Quest Journals
Journal of Research in Applied Mathematics
Volume 11 ~ Issue 3 (2025) pp: 94-104
ISSN (Online): 2394-0743 ISSN (Print): 2394-0735
www.questjournals.org

Review Paper

Quadratic Element Method of the Mixed Finite Element Method for Second-Order Elliptic Eigenvalue Problems

Qiuxia Tian, Xiaomin Cai, Shan Nie
(School of Mathematical Sciences, Guizhou Normal University)
Corresponding Author: Qiuxia Tian

ABSTRACT: *The general second-order elliptic eigenvalue problem is of great significance and is closely related to various fields such as fluid mechanics, quantum mechanics, and structural engineering analysis. Classical finite element methods have been successfully applied to solve such problems, but in some cases, particularly when dealing with complex boundary conditions and non-homogeneous media, the efficiency and accuracy of traditional methods may not meet the requirements. To improve both the solution accuracy, the mixed finite element method has been proposed and has achieved significant results in solving second order eigenvalue problems. The mixed finite element method introduces auxiliary variables (which generally also have practical physical significance), allowing for a reduction in the order of high-order differential equations, thereby relaxing the smoothness requirements of the finite element space. This article investigates the mixed finite element method for a general second-order elliptic eigenvalue problem and provides a-priori error estimates.*

KEYWORDS: *Second-Order eigenvalue problems, Mixed finite element method, A-priori error.*

R*eceived 09 Mar., 2025;* ***R****evised 21 Mar., 2025;* ***A****ccepted 23 Mar., 2025 © The author(s) 2025.*
P*ublished with open access at* ***www.questjournas.org***

I. INTRODUCTION

The second-order eigenvalue problem is widely applied in various fields such as vibration analysis, material mechanics, acoustics, and quantum mechanics. In the classical finite element method, eigenvalue problems are typically solved by discretizing the differential operators being approximated. However, these methods often rely on the direct solution of higher-order differential equations, making the treatment of boundary conditions, material inhomogeneity, and complex geometries relatively challenging. Especially for high-frequency problems with irregular boundaries, traditional methods may struggle to meet the demands of both computational efficiency and accuracy.

As an advanced branch of the finite element method, the mixed finite element method was initially established by Babuška and Brezzi in the early 1970s, who developed the general theory of the method [1, 2]. In the early 1980s, Falk and Osborn proposed an improved version of the method [3]. [4,5,6,7]provides extensive research on mixed problems, presenting numerous mixed finite element formulations, and further investigates the theoretical development and practical applications of the mixed finite element method.

There are several works for second order elliptic eigenvalue problems by the mixed formulation and their numerical methods such as Babuška and Osborn [8,9], Mercier, Osborn, Rappaz, and Raviart [11], etc. Based on the general theory of compact operators [10], Osborn [12], Mercier, Osborn, Rappaz, and Raviart [11] give abstract analysis for the eigenpair approximations by mixed/hybrid finite element methods. [13]discusses the $L^2(\Omega)$ norm and L^∞ norm estimates of eigenvalues and eigenfunctions for a more general class of eigenvalue problems. [14] propose a method to improve the convergence rate of the lowest order Raviart–Thomas mixed finite element approximations for the second order elliptic eigenvalue problem. [15] proposes a non-standard mixed finite element method for the Dirichlet boundary value problem of second-order elliptic equations.

This paper draws on the ideas from the aforementioned literature to study the a priori error estimation of the mixed finite element method combined with the quadratic finite element method for solving the general second-order elliptic eigenvalue problem. The mixed finite element method, combined with the quadratic method, is an effective numerical approach for solving second-order elliptic eigenvalue problems. By enhancing the accuracy and stability of the solution, it can be widely applied to elliptic equation problems with complex boundaries and physical properties, particularly in cases involving high gradients and curvatures, where it demonstrates its advantages. This method has significant practical value in engineering mechanics, physics, and other disciplines that require the solution of partial differential equations.

In the entire paper, denotes a general constant that is independent of the mesh size but sometimes depend on the eigenvalues of the problem (1).

II. BASIC THEORETICAL PREPARATION

Let $L^q(\Omega)$ be a standard Lebesgue space, where $1 \le q \le \infty, \Omega \subset R^2$,The corresponding norm is expressed by $\|\cdot\|_{L^q(\Omega)}$. In this paper, the norm of $L^q(\Omega)$ is represented by $\|\cdot\|_{\Omega}$.We also use $H^s(\Omega)$ to express the standard Hilbert Sobolev space of real functions defined at $\Omega \subset R^2$ with index $s \ge 0$ and the corresponding norm and semi-norm are $\|\cdot\|_{s,\Omega}$ and $|\cdot|_{s,\Omega}$ .Let Ω be the bounded open polygon region of R^2 ,and let $\partial\Omega$ represent its boundary. In this paper, we are concerned with the following second order elliptic eigenvalue problem:

$$\begin{cases} -\nabla \cdot (K(x,y)\nabla p) = \lambda p, & \text{in } \Omega, \\ p = 0, & \text{on } \partial\Omega, \end{cases} \tag{1}$$

where $K = (a_{ij})_{2\times 2}$ is a symmetric positive definite matrix with $a_{ij} \in W^{1,\infty}(\Omega)$ for $1 \le i, j \le 2$, $K^{-1} = (a_{ij})_{2\times 2}$ is also a symmetric positive definite matrix, $\Omega \subset R^2$ is a bounded domain with Lipschitz boundary $\partial\Omega$, ∇ and $\nabla\cdot$ denote the gradient and divergence operators.

III. MIXED FINITE ELEMENT METHOD

We define a new vector-valued function $\boldsymbol{\mu} = K\nabla p$.

Then (1) can be transformed into the following equivalent formulation

$$\begin{cases} K^{-1}\boldsymbol{\mu} - \nabla p = 0, & \text{in } \Omega, \\ -\nabla \cdot \boldsymbol{\mu} = \lambda p, & \text{on } \partial\Omega, \\ p = 0, & \text{on } \partial\Omega. \end{cases} \tag{2}$$

Next, define the spaces

$$W = L^2(\Omega),\ G = L^2(\Omega),\ \mathbf{H} = [L^2(\Omega)]^2,$$

$$\mathbf{V} = H(div,\Omega) = \left\{ \boldsymbol{\mu} \in [L^2(\Omega)]^2 : \nabla \cdot \boldsymbol{\mu} \in L^2(\Omega) \right\},$$

equipped with the norm

$$\|\boldsymbol{\mu}\|^2_{H(div,\Omega)} = (\|\boldsymbol{\mu}\|_0^2 + \|\nabla \cdot \boldsymbol{\mu}\|_0^2).$$

Then, the weak form for the problem (1) can be defined as follows:

Find $(\lambda, \boldsymbol{\mu}, p) \in R \times \mathbf{V} \times W$, $(\boldsymbol{\mu}, p) \ne (0,0)$, such that

$$\begin{cases} a(\boldsymbol{\mu}, \boldsymbol{\varphi}) - b(\boldsymbol{\varphi}, p) = 0, & \forall \boldsymbol{\varphi} \in \mathbf{V}, \\ b(\boldsymbol{\mu}, v) = \lambda(p, v), & \forall v \in W, \end{cases} \tag{3}$$

where $a(\cdot,\cdot)$, $b(\cdot,\cdot)$ are bilinear forms defined by

$$a(\boldsymbol{\mu}, \boldsymbol{\varphi}) = \int_\Omega \boldsymbol{\mu} \cdot K^{-1}\boldsymbol{\varphi}\, dx, \quad b(\boldsymbol{\varphi}, p) = -\int_\Omega div\boldsymbol{\varphi} \cdot p\, dx, \quad (p, v) = \int_\Omega pv\, dx.$$

Clearly, the bilinear forms $a(\cdot,\cdot)$ is symmetric and the bilinear forms defined above have the following characteristics:

$$|a(\boldsymbol{\mu}, \boldsymbol{\mu})| \le C_0 \|\boldsymbol{\mu}\|_{\mathbf{H}} \|\boldsymbol{\mu}\|_{\mathbf{H}}, \tag{4}$$

$$|a(\boldsymbol{\mu}, \boldsymbol{\varphi})| \le C_1 \|\boldsymbol{\mu}\|_{\mathbf{H}} \|\boldsymbol{\varphi}\|_{\mathbf{H}}, \tag{5}$$

$$|b(\boldsymbol{\varphi}, p)| \le C_2 \|\boldsymbol{\varphi}\|_{\mathbf{V}} \|p\|_W, \tag{6}$$

where $C_i (i = 0,1,2)$ represents a constant independent of h .

For the eigenvalue λ , there exists the following Rayleigh quotient expression

$$\lambda = \frac{-a(\boldsymbol{\mu}, \boldsymbol{\mu}) + 2b(\boldsymbol{\mu}, p)}{(p, p)}.$$

Form [9], we know eigenvalue problem (3) has an eigenvalue sequence $\{\lambda_j\}$:

$$0 \le \lambda_1 \le \lambda_2 \le \cdots \le \lambda_k \le \cdots, \lim_{k\to\infty} \lambda_k = \infty,$$

and the associated eigenfunctions

$$(\mathbf{\mu}_1, p_1), (\mathbf{\mu}_2, p_2), \cdots (\mathbf{\mu}_k, p_k) \cdots,$$

where $(p_i, p_j) = \delta_{ij}$.

Theorem 1. Let (λ, p) be an eigenpair of equation (1), $\mathbf{\mu} = K\nabla p$, then $(\lambda, \mathbf{\mu}, p)$ satisfies equation (3); if $(\lambda, \mathbf{\mu}, p)$ satisfies equation (3), then (λ, p) is an eigenpair of equation (1), and $\mathbf{\mu} = K\nabla p$.

Proof. From the derivation above, the first part of the theorem has been established. Now we will prove the second part of the theorem.

Let $(\lambda, \mathbf{\mu}, p)$ satisfy equation (3), and consider the auxiliary problem

$$\begin{cases} -\nabla \cdot (K(x,y)\nabla \tilde{p}) = \lambda p, & \text{in } \Omega, \\ \tilde{p} = 0, & \text{on } \partial\Omega. \end{cases} \tag{7}$$

Let $\tilde{\mathbf{\mu}} = K\nabla\tilde{p}$, then the mixed variational form of (7) is:

Find $(\lambda, \tilde{\mathbf{\mu}}, \tilde{p}) \in R \times \mathbf{V} \times W$, such that

$$\begin{cases} a(\tilde{\mathbf{\mu}}, \mathbf{\varphi}) - b(\mathbf{\varphi}, \tilde{p}) = 0, & \forall \mathbf{\varphi} \in \mathbf{V}, \\ b(\tilde{\mathbf{\mu}}, v) = \lambda(p, v), & \forall v \in W. \end{cases} \tag{8}$$

From the subtraction of (3) and (8), we get: find $(\mathbf{\mu} - \tilde{\mathbf{\mu}}, p - \tilde{p}) \in \mathbf{V} \times W$, such that

$$\begin{cases} a(\mathbf{\mu} - \tilde{\mathbf{\mu}}, \mathbf{\varphi}) - b(\mathbf{\varphi}, p - \tilde{p}) = 0, & \forall \mathbf{\varphi} \in \mathbf{V}, \\ b(\mathbf{\mu} - \tilde{\mathbf{\mu}}, v) = 0, & \forall v \in W. \end{cases} \tag{9}$$

Take form (9), let $\mathbf{\varphi} = \mathbf{\mu} - \tilde{\mathbf{\mu}}$, $v = p - \tilde{p}$, then

$$\begin{cases} a(\mathbf{\mu} - \tilde{\mathbf{\mu}}, \mathbf{\mu} - \tilde{\mathbf{\mu}}) - b(\mathbf{\mu} - \tilde{\mathbf{\mu}}, p - \tilde{p}) = 0 \\ b(\mathbf{\mu} - \tilde{\mathbf{\mu}}, p - \tilde{p}) = 0 \end{cases}$$

Add the above two equations, and we get $a(\mathbf{\mu} - \tilde{\mathbf{\mu}}, \mathbf{\mu} - \tilde{\mathbf{\mu}}) = 0$, this illustrates $\mathbf{\mu} = \tilde{\mathbf{\mu}}$.

Substitute $\mathbf{\mu} = \tilde{\mathbf{\mu}}$ into the first equation of (9), and we get $b(\mathbf{\varphi}, p - \tilde{p}) = 0$, i.e $\int_\Omega (p - \tilde{p}) \cdot div\mathbf{\varphi} dx = 0, \forall \mathbf{\varphi} \in \mathbf{V}$.

Take ω satisfied $\Delta\omega = p - \tilde{p}$, and let $\mathbf{\varphi} = \nabla\omega$, then by $div\mathbf{\varphi} = p - \tilde{p}$, pushed $p = \tilde{p}$.

This proves (λ, p) is an eigenpair of equation (1), and $\mathbf{\mu} = K\nabla p$.

We complete the proof.

Now, let's define the mixed finite element approximations of the problem (3). Let $\mathfrak{I}_h$ be a partition of Ω into finite elements(triangles),which is regular and has a mesh size h .Associated with the partition $\mathfrak{I}_h$,we define the finite dimensional spaces W_h and $\mathbf{V}_h$ (see[4]), where for any $\kappa \in \mathfrak{I}_h$, $P_n(\kappa)(n \geq 0)$ denotes the spaces of polynomial of degree not greater than n on κ.

Define

$$\mathbf{V}_h^L = \left\{ q_h \in \mathbf{V} \cap C^0(\bar{\Omega})^2 : q_h|_\kappa \in P_1(\kappa)^2, \forall \kappa \in \mathfrak{I}_h \right\},$$

for each $\kappa \in \mathfrak{I}_h$, and the barycentric coordinates $\lambda_1\lambda_2\lambda_3$ on κ, define

$$B = (span\left\{ \lambda_1\lambda_2\lambda_3\lambda_j : \kappa \in \mathfrak{I}_h, j = 1,2,3 \right\})^2,$$

and

$$\mathbf{V}_h = \mathbf{V}_h^L \oplus B.$$

Apparently, we have $\mathbf{V}_h^L \subset \mathbf{V}_h \subset \mathbf{V}$.

Afterward, define

$$W_h = \left\{ v_h \in H_0^1(\Omega) \cap C^0(\bar{\Omega}) : v_h|_\kappa \in P_2(\kappa), \forall \kappa \in \mathfrak{I}_h \right\}.$$

Apparently, we have $W_h \subset W$.

With the discrete spaces defined above, the mixed finite element approximation of (3) is given by:

Find $(\lambda_h, \mathbf{\mu}_h, p_h) \in R \times \mathbf{V}_h \times W_h$, $(\mathbf{\mu}_h, p_h) \neq (0,0)$, such that

$$\begin{cases} a(\mathbf{\mu}_h, \mathbf{\varphi}) - b(\mathbf{\varphi}, p_h) = 0, & \forall \mathbf{\varphi} \in \mathbf{V}_h, \\ b(\mathbf{\mu}_h, v) = \lambda_h(p_h, v), & \forall v \in W_h. \end{cases} \tag{10}$$

For the eigenvalue λ_h, there exists the following Rayleigh quotient expression

$$\lambda_h = \frac{-a(\boldsymbol{\mu}_h, \boldsymbol{\mu}_h) + 2b(\boldsymbol{\mu}_h, p_h)}{(p_h, p_h)}.$$

Form [9] the eigenvalue problem (10) has eigenvalues

$$0 \le \lambda_{1,h} \le \cdots \lambda_{k,h} \le \cdots \le \lambda_{N,h},$$

and the corresponding eigenfunctions

$$(\boldsymbol{\mu}_{1,h}, p_{1,h}), (\boldsymbol{\mu}_{2,h}, p_{2,h}), \cdots (\boldsymbol{\mu}_{k,h}, p_{k,h}) \cdots (\boldsymbol{\mu}_{N,h}, p_{N,h}),$$

where $(p_{i,h}, p_{j,h}) = \delta_{ij}, 1 \le i, j \le N, N = \dim W_h$.

For any $f \in L^2(\Omega)$, consider the boundary value problem corresponding to (3) and its mixed finite element approximation: find $(\boldsymbol{\sigma}, u) \in \mathbf{V} \times W, (\boldsymbol{\sigma}, u) \ne (0,0)$, such that

$$\begin{cases} a(\boldsymbol{\sigma}, \boldsymbol{\varphi}) - b(\boldsymbol{\varphi}, u) = 0, & \forall \boldsymbol{\varphi} \in \mathbf{V}, \\ b(\boldsymbol{\sigma}, v) = (f, v), & \forall v \in W. \end{cases} \tag{11}$$

find $(\boldsymbol{\sigma}_h, u_h) \in \mathbf{V}_h \times W_h, (\boldsymbol{\sigma}_h, u_h) \ne (0,0)$, such that

$$\begin{cases} a(\boldsymbol{\sigma}_h, \boldsymbol{\varphi}) - b(\boldsymbol{\varphi}, u_h) = 0, & \forall \boldsymbol{\varphi} \in \mathbf{V}_h, \\ b(\boldsymbol{\sigma}_h, v) = (f, v), & \forall v \in W_h. \end{cases} \tag{12}$$

IV. OPERATOR FORM AND ITS PROPERTIES

For any $f \in L^2(\Omega)$, assume that (11) has a unique solution $(\boldsymbol{\sigma}, u)$, and since $\mathbf{V}_h \subset \mathbf{V}, W_h \subset W$, it is known that (12) has a unique solution $(\boldsymbol{\sigma}_h, u_h)$ .Thus, a linear bounded operator can be defined

$$T: G \to W \subset G, Tf = u. \qquad T_h: G \to W_h \subset G, T_h f = u_h.$$
$$\mathbf{S}: G \to \mathbf{V} \subset \mathbf{H}, \mathbf{S}f = \boldsymbol{\sigma}. \qquad \mathbf{S}_h: G \to \mathbf{V}_h \subset \mathbf{H}, \mathbf{S}_h f = \boldsymbol{\sigma}_h.$$

Thus, the eigenvalue problems (3) and (10) have equivalent operator forms, respectively.

$$\begin{cases} \lambda T p = p \\ \mathbf{S}(\lambda p) = \boldsymbol{\mu} \end{cases} \tag{13}$$

$$\begin{cases} \lambda_h T_h p_h = p_h \\ \mathbf{S}_h(\lambda_h p_h) = \boldsymbol{\mu}_h \end{cases} \tag{14}$$

Therefore, solving for the eigenpair of (3) for $(\lambda, \boldsymbol{\mu}, p)$ can be reduced to solving for the eigenpair of the operator T for $((\tau = \lambda^{-1}), p)$ and $\boldsymbol{\mu} = \mathbf{S}(\lambda p)$; similarly, solving for the eigenpair of (10) for $(\lambda_h, \boldsymbol{\mu}_h, p_h)$ can be reduced to solving for the eigenpair of the operator T_h for $((\tau_h = \lambda_h^{-1}), p_h)$ and $\boldsymbol{\mu}_h = \mathbf{S}_h(\lambda_h p_h)$.

For the linear bounded operators T and $\mathbf{S}$ defined in above, for any $f \in L^2(\Omega)$, the following relations hold:

$$\begin{cases} a(\mathbf{S}f, \boldsymbol{\varphi}) - b(\boldsymbol{\varphi}, Tf) = 0, & \forall \boldsymbol{\varphi} \in \mathbf{V}, \\ b(\mathbf{S}f, v) = (f, v), & \forall v \in W. \end{cases} \tag{15}$$

For this elliptic problem, the following regularity estimate holds

$$\|Tf\|_{1+r_0} \le C \|f\|_0.$$

Where $\frac{1}{2} < r_0 \le 1$, depends on the shape of the domain.

For the discrete version of the linear bounded operators T_h and $\mathbf{S}_h$ defined in above, for any $f \in L^2(\Omega)$, the following relations hold:

$$\begin{cases} a(\mathbf{S}_h f, \boldsymbol{\varphi}) - b(\boldsymbol{\varphi}, T_h f) = 0, & \forall \boldsymbol{\varphi} \in \mathbf{V}_h, \\ b(\mathbf{S}_h f, v) = (f, v), & \forall v \in W. \end{cases} \tag{16}$$

Lemma 1. (Lemma 1 in [13]) T and T_h are self-adjoint operators.

Proof. For any $g \in L^2(\Omega)$, let $\boldsymbol{\sigma} = \mathbf{S}g, u = Tg$, similarly, we have

$$\begin{cases} a(\mathbf{S}g, \boldsymbol{\varphi}) - b(\boldsymbol{\varphi}, Tg) = 0, & \forall \boldsymbol{\varphi} \in \mathbf{V}, \\ b(\mathbf{S}g, v) = (g, v), & \forall v \in W. \end{cases} \tag{17}$$

By taking $\boldsymbol{\varphi} = \mathbf{S}g, v = Tg$ in (15), we get

$$\begin{cases} a(\mathbf{S}f,\mathbf{S}g)-b(\mathbf{S}g,Tf)=0 \\ b(\mathbf{S}f,Tg)=(f,Tg) \end{cases} \tag{18}$$

By taking $\boldsymbol{\varphi}=\mathbf{S}f, v=Tf$ in (17), we get

$$\begin{cases} a(\mathbf{S}g,\mathbf{S}f)-b(\mathbf{S}f,Tg)=0 \\ b(\mathbf{S}g,Tf)=(g,Tf) \end{cases} \tag{19}$$

From the symmetry of $a(\cdot,\cdot)$, (18) and (19) , we can obtain

$$(f,Tg)_G = b(\mathbf{S}f,Tg) = a(\mathbf{S}g,\mathbf{S}f) = a(\mathbf{S}f,\mathbf{S}g) = b(\mathbf{S}g,Tf) = (g,Tf)_G.$$

Thus, T is self-adjoint; similarly, it can be proven that T_h is self-adjoint.

V. A PRIORI ERROR ESTIMATE

5.1 A priori error estimate for eigenfunctions

Lemma 2. For any $\boldsymbol{\varphi}_h \in z_h$, there exists a constant α independent of h, such that the following inequality holds:

$$a(\boldsymbol{\varphi}_h,\boldsymbol{\varphi}_h) \geq \alpha \|\boldsymbol{\varphi}_h\|_{\mathbf{H}}^2, \forall \boldsymbol{\varphi}_h \in Z_h , \tag{20}$$

where $Z_h = \{\boldsymbol{\varphi}_h \in \mathbf{V}_h : b(\boldsymbol{\varphi}_h, v_h)=0, \forall v_h \in W_h\}$

Proof. The property (20) is obvious.

Lemma 3. (Lemma 2.27 in [4]) For any $\boldsymbol{\varphi} \in \mathbf{V}$ and $v \in H_0^1(\Omega)$, both

$$(\mathrm{div}\boldsymbol{\varphi}, v) = -(\boldsymbol{\varphi}, \nabla v) , \tag{21}$$

where $(\cdot,\cdot)$ represents the inner product of $L^2(\Omega)^2$.

Proof. The property (21) can be proven using the divergence theorem.

Corollary 1. For any $u \subset H^3(\Omega)$, we have

$$|u-\rho_h u|_1 \lesssim h^t |u|_{1+t}, 1 \le t \le 2 , \tag{22}$$

where $\rho_h : H^3(\Omega) \to W_h$ is the L^2-projection operator.

Proof. see [4] .

Let $Q_h : \mathbf{V} \to \mathbf{V}_h^L$ be the L^2-projection, such that for any $\boldsymbol{\varphi} \in \mathbf{V}$, it holds that

$$(\boldsymbol{\varphi} - Q_h\boldsymbol{\varphi}, \boldsymbol{\varphi}_h) = 0, \forall \boldsymbol{\varphi}_h \in \mathbf{V}_h^L , \tag{23}$$

then

$$\|\boldsymbol{\varphi} - Q_h\boldsymbol{\varphi}\|_0 \lesssim h^k |\boldsymbol{\varphi}|_k , \tag{24}$$

where $\boldsymbol{\varphi} \in [H^k(\Omega)]^2, 0 \le k \le 2$.

Define the operator $r_h : \mathbf{V} \to \mathbf{V}_h$, such that for any $\boldsymbol{\varphi} \in \mathbf{V}$, it holds that $r_h\boldsymbol{\varphi}|_\kappa = Q_h\boldsymbol{\varphi}|_\kappa + \sum_{j=1}^{3}\alpha_j\lambda_1\lambda_2\lambda_3\lambda_j, \forall \kappa \in \mathfrak{I}_h$,

where Q_h is defined by (23), and $\alpha_j \in R^2 (j=1,2,3)$ is an undetermined constant vector.

Assume that there exists an operator $r_h : \mathbf{V} \to \mathbf{V}_h$, such that for any $\boldsymbol{\varphi} \in \mathbf{V}$, it holds that

$$b(\boldsymbol{\varphi} - r_h\boldsymbol{\varphi}, v_h) = 0, \forall v_h \in W_h , \tag{25}$$

for any $v_h \in W_h$, since $\nabla v_h|_\kappa$ is a first-degree polynomial vector, without loss of generality, let $\nabla v_h|_\kappa = \sum_{i=1}^{3}\beta_i\lambda_i$, where $\beta_i (i=1,2,3)$ is a constant vector. Then, by Lemma 3, we have

$$\begin{aligned} b(\boldsymbol{\varphi} - r_h\boldsymbol{\varphi}, v_h) &= -(\mathrm{div}(\boldsymbol{\varphi} - r_h\boldsymbol{\varphi}), v_h) \\ &= (\boldsymbol{\varphi} - r_h\boldsymbol{\varphi}, \nabla v_h) \\ &= \sum_{\kappa\in\mathfrak{I}_h}\int_\kappa (\boldsymbol{\varphi} - r_h\boldsymbol{\varphi})\nabla v_h \mathrm{d}x \\ &= \sum_{\kappa\in\mathfrak{I}_h}\sum_{i=1}^{3}\beta_i\int_\kappa \lambda_i(\boldsymbol{\varphi} - r_h\boldsymbol{\varphi})\mathrm{d}x \end{aligned}$$

To make equation (25) hold, it is sufficient to $\int_\kappa \lambda_i(\boldsymbol{\varphi} - r_h\boldsymbol{\varphi})\mathrm{d}x = 0, i=1,2,3, \forall \kappa \in \mathfrak{I}_h$.

From the definition of r_h, it is enough to choose α_j, such that

$$\sum_{j=1}^{3}\alpha_j\int_\kappa \lambda_1\lambda_2\lambda_3\lambda_j\lambda_i \mathrm{d}x=\int_\kappa \lambda_i(\boldsymbol{\varphi}-Q_h\boldsymbol{\varphi})\mathrm{d}x, i=1,2,3, \tag{26}$$

upon calculation, the determinant of the coefficient matrix of the system (26) is non-zero, which implies that the system (26) has a unique solution $\alpha_j\,(j=1,2,3)$, ensuring that r_h satisfies equation (25).

Lemma 4. Existence of the operator $r_h:\mathbf{V}\to\mathbf{V}_h$, such that for any $\boldsymbol{\varphi}\in\mathbf{V}$, it holds that

$$b(\boldsymbol{\varphi}-r_h\boldsymbol{\varphi},v_h)=0,\forall v_h\in W_h,$$

when $\boldsymbol{\varphi}\in[H^k(\Omega)]^2, 0\le k\le 2$, we have

$$\|\boldsymbol{\varphi}-r_h\boldsymbol{\varphi}\|_{0,\Omega}\lesssim h^k|\boldsymbol{\varphi}|_k. \tag{27}$$

Proof. When solving for $\alpha_j\,(j=1,2,3)$ using the Grammer rule, from Section 1.4 of [4], specifically equations (1.4.9), (1.4.32), and Lemma 1.20–Lemma 1.22, we have

$$|\alpha_j|\lesssim h^{-1}\|\boldsymbol{\varphi}-Q_h\boldsymbol{\varphi}\|_{0,\kappa}, j=1,2,3. \tag{28}$$

From the definition of r_h, (28), *Hölder* inequality, we have

$$\begin{aligned}\|\boldsymbol{\varphi}-r_h\boldsymbol{\varphi}\|_{0,\kappa}&\lesssim\|\boldsymbol{\varphi}-Q_h\boldsymbol{\varphi}\|_{0,\kappa}+\sum_{j=1}^{3}|\alpha_j|\,\|\lambda_1\lambda_2\lambda_3\lambda_j\|_{0,\kappa}\\&\lesssim\|\boldsymbol{\varphi}-Q_h\boldsymbol{\varphi}\|_{0,\kappa}+h^{-1}\|\boldsymbol{\varphi}-Q_h\boldsymbol{\varphi}\|_{0,\kappa}\cdot[\mathrm{mes}(\kappa)]^{1/2}\\&\lesssim\|\boldsymbol{\varphi}-Q_h\boldsymbol{\varphi}\|_{0,\kappa}\end{aligned}$$

From (24) we can obtain

$$\|\boldsymbol{\varphi}-r_h\boldsymbol{\varphi}\|_{0,\Omega}=\left(\sum_{\kappa\in\mathfrak{J}_h}\|\boldsymbol{\varphi}-r_h\boldsymbol{\varphi}\|_{0,\kappa}^2\right)^{1/2}\lesssim\|\boldsymbol{\varphi}-Q_h\boldsymbol{\varphi}\|_{0,\Omega}\lesssim h^k|\boldsymbol{\varphi}|_k,$$

where $\boldsymbol{\varphi}\in[H^k(\Omega)]^2, 0\le k\le 2$.

Theorem 2. Assume that there exists an operator $r_h:\mathbf{V}\to\mathbf{V}_h$, such that for any $\boldsymbol{\varphi}\in\mathbf{V}$, it holds that

$$b(\boldsymbol{\varphi}-r_h\boldsymbol{\varphi},v_h)=0,\forall v_h\in W_h,$$

moreover, $(\boldsymbol{\sigma},u)\in\mathbf{V}\times W$ is the solution to problem (11), and $(\boldsymbol{\sigma}_h,u_h)\in\mathbf{V}_h\times W_h$ is the solution to problem (12). Then the following error estimate holds:

$$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_0\lesssim\|\boldsymbol{\sigma}-r_h\boldsymbol{\sigma}\|_0+|u-v_h|_1,\forall v_h\in W_h\cap H_0^1(\Omega), \tag{29}$$

$$\begin{aligned}\|u-u_h\|_0\lesssim&\sup_{0\ne d\in L^2(\Omega)}\frac{1}{\|d\|_0}[b(\boldsymbol{\lambda}_d-r_h\boldsymbol{\lambda}_d,u-v_h)+a(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,r_h\boldsymbol{\lambda}_d-\boldsymbol{\lambda}_d)\\&+b(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,y_d-\eta_h)],\forall v_h,\eta_h\in W_h.\end{aligned} \tag{30}$$

Where for any $d\in L^2(\Omega)$, the function pairs are defined in $(\boldsymbol{\lambda}_d,y_d)\in\mathbf{V}\times[H_0^1(\Omega)\cap H^{1+r}(\Omega)]$ and satisfies

$$\begin{cases}a(\boldsymbol{\lambda}_d,\boldsymbol{\varphi})-b(\boldsymbol{\varphi},y_d)=0, & \forall\boldsymbol{\varphi}\in\mathbf{V},\\ b(\boldsymbol{\lambda}_d,v)=(d,v), & \forall v\in W.\end{cases} \tag{31}$$

and we know a priori estimate:

$$\|\boldsymbol{\lambda}_d\|_{r_0}+\|y_d\|_{1+r_0}\lesssim\|d\|_0. \tag{32}$$

Where $\frac{1}{2}<r\le 2$, depends on the shape of the domain.

Proof. From equations (11) and (12), by subtracting the corresponding terms, we obtain the error equation

$$\begin{cases}a(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\boldsymbol{\varphi})-b(u-u_h,\boldsymbol{\varphi})=0, & \forall\boldsymbol{\varphi}\in\mathbf{V}_h,\\ b(v,\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)=0, & \forall v\in W_h.\end{cases} \tag{33}$$

From the error equation (33), $b(\boldsymbol{\varphi}-r_h\boldsymbol{\varphi},v_h)=0$, it follows that for any $v_h\in W_h$, we have

$$\begin{aligned}b(r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,v_h)&=b(r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}+\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,v_h)\\&=b(r_h\boldsymbol{\sigma}-\boldsymbol{\sigma},v_h)+b(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,v_h)\\&=0\end{aligned}$$

Thus

$$r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\in Z_h.$$

By combining inequality (20) and the error equation (33), we have

$$
\begin{aligned}
\alpha \| r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}_h \|_0^2 \le\ & a(r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\\
&= a(r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}, r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)+a(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\\
&= a(r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}, r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)+b(u-u_h, r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\\
&= a(r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}, r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)+b(r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}, u-u_h)+b(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, u-u_h)\\
&= a(r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}, r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)+b(r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}, u-v_h+v_h-u_h)+b(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, u-v_h+v_h-u_h)\\
&= a(r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}, r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)+b(r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}, u-v_h)+b(r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}, v_h-u_h)+b(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, u-v_h)+b(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, v_h-u_h)\\
&= I_1
\end{aligned}
$$

Also, since the operator r_h satisfies that for any $\boldsymbol{\varphi}\in \mathbf{V}, v_h \in W_h$, $b(\boldsymbol{\varphi}-r_h\boldsymbol{\varphi}, v_h)=0$ holds, we obtain

$$b(r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}, v_h-u_h)=0.$$

Next, from the error equation (33), we obtain

$$b(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, v_h-u_h)=0.$$

Thus, from the above two expressions, Lemma 3, and the $H\ddot{o}lder$ inequality, we have

$$
\begin{aligned}
I_1 &= a(r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}, r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)+b(r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}, u-v_h)+b(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, u-v_h)\\
&= a(r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}, r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)+b(r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, u-v_h)\\
&= a(r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}, r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)+(r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, \nabla(u-v_h))\\
&\le \| r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}\|_0 \| r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_0 + \| r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_0 \|\nabla(u-v_h)\|_0
\end{aligned}
$$

Therefore, we have

$$\| r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_0 \le \| r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}\|_0 + \|\nabla(u-v_h)\|_0, \forall v_h \in W_h \cap H_0^1(\Omega).$$

Using the triangle inequality and the above conclusions, we obtain

$$
\begin{aligned}
\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_0 &= \|\boldsymbol{\sigma}-r_h\boldsymbol{\sigma}+r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_0\\
&\le \|\boldsymbol{\sigma}-r_h\boldsymbol{\sigma}\|_0 + \| r_h\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_0\\
&\le \|\boldsymbol{\sigma}-r_h\boldsymbol{\sigma}\|_0 + \|\nabla(u-v_h)\|_0\\
&= \|\boldsymbol{\sigma}-r_h\boldsymbol{\sigma}\|_0 + |u-v_h|_1, \forall v_h \in W_h \cap H_0^1(\Omega)
\end{aligned}
$$

Thus, we obtain (29).

By combining inequality (31) and the error equation (33), for any $v_h \in W_h$, we have

$$
\begin{aligned}
(d, u-u_h) &= b(\boldsymbol{\lambda}_d, u-u_h)\\
&= b(\boldsymbol{\lambda}_d - r_h\boldsymbol{\lambda}_d + r_h\boldsymbol{\lambda}_d, u-u_h)\\
&= b(\boldsymbol{\lambda}_d - r_h\boldsymbol{\lambda}_d, u-u_h)+b(r_h\boldsymbol{\lambda}_d, u-u_h)\\
&= b(\boldsymbol{\lambda}_d - r_h\boldsymbol{\lambda}_d, u-v_h+v_h-u_h)+a(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, r_h\boldsymbol{\lambda}_d)\\
&= b(\boldsymbol{\lambda}_d - r_h\boldsymbol{\lambda}_d, u-v_h)+a(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, r_h\boldsymbol{\lambda}_d)+b(\boldsymbol{\lambda}_d - r_h\boldsymbol{\lambda}_d, v_h-u_h)\\
&= I_2
\end{aligned}
$$

Furthermore, since the operator r_h satisfies that for any $\boldsymbol{\varphi}\in \mathbf{V}, v_h \in W_h$, $b(\boldsymbol{\varphi}-r_h\boldsymbol{\varphi}, v_h)=0$ holds, we obtain

$$b(\boldsymbol{\lambda}_d - r_h\boldsymbol{\lambda}_d, v_h-u_h)=0.$$

For any $\eta_h \in W_h$, by combining the above expressions and (31), we get

$$
\begin{aligned}
I_2 &= b(\boldsymbol{\lambda}_d - r_h\boldsymbol{\lambda}_d, u-v_h)+a(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, r_h\boldsymbol{\lambda}_d)\\
&= b(\boldsymbol{\lambda}_d - r_h\boldsymbol{\lambda}_d, u-v_h)+a(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, r_h\boldsymbol{\lambda}_d-\boldsymbol{\lambda}_d)+a(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, \boldsymbol{\lambda}_d)\\
&= b(\boldsymbol{\lambda}_d - r_h\boldsymbol{\lambda}_d, u-v_h)+a(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, r_h\boldsymbol{\lambda}_d-\boldsymbol{\lambda}_d)+b(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, y_d)\\
&= b(\boldsymbol{\lambda}_d - r_h\boldsymbol{\lambda}_d, u-v_h)+a(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, r_h\boldsymbol{\lambda}_d-\boldsymbol{\lambda}_d)+b(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, y_d-\eta_h)+b(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, \eta_h)
\end{aligned}
$$

Then, from the error equation (33), we obtain $b(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, \eta_h)=0$. So we have

$$
\begin{aligned}
(d, u-u_h) &= b(\boldsymbol{\lambda}_d, u-u_h)\\
&= b(\boldsymbol{\lambda}_d - r_h\boldsymbol{\lambda}_d, u-v_h)+a(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, r_h\boldsymbol{\lambda}_d-\boldsymbol{\lambda}_d)+b(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, y_d-\eta_h)
\end{aligned}
$$

Substituting the above results into the following norm expression,

$$\| u-u_h\|_G = \sup_{0\ne d\in L^2(\Omega)} \frac{(d, u-u_h)}{\| d\|_G} = \sup_{0\ne d\in L^2(\Omega)} \frac{b(\boldsymbol{\lambda}_d, u-u_h)}{\| d\|_G}.$$

we can obtain

$$\|u-u_h\|_0 \lesssim \sup_{0\neq d\in L^2(\Omega)} \frac{1}{\|d\|_0}[b(\boldsymbol{\lambda}_d - r_h\boldsymbol{\lambda}_d, u-v_h) + a(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, r_h\boldsymbol{\lambda}_d - \boldsymbol{\lambda}_d) + b(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, y_d - \eta_h)], \forall v_h, \eta_h \in W_h,$$

we obtain equation (30).

Theorem 3.Suppos $(\boldsymbol{\sigma}, u) \in [H^r(\Omega)]^2 \times (H^{r+1}(\Omega) \cap H_0^1(\Omega))$ is the solution of problem (11),and $(\boldsymbol{\sigma}_h, u_h) \in \mathbf{V}_h \times W_h$ is the solution of problem (12), then the following error estimate holds:

$$h^{r_0}\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_0 + \|u-u_h\|_0 \lesssim h^{r_0+r}\|u\|_{1+r}. \tag{34}$$

Proof. From(29),(27) and (22), we can give the estimate for $\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_0$ as

$$\begin{aligned} \|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_0 &\lesssim \|\boldsymbol{\sigma}-r_h\boldsymbol{\sigma}\|_0 + |u-\rho_h u|_1 \\ &\lesssim h^r|\boldsymbol{\sigma}|_r + h^r|u|_{1+r} \\ &= h^r|K\nabla u|_r + h^r|u|_{1+r} \\ &\lesssim h^r\|u\|_{1+r}. \end{aligned} \tag{35}$$

Where $\rho_h : H^3(\Omega) \to W_h$ is the L^2-projection operator.

Next, we estimate the three terms on the right-hand side of the inequality in (30).

Here, $(\boldsymbol{\lambda}_d, y_d)$ is the solution to the auxiliary problem (31) introduced in Theorem 2.

From (21),(22),(27),(32) and *Hölder* inequality, we have

$$\begin{aligned} b(\boldsymbol{\lambda}_d - r_h\boldsymbol{\lambda}_d, u-\rho_h u) &= -(\mathrm{div}(\boldsymbol{\lambda}_d - r_h\boldsymbol{\lambda}_d), u-\rho_h u) \\ &= (\boldsymbol{\lambda}_d - r_h\boldsymbol{\lambda}_d, \nabla(u-\rho_h u)) \\ &\lesssim \|\boldsymbol{\lambda}_d - r_h\boldsymbol{\lambda}_d\|_{r_0}|u-\rho_h u|_1 \\ &\lesssim h^{r_0}\|\boldsymbol{\lambda}_d\|_{r_0}\cdot h^r|u|_{1+r} \\ &\lesssim h^{r_0+r}\|u\|_{1+r}\|\boldsymbol{\lambda}_d\|_{r_0} \\ &\lesssim h^{r_0+r}\|u\|_{1+r}\|d\|_0. \end{aligned} \tag{36}$$

From (27),(32),(35) and *Hölder* inequality, we have

$$\begin{aligned} a(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, r_h\boldsymbol{\lambda}_d - \boldsymbol{\lambda}_d) &\lesssim \|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_0\|r_h\boldsymbol{\lambda}_d - \boldsymbol{\lambda}_d\|_0 \\ &\lesssim h^r\|u\|_{1+r}\cdot h^{r_0}\|\boldsymbol{\lambda}_d\|_{r_0} \\ &\lesssim h^{r_0+r}\|u\|_{1+r}\|d\|_0. \end{aligned} \tag{37}$$

From (21),(32),(35) and *Hölder* inequality, we have

$$\begin{aligned} b(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, y_d - \rho_h y_d) &= -(\mathrm{div}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h), y_d - \rho_h y_d) \\ &= (\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, \nabla(y_d - \rho_h y_d)) \\ &\lesssim \|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_0|y_d - \rho_h y_d|_1 \\ &\lesssim h^r\|u\|_{1+r}\cdot h^{r_0}|y_d|_{1+r_0} \\ &\lesssim h^r\|u\|_{1+r}\cdot h^{r_0}\|y_d\|_{1+r_0} \\ &\lesssim h^{r_0+r}\|u\|_{1+r}\|d\|_0. \end{aligned} \tag{38}$$

Thus, substituting (36),(37) and (38) into (30) from Theorem 2, we can obtain

$$\begin{aligned} \|u-u_h\|_0 \lesssim & \sup_{0\neq d\in L^2(\Omega)} \frac{1}{\|d\|_0}[b(\boldsymbol{\lambda}_d - r_h\boldsymbol{\lambda}_d, u-\rho_h u) \\ & + a(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, r_h\boldsymbol{\lambda}_d - \boldsymbol{\lambda}_d) + b(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h, y_d - \rho_h y_d)] \\ \lesssim & h^{r_0+r}\|u\|_{1+r}. \end{aligned} \tag{39}$$

Finally combining (35) and (39), we can get the desired result (34).

Theorem 4. For the previously defined T and T_h, we have $\|T-T_h\|_G \to 0$, as $h \to 0$.

Proof. Let $Tf = u, T_h f = u_h$, then we have

$$\|T-T_h\|_G=\sup_{0\neq f\in L^2(\Omega)}\frac{\|Tf-T_hf\|_0}{\|f\|_0}=\sup_{0\neq f\in L^2(\Omega)}\frac{\|u-u_h\|_0}{\|f\|_0}$$
$$\lesssim \sup_{0\neq f\in L^2(\Omega)}\frac{h^{r_0+r_0}\|u\|_{1+r_0}}{\|f\|_0}\lesssim \sup_{0\neq f\in L^2(\Omega)}\frac{h^{r_0+r_0}\|f\|_0}{\|f\|_0}\lesssim h^{2r_0}\to 0,(h\to 0).$$

We complete the proof.

Corollary 2. Assume that $u\in H^3(\Omega)$, the following error estimate holds:

$$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{H(div,\Omega)}\lesssim h\|u\|_3 \tag{40}$$

Proof.

$$\begin{aligned}\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{H(div,\Omega)}&=\|\boldsymbol{\sigma}-I_h^1\boldsymbol{\sigma}+I_h^1\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{H(div,\Omega)}\\
&\lesssim h\|\boldsymbol{\sigma}\|_2+h^{-1}\|I_h^1\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_0\\
&=h\|\boldsymbol{\sigma}\|_2+h^{-1}\|I_h^1\boldsymbol{\sigma}-\boldsymbol{\sigma}+\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_0\\
&\lesssim h\|\boldsymbol{\sigma}\|_2+h^{-1}h^2\|\boldsymbol{\sigma}\|_2+h^{-1}\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_0\\
&\lesssim h\|\boldsymbol{\sigma}\|_2\\
&=h\|K\nabla u\|_2\\
&\lesssim h\|u\|_3 .\end{aligned}$$

Where I_h^1 is the linear finite element interpolation.

5.2 A priori error estimate for eigenvalues

Let $(\lambda,\boldsymbol{\mu},p)$ be an eigenpair of (3), and $(\lambda_h,\boldsymbol{\mu}_h,p_h)$ be an eigenpair of (10); $(\lambda_h,\boldsymbol{\mu}_h,p_h)$ approximates $(\lambda,\boldsymbol{\mu},p)$.
Let M_λ be the space spanned by the eigenfunctions $\{u_j\}$ corresponding to the eigenvalue λ of (3).

Lemma 5. (Lemma 2 in [13]) Suppose the multiplicity of the eigenvalue λ is *m*, then the following estimate holds:

$$|\lambda-\lambda_{l,h}|\lesssim C\left\{\left\|(\mathbf{S}-\mathbf{S}_h)|_{M_\lambda}\right\|^2_{G\to\mathbf{H}}+\left\|(\mathbf{S}-\mathbf{S}_h)|_{M_\lambda}\right\|_{G\to\mathbf{V}}\cdot\left\|(T-T_h)|_{M_\lambda}\right\|_{G\to W}+\left\|(T-T_h)|_{M_\lambda}\right\|^2_{G\to G}\right\} \tag{41}$$

Proof. Since the multiplicity of the eigenvalues of a self-adjoint operator is equal to the dimension of the eigenspace, let $u_1,u_2,\cdots u_m$ be an orthonormal basis of M_λ.
By Theorem 3 of [9] and the steepness $\alpha=1$ of the self-adjoint operator, we have the following estimate.

$$\left|\lambda^{-1}-\lambda_{l,h}^{-1}\right|\lesssim C\left\{\sum_{i,j=1}^m\left|((T-T_h)u_i,u_j)\right|+\left\|(T-T_h)|_{M_\lambda}\right\|^2_{G\to G}\right\}. \tag{42}$$

For any $f,g\in L^2(\Omega)$, let us consider $|(T-T_h)g,f|$.
By the two equations of (15), we obtain the following

$$(f,v)=-a(\mathbf{S}f,\boldsymbol{\varphi})+b(\boldsymbol{\varphi},Tf)+b(\mathbf{S}f,v),\forall(\boldsymbol{\varphi},v)\in\mathbf{V}\times W .$$

For $g\in L^2(\Omega)$, let $\boldsymbol{\varphi}=(\mathbf{S}-\mathbf{S}_h)g$, $v=(T-T_h)g$, then

$$(f,(T-T_h)g)=-a(\mathbf{S}f,(\mathbf{S}-\mathbf{S}_h)g)+b((\mathbf{S}-\mathbf{S}_h)g,Tf)+b(\mathbf{S}f,(T-T_h)g), \tag{43}$$

replacing $g\in L^2(\Omega)$ for f from (15), we have

$$\begin{cases}a(\mathbf{S}g,\boldsymbol{\varphi})-b(\boldsymbol{\varphi},Tg)=0, & \forall\boldsymbol{\varphi}\in\mathbf{V},\\ b(\mathbf{S}g,v)=(g,v), & \forall v\in W.\end{cases} \tag{44}$$

From (16), we have

$$\begin{cases}a(\mathbf{S}_hg,\boldsymbol{\varphi})-b(\boldsymbol{\varphi},T_hg)=0, & \forall\boldsymbol{\varphi}\in\mathbf{V}_h,\\ b(\mathbf{S}_hg,v)=(g,v), & \forall v\in W.\end{cases} \tag{45}$$

By subtracting (44) and (45), we have

$$\begin{cases}a((\mathbf{S}-\mathbf{S}_h)g,\boldsymbol{\varphi})-b(\boldsymbol{\varphi},(T-T_h)g)=0\\ -b((\mathbf{S}-\mathbf{S}_h)g,v)=0\end{cases}$$

Adding the two above equations yields the following.

$$a((\mathbf{S}-\mathbf{S}_h)g,\boldsymbol{\varphi})-b(\boldsymbol{\varphi},(T-T_h)g)-b((\mathbf{S}-\mathbf{S}_h)g,v)=0 . \tag{46}$$

Since $a(\cdot,\cdot)$ is symmetric, adding (43) and (46) gives the result

$$(f,(T-T_h)g) = -a((\mathbf{S}-\mathbf{S}_h)g, \boldsymbol{\varphi}-\mathbf{S}f) + b((\mathbf{S}-\mathbf{S}_h)g, Tf-v) + b(\mathbf{S}f-\boldsymbol{\varphi}, (T-T_h)g).$$

From equations (4) to (6), for any $\boldsymbol{\varphi} \in \mathbf{V}_h$, $v \in W_h$, we have

$$\begin{aligned} |(f,(T-T_h)g)| \lesssim\ & C_1\|(\mathbf{S}-\mathbf{S}_h)g\|_{\mathbf{H}}\|\boldsymbol{\varphi}-\mathbf{S}f\|_{\mathbf{H}} + C_2\|(\mathbf{S}-\mathbf{S}_h)g\|_{\mathbf{V}}\|v-Tf\|_W \\ & + C_2\|\boldsymbol{\varphi}-\mathbf{S}f\|_{\mathbf{V}}\|(T-T_h)g)\|_W . \end{aligned} \tag{47}$$

Taking $\boldsymbol{\varphi} = \mathbf{S}_h f, v = T_h f$ in (47), we get

$$\begin{aligned} |(f,(T-T_h)g)| \lesssim\ & C_1\|(\mathbf{S}-\mathbf{S}_h)g\|_{\mathbf{H}}\|(\mathbf{S}-\mathbf{S}_h)f\|_{\mathbf{H}} \\ & + C_2\|(\mathbf{S}-\mathbf{S}_h)g\|_{\mathbf{V}}\|(T-T_h)f\|_W \\ & + C_2\|(\mathbf{S}-\mathbf{S}_h)f\|_{\mathbf{V}}\|(T-T_h)g\|_W \end{aligned} \tag{48}$$

In (48), replacing u_i for g and u_j for f, we have

$$|(T-T_h)u_i, u_j)| \lesssim C_1\left\|(\mathbf{S}-\mathbf{S}_h)|_{M_\lambda}\right\|^2_{G\to\mathbf{H}} + 2C_2\left\|(\mathbf{S}-\mathbf{S}_h)|_{M_\lambda}\right\|_{G\to\mathbf{V}}\left\|(T-T_h)|_{M_\lambda}\right\|_{G\to W}. \tag{49}$$

Substituting (49) into (42), we arrive at (41).

The mixed discretised source problem is well-posed and has a unique solution when h is small enough. Based on (34) and (40), we can obtain the following a priori error estimate.

For any $f \in L^2(\Omega)$, the following hold:

$$\|Tf - T_h f\|_0 \lesssim h^{r_0+r}\|Tf\|_{1+r}, \frac{1}{2} < r \le 2. \tag{50}$$

$$\|\mathbf{S}f - \mathbf{S}_h f\|_0 \lesssim h^r\|Tf\|_{1+r}, \frac{1}{2} < r \le 2. \tag{51}$$

$$\|\mathbf{S}f - \mathbf{S}_h f\|_{\mathbf{V}} \lesssim h^2\|Tf\|_3, \frac{1}{2} < r \le 2. \tag{52}$$

If $f \in M_\lambda$, then $Tf = \lambda_l^{-1} f$, and we can obtain the following estimate:

$$\left\|(T-T_h)|_{M_\lambda}\right\|_0 \lesssim h^{r_0+r}, \text{ if } M_\lambda \subset H^{1+r}(\Omega).$$

$$\left\|(\mathbf{S}-\mathbf{S}_h)|_{M_\lambda}\right\|_0 \lesssim h^r, \text{ if } M_\lambda \subset H^{1+r}(\Omega).$$

$$\left\|(\mathbf{S}-\mathbf{S}_h)|_{M_\lambda}\right\|_{\mathbf{V}} \lesssim h^2, \text{ if } M_\lambda \subset H^3(\Omega).$$

Lemma 6. Let $(\lambda_h, \boldsymbol{\mu}_h, p_h)$ be a mixed finite element eigenpair of (10), then there exists an eigenpair $(\lambda, \boldsymbol{\mu}, p)$ of (3), such that the following a priori error estimate holds:

$$h^{r_0}\|\boldsymbol{\mu}-\boldsymbol{\mu}_h\|_0 + \|p-p_h\|_0 \lesssim h^{r_0+r} \tag{53}$$

$$|\lambda - \lambda_h| \le h^{2(r_0+r)} \tag{54}$$

$$\|\boldsymbol{\mu}_h - \boldsymbol{\mu}\|_{\mathbf{V}} \lesssim h^2 \tag{55}$$

VI. NUMERICAL RESULTS

In this section, we report some numerical experiments to demonstrate the effectiveness of our approach. Considering the problem (1), our program is compiled under the iFEM package. Consider the following three test domains: the L-shaped domain $\Omega_L = [-1,1]\times[-1,1]\setminus[-1,0]\times[0,1]$, the square domain Ω_S with vertices at (0,1), (0,0), (1,0), (1,1), and the crack structure domain $\Omega_{SL} = (-1,1)^2 \setminus \{0 \le x \le 1, y = 0\}$.

Since the exact eigenvalues are unknown, we take the reference eigenvalue $\lambda_1 = 9.6397238440219$ in the L-shaped domain, the reference eigenvalue $\lambda_2 = 13.6079200746419$ in the S-shaped domain, and the reference eigenvalue $\lambda_3 = 15.1958966562930$ in the crack structure domain Ω_{SL}.

Table 1: The eigenvalue numerical solution results for Ω_L.

Domain	h	*dof*	λ_1	*Error*	*rate*
	1/4	867	9.60951249620301	0.030211347818890	1.496692724851000
	1/8	3459	9.62901800530627	0.010705838715630	1.395693349101040

Ω_L	1/16	138277	9.63565496166262	0.004068882359279	1.352817216478830
	1/32	55299	9.63813076773640	0.001593076285500	1.338811829426640
	1/64	221187	9.63909402749373	0.000629816528170	1.334796392731570
	1/128	884739	9.63947415450950	0.000249689512399	

Table 2: The eigenvalue numerical solution results for Ω_S .

Domain	h	*dof*	λ_1	*Error*	*rate*
Ω_S	1/4	1155	12.9623099078139	0.6456101668280	0.9669064906662
	1/8	4611	13.2776246907373	0.3302953839046	0.9616941545662
	1/16	18435	13.4383287169528	0.1695913576891	0.9704453816743
	1/32	73731	13.5213693819604	0.0865506926815	0.9851167695416
	1/64	294915	13.5641959774398	0.0437240972021	

Table 3: The eigenvalue numerical solution results for Ω_{SL} .

Domain	h	*dof*	λ_1	*Error*	*rate*
Ω_{SL}	1/4	1155	14.6945067366663	0.5013899197267	1.0468984857407
	1/8	4611	14.9532201401681	0.2426765162249	1.0162037899259
	1/16	18435	15.0759135976490	0.1199830587440	1.0141303628434
	1/32	73731	15.1364898411833	0.0594068152097	1.0296934701136
	1/64	294915	15.1667983542083	0.0290983021847	

VII. CONCLUSION

The general second-order elliptic eigenvalue problem has wide applications in practical problems. This paper presents a mixed finite element method for solving the general second-order eigenvalue problem. To derive the a priori error estimate, the key is to prove that the discrete operator T_h converges to the Dirichlet operator T in the sense of the $L^2(\Omega)$ norm, also $\|T-T_h\|_{0,\Omega}\to 0$ .Numerical experiments were conducted on three test domains: Ω_L , Ω_S and Ω_{SL} .The numerical results show that our method can achieve optimal convergence rates for the eigenvalues and obtain optimal error estimates for the eigenfunctions. The numerical experiments demonstrate the effectiveness of the algorithm.

REFERENCES

[1]. Babuška I. Error-bounds for finite element method[J]. Numerische Mathematik, 1971, 16(4): 322-333.

[2]. Brezzi F. On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers[J]. Publications des séminaires de mathématiques et informatique de Rennes, 1974 (S4): 1-26.

[3]. Falk R S, Osborn J E. Error estimates for mixed methods[J]. RAIRO. Analyse numérique, 1980, 14(3): 249-277.

[4]. 罗振东.混合有限元法基础及其应用[M].科学出版社,2006.

[5]. 杨一都著. 特征值问题有限元方法[M]. 科学出版社,2012.

[6]. Shen J. A block finite difference scheme for second-order elliptic problems with discontinuous coefficients[J]. SIAM journal on numerical analysis, 1996, 33(2): 686-706.

[7]. Talay D, Burger M, Wahlbin L B .Book Review: The finite element method for elliptic problems[J].Bulletin of the American Mathematical Society, 2003, 1(5):800-803.

[8]. Babuška I, Osborn J E. Finite element-Galerkin approximation of the eigenvalues and eigenvectors of selfadjoint problems[J]. Mathematics of computation, 1989, 52(186): 275-297.

[9]. Babuska I, Osborn J. Eigenvalue problems[J]. Handbook of numerical analysis., 1991, 2: 641.

[10]. Chatelin F. Spectral approximation of linear operators[M]. Society for Industrial and Applied Mathematics, 2011.

[11]. Mercier B, Osborn J, Rappaz J, et al. Eigenvalue approximation by mixed and hybrid methods[J]. Mathematics of Computation, 1981, 36(154): 427-453.

[12]. Osborn J E. Approximation of the eigenvalues of a nonselfadjoint operator arising in the study of the stability of stationary solutions of the Navier–Stokes equations[J]. SIAM Journal on Numerical Analysis, 1976, 13(2): 185-197.

[13]. 罗贤兵.二阶椭圆特征值问题的混合有限元法误差分析[D].贵州师范大学,2005.

[14]. Chen H, Jia S, Xie H. Postprocessing and higher order convergence for the mixed finite element approximations of the eigenvalue problem[J].Elsevier Science Publishers B. V. 2011, 61(4):615-629.DOI:10.1016/j.apnum.2010.12.007.

[15]. 黑圆圆.一种非标准的混合有限元法[D].吉林大学,2020.DOI:10.27162/d.cnki.gjlin.2020.002344.