

## NUMERICAL QUADRATURES AND THEIR EFFECT ON THE SOLUTION OF BOUNDARY VALUE PROBLEMS

By Hoon Liong Ong

### 1. Introduction

Consider the second order elliptic boundary value problem (Birkhoff [1], Strang and Fix [9]), defined in an open domain  $\Omega$  with polygonal boundary  $\partial\Omega$  by

$$\begin{cases} Lu = -\nabla \cdot (p\nabla u) + qu = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \quad (1)$$

Under the assumptions  $p, q$  are smooth functions and  $p \geq p_{\min} > 0, q \geq 0$  in  $\Omega$ , the solution of problem (1) is equivalent to the solution of

$$(Lu, v) = (f, v) \quad \text{for all } v \in H_0^1; \quad (2)$$

where  $H_0^1$  is the space of all  $u$  in the Sobolev space  $H^1$  with  $u=0$  on  $\partial\Omega$  and  $(\cdot, \cdot)$  is the inner product defined by  $(u, v) = \int_{\Omega} uv \, du$ .

The Ritz-Galerkin solution to (2) is to replace  $H^1$  by a finite dimensional subspace  $S^h$  contained in  $H^1$ . The elements of  $S^h$  are of the form  $v^h = \sum_{i=1}^n \lambda_i \phi_i$ , where  $\phi_i$  are basis functions of  $S^h$ . In the finite element method, the region  $\Omega$  is subdivided into a set  $\tau$  of finite number of disjoint regions, say triangle elements, and the space  $S^h$  consists of piecewise polynomials defined on the elements in  $\tau$ . An approximate solution to (1) can be obtained by solving the following discrete finite element system (Bramble and Zlamal [2], Courant [4])

$$(Lu^h, \phi_i) = (f, \phi_i) \quad \text{for all } i=1, 2, \dots, n. \quad (3)$$

In general, the integrals  $F_i = (f, \phi_i)$  may not be calculated exactly and hence some numerical quadratures will be necessary to approximate these integrals. Herbold and Varga [5] have proposed some quadrature schemes for the numerical solution of a class of boundary value problems by variational technique and investigated the errors introduced in the approximate solutions by such quadrature formulas. In this paper we derive two numerical quadratures for the load vector  $F_i$  and analyze the errors introduced by these quadratures.

Error bounds for the quadratures are expressed in term of the Sobolev seminorm in the form  $ch^s|f|_{s,\Omega}$ . The order  $s$  and the constant factor  $c$  of the quadrature error are estimated. Experimental results for the two numerical quadratures are discussed in [7].

The barycentric coordinate system, which some engineers call the areal coordinate system, is one of the best local coordinate systems for a study of the finite element method (Ciarlet [3], Strang and Fix [9]). In Section 2 we state the underlying notations and develop some fundamental results of vector calculus in the barycentric coordinate system. In Section 3 two quadrature schemes for calculating the load vector  $F_i$  are derived. Quadrature errors and their effect on the finite element solution of boundary value problems using linear splines are analyzed in Section 4. For simplicity, we assume  $\Omega$  can be triangulated into uniform equilateral triangles and restrict our analysis to the uniform meshes only. The calculation can be generalized to non-uniform meshes through affine transformation.

## 2. Notation and preliminaries

Denote by  $\tau$  a triangulation of  $\Omega$ , by  $\Omega_h$  the set of all vertices of triangles  $T$  in  $\tau$ , and by  $\mathring{\Omega}_h$  the set of all interior nodes of  $\Omega_h$ .

A linear operator  $L_x: \Omega_h \rightarrow R$  is called an  $q$ -local operator for some integer  $q$  if the value of  $L_x(y)$  at  $y \in \Omega_h$  depends only on the values of  $y$  in a  $q$ -neighbour of  $x$ ; more precisely,  $L_x(y) = 0$  for all values of  $y$  in  $\Omega_h$  with  $|x-y| > q$ , where  $|x-y|$  is the minimum number of edges connecting the two nodes  $x$  and  $y$  along some sides of triangles in  $\tau$ , (see [8] for further detail).

Let  $T = A_0A_1A_2$  be a triangle in a polygonal domain  $\Omega$ . We will denote by  $\xi_i$ ,  $i=0, 1, 2$  the three affine functions defined by the equation  $\xi_i(A_j) = \delta_{ij}$ , where  $\delta$  is the Kronecker delta function. Since  $A_0, A_1, A_2$  are affinely independent, any point  $x \in \Omega$  can be uniquely represented as  $x = \xi_0(x)A_0 + \xi_1(x)A_1 + \xi_2(x)A_2$ ,  $\xi_0 + \xi_1 + \xi_2 = 1$ . We will refer to  $\xi_0, \xi_1, \xi_2$  the barycentric coordinates of  $x$  with respect to the triangle  $T = A_0A_1A_2$ . Any polynomial of degree  $n$  on  $\Omega$  can be expressed uniquely as a homogeneous polynomial of degree  $n$  in the barycentric coordinates with respect to a specific triangle  $T = A_0A_1A_2$  in  $\tau$ , or else as a polynomial of degree  $n$  in any two of the coordinates. Define the first order linear differential operators  $D_i$  with respect to the

coordinate  $\xi_i$  by

$$D_i(\xi_i)=0 \text{ and } D_i(\xi_{i\pm 1})=\mp 1.$$

That is, the counter clockwise normalized derivative of a function  $f$  in the direction parallel to the side opposite  $A_i$ .

If  $f$  is a differentiable function mapping  $\Omega$  into  $R$  then we have  $\sum_i D_i f=0$ .

The differential operator can be extended to any order by  $D^\alpha f=D_0^{\alpha_0} D_1^{\alpha_1} D_2^{\alpha_2} f$ , where  $\alpha=(\alpha_0, \alpha_1, \alpha_2)\in N^3$ ,  $D^0$  and  $D_i^0$  denote the identity operator. We denote by  $|\alpha|=\alpha_0+\alpha_1+\alpha_2$  the order of  $D^\alpha$ . Define  $d\mu_T$  and  $d\mu_\Omega$  as Lebesgue measure on  $T$  and  $\Omega$  normalized by  $\int_T d\mu_T=\int_\Omega d\mu_\Omega=1$ . If  $\Omega$  is a polygon and  $\tau$  is a triangulation of  $\Omega$ , then

$$\int_\Omega f d\mu_\Omega=\sum_{T\in\tau} \mu_\Omega(T) \int_T f d\mu_T$$

In particular, if  $\Omega$  is an equilateral triangle of unit side length and  $\tau$  is an equilateral triangulation of  $\Omega$ , then  $\mu_\Omega(T)=h^2$  for all triangles of side length  $h$  in  $\tau$ .

Define  $dx$  as Lebesgue measure along the edge  $A_{i+1}A_{i-1}$  of a triangle  $T=A_0A_1A_2$  normalized by  $\int_{A_{i+1}}^{A_{i-1}} dx=1$ .

Some results for lines and surfaces integrals are given below.

LEMMA 2.1. Let  $f : T \rightarrow R$ , if  $D_i$  exists on the side  $A_{i+1}A_{i-1}$ , then

$$\int_{A_{i+1}}^{A_{i-1}} D_i f(x) dx=f(A_{i-1})-f(A_{i+1})$$

LEMMA 2.2.  $\int_{A_{i+1}}^{A_{i-1}} g D_i f dx=(gf)(A_{i-1})-(gf)(A_{i+1})-\int_{A_{i+1}}^{A_{i-1}} f D_i g dx$

LEMMA 2.3.  $\int_T D_i f d\mu_T=2 \int_{A_{i-1}}^{A_i} f dx-2 \int_{A_i}^{A_{i+1}} f dx$

LEMMA 2.4.  $\int_T g D_i f d\mu_T=2 \int_{A_{i-1}}^{A_i} f g dx-2 \int_{A_i}^{A_{i+1}} f g dx-\int_T f D_i g d\mu_T$

LEMMA 2.5(Holand Bell [6]).

$$\int_T \xi_0^{s_0} \xi_1^{s_1} \xi_2^{s_2} d\mu_T=\frac{s_0!s_1!s_2!2!}{(s_0+s_1+s_2+2)!}$$

Denote by  $H^k$ ,  $k \geq 0$  the Sobolev space of real valued functions which together with their generalized derivatives up to the  $k^{th}$  order are square integrable

over  $\mathcal{Q}$ . A norm on a triangle  $T$  is given by  $\|u\|_{k,T} = \left\{ \sum_{i \leq k} |u|_{i,T}^2 \right\}^{1/2}$ , where  $|u|_{i,T} = \left\{ \sum_{|\alpha|=i} h^{-2i} \int_T (D^\alpha u)^2 du_T \right\}^{1/2}$  is the Sobolev semi-norm on the triangle  $T$ . The Sobolev norm for  $H^k$  can be represented as

$$\|u\|_{k,\mathcal{Q}} = \left\{ \sum_{T \in \tau} \mu_{\mathcal{Q}}(T) \|u\|_{k,T}^2 \right\}^{1/2}.$$

### 3. Numerical quadrature formulas

For every interior node  $\beta$  of  $\mathcal{Q}_h$ , let  $\phi_\beta$  be the trial function which equals one at  $\beta$  and zero at all other nodes. These pyramid functions  $\phi_\beta$  form a basis for the trial space  $S^h$ .

The basis function  $\phi_\beta(x)$  can be expressed as

$$\phi_\beta(x) = \begin{cases} \xi_\beta & \text{if } x \in H_\beta \\ 0 & \text{otherwise} \end{cases}$$

where  $H_\beta$  is the hexagon formed by the six triangles in  $\tau$  with the common vertex  $\beta$  and  $\xi_\beta$  is the barycentric coordinate of  $x$  with respect to any of the triangles in  $H_\beta$ .

In this section, we shall derive two quadratures for  $F_\beta = \int_{\mathcal{Q}} f \phi_\beta d\mu_{\mathcal{Q}}$ . The first one is a 0-local operator while the second one is a 1-local operator. A 0-local quadrature  $\tilde{F}_\beta$  for the integral  $F_\beta$  is given by  $\tilde{F}_\beta = af(\beta)$ . The constant  $a$  can be determined as follows:

$$\int_{\mathcal{Q}} \phi_\beta d\mu_{\mathcal{Q}} - a = 0.$$

It follows that

$$a = \int_{\mathcal{Q}} \phi_\beta d\mu_{\mathcal{Q}} = 6\mu_{\mathcal{Q}}(T) \int_T \xi_\beta d\mu_T = 2\mu_{\mathcal{Q}}(T).$$

It is easy to verify that the 0-local quadrature is exact for all polynomials of degrees less than or equal to 1.

To obtain numerical quadrature with higher order of accuracy, we require the following lemma:

LEMMA 3.1. Let  $x_j$ ,  $j=1,2,\dots,6$  be the six vertices of a hexagon  $H_{x_0}$  and  $\phi$  be a quadratic polynomial which takes the value 1 along the edges  $x_0x_1$  and  $x_3x_4$  and vanishes along the line  $x_2x_5$  of the hexagon  $H_{x_0}$ . Then

$$\int_{\mathcal{Q}} \phi \phi_{x_0} d\mu_{\mathcal{Q}} = \frac{\mu_{\mathcal{Q}}(T)}{3}$$

PROOF. Denote the barycentric coordinates of a point  $x \in H_{x_0}$  with respect to the triangle  $T_j = x_0 x_j x_{j+1}$  by  $(\xi, \eta, \kappa)$ . Then the polynomial  $\phi$  can be represented in terms of the barycentric coordinates of  $x$  with respect to the six triangles  $T_j$  as follows:

$$\phi(\xi, \eta, \kappa) = \begin{cases} \kappa^2 & \text{in } T_2 \text{ and } T_5 \\ \eta^2 + 2\eta\kappa + \kappa^2 & \text{in } T_3 \text{ and } T_6 \\ \eta^2 & \text{in } T_1 \text{ and } T_4. \end{cases}$$

It follows that

$$\begin{aligned} \int_{\mathcal{Q}} \phi \phi_{\beta} d\mu_{\mathcal{Q}} &= 4\mu_{\mathcal{Q}}(T) \int_T \xi(\eta^2 + \kappa^2 + \eta\kappa) d\mu_T \\ &= 4\mu_{\mathcal{Q}}(T) \left( \frac{2!2!}{5!} + \frac{2!2!}{5!} + \frac{2!}{5!} \right) \\ &= \frac{\mu_{\mathcal{Q}}(T)}{3}. \end{aligned}$$

From the result of Lemma 3.1, we observe that the 0-local numerical quadratures  $\tilde{F}_{\beta} = 2\mu_{\mathcal{Q}}(T)$  is not exact for all polynomials of degree 2. Now we shall derive another quadrature which is exact for polynomials of higher degree.

Consider the 1-local quadrature  $\tilde{F}_{\beta} = af(\beta) + b \sum_{|\beta-\gamma|=1} f(\gamma)$ , the two parameters  $a$  and  $b$  can be determined as follows:

$$\begin{cases} \int_{\mathcal{Q}} \phi_{\beta} d\mu_{\mathcal{Q}} - a - 6b = 0 \\ \int_{\mathcal{Q}} \phi \phi_{\beta} d\mu_{\mathcal{Q}} - 4b = 0. \end{cases}$$

It follows from Lemma 3.1 that  $b = \frac{\mu_{\mathcal{Q}}(T)}{12}$  and  $a = \frac{3}{2}\mu_{\mathcal{Q}}(T)$ .

#### 4. Quadrature errors and their effect on the numerical solution of boundary value problems

In this section, we analyze the errors of the two quadratures derived in Section 3 and discuss the effect of quadrature errors to the finite element solution of the boundary value problems using linear splines.

For each  $\beta \in \overset{\circ}{\mathcal{Q}}_h$  let  $E(f, \beta)$  denote the quadrature error  $F_{\beta} - \tilde{F}_{\beta}$ .

For simplicity, we denote by  $x_0$  the centre  $\beta$  of a hexagon  $H_{\beta}$  and by  $x_j, j=1, \dots, 6$  the six vertices of  $H_{\beta}$ .

THEOREM 4.1. If  $f \in H^2$  and  $\tilde{F}_\beta(f) = 2\mu_Q(T) f(\beta)$ , then

$$\left( \sum_{\beta \in \mathcal{Q}_k} |E(f, \beta)|^2 \right)^{1/2} \leq \left[ \frac{23}{560} \mu_Q(T) \right]^{1/2} h^2 |f|_{2, \mathcal{Q}}.$$

To prove the theorem, we need the following two auxiliary lemmas.

LEMMA 4.1. If  $f \in H^2$  and  $P(\xi_0)$  is a real valued function of  $\xi_0$  defined on each of the triangular elements  $T_j = x_0 x_j x_{j+1}$ , then

$$\sum_{j=1}^6 \int_{x_j}^{x_0} P(\xi_0) D_{x_j x_0} f dx = - \sum_{j=1}^6 \int_{T_j} \frac{1}{2} P(\xi_0) D_{00} f d\mu_{T_j}$$

PROOF. The result of the lemma follows from the decomposition of the derivatives  $D_{x_j x_0} f$  into the sum of the two derivatives  $D_0 f$  and  $D_1 f$ .

LEMMA 4.2. If  $f \in H^2$ , then the error of the 0-local numerical quadrature has a representation of the form:

$$\begin{aligned} E(f, \beta) &= \int_Q f \phi_\beta d\mu_Q - 2h^2 f(\beta) \\ &= \sum_{j=1}^6 \mu_Q(T) \int_{T_j} \left( \frac{\xi_0^2}{2} - \frac{\xi_0^3}{3} - \frac{\xi_0 \xi_1 \xi_2}{2} \right) D_{00} f d\mu_{T_j} \end{aligned}$$

$$\begin{aligned} \text{PROOF. } E(f, \beta) &= \int_Q f \phi_\beta d\mu_Q - 2h^2 f(\beta) \\ &= \sum_{j=1}^6 \mu_Q(T) \left[ \int_{T_j} f \xi_0 d\mu_{T_j} - \int_{T_j} f(x_0) \xi_0 d\mu_{T_j} \right] \\ &= \sum_{j=1}^6 \mu_Q(T) \int_{T_j} [f - f(\beta)] \xi_0 d\mu_{T_j} \\ &= \sum_{j=1}^6 \mu_Q(T) \int_{T_j} [f - f(\beta)] \left[ D_0 \frac{\xi_0}{2} (\xi_2 - \xi_1) \right] d\mu_{T_j} \end{aligned}$$

It follows from Lemma 2.4 and the fact  $\xi_1$  and  $\xi_2$  vanish on  $x_{j+1} x_0$  and  $x_0 x_j$  respectively that

$$\begin{aligned} E(f, \beta) &= \sum_{j=1}^6 \mu_Q(T) \left[ \int_{x_{j+1}}^{x_0} \xi_0 (1 - \xi_0) (f - f(\beta)) dx \right. \\ &\quad \left. + \int_{x_0}^{x_j} \xi_0 (1 - \xi_0) (f - f(\beta)) dx + \int_{T_j} D_0 \left( \frac{1}{2} \xi_0 \xi_1 \xi_2 \right) D_0 f d\mu_{T_j} \right] \\ &= \sum_{j=1}^6 \mu_Q(T) \left[ 2 \int_{x_{j+1}}^{x_0} \xi_0 (1 - \xi_0) (f - f(\beta)) dx + \int_{x_{j+1}}^x \xi_0 \xi_1 \xi_2 D_0 f dx \right. \\ &\quad \left. - \int_{x_0}^{x_j} \xi_0 \xi_1 \xi_2 D_0 f dx - \int_{T_j} \frac{1}{2} \xi_0 \xi_1 \xi_2 D_{00} f d\mu_{T_j} \right] \end{aligned}$$

$$\begin{aligned} &= \sum_{j=1}^6 \mu_{\mathcal{Q}}(T) \left[ 2 \left( \frac{\xi_0^2}{2} - \frac{\xi_0^3}{3} \right) (f - f(\beta)) \Big|_{x_j}^{x_0} - 2 \int_{x_j}^{x_0} \left( \frac{\xi_0^2}{2} - \frac{\xi_0^3}{3} \right) D_{x_j x_0} f dx \right. \\ &\quad \left. - \int_{T_j} \frac{1}{2} \xi_0 \xi_1 \xi_2 D_{00} f d\mu_{T_j} \right] \\ &= -\mu_{\mathcal{Q}}(T) \sum_{j=1}^6 \left[ \int_{x_j}^{x_0} \left( \frac{\xi_0^2}{2} - \frac{2}{3} \xi_0^3 \right) D_{x_j x_0} f dx + \int_{T_j} \frac{1}{2} \xi_0 \xi_1 \xi_2 D_{00} f d\mu_{T_j} \right] \end{aligned}$$

It follows from Lemma 4.1 that the above equation gives Lemma 4.2.

PROOF OF THEOREM 4.1. Application of the triangle inequality and the Cauchy-Schwarz inequality to the equation in Lemma 4.2, we get

$$\begin{aligned} |E(f, \beta)| &\leq \sum_{j=1}^6 \mu_{\mathcal{Q}}(T) \left[ \int_{T_j} \left( \frac{1}{2} \xi_0^2 - \frac{1}{3} \xi_0^2 - \frac{1}{2} \xi_0 \xi_1 \xi_2 \right)^2 d\mu_{T_j} \right]^{1/2} \|D_{00} f\|_{L^2(T_j)} \\ &= \left( \frac{23}{3360} \right)^{1/2} \mu_{\mathcal{Q}}(T) \sum_{j=1}^6 \|D_{00} f\|_{L^2(T_j)} \end{aligned}$$

It follows that

$$\begin{aligned} \sum_{\beta \in \mathcal{Q}_h} |E(f, \beta)|^2 &\leq \frac{23}{3360} \mu_{\mathcal{Q}}(T) \sum_{\beta \in \mathcal{Q}_h} \mu_{\mathcal{Q}}(T) \left[ \sum_{j=1}^6 \|D_{00} f\|_{L^2(T_j)} \right]^2 \\ &\leq \frac{23}{560} \mu_{\mathcal{Q}}(T) \sum_{\beta \in \mathcal{Q}_h} \mu_{\mathcal{Q}}(T) \left[ \sum_{j=1}^6 \|D_{00} f\|_{L^2(T_j)} \right]^2 \end{aligned}$$

We observe that for each  $T = x_0 x_1 x_2 \in \tau$  and each  $i$ , the term  $\|D_{ii} f\|_{L^2(T)}$  appears in the right hand side of the above inequality at most once; thus we have

$$\begin{aligned} \left( \sum_{\beta \in \mathcal{Q}_h} |E(f, \beta)|^2 \right)^{1/2} &\leq \left[ \frac{23}{560} \mu_{\mathcal{Q}}(T) \right]^{1/2} \left[ \sum_{T \in \tau} \mu_{\mathcal{Q}}(T) \sum_{i=1}^3 \|D_{ii} f\|_{L^2(T)}^2 \right]^{1/2} \\ &\leq \left[ \frac{23}{560} \mu_{\mathcal{Q}}(T) \right]^{1/2} h^2 |f|_{2, \mathcal{Q}}. \end{aligned}$$

LEMMA 4.2. If  $f \in H^4$  and  $\tilde{F}_{\beta}$  is the 1-local numerical quadrature of  $F_{\beta}$ , then

$$\left( \sum_{\beta \in \mathcal{Q}_h} |E(f, \beta)|^2 \right)^{1/2} \leq 0.07208 (\mu_{\mathcal{Q}}(T))^{1/2} h^4 |f|_{4, \mathcal{Q}}$$

To prove the theorem, we need the following auxiliary lemmas.

LEMMA 4.3. If  $f \in H^4$  and  $P(\xi_0)$  is a real valued function of  $\xi_0$  defined on each of the triangular elements  $T_j = x_0 x_j x_{j+1}$ , then,

$$\sum_{j=1}^6 \left[ \int_{x_0}^{x_j} P(\xi_0) D_{200} f dx - \int_{x_{j+1}}^{x_0} P(\xi_0) D_{100} f dx \right]$$

$$= \sum_{j=1}^6 \int_{T_j} \frac{1}{2} P(\xi_0) (D_{0000} f - \frac{1}{2} D_{0012} f) d\mu_{T_j}$$

PROOF. 
$$\begin{aligned} & \sum_{j=1}^6 \left[ \int_{x_0}^{x_j} P(\xi_0) D_{200} f dx - \int_{x_{j+1}}^{x_0} P(\xi_0) D_{100} f dx \right] \\ &= \sum_{j=1}^6 \left[ \int_{x_0}^{x_j} P(\xi_0) (-D_{000} f - D_{100} f) dx - \int_{x_{j+1}}^{x_0} P(\xi_0) (-D_{000} f - D_{200} f) dx \right] \\ &= \sum_{j=1}^6 \left[ \int_{x_{j+1}}^{x_0} P(\xi_0) D_{000} f dx - \int_{x_0}^{x_j} P(\xi_0) D_{000} f dx \right. \\ & \quad \left. + \int_{x_{j+1}}^{x_0} P(\xi_0) D_{200} f dx - \int_{x_0}^{x_j} P(\xi_0) D_{100} f dx \right] \\ &= \sum_{j=1}^6 \left[ \int_{T_j} \frac{1}{2} P(\xi_0) D_{0000} f d\mu_{T_j} + \int_{x_{j+1}}^x P(\xi_0) D_{200} f dx - \int_{x_0}^{x_j} P(\xi_0) D_{100} f dx \right] \end{aligned}$$

We note that the derivative  $-D_{100} f$  along  $x_j x_0$  with respect to  $T_j$  is the same as  $D_{220} f$  with respect to  $T_{j-1}$ , and  $D_{200} f$  along  $x_{j+1} x_0$  with respect to  $T_j$  is the same as  $-D_{110} f$  with respect to  $T_{j+1}$ . Thus, these derivatives can be divided into two equal parts, half of them will be added to the line integral of the adjacent triangle. It follows that

$$\begin{aligned} & \sum_{j=1}^6 \left[ \int_{x_0}^{x_j} P(\xi_0) D_{200} f dx - \int_{x_{j+1}}^{x_0} P(\xi_0) D_{100} f dx \right] \\ &= \sum_{i=1}^6 \left[ \int_{T_i} \frac{1}{2} P(\xi_0) D_{0000} f d\mu_{T_i} + \int_{x_{i+1}}^{x_0} \frac{1}{2} P(\xi_0) (D_{200} f + D_{220} f) dx \right. \\ & \quad \left. - \int_{x_0}^{x_j} \frac{1}{2} P(\xi_0) (D_{100} f + D_{110} f) dx \right] \\ &= \sum_{j=1}^6 \left[ \int_{T_j} \frac{1}{2} P(\xi_0) D_{0000} f d\mu_{T_j} - \int_{x_{j+1}}^{x_0} \frac{1}{2} P(\xi_0) D_{210} f dx \right. \\ & \quad \left. + \int_{x_0}^{x_j} \frac{1}{2} P(\xi_0) D_{120} f dx \right] \end{aligned}$$

which gives Lemma 4.3.

LEMMA 4.4. *If  $f \in H^4$ , then the error functional of the 1-local numerical quadrature has a representation of the form*

$$\begin{aligned} E(f, \beta) &= \int_{\mathcal{D}} f \phi_{\beta} d\mu_{\mathcal{D}} - \mu_{\mathcal{D}}(T) \left[ \frac{3}{2} f(\beta) + \frac{1}{12} \sum_{j=1}^6 f(x_j) \right] \\ &= \frac{1}{24} \mu_{\mathcal{D}}(T) \sum_{j=1}^6 \int_{T_j} \left\{ \left[ \left( \frac{1}{2} + \xi_0 - 8\xi_0^2 + 5\xi_0^3 + \xi_0 \xi_1 \xi_2 \right) \xi_1 \xi_2 \right. \right. \end{aligned}$$



$$\begin{aligned}
 & -\frac{\xi_0}{2} - \frac{\xi_0^2}{4} + 3\xi_0^3 - \frac{13}{4}\xi_0^4 + \xi_0^5 \Big] D_{0000} f \\
 & + \frac{1}{2} \left( \frac{\xi_0}{2} + \frac{\xi_0^2}{4} - 3\xi_0^3 + \frac{13}{4}\xi_0^4 - \xi_0^5 \right) D_{0012} f \Big) d\mu_{T_j}
 \end{aligned}$$

PROOF. We shall only give a brief proof for this lemma.

$$\begin{aligned}
 E(f, \beta) &= \left[ \int_{\mathcal{Q}} f \phi_{\beta} d\mu_{\mathcal{Q}} - 2h^2 f(\beta) \right] + \frac{\mu_{\mathcal{Q}}(T)}{12} \sum_{j=1}^6 [f(\beta) - f(x_j)] \\
 &= \left[ \int_{\mathcal{Q}} f \phi_{\beta} d\mu_{\mathcal{Q}} - 2h^2 f(\beta) \right] + \frac{1}{12} \mu_{\mathcal{Q}}(T) \sum_{j=1}^6 \int_{x_j}^{x_0} D_{x_j x_0} f dx
 \end{aligned}$$

It follows from Lemma 4.2 and Lemma 4.1 that

$$E(f, \beta) = \sum_{j=1}^6 \frac{\mu_{\mathcal{Q}}(T)}{24} \int_{T_j} (-1 + 12\xi_0^2 - 8\xi_0^3 - 12\xi_0 \xi_1 \xi_2) D_{00} f d\mu_{T_j}$$

which reduces to the following result

$$\begin{aligned}
 E(f, \beta) &= \frac{\mu_{\mathcal{Q}}(T)}{24} \sum_{j=1}^6 \left[ \int_{x_{j+1}}^{x_0} (-1 - 2\xi_0 + 16\xi_0^2 - 10\xi_0^3) \xi_2 D_{00} f dx \right. \\
 &+ \int_{x_0}^{x_j} (-1 - 2\xi_0 + 16\xi_0^2 - 10\xi_0^3) \xi_1 D_{00} f dx \\
 &- \int_{T_j} \frac{1}{2} (-1 - 2\xi_0 + 16\xi_0^2 - 10\xi_0^3) (\xi_2 - \xi_1) D_{000} f d\mu_{T_j} \\
 &\left. - \int_{T_j} D_0 (\xi_0 \xi_1 \xi_2^2) D_{000} f d\mu_{T_j} \right]
 \end{aligned}$$

After further evaluation, we have

$$\begin{aligned}
 E(f, \beta) &= \frac{\mu_{\mathcal{Q}}(T)}{24} \sum_{j=1}^6 \left[ \int_{x_0}^{x_j} \left( -\xi_0 - \frac{\xi_0^2}{2} + 6\xi_0^3 - \frac{13}{2}\xi_0^4 + 2\xi_0^5 \right) D_{200} f dx \right. \\
 &- \int_{x_{j+1}}^{x_0} \left( -\xi_0 - \frac{\xi_0^2}{2} + 6\xi_0^3 - \frac{13}{2}\xi_0^4 + 2\xi_0^5 \right) D_{100} f dx \\
 &\left. + \int_{T_j} \left( \frac{1}{2} + \xi_0 - 8\xi_0^2 + 5\xi_0^3 + \xi_0 \xi_1 \xi_2 \right) \xi_1 \xi_2 D_{0000} f d\mu_{T_j} \right]
 \end{aligned}$$

It follows from Lemma 4.3 that the above equation gives Lemma 4.4.

PROOF OF THEOREM 4.2. Applying the triangle inequality and the *Cauchy-Schwarz inequality* to the equation in Lemma 4.4, we get

$$|E(f, \beta)| \leq \frac{\mu_{\mathcal{Q}}(T)}{24} \sum_{j=1}^6 (c_0 \|D_{0000} f\|_{L^2(T_j)} + c_1 \|D_{0012} f\|_{L^2(T_j)})$$

where  $c_0 = 0.08495$  and  $c_1 = 0.04297$ . Applying the *Cauchy-Schwarz inequality*

to the above inequality, we get

$$|E(f, \beta)| \leq \frac{\mu_{\mathcal{Q}}(T)}{24} (6c_0^2 + 6c_1^2)^{1/2} \left\{ \sum_{j=1}^6 (\|D_{0000} f\|_{L^2(T_j)}^2 + \|D_{0012} f\|_{L^2(T_j)}^2) \right\}^{1/2}$$

It follows that

$$\begin{aligned} \left( \sum_{\beta \in \mathcal{Q}_h} |E(f, \beta)|^2 \right)^{1/2} &\leq 0.07208 (\mu_{\mathcal{Q}}(T))^{1/2} \left\{ \sum_{\beta \in \mathcal{Q}_h} \mu_{\mathcal{Q}}(T) \sum_{j=1}^6 (\|D_{0000} f\|_{L^2(T_j)}^2 \right. \\ &\quad \left. + \|D_{0012} f\|_{L^2(T_j)}^2) \right\}^{1/2} \end{aligned}$$

for each  $T \in \tau$ , since the terms  $\|D_{0000} f\|_{L^2(T_j)}^2$  and  $\|D_{0012} f\|_{L^2(T_j)}^2$  appear in the right hand side of the above inequality at most once, thus, we have

$$\begin{aligned} \left( \sum_{\beta \in \mathcal{Q}_h} |E(f, \beta)|^2 \right)^{1/2} &\leq 0.07208 (\mu_{\mathcal{Q}}(T))^{1/2} \left( \sum_{T \in \tau} \mu_{\mathcal{Q}}(T) \sum_{|\alpha|=4} \|D^{\alpha} f\|_{L^2(T)}^2 \right)^{1/2} \\ &= 0.07208 (\mu_{\mathcal{Q}}(T))^{1/2} h^4 |f|_{4, \mathcal{Q}}. \end{aligned}$$

If the load vector  $F_{\beta} = (f, \phi_{\beta})$  of (3) is approximated by a numerical quadrature  $\tilde{F}_{\beta}$ , then we are solving

$$(\tilde{u}^h, \phi_{\beta}) = \tilde{F}_{\beta}, \quad \beta \in \mathcal{Q}_h \quad (4)$$

where  $\tilde{u}^h = \sum_{\beta \in \mathcal{Q}_h} \tilde{\lambda}_{\beta} \phi_{\beta}$  is a solution to the linear system (4).

From (4), we get

$$(u^h - \tilde{u}^h, \phi_{\beta}) = F_{\beta} - \tilde{F}_{\beta} = E(F, \beta)$$

It follows that

$$(u^h - \tilde{u}^h, u^h - \tilde{u}^h) = \sum_{\beta \in \mathcal{Q}_h} (\lambda_{\beta} - \tilde{\lambda}_{\beta}) E(f, \beta)$$

which reduces to

$$\|u^h - \tilde{u}^h\|_a^2 \leq \sum_{\beta \in \mathcal{Q}_h} |\lambda_{\beta} - \tilde{\lambda}_{\beta}| |E(f, \beta)|$$

By applying the Cauchy-Schwarz inequality to the right hand side of the above inequality, we get

$$\|u^h - \tilde{u}^h\|_a^2 \leq \left( \sum_{\beta \in \mathcal{Q}_h} (\lambda_{\beta} - \tilde{\lambda}_{\beta})^2 \right)^{1/2} \left( \sum_{\beta \in \mathcal{Q}_h} |E(f, \beta)|^2 \right)^{1/2}. \quad (5)$$

To obtain an upper bound for  $\left( \sum_{\beta \in \mathcal{Q}_h} (\lambda_{\beta} - \tilde{\lambda}_{\beta})^2 \right)^{1/2}$  in terms of the  $L^2$  norm  $\|u^h - \tilde{u}^h\|_{L^2(\mathcal{Q})}$ , we get the following lemma.

LEMMA 4.5. If  $u(x) = \sum_{\beta \in \mathcal{Q}_h} \lambda_{\beta} \phi_{\beta}(x)$  vanishes on  $\partial\mathcal{Q}$ , then

$$\|u\|_{L^2(\Omega)}^2 \geq \frac{1}{2} \mu_{\Omega}(T) \sum_{\beta \in \mathcal{Q}_h} \lambda_{\beta}^2$$

PROOF.  $\|u\|_{L^2(\Omega)}^2 = \int_{\Omega} u^2 d\mu_{\Omega}$

$$= \mu_{\Omega}(T) \sum_{T \in \mathcal{T}} \int_T (\lambda_{\beta} \phi_{\beta} + \lambda_{\gamma} \phi_{\gamma} + \lambda_{\kappa} \phi_{\kappa})^2 d\mu_T$$

$$= \frac{1}{6} \mu_{\Omega}(T) \sum_{T \in \mathcal{T}} (\lambda_{\beta}^2 + \lambda_{\gamma}^2 + \lambda_{\kappa}^2 + \lambda_{\alpha} \lambda_{\beta} + \lambda_{\beta} \lambda_{\kappa} + \lambda_{\kappa} \lambda_{\beta})$$

where  $\lambda_{\beta}, \lambda_{\gamma}, \lambda_{\kappa}$  are the values of  $u$  at the three vertices of  $T = \beta\gamma\kappa$ .

Since  $\lambda_{\beta}^2 + \lambda_{\gamma}^2 + \lambda_{\kappa}^2 + 2(\lambda_{\beta}\lambda_{\gamma} + \lambda_{\gamma}\lambda_{\kappa} + \lambda_{\kappa}\lambda_{\beta}) \geq 0$  for all  $\lambda_{\beta}, \lambda_{\gamma}, \lambda_{\kappa} \in \mathbb{R}$ , we have

$$\sum_{T \in \mathcal{T}} (\lambda_{\beta}^2 + \lambda_{\gamma}^2 + \lambda_{\kappa}^2 + \lambda_{\beta}\lambda_{\gamma} + \lambda_{\gamma}\lambda_{\kappa} + \lambda_{\kappa}\lambda_{\beta}) \geq \sum_{T \in \mathcal{T}} \frac{1}{2} (\lambda_{\beta}^2 + \lambda_{\gamma}^2 + \lambda_{\kappa}^2) \tag{6}$$

Since  $\lambda_{\beta} = 0$  for all  $\beta \in \partial\Omega$  and for each  $\beta \in \mathring{\mathcal{Q}}_h$ , there are six triangles  $T$  in  $\tau$  with the common vertex  $\beta$ . Thus the right hand side of the inequality (6) can be written as  $3 \sum_{\beta \in \mathcal{Q}_h} \lambda_{\beta}^2$ . It follows that

$$\|u\|_{L^2(\Omega)}^2 \geq \frac{1}{2} \mu_{\Omega}(T) \sum_{\beta \in \mathcal{Q}_h} \lambda_{\beta}^2$$

Since  $u^h - \tilde{u}^h$  is a piecewise linear function on  $\Omega$  and vanishes on the boundary  $\partial\Omega$ , we can apply Lemma 4.5 to inequality (5) to get

$$\|u^h - \tilde{u}^h\|_a^2 \leq \left( \frac{2}{\mu_{\Omega}(T)} \right)^{1/2} \|u^h - \tilde{u}^h\|_{L^2(\Omega)} \left( \sum_{\beta \in \mathcal{Q}_h} |E(f, \beta)|^2 \right)^{1/2}$$

Since the energy norm  $\|\cdot\|_a$  is equivalent to the Sobolev norm  $\|\cdot\|_{1,\Omega}$ ; we have

$$\|u^h - \tilde{u}^h\|_a \geq c \|u^h - \tilde{u}^h\|_{1,\Omega} \geq c \|u^h - \tilde{u}^h\|_{L^2(\Omega)}$$

for some constant  $c > 0$ . It follows that

$$\|u^h - \tilde{u}^h\|_a \leq \frac{\sqrt{2}}{c(\mu_{\Omega}(T))^{1/2}} \left( \sum_{\beta \in \mathcal{Q}_h} |E(f, \beta)|^2 \right)^{1/2}$$

If the 0-local numerical quadrature is used, from Theorem 4.1, we have

$$\|u^h - \tilde{u}^h\|_a \leq \left( \frac{23}{280} \right)^{1/2} \frac{1}{c} h^2 |f|_{2,\Omega}$$

and if the 1-local numerical quadrature is used, from Theorem 4.2, we have

$$\|u^h - \tilde{u}^h\|_a \leq \frac{0.1019}{c} h^4 |f|_{4,\Omega}$$

If  $u$  is the exact solution to  $Lu = f$ , from the triangle inequality, we get

$$\|u - \tilde{u}^h\|_a \leq \|u - \tilde{u}^h\|_a + \|u^h - \tilde{u}^h\|_a$$

Since the energy norm  $\|u - \tilde{u}^h\|_a$  has an order of accuracy  $O(h^2)$  [9], and the energy norm  $\|u^h - \tilde{u}^h\|_a$  has an order of accuracy  $O(h^2)$  and  $O(h^4)$  for the 0-local and 1-local numerical quadrature respectively, both numerical quadratures are consistent in the energy norm. That is, the solution still has an order of accuracy  $O(h^2)$  in the energy norm for the 0-local and 1-local numerical quadratures.

Department of Industrial & Systems Engineering  
National University of Singapore  
Kent Ridge, Singapore 0511

#### REFERENCES

- [1] Birkhoff, G., *The numerical solution of elliptic equations*, SIAM, Regional Conference Series, Vol.1(1971).
- [2] Bramble, J.H. and Zlamal, M., *Triangular elements in the finite element method*, Math. of Comp. 24, 809-821(1970).
- [3] Ciarlet, P.G., *The finite element method for elliptic problems*, North-Holland, 1978.
- [4] Courant, R., *Variational method for the solution of problems of equilibrium and vibrations*, Bull. Amer. Math. Soc. 49, 1-23(1943).
- [5] Herbold, R.J. and Varga, R.S., *The effect of quadrature errors in the numerical solution of two-dimensional boundary value problems by variational techniques*, Aequationes Math. 7, 36-58(1972).
- [6] Holand, I., and Bell, K., eds., *Finite element methods in stress analysis*, Tapir, Trondheim, Norway, 1969.
- [7] Ong, H.L., *Finite element solution to the second order elliptic problems*, Journal of Computational and Applied Mathematics, Vol.10, 39-43(1984).
- [8] Ong, H.L., *Fast approximate solution of large scale sparse linear systems*, Journal of Computational and Applied Mathematics, Vol.10, 45-54(1984).
- [9] Strang, G. and Fix, G., *An analysis of the finite element method*, Prentice-Hall, Englewood Cliffs, New Jersey(1973).