslant \xi_i \leqslant 1$, $i = 1, 2$, is applied. Numerical results (Section 6) indicate convincingly that superconvergence does not set in if the condition (2.8) about finite elements is not satisfied. Under a further assumption on finite elements the superconvergence is proved if there is applied any symmetric formula of the type (2.16) with positive coefficients which integrates exactly all polynomials from $\hat{Q}(3)$ on C_2 or any formula (2.16) which integrates exactly all polynomials from $\hat{P}(4)$ on C_2 ($\hat{P}(k)$ and $\hat{Q}(k)$ denote the classes of polynomials of degree k and of degree k in each variable, respectively). This result shows that the superconvergence phenomenon is not closely connected with the reduced integration. However, Gauss' 2×2 formula has the smallest number of points among the above-mentioned formulas.

The theorem on superconvergence is true in three dimensions under the condition that the partitions are 3-strongly regular. In the last section there are introduced numerical results and an approach is proposed how to use the superconvergence in practical computations.

Received March 21, 1977; revised November 22, 1977.
AMS (MOS) subject classifications (1970). Primary 65N30, 65N15; Secondary 35A15.
Key words and phrases. Finite elements.

2. Preliminaries. Let Ω be a bounded domain with a sufficiently smooth boundary Γ. We consider the Dirichlet problem

$$(2.1)\qquad \begin{aligned} &Lu = f(x) \quad \forall x \in \Omega, \quad u|_\Gamma = 0,\\ &Lu = -\sum_{i,j=1}^{2} \frac{\partial}{\partial x_i}\left[a_{ij}(x)\frac{\partial u}{\partial x_j}\right]; \end{aligned}$$

here $x = (x_1, x_2)$. Let us remark at this point that we could add a term $a_0 u$ with $a_0 \geqslant 0$ in the definition (2.1) of the operator Lu. All that follows applies equally well to this case, with a straightforward supplementary analysis. To (2.1) there is associated the bilinear functional

$$(2.2)\qquad a(u, v) = \int_\Omega \sum_{i,j=1}^{2} a_{ij} \frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_j}\, dx.$$

We assume that the coefficients are defined on $\bar{\Omega}$ and that

$$(2.3)\quad a_{ij}(x) = a_{ji}(x), \quad \sum_{i,j=1}^{2} a_{ij}(x)\xi_i\xi_j \geqslant c_1 \sum_{i=1}^{2} \xi_i^2 \quad \forall x \in \Omega, \ c_1 = \text{const} > 0.$$

Hence $a(u, v)$ is $H_0^1(\Omega)$-elliptic.

The weak solution of the problem (2.1) is a function $u \in H_0^1(\Omega)$ which satisfies

$$(2.4)\qquad a(u, v) = (f, v)_{0,\Omega} \quad \forall v \in H_0^1(\Omega).$$

We are using the usual notation for the Sobolev spaces:

$$H^m(\Omega) = \{u \in L^2(\Omega), D^\alpha u \in L^2(\Omega) \ \forall |\alpha| \leqslant m\}, \qquad m = 0, 1, \dots,$$

$$H_0^1(\Omega) = \{u \in H^1(\Omega), u|_\Gamma = 0\}.$$

The norm in $H^m(\Omega)$ is denoted by $\|\cdot\|_{m,\Omega}$ and defined by

$$\|u\|_{m,\Omega} = \left\{\sum_{|\alpha|\leqslant m} \|D^\alpha u\|^2_{L^2(\Omega)}\right\}^{1/2},$$

the inner product in $H^m(\Omega)$ is denoted by $(\cdot, \cdot)_{m,\Omega}$. Often we shall use the seminorm

$$|u|_{m,\Omega} = \left\{\sum_{|\alpha|=m} \|D^\alpha u\|^2_{L^2(\Omega)}\right\}^{1/2}$$

(we set $|u|_{0,\Omega} = \|u\|_{0,\Omega}$).

To construct the finite element space V_h in which the approximate solution will lie let us "cover" Ω by in general curved quadrilateral quadratic elements of the Serendipity family. Denote by $\hat{P}$ the class of incomplete cubic polynomials of the form

$$(2.5)\qquad \alpha_1 + \alpha_2\xi_1 + \alpha_3\xi_2 + \alpha_4\xi_1^2 + \alpha_5\xi_1\xi_2 + \alpha_6\xi_2^2 + \alpha_7\xi_1^2\xi_2 + \alpha_8\xi_1\xi_2^2.$$

Evidently,

$$(2.6)\qquad \hat{P}(2) \subset \hat{P} \subset \hat{Q}(2).$$

Let $N_j(\xi_1, \xi_2)$ $(j = 1, \dots, 8)$ be polynomials introduced in Zienkiewicz [8, p. 109]. Then $\Sigma_{j=1}^8 v_j N_j(\xi_1, \xi_2)$ is the only polynomial from $\hat{P}$ assuming the given values v_j

at the nodes of the two-dimensional cube C_2: $-1 \leq \xi_1 \leq 1, -1 \leq \xi_2 \leq 1$, i.e. at the vertices and at the midpoints of the sides of C_2. This polynomial is a quadratic polynomial in each variable determined on every side of C_2 uniquely by its values at nodes of this side.

Now consider eight points (nodes) a_j $(j = 1, \ldots, 8)$ with coordinates (x_1^j, x_2^j) and the mapping

$$(2.7)\quad x_1 = x_1^e(\xi_1, \xi_2) \equiv \sum_{j=1}^{8} x_1^j N_j(\xi_1, \xi_2), \qquad x_2 = x_2^e(\xi_1, \xi_2) \equiv \sum_{j=1}^{8} x_2^j N_j(\xi_1, \xi_2).$$

If (2.7) maps the cube C_2 one-to-one on a closed domain e lying in the (x_1, x_2)-plane, we call e a quadratic quadrilateral element (curved or straight which depends on the choice of the nodes a_j).

We "cover" Ω by such elements, and we suppose that every "partition" of Ω by these elements is a 2-strongly regular partition. By a k-strongly regular partition we understand a partition with the following properties:

(a) for every element the mapping (2.7) is a C^{k+1} diffeomorphism (in particular, (2.7) is invertible).

(b) to every element e there is associated a positive parameter h_e, and the mapping (2.7) is such that on e

$$(2.8)\qquad |D^\alpha x_i^e| \leq C_1 h_e^{|\alpha|}, \qquad |\alpha| \leq k + 1, i = 1, 2,$$

$$(2.9)\qquad c_2^{-1} h_e^2 \leq |J_e| \leq c_2 h_e^2;$$

here $J_e(\xi_1, \xi_2)$ is the Jacobian of (2.7) and C_1, c_2 are positive constants independent of h_e as well as of the chosen partition. If h is defined by

$$h = \max_e h_e,$$

then the constants C_1, c_2 are independent of h, too.

We will consider a family of 2-strongly regular partitions of Ω such that $h \longrightarrow 0$. We denote by Ω_h the interior of the union of all elements of the given partition (in general $\Omega_h \neq \Omega$); Γ_h is its boundary.

Remark 1. The definition of a k-strongly regular partition is similar to the definition of a k-regular family of elements by Ciarlet and Raviart [4]. The main difference is that, instead of their requirement (2.17′) (p. 427), we ask (2.8). This is evidently a much stronger condition, and every domain Ω cannot be covered by such elements. However, numerical results (see Section 6) indicate convincingly that (2.8) with $k = 2$ is a necessary condition for superconvergence introduced later. In the following Ω is supposed to be such that there exists a family of 2-strongly regular partitions with $h \longrightarrow 0$.

Remark 2. The following simple condition is sufficient for a partition to satisfy (2.8) and (2.9) for h sufficiently small: to each element e of the partition there exists a parallelogram e' with sides h_e and k_e, $h_e \geq k_e$ (i.e., we denote the larger side by h_e), with angle ω_e and with nodes a_i' (the nodes corresponding to the midpoints of sides of C_2 must be midpoints of the sides of e') such that

(2.10) $$\frac{k_e}{h_e} \geqslant c_3 > 0, \qquad 0 < \omega_0 \leqslant \omega_e \leqslant \pi - \omega_0,$$

(2.11) $$\rho(a_i, a_i') \leqslant C_2 h_e^{k+1}, \qquad 1 \leqslant i \leqslant 8,$$

where $\rho(a_i, a_i')$ is the distance of a_i and a_i' and ω_0, c_3, C_2 are positive constants independent of h_e, and the given partition, i.e. independent of h, too. To prove it, write $x_i^e = \Sigma_{j=1}^8 x_i'^j N_j + \Sigma_{j=1}^8 (x_i^j - x_i'^j) N_j$ ($x_1'^j, x_2'^j$ are coordinates of a_j'). The mapping $x_i = \Sigma_{j=1}^8 x_i'^j N_j(\xi_1, \xi_2)$ $(i = 1, 2)$ is a mapping which maps C_2 on the parallelogram e' and midpoints on the midpoints of sides. Therefore, it is bilinear and we easily compute that $\partial x_i^{e'}/\partial \xi_j$ are constant and bounded by $|\partial x_i^{e'}/\partial \xi_j| \leqslant \frac{1}{2} h_e$ and $|J_{e'}| = \frac{1}{4} h_e k_e \sin \omega_e$. Hence, $|D^\alpha x_i^{e'}| \leqslant \frac{1}{2} h_e$ if $|\alpha| = 1$, $D^\alpha x_i^{e'} = 0$, if $|\alpha| \geqslant 2$ and $(c_3/4) \sin \omega_0 h_e^2 \leqslant |J_{e'}| \leqslant \frac{1}{4} h_e^2$. From (2.11) it easily follows that (2.8) and (2.9) are true for h_e sufficiently small.

Let us remark that the condition (b) is not as strong as (2.10) and (2.11) which effectively eliminate curved edges. E.g., consider a closed domain $\bar{\Omega}$ which is a map of a closed rectangle $\bar{R}$ and the corresponding mapping $x_i = \varphi_i(s_1, s_2)$, $i = 1, 2$, is such that $\varphi_i \in C^3(\bar{R})$ and $\partial(\varphi_1, \varphi_2)/\partial(s_1, s_2) \neq 0$ on $\bar{R}$. We construct a mesh on $\bar{\Omega}$ in the following simple way: Its nodes are maps of nodes of a rectangular mesh of $\bar{R}$. Consider a rectangular element of $\bar{R}$ and denote by h_e, k_e, the lengths of its sides, h_e being always the length of the larger one, and by s_1^0, s_2^0, the coordinates of its center. Let e be the element of $\bar{\Omega}$ which corresponds to this rectangular element. Then one can easily express the functions x_i^e from (2.7) and their Jacobian as follows (we may assume that R has sides parallel to coordinate axes):

$$x_i^e \equiv \sum_{j=1}^{8} x_i^j N_j(\xi_1, \xi_2) = \varphi_i(s_1^0, s_2^0) + \frac{1}{2} h_e \frac{\partial \varphi_i(s_1^0, s_2^0)}{\partial s_1} \xi_1 + \frac{1}{2} k_e \frac{\partial \varphi_i(s_1^0, s_2^0)}{\partial s_2} \xi_2$$

$$+ \frac{1}{8} h_e^2 \frac{\partial^2 \varphi_i(s_1^0, s_2^0)}{\partial s_1^2} \xi_1^2 + \frac{1}{4} h_e k_e \frac{\partial^2 \varphi_i(s_1^0, s_2^0)}{\partial s_1 \partial s_2} \xi_1 \xi_2 + \frac{1}{8} k_e^2 \frac{\partial^2 \varphi_i(s_1^0, s_2^0)}{\partial s_2^2} \xi_2^2$$

$$+ r_i(\xi_1, \xi_2), \qquad D^\alpha r_i = O(h_e^3) \quad \text{for } |\alpha| \geqslant 0.$$

$$J_e(\xi_1, \xi_2) = \frac{1}{4} h_e k_e \left. \frac{\partial(\varphi_1, \varphi_2)}{\partial(s_1, s_2)} \right|_{s_1 = s_1^0, s_2 = s_2^0} + O(h_e^3).$$

Let us now assume that the rectangular mesh of $\bar{R}$ is chosen in such a way that $k_e/h_e \geqslant c_3 > 0$ where c_3 is a positive constant independent of h_e and the given mesh. Then the condition (b) is evidently satisfied for $k = 2$. An example of the mapping $x_i = \varphi_i(s_1, s_2)$: polar coordinates.

Remark 3. The sign of J_e changes if the local ordering of nodes is taken in the opposite direction. Therefore, we may and we will assume that for every e

(2.12) $$J_e(\xi_1, \xi_2) > 0 \quad \forall \xi \in C_2.$$

The functions v from the finite element space V_h are defined piecewise:

(2.13) $$v(x_1, x_2) = \hat{v}[\xi_1^e(x_1, x_2), \xi_2^e(x_1, x_2)], \qquad \hat{v}(\xi_1, \xi_2) = \sum_{j=1}^{8} v_j N_j(\xi_1, \xi_2).$$

Here $\xi_i = \xi_i^e(x_1, x_2)$ is the inverse mapping to (2.7), and v_j are values of v at nodes of the element e. For the complete definition of V_h it remains to ask $v|_{\Gamma_h} = 0$ which is equivalent to the requirement that the values of v at nodes lying on Γ are equal to zero. Evidently,

$$V_h \subset C(\bar{\Omega}_h), \qquad V_h \subset H_0^1(\Omega_h). \tag{2.14}$$

To define the approximate solution of the problem (2.4) we proceed in a similar way as in [4]. We extend the solution $u \in H^4(\Omega)$ and the coefficients $a_{ij} \in H^3(\Omega)$ according to Calderon's extension theorem (see Nečas [6, p. 80]) to R^2 and denote these extensions by $\tilde{u}$ and $\tilde{a}_{ij}$, respectively. We also extend f as follows:

$$\tilde{f} = -\sum_{i,j=1}^{2} \frac{\partial}{\partial x_i}\left(\tilde{a}_{ij}\frac{\partial \tilde{u}}{\partial x_j}\right) \in H^2(\Omega).$$

Denote by $\tilde{a}(w, v)$ the bilinear functional $\int_{\Omega_h} \Sigma_{i,j=1}^2 \tilde{a}_{ij}(\partial w/\partial x_i)(\partial v/\partial x_j)dx$. Due to $v|_{\Gamma_h} = 0$ we get for any $v \in V_h$ by Green's theorem $\tilde{a}(\tilde{u}, v) = (\tilde{f}, v)_{0,\Omega_h}$. For simplicity of writing we will leave out the sign $\sim$ and write

$$(2.15) \qquad \begin{aligned} &a(u, v) = (f, v)_{0,\Omega_h} \quad \forall v \in V_h, \\ &a(u, v) = \int_{\Omega_h} \sum_{i,j=1}^{2} a_{ij}\frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_j}\, dx. \end{aligned}$$

This will not cause any confusion in the estimates carried out later. All constants will depend on $\|\tilde{u}\|_{4,\Omega_h}$ and $\|\tilde{f}\|_{2,\Omega_h}$. The first norm is bounded, according to Calderon's theorem, by $\|u\|_{4,\Omega}$. Evidently, also $\|\tilde{f}\|_{2,\Omega_h}$ is bounded by this norm. By (2.3) the matrix $A = \{a_{ij}\}_{i,j=1}^2$ is uniformly positive definite for $x \in \Omega_h$ and h sufficiently small if the extensions of the coefficients are continuous. Hence, under this condition

$$\sum_{i,j=1}^{2} a_{ij}(x)\xi_i\xi_j \geqslant c_1 \sum_{i=1}^{2} \xi_i^2 \quad \forall x \in \Omega_h, \tag{2.3'}$$

where c_1 is a positive constant independent on h.

We could define the approximate solution u_h as that function from V_h which satisfies $a(u_h, v) = (f, v)_{0,\Omega_h}$ $v \in V_h$. However, in general the values of $a(w, v)$ and $(f, v)_{0,\Omega_h}$ for $v, w \in V_h$ cannot be computed exactly. Numerical integration is the usual and only practical way out. To this end let us consider quadrature formulas $I(\varphi)$ for the cube C_2 of the form

$$I(\varphi) = \sum_r A_r \varphi(Q_r). \tag{2.16}$$

We make the assumption that the points Q_r of the formula belong to the interior of C_2 or are nodes of C_2. Then expressing $a(w, v)$ as a sum of integrals over the elements e, transforming these integrals by means of (2.7) in integrals over C_2 and using (2.16), we get the approximate value $a_h(w, v)$ of $a(w, v)$:

$$a_h(w, v) = \sum_e \sum_r A_r J_e(Q_r) \sum_{i,j=1}^{2} \hat{a}_{ij}(Q_r)\widehat{\frac{\partial w}{\partial x_i}}(Q_r)\widehat{\frac{\partial v}{\partial x_j}}(Q_r). \tag{2.17}$$

Here the following notation (in agreement with the notation in (2.13)) is used for any function g defined on $\bar{\Omega}_h$:

$$\hat{g}(\xi_1, \xi_2) = g[x_1^e(\xi_1, \xi_2), x_2^e(\xi_1, \xi_2)].$$

Similarly,

$$f_h(v) = \sum_e \sum_r A_r J_e(Q_r) \hat{f}(Q_r) \hat{v}(Q_r) \tag{2.18}$$

is the approximate value of $(f, v)_{0,\Omega_h}$. Our assumption concerning the points Q_r guarantees that, at least for h sufficiently small, we do not need for the computation of $a_h(w, v)$ and $f_h(v)$ values of data at other points than at points from $\bar{\Omega}$. Now the approximate solution $u_h \in V_h$ is defined by

$$a_h(u_h, v) = f_h(v) \quad \forall v \in V_h. \tag{2.19}$$

It is clear from the remark made above that u_h does not depend on extensions of the coefficients a_{ij} and the right-hand side f of the equation (2.1). In general, it is not true that u_h exists and is unique. We will consider the cases that $I(\varphi)$ is Gauss' product formula 2×2 or any symmetric formula with positive coefficients which integrates exactly all polynomials from $\hat{Q}(3)$ (Gauss' 2×2 is a special case of such formulas having the smallest number of points) or any formula which integrates exactly all polynomials from $\hat{P}(4)$. The existence and uniqueness of u_h will follow from Lemma 3.6.

3. Some Lemmas. In what follows we denote by C a generic positive constant not necessarily the same in any two places which does not depend on h_e, h and on some functions. It will be clear from the context of which functions the constant is independent.

LEMMA 3.1. *We have for any* $\hat{v} \in \hat{P}$

$$|\hat{v}|_{j,C_2} \leq C|\hat{v}|_{i,C_2}, \qquad 0 \leq i < j \leq 3 \ (|\hat{v}|_{0,C_2} \equiv \|\hat{v}\|_{0,C_2}), \tag{3.1}$$

$$\max_{C_2} |D^\alpha \hat{v}| \leq C|\hat{v}|_{|\alpha|,C_2}, \qquad |\alpha| \leq 3. \tag{3.2}$$

Proof. To prove (3.1) for $j = 1$ it is sufficient to realize that $\|\hat{v}\|_{0,C_2}^2$ is a positive definite quadratic form of the coefficients α_j ($j = 1, \ldots, 8$) and $|\hat{v}|_{1,C_2}^2$ is a bounded quadratic form of these coefficients. Applying (3.1) with $j = 1$ to partial derivatives of $\hat{v}$ we get (3.1) for $j = 2, 3$. (3.2) follows from equivalence of all norms of finite dimensional spaces.

LEMMA 3.2. *Let* $g \in H^i(\Omega_h)$, $0 \leq i \leq 3$. *Then*

$$|\hat{g}|_{i,C_2} \leq C h_e^{i-1} \|g\|_{i,e}. \tag{3.3}$$

Proof. We transform the integral $|\hat{g}|_{i,C_2}^2$ by means of the inverse mapping of the mapping (2.7). (3.3) follows from (2.8) and (2.9) (the Jacobian J_e^{-1} of the inverse mapping is bounded by $c_2 h_e^{-2}$).

Often, we shall make use of the Bramble-Hilbert lemma (see [2] and [3]) on linear functionals. In fact, this lemma will be applied for the domain C_2 only.

LEMMA 3.3 (SPECIAL CASE OF THE BRAMBLE-HILBERT LEMMA). *Let the linear functional $L(\varphi)$ be bounded on $H^{k+1}(C_2)$, $|L(\varphi)| \leqslant M\|\varphi\|_{k+1,C_2}$, and let it vanish for $\varphi \in \hat{P}(k)$. Then there exists a constant C independent on φ such that*

$$|L(\varphi)| \leqslant CM|\varphi|_{k+1,C_2} \quad \forall \varphi \in H^{k+1}(C_2). \tag{3.4}$$

If $L(\varphi)$ vanishes for $\varphi \in \hat{Q}(k)$, then

$$|L(\varphi)| \leqslant CM\left\{\left\|\frac{\partial^{k+1}\varphi}{\partial\xi_1^{k+1}}\right\|_{0,C_2} + \left\|\frac{\partial^{k+1}\varphi}{\partial\xi_2^{k+1}}\right\|_{0,C_2}\right\}. \tag{3.5}$$

Lemma 3.3 allows us to estimate the interpolation error for a given function. The interpolate φ_I of a function φ defined on C_2 is the polynomial $\Sigma_{j=1}^{8} \varphi_j N_j(\xi_1, \xi_2)$, where φ_j are the values of φ at the nodes of C_2. The interpolate g_I of a function g defined on $\overline{\Omega}_h$ is the function from V_h which assumes the same values at all nodes of the given partition as the function g.

LEMMA 3.4. *If $\varphi \in H^3(C_2)$, then*

$$\|\varphi - \varphi_I\|_{j,C_2} \leqslant C|\varphi|_{3,C_2}, \qquad j = 0, \ldots, 3. \tag{3.6}$$

Proof. We get (3.6) if we apply Lemma 3.3 to the functional $L(\varphi) = (\varphi - \varphi_I, w)_{j,C_2}$, and afterwards we set $\varphi - \varphi_I$ for w.

We shall need estimates of the error functional $E(\varphi) = \int_{C_2} \varphi d\xi - I(\varphi)$. Such estimates follow immediately from (3.5) and (3.4).

LEMMA 3.5. *Let $I(\varphi)$ be a formula which integrates exactly all polynomials from $\hat{Q}(3)$. Then*

$$|E(\varphi)| \leqslant C\left\{\left\|\frac{\partial^4\varphi}{\partial\xi_1^4}\right\|_{0,C_2} + \left\|\frac{\partial^4\varphi}{\partial\xi_2^4}\right\|_{0,C_2}\right\}. \tag{3.7}$$

If $I(\varphi)$ integrates exactly all polynomials from $\hat{P}(4)$, then

$$|E(\varphi)| \leqslant C|\varphi|_{5,C_2}. \tag{3.8}$$

The following is the main lemma from which, among other things, existence and uniqueness of the approximate solution u_h follows.

LEMMA 3.6. *Let $I(\varphi)$ be any symmetric formula with positive coefficients which integrates exactly all polynomials from $\hat{Q}(3)$ or any formula which integrates exactly all polynomials from $\hat{P}(4)$. Let the coefficients a_{ij} satisfy* (2.3′) *and let them be bounded and in the latter case be Lipschitz continuous on $\overline{\Omega}_h$. Finally, let the finite element partitions be* 1*-strongly regular* (*in fact, it is sufficient that* (2.8) *be true for $|\alpha| \leqslant 1$ and $|\alpha| \leqslant 2$, respectively*). *Then $|v|_h = \{a_h(v, v)\}^{1/2}$ is a norm on V_h equivalent uniformly with respect to h to the norm $|v|_{1,\Omega_h}$, i.e. there exists a constant c_4 independent of h such that*

$$c_4^{-1}|v|_{1,\Omega_h} \leqslant |v|_h \leqslant c_4|v|_{1,\Omega_h} \quad \forall v \in V_h. \tag{3.9}$$

Remark 4. Among formulas satisfying the assumptions of Lemma 3.6 Gauss' 2×2 formula has the smallest number of points.

Proof. (a) Let $I(\varphi)$ be a symmetric formula with positive coefficients which integrates exactly all polynomials from $\hat{Q}(3)$. Denote by $I^*(\varphi)$ the special case of Gauss' 2×2 formula. Denote by γ the value $\gamma = I(\xi_1^4) = I(\xi_2^4)$ and by α the value $\alpha = (45/16)(4/5 - \gamma)$. As $I^*(\xi_1^4) = I^*(\xi_2^4) = 4/9$, we easily find that if $\alpha \neq 1$, the formula

$$I^0(\varphi) = \frac{1}{1-\alpha}[I(\varphi) - \alpha I^*(\varphi)]$$

integrates exactly all polynomials from $\hat{P}(4)$. Hence

$$I(\varphi) = \alpha I^*(\varphi) + (1-\alpha) I^0(\varphi). \tag{3.10}$$

If $\alpha = 1$, then $I(\xi_i^4) = I^*(\xi_i^4)$ $(i = 1, 2)$, and $E^1(\varphi) = I(\varphi) - I^*(\varphi)$ satisfies (3.8). We have

$$I(\varphi) = I^*(\varphi) + E^1(\varphi). \tag{3.11}$$

Now consider the function $\psi = (\partial\hat{v}/\partial\xi_1)^2 + (\partial\hat{v}/\partial\xi_2)^2$, where $\hat{v} \in \hat{P}$. As $\psi \in \hat{P}(4)$ it follows from (3.10) and (3.11), respectively, that $I(\psi) = \alpha I^*(\psi) + (1-\alpha)\int_{C_2} \psi\, d\xi$. As $\hat{v}$ is of the form (2.5), $\int_{C_2} \psi\, d\xi$ must be of the form $\mathbf{z}^T A \mathbf{z}$ where $\mathbf{z} = (\alpha_2, \ldots, \alpha_8)^T$ and A is a symmetric 7×7 matrix. Further, we compute easily $I^*(\psi) = \mathbf{z}^T A \mathbf{z} - (16/45)(\alpha_7^2 + \alpha_8^2)$; hence

$$I(\psi) = \mathbf{z}^T A \mathbf{z} - \frac{16}{45}\alpha(\alpha_7^2 + \alpha_8^2) = \left(1 - \frac{4}{9}\alpha\right)\mathbf{z}^T A \mathbf{z} + \frac{4}{9}\alpha\left[\mathbf{z}^T A \mathbf{z} - \frac{4}{5}(\alpha_7^2 + \alpha_8^2)\right].$$

A direct computation gives

$$\mathbf{z}^T A \mathbf{z} - \frac{4}{5}(\alpha_7^2 + \alpha_8^2) = 4\left(\alpha_2^2 + \frac{2}{3}\alpha_2\alpha_8 + \frac{4}{9}\alpha_8^2\right) + 4\left(\alpha_3^2 + \frac{2}{3}\alpha_3\alpha_7 + \frac{4}{9}\alpha_7^2\right)$$

$$+ \frac{16}{3}\alpha_4^2 + \frac{8}{3}\alpha_5^2 + \frac{16}{3}\alpha_6^2 \geqslant 0.$$

As γ is always positive, α must be smaller than $9/4$; and setting $c = 1 - 4\alpha/9 > 0$, we have

$$I(\psi) \equiv I\left(\left(\frac{\partial\hat{v}}{\partial\xi_1}\right)^2 + \left(\frac{\partial\hat{v}}{\partial\xi_2}\right)^2\right) \geqslant c\int_{C_2} \psi\, d\xi \equiv c\,|\hat{v}|_{1,C_2}^2 \quad \forall \hat{v} \in \hat{P}. \tag{3.12}$$

This inequality will be used to prove the first part of Lemma 3.6.

(b) From (2.3′), (2.12), (2.17) (the coefficients A_r are positive) we get

$$a_h(v, v) \geqslant c_1 \sum_e I\left(J_e\left[\left(\widehat{\frac{\partial v}{\partial x_1}}\right)^2 + \left(\widehat{\frac{\partial v}{\partial x_2}}\right)^2\right]\right). \tag{3.13}$$

If Δ_x is the vector $(\widehat{\partial v/\partial x_1}, \widehat{\partial v/\partial x_2})^T$ and Δ_ξ the vector $(\partial\hat{v}/\partial\xi_1, \partial\hat{v}/\partial\xi_2)^T$, we have $\Delta_\xi = D\Delta_x$ where $D = \{\partial x_j^e/\partial\xi_i\}_{i,j=1}^2$. From (2.8) it follows $\|D\|^2 \leqslant Ch_e^2$. If we compute D^{-1} and take into account (2.9), we get $\|D^{-1}\|^2 \leqslant Ch_e^{-2}$. Now for any nonsingular matrix M the matrix $M^T M$ is positive definite and $\Delta_\xi^T M^T M \Delta_\xi \geqslant \|M^{-1}\|^{-2}\|\Delta_\xi\|^2$. Therefore,

$$J_e\left[\left(\widehat{\frac{\partial v}{\partial x_1}}\right)^2 + \left(\widehat{\frac{\partial v}{\partial x_2}}\right)^2\right] = J_e\|\Delta_x\|^2 = J_e\Delta_\xi^T(D^{-1})^T D^{-1}\Delta_\xi \geqslant J_e\|D\|^{-2}\|\Delta_\xi\|^2$$

$$\geqslant C\|\Delta_\xi\|^2 = C\psi;$$

and with respect to (3.13) and (3.12), $a_h(v, v) \geqslant c_1 C \Sigma_e I(\psi) \geqslant C\Sigma_e |\hat{v}|^2_{1,C_2}$.

On the other hand, $\|\Delta_\xi\|^2 = \Delta_x^T D^T D \Delta_x \geqslant \|D^{-1}\|^{-2}\|\Delta_x\|^2 \geqslant Ch_e^2\|\Delta_x\|^2$; hence

$$|\hat{v}|^2_{1,C_2} \geqslant Ch_e^2 \int_e J_e^{-1}\left[\left(\frac{\partial v}{\partial x_1}\right)^2 + \left(\frac{\partial v}{\partial x_2}\right)^2\right] dx \geqslant C\int_e \left[\left(\frac{\partial v}{\partial x_1}\right)^2 + \left(\frac{\partial v}{\partial x_2}\right)^2\right] dx,$$

i.e. $|\hat{v}|^2_{1,C_2} \geqslant C|v|^2_{1,e}$ $\forall \hat{v} \in \hat{P}$, and the final estimate is $a_h(v, v) \geqslant C\Sigma_e |v|^2_{1,e} = C|v|^2_{1,\Omega_h}$.

(c) Let $I(\varphi)$ integrate exactly all polynomials from $\hat{P}(4)$. Consider first the sum

$$S_e = J_e \sum_{i,j=1}^{2} \hat{a}_{ij} \widehat{\frac{\partial v}{\partial x_i}}\, \widehat{\frac{\partial v}{\partial x_j}}$$

(the values of this sum at Q_r appear in (2.17)). We have

$$S_e = \Delta_\xi^T B \Delta_\xi = \sum_{i,j=1}^{2} b_{ij} \frac{\partial \hat{v}}{\partial \xi_i}\frac{\partial \hat{v}}{\partial \xi_j}, \qquad B = J_e(D^{-1})^T \hat{A} D^{-1} \tag{3.14}$$

($\hat{A}$ is the matrix $\{\hat{a}_{ij}\}^2_{i,j=1}$). Elementary computations give the following expressions for the coefficients of the symmetric matrix B:

$$\begin{aligned}
b_{11} &= J_e^{-1}\left\{\left(\frac{\partial x_2^e}{\partial \xi_2}\right)^2 \hat{a}_{11} - 2\frac{\partial x_1^e}{\partial \xi_2}\frac{\partial x_2^e}{\partial \xi_2}\hat{a}_{12} + \left(\frac{\partial x_1^e}{\partial \xi_2}\right)^2 \hat{a}_{22}\right\},\\
b_{12} &= J_e^{-1}\left\{-\frac{\partial x_2^e}{\partial \xi_1}\frac{\partial x_2^e}{\partial \xi_2}\hat{a}_{11} + \left(\frac{\partial x_1^e}{\partial \xi_1}\frac{\partial x_2^e}{\partial \xi_2} + \frac{\partial x_2^e}{\partial \xi_1}\frac{\partial x_1^e}{\partial \xi_2}\right)\hat{a}_{12} - \frac{\partial x_1^e}{\partial \xi_1}\frac{\partial x_1^e}{\partial \xi_2}\hat{a}_{22}\right\},\\
b_{22} &= J_e^{-1}\left\{\left(\frac{\partial x_2^e}{\partial \xi_1}\right)^2 \hat{a}_{11} - 2\frac{\partial x_1^e}{\partial \xi_1}\frac{\partial x_2^e}{\partial \xi_1}\hat{a}_{12} + \left(\frac{\partial x_1^e}{\partial \xi_1}\right)^2 \hat{a}_{22}\right\}.
\end{aligned} \tag{3.15}$$

Let us denote by β any of the factors appearing at any of the coefficients $\hat{a}_{ij}$ on the right-hand sides of (3.15). We shall need later the following estimate of β:

$$|D^\alpha \beta| \leqslant Ch_e^{|\alpha|}, \qquad |\alpha| \geqslant 0 \tag{3.16}$$

(to prove (3.16) differentiate the identity $J_e J_e^{-1} = 1$ and prove by induction $D^\alpha J_e^{-1} = O(h_e^{-2+|\alpha|})$; (3.16) follows by Leibniz rule). At this time we use (3.16) with $|\alpha| \leqslant 1$. Lipschitz continuity of a_{ij} and (3.16) with $|\alpha| \leqslant 1$ give $b_{ij} = b_{ij}^0 + O(h_e)$ where b_{ij}^0 mean the value of b_{ij} at the center (0, 0). Therefore

$$S_e = \Delta_\xi^T B^0 \Delta_\xi + O(h_e)\|\Delta_\xi\|^2 = \sum_{i,j=1}^{2} b_{ij}^0 \frac{\partial \hat{v}}{\partial \xi_i}\frac{\partial \hat{v}}{\partial \xi_j} + O(h_e)|\hat{v}|^2_{1,C_2}, \qquad B^0 = \{b_{ij}^0\}^2_{i,j=1}.$$

As $(\partial\hat{v}/\partial\xi_i)(\partial\hat{v}/\partial\xi_j) \in \hat{P}(4)$, we get

$$\begin{aligned}
&\sum_r A_r J_e(Q_r) \sum_{i,j=1}^{2} \hat{a}_{ij}(Q_r) \widehat{\frac{\partial v}{\partial x_i}}(Q_r) \widehat{\frac{\partial v}{\partial x_j}}(Q_r)\\
&= \sum_r A_r \sum_{i,j=1}^{2} b_{ij}^0 \frac{\partial \hat{v}(Q_r)}{\partial \xi_i}\frac{\partial \hat{v}(Q_r)}{\partial \xi_j} + O(h_e)|\hat{v}|^2_{1,C_2}\\
&= \int_{C_2} \sum_{i,j=1}^{2} b_{ij}^0 \frac{\partial \hat{v}}{\partial \xi_i}\frac{\partial \hat{v}}{\partial \xi_j}\, d\xi + O(h_e)|\hat{v}|^2_{1,C_2}.
\end{aligned}$$

Further, $\Delta_\xi^T B \Delta_\xi \geqslant c_1 J_e \Delta_\xi^T (D^{-1})^T D^{-1} \Delta_\xi$ (because $B = J_e (D^{-1})^T \hat{A} D^{-1}$ and $\hat{A}$ satisfies (2.3')). We have proved before that $J_e \Delta_\xi^T (D^{-1})^T D^{-1} \Delta_\xi \geqslant C \|\Delta_\xi\|^2$. Therefore,

$$\Delta_\xi^T B^0 \Delta_\xi \geqslant C \left[\left(\frac{\partial \hat{v}}{\partial \xi_1} \right)^2 + \left(\frac{\partial \hat{v}}{\partial \xi_2} \right)^2 \right] \quad \text{and} \quad \int_{C_2} \sum_{i,j=1}^{2} b_{ij}^0 \frac{\partial \hat{v}}{\partial \xi_i} \frac{\partial \hat{v}}{\partial \xi_j} d\xi \geqslant C |\hat{v}|_{1,C_2}^2 .$$

Consequently, for h sufficiently small,

$$a_h(v, v) = \sum_e \sum_r A_r J_e(Q_r) \sum_{i,j=1}^{2} \hat{a}_{ij}(Q_r) \widehat{\frac{\partial v}{\partial x_i}}(Q_r) \widehat{\frac{\partial v}{\partial x_j}}(Q_r)$$

$$\geqslant C \sum_e |\hat{v}|_{1,C_2}^2 \geqslant C \sum_e |v|_{1,e}^2 = C |v|_{1,\Omega_h}^2 .$$

(d) We have

$$a_h(v, v) = \sum_e \sum_r A_r \sum_{i,j=1}^{2} b_{ij}(Q_r) \frac{\partial \hat{v}(Q_r)}{\partial \xi_i} \frac{\partial \hat{v}(Q_r)}{\partial \xi_j} .$$

As β and $\hat{a}_{ij}$ are bounded, so are bounded the coefficients b_{ij}. Hence, from (3.2) and (3.3) we easily get $a_h(v, v) \leqslant C \Sigma_e |\hat{v}|_{1,C_2}^2 \leqslant C |v|_{1,\Omega_h}^2$.

4. Superconvergence Theorems. First, the integration by Gauss' 2×2 formula is considered. Besides the special notation $I^*(\varphi)$, for this formula we will use the sign * for other quantities as, e.g., for u_h^*, $a_h^*(v, v)$, $f_h^*(v)$, $|\cdot|_h^*$, E^*, Q_r^*. The rate of convergence will not be expressed by means of the norm $|\cdot|_h^*$, because it depends on the coefficients a_{ij} of the operator Lu. We consider the norm $|\cdot|_h^*$ associated to the operator $Lu = -\Delta u$ and this norm is denoted by $\|\cdot\|_h$. Hence, (A_r^* are equal to 1)

$$\|v\|_h = \left\{ \sum_e \sum_{r=1}^{4} J_e(Q_r^*) \left[\left(\widehat{\frac{\partial v}{\partial x_1}}(Q_r^*) \right)^2 + \left(\widehat{\frac{\partial v}{\partial x_2}}(Q_r^*) \right)^2 \right] \right\}^{1/2} , \tag{4.1}$$

$$Q_r^* = \left(\pm \frac{\sqrt{3}}{3}, \pm \frac{\sqrt{3}}{3} \right).$$

The norm $\|\cdot\|_h$ is on V_h equivalent uniformly with respect to h to the norm $|v|_{1,\Omega_h}$:

$$c_4^{-1} |v|_{1,\Omega_h} \leqslant \|v\|_h \leqslant c_4 |v|_{1,\Omega_h} \quad \forall v \in V_h . \tag{4.2}$$

Ciarlet and Raviart proved (see [4, p. 462]) the following estimate for the discretization error $u - u_h$ where u is the solution of (2.4) and u_h the solution of (2.19):

$$\|u - u_h\|_{1,\Omega_h} \leqslant C h^2$$

(they consider 9 degrees of freedom elements; however the bound can be proved in the same way for 8 degrees of freedom elements considered here). We can say that in the sense of L_2-norm average error of the gradient is of the order $O(h^2)$. We shall prove that $\|u - u_h^*\|_h \leqslant C h^3$, and this is the reason that we speak about superconvergence. In fact, let us denote by N_G the number of all Gaussian points and by $E(P)$ the error of the gradient,

$$E(P) = \left[\left(\frac{\partial(u - u_h^*)(P)}{\partial x_1}\right)^2 + \left(\frac{\partial(u - u_h^*)(P)}{\partial x_2}\right)^2\right]^{1/2}.$$

We have

$$\text{meas}\ \Omega_h = \sum_e \int_e dx = \sum_e \int_{C_2} J_e \, d\xi \leqslant Ch^2 N_G;$$

therefore, $N_G \geqslant Ch^{-2}$, $C > 0$. By the Cauchy inequality we prove under the additional assumption $h_e/h \geqslant C > 0 \ \forall e$

$$N_G^{-1} \sum_{P \in G} E(P) \leqslant C \|u - u_h^*\|_h.$$

Hence, it follows that the arithmetic mean of errors of the gradient at Gaussian points is $O(h^3)$.

THEOREM 4.1. *Let the finite element partitions of* Ω *be* 2*-strongly regular. Further, assume the boundary* Γ *to be sufficiently smooth,*

$$u \in H^4(\Omega), \quad a_{ij} \in H^3(\Omega), \quad f \in H^3(\Omega), \tag{4.3}$$

and the operator Lu to be uniformly elliptic. Finally, let the quadrature formula (2.16) *be Gauss'* 2 × 2 *product formula. Then there exists a constant C independent on h (it is of the form* $C_1 \|u\|_{4,\Omega} + C_2 \|f\|_{3,\Omega}$ *where* C_1 *and* C_2 *do not depend both on h and u and f) such that*

$$\|u - u_h^*\|_h \leqslant Ch^3. \tag{4.4}$$

Proof. (a) Subtracting (2.19) from (2.15), we get

$$a_h^*(u - u_h^*, v) = S_h^*(v) - R_h^*(u, v) \quad \forall v \in V_h, \tag{4.5}$$

where

$$R_h^*(u, v) = a(u, v) - a_h^*(u, v), \qquad S_h^*(v) = (f, v)_{0,\Omega_h} - f_h^*(v). \tag{4.6}$$

Further,

$$a_h^*(u_I - u_h^*, v) = S_h^*(v) - R_h^*(u, v) - a_h^*(u - u_I, v) \quad \forall v \in V_h \tag{4.7}$$

(u_I is the interpolate of u). We prove later that

$$\left.\begin{aligned} &|a_h^*(u - u_I, v)| \leqslant Ch^3 |v|_{1,\Omega_h} &(4.8)\\ &|R_h^*(u, v)| \leqslant Ch^3 |v|_{1,\Omega_h} &(4.9)\\ &|S_h^*(v)| \leqslant Ch^3 |v|_{1,\Omega_h} &(4.10)\end{aligned}\right\} \quad \forall v \in V_h.$$

From these inequalities, (4.7) and (3.9) it follows that

$$|a_h^*(u_I - u_h^*, v)| \leqslant Ch^3 |v|_h^* \quad \forall v \in V_h. \tag{4.11}$$

Setting $v = u_I - u_h^* \in V_h$, we obtain

$$|u_I - u_h^*|_h^* \leqslant Ch^3. \tag{4.12}$$

Consequently, $|u - u_h^*|_h^* \leqslant |u - u_I|_h^* + |u_I - u_h^*|_h^* \leqslant |u - u_I|_h^* + Ch^3$. We also prove

$$(4.13) \qquad |u - u_I|_h^* \leqslant Ch^3.$$

The last two inequalities give

$$(4.14) \qquad |u - u_h^*|_h^* \leqslant Ch^3.$$

Now from (4.1), (2.17) (with $A_r = 1$, $Q_r = Q_r^*$, $r = 1, \ldots, 4$) and (2.3′) it follows for any function g piecewise differentiable in Ω_h that

$$(4.15) \qquad \|g\|_h \leqslant C|g|_h^*,$$

which together with (4.14) proves the theorem.

(b) We prove (4.8) and (4.13). We express $a_h^*(\omega, v)$, where $v \in V_h$ and ω is any function such that $a_h^*(\omega, v)$ is defined, as follows (see (3.14)):

$$a_h^*(\omega, v) = \sum_e \sum_{r=1}^{4} \sum_{i,j=1}^{2} b_{ij}(Q_r^*) \frac{\partial \hat{\omega}(Q_r^*)}{\partial \xi_i} \frac{\partial \hat{v}(Q_r^*)}{\partial \xi_j}.$$

The coefficients b_{ij} are bounded (it follows from (3.15) and (3.16) with $|\alpha| = 0$). Hence by (3.2),

$$\left| \sum_{i,j=1}^{2} b_{ij}(Q_r^*) \frac{\partial \hat{\omega}(Q_r^*)}{\partial \xi_i} \frac{\partial \hat{v}(Q_r^*)}{\partial \xi_j} \right| \leqslant C \left[\left(\frac{\partial \hat{\omega}(Q_r^*)}{\partial \xi_1} \right)^2 + \left(\frac{\partial \hat{\omega}(Q_r^*)}{\partial \xi_2} \right)^2 \right]^{1/2} |\hat{v}|_{1,C_2}$$

and

$$(4.16) \qquad |a_h^*(\omega, v)| \leqslant C \sum_e |\hat{v}|_{1,C_2} \sum_{r=1}^{4} \left[\left(\frac{\partial \hat{\omega}(Q_r^*)}{\partial \xi_1} \right)^2 + \left(\frac{\partial \hat{\omega}(Q_r^*)}{\partial \xi_2} \right)^2 \right]^{1/2}.$$

We estimate the functional $L(\hat{u}) = \partial \hat{\omega}(Q_r^*)/\partial \xi_1$ where $\hat{\omega} = \hat{u} - \hat{u}_I$. It is bounded on $H^4(C_2)$, it vanishes for $\hat{u} \in \hat{P}$ because $\hat{u}_I = \hat{u}$ in this case. If $\hat{u} = \xi_2^3$, then $\hat{u}_I = \xi_2$ and $\partial \hat{\omega}/\partial \xi_1 = 0$. If $\hat{u} = \xi_1^3$, then $\hat{u}_I = \xi_1$ and $\partial \hat{\omega}(Q_r^*)/\partial \xi_1 = (3\xi_1^2 - 1)_{\xi_1 = \pm\sqrt{3}/3} = 0$. Hence $L(\hat{u})$ vanishes for $\hat{u} \in \hat{P}(3)$ and, according to the Bramble-Hilbert lemma, $|\partial \hat{\omega}(Q_r^*)/\partial \xi_1| \leqslant C|\hat{u}|_{4,C_2}$; in the same way we get $|\partial \hat{\omega}(Q_r^*)/\partial \xi_2| \leqslant C|\hat{u}|_{4,C_2}$. From (4.16), (2.9) and (3.3) we obtain

$$|a_h^*(u - u_I, v)| \leqslant C \sum_e |\hat{u}|_{4,C_2} |\hat{v}|_{1,C_2} \leqslant C \sum_e h_e^3 \|u\|_{4,e} |v|_{1,e} \leqslant Ch^3 \|u\|_{4,\Omega_h} |v|_{1,\Omega_h}$$

The proof of (4.13) is similar.

(c) To prove (4.9) express $a(u, v)$ as follows:

$$(4.17) \qquad a(u, v) = \sum_e \int_{C_2} \sum_{i,j=1}^{2} J_e \hat{a}_{ij} \widehat{\frac{\partial u}{\partial x_i}} \widehat{\frac{\partial v}{\partial x_j}} \, d\xi.$$

With respect to definition of the error functional $E^*(\varphi)$ we get

$$(4.18) \qquad R_h^*(u, v) = \sum_e E^* \left(\sum_{i,j=1}^{2} J_e \hat{a}_{ij} \widehat{\frac{\partial u}{\partial x_i}} \widehat{\frac{\partial v}{\partial x_j}} \right).$$

We estimate $E^*(J_e \hat{a}_{11} (\widehat{\partial u/\partial x_1})(\widehat{\partial v/\partial x_1}))$. We have

$$J_e \widehat{\frac{\partial v}{\partial x_1}} = J_e\left[\frac{\partial \hat v}{\partial \xi_1}\frac{\partial \xi_1^e}{\partial x_1} + \frac{\partial \hat v}{\partial \xi_2}\frac{\partial \xi_2^e}{\partial x_1}\right] = \frac{\partial x_2^e}{\partial \xi_2}\frac{\partial \hat v}{\partial \xi_1} - \frac{\partial x_2^e}{\partial \xi_1}\frac{\partial \hat v}{\partial \xi_2},$$

so that

$$(4.19)\quad E^*\left(J_e \hat a_{11}\widehat{\frac{\partial u}{\partial x_1}}\widehat{\frac{\partial v}{\partial x_1}}\right) = E^*\left(\frac{\partial x_2^e}{\partial \xi_2}\hat a_{11}\widehat{\frac{\partial u}{\partial x_1}}\frac{\partial \hat v}{\partial \xi_1}\right) - E^*\left(\frac{\partial x_2^e}{\partial \xi_1}\hat a_{11}\widehat{\frac{\partial u}{\partial x_1}}\frac{\partial \hat v}{\partial \xi_2}\right).$$

To estimate the first term in (4.19) consider the functional $L(\sigma) = E^*(\sigma\partial\hat v/\partial\xi_1) - H(\sigma)\ \forall \sigma \in H^3(C_2)$ where

$$(4.20)\quad H(\sigma) = \frac{1}{45}\left\{\int_{-1}^{1}\frac{\partial^2\sigma(1,\xi_2)}{\partial\xi_2^2}\frac{\partial^2\hat v(1,\xi_2)}{\partial\xi_2^2}d\xi_2 - \int_{-1}^{1}\frac{\partial^2\sigma(-1,\xi_2)}{\partial\xi_2^2}\frac{\partial^2\hat v(-1,\xi_2)}{\partial\xi_2^2}d\xi_2\right\}.$$

If $\sigma = 1, \xi_1, \xi_2, \xi_1^2, \xi_1\xi_2$, then from $\hat v \in \hat P$ (i.e., $\hat v$ is of the form (2.5)) and from (3.7) (which is satisfied by $E^*(\varphi)$) it follows $E^*(\sigma\partial\hat v/\partial\xi_1) = 0$. Also $H(\sigma) = 0$, hence $L(\sigma) = 0$. If $\sigma = \xi_2^2$, then an easy calculation gives $E^*(\sigma\,\partial\hat v/\partial\xi_1) = (16/45)\,\alpha_8 = H(\sigma)$ (α_8 is the last coefficient in (2.5)), thus $L(\sigma) = 0$ for $\sigma \in \hat P(2)$. Further, from the explicit form of $L(\sigma)$,

$$L(\sigma) = \int_{C_2}\sigma\frac{\partial\hat v}{\partial\xi_1}d\xi - \sum_{r=1}^{4}\sigma(Q_r^*)\frac{\partial\hat v(Q_r^*)}{\partial\xi_1} - H(\sigma),$$

it follows by (3.2), the Sobolev lemma, the inequality $\int_{\partial C_2}\varphi^2 ds \leqslant C\|\varphi\|_{1,C_2}^2\ \forall\varphi \in H^1(C_2)$ and by (3.1) that $|L(\sigma)| \leqslant C|\hat v|_{1,C_2}\|\sigma\|_{3,C_2}$. Hence, the Bramble-Hilbert lemma gives

$$(4.21)\quad |L(\sigma)| \leqslant C|\sigma|_{3,C_2}|\hat v|_{1,C_2}.$$

Therefore,

$$\left|\sum_e E^*\left(\frac{\partial x_2^e}{\partial\xi_2}\hat a_{11}\widehat{\frac{\partial u}{\partial x_1}}\frac{\partial\hat v}{\partial\xi_1}\right)\right| \leqslant C\sum_e|\sigma|_{3,C_2}|\hat v|_{1,C_2} + \left|\sum_e H(\sigma)\right|,\qquad \sigma = \frac{\partial x_2^e}{\partial\xi_2}\hat a_{11}\widehat{\frac{\partial u}{\partial x_1}}.$$

However, the last sum in the above inequality is equal to zero. In this sum they appear either integrals over element sides which lie on Γ_h; and, as $v|_{\Gamma_h} = 0$, we have $\partial^2\hat v(\pm 1, \xi_2)/\partial\xi_2^2 = 0$. Or they appear couples of integrals over a common side of two adjacent elements taken in opposite directions with integrands which are the same. The functions $a_{11}\partial u/\partial x_1$ as well as v assume namely the same values on such side (they are continuous on $\bar\Omega_h$); also x_2^e assume the same values on such side because these are quadratic polynomials in one variable determined uniquely by values at the three nodes of that side. So $(\partial^2\sigma/\partial\xi_2^2)(\partial^2\hat v/\partial\xi_2^2)$ assume the same values on such side. We have come to the bound

$$(4.22)\quad \left|\sum_e E^*\left(\frac{\partial x_2^e}{\partial\xi_2}\hat a_{11}\widehat{\frac{\partial u}{\partial x_1}}\frac{\partial\hat v}{\partial\xi_1}\right)\right| \leqslant C\sum_e|\sigma|_{3,C_2}|\hat v|_{1,C_2},\qquad \sigma = \frac{\partial x_2^e}{\partial\xi_2}\hat a_{11}\widehat{\frac{\partial u}{\partial x_1}}.$$

The other term in (4.19) can be estimated in the same way. In this case we set

$$H(\sigma) = \frac{1}{45}\left\{\int_{-1}^{1} \frac{\partial^2\sigma(\xi_1, 1)}{\partial\xi_1^2}\, \frac{\partial^2\hat{v}(\xi_1, 1)}{\partial\xi_1^2}\, d\xi_1 - \int_{-1}^{1} \frac{\partial^2\sigma(\xi_1, -1)}{\partial\xi_1^2}\, \frac{\partial^2\hat{v}(\xi_1, -1)}{\partial\xi_1^2}\, d\xi_1\right\}.$$

We find out that for $\sigma = 1, \xi_1, \xi_2, \xi_1\xi_2, \xi_2^2$ both $E^*(\sigma\partial\hat{v}/\partial\xi_2)$ and $H(\sigma)$ vanish. For $\sigma = \xi_1^2$ we get $E^*(\sigma\,\partial\hat{v}/\partial\xi_2) = H(\sigma) = (16/45)\alpha_7$. The remaining arguments are the same and lead again to the bound (4.22) where now $\sigma = (\partial x_2^e/\partial\xi_1)\hat{a}_{11}\widehat{\partial u}/\partial x_1$.

We must estimate $|\sigma|_{3,C_2}$. Take $\sigma = (\partial x_2^e/\partial\xi_2)\hat{a}_{11}\widehat{\partial u}/\partial x_1$ and express $\widehat{\partial u}/\partial x_1$ by rule of differentiation of composite functions. We find out that σ is a linear combination of $\hat{a}_{11}\partial\hat{u}/\partial\xi_1$ and $\hat{a}_{11}\partial\hat{u}/\partial\xi_2$ with coefficients whose generic notation β was introduced before and which satisfy (3.16). So set $\sigma = \beta\hat{a}_{11}\partial\hat{u}/\partial\xi_i$. As $u \in H^4(\Omega)$ and the mapping (2.7) satisfies (2.8), we have by Sobolev's lemma (applied to $\Omega^0 \supset \bar{\Omega}_h$) $\max_{C_2}|D^\alpha\hat{u}| \leqslant Ch_e^{|\alpha|}\|u\|_{4,\Omega}$ for $|\alpha| \leqslant 2$. Further, from $a_{11} \in H^3(\Omega)$ it follows $a_{11} \in C^1(\bar{\Omega}_h)$ and $\max_{C_2}|D^\alpha(\beta\hat{a}_{11})| \leqslant Ch_e^{|\alpha|}$ for $|\alpha| \leqslant 1$ (β satisfies (3.16)). Therefore, (using (3.16) and (3.3))

$$\begin{aligned}\left|\beta\hat{a}_{11}\frac{\partial\hat{u}}{\partial\xi_i}\right|_{3,C_2} &\leqslant C\{h_e|\beta\hat{a}_{11}|_{3,C_2} + h_e^2|\beta\hat{a}_{11}|_{2,C_2}\}\|u\|_{4,\Omega_h} + C\{h_e|\hat{u}|_{3,C_2} + |\hat{u}|_{4,C_2}\}\\ &\leqslant C\{h_e^4\|\hat{a}_{11}\|_{0,C_2} + h_e^3|\hat{a}_{11}|_{1,C_2} + h_e^2|\hat{a}_{11}|_{2,C_2} + h_e|\hat{a}_{11}|_{3,C_2}\}\|u\|_{4,\Omega_h}\\ &\quad + C\{h_e|\hat{u}|_{3,C_2} + |\hat{u}|_{4,C_2}\}\\ &\leqslant Ch_e^3\{\|a_{11}\|_{3,e}\|u\|_{4,\Omega_h} + \|u\|_{4,e}\}.\end{aligned}$$

The same bound is true for $\sigma = (\partial x_2^e/\partial\xi_1)\hat{a}_{11}\widehat{\partial u}/\partial x_1$. Consequently,

$$\begin{aligned}\left|\sum_e E^*\left(J_e\,\hat{a}_{11}\frac{\widehat{\partial u}}{\partial x_1}\,\frac{\widehat{\partial v}}{\partial x_1}\right)\right| &\leqslant C\sum_e h_e^3\{\|a_{11}\|_{3,e}\|u\|_{4,\Omega_h} + \|u\|_{4,e}\}|v|_{1,e}\\ &\leqslant Ch^3\|u\|_{4,\Omega_h}|v|_{1,\Omega_h}.\end{aligned}$$

In the same way we can estimate the other terms in the first sum of (4.18). Thus, (4.9) is proved (see a remark following the equation (2.15)).

(d) We prove (4.10) for any formula with properties introduced in Lemma 3.6. Let us first observe that estimating each term of the functional $E(\sigma\hat{v})$ (we use (3.2) and Sobolev's lemma) we obtain $|E(\sigma\hat{v})| \leqslant C\|\hat{v}\|_{0,C_2}\|\sigma\|_{2,C_2}$. If $\sigma \in \hat{P}(1)$, then by (3.7) or (3.8) and by (2.6) $E(\sigma\hat{v}) = 0$. Therefore, by the Bramble-Hilbert lemma

$$(4.23)\qquad |E(\sigma\hat{v})| \leqslant C|\sigma|_{2,C_2}\|\hat{v}\|_{0,C_2}\quad \forall\sigma \in H^2(C_2), \forall\hat{v} \in \hat{P}.$$

For $\sigma \in H^3(C_2)$ we get a better estimate. We have $E(\sigma\hat{v}) = E([\sigma - \sigma_I]\hat{v}) + E(\sigma_I\hat{v})$. By (4.23) and (3.6) the bound of the first term is

$$|E([\sigma - \sigma_I]\hat{v})| \leqslant C|\sigma - \sigma_I|_{2,C_2}\|\hat{v}\|_{0,C_2} \leqslant C|\sigma|_{3,C_2}\|\hat{v}\|_{0,C_2}.$$

Further, by (3.7) or (3.8) and by (2.6), (3.1)

$$|E(\sigma_I\hat{v})| \leqslant C\{|\sigma_I|_{2,C_2}|\hat{v}|_{3,C_2} + |\sigma_I|_{3,C_2}|\hat{v}|_{2,C_2}\}\quad\text{or}\quad |E(\sigma_I\hat{v})| \leqslant C|\sigma_I|_{2,C_2}|\hat{v}|_{2,C_2},$$

thus in both cases

$$|E(\sigma_I \hat{v})| \leqslant C\{|\sigma - \sigma_I|_{2,C_2} + |\sigma|_{2,C_2}\}|\hat{v}|_{1,C_2} \leqslant C|\sigma|_{2,C_2}|\hat{v}|_{1,C_2}.$$

Hence,

$$(4.24)\quad |E(\sigma\hat{v})| \leqslant C|\sigma|_{2,C_2}|\hat{v}|_{1,C_2} + C|\sigma|_{3,C_2}\|\hat{v}\|_{0,C_2} \quad \forall \sigma \in H^3(C_2), \forall \hat{v} \in \hat{P}.$$

Now we must come back to the original notation: $\widetilde{f}$ is the extension of f defined on R^2 by $\widetilde{f} = -\sum_{i,j=1}^{2} \partial(\widetilde{a}_{ij}\partial\widetilde{u}/\partial x_j)/\partial x_i$. We express $S_h(v)$ as follows:

$$S_h(v) = \sum_e{}' E(J_e \hat{\widetilde{f}}\, \hat{v}) + \sum_e{}'' E(J_e \hat{f} \hat{v}).$$

In the first sum the summation is taken over the boundary elements, in the second one over the inner elements (if h is sufficiently small, the inner elements belong to Ω; hence we may write $\hat{f}$ in the second sum).

Let us use (4.23) with $\sigma = J_e \hat{\widetilde{f}}$. As $D^\alpha J_e = O(h_e^{|\alpha|+2})$ (it follows from (2.8)), we easily get

$$|J_e \hat{\widetilde{f}}|_{2,C_2} \leqslant C\{h_e^2|\hat{\widetilde{f}}|_{2,C_2} + h_e^3|\hat{\widetilde{f}}|_{1,C_2} + h_e^4\|\hat{\widetilde{f}}\|_{0,C_2}\} \leqslant Ch_e^3\|\widetilde{f}\|_{2,e}$$

and

$$\left|\sum_e{}' E(J_e \hat{\widetilde{f}}\hat{v})\right| \leqslant C \sum_e{}' h_e^3\|\widetilde{f}\|_{2,e}\|\hat{v}\|_{0,C_2}.$$

For the boundary elements Friedrichs' inequality gives $\|\hat{v}\|_{0,C_2} \leqslant C|\hat{v}|_{1,C_2}$ ($\hat{v}$ vanishes on one side of C_2). Therefore

$$\begin{aligned}\left|\sum_e{}' E(J_e \hat{\widetilde{f}}\hat{v})\right| &\leqslant C\sum_e{}' h_e^3\|\widetilde{f}\|_{2,e}|v|_{1,e} \leqslant Ch^3\|\widetilde{f}\|_{2,\Omega_h}|v|_{1,\Omega_h}\\ &\leqslant Ch^3\|\widetilde{u}\|_{4,\Omega_h}|v|_{1,\Omega_h} \leqslant Ch^3\|u\|_{4,\Omega}|v|_{1,\Omega_h}.\end{aligned}$$

For the inner elements we use (4.24) with $\sigma = J_e\hat{f}$, and we easily get

$$\left|\sum_e{}'' E(J_e \hat{f}\hat{v})\right| \leqslant Ch^3\|f\|_{3,\Omega}|v|_{1,\Omega_h}.$$

Thus,

$$(4.25)\quad \begin{aligned}&|S_h(v)| \leqslant Ch^3[\|u\|_{4,\Omega_h} + \|f\|_{3,\Omega}]\,|v|_{1,\Omega_h} \quad \forall v \in V_h,\\ &S_h(v) = (f, v)_{0,\Omega_h} - f_h(v).\end{aligned}$$

Remark 5. Relaxing one assumption of the theorem, changing namely the condition of 2-strong regularity into 1-strong regularity, one can prove in the same way that

$$\|u - u_h^*\|_{1,\Omega\cap\Omega_h} \leqslant Ch^2.$$

(2.8) is satisfied for $|\alpha| \leqslant 2$ if instead of (2.11) we require $\rho(a_i, a_i') \leqslant C_2h_e^2$ (see Remark 2).

Theorem 4.2 introduced below shows that the superconvergence phenomenon is not closely connected with Gauss' 2×2 formula. For Theorem 4.2 we need that the finite element partitions, besides being 2-strongly regular, are such that

$$\left| J_e^{-1} \frac{\partial x_i^e}{\partial \xi_1} \frac{\partial x_j^e}{\partial \xi_2} - J_{e^1}^{-1} \frac{\partial x_i^{e^1}}{\partial \xi_1} \frac{\partial x_j^{e^1}}{\partial \xi_2} \right| \leq Ch, \qquad i, j = 1, 2, \tag{4.26}$$

for any two adjacent elements e, e^1.

Remark 6. Condition (4.26) is of different nature than conditions (2.8) and (2.9) because it does not concern single elements. If (4.26) is satisfied and the coefficients a_{ij} are Lipschitz continuous, then the difference of values of the coefficient b_{12} (see (3.15)) on adjacent elements is $O(h)$. We give a sufficient condition that (4.26) be fulfilled. It is similar to conditions given in Remark 2.

We ask again (2.10) and instead of (2.11) a weaker condition

$$\rho(a_i, a_i') \leq C_2 h_e^2, \qquad 1 \leq i \leq 8. \tag{4.27}$$

However, we add a third condition. Let us denote by α_e and β_e the angles which make the sides $a_1'a_2'$ and $a_1'a_4'$, respectively, with the x_1-axis (if a suitable notation is used then $|\beta_e - \alpha_e| = \omega_e$). The condition reads

$$|\alpha_e - \alpha_{e^1}| \leq C_3 h, \qquad |\beta_e - \beta_{e^1}| \leq C_3 h \tag{4.28}$$

for any two adjacent elements e, e^1.

The proof that this condition is sufficient is simple: For the parallelogram e' corresponding to the element e one computes

$$\frac{\partial x_1^{e'}}{\partial \xi_1} = \frac{1}{2} h_e \cos \alpha_e, \quad \frac{\partial x_1^{e'}}{\partial \xi_2} = \frac{1}{2} k_e \cos \beta_e, \quad \frac{\partial x_2^{e'}}{\partial \xi_1} = \frac{1}{2} h_e \sin \alpha_e, \quad \frac{\partial x_2^{e'}}{\partial \xi_2} = \frac{1}{2} k_e \sin \beta_e.$$

Hence, $J_{e'} = \frac{1}{4} h_e k_e \sin \omega_e$ and, for instance,

$$J_{e'}^{-1} \frac{\partial x_1^{e'}}{\partial \xi_1} \frac{\partial x_1^{e'}}{\partial \xi_2} = \frac{\cos \alpha_e \cos \beta_e}{\sin \omega_e}.$$

From (4.27) and $k_e/h_e \geq C_3$ we easily get

$$J_e^{-1} \frac{\partial x_1^e}{\partial \xi_1} \frac{\partial x_1^e}{\partial \xi_2} = \frac{\cos \alpha_e \cos \beta_e}{\sin \omega_e} + O(h_e).$$

From (2.10) and (4.28) it follows (4.26) for $i = j = 1$.

THEOREM 4.2. *Let the quadrature formula* (2.16) *be either a symmetric formula with positive coefficients which integrates exactly all polynomials from* $\hat{Q}(3)$ *or a formula which integrates exactly all polynomials from* $\hat{P}(4)$. *Let the finite element partitions, besides being* 2*-strongly regular, satisfy* (4.26), *and let the remaining assumptions of Theorem* 4.1 *be fulfilled. Then there exists a constant C independent of h* (*it is of the form* $C_1\|u\|_{4,\Omega} + C_2\|f\|_{3,\Omega}$) *such that*

$$\|u - u_h\|_h \leq Ch^3. \tag{4.29}$$

Proof. (a) Distinguish two cases: $I(\varphi)$ is of the form (3.10) where $I^0(\varphi)$ integrates exactly all polynomials from $\hat{P}(4)$ or $I(\varphi)$ itself has this property. The other possibility is included in the preceding one because $I(\varphi)$ is again of the form (3.10) with $\alpha = 0$. The second case is that $I(\varphi)$ is of the form (3.11) where $E^1(\varphi)$ satisfies (3.8). Consider the first case. Then

$$a_h(u, v) = (f, v)_{0,\Omega_h} - \alpha R_h^*(u, v) - (1-\alpha)R_h^0(u, v) \quad \forall v \in V_h,$$

(4.30)

$$R_h^0(w, v) = a(w, v) - a_h^0(w, v)$$

and $a_h^0(w, v)$ is the approximate value of $a(w, v)$ computed by means of the formula $I^0(\varphi)$ (or $I(\varphi)$ if $\alpha = 0$). Hence, subtracting (2.19), we get

$$a_h(u - u_h, v) = S_h(v) - \alpha R_h^*(u, v) - (1-\alpha)R_h^0(u, v) \quad \forall v \in V_h,$$

where $S_h(v)$ is defined in (4.25). Adding $a_h(u_I - u, v)$ to both sides, we easily obtain

$$a_h(u_I - u_h, v) = S_h(v) - \alpha R_h^*(u, v) + \alpha a_h^*(u_I - u, v) - (1-\alpha)R_h^0(u_I, v)$$

(4.31)

$$+ (1-\alpha)a(u_I - u, v) \quad \forall v \in V_h.$$

Suppose that we prove

(4.32) $$|R_h^0(u_I, v)| \leqslant Ch^3|v|_{1,\Omega_h} \quad \forall v \in V_h,$$

(4.33) $$|a(u - u_I, v)| \leqslant Ch^3|v|_{1,\Omega_h} \quad \forall v \in V_h.$$

Then putting $v = u_I - u_h$ in (4.31) we get, by (4.25), (4.9), (4.8), (4.32), (4.33) and (3.9)

(4.34) $$|u_I - u_h|_{1,\Omega_h} \leqslant Ch^3.$$

Consequently, by (4.15), (4.13), (4.34) and (4.2)

$$\|u - u_h\|_h \leqslant \|u - u_I\|_h + \|u_I - u_h\|_h \leqslant Ch^3 + c_4|u_I - u_h|_{1,\Omega_h} \leqslant Ch^3.$$

(b) To prove (4.32) we express $a(u_I, v)$ as follows:

$$a(u_I, v) = \sum_e \int_{C_2} \sum_{i,j=1}^{2} b_{ij} \frac{\partial \hat{u}_I}{\partial \xi_i} \frac{\partial \hat{v}}{\partial \xi_j}\, d\xi$$

(b_{ij} are the coefficients (3.15)). Hence,

(4.35) $$R_h^0(u_I, v) = \sum_e E^0\left(\sum_{i,j=1}^{2} b_{ij} \frac{\partial \hat{u}_I}{\partial \xi_i} \frac{\partial \hat{v}}{\partial \xi_j}\right);$$

here $E^0(\varphi)$ is the error functional associated to $I^0(\varphi)$ and satisfying (3.8). The coefficients b_{ij} are linear combinations of terms of the form $\beta \hat{a}_{mn}$, where β satisfies (3.16). So it is sufficient to prove

$$\left|E^0\left(\beta \hat{a}_{mn} \frac{\partial \hat{u}_I}{\partial \xi_i} \frac{\partial \hat{v}}{\partial \xi_j}\right)\right| \leqslant Ch_e^3\|u\|_{3,e}|v|_{1,e},$$

and (4.32) follows immediately.

Set $\sigma = \beta \hat{a}_{mn} \partial \hat{u}_I/\partial \xi_i$. We have

$$E^0\left(\sigma \frac{\partial \hat{v}}{\partial \xi_j}\right) = E^0\left([\sigma - \sigma_I] \frac{\partial \hat{v}}{\partial \xi_j}\right) + E^0\left(\sigma_I \frac{\partial \hat{v}}{\partial \xi_j}\right).$$

We estimate the first term by means of (4.23):

$$\left|E^0\left([\sigma - \sigma_I]\frac{\partial\hat{v}}{\partial\xi_j}\right)\right| \leqslant C|\sigma - \sigma_I|_{2,C_2}|\hat{v}|_{1,C_2} \leqslant C|\sigma|_{3,C_2}|\hat{v}|_{1,C_2}.$$

Now we use (3.8) and get

$$\left|E^0\left(\sigma_I\frac{\partial\hat{v}}{\partial\xi_j}\right)\right| \leqslant C|\sigma_I|_{3,C_2}|\hat{v}|_{3,C_2} \leqslant C\{|\sigma - \sigma_I|_{3,C_2} + |\sigma|_{3,C_2}\}|\hat{v}|_{1,C_2}$$

$$\leqslant C|\sigma|_{3,C_2}|\hat{v}|_{1,C_2}.$$

Thus $|E^0(\sigma\,\partial\hat{v}/\partial\xi_j)| \leqslant C|\sigma|_{3,C_2}|\hat{v}|_{1,C_2}$. Further, from (3.16), (3.2) and $a_{mn} \in C^1(\bar{\Omega}_h)$ it follows

$$\left|\beta\hat{a}_{mn}\frac{\partial\hat{u}_I}{\partial\xi_i}\right|_{3,C_2} \leqslant C\left\{\left|\beta\frac{\partial\hat{u}_I}{\partial\xi_i}\right|_{3,C_2} + h_e\left|\beta\frac{\partial\hat{u}_I}{\partial\xi_i}\right|_{2,C_2} + [h_e|\hat{u}_I|_{1,C_2} + |\hat{u}_I|_{2,C_2}]|\hat{a}_{mn}|_{2,C_2}\right.$$

$$\left. + |\hat{u}_I|_{1,C_2}|\hat{a}_{mn}|_{3,C_2}\right\}$$

$$\leqslant C\{h_e^3|\hat{u}_I|_{1,C_2} + h_e^2|\hat{u}_I|_{2,C_2} + h_e|\hat{u}_I|_{3,C_2}$$

$$+ h_e[h_e|\hat{u}_I|_{1,C_2} + |\hat{u}_I|_{2,C_2}]|a_{mn}|_{2,e} + h_e^2|\hat{u}_I|_{1,C_2}|a_{mn}|_{3,e}\}.$$

We use the bound

$$|\hat{u}_I|_{j,C_2} \leqslant |\hat{u}_I - \hat{u}|_{j,C_2} + |\hat{u}|_{j,C_2} \leqslant C\{|\hat{u}|_{3,C_2} + |\hat{u}|_{j,C_2}\},$$

$j = 1, 2, 3,$ and the fact that, due to (4.3), $u \in C^2(\bar{\Omega}_h)$, and we get the final bound

$$\left|E^0\left(\beta\hat{a}_{mn}\frac{\partial\hat{u}_I}{\partial\xi_i}\frac{\partial\hat{v}}{\partial\xi_j}\right)\right| \leqslant Ch_e^3\|u\|_{4,e}|v|_{1,e}.$$

(c) The main problem is to prove (4.33). We have

$$a(\omega, v) = \sum_e \int_{C_2}\sum_{i,j=1}^{2} b_{ij}\frac{\partial\hat{\omega}}{\partial\xi_i}\frac{\partial\hat{v}}{\partial\xi_j}d\xi$$

(again $\omega = u - u_I$). We must estimate the terms $\int_{C_2} b_{ij}(\partial\hat{\omega}/\partial\xi_i)(\partial\hat{v}/\partial\xi_j)d\xi$. We may restrict ourselves to two cases: $i = j = 1, i = 1, j = 2$. Consider first $\int_{C_2} b_{11}(\partial\hat{\omega}/\partial\xi_1)\cdot(\partial\hat{v}/\partial\xi_1)d\xi$. From (3.16) and $a_{ij} \in C^1(\bar{\Omega}_h)$ it follows $b_{ij} = b_{ij}^0 + O(h_e)$, where b_{ij}^0 denotes again the value of b_{ij} at the center. Hence

$$\int_{C_2} b_{11}\frac{\partial\hat{\omega}}{\partial\xi_1}\frac{\partial\hat{v}}{\partial\xi_1}d\xi = b_{11}^0\int_{C_2}\frac{\partial\hat{\omega}}{\partial\xi_1}\frac{\partial\hat{v}}{\partial\xi_1}d\xi + \int_{C_2} O(h_e)\frac{\partial\hat{\omega}}{\partial\xi_1}\frac{\partial\hat{v}}{\partial\xi_1}d\xi.$$

As $|\hat{\omega}|_{1,C_2} \leqslant C|\hat{u}|_{3,C_2}$, the second term is bounded by $Ch_e|\hat{u}|_{3,C_2}|\hat{v}|_{1,C_2} \leqslant Ch_e^3\|u\|_{3,e}|v|_{1,e}$. To estimate the first term consider the functional $L(\hat{u}) = \int_{C_2}(\partial\hat{\omega}/\partial\xi_1)(\partial\hat{v}/\partial\xi_1)d\xi$. It vanishes for $\hat{u} \in \hat{P}$. If $\hat{u} = \xi_2^3$, then $\hat{\omega} = \xi_2^3 - \xi_2$, $\partial\hat{\omega}/\partial\xi_1 = 0$ and L vanishes. If $\hat{u} = \xi_1^3$, then $\partial\hat{\omega}/\partial\xi_1 = 2((3/2)\xi_1^2 - ½)$, i.e. $\partial\hat{\omega}/\partial\xi_1$ is a multiple of the Legendre polynomial $P_2(\xi_1)$; and as $\partial\hat{v}/\partial\xi_1$ is a linear polynomial of ξ_1 and integration with respect to ξ_1 is done over the interval $\langle -1, 1\rangle$, L vanishes, too. The

Bramble-Hilbert lemma implies $|L(\hat{u})| \leqslant C|\hat{u}|_{4,C_2}|\hat{v}|_{1,C_2} \leqslant Ch_e^3\|u\|_{4,e}|v|_{1,e}$. Hence,

$$\left|\sum_e \int_{C_2} \sum_{i=1}^{2} b_{ii} \frac{\partial\hat{\omega}}{\partial\xi_i}\frac{\partial\hat{v}}{\partial\xi_i}d\xi\right| \leqslant C\sum_e h_e^3\|u\|_{4,e}|v|_{1,e} \leqslant Ch^3\|u\|_{4,\Omega_h}|v|_{1,\Omega_h}.$$

Now let us consider the integral

$$\int_{C_2} b_{12}\frac{\partial\hat{\omega}}{\partial\xi_1}\frac{\partial\hat{v}}{\partial\xi_2}d\xi = b_{12}^0\int_{C_2}\frac{\partial\hat{\omega}}{\partial\xi_1}\frac{\partial\hat{v}}{\partial\xi_2}d\xi + \int_{C_2} O(h_e)\frac{\partial\hat{\omega}}{\partial\xi_1}\frac{\partial\hat{v}}{\partial\xi_2}d\xi.$$

The second term and the sum of these terms are bounded as above. To estimate the first term we introduce the functional

$$L(\hat{u}) = \int_{C_2}\frac{\partial\hat{\omega}}{\partial\xi_1}\frac{\partial\hat{v}}{\partial\xi_2}d\xi - H(\hat{u}),$$

(4.36)

$$H(\hat{u}) = \frac{1}{15}\left\{\int_{-1}^{1}\frac{\partial^2\hat{\omega}(\xi_1, 1)}{\partial\xi_1^2}\frac{\partial\hat{v}(\xi_1, 1)}{\partial\xi_1}d\xi_1 - \int_{-1}^{1}\frac{\partial^2\hat{\omega}(\xi_1, -1)}{\partial\xi_1^2}\frac{\partial\hat{v}(\xi_1, -1)}{\partial\xi_1}d\xi_1\right\}.$$

If $\hat{u} \in \hat{P}$, then $\hat{\omega} = 0$ and $L(\hat{u}) = 0$. If $\hat{u} = \xi_2^3$, then $\partial\hat{\omega}/\partial\xi_1 = 0$ and $L(\hat{u}) = 0$. If $\hat{u} = \xi_1^3$, then $\int_{C_2}(\partial\hat{\omega}/\partial\xi_1)(\partial\hat{v}/\partial\xi_2)d\xi = (16/15)\alpha_7 = H(\hat{u})$ and $L(\hat{u}) = 0$. The Bramble-Hilbert lemma implies $|L(\hat{u})| \leqslant C|\hat{u}|_{4,C_2}|\hat{v}|_{1,C_2} \leqslant Ch_e^3\|u\|_{4,e}|v|_{1,e}$. Thus,

$$\left|\sum_e b_{12}^0\int_{C_2}\frac{\partial\hat{\omega}}{\partial\xi_1}\frac{\partial\hat{v}}{\partial\xi_2}d\xi\right| \leqslant C\sum_e h_e^3\|u\|_{4,e}|v|_{1,e} + \left|\sum_e b_{12}^0 H(\hat{u})\right|$$
$$\leqslant Ch^3\|u\|_{4,\Omega_h}|v|_{1,\Omega_h} + \left|\sum_e b_{12}^0 H(\hat{u})\right|.$$

In the sum $\Sigma_e b_{12}^0 H(\hat{u})$ they appear either integrals over element sides which lie on Γ_h; and because $v|_{\Gamma_h} = 0$, thus $\partial\hat{v}(\xi_1, \pm 1)/\partial\xi_1 = 0$, these integrals vanish. Or they appear couples of integrals over a common side of two adjacent elements taken in opposite directions with integrands which are the same (ω and v are continuous on $\bar{\Omega}_h$, hence they assume the same values on each side). The factors b_{12}^0 need not be the same; however, their difference is $O(h)$ according to Remark 6. Therefore, (we use again the inequality $\int_{\partial C_2}\varphi^2 ds \leqslant C\|\varphi\|_{1,C_2}^2$) by (3.2), (3.1) and (3.6)

$$\left|\sum_e b_{12}^0 H(\hat{u})\right| \leqslant Ch\sum_e\left|\int_{-1}^{1}\frac{\partial^2\hat{\omega}(\xi_1, 1)}{\partial\xi_1^2}\frac{\partial\hat{v}(\xi_1, 1)}{\partial\xi_1}d\xi_1\right|$$
$$\leqslant Ch\sum_e\left\|\frac{\partial^2\hat{\omega}}{\partial\xi_1^2}\frac{\partial\hat{v}}{\partial\xi_1}\right\|_{1,C_2} \leqslant Ch\sum_e\|\hat{\omega}\|_{3,C_2}|\hat{v}|_{1,C_2} \leqslant Ch\sum_e|\hat{u}|_{3,C_2}|\hat{v}|_{1,C_2}$$
$$\leqslant Ch\sum_e h_e^2\|u\|_{3,e}|v|_{1,e} \leqslant Ch^3\|u\|_{3,\Omega_h}|v|_{1,\Omega_h};$$

and the proof of (4.33) is finished.

(d) Let us consider the second case of $I(\varphi)$, namely $I(\varphi)$ is of the form (3.11). Then $a_h(w, v) = a_h^*(w, v) + R_h^1(w, v)$, where

(4.37) $$R_h^1(w, v) = \sum_e E^1\left(\sum_{i,j=1}^{2} b_{ij} \frac{\partial \hat{w}}{\partial \xi_i} \frac{\partial \hat{v}}{\partial \xi_j}\right).$$

We have $a_h(u, v) = (f, v)_{0,\Omega_h} - a(u, v) + a_h(u, v) = (f, v)_{0,\Omega_h} - R_h^*(u, v) + R_h^1(u, v)$, and we easily get

(4.38) $$a_h(u_I - u_h, v) = S_h(v) - R_h^*(u, v) + R_h^1(u_I, v) + a_h^*(u_I - u, v) \quad \forall v \in V_h.$$

If we show that

(4.39) $$|R_h^1(u_I, v)| \leqslant Ch^3 |v|_{1,\Omega_h} \quad \forall v \in V_h,$$

then using the arguments of part (a) of the proof we come to (4.29). The proof of (4.39) is the same as that of (4.32) because $R_h^1(u_I, v)$ and $R_h^0(u_I, v)$ have the same form (see (4.35) and (4.37)) and $E^1(\varphi)$ has the same property as $E^0(\varphi)$, namely it satisfies (3.8).

5. **Superconvergence in Three Dimensions.** A three-dimensional isoparametric quadratic element of the Serpendipity family has 20 nodes corresponding to 20 nodes of the three-dimensional cube C_3: $-1 \leqslant \xi_i \leqslant 1$, $i = 1, 2, 3$. The nodes of C_3 are vertices of C_3 and midpoints of sides. The space $\hat{P}$ consists of incomplete quartic polynomials of the form

(5.1) $$\begin{aligned} &\alpha_1 + \alpha_2\xi_1 + \alpha_3\xi_2 + \alpha_4\xi_3 + \alpha_5\xi_1^2 + \alpha_6\xi_1\xi_2 + \alpha_7\xi_1\xi_3 + \alpha_8\xi_2^2 + \alpha_9\xi_2\xi_3 \\ &+ \alpha_{10}\xi_3^2 + \alpha_{11}\xi_1^2\xi_2 + \alpha_{12}\xi_1^2\xi_3 + \alpha_{13}\xi_1\xi_2^2 + \alpha_{14}\xi_1\xi_3^2 + \alpha_{15}\xi_2^2\xi_3 \\ &+ \alpha_{16}\xi_2\xi_3^2 + \alpha_{17}\xi_1\xi_2\xi_3 + \alpha_{18}\xi_1^2\xi_2\xi_3 + \alpha_{19}\xi_1\xi_2^2\xi_3 + \alpha_{20}\xi_1\xi_2\xi_3^2. \end{aligned}$$

$\hat{P}$ satisfies again (2.6). The functions $N_j(\xi_1, \xi_2, \xi_3)$ can be found in [8, p. 121]. The definition of a k-strongly regular partition is the same as in two dimensions with one exception: instead of (2.9) we have to require

(5.2) $$c_2^{-1} h_e^3 \leqslant |J_e| \leqslant c_2 h_e^3.$$

With exception of these changes the definition of V_h is the same as before. The definition of u_h^* is $a_h^*(u_h^*, v) = f_h^*(v) \, \forall v \in V_h$, where

(5.3) $$\begin{aligned} a_h^*(w, v) &= \sum_e \sum_{r=1}^{8} J_e(Q_r^*) \sum_{i,j=1}^{3} \hat{a}_{ij} \widehat{\frac{\partial w}{\partial x_i}}(Q_r^*) \widehat{\frac{\partial v}{\partial x_j}}(Q_r^*), \\ f_h^*(v) &= \sum_e \sum_{r=1}^{8} J_e(Q_r^*) \hat{f}(Q_r^*) \hat{v}(Q_r^*), \\ Q_r^* &= \left(\pm\frac{\sqrt{3}}{3}, \pm\frac{\sqrt{3}}{3}, \pm\frac{\sqrt{3}}{3}\right). \end{aligned}$$

The same superconvergence theorem as in two dimensions is true with one change concerning the assumptions. We have to assume that the partitions are 3-strongly regular. The reason is the following: in two dimensions the derivatives $D^\alpha x_i^e$ of 2-

strongly partitions are bounded by $h_e^{|\alpha|}$ not only for $|\alpha| \leq 3$ but for all α because x_i^e are cubic polynomials. We need this fact to prove (3.16). In three dimensions x_i^e are quartic polynomials, so we have to assume the 3-strong regularity. The proof of the theorem is similar to the proof of Theorem 4.1, and we leave it out.

6. Numerical Results and Application of Superconvergence in Practical Computations. The following problem was solved:*

$$(6.1)\quad \begin{aligned} -\Delta u &= -2y + 54xy - 12xy^2 + 16y^2 - 14y^3 - 4x^3 - 12x + 16x^2 - 42x^2y \text{ in } \Omega, \\ u|_\Gamma &= 0, \quad \Omega\colon\ 0 < x < 1, 0 < y < 1. \end{aligned}$$

The exact solution is $u(x, y) = x(1 - x)y(1 - y)(1 + 2x + 7y)$. We used partitions consisting of square elements with vertices $\{(ih, jh)\}_{i,j=0}^{M}$, $M = h^{-1}$, $h = 1/4, 1/5, 1/6, 1/7, 1/8$ and Gauss' 2×2 formula. The norm $\|u - u_h^*\|_h$ is denoted by E_G and is equal in this case to

$$(6.2)\qquad E_G = \left\{ N_G^{-1} \sum_{P \in G} \left[\left(\frac{\partial(u - u_h^*)(P)}{\partial x} \right)^2 + \left(\frac{\partial(u - u_h^*)(P)}{\partial y} \right)^2 \right] \right\}^{1/2}.$$

Here $N_G = 4h^{-2}$ is the number of Gaussian points. Also, the gradient at vertices of square elements was computed (the unique values of the gradient were won by averaging); and, as a measure of the error, the number

$$(6.3)\qquad E_V = \left\{ N_V^{-1} \sum_{P \in V} \left[\left(\frac{\partial(u - u_h^*)(P)}{\partial x} \right)^2 + \left(\frac{\partial(u - u_h^*)(P)}{\partial y} \right)^2 \right] \right\}^{1/2}$$

is taken. The set V consists of all vertices of square elements with exception of the vertices of Ω. Table 1 shows on one hand the big difference between the magnitudes of E_G and E_V and the superconvergence at Gaussian points; on the other hand it shows that E_V goes to zero just as fast as h^2, i.e. $h^{-2}E_V \longrightarrow \text{const} > 0$.

TABLE 1

h	$E_G \times 10^3$	$E_V \times 10^3$	$h^{-3}E_G$	$h^{-2}E_V$
$\frac{1}{4}$	9.0	41	0.57	0.65
$\frac{1}{5}$	4.2	25	0.52	0.64
$\frac{1}{6}$	2.2	18	0.49	0.63
$\frac{1}{7}$	1.3	13	0.46	0.63
$\frac{1}{8}$	0.85	10	0.43	0.63

*The author is indebted to M. Kovaříková who carried out all computations on the computer DATASAAB D21.

Gauss' 3×3 formula and Čebyšev's product formula with nine points were also applied. The values E_G and E_V differ less than 0.2% from values given in Table 1 (let us emphasize that whatever formula is applied, the set G is the set of maps of points Q_r^* $(r = 1, \ldots, 4)$ of Gauss' 2×2 formula).

The problem (6.1) was solved by curved elements not satisfying (2.8). The square elements were distorted into curved ones in that the midpoints of two sides of each element were moved in the x-direction and y-direction, respectively. The length of these displacements was always the same: $\frac{1}{2} h^2$. Such elements do not satisfy (2.8) as $\partial^3 x_1^e/\partial\xi_1 \partial\xi_2^2$ and $\partial^3 x_2^e/\partial\xi_1^2 \partial\xi_2$ are in absolute value equal to $\frac{1}{2}h^2$. The number E_G is not equal to $\|u - u_h^*\|_h$; however, it is an equivalent norm; in addition $E_G^{-1} \|u - u_h^*\|_h = 1 + O(h)$. Table 2 indicates convincingly that $E_G \geqslant ch^2$, $c > 0$, i.e. there is no superconvergence. Nevertheless, E_G is still substantially smaller than E_V.

TABLE 2

h	$E_G \times 10^2$	$h^{-2}E_G$	$E_V \times 10^2$
$\frac{1}{4}$	4.4	0.70	15
$\frac{1}{5}$	3.1	0.76	5.7
$\frac{1}{6}$	2.3	0.83	5.9
$\frac{1}{7}$	2.0	0.97	3.9
$\frac{1}{8}$	1.3	0.80	3.4

Gauss' 3×3 formula gives values of E_G which differ less than 4% from values given in Table 2.

In general, if we compute the gradient at Gaussian points we always can expect much more accurate values than at vertices. If a greater part of the elements differ little from parallelepipeds we reach even a greater improvement of accuracy. The question is what integration formula to choose. Theorems 4.1 and 5.1 and Remark 5 show that very often Gauss' 2×2 or $2 \times 2 \times 2$ formula can be sufficient. However, Gauss' 3×3 or $3 \times 3 \times 3$ formula guarantees that, both in case of superconvergence as well as in case that superconvergence does not set in (see [4, pp. 462–463]), we retain the highest order of accuracy which is possible.

The usual requirement of users of finite element codes is to get the values of gradient at vertices of elements. These values must be interpolated from values at Gaussian points. If there is no superconvergence, i.e. the rate of convergence in the $\|\cdot\|_h$-norm is $O(h^2)$ and not better, then interpolation from four Gaussian points on each element by a linear isoparametric shape function is sufficient. Evidently, such interpolation would make worse the accuracy won in case of superconvergence. Therefore, a better

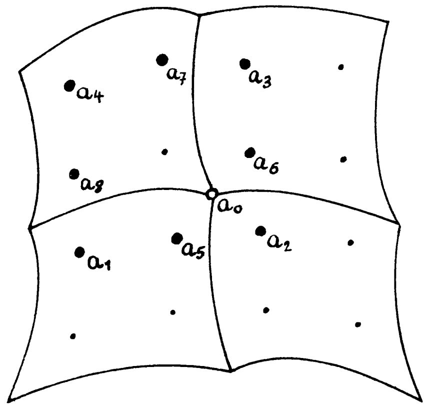

FIGURE 1

way is to use the quadratic isoparametric shape function. We choose eight Gaussian points $a_j(x_1^j, x_2^j)$, $j = 1, \ldots, 8$ (see Figure 1). The interpolation of the partial derivatives $\partial u_h/\partial x_i$ at the point $a_0 = (x_1^0, x_2^0)$ is done by the formula

$$\sum_{j=1}^{8} \frac{\partial u_h(x_1^j, x_2^j)}{\partial x_i} N_j(\xi_1^0, \xi_2^0).$$

The coordinates ξ_1^0, ξ_2^0 are computed by solving the system of two nonlinear equations $x_i^0 = \Sigma_{j=1}^8 x_i^j N_j(\xi_1, \xi_2)$, $i = 1, 2$, by Newton's method. As the initial guess, we choose the point (0, 0). The convergence is very fast, and it is entirely sufficient to stop after three iterations.

Computing Center of the Technical University
Obránců míru 21
602 00 Brno, Czechoslovakia

1. J. BARLOW, "Optimal stress locations in finite element models," *Internat. J. Numer. Methods.*, v. 10, 1976, pp. 243–251.

2. 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.

3. J. H. BRAMBLE & S. R. HILBERT, "Bounds for a class of linear functionals with applications to Hermite interpolation," *Numer. Math.*, v. 16, 1971, pp. 362–369.

4. P. G. CIARLET & P. A. RAVIART, "The combined effect of curved boundaries and numerical integration in isoparametric finite element methods," *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations* (A. K. Aziz, Editor), Academic Press, New York, 1972, pp. 409–474.

5. B. M. IRONS & A. RAZZAQUE; "Experience with the patch test for convergence of finite elements," *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations* (A. K. Aziz, Editor), Academic Press, New York, 1972, pp. 557–587.

6. J. NEČAS, *Les Méthodes Directes en Théorie des Equations Elliptiques*, Academia, Prague, 1967.

7. D. A. VERYARD, *Problems Associated with the Convergence of Isoparametric and Mixoparametric Finite Elements*, M. Sc. Thesis, University of Wales, 1971.

8. O. C. ZIENKIEWICZ, *The Finite Element Method in Engineering Science*, McGraw-Hill, London, 1972.

9. M. ZLÁMAL, "Some superconvergence results in the finite element method," *Mathematical Aspects of Finite Element Methods*, Springer-Verlag, Berlin, Heidelberg, New York, 1977, pp. 351–362.