IMA Journal of Numerical Analysis (1985) 5, 407–427

Superconvergent Recovery of the Gradient from Piecewise Linear Finite-element Approximations

NICK LEVINE

Department of Mathematics, University of Reading, Whiteknights, Reading RG6 2AX

[Received 27 January 1984 and in revised form 14 March 1985]

We propose and justify the use of a simple scheme which recovers gradients from the piecewise linear finite-element approximation on triangular elements to the solution of a second-order elliptic problem. The recovered gradient is a superconvergent estimate of the true gradient at the midpoints of element edges. A related scheme recovers the gradient at the element centroids.

1. Introduction

FOR QUADRILATERAL ELEMENTS, gradient superconvergence has been well established since Veryard (1971) improved the accuracy of the gradients of biquadratic Galerkin approximations by sampling them at the second-order Gauss points in each element. Such points of exceptional accuracy of derivatives of finite element approximations have come to be known as "stress points" and their existence as an example of the phenomenon of "superconvergence"; we associate these terms here with the sampling or recovery of gradients to one order of accuracy higher than is globally possible.

Stress points are located (Strang & Fix, 1973, pp. 168–169; Barlow, 1976) by the property that the derivative of the polynomial which dominates the error expansion coincides with its approximation (i.e. the derivative of a lower-degree polynomial). This idea is at the heart of any superconvergence result, for it leads us directly to the stress points of the "unknown" function's interpolant. Further, the gradient of the interpolant is at all points a superconvergent approximation to the gradient of the finite-element solution. Therefore the stress points for these two functions are identical. (For proofs of this result on quadrilateral elements, see Zlámal, 1977, 1978 and Lesaint & Zlámal, 1979.)

It has been suggested (e.g. Moan, 1974) that the Galerkin least-squares approximation to gradients is "almost local" and can therefore be analysed in one element in complete isolation from all others. Although this reasoning is fortuitously successful for quadrilaterals it fails on linear triangles, for it implies that their centroids are stress points. On the other hand the interpolant method given above predicts for these elements that "midpoints of an (element) edge seem ... to be exceptional for derivatives along the edge but not for stresses in the direction of the normal" (Strang & Fix, 1973, p. 169).

In this paper we consider piecewise linear approximations on triangular elements to a model Dirichlet problem. We prove (in Section 3) that element edge midpoints

0272-4979/85/040407+21 \$03.00/0

are indeed tangential derivative stress points for the interpolant and complete the above outline by returning to the finite-element method in Sections 4 and 5. We must, however, impose the following condition on the mesh: the nodes are given by a smooth transformation of a square grid and the triangulation is topologically equivalent to a triangulation on this grid by identical triangles, arranged so that six elements surround each internal node. The last requirement leads to the cancellation of error contributions between neighbouring elements; it demonstrates the non-local nature of the Galerkin approximation. In Sections 3–5 we simplify the argument somewhat by restricting the mesh further. We take up the more general case in Section 6 and predict an alternative superconvergence result which may hold even when the domain of the problem is incompatible with the more relaxed conditions. We do not prove this last result here, but give a numerical demonstration of this and other aspects of superconvergence in Section 7.

Separate components of a vector at different points are not what is usually required: we have established superconvergence only for the stress component tangential to the edge on which the sampling point lies. However, if we average the approximate gradient (a piecewise constant vector) between the two elements neighbouring this point, then the interpolant method above indicates that this "recovered" gradient is a superconvergent approximation to both components of the derivative at the stress point. We prove this in Section 3; Lin Qun, Lu Tao & Shen Shumin (1983) and Křížek & Neittaanmäki (1983) have obtained restricted forms of this result. There is again a geometrical condition on the mesh: it must obey the restrictions already imposed in Sections 3–5. (For the more general meshes of Section 6 we must modify the recovery scheme.) Finally, we can recover the gradient at an element centroid by averaging recovered values from the three stress points for that element; in Section 7 we present a comparison of this scheme with the corresponding superconvergence result on bilinear elements (Lesaint & Zlámal, 1979).

2. Preliminaries and Triangulation

The results of this paper are presented in the context of Sobolev spaces; we introduce here the relevant notation and a key lemma. We work only with bounded open regions in $\mathbb{R}^2$ which have the strong cone property (see e.g. Bramble & Hilbert, 1970). Let R be such a region: typically this will be either the problem domain Ω or a small patch of elements. We denote by $W_p^m(R)$ $(m = 0, 1, \ldots)$ the Sobolev space of functions which together with their generalized derivatives up to order m inclusive are in $L_p(R)$. The norm and seminorm are given by

$$\|w\|_{W_p^m(R)} = \left(\iint_R \sum_{|\alpha| \leq m} |D^\alpha w|^p\right)^{1/p},$$

$$|w|_{W_p^m(R)} = \left(\iint_R \sum_{|\alpha| = m} |D^\alpha w|^p\right)^{1/p},$$

respectively for $p < \infty$, with the usual modification when $p = \infty$. For the most part we take $p = 2$ and write H^m, $\|\cdot\|_{m,R}$ and $|\cdot|_{m,R}$ for W_2^m, $\|\cdot\|_{W_2^m(R)}$ and $|\cdot|_{W_2^m(R)}$.

In all that follows the letter c stands for a generic, positive number, different at each appearance but "constant" in that it is independent of the functions denoted below by f, u, v or w, the element(s) under consideration and the discretization parameter h.

We will make frequent use of the Bramble–Hilbert lemma. Based on Taylor's theorem, this is a multi-dimensional non-constructive generalization of the Peano Kernel theorem. We give here a simplified form; see Bramble & Hilbert (1970) for the full result and proof.

LEMMA 2.1 *Let F be a linear functional on $W_p^m(R)$ such that*

(i) $|F(w)| \leqslant c\|w\|_{W_p^m(R)}$, $\forall\ w \in W_p^m(R)$;

(ii) $F(w) = 0$ *if w is a polynomial on R of degree less than m.*

Then $|F(w)| \leqslant c|w|_{W_p^m(R)}$, $\forall\ w \in W_p^m(R)$; *the constant c depends on F, p, m and* diam (R) *only.*

We now describe a particular triangulation for which our results hold; the general form is given in Section 6. Let Ω be a bounded open domain in $\mathbf{R}^2$ with the strong cone property in which (x, y) are rectangular Cartesian co-ordinates. For decreasing values of the parameter h we triangulate Ω in the following way. We choose a pair of functions (X, Y) of (x, y) (and, if necessary, of h) which can be used as a co-ordinate system on Ω and its neighbourhood. (For example, see (7.1) and Fig. 11 below.) We require that (X, Y) be smooth in the sense that, uniformly in h, the global mapping $(x, y) \to (X, Y)$ is a W_∞^2-diffeomorphism with Jacobian satisfying

$$c^{-1} \leqslant |\partial(X, Y)/\partial(x, y)| \leqslant c. \tag{2.1}$$

Note that the two norms $\|\cdot\|_{3,\Omega}$ based on (x, y)- and (X, Y)-derivatives are not equivalent. We will make no use of (X, Y)-norms.

We place grid-points on Ω so as to give a uniform, square grid in the (X, Y) plane with mesh size h. We triangulate the region in the (X, Y) plane by means of horizontals, verticals and diagonals of slope $+1$ between the grid points and then in the (x, y) plane with straight lines topologically corresponding to the (X, Y) links. When we refer to elements we will mean the (non-curved) triangles in the (x, y) plane; we call the union of elements Ω_h. (See Fig. 1.)

We require the triangulation to approximate the boundary $\partial\Omega$ of Ω well, in that all the nodes on $\partial\Omega_h$ lie on $\partial\Omega$ and the intersection of each element with Ω contains an open disc of diameter $\geqslant ch$ which itself contains the centroid of that element. (It is clear that triangulation functions (X, Y) satisfying all the above conditions do not exist for general regions Ω; we discuss this problem in Section 6.)

We denote the elements by T_k $(k = 1, \ldots, K)$. For each k, let $N_{k0} = (X_k, Y_k)$ and $N_{k1} = (X_k + h, Y_k)$ be the nodes with the same Y co-ordinate (see Fig. 2); let N_{k2} be the third node. We introduce local co-ordinates (ξ_k, η_k) by means of a linear transformation t_k of (x, y) which maps T_k onto the triangle τ_k and the nodes N_{k0}, N_{k1}, N_{k2} to $(0, 0)$, $(1, 0)$, (ξ_{k2}, η_{k2}) respectively such that

$$|\xi_{k2}| \leqslant c \quad \text{and} \quad c^{-1} \leqslant h^2|\partial(\xi_k, \eta_k)/\partial(x, y)| \leqslant c. \tag{2.2}$$

We adopt the notation that functions of (ξ_k, η_k) are distinguished by a hat from their

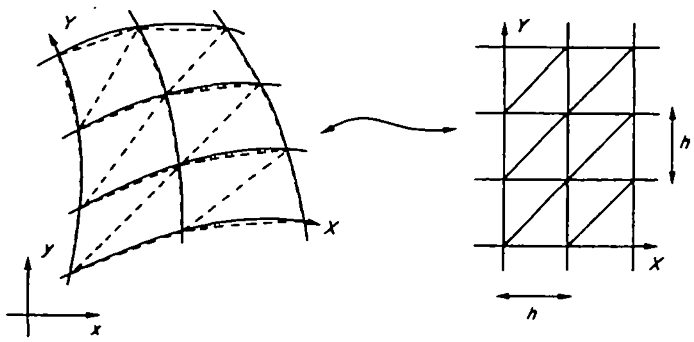

FIG. 1. The global transformation $(x, y) \leftrightarrow (X, Y)$. The solid lines indicate (X, Y) contours, elements (in the (x, y) plane) are shown by broken lines.

counterpart functions of (x, y) or (X, Y). Thus

$$\hat{w}(\xi_k, \eta_k) = w(t_k^{-1}(\xi_k, \eta_k)),$$

etc.

LEMMA 2.2 *Let ∇_m denote the tensor of mth derivatives with respect to (X, Y) and ∇ denote gradient with respect to (x, y). The following estimates hold:*

(i) $c^{-1} \leqslant h^{-2}$ meas $(T_k) \leqslant c$, $c^{-1} \leqslant$ meas $(\tau_k) \leqslant c$;

(ii) *each element of the matrix $\partial(x, y)/\partial(\xi_k, \eta_k)$ is bounded in modulus by ch;*

(iii) $|\nabla \xi| \leqslant ch^{-1}$, $|\nabla_m \xi| \leqslant ch^{-1}$ $(m = 1, 2)$, *similarly for derivatives of* η;

(iv) $|\hat{w}|_{m, \tau_k} \leqslant ch^{m-1} \|w\|_{m, T_k}$, $\forall\, w \in H^m(T_k)$ $(m = 0, \ldots, 3)$.

Proofs. These follow from (2.1), (2.2) and the linearity of the t_k. For (iii) we use

$$\partial \xi_k/\partial x = \partial y/\partial \eta_k \cdot \det\,(\partial(\xi_k, \eta_k)/\partial(x, y)),$$

etc.; the ∇_m bounds follow by the chain rule (and the linearity of the t_k).

There usually exists another triangle, $T_{k'}$, with the nodes N_{k0} and N_{k1} in common with T_k. We will map $T_{k'}$ into the (ξ_k, η_k) plane with the same transformation (i.e. t_k) as T_k and refer to the quadrilaterals $T_k \cup T_{k'}$ and $\tau_k \cup \tau_{k'}$ as A_k and α_k (see Fig. 3). If such a $T_{k'}$ does not exist, then N_{k0} and N_{k1} lie on $\partial\Omega_h$ and we denote T_k by B_k (see

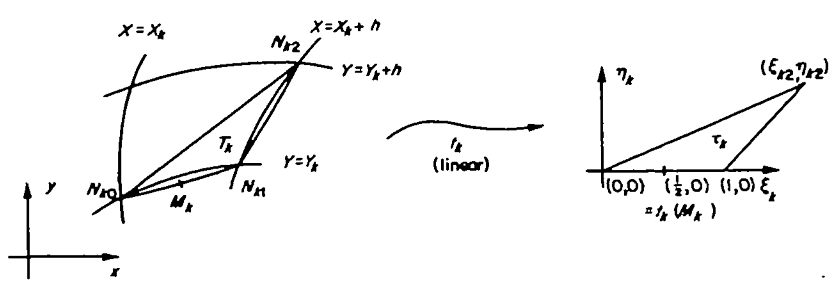

FIG. 2. The local transformation of a triangle.

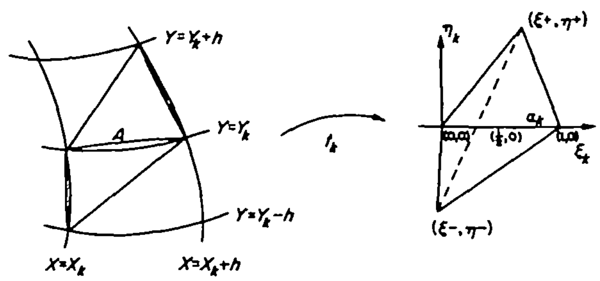

Fig. 3. The transformation of a triangle pair $A \to \alpha$. The midpoints of the diagonals of α are at $(\frac{1}{2}, 0)$ and $(\frac{1}{2}(\xi_+ + \xi_-), \frac{1}{2}(\eta_+ + \eta_-))$; their separation is $O(h)$.

Fig. 4). In other words we have the decompositions

$$\Omega_h = \bigcup_{k=1}^{K} T_k = \left(\bigcup_{k=1}^{K_A} A_k\right) \cup \left(\bigcup_{k=1}^{K_B} B_k\right).$$

In each T_k, A_k or B_k, we refer to the midpoint of the edge $N_{k0}N_{k1}$ as M_k (i.e. M_k is the point $t_k^{-1}(\frac{1}{2}, 0)$). (See Fig. 2.)

We now show that each α_k is close to a parallelogram, in the sense that the midpoints of its diagonals have separation $\leqslant ch$. This lemma is the link between the triangulation conditions and the superconvergence results that follow.

LEMMA 2.3 *Let $k \leqslant K_A$ be fixed (for convenience we drop the subscript k) and consider the quadrilateral α_k with vertices labelled as in Fig. 3. Then*

$$|\xi_+ + \xi_- - 1| \leqslant ch \quad \text{and} \quad |\eta_+ + \eta_-| \leqslant ch.$$

Proof. We can view ξ_k as a twice differentiable (i.e. W^2_∞) function of (X, Y) and consider the functional

$$\xi(X_k + h, Y_k + h) + \xi(X_k, Y_k - h) - \xi(X_k + h, Y_k) - \xi(X_k, Y_k).$$

This vanishes for linear ξ_k and so the first bound follows from Lemma 2.1 (or Taylor's theorem!) and Lemma 2.2; the second is derived identically.

3. Recovering the Gradient from the Interpolant

We are now in a position to define interpolation on the mesh and derive some error bounds. Let $S^h(\Omega_h) \subset H^1(\Omega_h)$ be the space of continuous piecewise linears on the triangulation of Ω_h. Let u be any function satisfying $u \in H^3(\Omega_h)$ (so that u and ∇u

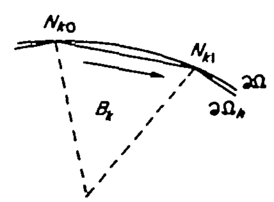

FIG. 4. A boundary triangle B_k. The arrow gives the direction of $\partial/\partial\xi_k$.

are defined everywhere in Ω_k). Let $u_I \in S^h(\Omega_k)$ interpolate the values of u at all the nodes of Ω_k. It is well known that $|\nabla(u_I-u)| = O(h)$; since u can (just) fail to be in W^2_∞ the result takes this form.

LEMMA 3.1 *Let Q be a point in T_k and let $[\cdot]_Q$ stand for point sampling at Q. Then*

$$|[\nabla(u_I-u)]_Q| \leqslant c\|u\|_{3,T_k}.$$

Proof. Let $\mathbf{F}(u) = [\nabla(u_I-u)]_Q$ and $\hat{\mathbf{F}}(\hat{u})$ be the corresponding functional-pair on $H^3(\tau_k)$. When u is linear on T_k, $u_I = u$ and so by Lemma 2.2, the Sobolev lemma and Lemma 2.1,

$$|\mathbf{F}(u)| \leqslant ch^{-1}|\hat{\mathbf{F}}(\hat{u})| \leqslant ch^{-1}|\hat{u}|_{W^1_{2+\varepsilon}(\tau_k)}$$

for any fixed $\varepsilon > 0$. By the Sobolev lemma again,

$$|\mathbf{F}(u)| \leqslant ch^{-1}(|\hat{u}|_{2,\tau_k} + |\hat{u}|_{3,\tau_k});$$

the result now follows from Lemma 2.2.

Although we use this lemma later, our aim in this section is to suppose we know u_I and obtain from it values of ∇u, correct to $O(h^2)$. To estimate both components of this gradient at a single point, we will use the recovery scheme introduced in section 1. Note that if $v \in S^h(\Omega_k)$, then

$$\partial\hat{v}/\partial\xi_k \text{ is constant over } \alpha_k; \tag{3.1}$$

this is a constant multiple of the derivative of v in the direction of the edge common to elements T_k and $T_{k'}$. So, for this component of the gradient, our scheme is equivalent to point sampling at M_k, the midpoint of this shared edge.

For each A_k, we define the recovery operator $\mathbf{D}_k$ on $S^h(\Omega_k)$ by

$$\mathbf{D}_k(v) = \tfrac{1}{2}([\nabla v]_{T_k} + [\nabla v]_{T_{k'}}). \tag{3.2}$$

LEMMA 3.2 *Let $k \leqslant K_A$ be fixed (for convenience we drop the subscript k). Then*

$$|\mathbf{D}u_I - [\nabla u]_M| \leqslant ch\|u\|_{3,A}.$$

Proof. Let $F_1(u)$ and $F_2(u)$ be the tangential (ξ) and normal (η) components, respectively, of the recovery error $\mathbf{D}u_I - [\nabla u]_M$. Then

$$\hat{F}_1(\hat{u}) = [\partial\hat{u}_I/\partial\xi]_\alpha - [\partial\hat{u}/\partial\xi]_{(\frac{1}{2},0)}$$

and

$$\hat{F}_2(\hat{u}) = \tfrac{1}{2}([\partial\hat{u}_I/\partial\eta]_{\tau_+} + [\partial\hat{u}_I/\partial\eta]_{\tau_-}) - [\partial\hat{u}/\partial\eta]_{(\frac{1}{2},0)},$$

where τ_+, τ_- are the triangles above and below the ξ-axis which comprise α (with vertices labelled as in Fig. 3).

Now, if

$$\left.\begin{aligned} &\hat{u} = \xi^2 \text{ in } \alpha, &&\text{then} \quad \hat{u}_I = \xi + \eta(\xi^2_\pm - \xi_\pm)/\eta_\pm \text{ in } \tau_\pm;\\ \text{if}\quad &\hat{u} = \xi\eta \text{ in } \alpha, &&\text{then} \quad \hat{u}_I = \eta\xi_\pm \text{ in } \tau_\pm;\\ \text{if}\quad &\hat{u} = \eta^2 \text{ in } \alpha, &&\text{then} \quad \hat{u}_I = \eta\eta_\pm \text{ in } \tau_\pm. \end{aligned}\right\} \tag{3.3}$$

Also, $\hat{u}_I = \hat{u}$ for all linear $\hat{u}$ on α. So, without reference to the mesh geometry, $\hat{F}_1$ vanishes for a basis of quadratics on α (and hence for all quadratics). Therefore, by the Sobolev lemma and Lemma 2.1,

$$|\hat{F}_1(\hat{u})| \leqslant c|\hat{u}|_{3,\alpha}. \tag{3.4}$$

Unfortunately F_2 is less easy to bound. From (3.3) we have

$$\begin{aligned}\hat{F}_2(\xi^2) &= \tfrac{1}{2}((\xi_+^2-\xi_+)/\eta_+ + (\xi_-^2-\xi_-)/\eta_-),\\ &= \tfrac{1}{2}(\xi_+ + \xi_- - 1)(\xi_+ - \xi_-)/\eta_+ +\\ &\quad \tfrac{1}{2}(\eta_+ + \eta_-)\xi_-(\xi_- - 1)/\eta_+\eta_-\end{aligned} \tag{3.5}$$

and

$$\hat{F}_2(\xi\eta) = \tfrac{1}{2}(\xi_+ + \xi_- - 1)$$
$$\hat{F}_2(\eta^2) = \tfrac{1}{2}(\eta_+ + \eta_-).$$

If α (or equivalently A) were a parallelogram, then $\hat{F}_2$ would vanish for quadratic $\hat{u}$ and be bounded exactly as $\hat{F}_1$. However, α is only close to a parallelogram and $\hat{F}_2$ is as a result only close to a functional—$\hat{F}_2(\hat{u}-\hat{R}\hat{u})$ below—which vanishes for quadratic $\hat{u}$.

To be precise, let $\hat{R}$ be the projection given by

$$\hat{R}\hat{u} = \frac{1}{2m(\alpha)}\cdot\left(\xi^2\iint_\alpha \frac{\partial^2\hat{u}}{\partial\xi^2} + 2\xi\eta\iint_\alpha \frac{\partial^2\hat{u}}{\partial\xi\,\partial\eta} + \eta^2\iint_\alpha \frac{\partial^2\hat{u}}{\partial\eta^2}\right), \tag{3.6}$$

where $m(\alpha)$ is the measure of α. Then, by Lemmas 2.2 and 2.3, (2.2) and (3.5),

$$\begin{aligned}|\hat{F}_2(\hat{R}\hat{u})| &\leqslant \frac{c}{m(\alpha)}\cdot\left(|\hat{F}_2(\xi^2)|\cdot\left|\iint_\alpha\frac{\partial^2\hat{u}}{\partial\xi^2}\right| + |\hat{F}_2(\xi\eta)|\cdot\left|\iint_\alpha\frac{\partial^2\hat{u}}{\partial\xi\,\partial\eta}\right| + |\hat{F}_2(\eta^2)|\cdot\left|\iint_\alpha\frac{\partial^2\hat{u}}{\partial\eta^2}\right|\right)\\ &\leqslant ch|\hat{u}|_{2,\alpha}.\end{aligned} \tag{3.7}$$

Also, by the Sobolev lemma and (3.7),

$$\begin{aligned}|\hat{F}_2(\hat{u}-\hat{R}\hat{u})| &\leqslant |\hat{F}_2(\hat{u})| + |\hat{F}_2(\hat{R}\hat{u})|\\ &\leqslant c(\|\hat{u}\|_{3,\alpha} + h|\hat{u}|_{2,\alpha})\\ &\leqslant c\|\hat{u}\|_{3,\alpha}.\end{aligned} \tag{3.8}$$

But if $\hat{u}$ is linear, $\hat{R}\hat{u}=0$ and $\hat{u}_I=\hat{u}$ (so that $\hat{F}_2(\hat{u}-\hat{R}\hat{u})=0$); if $\hat{u}$ is one of ξ^2, $\xi\eta$, η^2, then $\hat{R}\hat{u}=\hat{u}$ and so this functional vanishes for all quadratic $\hat{u}$.

Thus by (3.8) and Lemma 2.1

$$|\hat{F}_2(\hat{u}-\hat{R}\hat{u})| \leqslant c|\hat{u}|_{3,\alpha},$$

whence by (3.7)

$$|\hat{F}_2(\hat{u})| \leqslant c(|\hat{u}|_{3,\alpha} + h|\hat{u}|_{2,\alpha}). \tag{3.9}$$

Finally, (3.4), (3.9) and Lemma 2.2 give

$$\begin{aligned}|\mathbf{D}u_I - [\nabla u]_M| &\leqslant ch^{-1}(|\hat{F}_1(\hat{u})| + |\hat{F}_2(\hat{u})|)\\ &\leqslant c(h^{-1}|\hat{u}|_{3,\alpha} + |\hat{u}|_{2,\alpha})\\ &\leqslant ch\|u\|_{3,A} \quad \text{as desired.}\end{aligned}$$

We now turn to the central question of this paper. Given $u_h \in S^h(\Omega_h)$ (a finite element approximation to an unknown function u), how do we estimate ∇u? The answer is that since $\nabla(u_h - u_I)$ is constant over each element, we should use the same recovery procedure with u_h as with u_I.

THEOREM 3.1 *Let $u \in H^3(\Omega_h)$, $u_I \in S^h(\Omega_h)$ and $\mathbf{D}_k$ $(k = 1, \ldots, K_A)$ be as above and let u_h be any member of $S^h(\Omega_h)$. Then*

$$h\left(\sum_{k=1}^{K_A} |\mathbf{D}_k u_h - [\nabla u]_{M_k}|^2\right)^{\frac{1}{2}} \leqslant c(|u_I - u_h|_{1,\Omega_h} + h^2\|u\|_{3,\Omega_h}).$$

Remark. Although this result only bounds recovered derivatives on element edges between nodes with the same Y co-ordinate (see Fig. 3), it is clear that edges linking nodes with the same value of X (or $Y - X$) can be included in the average. So Theorem 3.1 states that the mean-square average error of the recovered gradient over all internal edges is bounded by $|u_I - u_h|_{1,\Omega_h} + O(h^2)$. (The average can also include tangential derivatives at midpoints of edges on $\partial\Omega_h$.) This comment also applies to Theorems 4.1 and 5.2 below.

Proof. (Zlámal, 1977, has proved a similar result.)

By Lemma 2.2 the operators $\mathbf{D}_k$ are bounded thus:

$$|\mathbf{D}_k(v)| \leqslant ch^{-1}|v|_{1,A_k}, \quad \forall\, v \in S^h.$$

Squaring and summing,

$$\sum_k |\mathbf{D}_k(v)|^2 \leqslant ch^{-2}\sum_k |v|^2_{1,A_k} \leqslant ch^{-2}|v|^2_{1,\Omega_h}, \quad \forall\, v \in S^h.$$

Squaring and summing the result of Lemma 3.2,

$$\sum_k |\mathbf{D}_k u_I - [\nabla u]_{M_k}|^2 \leqslant ch^2\|u\|^2_{3,\Omega_h}.$$

Then, setting $v = u_h - u_I$ in the above,

$$\sum_k |\mathbf{D}_k u_h - [\nabla u]_{M_k}|^2 \leqslant ch^{-2}|u_I - u_h|^2_{1,\Omega_h} + ch^2\|u\|^2_{3,\Omega_h},$$

whence the result as required.

4. The Finite-element Approximation and Numerical Quadrature

From now on, we take u to be the (unknown) solution to the model problem:

$$\left.\begin{aligned} Lu &= f \text{ in } \Omega, \qquad u = g \text{ on } \partial\Omega \\ Lu &= -\frac{\partial}{\partial x}\left(a_{11}\frac{\partial u}{\partial x} + a_{12}\frac{\partial u}{\partial y}\right) - \frac{\partial}{\partial y}\left(a_{21}\frac{\partial u}{\partial x} + a_{22}\frac{\partial u}{\partial y}\right)\end{aligned}\right\} \tag{4.1}$$

where (the second line above)

satisfies the classical ellipticity condition and $a_{12} = a_{21}$. (We could add the term $a_0 u$—with $a_0 \geqslant 0$, $a_0 \in H^2(\Omega)$—to L with a straightforward supplementary analysis.) The goal of the next two sections is to apply Theorem 3.1 to a finite element approximation u_h to u by showing that $|u_I - u_h|_{1,\Omega_h} = O(h^2)$.

We associate with (4.1) the bilinear form on $[H^1(\Omega)]^2$:

$$a_\Omega(w, v) = \iint_\Omega \left(a_{11} \frac{\partial w}{\partial x} \frac{\partial v}{\partial x} + a_{12} \left(\frac{\partial w}{\partial x} \frac{\partial v}{\partial y} + \frac{\partial w}{\partial y} \frac{\partial v}{\partial x} \right) + a_{22} \frac{\partial w}{\partial y} \frac{\partial v}{\partial y} \right) dx\, dy. \quad (4.2)$$

We will also use the inner product

$$(w, v)_\Omega = \iint_\Omega wv\, dx\, dy. \quad (4.3)$$

Let $H^1_E(\Omega) \subset H^1(\Omega)$ be the set whose members satisfy the condition $w = g$ on $\partial\Omega$; similarly, let $w = 0$ on $\partial\Omega$ for all $w \in H^1_0(\Omega)$. Then the weak solution of (4.1) is a function $u \in H^1_E(\Omega)$ satisfying

$$a_\Omega(u, v) = (f, v)_\Omega, \quad \forall\ v \in H^1_0(\Omega); \quad (4.4)$$

in fact we require the additional smoothness

$$\left.\begin{aligned} &u \in H^1_E(\Omega) \cap H^3(\Omega), \\ &a_{ij} \in W^2_{2+\varepsilon}(\Omega) \quad (i, j = 1, 2) \text{ for some } \varepsilon > 0 \\ \text{and} \quad &f \in H^2(\Omega). \end{aligned}\right\} \quad (4.5)$$

Now, the finite element approximation u_h which we introduce below is defined on Ω_h, which is not necessarily contained in Ω. Although u_h is computed using values of functions on Ω, it will simplify our analysis to extend these functions to Ω_h. Indeed, since Ω has the strong cone property we can use Calderon's theorem (Calderon, 1961, Theorem 12 or Babich, 1953) to give extensions of u and a_{ij} $(i, j = 1, 2)$ in the Sobolev spaces of (4.5) (to $\mathbf{R}^2$ as opposed to Ω). The restriction back to Ω of the extension operator yields the identity and so we can use a single symbol for a function and its extension. We have

$$\|u\|_{3,\Omega_h} \leqslant c\|u\|_{3,\Omega} \quad \text{and} \quad \|a_{ij}\|_{W^2_{2+\varepsilon}(\Omega_h)} \leqslant c, \quad (i, j = 1, 2). \quad (4.6)$$

We extend f as follows:

$$f = Lu \in H^1(\Omega_h) \cap H^2(\Omega), \quad (4.7)$$

where L is the operator of (4.1) and u and the coefficients of L are extended as above. Then by Green's theorem,

$$a_{\Omega_h}(u, v) = (f, v)_{\Omega_h}, \quad \forall\ v \in H^1_0(\Omega_h), \quad (4.8)$$

where $a_{\Omega_h}(\cdot, \cdot)$ and $(\cdot, \cdot)_{\Omega_h}$ correspond to the forms (4.2) and (4.3) with integration over Ω_h.

In all practical computations, $a_{\Omega_h}(\cdot, \cdot)$ and $(\cdot, \cdot)_{\Omega_h}$ will be evaluated by numerical quadrature. The centroid rule is sufficient for our purposes; its use is denoted thus: $a^*_{\Omega_h}(\cdot, \cdot)$, $(\cdot, \cdot)^*_{\Omega_h}$. (Our results can be modified to apply to any other rule, provided it integrates linear functions exactly in each element.)

If h is sufficiently small, the ellipticity of L is passed onto its extension so that (Ciarlet, 1978, Theorem 4.1.2)

$$|v|^2_{1,\Omega_h} \leqslant c a^*_{\Omega_h}(v, v), \quad \forall\ v \in S^h. \quad (4.9)$$

This discrete coercivity condition implies the existence and uniqueness of the finite element approximation which we now define.

Let $S_E^h(\Omega_h)$ be the subset of $S^h(\Omega_h)$ whose members interpolate nodal values of g on the boundary $\partial\Omega_h$; similarly, let $v=0$ on $\partial\Omega_h$ for all $v \in S_0^h(\Omega_h)$. [Note that $S_E^h(\Omega_h) \not\subset H_E^1(\Omega_h)$ but $S_0^h(\Omega_h) \subset H_0^1(\Omega_h)$.] We define $u_h \in S_E^h(\Omega_h)$ by

$$a^*_{\Omega_h}(u_h, v) = (f, v)^*_{\Omega_h}, \quad \forall\ v \in S_0^h(\Omega_h). \tag{4.10}$$

Also let $u_I \in S_E^h(\Omega_h)$ interpolate the values of u at all the nodes of Ω_h. (This corresponds to the u_I of Section 3, except that u, and therefore u_I, are no longer "known" functions.)

Contributions to $|u_I - u_h|_{1,\Omega_h}$ come from a variety of sources; we examine in this section the error due to numerical quadrature. Note that the mesh geometry, i.e. Lemma 2.3, plays no further part until Lemma 5.3.

We expand the numerical approximation to $\iint_{\Omega_h} w\,dx\,dy$ in the form $\sum_{k=1}^{K} [w]_{G_k} \cdot m(T_k)$, where G_k and $m(T_k)$ are the centroid and measure of element T_k. We define the local error functional over T_k:

$$E_k(w) = \iint_{T_k} w\,dx\,dy - [w]_{G_k} \cdot m(T_k).$$

LEMMA 4.1 *Let k be fixed (again we drop the subscript). Then*

(i) $$|E(w)| \leqslant ch^3 \|w\|_{2,T}, \quad \forall\ w \in H^2(\Omega_h);$$

(ii) $$|E(w)| \leqslant ch^2(\|w\|_{1,T} + \|w\|_{2,T\cap\Omega}), \quad \forall\ w \in H^1(\Omega_h) \cap H^2(\Omega).$$

Proof. By (2.2) and Lemma 2.2,

$$|E(w)| \leqslant ch^2 |\hat{E}(\hat{w})|. \tag{4.11}$$

Since the quadrature scheme is exact for linears, (i) follows immediately from (4.11) and Lemmas 2.1 and 2.2. For (ii) we proceed with caution, for $|\hat{E}(\hat{w})|$ is not bounded by $\|\hat{w}\|_{1,\tau}$. We recall that $T \cap \Omega$ contains an open disc which itself contains the centroid G and whose image, τ_* say, in the (ξ, η)-plane has measure $\geqslant c$ (by (2.2)).

As in Lemma 3.2, we introduce a projection operator:

$$\hat{R}\hat{w} = \begin{cases} \text{meas}\,(\tau_*)^{-1} \displaystyle\iint_{\tau_*} \hat{w}\,d\xi\,d\eta & \text{in } \tau_* \\ \hat{w} & \text{in } \tau \backslash \tau_* \end{cases}$$

Then

$$\hat{E}(\hat{R}\hat{w}) = \iint_{\tau} \hat{w}\,d\xi\,d\eta - \frac{\text{meas}\,(\tau)}{\text{meas}\,(\tau_*)} \iint_{\tau_*} \hat{w}\,d\xi\,d\eta;$$

this is bounded in $H^1(\tau)$ and vanishes for $\hat{w}$ constant on τ. Therefore by the Sobolev lemma and Lemma 2.1,

$$|\hat{E}(\hat{R}\hat{w})| \leqslant c|\hat{w}|_{1,\tau}.$$

The remainder, $\hat{E}(\hat{w} - \hat{R}\hat{w})$ is bounded in $W^1_{2+\varepsilon}(\tau_*)$ for fixed $\varepsilon > 0$; it vanishes for $\hat{w}$ constant on τ_* and is thus bounded by $c(|\hat{w}|_{1,\tau_*} + |\hat{w}|_{2,\tau_*})$. (This is similar to the proof of Lemma 3.1.) So

$$|\hat{E}(\hat{w})| \leqslant |\hat{E}(\hat{R}\hat{w})| + |\hat{E}(\hat{w}-\hat{R}\hat{w})| \leqslant c(|\hat{w}|_{1,\tau}+|\hat{w}|_{2,\tau})$$

and (ii) follows from (4.11) and Lemma 2.2.

We apply Lemma 4.1 to give global estimates of the quadrature error:

LEMMA 4.2 *Let* $v \in S_0^*(\Omega_h)$. Then

(i) $$|(f, v)_{\Omega_h} - (f, v)^*_{\Omega_h}| \leqslant ch^2(\|f\|_{2,\Omega} + \|f\|_{1,\Omega_h})\|v\|_{1,\Omega_h};$$

(ii) $$|a_{\Omega_h}(u, v) - a^*_{\Omega_h}(u, v)| \leqslant ch^2\|u\|_{3,\Omega_h}\|v\|_{1,\Omega_h}.$$

Proof. We again employ a projection method. For (i) we write

$$|(f, v)_{\Omega_h} - (f, v)^*_{\Omega_h}| \leqslant \sum_{k=1}^{K} \left(|E_k(f[v]_{G_k})| + |E_k(f(v-[v]_{G_k}))|\right). \tag{4.12}$$

To bound the first term when T_k is wholly contained in Ω (so that $\hat{f} \in H^2(\tau_k)$) we note that

$$\sup_{T_k}(|v| + |\nabla v|) \leqslant ch^{-1}\|v\|_{1,T_k}, \quad \forall\ v \in S^h \tag{4.13}$$

and use Lemma 4.1(i) to obtain

$$|E_k(f[v]_{G_k})| \leqslant |E_k(f)| \cdot |[v]_{G_k}| \leqslant ch^2\|f\|_{2,T_k}\|v\|_{1,T_k}.$$

Alternatively, T_k is a boundary element and by Lemma 4.1(ii) (recall (4.7))

$$|E_k(f[v]_{G_k})| \leqslant ch^2(\|f\|_{1,T_k} + \|f\|_{2,T_k\cap\Omega}) \cdot |[v]_{G_k}|.$$

We cannot use (4.13) here without losing an order of h. But since $v = 0$ on $\partial\Omega$,

$$|[v]_{G_k}| \leqslant ch^{-1}\|v\|_{0,T_k} \leqslant c|v|_{1,T_k}.$$

Therefore, for all T_k,

$$|E_k(f[v]_{G_k})| \leqslant ch^2(\|f\|_{1,T_k} + \|f\|_{2,T_k\cap\Omega})\|_{1,T_k}.$$

To bound the other term in (4.12) we write γ_k for the centroid of τ_k and note that since v is linear, $\hat{E}_k = 0$ when $\hat{f}$ is a constant on τ_k. So by (2.2) and Lemmas 2.1 and 2.2,

$$\begin{aligned}|E_k(f(v-[v]_{G_k})) &\leqslant ch^2|\hat{E}_k(\hat{f}(\hat{v}-[\hat{v}]_{\gamma_k}))| \\ &\leqslant ch^2\left|\iint_{\tau_k} \hat{f}(\hat{v}-[\hat{v}]_{\gamma_k})\right| \\ &\leqslant ch^2|\hat{f}|_{1,\tau_k}|\hat{v}|_{1,\tau_k} \\ &\leqslant ch^2\|f\|_{1,T_k}\|v\|_{1,T_k}.\end{aligned}$$

We now obtain (i) from the above estimates and the Cauchy–Schwarz inequality.

For (ii) we recall that ∇v is constant over each element and write

$$\begin{aligned}&\left|\iint_{T_k} a_{11}\frac{\partial u}{\partial x}\frac{\partial v}{\partial x} - \left[a_{11}\frac{\partial u}{\partial x}\frac{\partial v}{\partial x}\right]_{G_k}\cdot m(T_k)\right| \\ &\qquad \leqslant \left(\left|E_k\left(\frac{\partial u}{\partial x}[a_{11}]_{G_k}\right)\right| + \left|E_k\left(\frac{\partial u}{\partial x}(a_{11}-[a_{11}]_{G_k})\right)\right|\right)\cdot\left|\left[\frac{\partial v}{\partial x}\right]_{T_k}\right|.\end{aligned}$$

We bound the first term directly from Lemma 4.1(i); the bound on the second is similar to the corresponding term in (i) above. We use an identical method for the a_{12} and a_{22} terms of $a_\Omega(\cdot, \cdot)$ and Cauchy–Schwarz to sum over the elements T_k.

We collect the above results:

THEOREM 4.1 *Let u, u_h, u_I, f and $\mathbf{D}_k$ ($k = 1, \ldots, K_A$) be as defined above. Then*

$$h\left(\sum_{k=1}^{K_A} |\mathbf{D}_k u_h - [\nabla u]_{M_k}|^2\right)^{\frac{1}{2}}$$
$$\leqslant c \sup_{v \in S_0^h} |a_{\Omega_h}^*(u_I - u, v)| / \|v\|_{1,\Omega_h} + ch^2(\|u\|_{3,\Omega} + \|f\|_{2,\Omega}).$$

Proof. If $|u_I - u_h|_{1,\Omega_h} = 0$, the result follows directly from Theorem 3.1. Otherwise we substitute $v = u_I - u_h \in S_0^h$ into (4.8)–(4.10). Then, by Lemma 4.2,

$$\begin{aligned}|u_I - u_h|_{1,\Omega_h}^2 &\leqslant c a_{\Omega_h}^*(u_I - u_h, v)\\
&\leqslant c(|a_{\Omega_h}^*(u_I - u, v)| + |a_{\Omega_h}^*(u, v) - a_{\Omega_h}(u, v)| + |(f, v)_{\Omega_h} - (f, v)_{\Omega_h}^*|)\\
&\leqslant c|a_{\Omega_h}^*(u_I - u, v)| + ch^2(\|u\|_{3,\Omega_h} + \|f\|_{2,\Omega} + \|f\|_{1,\Omega_h})\|v\|_{1,\Omega_h}.\end{aligned}$$

By (4.7), $\|f\|_{1,\Omega_h} \leqslant c\|u\|_{3,\Omega_h}$; by Friedrich's inequality $\|v\|_{1,\Omega_h} \leqslant c|v|_{1,\Omega_h}$. So

$$|u_I - u_h|_{1,\Omega_h} \leqslant \frac{c|a_{\Omega_h}^*(u_I - u, v)|}{\|v\|_{1,\Omega_h}} + ch^2(\|u\|_{3,\Omega_h} + \|f\|_{2,\Omega}).$$

The result now follows from Theorem 3.1 and (4.6).

It remains to show that the term $a_{\Omega_h}^*(u_I - u, v)$ is small enough to justify use of the recovery scheme proposed earlier. We devote the next section to this.

5. $a_{\Omega_h}^*(u_I - u, v) = O(h^2)$, $\forall\ v \in S_0^h$

This bound was derived independently by Oganesjan & Ruchovec (1969), though only for the case of a fully "uniform" grid (i.e. $X \equiv x$, $Y \equiv y$) and without the application to superconvergence. The result here is similar to Lemma 3.2; the principal difference is that the interpolation error

$$e = u_I - u$$

is now averaged over each A_k instead of in the neighbourhood of each sampling point M_k. Further complications arise from non-uniformity of the mesh and variability of the coefficients a_{ij}. These are essentially perturbations to the superconvergence effect and we deal with them first.

We use the notation:

$$b_{11} = a_{11}X_x + a_{12}X_y, \qquad b_{12} = a_{11}Y_x + a_{12}Y_y,$$
$$b_{21} = a_{12}X_x + a_{22}X_y, \qquad b_{22} = a_{12}Y_x + a_{22}Y_y$$

and

$$\mathbf{b} = \begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix}.$$

By the chain rule,

$$\left.\begin{aligned} a^*_{\Omega_k}(e, v) &= \sum_{k=1}^{K} [w]_{G_k} \cdot m(T_k) \\ \text{where} \quad w &= \left(\frac{\partial e}{\partial x} b_{11} + \frac{\partial e}{\partial y} b_{21}\right) \frac{\partial v}{\partial X} + \left(\frac{\partial e}{\partial x} b_{12} + \frac{\partial e}{\partial y} b_{22}\right) \frac{\partial v}{\partial Y}. \end{aligned}\right\} \tag{5.1}$$

We will examine the first term in (5.1) in detail:

$$\sum_{k=1}^{K} \left[\frac{\partial v}{\partial X} (\mathbf{b} \cdot \nabla e)\right]_{G_k} \cdot m(T_k), \tag{5.2}$$

considering first the extent to which $\partial v/\partial X$ differs from the constant $h^{-1}(\partial \hat{v}/\partial \xi_k)$ in element T_k. (Recall $v \in S_0^*(\Omega_k)$; see Fig. 3.)

LEMMA 5.1 *Let*

$$S_1 = \sum_{k=1}^{K} \left(\left[\frac{\partial v}{\partial X}\right]_{G_k} - h^{-1}\left[\frac{\partial \hat{v}}{\partial \xi_k}\right]_{\tau_k}\right) [(\mathbf{b} \cdot \nabla e)]_{G_k} \cdot m(T_k).$$

Then

$$|S_1| \leqslant ch^2 \|u\|_{3, \Omega_k} \|v\|_{1, \Omega_k}.$$

Proof. Consider a single element T_k. (As before we drop the subscript k.) In this element,

$$\frac{\partial \hat{v}}{\partial \xi} = h \frac{\partial v}{\partial X} + \hat{F}(\hat{X}) \frac{\partial v}{\partial X} + \hat{F}(\hat{Y}) \frac{\partial v}{\partial Y}, \tag{5.3}$$

where

$$\hat{F}(\cdot) = \frac{\partial}{\partial \xi} - [\cdot]_{(1,0)} + [\cdot]_{(0,0)},$$

$(\hat{X}, \hat{Y})$ are viewed as functions of (ξ, η) (hence the $\hat{}$) and everything is evaluated at some fixed point $(\bar{\xi}, \bar{\eta}) = t(\bar{x}, \bar{y})$ unless otherwise stated. (Recall that if $(\hat{X}, \hat{Y})(0, 0) = (X_0, Y_0)$, then $(\hat{X}, \hat{Y})(1, 0) = (X_0 + h, Y_0)$.

Now if $\hat{X}$ is linear in (ξ, η), then $\hat{F}(\hat{X}) = 0$. So by the Sobolev lemma and Lemmas 2.1 and 2.2,

$$|\hat{F}(\hat{X})| \leqslant ch^2 |X|_{W^2_\infty(T)} \leqslant ch^2.$$

Similarly, $|\hat{F}(\hat{Y})| \leqslant ch^2$ and (5.3) becomes

$$\begin{aligned} \left|\left[\frac{\partial v}{\partial X}\right]_{(\bar{x}, \bar{y})} - h^{-1} \frac{\partial \bar{v}}{\partial \xi}\right| &\leqslant h^{-1} \sup_{\substack{(\xi, \eta) \in \tau \\ (\bar{x}, \bar{y}) \in T}} \left|\hat{F}(\hat{X}) \frac{\partial v}{\partial X} + \hat{F}(\hat{Y}) \frac{\partial v}{\partial Y}\right| \\ &\leqslant ch \sup_T |\nabla v| \\ &\leqslant c\|v\|_{1, T}. \end{aligned} \tag{5.4}$$

We sum over k. By Lemmas 2.2 and 3.1 and Cauchy–Schwarz,

$$\begin{aligned} |S_1| &\leqslant \sum_{k=1}^{K} \left|\left[\frac{\partial v}{\partial X}\right]_{G_k} - h^{-1} \left[\frac{\partial \hat{v}}{\partial \xi_k}\right]_{\tau_k}\right| \cdot |[\mathbf{b}]_{G_k}| \cdot |[\nabla e]_{G_k}| \cdot m(T_k) \\ &\leqslant \sum_{k=1}^{K} c\|v\|_{1, T_k} \cdot c \cdot c\|u\|_{3, T_k} \cdot ch^2 \\ &\leqslant ch^2 \|u\|_{3, \Omega_k} \|v\|_{1, \Omega_k}. \end{aligned}$$

The next step is to bound the variation of $\mathbf{b}$ over each element. Recall M_k $(k = 1, \ldots, K)$ are the stress points (as in Fig. 2).

LEMMA 5.2 *Let*

$$S_2 = \sum_{k=1}^{K} h^{-1} \left[\frac{\partial V}{\partial \xi_k}\right]_{\tau_k} ([\mathbf{b} \cdot \nabla e]_{G_k} - [\mathbf{b}]_{M_k} \cdot [\nabla e]_{G_k}) \cdot m(T_k).$$

Then

$$|S_2| \leqslant ch^2 \|u\|_{3,\Omega_h} \|v\|_{1,\Omega_h}.$$

Proof. Again, we consider a single element, T_k, and drop the subscript k. We denote by Z the matrix whose (i, j)th element is a_{ij} and write

$$[\mathbf{b}]_G - [\mathbf{b}]_M = \tfrac{1}{2}(([Z]_G + [Z]_M)([\nabla X]_G - [\nabla X]_M) + ([Z]_G - [Z]_M)([\nabla X]_G + [\nabla X]_M)).$$

Now $X \in W_\infty^2$ and each element of Z is in $W_{2+\varepsilon}^2$. Therefore both $|Z|_{W_\infty^1(T)}$ and $|\nabla X|_{W_\infty^1(T)}$ are bounded and we obtain

$$|[\mathbf{b}]_G - [\mathbf{b}]_M| \leqslant ch$$

from Lemmas 2.1 and 2.2. Furthermore, from (5.4) we have

$$|[\partial \hat{v}/\partial \xi]_\tau| \leqslant c\|v\|_{1,T}.$$

The result now follows directly from Lemmas 2.2 and 3.1 and Cauchy–Schwarz.

For the next result we recall (3.1) and group the elements into triangle pairs A_k, rewriting numerical quadrature of ∇e as a functional pair:

$$\mathbf{F}_k(u) = [\nabla e]_{G_k} m(T_k) + [\nabla e]_{G_{k'}} m(T_{k'}), \quad (k = 1, \ldots, K_A).$$

LEMMA 5.3 *Let*

$$S_3 = \sum_{k=1}^{K_A} h^{-1} \left[\frac{\partial \hat{v}}{\partial \xi_k}\right]_{\alpha_k} [\mathbf{b}]_{M_k} \cdot \mathbf{F}_k(u).$$

Then

$$|S_3| \leqslant ch^2 \|u\|_{3,\Omega_h} \|v\|_{1,\Omega_h}.$$

Proof. This is essentially the same as that of Lemma 3.2. We fix k initially, dropping the subscript. As before we write τ_+, τ_- for the triangles with vertices $\{(0, 0), (1, 0), (\xi_+, \eta_+)\}$, $\{(0, 0), (\xi_-, \eta_-), (1, 0)\}$, with $\tau_+ \cup \tau_- = \alpha$. (See Fig. 3.) Then the relations (3.3) hold and we have

$$\left.\begin{aligned} \hat{\mathbf{F}}(\xi^2) &= \begin{pmatrix} \tfrac{1}{6}((2\xi_- - 1)\eta_- - (2\xi_+ - 1)\eta_+) \\ \tfrac{1}{2}(\xi_-(1-\xi_-) - \xi_+(1-\xi_+)) \end{pmatrix}, \\ \hat{\mathbf{F}}(\xi_\eta) &= \begin{pmatrix} \tfrac{1}{6}(\eta_-^2 - \eta_+^2) \\ \tfrac{1}{6}((1-2\xi_-)\eta_- - (1-2\xi_+)\eta_+) \end{pmatrix} \\ \text{and} \quad \hat{\mathbf{F}}(\eta^2) &= \begin{pmatrix} 0 \\ \tfrac{1}{6}(\eta_+^2 - \eta_-^2) \end{pmatrix}. \end{aligned}\right\} \tag{5.5}$$

If α were a parallelogram, $\hat{\mathbf{F}}$ would vanish for quadratic $\hat{u}$. But α is only close to a parallelogram, so we recall the projection $\hat{R}$ of (3.6) and, exactly as in Lemma 3.2, use (5.5) to obtain

$$|\hat{\mathbf{F}}(\hat{R}\hat{u})| \leqslant ch|\hat{u}|_{2,\alpha}$$

and

$$|\hat{\mathbf{F}}(\hat{u}-\hat{R}\hat{u})| \leqslant c|\hat{u}|_{3,\alpha}.$$

So by Lemma 2.2

$$|F(u)| \leqslant ch|\hat{\mathbf{F}}(\hat{u})| \leqslant ch^3\|u\|_{3,A}$$

and the result is obtained, as in the last two lemmas, by summing over k.

We now complete the superconvergence proof with:

THEOREM 5.1 *Let u and u_I be as defined and let $v \in S_0^h(\Omega_h)$. Then*

$$|a^*_{\Omega_h}(u_I-u, v)| \leqslant ch^2\|u\|_{3,\Omega_h}\|v\|_{1,\Omega_h}.$$

THEOREM 5.2 *As an immediate consequence of* (4.6) *and Theorems* 4.1 *and* 5.1,

$$h\left(\sum_{k=1}^{K_A} |\mathbf{D}_k u_h - [\nabla u]_{M_k}|^2\right)^{\frac{1}{2}} \leqslant ch^2(\|u\|_{3,\Omega}+\|f\|_{2,\Omega}).$$

Proof of Theorem 5.1 We note that $\partial\hat{v}/\partial\xi_k = 0$ in each B_k. For (see Fig. 4) in every B_k the nodes N_{k0} and N_{k1} lie on $\partial\Omega_h$. But $v \in S_0^h(\Omega_h)$ and so $v = 0$ at N_{k0} and N_{k1} and varies linearly between them. Hence the component of ∇v parallel to $N_{k0}N_{k1}$ is zero.

Therefore by Lemmas 5.1–5.3,

$$\left|\sum_{k=1}^{k}\left[\frac{\partial v}{\partial X}(\mathbf{b}\,.\,\nabla e)\right]_{G_k} m(T_k)\right| = |S_1+S_2+S_3| \leqslant ch^2\|u\|_{3,\Omega_h}\|v\|_{1,\Omega_h}.$$

Returning to (5.1), we have bounded the first term; the second is bounded similarly. We have now derived l_2 superconvergence of the recovered gradient at element edge midpoints.

6. The Mesh Geometry

In this section we relax the triangulation conditions of Section 2, but on general curved regions Ω we do not then expect superconvergence in the global sense of Sections 3–5. For these cases we propose a local form of the superconvergence property.

We refer to any region for which Theorem 5.2 holds ($\forall\, u \in H^3(\Omega)$, $f \in H^2(\Omega)$) as "superconvergent". For example, it is clear that this includes any region whose topologically equivalent triangulation in the (X, Y) plane is the mesh shown in Fig. 5(a).

The region shown in Fig. 5(b) does not satisfy the triangulation specifications: however, with a modified recovery scheme it is superconvergent. This property is true for all (sufficiently smooth) meshes with exactly six elements meeting at each internal node. We call them "chevron meshes" (Fig. 6); their definition is sufficient to ensure that Ω_h can be exactly partitioned into "bands" of triangles. A "band" consists of one or more adjacent, entire columns (or rows) of the squares which make up the triangulation Ω_h, plus any left-over triangles (B_k) at the two ends (i.e. on $\partial\Omega_h$). Ω is triangulated as before, except that all the diagonals in a band may have (X, Y)-slope -1 instead of $+1$.

This generalization affects only two stages of the superconvergence proofs in the previous sections. We recall that Lemma 5.3 requires the sum (5.1) to be partitioned

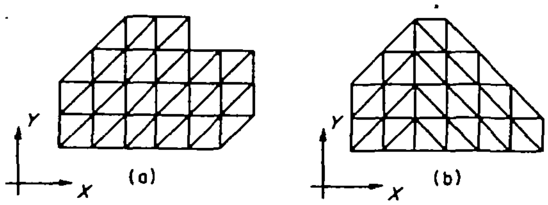

FIG. 5. These two meshes are superconvergent.

into two terms. The first term (5.2) has as a factor the component of ∇v which is (almost) constant over triangle pairs with common edge (almost) parallel to the X-axis—see Lemma 5.1—similarly for the second term and the Y-axis. It is this partition which we modify here, dealing with each band separately. We consider without loss of generality a region Ω with band-boundaries (almost) parallel to the X-axis (as in Fig. 6) and a band for which each hypotenuse has (X, Y)-slope ± 1; instead of (5.1) we write

$$w=\left(\frac{\partial e}{\partial x}b_{11}+\frac{\partial e}{\partial y}b_{21}\right)\left(\frac{\partial v}{\partial X}\pm\frac{\partial v}{\partial Y}\right)+\left(\frac{\partial e}{\partial x}(b_{12}\pm b_{11})+\frac{\partial e}{\partial y}(b_{22}\mp b_{21})\right)\frac{\partial v}{\partial Y}. \tag{6.1}$$

Now $(\partial v/\partial X \pm \partial v/\partial Y)$ is (almost) constant over triangle pairs with common edge of slope ± 1, i.e. over the squares which comprise the band. Therefore there are no unpaired triangles B_k on the "long" edges of the band (internal to Ω_h) on which $v=0$ is not guaranteed. So with the decompositions (6.1) we can write $a^*_{\Omega_h}(e, v)$ as a sum of contributions from each band and Theorem 5.1 proceeds as before.

The other aspect of superconvergence which is sensitive to mesh geometry is the recovery of the stress component normal to element edges. The tangential component (recall the bounding of $\hat{F}_1$ in Lemma 3.2) is not affected. So Theorem 5.2 holds for all chevron meshes, if we remove from the average those normal stress components which are directed through band-boundaries. We can, however, recover the full stress at these points by means of a modified scheme: we average the gradient over four elements, as shown in Fig. 7. The resulting error of this well-centred difference scheme is bounded analogously to Lemma 3.2; we conclude that with this new recovery scheme all chevron meshes are superconvergent.

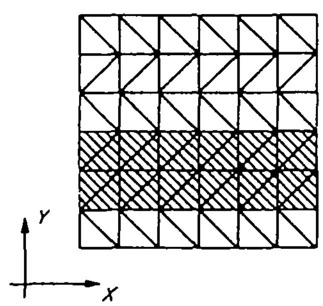

FIG. 6. A chevron mesh. The shaded region is a band with diagonals of (X, Y)-slope $+1$. $(\partial v/\partial X + \partial v/\partial Y$ is (almost) constant in each "square".)

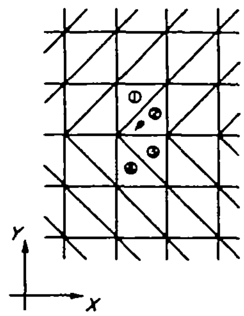

FIG. 7. To recover the full gradient at the point P—note that the union of elements is not (even close to) a parallelogram—average the approximate gradient over elements 1, 2, 3, 4.

We note here that the criss-cross mesh (see Fig. 8) required for derivative superconvergence in the mixed method of Fix, Gunzburger & Nicolaides (1981) does not have six elements surrounding each node and cannot be arranged into bands. So (this is independent of the choice of recovery scheme) Lemma 5.3 cannot be applied; the mesh is not superconvergent. It is because of this necessary restriction that no simple rectangular mesh on an octagonal region is superconvergent [Fig. 9(a)]. This case is qualitatively equivalent to the mesh shown in Fig. 9(b) where we obtain at best $|a^*_{\Omega_h}(e, v)| = O(h^{\frac{3}{2}})$. This $O(h^{\frac{1}{2}})$ drop in accuracy is confirmed numerically in Section 7.

We conclude that, with the introduction of chevron meshes, the conditions for superconvergence can be satisfied on a wide variety of practical problem domains. We are, however, still a long way from superconvergence on general regions; there is a theoretical barrier (the "six element" condition) to further progress. In the next section we present evidence supporting an alternative result: that superconvergence holds in those subregions of Ω_h which are bounded away from areas where the mesh conditions (or, for example, the smoothness of u) break down. A proof of this, combining Sections 3–5 above with approximation properties of Green's functions (Rannacher & Scott, 1982) is discussed elsewhere (Levine, 1985). (Incidentally, this new result gives pointwise superconvergence, i.e. without the need (as in e.g. Theorem 5.2) to take a global average over the stress points.) Also, see Levine (1985) for further ideas on triangulation methods.

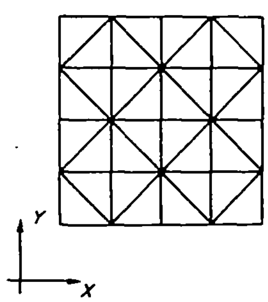

FIG. 8. A criss-cross mesh similar to that used by Fix *et al.* (1981).

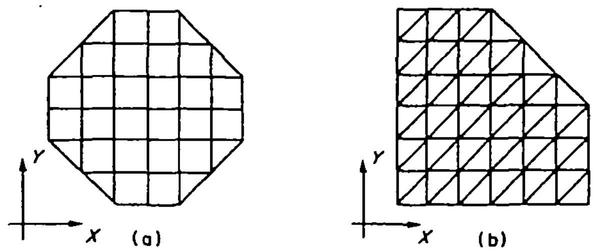

FIG. 9. (a) The triangulation cannot be completed so as to give a superconvergent mesh. (b) This mesh too is not superconvergent. Note that there are internal nodes which are not surrounded by six elements.

7. Numerical Results

7.1 *Centroid Recovery*

The recovery scheme considered above is that of averaging the approximate gradient between neighbouring elements; this yields an $O(h^2)$ estimate of the true gradient at the midpoint of the shared edge. We denote the root-mean-square error of this recovery scheme (averaging over all possible edges) by E_{mid}. A simpler though generally less useful procedure is to sample just the tangential component of the gradient at the midpoint of each element edge; this is an $O(h^2)$ estimate of that component of the true gradient. We denote this averaged error by E_{tgt}.

We can also recover the gradient at the centroids, simply and to $O(h^2)$: we first recover the gradient at the midpoint of each of the edges of a triangle and then average these three gradients to obtain an approximation to the gradient at the centroid. (To prove that this scheme leads to superconvergence we either regard it as the result of a linear fit to the recovered gradient at these three stress points or make a straightforward change to Lemma 3.2.) We then have a weighted averaging scheme between four elements (see Fig. 10); we denote the average error by E_{rec}. (See Levine, 1985 for recovery schemes for general points in Ω.)

We now recall the claim that the gradient can be sampled to high accuracy at the centroid of each element. That this cannot be to $O(h^2)$ follows simply from Theorem 5.2 and Taylor's theorem (for details see Levine, 1982, 1985). We denote the average error for this sampling procedure by E_{cent}.

To compare these four measures of error, we considered Poisson's equation on the unit square, $\Omega = (0, 1) \times (0, 1)$, with exact solution

$$u = x(1-x)y(1-y)(1+2x+7y).$$

We triangulated Ω with a uniform mesh separation $(x \equiv X, y \equiv Y)$ taking successively $h = \frac{1}{4}, \frac{1}{6}, \frac{1}{8}, \frac{1}{10}, \frac{1}{12}$. We set $f = -\Delta u$ on Ω and $g = u$ on $\partial\Omega$ and solved (4.10) to obtain u_h for each h. We found that each error was within 10% (and usually 5%) of its asymptotic rate for $h \leq \frac{1}{8}$; these rates were

$$E_{\text{tgt}} \cong 1{\cdot}4h^2,$$
$$E_{\text{mid}} \cong 3{\cdot}0h^2,$$
$$E_{\text{rec}} \cong 3{\cdot}0h^2$$

and

$$E_{\text{cent}} \cong 1{\cdot}2h.$$

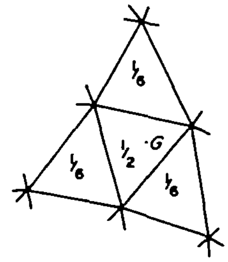

FIG. 10. Weights of ∇u_h in four elements, yielding a superconvergent approximation to ∇u at the centroid G.

Remarks

(a) The same problem has been solved with bilinear elements, where sampling at centroids leads to superconvergence. Lesaint & Zlámal (1979) gave the result $E_{\text{cent}} \cong 0{\cdot}91h^2$.

(b) To investigate the error introduced by numerical quadrature we solved the above problem using exact integration instead of the centroid rule. Only a slight improvement occurred ($E_{\text{rec}} \cong 2{\cdot}8h^2$).

(c) We solved this problem using the criss-cross mesh (Fig. 8) for which superconvergence is not expected (under any recovery scheme). We obtained

$$E_{\text{rec}} \cong 0{\cdot}45h$$

and

$$E_{\text{cent}} \cong 1{\cdot}0h,$$

indicating that there may be some value in using the recovery algorithm even when superconvergence is absent.

(d) We considered a curved mesh, distorting Ω into the sector shown in Fig. 11 by the transformation

$$\left.\begin{aligned} x &= (X+2)/(1+Y^2/4)^{\frac{1}{2}}-2, \\ y &= Y(1+x/2). \end{aligned}\right\} \qquad (7.1)$$

Superconvergence was again observed with

$$E_{\text{tgt}} \cong 1{\cdot}2h^2,$$

$$E_{\text{mid}} \cong 3{\cdot}2h^2$$

and

$$E_{\text{rec}} \cong 3{\cdot}0h^2.$$

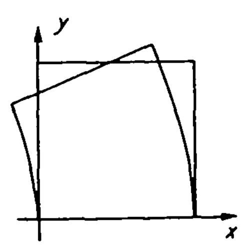

FIG. 11. Distortion of Ω by (7.1).

7.2 *Local Superconvergence*

We took Ω to be the truncated unit square triangulated as in Fig. 9(b), $h = \frac{1}{4}, \ldots, \frac{1}{12}$ and solved (4.10) for Poisson's equation as above, with exact solution

$$u = (x-1)^2 + y^2.$$

(This is a function for which there is zero error on uniform superconvergent triangulations: we chose it to highlight asymptotic behaviour for computationally reasonable values of h. It has been our experience that when breakdown of superconvergence is due to effects from a subdomain of Ω, such as the neighbourhood of a line, the error is somewhat smaller than expected and the asymptotic rate is not attained for practical values of h.)

As expected, we obtained

$$E_{\text{rec}} \cong 1{\cdot}9h^{\frac{3}{2}} \quad \text{and} \quad E_{\text{cent}} \cong 0{\cdot}47h.$$

However, when we restricted the averages to elements in the subdomain $(0, \frac{1}{2}) \times (0, \frac{1}{2})$ (this is bounded away from the region where the mesh conditions break down) we obtained

$$E_{\text{rec}} \cong 1{\cdot}9h^2 \quad (\text{and } E_{\text{cent}} \cong 0{\cdot}47h).$$

This is the local superconvergence effect predicted in the last section. Its implication, which is of practical significance, is that for any Ω there exists a series of triangulations such that our superconvergence results hold in all elements bounded away from $\partial\Omega$.

I would like to thank Professor K. W. Morton for invaluable discussions, Paul Samuels for the use of his finite element code and the Science & Engineering Research Council and the University of Reading for their financial support.

References

BABICH, V. M. 1953 The problem of the extension of functions. *Usp. mat. Nauk.* **7**, 111–113.

BARLOW, J. 1976 Optimal stress locations in finite-element models. *Int. J. num. Meth. Engng* **10**, 243–251.

BRAMBLE, J. H. & HILBERT, S. R. 1970 Estimation of linear functions on Sobolev spaces with applications to Fourier transforms and spline interpolation. *SIAM J. num. Analysis* **7**, 113–124.

CALDERON, A. P. 1961 Lebesgue spaces of differentiable functions and distributions. *Proc. Symp. pure Math.* **IV**, 33–49.

CIARLET, P. G. 1978 *The Finite Element Method for Elliptic Problems*. Amsterdam: North-Holland.

FIX, G. J., GUNZBURGER, M. D. & NICOLAIDES, R. A. 1981 On mixed finite element methods for first order elliptic systems. *Num. Math.* **37**, 29–48.

KŘÍŽEK, M. & NEITTAANMÄKI, P. 1983 Superconvergence phenomena arising in averaged gradients. Preprint 23, University of Jyväskylä.

LESAINT, P. & ZLÁMAL, M. 1979 Superconvergence of the gradient of finite element solutions. *Rev. fr. Autom. Inf. Rech. opérat. Sér. Rouge Analyse Num.* **13**, 139–166.

LEVINE, N. D. 1982 Stress sampling points for linear triangles in the finite element method. Numerical Analysis Report 10/82, University of Reading.

LEVINE, N. D. 1985 Superconvergent estimation of the gradient from linear finite element approximations on triangular elements. Numerical Analysis Report 3/85. Ph.D. Thesis, University of Reading.

LIN QUN, LU TAO & SHEN SHUMIN 1983 Asymptotic expansion for finite element approximations. Research Report 11, Institute of Mathematical Sciences, Chengdu branch of Academia Sinica.

MOAN, T. 1974 Experiences with orthogonal polynomials and 'best' numerical integration formulas on a triangle. *Z. angew. Math. Mech.* **54**, 501–508.

ORGANESJAN, L. A. & RUCHOVEC, L. A. 1969 Investigation of the convergence rate of variational-difference schemes for elliptic second order equations in a two-dimensional domain with a smooth boundary. *Z. Vychisl. Mat. i Mat. Fiz.* **9**, 1102–1120. (Translation in *USSR Comput. Math. math. Phy.* **9** (1969), 153–188).

RANNACHER, R. & SCOTT, R. 1982 Some optimal error estimates for piecewise linear finite element approximations. *Maths Comput.* **38**, 437–445.

STRANG, G. & FIX, G. 1973 *An Analysis of the Finite Element Method.* Englewood Cliffs, N.J.: Prentice Hall.

VERYARD, D. A. 1971 Problems associated with the convergence of isoparametric and mixoparametric finite elements. M.Sc. Thesis, University of Wales.

ZLÁMAL, M. 1977 Some superconvergence results in the finite element method. In *Mathematical Aspects of Finite Element Methods* (Galligani, Magenes, Eds), pp. 351–362. New York: Springer-Verlag.

ZLÁMAL, M. 1978 Superconvergence and reduced integration in the finite element method. *Maths Comput.* **32**, 663–685.