On the fourth order accuracy of the finite difference implementation of $C^0$-$Q^2$ finite element method for elliptic equations

Hao Li † and Xiangxiong Zhang ‡

Department of Mathematics, Purdue University, 150 N. University Street, West Lafayette, IN 47907-2067, USA

[Received on ]

The classical continuous finite element method with Lagrangian $Q^2$ basis reduces to a finite difference scheme when all the integrals are replaced by the $3 \times 3$ Gauss-Lobatto quadrature. By deriving an explicit representation of the quadrature error, we prove that this finite difference scheme is fourth order accurate in the discrete 2-norm for an elliptic equation $-\nabla(a\nabla u) + b \cdot \nabla u + cu = f$ with Dirichlet boundary conditions, which is a superconvergence result of function values.

Keywords: Superconvergence, fourth order accurate discrete Laplacian, elliptic equations, finite difference implementation of finite element method, $3 \times 3$ Gauss-Lobatto quadrature.

1. Introduction

1.1 Motivation

In this paper we consider solving a two-dimensional elliptic equation with smooth coefficients on a rectangular domain with continuous finite element method using tensor product polynomials of degree two on a rectangular mesh. Consider the following model problem as an example: a variable coefficient Poisson equation $-\nabla(a(x)\nabla u) + f, a(x) > 0$ on a square domain $\Omega = (0, 1) \times (0, 1)$ with homogeneous Dirichlet boundary conditions. The variational form is to find $u \in H^1_0(\Omega) = \{v \in H^1(\Omega) : v|_{\partial \Omega} = 0\}$ satisfying

$$A(u, v) = (f, v), \quad \forall v \in H^1_0(\Omega),$$

where $A(u, v) = \int_\Omega a\nabla u \cdot \nabla v dx dy, (f, v) = \int_\Omega f v dx dy$. Let $h$ be the mesh size of an uniform rectangular mesh and $V^h_0 \subseteq H^1_0(\Omega)$ be the continuous finite element space consisting of piecewise $Q^k$ polynomials (i.e., tensor product of piecewise polynomials of degree $k$), then the $C^0$-$Q^k$ finite element solution is defined as $u_h \in V^h_0$ satisfying

$$A(u_h, v_h) = (f, v_h), \quad \forall v_h \in V^h_0. \quad (1.1)$$

Standard error estimates of (1.1) are $\|u - u_h\|_1 \leq Ch^k\|u\|_{k+1}$ and $\|u - u_h\|_0 \leq Ch^{k+1}\|u\|_{k+1}$ where $\| \cdot \|_k$ denotes $H^k(\Omega)$-norm, see [Ciarlet (1991)]. For $k \geq 2$, $O(h^{k+1})$ superconvergence for the gradient at Gauss quadrature points and $O(h^{k+2})$ superconvergence for functions values at Gauss-Lobatto quadrature points were proven for one-dimensional case in [Lesaint & Zlamal (1979); Chen (1979); Bakker (1982)] and for two-dimensional case in [Douglas et al. (1974); Wahlbin (2006); Chen (2001); Lin & Yan (1996)].

†Email: li2497@purdue.edu
‡Corresponding author. Email: zhan1966@purdue.edu
To implement the scheme (1.1), integrals are usually approximated by quadrature. In practice the most convenient choice of quadrature for $Q^2$ element is to use $3 \times 3$ Gauss-Lobatto quadrature rule since the quadrature points are also the exact degree of freedoms to represent the Lagrangian $Q^2$ basis, see Figure 1. Such a quadrature scheme can be denoted as finding $u_h \in V_h^0$ satisfying

$$A_h(u_h, v_h) = \langle f, v_h \rangle_h, \quad \forall v_h \in V_h^0,$$

(1.2)

where $A_h(u_h, v_h)$ and $\langle f, v_h \rangle_h$ denote using tensor product of 3-point Gauss Lobatto quadrature for integrals $A(u_h, v_h)$ and $\langle f, v_h \rangle$ respectively.

![Figure 1. An illustration of Lagrangian $Q^2$ element and the $3 \times 3$ Gauss-Lobatto quadrature.](image)

It is well-known that many classical finite difference schemes are exactly finite element methods with specific quadrature scheme, see Ciarlet (1991). The scheme (1.2) becomes a finite difference scheme, which will be explained in Section 7. On the one hand, such a finite difference implementation provides an efficient way for assembling the stiffness matrix especially for a variable coefficient problem. On other hand, (1.2) is the variational approach to construct a high order accurate finite difference scheme with advantages inherited from the variational formulation such as symmetry of stiffness matrix and easiness of handling boundary conditions in high order schemes.

Classical quadrature error estimates imply that standard finite element error estimates still hold for (1.2), see Ciarlet & Raviart (1972); Ciarlet (1991). The focus of this paper is to prove that the superconvergence of function values at Gauss-Lobatto points still holds. To be more specific, for Dirichlet type boundary conditions, we will show that (1.2) is a fourth order accurate finite difference scheme in the discrete 2-norm under suitable smoothness assumptions on the exact solution and the coefficients.

1.2 Related work and difficulty in using standard tools

The finite element method with Lagrangian quadratic polynomial basis for solving $-\Delta u = f$ on a regular triangular mesh (two adjacent triangles form a rectangle) is equivalent to a finite difference scheme (Whiteman (1975)) since the quadrature using three edge centers and three vertices on a triangle is exact for integrating quadratic polynomials over this triangle thus the quadrature is exact for the bilinear form in the finite element method. Superconvergence of function values in $C^0-P^2$ finite element method at the three vertices and three edge centers can also be proven for solving $-\Delta u = f$ (Chen (2001); Wahlbin (2006)). See also Huang & Xu (2008) for superconvergence of $P^2$ finite element method. Thus one can also construct a fourth order accurate finite difference scheme by using $P^2$ finite element method discussed in Whiteman (1975) for solving $-\Delta u = f$. Since the quadrature is exact for the bilinear form,
the superconvergence results for \(C^0-P^2\) finite element method hold trivially after using the quadrature in the bilinear form, but only for solving \(-\Delta u = f\). For a variable coefficient Poisson equation or a general elliptic problem, since such a quadrature is only third order accurate, we do not expect the fourth order accuracy in the corresponding finite difference scheme.

For computing the bilinear form in the scheme \(1.1\), another convenient implementation is to replace the smooth coefficient \(a(x,y)\) by a piecewise \(Q^2\) polynomial \(a_I(x,y)\) obtained by interpolating \(a(x,y)\) at the quadrature points in each cell shown in Figure 1. Then one can compute the integrals in the bilinear form exactly since the integrand is a polynomial. Superconvergence of function values for such an approximated coefficient scheme was proven in Li & Zhang (2019b) and the proof can be easily extended to higher order polynomials and three-dimensional cases. This result might seem surprising since interpolation error \(a(x,y) - a_I(x,y)\) is of third order. On the other hand, all the tools used in Li & Zhang (2019b) are standard in the literature.

From a practical point of view, \(1.2\) is more interesting since it gives a genuine finite difference scheme. It is straightforward to use standard tools in the literature for showing superconvergence still holds for accurate enough quadrature. Even though the \(3 \times 3\) Gauss-Lobatto quadrature is fourth order accurate, the standard quadrature error estimates cannot be used to establish the fourth order accuracy of \(1.2\). To be specific, in order to extend standard superconvergence proof to the scheme \(1.2\), it is necessary to establish the following consistency estimate:

\[
A(u,v_h) - A_h(u,v_h) = O(h^4 \|u\|_5 \|v_h\|_2).
\]

As will be explained in Remark 3.3 in Section 3.3, such an estimate cannot be obtained by using standard quadrature estimating tools, i.e., the Bramble-Hilbert Lemma. The Bramble-Hilbert Lemma gives a sharp quadrature error estimate for each cell but not for the whole bilinear form since it does not count in the cancellation of some quadrature errors between neighboring cells. In order to obtain a sharp estimate of \(A(u,v_h) - A_h(u,v_h)\), we will derive an explicit error term of the Gauss-Lobatto quadrature, with which standard superconvergence proof can be applied.

We can also rewrite \(1.2\) as a standard finite difference scheme and try to apply traditional finite difference approaches to analyze its convergence order. However, the local truncation error is only second order as will be shown in Section 7.4. The phenomenon that truncation errors have lower orders was named supraconvergence in the literature. The second order local truncation error makes it extremely difficult to establish the fourth order accuracy following any traditional finite difference analysis approaches.

1.3 Contributions and organization of the paper

The main contribution of this paper is to establish the fourth order accuracy of the simple scheme \(1.2\) for a general elliptic equation \(-\nabla(a\nabla u) + b \cdot \nabla u + cu = f\) with Dirichlet boundary conditions. The same proof also applies to Neumann type boundary condition but only 3.5 order of accuracy can be proven, even though fourth order accuracy holds in numerical tests.

This paper is organized as follows. In Section 2 we introduce our notations and assumptions. In Section 3 standard quadrature estimates are reviewed and an explicit 3-point Gauss-Lobatto quadrature error is proposed. Superconvergence of bilinear forms with quadrature is shown in Section 4. Then we prove the main result for homogeneous Dirichlet boundary conditions in Section 5 and for nonhomogeneous Dirichlet boundary conditions in Section 6. Section 7 provides a simple finite difference implementation of the discussed scheme. Section 8 contains numerical tests. Concluding remarks are given in Section 9.
2. Notations and assumptions

2.1 Notations and basic tools

We will use the same notations as in Li & Zhang (2019b):

- We only consider a rectangular domain $\Omega = (0, 1) \times (0, 1)$ with its boundary denoted as $\partial \Omega$.
- Only for convenience, we assume $\Omega_h$ is a uniform rectangular mesh for $\bar{\Omega}$ and $e = [x_e - h, x_e + h] \times [y_e - h, y_e + h]$ denotes any cell in $\Omega_h$ with cell center $(x_e, y_e)$. The assumption of a uniform mesh is not essential to the discussion of superconvergence.
- $Q^k(e) = \left\{ p(x, y) = \sum_{i=0}^{k} \sum_{j=0}^{k} p_{ij} x^i y^j, (x, y) \in e \right\}$ is the set of tensor product of polynomials of degree $k$ on a cell $e$.
- $V^k = \{ p(x, y) \in C^0(\Omega_h) : p|_e \in Q^2(e), \; \forall e \in \Omega_h \}$ denotes the continuous piecewise $Q^2$ finite element space on $\Omega_h$.
- $V_0^k = \{ v_h \in V^k : v_h = 0 \; \text{on} \; \partial \Omega \}$.
- The norm and seminorms for $W^{k,p}(\Omega)$ and $1 \leq p < +\infty$, with standard modification for $p = +\infty$:

$$
\|u\|_{k,p,\Omega} = \left( \sum_{i+j=k} \int_{\Omega} \left| \partial_x^i \partial_y^j u(x,y) \right|^p dx dy \right)^{1/p},
$$

$$
|u|_{k,p,\Omega} = \left( \sum_{i+j=k} \int_{\Omega} \left| \partial_x^i \partial_y^j u(x,y) \right|^p dx dy \right)^{1/p},
$$

$$
[u]_{k,p,\Omega} = \left( \int_{\Omega} \left| \partial_x^k u(x,y) \right|^p dx dy + \int_{\Omega} \left| \partial_y^k u(x,y) \right|^p dx dy \right)^{1/p}.
$$

Notice that $[u]_{k+1,p,\Omega} = 0$ if $u$ is a $Q^k$ polynomial.

- For simplicity, sometimes we may use $\|u\|_{k,\Omega}$, $|u|_{k,\Omega}$ and $[u]_{k,\Omega}$ denote norm and seminorms for $H^k(\Omega) = W^{k,2}(\Omega)$.

- When there is no confusion, $\Omega$ may be dropped in the norm and seminorms, e.g., $\|u\|_k = \|u\|_{k,2,\Omega}$.

- For any $v_h \in V^k$, $1 \leq p < +\infty$ and $k \geq 1$,

$$
\|v_h\|_{k,p,\Omega} := \left( \sum_{e} \|v_h\|_{k,p,e}^p \right)^{1/p}, \quad [v_h]_{k,p,\Omega} := \left( \sum_{e} [v_h]_{k,p,e}^p \right)^{1/p}.
$$

- Let $Z_{0,e}$ denote the set of $3 \times 3$ Gauss-Lobatto points on a cell $e$.
- $Z_0 = \bigcup_{e} Z_{0,e}$ denotes all Gauss-Lobatto points in the mesh $\Omega_h$. 
• Let $\|u\|_{2,Z_0}$ and $\|u\|_{\infty,Z_0}$ denote the discrete 2-norm and the maximum norm over $Z_0$ respectively:

$$\|u\|_{2,Z_0} = \left( \sum_{(x,y) \in Z_0} |u(x,y)|^2 \right)^{\frac{1}{2}}, \quad \|u\|_{\infty,Z_0} = \max_{(x,y) \in Z_0} |u(x,y)|.$$ 

• For a continuous function $f(x,y)$, let $f_1(x,y)$ denote its piecewise $Q^2$ Lagrange interpolant at $Z_{0,e}$ on each cell $e$, i.e., $f_1 \in V^h$ satisfies:

$$f(x,y) = f_1(x,y), \quad \forall (x,y) \in Z_0.$$ 

• $P^k(t)$ denotes the polynomial of degree $k$ of variable $t$.

• $(f,v)_e$ denotes the inner product in $L^2(e)$ and $(f,v)$ denotes the inner product in $L^2(\Omega)$:

$$(f,v)_e = \int_e f v dx dy, \quad (f,v) = \int_\Omega f v dx dy = \sum_e (f,v)_e.$$ 

• $(f,v)_{e,h}$ denotes the approximation to $(f,v)_e$ by using $3 \times 3$-point Gauss Lobatto quadrature for integration over cell $e$.

• $(f,v)_h$ denotes the approximation to $(f,v)$ by using $(k+1) \times (k+1)$-point Gauss Lobatto quadrature for integration over each cell $e$.

• $\hat{K} = [-1,1] \times [-1,1]$ denotes a reference cell.

• For $f(x,y)$ defined on $e$, consider $\hat{f}(s,t) = f(sh + xe, th + ye)$ defined on $\hat{K}$. Let $\hat{f}$ denote the $Q^2$ Lagrange interpolation of $\hat{f}$ at the $3 \times 3$ Gauss Lobatto quadrature points on $\hat{K}$.

• $(\hat{f},\hat{v})_{\hat{K}} = \int_{\hat{K}} \hat{f} \hat{v} ds dt$.

• $(\hat{f},\hat{v})_{\hat{K}}$ denotes the approximation to $(\hat{f},\hat{v})_{\hat{K}}$ by using $3 \times 3$-point Gauss Lobatto quadrature.

On the reference cell $\hat{K}$, for convenience we use the superscript $h$ over the $ds$ or $dt$ to denote we use 3-point Gauss-Lobatto quadrature on the corresponding variable. For example,

$$\int_{\hat{K}} \hat{f} \hat{v} ds dt = \frac{1}{3} \int_{-1}^1 [\hat{f}(-1,t) + 4 \hat{f}(0,t) + \hat{f}(1,t)] dt.$$ 

Since $(\hat{f}\hat{v})_I$ coincides with $\hat{f}\hat{v}$ at the quadrature points, we have

$$\int_{\hat{K}} (\hat{f}\hat{v})_I dx dy = \int_{\hat{K}} (\hat{f}\hat{v})_I ds ds = \int_{\hat{K}} \hat{f} \hat{v} ds ds = (\hat{f},\hat{v})_{\hat{K}}.$$ 

The following are commonly used tools and facts:

• For two-dimensional problems,

$$h^{k-2/p} |v|_{k,p,e} = |\hat{v}|_{k,p,\hat{K}}, \quad h^{k-2/p} |v|_{k,p,e} = |\hat{v}|_{k,p,\hat{K}}, \quad 1 \leq p \leq \infty.$$
• Inverse estimates for polynomials:
  \[ \|v_h\|_{k+1,e} \leq C h^{-1} \|v_h\|_{k,e}, \quad \forall v_h \in V_h, k \geq 0. \quad (2.1) \]

• Sobolev’s embedding in two and three dimensions: \( H^2(\hat{\Omega}) \hookrightarrow C^0(\hat{\Omega}) \).

• The embedding implies
  \[ \|\hat{f}\|_{0,\infty,\hat{\Omega}} \leq C \|\hat{f}\|_{k,2,\hat{\Omega}}, \quad \forall \hat{f} \in H^k(\hat{\Omega}), k \geq 2, \]
  \[ \|\hat{f}\|_{1,\infty,\hat{\Omega}} \leq C \|\hat{f}\|_{k+1,2,\hat{\Omega}}, \quad \forall \hat{f} \in H^{k+1}(\hat{\Omega}), k \geq 2. \]

• Cauchy-Schwarz inequalities in two dimensions:
  \[ \sum_e \|\alpha\|_{k,e}\|v\|_{k,e} \leq \left( \sum_e \|\alpha\|_{k,e}^2 \right)^{1/2} \left( \sum_e \|v\|_{k,e}^2 \right)^{1/2}, \quad \|u\|_{k,1,e} = O(h) \|u\|_{k,2,e}. \]

• Poincaré inequality: let \( \bar{u} \) be the average of \( u \in H^1(\Omega) \) on \( \Omega \), then
  \[ |u - \bar{u}|_{0,p,\Omega} \leq C \|\nabla u\|_{0,p,\Omega}, \quad p \geq 1. \]

  If \( \bar{u} \) is the average of \( u \in H^1(e) \) on a cell \( e \), we have
  \[ |u - \bar{u}|_{0,p,e} \leq C h \|\nabla u\|_{0,p,e}, \quad p \geq 1. \]

• For \( k \geq 2 \), the \((k+1) \times (k+1)\) Gauss-Lobatto quadrature is exact for integration of polynomials of degree \(2k - 1 \geq k + 1\) on \( \hat{\Omega} \).

• Define the projection operator \( \Pi_h : \hat{\Omega} \to \hat{\Omega} \) by
  \[ \int_{\hat{\Omega}} (\Pi_h u) \, wdsdt = \int_{\hat{\Omega}} \bar{u} wdsdt, \forall w \in Q^1(\hat{\Omega}). \quad (2.2) \]

Notice that all degree of freedoms of \( \Pi_h \) can be represented as a linear combination of \( \int_{\hat{\Omega}} \hat{u}(s,t) p(s,t) dsdt \) for \( p(s,t) = 1, s,t \in \Omega \), thus \( \Pi_h \) is a continuous linear mapping from \( L^2(\hat{\Omega}) \) to \( H^1(\hat{\Omega}) \) (or \( H^2(\hat{\Omega}) \)) by Cauchy-Schwarz inequality.

2.2 Elliptic regularity and \( V^h \) ellipticity

We consider the elliptic variational problem of finding \( u \in H^1_0(\Omega) \) to satisfy

\[ A(u, v) := \int_{\Omega} (\nabla v^T a \nabla u + b \nabla v + cu v) \, dx \, dy = (f, v), \forall v \in H^1_0(\Omega), \quad (2.3) \]

where \( a = \begin{pmatrix} a^{11} & a^{12} \\ a^{21} & a^{22} \end{pmatrix} \) is positive definite and \( b = [b^1 \quad b^2] \). Assume the coefficients \( a \), \( b \) and \( c \) are smooth with uniform upper bounds, thus \( A(u, v) \leq C \|u\|_1 \|v\|_1 \) for any \( u, v \in H^1_0(\Omega) \). Assume the eigenvalues of \( a \) have a uniform positive lower bound and \( \nabla \cdot b \leq 2c \), so that coercivity of the bilinear form can be easily achieved.
Since
\[ (b \cdot \nabla u, v) = \int_{\partial \Omega} uv b \cdot n ds - (\nabla \cdot (v b), u) = \int_{\partial \Omega} uv b \cdot n ds - (b \cdot \nabla v, u) - (v \nabla \cdot b, u), \]
we have
\[ 2(b \cdot \nabla v, v) + 2(cv, v) = \int_{\partial \Omega} v^2 b \cdot n ds + ((2c - \nabla \cdot b)v, v) \geq 0, \quad \forall v \in H^1_0(\Omega). \] (2.4)

By the equivalence of two norms $| \cdot |$ and $\| \cdot \|_1$ for the space $H^1_0(\Omega)$ (see [Ciarlet 1991]), we conclude that the bilinear form $A(u, v) = (a \nabla u, \nabla v) + (b \cdot \nabla u, v) + (c u, v)$ satisfies coercivity $A(v, v) \geq C\|v\|_1$ for any $v \in H^1_0(\Omega)$.

We need to make two additional assumptions for the general elliptic operator (2.3):

1. The elliptic regularity holds for the dual problem. Let $A^*$ be the dual operator of $A$, i.e., $A^*(u, v) = A(v, u)$. We assume the elliptic regularity $\|w\|_{2} \leq C\|f\|_{0}$ holds for the exact dual problem of finding $w \in H^1_0(\Omega)$ satisfying $A^*(w, v) = (f, v)$, $\forall v \in H^1_0(\Omega)$. See [Savaré 1998] and [Grisvard 2011] for the elliptic regularity with Lipschitz continuous coefficients on a Lipschitz domain.

2. The bilinear form $A_h$ satisfies the $V_h$-ellipticity:
\[ \forall v_h \in V^h, \quad C\|v_h\|^2 \leq A_h(v_h, v_h). \] (2.5)

In Section [5.1] we will show that $V_h$-ellipticity holds if $h$ is small enough.

3. Quadrature error estimates

In the following, we will use $\hat{\cdot}$ for a function to emphasize the function is defined on or transformed to the reference cell $\hat{K} = [-1, 1] \times [-1, 1]$ from a mesh cell.

3.1 Standard estimates

The Bramble-Hilbert Lemma for $Q^k$ polynomials can be stated as follows, see Exercise 3.1.1 and Theorem 4.1.3 in [Ciarlet 2002]:

**Theorem 3.1** If a continuous linear mapping $\hat{\Pi} : H^{k+1}(\hat{K}) \to H^{k+1}(\hat{K})$ satisfies $\hat{\Pi} \hat{v} = \hat{v}$ for any $\hat{v} \in Q^k(\hat{K})$, then
\[ \| \hat{v} - \hat{\Pi} \hat{v} \|_{k+1, \hat{K}} \leq C\| \hat{v} \|_{k+1, \hat{K}}, \quad \forall \hat{v} \in H^{k+1}(\hat{K}). \] (3.1)

Thus if $l(\cdot)$ is a continuous linear form on the space $H^{k+1}(\hat{K})$ satisfying $l(\hat{v}) = 0, \forall \hat{v} \in Q^k(\hat{K})$, then
\[ |l(\hat{u})| \leq C\| l \|_{k+1, \hat{K}} \| \hat{u} \|_{k+1, \hat{K}}, \quad \forall \hat{u} \in H^{k+1}(\hat{K}), \]
where $\| l \|_{k+1, \hat{K}}$ is the norm in the dual space of $H^{k+1}(\hat{K})$.

By applying Bramble-Hilbert Lemma, we have the following standard quadrature estimates. See [Li & Zhang 2019b] for the detailed proof.

**Theorem 3.2** For a sufficiently smooth function $a(x, y)$ and its $Q^2$ interpolation $a_I$, let $m$ is an integer satisfying $2 \leq m \leq 4$, we have
\[ \int_{\hat{K}} a(x, y) dx dy - \int_{\hat{K}} a_I(x, y) dx dy = O(h^{m+1})[a]_{m,e} = O(h^{m+2})[a]_{m,m,e}. \]
THEOREM 3.3 If \( f \in H^4(\Omega) \), \( (f, v_h) - \langle f, v_h \rangle_h = O(h^4)\|f\|_4\|v_h\|_2 \), \( \forall v_h \in V_h \).

REMARK 3.1 By the theorems above, on the reference cell \( \hat{K} \), we have
\[
\iint_{\hat{K}} \hat{a}(s, t) - \hat{a}(s, t)\, ds\, dt \leq C|\hat{a}|_{4, \hat{K}} \leq C|\hat{a}|_{4, \infty, \hat{K}},
\]
and
\[
|\hat{a} - \hat{a}|_{3, \hat{K}} \leq C|\hat{a}|_{3, \hat{K}}.
\]

The following two results are also standard estimates obtained by applying the Bramble-Hilbert Lemma.

LEMMA 3.1 If \( f \in H^2(\Omega) \) or \( f \in V_h \), we have \( (f, v_h) - \langle f, v_h \rangle_h = O(h^2)\|f\|_2\|v_h\|_0 \), \( \forall v_h \in V_h \).

Proof. For simplicity, we ignore the subscript in \( v_h \). Let \( E(f) \) denote the quadrature error for integrating \( f(x, y) \) on \( e \). Let \( \hat{E}(\hat{f}) \) denote the quadrature error for integrating \( \hat{f}(s, t) = f(x_c + sh, y_c + th) \) on the reference cell \( \hat{K} \). Due to the embedding \( H^2(\hat{K}) \hookrightarrow C^0(\hat{K}) \), we have
\[
|\hat{E}(\hat{f})| \leq C\|\hat{f}\|_{0, \hat{K}} \leq C\|\hat{f}\|_{0, \infty, \hat{K}} \leq C\|\hat{f}\|_{2, \hat{K}}\|\hat{v}\|_{0, \hat{K}}.
\]
Thus the mapping \( \hat{f} \rightarrow E(\hat{f}) \) is a continuous linear form on \( H^2(\hat{K}) \) and its norm is bounded by \( C\|\hat{v}\|_{0, \hat{K}} \).

If \( \hat{f} \in \hat{Q}^1(\hat{K}) \), then we have \( \hat{E}(\hat{f}) = 0 \). By the Bramble-Hilbert Lemma Theorem 3.1 on this continuous linear form, we get
\[
|\hat{E}(\hat{f})| \leq C\|\hat{f}\|_{2, \hat{K}}\|\hat{v}\|_{0, \hat{K}}.
\]
So on a cell \( e \), we get
\[
E(fv) = h^2\hat{E}(\hat{f}\hat{v}) \leq Ch^2\|\hat{f}\|_{2, \hat{K}}\|\hat{v}\|_{0, \hat{K}} \leq Ch^2\|f\|_{2, e}\|v\|_{0, e}.
\]

Summing over all elements and use Cauchy-Schwarz inequality, we get the desired result. \( \square \)

THEOREM 3.4 Assume all coefficients of \( (2.3) \) are in \( W^{2, \infty}(\Omega) \). We have
\[
A(z_h, v_h) - A_h(z_h, v_h) = O(h)|v_h|_2|z_h|_1, \quad \forall v_h, z_h \in V_h.
\]

Proof. By setting \( f = a^{11}(v_h)_x \) in (3.4), we get
\[
|\langle a^{11}(z_h)_x, (v_h)_x \rangle - \langle a^{11}(z_h)_x, (v_h)_x \rangle_h| = Ch^2\|a^{11}(v_h)_x\|_2(|z_h|_1)\|v_h|_0\|z_h|_1 \leq Ch\|a^{11}\|_{2, \infty}\|v_h\|_2|z_h|_1,
\]
where the inverse estimate (2.1) is used in the last inequality. Similarly, we have
\[
\langle a^{12}(z_h)_x, (v_h)_y \rangle - \langle a^{12}(z_h)_x, (v_h)_y \rangle_h = Ch\|a^{12}\|_{2, \infty}\|v_h\|_2|z_h|_1,
\]
\[
(a^{22}(z_h)_y, (v_h)_y) - \langle a^{22}(z_h)_y, (v_h)_y \rangle_h = Ch\|a^{22}\|_{2, \infty}\|v_h\|_2|z_h|_1,
\]
\[
(b^1(z_h)_x, v_h) - (b^1(z_h)_x, v_h)_h = Ch\|b^1\|_{2, \infty}\|v_h\|_2|z_h|_0,
\]
\[
(b^2(z_h)_y, v_h) - (b^2(z_h)_y, v_h)_h = Ch\|b^2\|_{2, \infty}\|v_h\|_2|z_h|_0,
\]
\[
(cz_h, v_h) - (cz_h, v_h)_h = Ch\|c\|_{2, \infty}\|v_h\|_2|z_h|_0,
\]
which implies
\[
A(z_h, v_h) - A_h(z_h, v_h) = O(h)|v_h|_2|z_h|_1.
\]
\( \square \)
3.2 Explicit quadrature error terms

Define \( p(t) = \frac{1}{24} t^4 - \frac{1}{9} t^3 + \frac{1}{12} t^2 - \frac{1}{12} \) and let \( \tilde{p}(t) \) denote its even extension:

\[
\tilde{p}(t) = \begin{cases} 
\frac{1}{24} t^4 - \frac{1}{9} t^3 + \frac{1}{12} t^2 - \frac{1}{12}, & t \geq 0, \\
\frac{1}{24} t^4 + \frac{1}{9} t^3 + \frac{1}{12} t^2 - \frac{1}{12}, & t < 0.
\end{cases}
\]

**Lemma 3.2** If \( \hat{g} \in W^{4,1}([0, 1]) \) with \( \hat{g}''(0) = \hat{g}^{(3)}(0) = 0 \), then

\[
\int_{0}^{1} \hat{g}(t)dt = \frac{2}{3} \hat{g}(1) + \frac{2}{3} \hat{g}(0) + \int_{0}^{1} p(t)\hat{g}^{(4)}(t)dt. \tag{3.5}
\]

**Proof.** First we assume that \( \hat{g} \in C^4([0, 1]) \), then it can be shown through integration by parts. It is straightforward to check

\[
\int_{0}^{1} \hat{g}(t)dt = \frac{2}{3} \hat{g}(1) + \frac{2}{3} \hat{g}(0) - \int_{0}^{1} (t - \frac{2}{3})\hat{g}'(t)dt.
\]

And we have

\[
- \int_{0}^{1} (t - \frac{2}{3})\hat{g}'(t)dt = - \int_{0}^{1} \hat{g}'(t)d(t^2 - \frac{2t}{3} + \frac{1}{6})

= - \left[ \frac{t^2}{2} - \frac{2t}{3} + \frac{1}{6} \right] \hat{g}'(t) \bigg|_{0}^{1} + \int_{0}^{1} \left( \frac{t^2}{2} - \frac{2t}{3} + \frac{1}{6} \right) \hat{g}''(t)dt

= \int_{0}^{1} \hat{g}''(t)d\left( \frac{t^3}{6} - \frac{t^2}{3} + \frac{t}{6} \right)

= \left[ \frac{t^3}{6} - \frac{t^2}{3} + \frac{t}{6} \right] \hat{g}''(t) \bigg|_{0}^{1} - \int_{0}^{1} \left( \frac{t^3}{6} - \frac{t^2}{3} + \frac{t}{6} \right) \hat{g}^{(3)}(t)dt

= - \int_{0}^{1} \hat{g}^{(3)}(t)d\left( \frac{1}{24} t^4 - \frac{1}{9} t^3 + \frac{1}{12} t^2 - \frac{1}{12} \right)

= \left[ \frac{1}{24} t^4 - \frac{1}{9} t^3 + \frac{1}{12} t^2 - \frac{1}{12} \right] \hat{g}^{(3)}(t) \bigg|_{0}^{1} + \int_{0}^{1} p(t)\hat{g}^{(4)}(t)dt

= \int_{0}^{1} p(t)\hat{g}^{(4)}(t)dt.
\]

By standard global approximation to \( \hat{g} \) by smooth functions we know the result holds for \( \hat{g} \in W^{4,1}([0, 1]). \)

By Lemma 3.2 it is straightforward to show the following result.

**Lemma 3.3** Suppose \( \hat{f} \in W^{4,1}([-1, 1]) \), then

\[
\int_{-1}^{1} \hat{f}(t)dt = \frac{2}{3} [\hat{f}(-1) + 4\hat{f}(0) + \hat{f}(1)] + \int_{-1}^{1} \tilde{p}(t)\hat{f}^{(4)}dt \tag{3.6}

= \int_{-1}^{1} \hat{f}(t)dt + \int_{-1}^{1} \tilde{p}(t)\hat{f}^{(4)}dt. \tag{3.7}
\]
Proof. Let $\tilde{g}(t) = f(t) + \hat{f}(t)$ then $\tilde{g}'(0) = \hat{g}'(0) = 0$. Apply Lemma 3.2 to $\tilde{g}(t)$, we have
\[
\int_{-1}^{1} \tilde{g}'(t) dt = \int_{0}^{1} \hat{f}(t) dt + \int_{0}^{1} \hat{f}(t) dt = \int_{0}^{1} \hat{g}(t) dt = \frac{1}{3} \hat{g}(1) + \frac{2}{3} \hat{g}(0) + \int_{0}^{1} p(t) \hat{g}^{(4)}(t) dt.
\]
\[
\int_{0}^{1} p(t) \hat{g}^{(4)}(t) dt = \int_{0}^{1} p(t) [\hat{f}^{(4)}(t) + \hat{f}^{(4)}(-t)] dt = \int_{-1}^{1} \hat{p}(t) \hat{f}^{(4)}(t) dt.
\]
\[\square\]

Remark 3.2 For a function $u(x) \in W^{4,1}([x_e - h, x_e + h])$, if we map it from cell $e = [x - x_e, x + x_e]$ to the reference cell $\hat{K} = [-1, 1]$ then apply Lemma 3.3 and map it back, we get the error estimation of 3-point Gauss-Lobatto quadrature:
\[
\int_{x_e - h}^{x_e + h} u(x) dx = \frac{h}{3} [u(x_e - h) + 4u(x_e) + u(x_e + h)] + h^4 \int_{x_e - h}^{x_e + h} \hat{p}(x) u^{(4)}(x) dx. \tag{3.8}
\]

3.3 A refined consistency error

In this subsection, we will show how to establish the desired consistency error estimate for smooth enough coefficients:
\[
A(u, v_h) - A_h(u, v_h) = \begin{cases}
O(h^4) \| u \|_5 \| v_h \|_2, & \forall v_h \in V_0^h, \\
O((h^3)^4) \| u \|_5 \| v_h \|_2, & \forall v_h \in V^h.
\end{cases}
\]

Theorem 3.5 Assume $a(x, y) \in W^{4,\infty}(\Omega)$, $u \in H^5(\Omega)$, then
\[
(a\partial_x u, \partial_y v_h) - (a\partial_x u, \partial_y v_h)_h = \begin{cases}
O(h^4) \| a \|_{4,\infty} \| u \|_5 \| v_h \|_2, & \forall v_h \in V_0^h, \\
O((h^3)^4) \| a \|_{4,\infty} \| u \|_5 \| v_h \|_2, & \forall v_h \in V^h,
\end{cases} \tag{3.9a}
\]
\[
(a\partial_x u, \partial_y v_h) - (a\partial_x u, \partial_y v_h)_h = \begin{cases}
O(h^4) \| a \|_{4,\infty} \| u \|_5 \| v_h \|_2, & \forall v_h \in V_0^h, \\
O((h^3)^4) \| a \|_{4,\infty} \| u \|_5 \| v_h \|_2, & \forall v_h \in V^h,
\end{cases} \tag{3.10a}
\]
\[
(a\partial_x u, v_h) - (a\partial_x u, v_h)_h = O(h^4) \| a \|_{4,\infty} \| u \|_4 \| v_h \|_2, & \forall v_h \in V_0^h, \\
O((h^3)^4) \| a \|_{4,\infty} \| u \|_4 \| v_h \|_2, & \forall v_h \in V^h, \tag{3.11}
\]
\[
(a, v_h) - (a, v_h)_h = O(h^4) \| a \|_{4,\infty} \| u \|_4 \| v_h \|_2, & \forall v_h \in V_0^h. \tag{3.12}
\]

Remark 3.3 We emphasize that Theorem 3.5 cannot be proven by applying the Bramble-Hilbert Lemma. Consider the constant coefficient case $a(x, y) \equiv 1$ as an example,
\[
(a, v_h) - (a, v_h)_h = \sum_e \left( \int_e \int u_x(v_h)_x dx dy - \int_e \int u_x(v_h)_x d^h x d^h y \right).
\]

Since the $3 \times 3$ Gauss-Lobatto quadrature is exact for integrating $O^3$ polynomials, by Theorem 3.1 we have
\[
\left| \int_e \int u_x(v_h)_x dx dy - \int_e \int u_x(v_h)_x d^h x d^h y \right| = \left| \int_K \int u_x(v_h)_x ds dt - \int_K \int u_x(v_h)_x d^h s d^h t \right| \leq C \| \bar{u}_x(v_h) \|_{4,\hat{K}}.
\]
Notice that $\hat{v}_h$ is $Q^2$ thus $(\hat{v}_h)_{tt}$ does not vanish and $[(\hat{v}_h)_t]_{4,R} \leq C[\hat{v}_h]_{3,R}$. So by Bramble-Hilbert Lemma for $Q^k$ polynomials, we can only get

$$\int_{\Omega} u_s(v_h)_s dx dy - \int_{\Omega} u_s(v_h)_s d^h y = O(h^4)\|u\|_{5,\varepsilon}\|v_h\|_{3,\varepsilon}. $$

Thus by Cauchy-Schwarz inequality after summing over $e$, we only have

$$(\partial_x u, \partial_x v_h) - (\partial_x u, \partial_x v_h)_h = O(h^4)\|u\|_s\|v_h\|_3.$$  

In order to get the desired estimate involving only the $H^2$-norm of $v_h$, we propose to derive the explicit error term of the Gauss-Lobatto quadrature.

Proof. For simplicity, we ignore the subscript $h$ of $v_h$ in this proof and all the following $\hat{v}$ are in $V_h$ which are $Q^2$ polynomials in each cell. First, by Theorem\[3.3] we easily obtain \[3.11\] and \[3.12\]:

\[ (au_s, v) - (au_s, v)_h = O(h^4)\|au_s\|_4\|v\|_2 = O(h^4)\|a\|_{4,\infty}\|u\|_5\|v\|_2, \]

\[ (au, v) - (au, v)_h = O(h^4)\|au\|_4\|v\|_2 = O(h^4)\|a\|_{4,\infty}\|u\|_4\|v\|_2. \]

We will only discuss $(au_s, v_s) - (au_s, v_s)_h$ and the same discussion also applies to derive \[3.10a\] and \[3.10b\].

Since we have

$$\int_{\Omega} u_s(v_h)_s dx dy - \int_{\Omega} u_s(v_h)_s d^h y = \sum_e \left( \int_{\Omega} au_s v_s dx dy - \int_{\Omega} au_s v_s d^h y \right)$$

$$= \sum_e \left( \int_{\hat{K}} \hat{a} u_s \hat{v}_s dsdt - \int_{\hat{K}} \hat{a} u_s \hat{v}_s d^h t \right) = \sum_e \left( \int_{\hat{K}} \hat{a} u_s \hat{v}_s dsdt - \int_{\hat{K}} \hat{a} (\hat{a}_s)_t \hat{v}_s d^h t \right).$$

For fixed $t$, $(\hat{a}_s)_t \hat{v}_s$ is a polynomial of degree 3 w.r.t. variable $s$. Thus on $\hat{K}$ the 3-point Gauss-Lobatto quadrature is exact for $s$-integration:

$$\int_{\hat{K}} \hat{a} (\hat{a}_s)_t \hat{v}_s d^h t = \int_{\hat{K}} \hat{a} (\hat{a}_s)_t \hat{v}_s d^h t.$$  

We apply Lemma\[3.3] to \( \int_{\hat{K}} \hat{a} (\hat{a}_s)_t \hat{v}_s dsdt \) on $t$-integration:

$$\int_{\hat{K}} \hat{a} u_s \hat{v}_s dsdt - \int_{\hat{K}} \hat{a} u_s \hat{v}_s d^h t = \int_{\hat{K}} \hat{a} u_s \hat{v}_s dsdt - \int_{\hat{K}} \hat{a} (\hat{a}_s)_t \hat{v}_s d^h t$$

$$= \int_{\hat{K}} \hat{a} u_s \hat{v}_s dsdt - \int_{\hat{K}} \hat{a} (\hat{a}_s)_t \hat{v}_s dsdt + \int_{\hat{K}} \hat{p}(t) \hat{a} (\hat{a}_s)_t \hat{v}_s dsdt. \quad (3.13)$$

Let $\overline{v}_e$ be the cell average of $\hat{v}_s$ on $\hat{K}$, then for the first two terms in \(3.13\) we have

$$\int_{\hat{K}} \hat{a} u_s \hat{v}_s dsdt - \int_{\hat{K}} \hat{a} (\hat{a}_s)_t \hat{v}_s dsdt = \int_{\hat{K}} (\hat{a} u_s - (\hat{a} u_s)_t) \overline{v}_s dsdt + \int_{\hat{K}} (\hat{a} u_s - (\hat{a} u_s)_t) \overline{v}_s dsdt.$$

By \[3.2\], we have

$$\left| \int_{\hat{K}} (\hat{a} u_s - (\hat{a} u_s)_t) \overline{v}_s dsdt \right| \leq C[\hat{a} u_s]_{4,R} \overline{v}_s = O(h^4)\|\hat{a}\|_{4,\infty,\varepsilon}\|\hat{u}\|_{5,\varepsilon}\|\overline{v}_s\|_{1,\varepsilon}. $$

By Cauchy-Schwarz inequality, Bramble-Hilbert Lemma on interpolation error and Poincaré inequality, we have
\[
\left| \int_K (\hat{a} \hat{u}_s) (\hat{v}_s - \bar{v}_s) ds dt \right| \leq |\hat{a} \hat{u}_s - (\hat{a} \hat{u}_s)_I|_{0, \hat{K}} |\hat{v}_s - \bar{v}_s|_{0, \hat{K}} \leq C[\hat{a} \hat{u}_s]_{1, \hat{K}} |\hat{v}|_{2, \hat{K}} = O(h^4) \|a\|_{\infty, e} \|u\|_{4, e} \|v\|_{2, e}.
\]
Thus we have
\[
\int_K \hat{a} \hat{u}_s \hat{v}_s ds dt - \int_K (\hat{a} \hat{u}_s)_I \hat{v}_s ds dt = O(h^4) \|a\|_{\infty, e} \|u\|_{4, e} \|v\|_{2, e}.
\]

For the last term in (3.13), notice that \(v\) is a \(Q^2\) polynomial on \(e\) thus some of its high order derivatives vanish. With product rule and integration by parts on \(s\), we get
\[
\int_K \hat{p}(t) \hat{a} \hat{u}_s \hat{v}_s ds dt = 6 \int_K \hat{p}(t) \hat{a} \hat{u}_s \hat{v}_s ds dt = 6 \int_{-1}^1 \hat{p}(t) \hat{a} \hat{u}_s \hat{v}_s dt \bigg|_{t=-1} - \int \hat{p}(t) \hat{a} \hat{u}_s \hat{v}_s ds dt.
\]
By Bramble-Hilbert Lemma (Theorem 3.1), we have
\[
\int_K \hat{p}(t) \hat{a} \hat{u}_s \hat{v}_s ds dt = \int_K \hat{p}(t) \hat{a} \hat{u}_s \hat{v}_s ds dt + \int_K \hat{p}(t) \hat{a} \hat{u}_s \hat{v}_s ds dt \\
\leq C[\hat{a} \hat{u}_s]_{1, \hat{K}} |\hat{v}|_{2, \hat{K}} + C[\hat{a} \hat{u}_s]_{1, \hat{K}} |\hat{v}|_{2, \hat{K}} \\
\leq C[\hat{a} \hat{u}_s]_{1, \hat{K}} |\hat{v}|_{2, \hat{K}} + C[\hat{a} \hat{u}_s]_{1, \hat{K}} |\hat{v}|_{2, \hat{K}} = O(h^4) \|a\|_{\infty, e} \|u\|_{4, e} \|v\|_{2, e}.
\]
Now we only need to discuss the line integral term.
\[
\int_{-1}^1 \hat{p}(t) \hat{a} \hat{u}_s \hat{v}_s dt \bigg|_{t=1} - \int_{-1}^1 \hat{p}(t) \hat{a} \hat{u}_s \hat{v}_s dt \bigg|_{t=-1} = \int_{-1}^1 \hat{p}(t) \hat{a} \hat{u}_s \hat{v}_s dt \bigg|_{t=1} + \int_{-1}^1 \hat{p}(t) \hat{a} \hat{u}_s \hat{v}_s dt \bigg|_{t=-1}.
\]
Notice that \(\hat{v}_s^2(s, t)\) is a quartic polynomial thus its integral over \(\hat{K}\) is equal to using 4-point Gauss-Lobatto quadrature for the \(s\)-variable. Therefore, by considering 4-point Gauss-Lobatto quadrature for the \(s\)-variable in the integral \(\int_K \hat{v}_s^2 ds dt\), we can obtain
\[
\int_{-1}^1 \hat{v}_s^2(\pm 1, t) dt \leq C \int_{-1}^1 \hat{v}_s^2(s, t) ds dt. \tag{3.14}
\]
Thus by Cauchy-Schwarz inequality, trace inequality and Theorem 3.1, we have
\[
\int_{-1}^1 \hat{p}(t) \hat{a} \hat{u}_s \hat{v}_s dt \bigg|_{t=1} = C[\hat{v}_s]_{0, \hat{K}} |\hat{a} \hat{u}_s|_{1, \hat{K}} |\hat{v}_s|_{0, \hat{K}} \\
\leq C[\hat{v}_s]_{2, \hat{K}} |\hat{a} \hat{u}_s|_{1, \hat{K}} |\hat{v}_s|_{0, \hat{K}} = C[\hat{v}_s]_{2, \hat{K}} |\hat{a} \hat{u}_s|_{1, \hat{K}} |\hat{v}_s|_{0, \hat{K}} \\
\leq C[\hat{v}_s]_{2, \hat{K}} |\hat{a} \hat{u}_s|_{1, \hat{K}} = O(h^4) \|a\|_{\infty, e} \|u\|_{4, e} \|v\|_{2, e}.
\]
After mapping back to the cell \(e\), we have
\[
\int_{-1}^1 \hat{p}(t) \hat{a} \hat{u}_s \hat{v}_s dt \bigg|_{t=1} = h^4 \int_{y_e-h}^{y_e+h} \hat{p}(\frac{y - y_e}{h}) \left[ \hat{a} \hat{u}_s \hat{v}_s \right] dy \bigg|_{y_e-h}^{y_e+h}.
\]
Let $L_2$ and $L_4$ denote the left and right boundary of $\Omega$ and let $l_2$ and $l_4$ denote the left and right edge of element $e$. Since $\partial_L^2 (au_e)$ and $v_{yy}$ are continuous across $L_2$ and $L_4$, after summing over all elements $e$, the line integrals along the inner edges are cancelled out and only the line integrals on $L_2$ and $L_4$ remain.

$$
\sum_e h^4 \int_{y_e-h}^{y_e+h} \bar{p}(\frac{Y-y_e}{h}) \partial^2_y (au_e) v_{yy} dy \Bigg|_{y=x-h}^{x=x+h} = \sum_{e' \in L_2 \neq \emptyset} h^4 \int_{y_e-h}^{y_e+h} \bar{p}(\frac{Y-y_e}{h}) \partial^2_y (au_e) v_{yy} (1,y) dy - \sum_{e' \in L_2 \neq \emptyset} h^4 \int_{y_e-h}^{y_e+h} \bar{p}(\frac{Y-y_e}{h}) \partial^2_y (au_e) v_{yy} (0,y) dy.
$$

Since $(v_{yy})^2$ is a quartic polynomial, by considering 4-point Gauss-Lobatto quadrature for $x$-integration in $\int_L v_{yy}^2 dx dy$, we get

$$
h \int_{y_e-h}^{y_e+h} (v_{yy})^2 (x_e - h,y) dy \leq C \int_e (v_{yy})^2 (x,y) dx dy.
$$

For the line integrals along $L_2$, we have

$$
\sum_{e' \in L_2 \neq \emptyset} h^4 \int_{y_e-h}^{y_e+h} \bar{p}(\frac{Y-y_e}{h}) (au_e)_{yy} v_{yy} (0,y) dy
$$

$$= O(h^4) \sum_{e' \in L_2 \neq \emptyset} \sqrt{\int_{y_e-h}^{y_e+h} ((au_e)_{yy})^2 (0,y) dy} \sqrt{\int_{y_e-h}^{y_e+h} v_{yy}^2 (0,y) dy}
$$

$$= O(h^{3.5}) ||a||_{2,\infty} \sum_{e' \in L_2 \neq \emptyset} ||u||_{3,2,\ell_4} \sqrt{h \int_{y_e-h}^{y_e+h} v_{yy}^2 (0,y) dy}
$$

$$= O(h^{3.5}) ||a||_{2,\infty} \sum_{e' \in L_2 \neq \emptyset} ||u||_{3,2,\ell_4} ||v||_{2,e}
$$

$$= O(h^{3.5}) ||a||_{2,\infty} ||u||_{3,2,\ell_4} ||v||_{2,\Omega} = O(h^{3.5}) ||a||_{2,\infty} ||u||_4 ||v||_2.
$$

where trace inequality $||u||_{3,\partial \Omega} \leq C ||u||_{4,\Omega}$ is applied.

With the same argument we have

$$
h^4 \sum_{e' \in L_4 \neq \emptyset} \int_{y_e-h}^{y_e+h} \bar{p}(\frac{Y-y_e}{h}) (au_e)_{yy} v_{yy} (1,y) dy = O(h^{3.5}) ||a||_{2,\infty} ||u||_4 ||v||_2.
$$

Combine all the estimates above, we get (3.9b). Since the $\frac{1}{4}$ order loss is only due to the line integral along the boundary $\partial \Omega$. If $v \in V_0^h$, $v_{yy} = 0$ on $L_2$ and $L_4$ so we have (3.9a).

4. Superconvergence of bilinear forms

The M-type projection in Chen (1981, 2001) is a very convenient tool for discussing the superconvergence of function values. Let $u_p$ be the M-type $Q^2$ projection of the smooth exact solution $u$ and its definition will be given in the following subsection. To establish the superconvergence of the original finite element method (1.1) for a generic elliptic problem (2.3) with smooth coefficients, one can show the
following superconvergence of bilinear forms, see Chen (2001); Lin & Yan (1996) (see also Li & Zhang (2019b) for a detailed proof):

\[
A(u - u_p, v_h) = \begin{cases} \mathcal{O}(h^{3.5})||u||_S||v_h||_2, & \forall v_h \in V^h, \\ \mathcal{O}(h^4)||u||_S||v_h||_2, & \forall v_h \in V_0^h. \end{cases}
\]

In this section we will show the superconvergence of the bilinear form \( A_h \):

\[
A_h(u - u_p, v_h) = \begin{cases} \mathcal{O}(h^{3.5})||u||_S||v_h||_2, & \forall v_h \in V^h, \\ \mathcal{O}(h^4)||u||_S||v_h||_2, & \forall v_h \in V_0^h. \end{cases}
\]

### 4.1 Definition of M-type projection

We first recall the definition of M-type projection. More detailed definition can also be found in Li & Zhang (2019b). Legendre polynomials on the reference interval \([-1, 1]\) are given as

\[
l_k(t) = \frac{1}{2^k k!} \frac{d^k}{dt^k} (t^2 - 1)^k : l_0(t) = 1, l_1(t) = t, l_2(t) = \frac{1}{2} (3t^2 - 1), \ldots,
\]

which are \(L^2\)-orthogonal to one another. Define their antiderivatives as M-type polynomials:

\[
M_{k+1}(t) = \frac{1}{2^k k!} \frac{d^{k+1}}{dt^{k+1}} (t^2 - 1)^k : M_0(t) = 1, M_1(t) = t, M_2(t) = \frac{1}{2} (4t^2 - 1), M_3(t) = \frac{1}{2} (6t^3 - t), \ldots
\]

Since Legendre polynomials form a complete orthogonal basis for \(L^2([-1, 1])\), for any \(\hat{f}(t) \in H^1([-1, 1])\), its derivative \(\hat{f}'(t)\) can be expressed as Fourier-Legendre series

\[
\hat{f}'(t) = \sum_{j=0}^{\infty} \hat{b}_j l_j(t), \quad \hat{b}_{j+1} = (j + \frac{1}{2}) \int_{-1}^{1} \hat{f}'(t) l_j(t) dt.
\]

The one-dimensional M-type projection is defined as

\[
\hat{f}_K(t) = \sum_{j=0}^{k} \hat{b}_j M_j(t),
\]

where \(\hat{b}_0 = \frac{\hat{f}(1) + \hat{f}(-1)}{2}\) is determined by \(\hat{b}_1 = \frac{\hat{f}(1) - \hat{f}(-1)}{2}\) so that \(\hat{f}_K(\pm 1) = \hat{f}(\pm 1)\). We have \(\hat{f}(t) = \lim_{k \to \infty} \hat{f}_k(t) = \sum_{j=0}^{\infty} \hat{b}_j M_j(t)\). The remainder \(\mathcal{R}[\hat{f}]_k(t)\) of one-dimensional M-type projection is

\[
\mathcal{R}[\hat{f}]_k(t) = \hat{f}(t) - \hat{f}_k(t) = \sum_{j=k+1}^{\infty} \hat{b}_j M_j(t).
\]

For a function \(\hat{f}(s, t) \in H^2(\hat{K})\) on the reference cell \(\hat{K} = [-1, 1] \times [-1, 1]\), its two-dimensional M-type expansion is given as

\[
\hat{f}(s, t) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \hat{b}_{i,j} M_i(s) M_j(t),
\]

where \(\hat{b}_{i,j} \) is determined by

\[
\hat{b}_{i,j} \left( \int_{-1}^{1} \int_{-1}^{1} \hat{f}(s, t) M_i(s) M_j(t) ds dt \right) = \int_{-1}^{1} \int_{-1}^{1} \hat{f}(s, t) M_i(s) M_j(t) ds dt.
\]
where

\[ \hat{b}_{0,0} = \frac{1}{4} [\hat{f}(-1, -1) + \hat{f}(-1, 1) + \hat{f}(1, -1) + \hat{f}(1, 1)], \]

\[ \hat{b}_{0,j}, \hat{b}_{1,j} = \frac{2j - 1}{4} \int_{-1}^{1} [\hat{f}(1, t) \pm \hat{f}(-1, t)]l_{j-1}(t)dt, \quad j \geq 1, \]

\[ \hat{b}_{i,0}, \hat{b}_{i,1} = \frac{2i - 1}{4} \int_{-1}^{1} [\hat{f}(s, 1) \pm \hat{f}(s, -1)]l_{i-1}(s)ds, \quad i \geq 1, \]

\[ \hat{b}_{i,j} = \frac{(2i - 1)(2j - 1)}{4} \int_{-1}^{1} \hat{f}_x(s,t)l_{i-1}(s)l_{j-1}(t)dsdt, \quad i, j \geq 1. \]

The M-type \( Q^2 \) projection of \( \hat{f} \) on \( \hat{K} \) and its remainder are defined as

\[ \hat{f}_{2,2}(s,t) = \sum_{i=0}^{2} \sum_{j=0}^{2} \hat{b}_{i,j} M_i(s) M_j(t), \quad \mathcal{R}[\hat{f}]_{2,2}(s,t) = \hat{f}(s,t) - \hat{f}_{2,2}(s,t). \]

The M-type \( Q^k \) projection is equivalent to the point-line-plane interpolation used in Lin et al. (1991); Lin & Yan (1996). See Li & Zhang (2019b) for the proof of the following fact:

**THEOREM 4.1** The M-type \( Q^k \) projection is equivalent to the \( Q^k \) point-line-plane projection \( \Pi \) defined as follows:

1. \( \Pi \hat{u} = \hat{u} \) at four corners of \( \hat{K} = [-1, 1] \times [-1, 1] \).
2. \( \Pi \hat{u} - \hat{u} \) is orthogonal to polynomials of degree \( k - 2 \) on each edge of \( \hat{K} \).
3. \( \Pi \hat{u} - \hat{u} \) is orthogonal to any \( \hat{v} \in Q^{k-2}(\hat{K}) \) on \( \hat{K} \).

For \( f(x,y) \) on \( e = [x_r - h, x_r + h] \times [y_r - h, y_r + h] \), let \( \hat{f}(s,t) = f(sh + x_r, th + y_r) \) then the M-type \( Q^k \) projection of \( f \) on \( e \) and its remainder are defined as

\[ \hat{f}_{k,k}(x,y) = \hat{f}_{k,k}(\frac{x - x_r}{h}, \frac{y - y_r}{h}), \quad \mathcal{R}[\hat{f}]_{k,k}(x,y) = f(x,y) - \hat{f}_{k,k}(x,y). \]

Now consider a function \( u(x,y) \in H^{k+2}(\Omega) \), let \( u_p(x,y) \) denote its piecewise M-type \( Q^k \) projection on each element \( e \) in the mesh \( \Omega_h \). The first two properties in Theorem 4.1 imply that \( u_p(x,y) \) on each edge of \( e \) is uniquely determined by \( u(x,y) \) along that edge. So \( u_p(x,y) \) is a piecewise continuous \( Q^k \) polynomial on \( \Omega_h \).

M-type projection has the following properties. See Li & Zhang (2019b) for the proof.

**THEOREM 4.2**

\[ \| u - u_p \|_{2,Z_0} = O(h^{k+2})\| u \|_{k+2}, \quad \forall u \in H^{k+2}(\Omega). \]

\[ \| u - u_p \|_{\infty,Z_0} = O(h^{k+2})\| u \|_{k+2,\infty}, \quad \forall u \in W^{k+2,\infty}(\Omega). \]

**LEMMA 4.1** For \( \hat{f} \in H^{k+1}(\hat{K}), k \geq 2 \):

1. \( |\hat{R}[\hat{f}]_{k,k}|_{0,\infty,\hat{K}} \leq C|\hat{f}|_{k+1,\hat{K}}, \quad |\partial_n \hat{R}[\hat{f}]_{k,k}|_{0,\infty,\hat{K}} \leq C|\hat{f}|_{k+1,\hat{K}}. \)
2. \( \hat{R}[\hat{f}]_{k+1,k+1} - \hat{R}[\hat{f}]_{k,k} = M_{k+1}(1) \sum_{j=0}^{k} \hat{b}_{j,k+1} M_j(s) + M_{k+1}(s) \sum_{j=0}^{k} \hat{b}_{k+1,j} M_j(t). \)
3. \( |\hat{b}_{k+1}| \leq C_k|\hat{f}|_{k+1,2,\hat{K}}, |\hat{b}_{k+1,i}| \leq C_k|\hat{f}|_{k+1,2,\hat{K}}, \quad 0 \leq i \leq k + 1. \)
4. If \( \hat{f} \in H^{k+2}(\hat{K}) \), then \( |\hat{b}_{k+1}| \leq C_k|\hat{f}|_{k+2,2,\hat{K}}, \quad 1 \leq i \leq k + 1. \)
4.2 Estimates of M-type projection with quadrature

**Lemma 4.2** Assume \( \hat{f}(s,t) \in H^3(\hat{K}) \),

\[
\langle \hat{R}[\hat{f}]_{3,3} - \hat{R}[\hat{f}]_{2,2}, 1 \rangle_{\hat{K}} = 0,
\quad |\langle \partial_j \hat{R}[\hat{f}]_{3,3}, 1 \rangle_{\hat{K}}| \leq C|\hat{f}|_{5,\hat{K}}.
\]

**Proof.** First, we have

\[
\langle \hat{R}[\hat{f}]_{3,3} - \hat{R}[\hat{f}]_{2,2}, 1 \rangle_{\hat{K}} = \langle M_3(t) \sum_{i=0}^{2} \hat{b}_{i,3} M_i(s) + M_3(s) \sum_{j=0}^{3} \hat{b}_{3,j} M_j(t), 1 \rangle_{\hat{K}} = 0
\]

due to the fact that \( M_3(0) = M_3(1) = 0 \).

We have

\[
\langle \partial_j \hat{R}[\hat{f}]_{3,3}, 1 \rangle_{\hat{K}} = \langle \partial_j \hat{R}[\hat{f}]_{4,4}, 1 \rangle_{\hat{K}} + \langle \partial_j (\hat{R}[\hat{f}]_{4,4} - \hat{R}[\hat{f}]_{3,3}), 1 \rangle_{\hat{K}}
\]

\[
= \langle \partial_j \hat{R}[\hat{f}]_{4,4}, 1 \rangle_{\hat{K}} + \langle M_4(t) \sum_{i=0}^{3} \hat{b}_{i,4} M_i(s) + M_4(s) \sum_{j=0}^{4} \hat{b}_{4,j} M_j(t), 1 \rangle_{\hat{K}}
\]

\[
= \langle \partial_j \hat{R}[\hat{f}]_{4,4}, 1 \rangle_{\hat{K}} + \langle M_4(t) \sum_{i=0}^{2} \hat{b}_{i+1,4} M_i(s), 1 \rangle_{\hat{K}} + \langle M_4(s) \sum_{j=0}^{4} \hat{b}_{4,j} M_j(t), 1 \rangle_{\hat{K}}.
\]

Then by Lemma \(4.1\)

\[
|\langle \partial_j \hat{R}[\hat{f}]_{4,4}, 1 \rangle_{\hat{K}}| \leq C|\hat{f}|_{5,\hat{K}}.
\]

Notice that we have \( \langle l_3(s) \sum_{i=0}^{4} \hat{b}_{4,i} M_i(t), 1 \rangle_{\hat{K}} = 0 \) since the 3-point Gauss-Lobatto quadrature for s-integration is exact and \( l_3(s) \) is orthogonal to 1. Lemma \(4.1\) implies \( |\hat{b}_{i+1,4}| \leq C|\hat{f}|_{5,\hat{K}} \) for \( i = 0,1,2 \), thus we have \( |\langle M_4(t) \sum_{i=0}^{2} \hat{b}_{i+1,4} M_i(s), 1 \rangle_{\hat{K}}| \leq C|\hat{f}|_{5,\hat{K}} \).

**Lemma 4.3** Assume \( a(x,y) \in W^{2,\infty}(\Omega) \). Then

\[
\langle a(u-u_p), v_h \rangle_{e,h} = \mathcal{O}(h^4||a||_{2,\infty}||u||_{3,\infty}||v_h||_2), \quad \forall v_h \in V_h.
\]

**Proof.** As before, we ignore the subscript of \( v_h \) for simplicity. We have

\[
\langle a(u-u_p), v_h \rangle_{e,h} = \sum_e \langle a(u-u_p), v_h \rangle_{e,h},
\]

and on each cell \( e \),

\[
\langle a(u-u_p), v_h \rangle_{e,h} = \langle (R[u]_{3,3}), av_h \rangle_{e,h} = \langle (\hat{R}[\hat{u}]_{3,3}), \hat{a} \hat{v}_h \rangle_{\hat{K}}
\]

\[
= \langle (\hat{R}[\hat{u}]_{3,3}), \hat{a} \hat{v}_h \rangle_{\hat{K}} + \langle (\hat{R}[\hat{u}]_{2,2} - \hat{R}[\hat{u}]_{3,3}), \hat{a} \hat{v}_h \rangle_{\hat{K}}. \tag{4.2}
\]

For the first term in (4.2), we have

\[
\langle (\hat{R}[\hat{u}]_{3,3}), \hat{a} \hat{v}_h \rangle_{\hat{K}} = \langle (\hat{R}[\hat{u}]_{3,3}), \hat{a} \hat{v}_s \rangle_{\hat{K}} + \langle (\hat{R}[\hat{u}]_{3,3}), \hat{a} (\hat{v}_s - \hat{v}_e) \rangle_{\hat{K}}.
\]

By Lemma 4.2

\[
\langle (\hat{R}[\hat{u}]_{3,3}), \hat{a} \hat{v}_s \rangle_{\hat{K}} \leq C|\hat{a}|_{0,\infty} |\hat{u}|_{5,\hat{K}} |\hat{v}_s|_{1,\hat{K}}.
\]
By Lemma 4.1

\[ |(\hat{R}[\hat{u}]_{3,3})_s|_{0,\infty,K} \leq C|\hat{a}|_{4,2,K}. \]

By Bramble-Hilbert Lemma Theorem 3.1, we have

\[
(\hat{R}[\hat{u}]_{3,3}, \hat{a}\nabla_s)_{K} = (\hat{R}[\hat{u}]_{3,3}, \hat{a}\nabla_s)_{K} + (\hat{R}[\hat{u}]_{3,3}, (\hat{a} - \overline{\hat{a}})\nabla_s)_{K} \\
\leq C(|\hat{a}|_{0,\infty,K} |\nabla_1|_{1,K} + |\hat{a} - \overline{\hat{a}}|_{0,\infty,K} |\nabla_1|_{1,K}) \\
\leq C(|\hat{a}|_{0,\infty,K} |\nabla_1|_{1,K} + |\hat{a}|_{1,\infty,K} |\nabla_1|_{1,K}) = \mathcal{O}(h^4)\|a\|_{1,\infty,e}\|u\|_{5,1,e}\|v\|_{1,e},
\]

and

\[
(\hat{R}[\hat{u}]_{3,3}), \hat{a}\nabla_s - \overline{\hat{a}}\nabla_s)_{K} \leq C|\hat{a}|_{4,2,K} |\nabla_1|_{0,\infty,K} |\nabla_1|_{0,\infty,K} \\
\leq C|\hat{a}|_{4,2,K} |\nabla_1|_{0,\infty,K} |\nabla_1|_{0,2,K} = \mathcal{O}(h^4)\|a\|_{4,2,e}\|u\|_{2,2,e}.
\]

Thus,

\[
(\hat{R}[\hat{u}]_{3,3}, \hat{a}\nabla_s)_{K} = \mathcal{O}(h^4)\|a\|_{1,\infty,e}\|u\|_{5,2,e}\|v\|_{2,2,e}.
\] (4.3)

For the second term in (4.2), we have

\[
(\langle \hat{R}[\hat{a}]_{3,3} - \hat{R}[\hat{a}]_{3,3}, \hat{a}\nabla_s \rangle_{K} = \langle (M_3(t) \sum_{i=0}^{2} \hat{b}_{i,3}M_i(s) + M_3(s) \sum_{j=0}^{3} \hat{b}_{3,j}M_j(t)), \hat{a}\nabla_s \rangle_{K} \\
= \langle M_3(t) \sum_{i=0}^{2} \hat{b}_{i,3}I_i(s) + I_2(s) \sum_{j=0}^{3} \hat{b}_{3,j}M_j(t), \hat{a}\nabla_s \rangle_{K} \\
= \langle M_3(t) \sum_{i=0}^{2} \hat{b}_{i,3}I_i(s), \hat{a}\nabla_s \rangle_{K} + \langle I_2(s) \sum_{j=0}^{3} \hat{b}_{3,j}M_j(t), \hat{a}\nabla_s \rangle_{K}.
\] (4.4)

Since \(M_3(t) = \frac{1}{4} (t^3 - t)\) vanishes at \(t = 0, \pm 1\), we have

\[
\langle M_3(t) \sum_{j=0}^{1} \hat{b}_{j+1,3}I_j(s), \hat{a}\nabla_s \rangle_{K} = 0.
\]

For the second term in (4.4),

\[
\langle I_2(s) \sum_{j=0}^{3} \hat{b}_{3,j}M_j(t), \hat{a}\nabla_s \rangle_{K} = \langle I_2(s) \sum_{j=0}^{3} \hat{b}_{3,j}M_j(t), \hat{a}\nabla_s \rangle_{K} + \langle I_2(s) \sum_{j=0}^{3} \hat{b}_{3,j}M_j(t), \hat{a}\nabla_s - \overline{\hat{a}}\nabla_s \rangle_{K} \\
= \langle I_2(s) \sum_{j=0}^{3} \hat{b}_{3,j}M_j(t), (\hat{a} - \overline{\hat{a}})\nabla_s \rangle_{K} + \langle I_2(s) \sum_{j=0}^{3} \hat{b}_{3,j}M_j(t), (\hat{a} - \overline{\hat{a}})(\hat{a}\nabla_s - \overline{\hat{a}}\nabla_s) \rangle_{K} \\
+ \langle I_2(s) \sum_{j=0}^{3} \hat{b}_{3,j}M_j(t), (\hat{a} - \overline{\hat{a}})(\hat{a}\nabla_s - \overline{\hat{a}}\nabla_s) \rangle_{K} + \langle I_2(s) \sum_{j=0}^{3} \hat{b}_{3,j}M_j(t), (\hat{a} - \overline{\hat{a}})(\hat{a}\nabla_s - \overline{\hat{a}}\nabla_s) \rangle_{K} \\
= \langle I_2(s) \sum_{j=0}^{3} \hat{b}_{3,j}M_j(t), (\hat{a} - \overline{\hat{a}})\nabla_s \rangle_{K} + \langle I_2(s) \sum_{j=0}^{3} \hat{b}_{3,j}M_j(t), (\hat{a} - \overline{\hat{a}})(\hat{a}\nabla_s - \overline{\hat{a}}\nabla_s) \rangle_{K}.
\]
where the last step is due to the fact that \( \Pi_s \) and \( \tilde{\psi}_s - \overline{\psi}_s \) are linear functions with respect to variable \( s \), the 3-point Gauss-Lobatto quadrature on \( s \)-integration is exact for polynomial of degree 3, and \( I_2(s) \) is orthogonal to linear functions. By Lemma 4.1 we have

\[
(l_2(s) \sum_{j=0}^3 \hat{b}_{3,j} M_j(t), \hat{a} \hat{v}_s)_K \leq C|\hat{a}|_{3,2,K} (|\hat{a}|_{2,\infty,1,K} + |\hat{a}|_{1,\infty,2,K}) = O(h^4)|\hat{a}|_{2,\infty}||v||_{2,e}.
\]  

(4.5)

Combined with (4.3), we have proved the estimate. \( \square \)

**Lemma 4.4** Assume \( c(x,y) \in W^{2,m}(\Omega) \). Then

\[
\langle a(u-u_p), v_h \rangle_h = O(h^4)||a||_{2,\infty}||v||_{h}, \quad \forall v_h \in V^h.
\]

**Proof.** As before, we ignore the subscript of \( v_h \) for simplicity and

\[
\langle a(u-u_p), v\rangle_h = \sum_e \langle a(u-u_p), v\rangle_{e,h}.
\]

On each cell \( e \) we have

\[
\langle a(u-u_p), v\rangle_{e,h} = \langle R|u|_{2,2}, av\rangle_{e,h} = h^2 \langle \hat{R}[\hat{u}]_{2,2}, \hat{a} \hat{v} - \overline{\hat{a} \hat{v}} \rangle_K = h^2 \langle \hat{R}[\hat{u}]_{2,2}, \hat{a} \hat{v} - \overline{\hat{a} \hat{v}} \rangle_K + h^2 \langle \hat{R}[\hat{u}]_{2,2}, \overline{\hat{a} \hat{v}} \rangle_K.
\]  

(4.6)

For the first term in (4.6), due to the embedding \( H^{3}(\hat{K}) \to C^{0}(\hat{K}) \) and Bramble-Hilbert Lemma Theorem 3.1 and Lemma 4.1 we have

\[
h^2 \langle \hat{R}[\hat{u}]_{2,2}, \hat{a} \hat{v} - \overline{\hat{a} \hat{v}} \rangle_K \leq C h^2 |\hat{R}[\hat{u}]_{2,2}| \langle |\hat{a} \hat{v} - \overline{\hat{a} \hat{v}}|_{\infty} \leq C \langle |\hat{a}|_{3,K} \hat{a} \hat{v} - \overline{\hat{a} \hat{v}} \rangle_{2,K}
\]

\[
\leq C \langle \hat{a}|_{3,K} \rangle (\|\hat{a} \hat{v} - \overline{\hat{a} \hat{v}}\|_{L^2(\hat{K})} + |\hat{a} \hat{v}|_{1,\hat{K}} + |\hat{a} \hat{v}|_{2,\hat{K}})
\]

\[
\leq C \langle \hat{a}|_{3,K} \rangle (|\hat{a} \hat{v}|_{1,\hat{K}} + |\hat{a} \hat{v}|_{2,\hat{K}}) = O(h^4)|a|_{2,\infty,\epsilon,3,e}||u||_{3,e}||v||_{2,e}.
\]

For the second term in (4.6), we have

\[
h^2 \langle \hat{R}[\hat{u}]_{2,2}, \overline{\hat{a} \hat{v}} \rangle_K = h^2 \langle \hat{R}[\hat{u}]_{3,3}, \overline{\hat{a} \hat{v}} \rangle_K - h^2 \langle \hat{R}[\hat{u}]_{3,3}, \hat{R}[\hat{u}]_{2,2}, \overline{\hat{a} \hat{v}} \rangle_K.
\]

By Lemma 4.1 and Lemma 4.2 we have

\[
h^2 \langle \hat{R}[\hat{u}]_{3,3}, \overline{\hat{a} \hat{v}} \rangle_K \leq C \langle \hat{R}[\hat{u}]_{4,4} \rangle |\hat{a} \hat{v}|_{0,K} = O(h^4)||a||_{0,\infty,e}||u||_{4,e}||v||_{0,e},
\]

and

\[
h^2 \langle \hat{R}[\hat{u}]_{3,3} - \hat{R}[\hat{u}]_{2,2}, \overline{\hat{a} \hat{v}} \rangle_K = 0.
\]

Thus, we have \( \langle a(u-u_p), v_h \rangle_h = O(h^4)||a||_{2,\infty}||v||_{h} \). \( \square \)

**Lemma 4.5** Assume \( a(x,y) \in W^{2,m}(\Omega) \). Then

\[
\langle a(u-u_p), v_h \rangle_h = O(h^4)||a||_{2,\infty}||u||_{s}||v_h||_{2}, \quad \forall v_h \in V^h.
\]

**Proof.** As before, we ignore the subscript in \( v_h \) and we have

\[
\langle a(u-u_p), v\rangle_h = \sum_{e} \langle a(u-u_p), v\rangle_{e,h}.
\]
On each cell \( e \), we have

\[
\langle (u - u_p)_s, v \rangle_{e,h} = \langle (R[u]_{2,2})_s, bv \rangle_{e,h} = h(\hat{R}[\hat{u}]_{2,2})_s, \hat{a} \hat{v} \rangle_{\hat{K}}
\]

\[
= h(\hat{R}[\hat{u}]_{3,3}, \hat{a} \hat{v})_{\hat{K}} - h(\hat{R}[\hat{u}]_{3,3} - \hat{R}[\hat{u}]_{2,2})_s, \hat{a} \hat{v} \rangle_{\hat{K}}.
\]  
(4.7)

For the first term in (4.7), we have

\[
\langle (\hat{R}[\hat{u}]_{3,3}, \hat{a} \hat{v}) \rangle_{\hat{K}} \leq \langle (\hat{R}[\hat{u}]_{3,3}, \hat{a} \hat{v} - \hat{a} \hat{o}) \rangle_{\hat{K}} + \langle (\hat{R}[\hat{u}]_{3,3}, \hat{a} \hat{o} - \hat{a} \hat{v}) \rangle_{\hat{K}}
\]

Due to Lemma 4.2

\[
h\langle (\hat{R}[\hat{u}]_{3,3}, \hat{a} \hat{o}) \rangle_{\hat{K}} \leq Ch\|a\|_{0,\infty}\|u\|_{5,\hat{K}}\|v\|_{0,\hat{K}} = \mathcal{O}(h^4)\|a\|_{0,\infty}\|u\|_{5,\hat{e}}\|v\|_{0,\hat{e}},
\]

and by the same arguments as in the proof of Lemma 4.4 we have

\[
h\langle (\hat{R}[\hat{u}]_{3,3}, \hat{a} \hat{v} - \hat{a} \hat{o}) \rangle_{\hat{K}} \leq Ch\|R[a]_{3,3}\|\leq Ch\|a\|_{0,\infty}\|\hat{a} \hat{v} - \hat{a} \hat{o}\|_{2,\hat{K}}
\]

\[
\leq h\|\hat{a} \hat{v} - \hat{a} \hat{o}\|_{2,\hat{K}} + |\hat{a} \hat{o}|_{1,\hat{K}} + |\hat{a} \hat{o}|_{2,\hat{K}} = \mathcal{O}(h^4)\|a\|_{2,\infty}\|u\|_{4,\hat{e}}\|v\|_{2,\hat{e}}
\]

Thus

\[
h\langle (\hat{R}[\hat{u}]_{3,3}, \hat{a} \hat{o}) \rangle_{\hat{K}} = \mathcal{O}(h^4)\|a\|_{2,\infty}\|u\|_{5,\hat{e}}\|v\|_{2,\hat{e}}.
\]  
(4.8)

For the second term in (4.7), we have

\[
\langle (\hat{R}[\hat{u}]_{2,2} - \hat{R}[\hat{u}]_{3,3}, \hat{a} \hat{v}) \rangle_{\hat{K}}
\]

\[
= h\langle (M_3(t) \sum_{j=0}^{1} \hat{b}_{j,3} M_j(s) + M_3(s) \sum_{j=0}^{3} \hat{b}_{3,j} M_j(t), \hat{a} \hat{v}) \rangle_{\hat{K}}
\]

\[
= h\langle (M_3(t) \sum_{j=0}^{1} \hat{b}_{j,1,3} l_j(s) + l_2(s) \sum_{j=0}^{3} \hat{b}_{3,j} M_j(t), \hat{a} \hat{v}) \rangle_{\hat{K}}
\]

\[
= h\langle (M_3(t) \sum_{j=0}^{1} \hat{b}_{j,1,3} l_j(s), \hat{a} \hat{o}) \rangle_{\hat{K}} + l_2(s) \sum_{j=0}^{3} \hat{b}_{3,j} M_j(t), \hat{a} \hat{v}) \rangle_{\hat{K}}
\]

\[
= (l_2(s) \sum_{j=0}^{3} \hat{b}_{3,j} M_j(t), \hat{a} \hat{v}) \rangle_{\hat{K}},
\]

where the last step is due to that \( M_3(t) = \frac{1}{4}(t^3 - t) \) vanishes at \( t = 0, \pm 1 \). Then

\[
\langle (\hat{R}[\hat{u}]_{2,2} - \hat{R}[\hat{u}]_{3,3}, \hat{a} \hat{v}) \rangle_{\hat{K}} = (l_2(s) \sum_{j=0}^{3} \hat{b}_{3,j} M_j(t), \hat{a} \hat{v}) \rangle_{\hat{K}}
\]

\[
= (l_2(s) \sum_{j=0}^{3} \hat{b}_{3,j} M_j(t), \hat{a} \hat{v} - \hat{I}_1(\hat{a} \hat{v}) \rangle_{\hat{K}} + (l_2(s) \sum_{j=0}^{3} \hat{b}_{3,j} M_j(t), \hat{I}_1(\hat{a} \hat{v}) \rangle_{\hat{K}}
\]

\[
= (l_2(s) \sum_{j=0}^{3} \hat{b}_{3,j} M_j(t), \hat{a} \hat{v} - \hat{I}_1(\hat{a} \hat{v}) \rangle_{\hat{K}},
\]

where the last step is due to the facts that \( \hat{I}_1(\hat{a} \hat{v}) \) is a linear function on \( s \)-integration thus the 3-point Gauss-Lobatto quadrature on \( s \)-variable is exact, and \( l_2(s) \) is orthogonal to linear functions.
For simplicity, define

\[ \hat{\Pi_1}(\hat{a}v) \]

and

\[ e \]

By (4.8) and (4.9) and sum up over all the cells, we get the desired estimate.

Thus

\[ h\langle \hat{R}[\hat{a}]_{2,2} - \hat{R}[\hat{a}]_{3,3}, \hat{a}v \rangle_{R} = \mathcal{O}(h^4)||a||_{2,\infty}||u||_{3,e}||v||_{2,e}. \]  \hspace{1cm} (4.9)

By (4.8) and (4.9) and sum up over all the cells, we get the desired estimate. □

**Lemma 4.6** Assume \( a(x,y) \in W^{4,\infty}(\Omega) \). Then

\[ \langle a(u-u_p), (\nu_h)_y \rangle_h = \begin{cases} \mathcal{O}(h^{3.5})||a||_{4,\infty}||u||_{5,h}||v_h||_{2,}\forall v_h \in V_h, \\ \mathcal{O}(h^4)||a||_{4,\infty}||u||_{5,h}||v_h||_{2,}\forall v_h \in V_h^0. \end{cases} \]  \hspace{1cm} (4.10a, 4.10b)

**Proof.** We ignore the subscript in \( \nu_h \) and we have

\[ \langle a(u-u_p), \nu_h \rangle_h = \sum_{e} \langle a(u-u_p), (\nu_h)_e \rangle_{e,h}, \]

and on each cell \( e \)

\[ \langle a(u-u_p), (\nu_h)_e \rangle_{e,h} = \langle (R[a]_{2,2}), a\nu \rangle_{e,h} = \langle (\hat{R}[\hat{a}]_{2,2}), \hat{a}\nu \rangle_{R} \]

\[ = \langle (\hat{R}[\hat{a}]_{3,3}), \hat{a}\nu \rangle_{R} + \langle (\hat{R}[\hat{a}]_{2,2} - \hat{R}[\hat{a}]_{3,3}), \hat{a}\nu \rangle_{R}. \]  \hspace{1cm} (4.11)

By the same arguments as in the proof of Lemma 4.3 we have

\[ \langle (\hat{R}[\hat{a}]_{3,3}), \hat{a}\nu \rangle_{R} = \mathcal{O}(h^4)||a||_{1,\infty}||u||_{5,2,e}||v||_{2,e}. \]  \hspace{1cm} (4.12)

and

\[ \langle (\hat{R}[\hat{a}]_{2,2} - \hat{R}[\hat{a}]_{3,3}), \hat{a}\nu \rangle_{R} = \langle l_2(s) \sum_{j=0}^{3} \hat{b}_3, M_j(t), \hat{a}\nu \rangle_{R}. \]

For simplicity, define

\[ \hat{b}_3(t) := \sum_{j=0}^{3} \hat{b}_3, M_j(t), \]

then by the third and fourth estimates in Lemma 4.1 we have

\[ |\hat{b}_3(t)| \leq C \sum_{j=0}^{3} |\hat{b}_3,j| \leq C|\hat{a}|_{3,R}, \hspace{1cm} |\hat{b}_1^{(2)}(t)| \leq C \sum_{j=1}^{4} |\hat{b}_3,j| \leq C|\hat{a}|_{4,R}, \]

\[ |\hat{b}_3^{(2)}(t)| \leq C \sum_{j=2}^{3} |\hat{b}_3,j| \leq C|\hat{a}|_{4,R}, \hspace{1cm} |\hat{b}_3^{(3)}(t)| \leq C|\hat{b}_3,3| \leq C|\hat{a}|_{4,R}. \]
We use the same technique in the proof of Theorem 3.5:

\[ \langle \tilde{R}[\tilde{u}]_{2,2} - \tilde{R}[\tilde{u}]_{3,3} \rangle_t \tilde{v}_t \| u \|_2, \quad \forall v \in V^h. \]

which is exactly the same integral estimated in the proof of Lemma 3.7 in Li & Zhang (2019b). By the same proof of Lemma 3.7 in Li & Zhang (2019b), after summing over all elements, we have the estimate for the term \( \int_K \int_t (l_2 s) \tilde{b}_3(t) \tilde{a} \tilde{v}_t ds dt \):

\[ \sum_{e} \int_K \int_t (l_2 s) \tilde{b}_3(t) \tilde{a} \tilde{v}_t ds dt = \begin{cases} O(h^3) \| a \|_{4,n} \| s \| \| v \|_2, & \forall v \in V^h, \\ O(h^4) \| a \|_{4,n} \| s \| \| v \|_2, & \forall v \in V_0^h. \end{cases} \]

So we have finished estimating

\[ \int_K \int_t (l_2 \tilde{b}_3(t)) \tilde{v}_t ds dt = \int_K \int_t (l_2 \tilde{b}_3(t)) \tilde{v}_t ds dt + \int_K l_2 \tilde{b}_3(t) \tilde{v}_t ds dt. \]

We only need to estimate the term \( \int_K \tilde{p}(s) \tilde{a} \tilde{v}_t \| l_2 \tilde{b}_3(t) \|_{l_2} \dot{v}_t ds dt \). By product rule of derivatives on the polynomial \( l_2 \tilde{b}_3(t) \tilde{v}_t \) and integration by parts, we have

\[ - \int_K \tilde{p}(s) \tilde{a} \tilde{v}_t \| l_2 \tilde{b}_3(t) \|_{l_2} \dot{v}_t ds dt = -6 \int_K \tilde{p}(s) \tilde{a} \| l_2 \tilde{b}_3(t) \|_{l_2} \| \dot{v}_t \|_2 ds dt 
\]

\[ = 6 \int_K \tilde{p}(s) \tilde{a} \| l_2 \tilde{b}_3(t) \|_{l_2} \| \dot{v}_t \|_2 ds dt - 6 \int_{-1}^1 \tilde{p}(s) \tilde{a} \| l_2 \tilde{b}_3(t) \|_{l_2} \| \dot{v}_t \|_2 ds dt \]
By Cauchy-Schwarz inequality and the estimate (3.3), we have
\[
\int\int_K \tilde{p}(s) \partial_x^2 (l_2 \hat{b}_3 \hat{a}) \hat{v}_{ss} dsdt = \int\int_K \tilde{p}(s) \partial_x^2 [(l_2 \hat{b}_3 \hat{a})_T - l_2 \hat{b}_3 \hat{a}] \hat{v}_{ss} dsdt + \int\int_K \tilde{p}(s) \partial_x^2 (l_2 \hat{b}_3 \hat{a}) \hat{v}_{ss} dsdt
\]
\[
\leq C \|l_2 \hat{b}_3 \hat{a}\|_{3,K} \|\hat{v}\|_{2,K} + C \|\partial_x^2 (l_2 \hat{b}_3 \hat{a})\|_{0,K} \|\hat{v}\|_{2,K}
\]
which gives the estimate \(O(h^4)\|a\|_{3,w,e} \|u\|_{4,w} \|v\|_{2,e}\).

Now we only need to discuss the line integral term.
\[
\int_{t=-1}^1 \tilde{p}(s) \partial_x^2 (l_2 \hat{b}_3 \hat{a}) \hat{v}_{ss} ds \bigg|_{t=-1}^{t=1} = \int_{t=-1}^1 \tilde{p}(s) \partial_x^2 [(l_2 \hat{b}_3 \hat{a})_T - l_2 \hat{b}_3 \hat{a}] \hat{v}_{ss} ds \bigg|_{t=-1}^{t=1} + \int_{t=-1}^1 \tilde{p}(s) \partial_x^2 (l_2 \hat{b}_3 \hat{a}) \hat{v}_{ss} ds \bigg|_{t=-1}^{t=1}.
\]

By (3.14), trace inequality and Theorem 3.1 we have
\[
\left| \int_{t=-1}^1 \tilde{p}(s) \partial_x^2 [(l_2 \hat{b}_3 \hat{a})_T - l_2 \hat{b}_3 \hat{a}] \hat{v}_{ss} ds \bigg|_{t=-1}^{t=1} \right| \leq C \|\hat{v}_{ss}\|_{0,K} \|\partial_x^2 [(l_2 \hat{b}_3 \hat{a})_T - l_2 \hat{b}_3 \hat{a}]\|_{0,0,K}
\]
\[
\leq C \|\hat{v}\|_{2,K} \|\partial_x^2 [(l_2 \hat{b}_3 \hat{a})_T - l_2 \hat{b}_3 \hat{a}]\|_{1,K} \leq C \|\hat{v}\|_{2,K} \|l_2 \hat{b}_3 \hat{a}\|_{3,K}
\]
\[
\leq C \|\hat{v}\|_{2,K} \|l_2 \hat{b}_3 \hat{a}\|_{3,K} = O(h^4)\|a\|_{3,w,e} \|u\|_{4,w} \|v\|_{2,e}.
\]

After mapping back to original cell \(e\), we have
\[
\int_{t=-1}^1 \tilde{p}(s) \partial_x^2 (l_2(s) \hat{b}_3(t) \hat{a}) \hat{v}_{ss} ds \bigg|_{t=-1}^{t=1} = h^3 \int_{x=-h}^{x=h} \tilde{p}(\frac{x-x_e}{h}) \partial_x^2 \left( l_2(\frac{x-x_e}{h}) \hat{b}_3(\frac{y-y_e}{h})a \right) v_{ss} dx \bigg|_{y=y_e-h}^{y=y_e+h}.
\]

Notice that we have
\[
\hat{b}_3(1) = \sum_{j=0}^3 \hat{b}_{3,j} M_j(1) = \hat{b}_{3,0} + \hat{b}_{3,1} = \frac{5}{2} \int_{-h}^{x=h} \partial_x \hat{u}(s,1) l_2(s) ds = \frac{5}{2} \int_{-h}^{x=h} \partial_x u(x,y_e+h) l_2(\frac{x-x_e}{h}) dx,
\]
and similarly we get
\[
\hat{b}_3(-1) = \frac{5}{2} \int_{x=-h}^{x=h} \partial_x u(x,y_e-h) l_2(\frac{x-x_e}{h}) dx.
\]

Thus the term \(l_2(\frac{x-x_e}{h}) \hat{b}_3(\frac{y-y_e}{h})a\) is continuous across the top/bottom edge of cells, and so is term \(\partial_x^2 (l_2(\frac{x-x_e}{h}) \hat{b}_3(\frac{y-y_e}{h})a)\). Therefore, if summing over all elements \(e\), the line integral on the inner edges are cancelled out. Let \(L_1\) and \(L_3\) denote the top and bottom boundary edges of \(\Omega\). Then the line integral after summing over \(e\) consists of two line integrals along \(L_1\) and \(L_3\). We only need to discuss one of them.

Let \(l_1\) and \(l_3\) denote the top and bottom edge of \(e\). First, after integration by parts twice, we get
\[
\hat{b}_3(1) = \frac{5}{2} \int_{-1}^1 \partial_x \hat{u}(s,1) l_2(s) ds = \frac{5}{2} \int_{-1}^1 \frac{\partial^3}{\partial s^3} \hat{u}(s,1) \frac{1}{8} (s^2 - 1)^2 ds,
\]
thus by Cauchy-Schwarz inequality we get
\[
|\hat{b}_3(1)| \leq C \sqrt{\int_{-1}^1 \left( \frac{\partial^3}{\partial s^3} \hat{u}(s,1) \right)^2 ds} \leq C h^2 |u|_{3,2,1}.
\]
Second, by (3.14), we get
\[
\int_{-1}^{1} |\hat{v}_{ss}(s,1)| ds \leq 2 \int_{-1}^{1} |\hat{v}_{ss}(s,1)|^2 ds \leq C|\hat{v}_{ss}|_{0,K}.
\]

The line integral along $L_1$ can be estimated by considering each $e$ adjacent to $L_1$ in the reference cell:
\[
\sum_{e \cap L_1 \neq \emptyset} \left| \int_{-1}^{1} \tilde{\varphi}(s) \hat{b}_{3}(1) \hat{v}_{ss}(s,1) ds \right| = \sum_{e \cap L_1 \neq \emptyset} \left| \int_{-1}^{1} \tilde{\varphi}(s) \hat{b}_{3}(1) (l_2(s) \hat{a})_{ss} \hat{v}_{ss}(s,1) ds \right| \leq \sum_{e \cap L_1 \neq \emptyset} C \| \tilde{\varphi} \|_{2,\infty,K} \| \hat{b}_{3}(1) \| \| \hat{v}_{ss} \|_{0,K} = O(h^{3.5}) \sum_{e \cap L_1 \neq \emptyset} \| a \|_{2,\infty} \| v \|_{2,\infty} \leq C \| u \|_{4,\Omega} \| v \|_{2,\Omega},
\]
where the trace inequality $\| u \|_{3,2,\Omega} \leq C \| u \|_{4,\Omega}$ is used.

Combine all the estimates above, we get (4.10a). Since the $\frac{1}{2}$ order loss is only due to the line integral along $L_1$ and $L_3$, on which $v_{ss} = 0$ if $v \in V_h^0$, we get (4.10b).

By all the discussions in this subsection, we have proven (4.1a) and (4.1b).

5. Homogeneous Dirichlet Boundary Conditions

5.1 $V^h$-ellipticity

In order to discuss the scheme (1.2), we need to show $A_h$ satisfies $V^h$-ellipticity (2.5). We first consider the $V_h$-ellipticity for the case $b \equiv 0.$

**Lemma 5.1** Assume the coefficients in (2.3) satisfy that $b \equiv 0,$ both $c(x,y)$ and the eigenvalues of $a(x,y)$ have a uniform upper bound and a uniform positive lower bound, then there exists two constants $C_1, C_2 > 0$ independent of mesh size $h$ such that
\[
\forall v_h \in V_h^0, \quad C_1 \| v_h \|_1^2 \leq A_h(v_h,v_h) \leq C_2 \| v_h \|_1^2.
\]

**Proof.** Let $Z_{0,K}$ denote the set of $3 \times 3$ Gauss-Lobatto points on the reference cell $K$. First we notice that the set $Z_{0,K}$ is a $Q^2(K)$-unisolvent subset. Since the Gauss-Lobatto quadrature weights are strictly positive, we have
\[
\forall \hat{p} \in Q^2(K), \quad \sum_{i=1}^{n} \langle \partial_i \hat{p}, \partial_i \hat{p} \rangle_K = 0 \implies \partial_i \hat{p} = 0 \text{ at quadrature points},
\]
where $i = 1, \ldots, n$ representing the spatial derivative on variable $x_i$ respectively. Since $\partial_i \hat{p} \in Q^2(K)$ and it vanishes on a $Q^2(K)$-unisolvent subset, we have $\partial_i \hat{p} \equiv 0$. As a consequence, $\sqrt{\sum_{i=1}^{n} \langle \partial_i \hat{p}, \partial_i \hat{p} \rangle_K}$ defines a norm over the quotient space $Q^2(K)/Q^0(K)$. Since that $\| \cdot \|_{1,K}$ is also a norm over the same quotient space, by the equivalence of norms over a finite dimensional space, we have
\[
\forall \hat{p} \in Q^2(K), \quad C_1 \hat{p}_{1,K}^2 \leq \sum_{i=1}^{n} \langle \partial_i \hat{p}, \partial_i \hat{p} \rangle_K \leq C_2 \hat{p}_{1,K}^2.
\]
On the reference cell $\hat{K}$, by the assumption on the coefficients, we have

$$C_1\|\hat{v}_h\|^2_{\hat{K}} \leq C_2 \sum_{i} \langle \partial_i \hat{v}_h, \partial_i \hat{v}_h \rangle_{\hat{K}} \leq \sum_{i,j=1}^n \langle (\partial_{ij} \hat{v}_h, \partial_j \hat{v}_h)_{\hat{K}} + (\partial_i \hat{v}_h, \partial_i \hat{v}_h)_{\hat{K}} \rangle \leq C_2 \|\hat{v}_h\|_{1,\hat{K}}^2$$

Mapping these back to the original cell $e$ and summing over all elements, by the equivalence of two norms $\| \cdot \|_1$ and $\| \cdot \|_1$ for the space $H^1_0(\Omega) \supset V_h^0$ (see Ciarlet (1991)), we get $C_1\|v_h\|^2_1 \leq A_h(v_h, v_h) \leq C_2\|v_h\|_1^2$. 

**Remark 5.1** For discussing continuity of $A_h(\cdot, \cdot)$ when $b$ is nonzero, if $b$ has a uniform upper bound, we have

$$\sum_i \langle b_i \partial_i \hat{v}_h, \hat{v}_h \rangle_{\hat{K}} \leq \sum_i |b_i| \langle (\partial_i \hat{v}_h, \partial_i \hat{v}_h)_{\hat{K}} \rangle \leq C\|v_h\|_1^2.$$ 

**Lemma 5.2** Assume $b \in W^{4,\infty}(\hat{\Omega})$,

$$(b \cdot \nabla v_h, v_h) + (c v_h, v_h) = (b \cdot \nabla v_h, v_h) - (c v_h, v_h) = o(h)\|v_h\|_1^2, \quad \forall v_h \in V_h^0.$$

**Proof.** By Bramble-Hilbert Lemma Theorem 3.1 and inverse estimates (2.1), on each cell $e$ we have

$$(b \cdot \nabla v_h, v_h)_e + (c v_h, v_h)_e - (c v_h, v_h)_e = o(h)\|v_h\|_1^2.$$ 

By standard estimates on the dual problem $5.2$ assume $v_h \in H^1_0(\Omega), \quad (4.4)$ still holds for $v = v_h$, thus $b \cdot \nabla v_h, v_h \geq 0, \forall v_h \in V_h^0$. So combine Lemma 5.1 and Lemma 5.2 we get $A_h(v_h, v_h) \geq C_3\|v_h\|_1^2 - C_4\|v_h\|_2^2$ where $C_3, C_4 > 0$ is from the uniform lower bound of eigenvalues of $a$ and $C_4$ is from Lemma 5.2. Thus if $h$ is small enough, (2.5) still holds for the variable coefficient $b$.

To discuss $V^h$-ellipticity for variable coefficient $b$ with arbitrary mesh size $h$, it depends on whether (2.4) still holds with quadrature. We do not discuss this matter in this paper.

### 5.2 Standard estimates for the dual problem

In order to apply the Aubin-Nitsche duality argument for establishing superconvergence of function values, we need certain estimates on a proper dual problem. Define $\theta_h := u_h - u_p$. Then we consider the dual problem: find $w \in H^1_0(\Omega)$ satisfying

$$A^*(w, v) = (\theta_h, v), \quad \forall v \in H^1_0(\Omega),$$

where $A^*(\cdot, \cdot)$ is the adjoint bilinear form of $A(\cdot, \cdot)$ such that $A^*(u, v) = A(v, u)$. Let $w_h \in V_h^0$ be the solution to

$$A_h^*(w_h, v_h) = (\theta_h, v_h), \quad \forall v_h \in V_h^0.$$

Notice that the right hand side of (5.2) is different from the right hand side of the scheme (1.2).

We need the following standard estimates on $w_h$ for the dual problem.

**Theorem 5.1** Assume all coefficients in (2.3) are in $W^{4,\infty}(\Omega)$, elliptic regularity and $V^h$ ellipticity holds, we have

\[
\|w - w_h\|_1 \leq C h \|w\|_2, \quad \|w_h\|_2 \leq C \|	heta_h\|_0.
\]
By $V^h$ ellipticity, we have $C_1\|w_h - v_h\|^2 \leq A^*_h(w_h - v_h, w_h - v_h)$. By the definition of the dual problem, we have

$$A^*_h(w_h, w_h - v_h) = \langle \theta_h, w_h - v_h \rangle = A^*(w, w_h - v_h), \quad \forall v_h \in V^h_0.$$ 

Thus for any $v_h \in V^h_0$, by Theorem 3.4 we have

$$C_1\|w_h - v_h\|^2 \leq A^*_h(w_h - v_h, w_h - v_h) = A^*(w - v_h, w_h - v_h) + [A^*_h(w_h, w_h - v_h) - A^*(w, w_h - v_h)] + [A^*(v_h, w_h - v_h) - A^*_h(v_h, w_h - v_h)] \leq C\|w - v_h\|_1\|w_h - v_h\|_1 + C\|w - v_h\|_1 + C\|w_h - v_h\|_2.$$ 

Thus

$$\|w - w_h\|_1 \leq \|w - v_h\|_1 + \|w_h - v_h\|_1 \leq C\|w - v_h\|_1 + C\|w_h - v_h\|_2.$$ 

Now consider $\Pi_1 w \in V^h_0$ where $\Pi_1$ is the piecewise $Q^1$ projection and its definition on each cell is defined through (2.2) on the reference cell. By the Bramble Hilbert Lemma Theorem 3.1 on the projection error, we have

$$\|w - \Pi_1 w\|_1 \leq C\|w - \Pi_1 w\|_2, \quad \|w - \Pi_1 w\|_2 \leq C\|w\|_2,$$ 

thus $\|\Pi_1 w\|_2 \leq \|w\|_2 + \|w - \Pi_1 w\|_2 \leq C\|w\|_2$. By setting $v_h = \Pi_1 w$, from (5.3) we have

$$\|w - w_h\|_1 \leq C\|w - \Pi_1 w\|_1 + C\|\Pi_1 w\|_2 \leq C\|w\|_2.$$ 

By the inverse estimate on the piecewise polynomial $w_h - \Pi_1 w$, we get

$$\|w_h\|_2 \leq \|w_h - \Pi_1 w\|_2 + \|\Pi_1 w - w\|_2 \leq C^{-1}\|w_h - \Pi_1 w\|_1 + C\|w\|_2.$$ 

By (5.4) and (5.5), we also have

$$\|w_h - \Pi_1 w\|_1 \leq \|w - \Pi_1 w\|_1 + \|w - w_h\|_1 \leq C\|w\|_2.$$ 

With (5.6), (5.7) and the elliptic regularity $\|w\|_2 \leq C\|\theta_0\|_0$, we get

$$\|w_h\|_2 \leq C\|w\|_2 \leq C\|\theta_0\|_0.$$ 

5.3 Superconvergence of function values

**Theorem 5.2** Assume $a_{ij}, b_i, c \in W^{4,\infty}(\Omega)$ and $u(x,y) \in H^3(\Omega)$, $f(x,y) \in H^1(\Omega)$. Assume $h$ is small enough so that $V^h$ ellipticity holds. Then $u_h$ is a fourth order accurate approximation to $u$ in the discrete 2-norm over all the $3 \times 3$ Gauss-Lobatto points:

$$\|u_h - u\|_{2,Z_0} = O(h^4)(\|u\|_5 + \|f\|_4).$$

**Proof.** By Theorem 3.5 and Theorem 3.3 for any $v_h \in V^h_0$

$$A_h(u - u_h, v_h) = [A(u, v_h) - A_h(u, v_h)] + [A_h(u, v_h) - A(u, v_h)] = A(u, v_h) - A_h(u, v_h) + O(h^4)\|u\|_{4,\infty}\|u\|_5\|v_h\|_2 = [f, v_h] - (f, v_h)h + O(h^4)\|u\|_5\|v_h\|_2 = O(h^4)(\|u\|_5 + \|f\|_4)\|v_h\|_2.$$

□
Let \( \theta_h = u_h - u_p \), then \( \theta_h \in V_0^h \) due to the properties of the M-type projection. So by (4.1b) and Theorem 5.1 we get
\[
\| \theta_h \|_0^2 = (\theta_h, \theta_h) = A_h(\theta_h, w_h) = A_h(u_h - u, w_h) + A_h(u - u_p, w_h)
\]

Thus
\[
A_h(u - u_p, w_h) + \mathcal{O}(h^4)(\|u\|_5 + \|f\|_4)\|w_h\|_2 = \mathcal{O}(h^4)(\|u\|_5 + \|f\|_4)\|\theta_h\|_0.
\]

Finally, by the equivalence of the discrete 2-norm on \( Z_0 \) and the \( L^2(\Omega) \) norm in finite-dimensional space \( V^h \) and Theorem 4.2 we obtain
\[
\|u_h - u_p\|_0 = \mathcal{O}(h^4)(\|u\|_5 + \|f\|_4).
\]

**Remark 5.2** To extend the discussions to Neumann type boundary conditions, due to (4.1a) and Lemma 5.5 we can only prove 3.5-th order accuracy:
\[
\|u_h - u\|_{2, Z_0} = \mathcal{O}(h^{3.5})(\|u\|_5 + \|f\|_4).
\]

**Remark 5.3** All key discussions can be extended to three-dimensional cases.

### 6. Nonhomogeneous Dirichlet Boundary Conditions

We consider a two-dimensional elliptic problem on \( \Omega = [0, 1]^2 \) with nonhomogeneous Dirichlet boundary condition,
\[
-\nabla(a\nabla u) + b \cdot \nabla u + cu = f \quad \text{on } \Omega
\]
\[
= g \quad \text{on } \partial \Omega. \tag{6.1}
\]

Assume there is a function \( \tilde{g} \in H^1(\Omega) \) as a smooth extension of \( g \) so that \( \tilde{g}|_{\partial \Omega} = g \). The variational form is to find \( \tilde{u} = u - \tilde{g} \in H^1_0(\Omega) \) satisfying
\[
A(\tilde{u}, v) = (f, v) - A(\tilde{g}, v), \quad \forall v \in H^1_0(\Omega). \tag{6.2}
\]

In practice, \( \tilde{g} \) is not used explicitly. By abusing notations, the most convenient implementation is to consider
\[
g(x, y) = \begin{cases} 
0, & \text{if } (x, y) \in (0, 1) \times (0, 1), \\
g(x, y), & \text{if } (x, y) \in \partial \Omega,
\end{cases}
\]
and \( g_I \in V^h \) which is defined as the \( Q^2 \) Lagrange interpolation at \( 3 \times 3 \) Gauss-Lobatto points for each cell on \( \Omega \) of \( g(x, y) \). Namely, \( g_I \in V^h \) is the piecewise quadratic interpolation of \( g \) along the boundary grid points and \( g_I = 0 \) at the interior grid points. The numerical scheme is to find \( \tilde{u}_h \in V_0^h \), s.t.
\[
A_h(\tilde{u}_h, v_h) = (f, v_h) - A_h(g_I, v_h), \quad \forall v_h \in V_0^h. \tag{6.3}
\]

Then \( u_h = \tilde{u}_h + g_I \) will be our numerical solution for (6.1). Notice that (6.3) is not a straightforward approximation to (6.2) since \( \tilde{g} \) is never used. Assuming elliptic regularity and \( V^h \) ellipticity hold, we will show that the numerical solution \( u_h - u \) is of fourth order in the discrete 2-norm over all \( 3 \times 3 \) Gauss-Lobatto points.
6.1 Auxiliary schemes

In order to discuss the superconvergence of \((6.3)\), we need to prove the superconvergence of two auxiliary schemes. Notice that we discuss these two auxiliary schemes only for proving the accuracy of \((6.3)\). In practice one should not implement the auxiliary schemes since \((6.3)\) is a much more convenient implementation and they all have the same accuracy.

The first auxiliary scheme is to find \(\tilde{u}_h^* \in V_0^h\) satisfying

\[
A_h(\tilde{u}_h^*, v_h) = (f, v_h) - A_h(g_p, v_h), \quad \forall v_h \in V_0^h, \tag{6.4}
\]

where \(g_p \in V^h\) is the piecewise M-type \(Q^2\) projection of the smooth extension function \(\tilde{g}\). Then \(u_h^{**} = \tilde{u}_h^* + g_p\) is the numerical solution of scheme \((6.4)\) for problem \((6.2)\).

Define \(\theta_h = u_h^{**} - u_p\), then by Theorem 4.1 we have \(\theta_h \in V_0^h\). Following Section 5.2, define the following dual problem: find \(w_h \in H^1_0(\Omega)\) satisfying

\[
A^*(w, v) = (\theta_h, v), \quad \forall v \in H^1_0(\Omega). \tag{6.5}
\]

Let \(w_h \in V^h\) be the solution to

\[
A^*_h(w_h, v_h) = (\theta_h, v_h), \quad \forall v_h \in V_0^h. \tag{6.6}
\]

Notice that the dual problem has homogeneous Dirichlet boundary conditions. By Theorem 3.3, for any \(v_h \in V_0^h\),

\[
A_h(u - u_h^{**}, v_h) = [A(u, v_h) - A_h(u_h^{**}, v_h)] + [A_h(u, v_h) - A(u, v_h)]
= A(u, v_h) - A_h(u_h^{**}, v_h) + O(h^4)|||u|||_s ||v||_2
= [(f, v_h) - (f, v_h)_h] + O(h^4)|||u|||_s ||v||_2
= O(h^4)(|||u|||_s + ||f||_4) ||v||_2.
\]

By (4.1b) and Theorem 5.1, we get

\[
\|\theta_h\|_0 = (\theta_h, \theta_h) = A_h(\theta_h, w_h) = A_h(u_h^{**} - u, w_h) + A_h(u - u_p, w_h)
= O(h^4)|||u|||_s + ||f||_4) ||w||_2
= O(h^4)(|||u|||_s + ||f||_4) ||\theta_h||_0,
\]

thus \(\|u_h^{**} - u_p\|_0 = \|\theta_h\|_0 = O(h^4)(|||u|||_s + ||f||_4)\). So Theorem 5.2 still holds for the first auxiliary scheme \((6.4)\).

Next define \(g_p \in V^h\) as \(g_p = \tilde{g}_p\) on \(\partial \Omega\) and \(g_p = 0\) at all the inner grids. The second auxiliary scheme is to find \(\tilde{u}_h^* \in V_0^h\) satisfying

\[
A_h(\tilde{u}_h^*, v_h) = (f, v_h) - A_h(g_p, v_h), \quad \forall v_h \in V_0^h. \tag{6.7}
\]

Then \(u_h^* = \tilde{u}_h^* + g_p\) is the numerical solution. We have

\[
A_h(u_h^{**} - u_h^*, v_h) = 0, \quad \forall v_h \in V_0^h.
\]

Since \(u_h^{**} - u_h^* \in V_0^h \subset H^1_0(\Omega)\), by \(V^h\)-ellipticity we have

\[
\|u_h^{**} - u_h^*\|_{1,\Omega}^2 \leq CA_h(u_h^{**} - u_h^*, u_h^{**} - u_h^*).
\]

Thus we get \(u_h^{**} = u_h^*\). So numerical solutions from \((6.4)\) and \((6.7)\) are the same. Thus Theorem 5.2 also holds for \(u_h^*\):

\[
\|u_h^* - u\|_{2,\Omega} = O(h^4)(|||u|||_s + ||f||_4). \tag{6.8}
\]
6.2 The main result
In order to extend Theorem 5.2 to (6.3), we only need to prove
\[ \|u_h - u_h^p\| = O(h^4). \]

The difference between (6.7) and (6.3) is
\[ A_h(\tilde{u}_h^p - \tilde{u}_h, v_h) = A_h(g_l - g_p, v_h), \quad \forall v_h \in V_h^0. \]  

We need the following Lemma.

**Lemma 6.1**
\[ A_h(g_l - g_p, v_h) = O(h^4)\|u\|_{5,\Omega}\|v_h\|_{2,\Omega}, \quad \forall v_h \in V_h^0. \]  

**Proof.** For simplicity, we ignore the subscript \( h \) of \( v_h \) in this proof and all the following \( v \) are in \( V_h \).

Notice that \( g_l - g_p \equiv 0 \) in interior cells thus we only need to consider cells adjacent to \( \partial \Omega \). Let \( L_1, L_2, L_3 \) and \( L_4 \) denote the top, left, bottom and right boundary edges of \( \hat{\Omega} = [0, 1] \times [0, 1] \) respectively. Let \( L_1, L_2, L_3, L_4 \) denote the top, left, bottom and right boundary edges of \( e \) respectively. Without loss of generality, we only consider a cell \( e = [x_e - h, x_e + h] \times [y_e - h, y_e + h] \) adjacent to the left boundary \( L_2 \), i.e., \( x_e - h = 0 \). On \( L_2 \subset L_2 \), we have \( g_p(0, y_e + h) = g(0, y_e + h) \), \( g_p(0, y_e - h) = g(0, y_e - h) \) and
\[ \int_{y_e - h}^{y_e + h} gdy = \int_{y_e - h}^{y_e + h} g_pdy = \frac{h}{3}[g_p(0, y_e - h) + 4g_p(0, y_e) + g_p(0, y_e + h)]. \]

Thus
\[ g_p(0, y_e) = \frac{3}{4h} \int_{y_e - h}^{y_e + h} gdy - \frac{1}{4}g(0, y_e + h) - \frac{1}{4}g(0, y_e - h). \]

By (5.8), we have
\[ g_p(0, y_e) - g(0, y_e) = \frac{3}{4h} \left[ \int_{y_e - h}^{y_e + h} gdy - \frac{h}{3}g(0, y_e - h) + \frac{4h}{3}g(0, y_e) + \frac{h}{3}g(0, y_e + h) \right] = O(h^3)\|g\|_{4,1,2} = O(h^3)\|u\|_{4,1,2} = O(h^3.5). \]

where the last step is by Cauchy-Schwartz inequality.

We first consider the case that the cell \( e \) is not adjacent to \( L_1 \) or \( L_3 \). For this case, \( g_l - g_p \) is nonzero at \( (0, y_e) \) and \( g_l - g_p = 0 \) at other \( 3 \times 3 \) Gauss-Lobatto points. Let \( \lambda = g_l(0, y_e) - g_p(0, y_e) \), then \( g_l - g_p \) is a \( Q^2 \) polynomial on cell \( e \) satisfying that \( q(0, y_e) = 1 \) and \( q(x, y) = 0 \) at other \( 3 \times 3 \) Gauss-Lobatto points.

Next we estimate \( (a(g_l - g_p), v_e) \),
\[ \langle a(g_l - g_p), v_e \rangle_e = \langle a\lambda q_s, v_e \rangle_e = \lambda \langle \tilde{a} \tilde{q}_s, \tilde{v}_s \rangle_K = \lambda \int_K \tilde{a} \tilde{q}_s \tilde{v}_s d^h t = \lambda \int_K (\hat{a} - \tilde{a}) \hat{q}_s \hat{v}_s d^h t + \lambda \int_K \tilde{a} \tilde{q}_s \tilde{v}_s d^h t. \]

By Theorem 5.1 and the equivalence of norms on finite-dimensional space, we have
\[ \int_K (\hat{a} - \tilde{a}) \hat{q}_s \hat{v}_s d^h t \leq C |\hat{a} - \tilde{a}|_{|e|, K} |\hat{q}_s|_{|e|, K} |\hat{v}_s|_{|e|, K} \leq C |a|_{1, \infty, K} \hat{v}_s |_{0, K} = O(h) |a|_{1, \infty} |v|_{1,e}. \]  

(6.12)
We have
\[\int_{-1}^{1} \hat{g}_s \hat{v}_s d^h t = \tilde{a} \int_{-1}^{1} \hat{g}_s \hat{v}_s d^h t = \tilde{a} \int_{-1}^{1} \hat{g}_s \hat{v}_s d^h t \]
\[= \tilde{a} \int_{-1}^{1} \hat{g}_s(-1,t) \hat{v}_s(-1,t) d^h t - \tilde{a} \int_{K} \hat{g}_s \hat{v}_s d^h t \]
\[\leq C\tilde{a}(\|\hat{v}_s(-1,0)\| + |\hat{v}_s|_{0,K}) \leq C\tilde{a}|\hat{v}_s(-1,0)| + |\hat{v}_s|_{0,K}.\]

By equivalence of norms for the finite dimensional Banach space consisting of all quadratic polynomials on \([-1, 1]\), we have
\[|\hat{v}_s(-1,0)| \leq \max_{t \in [-1,1]} |\hat{v}_s(-1,t)| \leq C \left( \int_{-1}^{1} |\hat{v}_s(-1,t)|^2 dt \right)^{\frac{1}{2}}.\]

So we have
\[\int_{-1}^{1} \tilde{a} \hat{g}_s \hat{v}_s (d^h t) \leq C |\hat{a}|_{\infty, K} \left[ \left( \int_{-1}^{1} |\hat{v}_s(-1,t)|^2 dt \right)^{\frac{1}{2}} + |\hat{v}_s|_{2, K} \right] \leq C |a|_{\infty, e}(h^2 |v|_{0,J_2} + h^4 |v|_{2,e}). \quad (6.13)\]

From (6.12) and (6.13), we get
\[\langle a(g_l - g_p), v_e \rangle_e = \mathcal{O}(h^{0.5}) \lambda \|a\|_{1,\infty}(|v|_{1,J_2} + h^{0.5} \|v\|_{2,e}) = \mathcal{O}(h^4) \|a\|_{1,\infty} \|u\|_{4,J_2}(|v|_{1,J_2} + h^{0.5} \|v\|_{2,e}).\]

For the case that the cell \(e\) is also adjacent to \(L_1\) or \(L_3\). Without loss of generality, assume \(e\) is adjacent to \(L_3\), then \(v_e - h = 0\) and \(g_l - g_p\) are nonzero only at two of the nine Gauss-Lobatto points \((x_e - h, y_e) = (0, y_e)\) and \((x_e, y_e - h) = (x_e, 0)\). Let \(\lambda = g_l(0, y_e) - g_p(0, y_e)\) and \(\mu = g_l(x_e, 0) - g_p(x_e, 0)\). Then \((g_l - g_p)|_e = \lambda q(x, y) + \mu p(x, y)\), where \(p(x, y)\) is a \(Q^2\) polynomial on cell \(e\) satisfying \(p(x_e, 0) = 1\) and \(p(x_e, y) = 0\) at other \(3 \times 3\) Gauss-Lobatto points.

Similar to (6.11), we can derive \(\mu = \mathcal{O}(h^{5.5}) \|u\|_{4, J_3}\).

We have
\[\langle a(g_l - g_p), v_e \rangle_e = \langle \lambda q(x_e, v_e) + \mu p(x_e, v_e) \rangle_K = \lambda \langle \hat{a} \hat{q}_s, \hat{v}_s \rangle_K + \mu \langle \hat{a} \hat{p}_s, \hat{v}_s \rangle_K.\]

We only need to estimate \(\langle \hat{a} \hat{p}_s, \hat{v}_s \rangle_K = \langle \langle \hat{a} - \bar{a} \rangle \hat{p}_s, \hat{v}_s \rangle_K + \langle \hat{a} \hat{p}_s, \hat{v}_s \rangle_K.\) By similar discussions as above, we have
\[\int_{K} (\hat{a} - \bar{a}) \hat{p}_s \hat{v}_s d^h t = \mathcal{O}(h) \|a\|_{1,\infty} \|v\|_{1,e},\]
and
\[\int_{K} \bar{a} \hat{p}_s \hat{v}_s d^h t = \mathcal{O}(h) \|a\|_{1,\infty} \|v\|_{1,e} \leq C |\hat{a}|_{\infty, K} \|v\|_{0,K}.\]

where the fact that \(\hat{a}(-1, t) = \hat{a}(1, t) = 0\) is used. Thus for the left lower corner cell \(e\), we have
\[\langle a(g_l - g_p), v_e \rangle_e = \mathcal{O}(h^4) \|a\|_{1,\infty} \|u\|_{4,J_2}(|v|_{1,J_2} + h^{0.5} \|v\|_{2,e}) + \mathcal{O}(h^{4.5}) \|a\|_{1,\infty} \|u\|_{4,J_3} \|v\|_{2,e}.\]

We can get similar estimates for all boundary cells. Sum up over all the boundary elements, by Cauchy-Schwartz inequality we have
\[a(g_l - g_p), v_e \rangle_h = \mathcal{O}(h^4) \|a\|_{1,\infty} \|u\|_{4,\partial \Omega}(|v|_{1,\Omega} + h^{0.5} \|v\|_{2,\Omega}).\]
With trace inequality \( \|u\|_{4,\partial\Omega} \leq C\|u\|_{5,\Omega} \), we get
\[
\langle a(g_1-g_p), v \rangle_h = \mathcal{O}(h^4)\|a\|_{1,\infty}\|u\|_s\|v\|_2, \quad \forall v \in V_0^h. \tag{6.14}
\]
Similarly, for any \( v \in V_0^h \), we have
\[
\langle a(g_1-g_p), v \rangle_h = \mathcal{O}(h^4)\|a\|_{1,\infty}\|u\|_s\|v\|_2, \\
\langle b \cdot \nabla (g_1-g_p), v \rangle_h = \mathcal{O}(h^4)\|b\|_{1,\infty}\|u\|_s\|v\|_2, \\
\langle c(g_1-g_p), v \rangle_h = \mathcal{O}(h^4)\|c\|_{1,\infty}\|u\|_s\|v\|_2.
\]
Thus we conclude that
\[
A_h(g_1-g_p, v_h) = \mathcal{O}(h^4)\|u\|_s\|v_h\|_2, \quad \forall v_h \in V_0^h. \tag{6.15}
\]
By (6.9) and Lemma 6.1 we have
\[
A_h(\tilde{u}_h^* - \tilde{u}_h, v_h) = \mathcal{O}(h^4)\|u\|_s\|v_h\|_2, \quad \forall v_h \in V_0^h. \tag{6.16}
\]
Let \( \theta_h = \tilde{u}_h^* - \tilde{u}_h \in V_0^h \). Following Section 5.2 define the following dual problem: find \( w \in H_0^1(\Omega) \) satisfying
\[
A^*(w, v) = (\theta_h, v), \quad \forall v \in H_0^1(\Omega). \tag{6.17}
\]
Let \( w_h \in V_0^h \) be the solution to
\[
A_h^*(w_h, v_h) = (\theta_h, v_h), \quad \forall v_h \in V_0^h. \tag{6.18}
\]
By (6.15) and Theorem 5.1 we get
\[
\|\theta_h\|_0 = (\theta_h, \theta_h) = A_h^*(w_h, \theta_h) = A_h(\tilde{u}_h^* - \tilde{u}_h, w_h) = \mathcal{O}(h^4)\|u\|_s\|w_h\|_2 = \mathcal{O}(h^4)\|u\|_s\|\theta_h\|_0,
\]
thus \( \|\tilde{u}_h^* - \tilde{u}_h\|_0 = \|\theta_h\|_0 = \mathcal{O}(h^4)\|u\|_s \). By equivalence of norms for polynomials, we have
\[
\|\tilde{u}_h^* - \tilde{u}_h\|_{2, Z_0} \leq C\|\tilde{u}_h^* - \tilde{u}_h\|_0 = \mathcal{O}(h^4)\|u\|_{5,\Omega}. \tag{6.18}
\]
Notice that both \( \tilde{u}_h \) and \( \tilde{u}_h^* \) are constant zero along \( \partial\Omega \), and \( u_h|_{\partial\Omega} = g_1 \) is the Lagrangian interpolation of \( g \) along \( \partial\Omega \). With (6.8), we have proven the following main result.

**Theorem 6.1** For a nonhomogeneous Dirichlet boundary problem (6.1), with suitable smoothness assumptions \( a_{ij}, b, c \in W^{4,\infty}(\Omega), u(x,y) \in H^3(\Omega) \) and \( f(x,y) \in H^4(\Omega) \), the numerical solution \( u_h \) by scheme (6.3) is a fourth order accurate approximation to \( u \) in the discrete 2-norm over all the \( 3 \times 3 \) Gauss-Lobatto points:
\[
\|u_h - u\|_{2, Z_0} = \mathcal{O}(h^4)(\|u\|_s + \|f\|_4).
\]

**7. Finite difference implementation**

In this section we present the finite difference implementation of the scheme (6.3) on a uniform mesh. The finite difference implementation of the nonhomogeneous Dirichlet boundary value problem is based on a homogeneous Neumann boundary value problem, which will be discussed first. We demonstrate how it is derived for the one-dimensional case then give the two-dimensional implementation. It provides efficient assembling of the stiffness matrix and one can easily implement it in MATLAB.
7.1 One-dimensional case

Consider a homogeneous Neumann boundary value problem $-(au')' = f$ on $[0, 1], u'(0) = 0, u'(1) = 0$, and its variational form is to seek $u \in H^1([0, 1])$ satisfying

$$(au', v') = (f, v), \quad \forall v \in H^1([0, 1]).$$

Let $I_k = [x_{2k}, x_{2k+2}]$ for $k = 0, \ldots, N-1$ as a finite element mesh for $P^2$ basis. Define

$$V_h = \{ v \in C^0([0, 1]) : v|_{I_k} \in P^2(I_k), k = 0, \ldots, N-1 \}.$$ 

Let $\{v_i\}_{i=0}^{n+1} \subset V_h$ be a basis of $V_h$ such that $v_i(x_j) = \delta_{ij}, i, j = 0, 1, \ldots, n+1$. With 3-point Gauss-Lobatto quadrature, the $C^0$-$P^2$ finite element method for (7.1) is to seek $u_h \in V_h$ satisfying

$$\langle au_h', v_i \rangle_h = \langle f, v_i \rangle_h, \quad i = 0, 1, \ldots, n+1. \tag{7.2}$$

Let $u_j = u_h(x_j), a_j = a(x_j)$ and $f_j = f(x_j)$ then $u_h(x) = \sum_{j=0}^{n+1} u_j v_j(x)$. We have

$$\sum_{j=0}^{n+1} u_j \langle a v_j', v_i \rangle_h = \langle au_h', v_i \rangle_h = \langle f, v_i \rangle_h = \sum_{j=0}^{n+1} f_j \langle v_j, v_i \rangle_h, \quad i = 0, 1, \ldots, n+1.$$ 

The matrix form of this scheme is $\tilde{S} \tilde{u} = \tilde{M} \tilde{f}$, where

$$\tilde{u} = [u_0, u_1, \ldots, u_n, u_{n+1}]^T, \quad \tilde{f} = [f_0, f_1, \ldots, f_n, f_{n+1}]^T,$$

the stiffness matrix $\tilde{S}$ is has size $(n+2) \times (n+2)$ with $(i, j)$-th entry as $\langle a v_j', v_i \rangle_h$, and the lumped mass matrix $M$ is a $(n+2) \times (n+2)$ diagonal matrix with diagonal entries $h \left( \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots, \frac{n+2}{n+2} \right)$.

Next we derive an explicit representation of the matrix $\tilde{S}$. Since basis functions $v_i \in V_h$ and $u_h(x)$ are not $C^1$ at the knots $x_{2k}$ ($k = 1, 2, \ldots, N-1$), their derivatives at the knots are double valued. We will use superscripts $+$ and $-$ to denote derivatives obtained from the right and from the left respectively, e.g., $v_{2k}^+$ and $v_{2k+2}^-$ denote the derivatives of $v_{2k}$ and $v_{2k+2}$ respectively in the interval $I_k = [x_{2k}, x_{2k+2}]$. Then in the interval $I_k = [x_{2k}, x_{2k+2}]$ we have the following representation of derivatives

$$
\begin{bmatrix}
  v_{2k}^+(x) \\
  v_{2k+2}^-(x) \\
  v_{2k+1}^+(x) \\
  v_{2k+2}^+(x)
\end{bmatrix} = \frac{1}{2h} \begin{bmatrix}
  -3 & 4 & -1 \\
  -1 & 0 & 1 \\
  1 & -4 & 3 \\
  1 & 0 & -1
\end{bmatrix} 
\begin{bmatrix}
  v_{2k}(x) \\
  v_{2k+1}(x) \\
  v_{2k+1}(x) \\
  v_{2k+2}(x)
\end{bmatrix}.
\tag{7.3}
$$

By abusing notations, we use $(v_i)_{2k}^+$ to denote the average of two derivatives of $v_i$ at the knots $x_{2k}$:

$$(v_i)_{2k}^+ = \frac{1}{2}[(v_i)_{2k}^- + (v_i)_{2k}^+].$$

Let $[v_i]$ denote the difference between the right derivative and left derivative:

$$[v_i]_0 = [v_i]_{n+2} = 0, \quad [v_i]_{2k} := (v_i)_{2k}^+ - (v_i)_{2k}^-, \quad k = 1, 2, \ldots, N-1.$$
Then at the knots, we have
\[
(v'_j)_{2k}(v'_j)_{2k} + (v'_j)_{2k}^2(v'_j)_{2k} = 2(v'_j)_{2k}(v'_j)_{2k} + \frac{1}{2}[v_i]_{2k}[v_j]_{2k}.
\] (7.4)

We also have
\[
\langle av'_j, v'_j \rangle_{I_{2k}} = h \left[ \frac{1}{3}a_{2k}(v'_j)_{2k} + \frac{4}{3}a_{2k+1}(v'_j)_{2k+1} + \frac{1}{3}a_{2k+2}(v'_j)_{2k+2} \right].
\] (7.5)

Let \(v_i\) denote a column vector of size \(n + 2\) consisting of grid point values of \(v_i(x)\). Plugging (7.4) into (7.5), with (7.3), we get
\[
\langle av'_j, v'_j \rangle_h = \sum_{k=0}^{N-1} \langle av'_j, v'_j \rangle_{I_{2k}} = \frac{1}{h} v'_j (D^T WAD + E^T WAE) v_j,
\]
where \(A\) is a diagonal matrix with diagonal entries \(a_0, a_1, \ldots, a_n, a_{n+1}\), and
\[
D = \frac{1}{2} \begin{pmatrix}
-\frac{3}{2} & 4 & -1 & & & & \\
\frac{1}{2} & -2 & 0 & 2 & -\frac{1}{2} & & \\
& -1 & 0 & 1 & & & \\
& & \ddots & \ddots & \ddots & & \\
& & & -\frac{1}{2} & 0 & 2 & -\frac{1}{2} & \\
& & & & -1 & 0 & 1 & \\
& & & & & \frac{1}{2} & -2 & 0 & 2 & -\frac{1}{2} & \\
& & & & & & -1 & 0 & 1 & 3
\end{pmatrix}_{(n+2) \times (n+2)}
\]
\[
E = \frac{1}{2} \begin{pmatrix}
0 & 0 & 0 & & & & \\
-\frac{1}{2} & -2 & 3 & -\frac{1}{2} & & & \\
& 0 & 0 & & & & \\
& & \ddots & \ddots & \ddots & & \\
& & & -\frac{1}{2} & -2 & 3 & -\frac{1}{2} & \\
& & & & 0 & 0 & 0 & \\
& & & & & \frac{1}{2} & -2 & 3 & -\frac{1}{2} & \\
& & & & & & 0 & 0 & 0 & 0
\end{pmatrix}_{(n+2) \times (n+2)}
\]

Since \(\{v_i\}_{i=0}^n\) are the Lagrangian basis for \(V^h\), we have
\[
\bar{S} = \frac{1}{h} (D^T WAD + E^T WAE).
\] (7.6)

Now consider the one-dimensional Dirichlet boundary value problem:
\[
-(au')' = f \text{ on } [0, 1],
\]
\[
u(0) = \sigma_1, \quad u(1) = \sigma_2.
\]

Consider the same mesh as above and define
\[
V^h_0 = \{ v \in C^0([0, 1]) : v|_{I_k} \in P^2(I_k), k = 0, \ldots, N - 1; v(0) = v(1) = 0 \}.
\]

Then \(\{v_i\}_{i=1}^n \subset V^h_0\) is a basis of \(V^h_0\) for \(\{v_i\}_{i=0}^{n+1}\) defined above. The one-dimensional version of (6.3) is to seek \(u_h \in V^h_0\) satisfying
\[
\langle au'_h, v'_i \rangle_h = \langle f, v'_i \rangle_h - \langle ag'_i, v'_i \rangle_h, \quad i = 1, 2, \ldots, n,
\]
\[
g_I(x) = \sigma_0 v_0(x) + \sigma_1 v_{n+1}(x).
\] (7.7)

Notice that we can obtain (7.7) by simply setting \(u_h(0) = \sigma_0\) and \(u_h(1) = \sigma_1\) in (7.2). So the finite difference implementation of (7.7) is given as follows:
1. Assemble the \((n+2) \times (n+2)\) stiffness matrix \(\mathcal{S}\) for homogeneous Neumann problem as in (7.6).

2. Let \(S\) denote the \(n \times n\) submatrix \(\mathcal{S}(2 : n+1, 2 : n+1)\), i.e., \(\mathcal{S}_{ij}\) for \(i, j = 2, \ldots, n+1\).

3. Let \(I\) denote the \(n \times 1\) submatrix \(\mathcal{S}(2 : n+1, 1)\) and \(r\) denote the \(n \times 1\) submatrix \(\mathcal{S}(2 : n+1, n+2)\), which correspond to \(v_0(x)\) and \(v_{n+1}(x)\).

4. Let \(u = [u_1 \ u_2 \ \cdots \ u_n]^T\) and \(f = [f_1 \ f_2 \ \cdots \ f_n]^T\). Define \(w = \left[\frac{1}{3}, \frac{2}{3}, \frac{1}{3}, \frac{2}{3}, \ldots, \frac{1}{3}, \frac{2}{3}\right]\) as a column vector of size \(n\). The scheme (7.7) can be implemented as

\[
\mathcal{S}u = hw^T f - \sigma_0 I - \sigma_1 r.
\]

7.2 Notations and tools for the two-dimensional case

We will need two operators:

- Kronecker product of two matrices: if \(A\) is \(m \times n\) and \(B\) is \(p \times q\), then \(A \otimes B\) is \(mp \times nq\) give by

\[
A \otimes B = \begin{pmatrix}
a_{11}B & \cdots & a_{1n}B \\
\vdots & \ddots & \vdots \\
a_{m1}B & \cdots & a_{mn}B
\end{pmatrix}.
\]

- For a \(m \times n\) matrix \(X\), \(\text{vec}(X)\) denotes the vectorization of the matrix \(X\) by rearranging \(X\) into a vector column by column.

The following properties will be used:

1. \((A \otimes B)(C \otimes D) = AC \otimes BD\).

2. \((A \otimes B)^{-1} = A^{-1} \otimes B^{-1}\).

3. \((B^T \otimes A)\text{vec}(X) = \text{vec}(AXB)\).

4. \((A \otimes B)^T = A^T \otimes B^T\).

Consider a uniform grid \((x_i, y_j)\) for a rectangular domain \(\Omega = [0, 1] \times [0, 1]\) where \(x_i = ih_x, \ i = 0, 1, \ldots, n_x + 1, \ h_x = \frac{1}{n_x+1}\) and \(y_j = jh_y, \ j = 0, 1, \ldots, n_y + 1, \ h_y = \frac{1}{n_y+1}\).

Assume \(n_x\) and \(n_y\) are odd and let \(N_x = \frac{n_x + 1}{2}\) and \(N_y = \frac{n_y + 1}{2}\). We consider rectangular cells \(e_{kl} = [x_{2k}, x_{2k+2}] \times [y_{2l}, y_{2l+2}]\) for \(k = 0, \ldots, N_x - 1\) and \(l = 0, \ldots, N_y - 1\) as a finite element mesh for \(Q^2\) basis. Define

\[
\mathcal{V}_h = \{ v \in C^0(\Omega) : v|_{e_{kl}} \in Q^2(e_{kl}), k = 0, \ldots, N_x - 1, l = 0, \ldots, N_y - 1 \},
\]

\[
\mathcal{V}_0^h = \{ v \in C^0(\Omega) : v|_{e_{kl}} \in Q^2(e_{kl}), k = 0, \ldots, N_x - 1, l = 0, \ldots, N_y - 1; v|_{\partial \Omega} \equiv 0 \}.
\]

For the coefficients \(a(x, y) = \begin{pmatrix} a^{11} & a^{12} \\ a^{21} & a^{22} \end{pmatrix}, \ b = [b^1 \ b^2]\) and \(c\) in the elliptic operator (2.2), consider their grid point values in the following form:

\[
A_{kl} = \begin{pmatrix}
a_{00} & a_{01} & \cdots & a_{0,n_x+1} \\
a_{10} & a_{11} & \cdots & a_{1,n_x+1} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n_y+1,0} & a_{n_y+1,1} & \cdots & a_{n_y+1,n_x+1}
\end{pmatrix}_{(n_y+2) \times (n_x+2)}, \ a_{ij} = a_{kl}(x_j, y_l), \ k, l = 1, 2, \ldots.
and its matrix representation is 

\[
\mathbf{B}^m = \begin{pmatrix}
    b_{00} & b_{01} & \cdots & b_{0,n+1} \\
    b_{10} & b_{11} & \cdots & b_{1,n+1} \\
    \vdots & \vdots & & \vdots \\
    b_{n+1,0} & b_{n+1,1} & \cdots & b_{n+1,n+1}
\end{pmatrix}, \quad b_{ij} = B^m(x_j,y_i), \quad m = 1, 2,
\]

Its adjoint is a restriction operator 

\[
\text{Res} : \mathbb{R}^{(n+2)\times(n+2)} \rightarrow \mathbb{R}^{n_x\times n_x}
\]

Define an inflation operator

\[
\text{diag}(x) = \begin{pmatrix}
    x_1 & 0 & \cdots & 0 \\
    0 & x_2 & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & \cdots & x_{n+1}
\end{pmatrix}, \quad c_{ij} = c(x_j,y_i).
\]

Let \(\text{diag}(x)\) denote a diagonal matrix with the vector \(x\) as diagonal entries and define

\[
\tilde{W}_x = \text{diag} \left( \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \ldots, \frac{1}{3}, \frac{1}{3} \right)_{(n_x+2)\times(n_x+2)},
\]

\[
\tilde{W}_y = \text{diag} \left( \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \ldots, \frac{1}{3}, \frac{1}{3} \right)_{(n_y+2)\times(n_y+2)},
\]

\[
\tilde{W}_x = \text{diag} \left( \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \ldots, \frac{1}{3}, \frac{1}{3} \right)_{n_x\times n_x}, \quad \tilde{W}_y = \text{diag} \left( \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \ldots, \frac{1}{3}, \frac{1}{3} \right)_{n_y\times n_y}.
\]

Let \(s = x\) or \(y\), we define the \(D\) and \(E\) matrices with dimension \((n_s+2)\times(n_s+2)\) for each variable:

\[
D_s = \frac{1}{2} \begin{pmatrix}
    -3 & 4 & -1 \\
    -1 & 0 & 1 \\
    \frac{1}{2} & 2 & -\frac{1}{2} \\
    -1 & 0 & 1 \\
    \frac{1}{2} & 2 & -\frac{1}{2} \\
    \vdots & \vdots & \vdots \\
    -1 & 0 & 1 \\
    \frac{1}{2} & 2 & -\frac{1}{2} \\
    -1 & 0 & 1
\end{pmatrix}, \quad E_s = \frac{1}{2} \begin{pmatrix}
    0 & 0 & 0 \\
    0 & 2 & -3 \frac{1}{2} \\
    -\frac{1}{2} & 2 & -3 \frac{1}{2} \\
    0 & 0 & 0 \\
    -\frac{1}{2} & 2 & -3 \frac{1}{2} \\
    \vdots & \vdots & \vdots \\
    0 & 0 & 0 \\
    -\frac{1}{2} & 2 & -3 \frac{1}{2} \\
    0 & 0 & 0
\end{pmatrix}.
\]

Define an inflation operator

\[
\text{Inf}(U) = \begin{pmatrix}
    0 & \cdots & 0 \\
    \vdots & U & \vdots \\
    0 & \cdots & 0
\end{pmatrix}_{(n_y+2)\times(n_x+2)}
\]

and its matrix representation is given as \(\tilde{I}_x \otimes \tilde{I}_y\) where

\[
\tilde{I}_x = \begin{pmatrix}
    0 \\
    0 \\
    \vdots \\
    0
\end{pmatrix}_{n_x\times n_x}, \quad \tilde{I}_y = \begin{pmatrix}
    0 \\
    0 \\
    \vdots \\
    0
\end{pmatrix}_{n_y\times n_y}.
\]

Its adjoint is a restriction operator

\[
\text{Res} : \mathbb{R}^{(n_x+2)\times(n_y+2)} \rightarrow \mathbb{R}^{n_x\times n_x}
\]

and its matrix representation is \(\tilde{I}_x^T \otimes \tilde{I}_y^T\).
7.3 Two-dimensional case

For \( \tilde{Q} = [0, 1]^2 \) we first consider an elliptic equation with homogeneous Neumann boundary condition:

\[
-\nabla \cdot (a\nabla u) + b\nabla u + cu = f \quad \text{on } \Omega, \\
\quad a\nabla u \cdot n = 0 \quad \text{on } \partial \Omega. 
\]

The variational form is to find \( u \in H^1(\Omega) \) satisfying

\[
A(u, v) = (f, v), \quad \forall v \in H^1(\Omega). 
\]

The \( C^0-Q^2 \) finite element method with \( 3 \times 3 \) Gauss-Lobatto quadrature is to find \( u_h \in V^h \) satisfying

\[
\langle a\nabla u_h, \nabla v \rangle_h + \langle b\nabla u_h, v \rangle_h + \langle cu_h, v \rangle_h = \langle f, v \rangle_h, \quad \forall v \in V^h, 
\]

Let \( \tilde{U} \) be a \( (n_x + 2) \times (n_y + 2) \) matrix such that its \( (j,i) \)-th entry is \( \tilde{U}(j,i) = u_h(x_{i-1}, y_{j-1}), \) \( i = 1, \ldots, n_x + 2, j = 1, \ldots, n_y + 2. \) Let \( \tilde{F} \) be a \( (n_x + 2) \times (n_y + 2) \) matrix such that its \( (j,i) \)-th entry is \( \tilde{F}(j,i) = f(x_{i-1}, y_{j-1}). \) Then the matrix form of (7.11) is

\[
\bar{S}\text{vec}(\tilde{U}) = \bar{M}\text{vec}(\tilde{F}), \quad \bar{M} = h_x h_y \bar{W}_x \otimes \bar{W}_y, \quad \bar{S} = \sum_{k,l=1}^{2} S_{kl}^{11} + \sum_{m=1}^{2} S_{m}^{11} + S_c, 
\]

where

\[
S_{a}^{11} = \frac{h_y}{h_x} (D_y^T \otimes I_x) \text{diag}(\text{vec}(\bar{W}_x A^{11} \bar{W}_x))(D_x \otimes I_y) + \frac{h_x}{h_y} (E_x^T \otimes I_y) \text{diag}(\text{vec}(\bar{W}_y A^{11} \bar{W}_y))(E_x \otimes I_y), \\
S_{b}^{11} = (D_y^T \otimes I_x) \text{diag}(\text{vec}(\bar{W}_x A^{12} \bar{W}_x))(I_x \otimes D_y) + (E_x^T \otimes I_y) \text{diag}(\text{vec}(\bar{W}_y A^{12} \bar{W}_y))(I_x \otimes E_y), \\
S_{a}^{21} = (I_x \otimes D_y^T) \text{diag}(\text{vec}(\bar{W}_x A^{21} \bar{W}_x))(D_x \otimes I_y) + (I_x \otimes E_y^T) \text{diag}(\text{vec}(\bar{W}_y A^{21} \bar{W}_y))(E_x \otimes I_y), \\
S_{a}^{22} = \frac{h_x}{h_y} (I_x \otimes D_y^T) \text{diag}(\text{vec}(\bar{W}_x A^{22} \bar{W}_x))(I_x \otimes D_y) + \frac{h_y}{h_x} (I_x \otimes E_y^T) \text{diag}(\text{vec}(\bar{W}_y A^{22} \bar{W}_y))(E_x \otimes I_y), \\
S_{b}^{12} = h_y \text{diag}(\text{vec}(\bar{W}_y B^1 \bar{W}_x))(D_x \otimes I_y), \quad S_{b}^{22} = h_x \text{diag}(\text{vec}(\bar{W}_x B^2 \bar{W}_x))(I_x \otimes D_y), \quad S_c = h_x h_y \text{diag}(\text{vec}(\bar{W}_x C \bar{W}_x)).
\]

Now consider the scheme (6.3) for nonhomogeneous Dirichlet boundary conditions. Its numerical solution can be represented as a matrix \( U \) of size \( ny \times nx \) with \( (j,i) \)-entry \( U(j,i) = u_h(x_i, y_j) \) for \( i = 1, \ldots, nx; j = 1, \ldots, ny. \) Similar to the one-dimensional case, its stiffness matrix can be obtained as the submatrix of \( \tilde{S} \) in (7.12). Let \( \tilde{G} \) be a \( (n_y + 2) \times (n_x + 2) \) matrix with \( (j,i) \)-th entry as \( \tilde{G}(j,i) = g(x_{i-1}, y_{j-1}) \), where

\[
g(x,y) = \begin{cases} 
0, & \text{if } (x,y) \in (0,1) \times (0,1), \\
g(x,y), & \text{if } (x,y) \in \partial \Omega.
\end{cases}
\]

In particular, \( \tilde{G}(j+1,i+1) = 0 \) for \( j = 1, \ldots, ny; i = 1, \ldots, nx. \) Let \( F \) be a matrix of size \( ny \times nx \) with \( (j,i) \)-entry as \( F(j,i) = f(x_i, y_j) \) for \( i = 1, \ldots, nx; j = 1, \ldots, ny. \) Then the scheme (6.3) becomes

\[
(\bar{I}_x \otimes \bar{I}_y)\tilde{S}(I_x \otimes I_y)\text{vec}(U) = (W_x \otimes W_y)\text{vec}(F) - (\bar{I}_x \otimes \bar{I}_y)\bar{S}\text{vec}(\tilde{G}).
\]
Even though the stiffness matrix is given as $S = (\tilde{T}_x \otimes \tilde{T}_y) S(I_x \otimes I_y)$, $S$ should be implemented as a linear operator in iterative linear system solvers. For example, the matrix vector multiplication $(\tilde{T}_x \otimes \tilde{T}_y) S(I_x \otimes I_y) vec(U)$ is equivalent to the following linear operator from $\mathbb{R}^{ny \times nx}$ to $\mathbb{R}^{ny \times nx}$:

$$
\frac{h}{h_x} \tilde{T}_x \{ I_y ( [\tilde{W}_y A^{11} \tilde{W}_y] \circ [I_y (\tilde{T}_x I_y) D_x^T]) D_x + I_y ([\tilde{W}_y A^{11} \tilde{W}_y] \circ [I_y (\tilde{T}_x I_y) E_x^T]) E_x \} \tilde{I}_x,
$$

where $\circ$ is the Hadamard product (i.e., entrywise multiplication).

7.4 The Laplacian case

For one-dimensional constant coefficient case with homogeneous Dirichlet boundary condition, the scheme can be written as a classical finite difference scheme $Hu = f$ with

$$
H = M^{-1} S = \frac{1}{h^2} \begin{pmatrix}
2 & -1 \\
-2 & 2 & -1 \\
& -2 & 2 & -1 \\
& & -2 & 2 & -1 \\
& & & -2 & 2 & -1 \\
& & & & -2 & 2 & -1 \\
& & & & & \ddots & \ddots & \ddots \\
& & & & & & -2 & 2 & -1 \\
& & & & & & & -2 & 2 & -1 \\
& & & & & & & & \ddots & \ddots & \ddots \\
& & & & & & & & & -2 & 2 & -1 \\
& & & & & & & & & & -2 & 2 & -1 \\
\end{pmatrix}
$$

In other words, if $x_i$ is a cell center, the scheme is

$$
\frac{-u_{i-1} + 2u_i - u_{i+1}}{h^2} = f_i,
$$

and if $x_i$ is a knot away from the boundary, the scheme is

$$
\frac{u_{i-2} - 8u_{i-1} + 14u_i - 8u_{i+1} + u_{i+2}}{4h^2} = f_i.
$$

It is straightforward to verify that the local truncation error is only second order.

For the two-dimensional Laplacian case homogeneous Dirichlet boundary condition, the scheme can be rewritten as

$$(H_x \otimes I_y) + (I_x \otimes H_y) vec(U) = vec(F),$$

where $H_x$ and $H_y$ are the same $H$ matrix above with size $n_x \times n_x$ and $n_y \times n_y$ respectively. The inverse of $(H_x \otimes I_y) + (I_x \otimes H_y)$ can be efficiently constructed via the eigen-decomposition of small matrices $H_x$ and $H_y$:

1. Compute eigen-decomposition of $H_x = T_x A_x T_x^{-1}$ and $H_y = T_y A_y T_y^{-1}$.

2. The properties of Kronecker product imply that

$$(H_x \otimes I_y) + (I_x \otimes H_y) = (T_x \otimes T_y)(A_x \otimes I_y + I_x \otimes A_y)(T_x^{-1} \otimes T_y^{-1}),$$

thus

$$
[(H_x \otimes I_y) + (I_x \otimes H_y)]^{-1} = (T_x \otimes T_y)(A_x \otimes I_y + I_x \otimes A_y)^{-1}(T_x^{-1} \otimes T_y^{-1}).
$$
3. It is nontrivial to determine whether $H$ is diagonalizable. In all our numerical tests, $H$ has no repeated eigenvalues. So if assuming $\Lambda_x$ and $\Lambda_y$ are diagonal matrices, the matrix vector multiplication $[(H_x \otimes I_y) + (I_x \otimes H_y)]^{-1} \text{vec}(F)$ can be implemented as a linear operator on $F$:

$$T_y([T_y^{-1}F(T_y^{-1})^T]/\Lambda)T_x^T,$$

where $\Lambda$ is a $n_x \times n_x$ matrix with $(i,j)$-th entry as $\Lambda(i,j) = \Lambda_x(i,i) + \Lambda_y(j,j)$ and $/\Lambda$ denotes entry-wise division for two matrices of the same size.

For the 3D Laplacian, the matrix can be represented as $H_x \otimes I_y \otimes I_z + I_x \otimes H_y \otimes I_z + I_x \otimes I_y \otimes H_z$ thus can be efficiently inverted through eigen-decomposition of small matrices $H_x$, $H_y$ and $H_z$ as well.

Since the eigen-decomposition of small matrices $H_x$ and $H_y$ can be precomputed, and (7.14) costs only $O(n^3)$ for a 2D problem on a mesh size $n \times n$, in practice (7.14) can be used as a simple preconditioner in conjugate gradient solvers for the following linear system equivalent to (7.13):

$$(W_x^{-1} \otimes W_y^{-1})(\tilde{T}_x \otimes \tilde{T}_y)\tilde{S}(I_x \otimes I_y)\text{vec}(U) = \text{vec}(F) - (W_x^{-1} \otimes W_y^{-1})(\tilde{T}_x \otimes \tilde{T}_y)\tilde{S}\text{vec}(G),$$

even though the multigrid method as reviewed in Xu & Zikatanov (2017) is the optimal solver in terms of computational complexity.

8. Numerical results

In this section we show a few numerical tests verifying the accuracy of the scheme (6.3) implemented as a finite difference scheme on a uniform grid. We first consider the following two dimensional elliptic equation:

$$-\nabla \cdot (a \nabla u) + cu = f \text{ on } [0,1] \times [0,2]$$

where $a = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $a_{11} = 10 + 30y^5 + x \cos y + y$, $a_{12} = a_{21} = 2 + 0.5(\sin(\pi x) + x^3)(\sin(\pi y) + y^3) + \cos(x^4 + y^3)$, $a_{22} = 10 + x^5 + 1 + x^4y^3$, with an exact solution

$$u(x,y) = 0.1(\sin(\pi x) + x^3)(\sin(\pi y) + y^3) + \cos(x^4 + y^3).$$

The errors at grid points are listed in Table 8 for purely Dirichlet boundary condition and Table 8 for purely Neumann boundary condition. We observe fourth order accuracy in the discrete 2-norm for both tests, even though only $O(h^{3.5})$ can be proven for Neumann boundary condition as discussed in Remark 5.2. Regarding the maximum norm of the superconvergence of the function values at Gauss-Lobatto points, one can only prove $O(h^{3.5}\log h)$ even for the full finite element scheme (1.1) since discrete Green’s function is used, see Chen (2001).

Next we consider a three-dimensional problem $-\Delta u = f$ with homogeneous Dirichlet boundary conditions on a cube $[0,1]^3$ with the following exact solution

$$u(x,y,z) = \sin(\pi x)\sin(2\pi y)\sin(3\pi z) + (x-x^3)(y^2-y^4)(z-z^2).$$

See Table 8 for the performance of the finite difference scheme. There is no essential difficulty to extend the proof to three dimensions, even though it is not very straightforward. Nonetheless we observe that the scheme is indeed fourth order accurate. The linear system is solved by the eigenvector method shown in Section 7.4. The discrete 2-norm over the set of all grid points $Z_0$ is defined as $\|u\|_{2,Z_0} = [h^3\sum_{(x,y,z) \in Z_0} |u(x,y,z)|^2]^{1/2}$. 

FOURTH ORDER ACCURACY OF THE FD IMPLEMENTATION OF $C^0$-$Q^2$ FEM

37 of 40
Table 1. A 2D elliptic equation with Dirichlet boundary conditions. The first column is the number of regular cells in a finite element mesh. The second column is the number of grid points in a finite difference implementation, i.e., number of degree of freedoms.

| FEM Mesh | FD Grid | $l^2$ error | order | $l^\infty$ error | order |
|----------|---------|-------------|-------|------------------|-------|
| 2 × 4    | 3 × 7   | 3.94E-2     | -     | 7.15E-2          | -     |
| 4 × 8    | 7 × 15  | 1.23E-2     | 1.67  | 3.28E-2          | 1.12  |
| 8 × 16   | 15 × 31 | 1.46E-3     | 3.08  | 5.42E-3          | 2.60  |
| 16 × 32  | 31 × 63 | 1.14E-4     | 3.68  | 3.96E-4          | 3.78  |
| 32 × 64  | 63 × 127| 7.75E-5     | 3.08  | 2.62E-5          | 3.92  |
| 64 × 128 | 127 × 255| 1.23E-2   | 3.95  | 1.73E-6          | 3.92  |
| 128 × 256| 255 × 511| 3.23E-8   | 3.96  | 1.13E-7          | 3.94  |

Table 2. A 2D elliptic equation with Neumann boundary conditions.

| FEM Mesh | FD Grid | $l^2$ error | order | $l^\infty$ error | order |
|----------|---------|-------------|-------|------------------|-------|
| 2 × 4    | 5 × 9   | 1.38E0      | -     | 2.27E0           | -     |
| 4 × 8    | 9 × 17  | 1.46E-1     | 3.24  | 2.52E-1          | 3.17  |
| 8 × 16   | 17 × 33 | 7.49E-3     | 4.28  | 1.64E-2          | 3.94  |
| 16 × 32  | 33 × 65 | 4.31E-4     | 4.12  | 1.02E-3          | 4.01  |
| 32 × 64  | 65 × 129| 2.61E-5     | 4.04  | 7.47E-5          | 3.78  |

Last we consider a two dimensional elliptic equation with convection term and the coefficients $b$ is incompressible $\nabla \cdot b = 0$:

$$-\nabla \cdot (a \nabla u) + b \cdot \nabla u + cu = f \quad \text{on } [0, 1] \times [0, 2]$$

where $a = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $a_{11} = 100 + 30y^5 + x\cos y + y$, $a_{12} = a_{21} = 2 + 0.5(\sin(\pi x) + x^3)(\sin(\pi y) + y^3) + \cos(x^4 + y^3)$, $a_{22} = 100 + x^5$, $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$, $b_1 = \psi y$, $b_2 = -\psi x$, $\psi = \exp(x^2 + y)$, $c = 1 + x^4 y^3$, with an exact solution

$$u(x, y) = 0.1(\sin(\pi x) + x^3)(\sin(\pi y) + y^3) + \cos(x^4 + y^3).$$

The errors at grid points are listed in Table 4 for Dirichlet boundary conditions.

9. Concluding remarks

In this paper we have proven the superconvergence of function values in the simplest finite difference implementation of $C^0, Q^2$ finite element method for elliptic equations. In particular, the scheme (6.3) can be easily implemented as a fourth order accurate finite difference scheme as shown in Section 7. It provides only only an convenient approach for constructing fourth order accurate finite difference schemes but also an efficient implementation of $C^0, Q^2$ finite element method without losing superconvergence of function values. In a follow up paper Li & Zhang (2019a), we will show that discrete maximum principle can be proven for the scheme (6.3) solving a variable coefficient Poisson equation.
Table 3. $-\Delta u = f$ in 3D with homogeneous Dirichlet boundary condition.

| Finite Difference Grid | $l^2$ error | order | $l^\infty$ error | order |
|------------------------|-------------|-------|-----------------|-------|
| $7 \times 7 \times 7$  | 1.51E-2     | -     | 4.87E-2         | -     |
| $15 \times 15 \times 15$ | 9.23E-4     | 4.04  | 3.12E-3         | 3.96  |
| $31 \times 31 \times 31$ | 5.68E-5     | 4.02  | 1.95E-4         | 4.00  |
| $63 \times 63 \times 63$ | 3.54E-6     | 4.01  | 1.22E-5         | 4.00  |
| $127 \times 127 \times 127$ | 2.21E-7    | 4.00  | 7.59E-7         | 4.00  |

Table 4. A 2D elliptic equation with convection term and Dirichlet boundary conditions.

| FEM Mesh | FD Grid | $l^2$ error | order | $l^\infty$ error | order |
|----------|---------|-------------|-------|-----------------|-------|
| $2 \times 4$ | $3 \times 7$ | 1.26E-1 | -     | 2.71E-1        | -     |
| $4 \times 8$ | $7 \times 15$ | 2.85E-2 | 2.15  | 9.70E-2        | 1.48  |
| $8 \times 16$ | $15 \times 31$ | 1.89E-3 | 3.92  | 7.23E-3        | 3.74  |
| $16 \times 32$ | $31 \times 63$ | 1.17E-4 | 4.01  | 4.01E-4        | 4.17  |
| $32 \times 64$ | $63 \times 127$ | 7.41E-6 | 3.98  | 2.54E-5        | 3.98  |

Acknowledgments

H. Li and X. Zhang were supported by the NSF grant DMS-1522593.

References

BAKKER, M. (1982) A note on $C^0$ Galerkin methods for two-point boundary problems. *Numerische Mathematik*, 38, 447–453.

CHEN, C. (1979) Superconvergent points of Galerkin’s method for two point boundary value problems. *Numerical Mathematics A Journal of Chinese Universities*, 1, 73–79.

CHEN, C. (1981) Superconvergence of finite element solutions and their derivatives. *Numerical Mathematics A Journal of Chinese Universities*, 3, 118–125.

CHEN, C. (2001) *Structure theory of superconvergence of finite elements (In Chinese)*. Hunan Science and Technology Press, Changsha.

CIARLET, P. G. (1991) Basic error estimates for elliptic problems. *Handbook of Numerical Analysis*, 2, 17–351.

CIARLET, P. G. (2002) *The Finite Element Method for Elliptic Problems*. Society for Industrial and Applied Mathematics.

CIARLET, P. G. & RAVIART, P.-A. (1972) 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*. Elsevier, pp. 409–474.

DOUGLAS, J., DUPONT, T. & WHEELER, M. F. (1974) An $L^\infty$ estimate and a superconvergence result for a Galerkin method for elliptic equations based on tensor products of piecewise polynomials.

GRISVARD, P. (2011) *Elliptic problems in nonsmooth domains*, vol. 69. SIAM.
REFERENCES

HUANG, Y. & XU, J. (2008) Superconvergence of quadratic finite elements on mildly structured grids. Mathematics of computation, 77, 1253–1268.

LESAIN, P. & ZLAMAL, M. (1979) Superconvergence of the gradient of finite element solutions. RAIRO. Analyse numérique, 13, 139–166.

LI, H. & ZHANG, X. (2019a) On the monotonicity and discrete maximum principle of the finite difference implementation of $C^0$-$Q^2$ finite element method. in preparation.

LI, H. & ZHANG, X. (2019b) Superconvergence of $C^0$-$Q^k$ finite element method for elliptic equations with approximated coefficients. arXiv preprint arXiv:1902.00945.

LIN, Q., YAN, N. & ZHOU, A. (1991) A rectangle test for interpolated finite elements. Proc. Sys. Sci. and Sys. Eng.(Hong Kong), Great Wall Culture Publ. Co. Proc. Sys. Sci. and Sys. Eng.(Hong Kong), Great Wall Culture Publ. Co., pp. 217–229.

LIN, Q. & YAN, N. (1996) Construction and Analysis for Efficient Finite Element Method (In Chinese). Hebei University Press.

SAVARÉ, G. (1998) Regularity results for elliptic equations in lipschitz domains. Journal of Functional Analysis, 152, 176–201.

WAHLBIN, L. (2006) Superconvergence in Galerkin finite element methods. Springer.

WHITEMAN, J. (1975) Lagrangian finite element and finite difference methods for poisson problems. Numerische Behandlung von Differentialgleichungen. Springer, pp. 331–355.

XU, J. & ZIKATANOV, L. (2017) Algebraic multigrid methods. Acta Numerica, 26, 591–721.