MATHEMATICS OF COMPUTATION
Volume 79, Number 271, July 2010, Pages 1621–1646
S 0025-5718(09)02308-4
Article electronically published on September 17, 2009

OPTIMAL ERROR ESTIMATES IN JACOBI-WEIGHTED SOBOLEV SPACES FOR POLYNOMIAL APPROXIMATIONS ON THE TRIANGLE

HUIYUAN LI AND JIE SHEN

Abstract. Spectral approximations on the triangle by orthogonal polynomials are studied in this paper. Optimal error estimates in weighted semi-norms for both the L^2- and H_0^1-orthogonal polynomial projections are established by using the generalized Koornwinder polynomials and the properties of the Sturm-Liouville operator on the triangle. These results are then applied to derive error estimates for the spectral-Galerkin method for second- and fourth-order equations on the triangle. The generalized Koornwinder polynomials and approximation results developed in this paper will be useful for many other applications involving spectral and spectral-element approximations in triangular domains.

1. Introduction

Polynomial approximations on the triangle play an important role for the spectral-element methods and the *hp* finite-element methods in complex geometries [19, 7, 24, 23, 16, 15, 14]. Optimal error estimates for polynomial approximations were first established for Jacobi polynomials in one dimension [4, 9, 18, 11], and extended to multi-dimensional rectangular domains by a standard tensor product argument. However, only limited efforts have been devoted to establishing analogous results on triangular domains. Braess and Schwab [5] derived sharp error estimates for the orthogonal approximations on a simplex in the norm defined by a second-order self-adjoint differential operator under barycentric coordinates, but their results do not imply the standard H^k-error estimates ($k = 1, 2$), which are pivotal to the numerical analysis of spectral methods for PDEs. Schwab [21] (see also Canuto et al. [6]) derived error estimates in the usual Sobolev space by padding the triangle into a rectangle and using the standard approximation results. On the other hand, Guo and Wang [13] obtained some polynomial approximation results on the triangle in non-uniformly Jacobi-weighted Sobolev spaces by using the warped product based on the one-dimensional Jacobi approximations. However, their results are derived under a rather restrictive condition on the numbers of modes

Received by the editor August 12, 2008 and, in revised form, June 1, 2009.
2000 *Mathematics Subject Classification.* Primary 65N35, 65N22, 65F05, 35J05.
Key words and phrases. Orthogonal polynomials, Koornwinder polynomials, error estimate, spectral method.
The first author was partially supported by the NSFC grants 10601056, 10431050 and 60573023.
The second author was partially supported by the NFS grant DMS-0610646 and AFOSR FA9550-08-1-0416.

used in each direction. We should add that a fully tensorial spectral method using rational functions has been developed in [22].

The purpose of this paper is to establish optimal error estimates in weighted semi-norms for both the L^2- and H^1-norms for polynomial approximations on the triangle

$$T = \{(x, y) : 0 < x, y, x + y < 1\}. \tag{1.1}$$

These estimates share the same essential characteristics with the one-dimensional Jacobi approximation results and are directly applicable to spectral methods for partial differential equations on the triangle.

As in the one-dimensional case, properties of orthogonal polynomials on the triangle play important roles in their analysis and applications. There exist several families of classical orthogonal polynomials on the triangle T, among which the monomial basis [25, 10], Appell polynomials [3, 2] and Koornwinder polynomials [17] are particularly interesting. The monomial basis and Appell polynomials are bi-orthogonal to each other [8], while the Koornwinder polynomials are fully orthogonal, and its simplest family, the so-called Dubiner polynomials [7] have been used frequently in spectral-element methods (cf. [16]).

We point out that, similar to the classical Jacobi polynomials, which are defined for indices $\alpha, \beta > -1$, the classical orthogonal polynomials on the triangle are defined for a set of triple real numbers $\alpha_1, \alpha_2, \alpha_3 > -1$. Recently, Guo, Shen and Wang [11, 12] introduced generalized Jacobi polynomials for indices $\alpha \leq -1$ and/or $\beta \leq -1$. The use of generalized Jacobi polynomials not only simplifies the numerical analysis for the spectral approximations of differential equations, but also leads to very efficient numerical algorithms with well-conditioned linear systems. Therefore, we shall introduce a similar generalization to the orthogonal polynomials on the triangle.

It is also worthy to note that any families of orthogonal polynomials on the triangle are eigenfunctions of a second-order Sturm-Liouville operator, and it is well known that for a selfadjoint and positive definite differential operator, the eigenfunctions associated with different eigenvalues are mutually orthogonal. Inspired by this fact, Owens [20] constructed a Sturm-Liouville operator on T which possesses distinct eigenvalues, and then derived a family of orthogonal polynomials in [20]. More importantly, the use of the Sturm-Liouville operator usually leads to a direct and simple way to establish optimal approximation results in the norm induced by the Sturm-Liouville operator.

Therefore, we shall combine in this paper the two approaches: the Sturm-Liouville operator and generalized orthogonal polynomials. We shall first extend the Koornwinder polynomials (extension of the Jacobi polynomials to the triangle T) to negative integer indices, and then we construct an arbitrarily high-order Sturm-Liouville differential operator for orthogonal polynomials defined on T. Using this operator, we define non-uniformly Jacobi-weighted Sobolev spaces and derive optimal error estimates for L^2- and H_0^1-orthogonal approximations. These results can be directly used in numerical analysis and algorithm design of spectral methods for partial differential equations on the triangle.

The remainder of the paper is organized as follows. In §2, we recall the definition and basic approximation results for the generalized Jacobi polynomials, and then we introduce the Koornwinder polynomials and define the generalized Koornwinder polynomials on the triangle. The main approximation results are proved in §3.1. As

examples of applications, we provide in §4 the error estimates and implementation details for the spectral-Galerkin approximation to the second-order and the fourth-order partial differential equations. We present some illustrative numerical results in §5.

2. Koornwinder and generalized Koornwinder polynomials on the triangle

In this section, we recall the definition of Koornwinder polynomials which are the extension of the Jacobi polynomials on the triangle and form an orthogonal basis in weighted Sobolev space. We then extend the definition of Koornwinder polynomials to indices with negative integers.

2.1. **Notation and conventions.**

We describe below some notation and conventions which will be used throughout the paper.

Let ϖ be a generic positive weight function on a bounded domain Ω, which is not necessary to be $L^1(\Omega)$. Denote by $(u,v)_{\varpi,\Omega} := \int_\Omega uv\varpi d\Omega$ the inner product of $L^2_\varpi(\Omega)$ whose norm is denoted by $\|\cdot\|_{\varpi,\Omega}$. We use $H^m_\varpi(\Omega)$ and $H^m_{0,\varpi}(\Omega)$ to denote the usual weighted Sobolev spaces, whose norm and semi-norms are denoted by $\|u\|_{m,\varpi,\Omega}$ and $|u|_{m,\varpi,\Omega}$, respectively. In cases where no confusion would arise, ϖ (if $\varpi \equiv 1$) and Ω (if $\Omega = T$) may be dropped from the notation.

Let T be the reference triangle defined by (1.1). For any $\alpha = (\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}^3$, we introduce the following multi-index notation:

$$|\alpha| = \alpha_1 + \alpha_2 + \alpha_3, \quad \alpha^* = (\alpha_1 + \alpha_3, \alpha_2 + \alpha_3, \alpha_1 + \alpha_2), \tag{2.1}$$

and denote

$$\chi^\alpha := \chi^\alpha(x,y) = x^{\alpha_1} y^{\alpha_2} (1-x-y)^{\alpha_3}. \tag{2.2}$$

We define the weighted inner product and the corresponding $L^2_{\chi^\alpha}$-norm by

$$(u,v)_{\chi^\alpha} = \iint_T uv\chi^\alpha dxdy, \quad \|u\|_{\chi^\alpha} = \sqrt{(u,u)_{\chi^\alpha}}. \tag{2.3}$$

Let $\mathbb{Z}^+$, $\mathbb{N}_0$ and $\mathbb{Z}^-$ be the collections of the positive integers, the non-negative integers and the negative integers, respectively. We introduce the index sets

$$\aleph_+ = (-1, \infty), \quad \aleph_- = \mathbb{Z}^-, \quad \aleph = \aleph_- \cup \aleph_+. \tag{2.4}$$

For any $\alpha = (\alpha_1, \alpha_2, \alpha_3) \in \aleph^3$, we define

$$\hat{\alpha}_i = \begin{cases} -\alpha_i, & \alpha_i \in \aleph_-, \\ 0, & \alpha_i \in \aleph_+, \end{cases} \qquad \hat{\alpha} = (\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3). \tag{2.5}$$

For convenience, we abbreviate the partial differential operators $\dfrac{\partial}{\partial x}$ and $\dfrac{\partial}{\partial y}$ to ∂_x and ∂_y, respectively. For $n = (n_1, n_2, n_3) \in \mathbb{N}_0^3$, we define the differential operator

$$D^n = \partial_x^{n_1} \partial_y^{n_2} (\partial_y - \partial_x)^{n_3}. \tag{2.6}$$

Different from the standard two-dimensional differential operators, we have introduced the third directional derivative $\partial_y - \partial_x$ in the definition of D^n since this direction is parallel to the hypotenuse and plays as important a role as the other two directions (x- and y-directions).

We denote by c a generic positive constant independent of any function and of any discretization parameters. We use the expression $A \lesssim B$ to mean that $A \le cB$.

Let $\mathbb{P}_N(\Omega)$ be the collection of all the algebraic polynomials on Ω with total degree no greater than N, i.e.,

$$\mathbb{P}_N = \mathbb{P}_N(T) = \left\{x^k y^l : k, l \in \mathbb{N}_0,\ k+l \le N\right\}. \tag{2.7}$$

Definition 2.1. We shall say that $q \in \mathbb{P}_N$ is an orthogonal polynomial of degree N with respect to $(\cdot,\cdot)_{\chi^\alpha}$ if

$$(q,p)_{\chi^\alpha} = 0, \quad \forall p \in \mathbb{P}_{N-1},$$

and denote it by $q \in \mathbb{R}_N^\alpha$.

2.2. **Jacobi and generalized Jacobi polynomials.** The classical Jacobi polynomials $J_k^{\alpha_1,\alpha_2}(\zeta)$, $k \ge 0$ with $\alpha_1, \alpha_2 > -1$ are mutually orthogonal with respect to the Jacobi weight function $\omega^{\alpha_1,\alpha_2} := \omega^{\alpha_1,\alpha_2}(\zeta) = (1-\zeta)^{\alpha_1}(1+\zeta)^{\alpha_2}$ on $I = (-1,1)$,

$$\int_{-1}^{1} J_m^{\alpha_1,\alpha_2}(\zeta) J_n^{\alpha_1,\alpha_2}(\zeta)\omega^{\alpha_1,\alpha_2}(\zeta)d\zeta = h_n^{\alpha_1,\alpha_2}\delta_{m,n}, \quad m,n \ge 0, \tag{2.8}$$

where $\delta_{m,n}$ is the Kronecker delta, and

$$h_n^{\alpha_1,\alpha_2} := \left\|J_n^{\alpha_1,\alpha_2}\right\|^2_{\omega^{\alpha_1,\alpha_2},I} = \frac{2^{\alpha_1+\alpha_2+1}\Gamma(n+\alpha_1+1)\Gamma(n+\alpha_2+1)}{(2n+\alpha_1+\alpha_2+1)\Gamma(n+1)\Gamma(n+\alpha_1+\alpha_2+1)}. \tag{2.9}$$

The classical Jacobi polynomials were originally defined for $\alpha_1, \alpha_2 > -1$. Recently Guo, Shen and Wang [11] extended the definition of Jacobi polynomials to allow α_1 and/or α_2 to be negative integers. The generalized Jacobi polynomials with negative indices not only simplify the numerical analysis for the spectral approximations of differential equations, but also lead to very efficient numerical algorithms [11, 12]. We recall that, given $(\alpha_1,\alpha_2) \in \aleph^2 \backslash \aleph_+^2$, the generalized Jacobi polynomials (still denoted by $J_n^{\alpha_1,\alpha_2}$) for $n \ge n_0 := \hat{\alpha}_1 + \hat{\alpha}_2$ can be defined as follows:

$$\begin{aligned} J_n^{\alpha_1,\alpha_2}(\zeta) &= \begin{cases} \left(\frac{\zeta-1}{2}\right)^{-\alpha_1}\left(\frac{\zeta+1}{2}\right)^{-\alpha_2} J_{n+\alpha_1+\alpha_2}^{-\alpha_1,-\alpha_2}(\zeta), & \alpha_1 \in \aleph_-,\ \alpha_2 \in \aleph_-, \\ \frac{\Gamma(n+\alpha_1+1)\Gamma(n+\alpha_2+1)}{\Gamma(n+1)\Gamma(n+\alpha_1+\alpha_2+1)}\left(\frac{\zeta-1}{2}\right)^{-\alpha_1} J_{n+\alpha_1}^{-\alpha_1,\alpha_2}(\zeta), & \alpha_1 \in \aleph_-,\ \alpha_2 \in \aleph_+, \\ \frac{\Gamma(n+\alpha_1+1)\Gamma(n+\alpha_2+1)}{\Gamma(n+1)\Gamma(n+\alpha_1+\alpha_2+1)}\left(\frac{\zeta+1}{2}\right)^{-\alpha_2} J_{n+\alpha_2}^{\alpha_1,-\alpha_2}(\zeta), & \alpha_1 \in \aleph_+,\ \alpha_2 \in \aleph_-, \end{cases} \\ &= c_n^{\alpha_1,\alpha_2}\omega^{\hat{\alpha}_1,\hat{\alpha}_2}(\zeta) J_{n-n_0}^{\alpha_1+2\hat{\alpha}_1,\alpha_2+2\hat{\alpha}_2}(\zeta). \end{aligned} \tag{2.10}$$

It can be readily checked that (2.8) still holds for any $m,n \ge n_0$. We also point out that the generalized Jacobi polynomials defined by (2.10) satisfy most essential recurrence relations of the classic Jacobi polynomials (cf. [22] and Appendix A).

We now define the orthogonal projection $\pi_N^{\alpha_1,\alpha_2}$: $L^2_{\omega^{\alpha_1,\alpha_2}}(I) \mapsto \mathbb{P}_N(I) \cap L^2_{\omega^{\alpha_1,\alpha_2}}(I)$ by

$$(\pi_N^{\alpha_1,\alpha_2}u - u, v)_{\omega^{\alpha_1,\alpha_2},I} = 0, \quad \forall v \in \mathbb{P}_N(I) \cap L^2_{\omega^{\alpha_1,\alpha_2}}(I). \tag{2.11}$$

Note that for $\alpha_1, \alpha_2 > -1$, we have $\mathbb{P}_N(I) \cap L^2_{\omega^{\alpha_1,\alpha_2}}(I) = \mathbb{P}_N(I)$, while for α_1 and/or α_2 being negative integers, certain boundary conditions are involved for any function $v \in \mathbb{P}_N(I) \cap L^2_{\omega^{\alpha_1,\alpha_2}}(I)$. For instance, $\mathbb{P}_N(I) \cap L^2_{\omega^{-1,-1}}(I) = \mathbb{P}_N(I) \cap H_0^1(I)$.

We introduce the Jacobi-weighted Sobolev space $B^{\varrho}_{\alpha_1,\alpha_2}(I)$,

$$B^{\varrho}_{\alpha_1,\alpha_2}(I) := \Big\{u:\ u \text{ is measurable on } I \text{ and } \|u\|_{B^{\varrho}_{\alpha_1,\alpha_2}(I)} < \infty\Big\},$$
$$\|u\|_{B^{\varrho}_{\alpha_1,\alpha_2}(I)} = \Big(\sum_{k=0}^{\varrho} \big\|\partial_\zeta^k u\big\|^2_{\omega^{\alpha_1+k,\alpha_2+k}}\Big)^{\frac{1}{2}}.$$

Then the following theorem is a direct extension of the same result for $\alpha_1, \alpha_2 \in \aleph_+$.

Theorem 2.1 (cf. [11]). *Let $\alpha_1, \alpha_2 \in \aleph$. Then for any $u \in B^{\varrho}_{\alpha_1,\alpha_2}(I)$ and $0 \le l \le \varrho \le N$,*

$$\|\pi_N^{\alpha_1,\alpha_2} u - u\|_{B^{l}_{\alpha_1,\alpha_2}(I)} \lesssim N^{l-\varrho} \big|\partial_\zeta^{\varrho} u\big|_{\omega^{\alpha_1+\varrho,\alpha_2+\varrho}}. \tag{2.12}$$

We shall introduce below the generalized orthogonal polynomials on the triangle and derive corresponding approximation results on the triangle similar to the above theorem.

2.3. **Koornwinder and generalized Koornwinder polynomials.** Consider the following one-to-one transformation from the reference square $Q = (-1,1)^2$ onto the triangle T,

$$x = \frac{1-\xi}{2}\frac{1-\eta}{2}, \qquad y = \frac{1+\xi}{2}\frac{1-\eta}{2} \tag{2.13}$$

with the inverse transform

$$\zeta = \frac{y-x}{y+x}, \quad \eta = 1-2x-2y. \tag{2.14}$$

This *collapsed* transformation maps polynomials in T to polynomials in Q.

We would like to point out that the transform (2.13) can be related to the one used in [7, 16] by a rotation

$$(x, y, 1-x-y) \to (y, 1-x-y, x).$$

The transform (2.13) collapses the edge: $\eta = 1, -1 < \xi < 1$, of the rectangle Q into the vertex $(0,0)$ instead of $(0,1)$. Thus, the formulas under the transform (2.13) are more symmetric with respect to the variables x and y.

For $\alpha \in \aleph_+^3$, Koornwinder polynomials are two-variable analogues of Jacobi polynomials defined through the collapsed transformation (2.13),

$$\begin{aligned}\mathcal{J}_l^{\alpha}(x,y) &= J_{l_1}^{\alpha_1,\alpha_2}(\xi)\Big(\frac{1-\eta}{2}\Big)^{l_1} J_{l_2}^{2l+\alpha_1+\alpha_2+1,\alpha_3}(\eta)\\ &= (y+x)^{l_1} J_{l_1}^{\alpha_1,\alpha_2}\left(\frac{y-x}{y+x}\right) J_{l_2}^{2l_1+\alpha_1+\alpha_2+1,\alpha_3}(1-2x-2y), \quad l \in \mathbb{N}_0^2.\end{aligned} \tag{2.15}$$

By (2.13) and (2.9), we get that

$$\begin{aligned}\iint_T \mathcal{J}_l^{\alpha}(x,y)\mathcal{J}_k^{\alpha}(x,y)\chi^{\alpha}dxdy &= \frac{1}{4}\iint_Q \Big(\frac{1-\xi}{2}\Big)^{\alpha_1}\Big(\frac{1+\xi}{2}\Big)^{\alpha_2}\Big(\frac{1-\eta}{2}\Big)^{l_1+k_2+\alpha_1+\alpha_2+1}\\ &\times\Big(\frac{1+\eta}{2}\Big)^{\alpha_3}\mathcal{J}_{l_1}^{\alpha_1,\alpha_2}(\xi)\mathcal{J}_{k_1}^{\alpha_1,\alpha_2}(\xi)\mathcal{J}_{l_2}^{2l_1+\alpha_1+\alpha_2+1,\alpha_3}(\eta)\mathcal{J}_{k_2}^{2k_1+\alpha_1+\alpha_2+1,\alpha_3}(\eta)d\xi d\eta\\ &= \gamma_l^{\alpha}\delta_{l,k}, \qquad l,k, \in \mathbb{N}_0^2,\end{aligned} \tag{2.16}$$

where

$$\gamma_l^\alpha = \frac{\Gamma(l_1+\alpha_1+1)\Gamma(l_1+\alpha_2+1)\Gamma(l_2+\alpha_3+1)\Gamma(|l|+l_1+\alpha_1+\alpha_2+2)}{\Gamma(l_1+1)\Gamma(l_2+1)\Gamma(l_1+\alpha_1+\alpha_2+1)\Gamma(|l|+l_1+|\alpha|+2)} \times \frac{1}{(2l_1+\alpha_1+\alpha_2+1)(2|l|+|\alpha|+2)}.$$

These orthogonal polynomials have been used in developing spectral-element methods for complex geometry; e.g., the basis functions used in [7, 16] become simply the basic family $\mathcal{J}_{l_1,l_2}^{0,0,0}$ through another affine transform from T to T.

Since the generalized Jacobi polynomials $J_l^{\alpha_1,\alpha_2}$ are well defined for α_1 and α_2 to be negative integers, we can still use (2.15) to define the generalized Koornwinder polynomials $\mathcal{J}_l^\alpha$. More precisely, for any $\alpha \in \aleph^3 \backslash \aleph_+^3$, by using the definition of the generalized Jacobi polynomials, we define

(2.17)

$$\begin{aligned}\mathcal{J}_l^\alpha(x,y) &= (y+x)^{l_1} J_{l_1}^{\alpha_1,\alpha_2}\left(\frac{y-x}{y+x}\right) J_{l_2}^{2l_1+\alpha_1+\alpha_2+1,\alpha_3}(1-2x-2y)\\ &= (y+x)^{l_1-\hat\alpha_1-\hat\alpha_2} J_{l_1-\hat\alpha_1-\hat\alpha_2}^{\alpha_1+2\hat\alpha_1,\alpha_2+2\hat\alpha_2}\left(\frac{y-x}{y+x}\right) J_{l_2-\hat\alpha_3}^{2l_1+\alpha_1+\alpha_2+1,\alpha_3+2\hat\alpha_3}(1-2x-2y)\\ &\quad \times c\, x^{\hat\alpha_1} y^{\hat\alpha_2}(1-x-y)^{\hat\alpha_3}\\ &= c\,\chi^{\hat\alpha}(x,y)\mathcal{J}_{l_1-\hat\alpha_1-\hat\alpha_2,l_2-\hat\alpha_3}^{\alpha+2\hat\alpha}(x,y).\end{aligned}$$

Hence, $\mathcal{J}_l^\alpha$ is well defined for $l \in \mathbb{N}_0^2$ with $l_1 \ge \hat\alpha_1+\hat\alpha_2$ and $l_2 \ge \hat\alpha_3$, and it can be checked that $\{\mathcal{J}_l^\alpha\}_{l_1\ge\hat\alpha_1+\hat\alpha_2,\, l_2\ge\hat\alpha_3}$ are mutually orthogonal in $L^2_{\chi^\alpha}(T)$ and form a complete basis for $L^2_{\chi^\alpha}(T)$.

3. Polynomial approximations on the triangle

In this section, we study the orthogonal polynomial approximations on the triangle T. For this purpose, we consider the eigenproblem of a selfadjoint partial differential equation for both the classic and the generalized orthogonal polynomials. Then we shall give an optimal error estimate in Jacobi-weighted Sobolev spaces.

3.1. Sturm-Liouville operator for classical orthogonal polynomials.

Lemma 3.1. *Let $\alpha \in \aleph_+^3$ and $n \in \mathbb{N}_0^3$. Then, the differential operator*

$$(-1)^{|n|}\chi^{-\alpha} D^n \chi^{n^*+\alpha} D^n \quad (n^* \text{ and } D^n \text{ are defined in (2.1) and (2.6))}$$

is selfadjoint and positive definite with respect to the inner product $(\cdot,\cdot)_{\chi^\alpha}$. It maps $\mathbb{P}_m$ into $\mathbb{P}_m$, and $\mathbb{R}_m^\alpha$ into $\mathbb{R}_m^\alpha$.

Proof. When applying $(-1)^{|n|}\chi^{-\alpha} D^n \chi^{n^*+\alpha} D^n$ to a polynomial p, the augment of the degree of p by multiplying by $\chi^{n^*+\alpha}$ and then $\chi^{-\alpha}$ is compensated by two differentiations since $|n^*| = 2|n|$. Thus we deduce that $(-1)^{|n|}\chi^{-\alpha} D^n \chi^{n^*+\alpha} D^n$ maps $\mathbb{P}_m$ into $\mathbb{P}_m$.

By Leibniz's rule, we derive that for a sufficiently smooth function g,

$$D^m\big(\chi^{n^*+\alpha} g\big) = \sum_{k_3=0}^{m_3}\sum_{k_2=0}^{m_2}\sum_{k_1=0}^{m_1}\binom{m_1}{k_1}\binom{m_2}{k_2}\binom{m_3}{k_3}\big(D^k\chi^{n^*+\alpha}\big)\big(D^{m-k}g\big).$$

Furthermore, it is easy to see that for any $\alpha_1, \alpha_2, \alpha_3 > -1$,

$$\begin{cases} D^k \chi^{n^*+\alpha}(0,y) = 0, & \text{if } k_1 + k_3 < n_1 + n_3, \\ D^k \chi^{n^*+\alpha}(x,0) = 0, & \text{if } k_2 + k_3 < n_2 + n_3, \\ D^k \chi^{n^*+\alpha}(x,1-x) = 0, & \text{if } k_1 + k_2 < n_1 + n_2, \end{cases}$$

which states that for any $\alpha \in \aleph_+^3$ and $0 \le m_i \le n_i$ with $i = 1, 2, 3$,

$$\begin{cases} D^m\big(\chi^{n^*+\alpha} g\big)(0,y) = 0 \text{ and } D^m\big(\chi^{n^*+\alpha} g\big)(1-y,y) = 0, & \text{if } m_1 < n_1, \\ D^m\big(\chi^{n^*+\alpha} g\big)(x,0) = 0 \text{ and } D^m\big(\chi^{n^*+\alpha} g\big)(x,1-x) = 0, & \text{if } m_2 < n_2, \\ D^m\big(\chi^{n^*+\alpha} g\big)(0,y) = 0 \text{ and } D^m\big(\chi^{n^*+\alpha} g\big)(x,0) = 0, & \text{if } m_3 < n_3. \end{cases}$$

Therefore no boundary terms occur when performing integration by parts in the following expression:

$$(-1)^{|n|} \iint_T D^n(\chi^{n^*+\alpha} D^n f)\, g \, dxdy = \iint_T \chi^{n^*+\alpha} (D^n f)\, (D^n g)\, dxdy. \tag{3.1}$$

This implies that the operator $(-1)^{|n|}\chi^{-\alpha} D^n \chi^{n^*+\alpha} D^n$ is positive definite with respect to the inner product $(\cdot,\cdot)_{\chi^\alpha}$. From the symmetry of (3.1) we obtain

$$\big((-1)^{|n|}\chi^{-\alpha} D^n(\chi^{n^*+\alpha} D^n f),\, g\big)_{\chi^\alpha} = \big(f,\, (-1)^{|n|}\chi^{-\alpha} D^n(\chi^{n^*+\alpha} D^n g)\big)_{\chi^\alpha}, \tag{3.2}$$

which clearly states that $(-1)^{|n|}\chi^{-\alpha} D^n \chi^{n^*+\alpha} D^n$ is selfadjoint with respect to $(\cdot,\cdot)_{\chi^\alpha}$.

Finally let $p \in \mathbb{R}_m^\alpha$ and $q \in \mathbb{P}_{m-1}$. Since $\chi^{-\alpha} D^n \chi^{n^*+\alpha} D^n q \in \mathbb{P}_{m-1}$ and $p \in \mathbb{P}_{m-1}^\perp$, we deduce from (3.2) that

$$\iint_T q \cdot (\chi^{-\alpha} D^n \chi^{n^*+\alpha} D^n p) \cdot \chi^\alpha \, dxdy = \iint_T p \cdot (\chi^{-\alpha} D^n \chi^{n^*+\alpha} D^n q) \cdot \chi^\alpha \, dxdy = 0.$$

Hence, we conclude that $(-1)^{|n|}\chi^{-\alpha} D^n \chi^{n^*+\alpha} D^n$ maps $\mathbb{R}_m^\alpha$ into $\mathbb{R}_m^\alpha$. □

As an analogue of the multinomial theorem [1], we have the following lemma.

Lemma 3.2. *Let* $e_1 = (1,0,1)$, $e_2 = (0,1,1)$ *and* $e_3 = (1,1,0)$. *Then for any* $\alpha \in \mathbb{R}^3$,

$$\begin{aligned} \sum_{|n|=\varrho+1} \frac{(\varrho+1)!}{n_1!n_2!n_3!} D^n \chi^{n^*+\alpha} D^n &= \partial_x \Big(\sum_{|n|=\varrho} \frac{\varrho!}{n_1!n_2!n_3!} D^n \chi^{n^*+\alpha+e_1} D^n \Big) \partial_x \\ &\quad + \partial_y \Big(\sum_{|n|=\varrho} \frac{\varrho!}{n_1!n_2!n_3!} D^n \chi^{n^*+\alpha+e_2} D^n \Big) \partial_y \\ &\quad + (\partial_y - \partial_x) \Big(\sum_{|n|=\varrho} \frac{\varrho!}{n_1!n_2!n_3!} D^n \chi^{n^*+\alpha+e_3} D^n \Big) (\partial_y - \partial_x). \end{aligned} \tag{3.3}$$

Now we define the differential operator

$$\mathcal{L}_\alpha^\varrho = (-1)^\varrho \chi^{-\alpha} \sum_{|n|=\varrho} \frac{\varrho!}{n_1!n_2!n_3!} D^n \chi^{n^*+\alpha} D^n,$$

which can be regarded as a 2ϱ-th order Sturm-Liouville operator on the triangle.

Theorem 3.1. *Let $\alpha \in \aleph_+^3$ and $\varrho \in \mathbb{Z}^+$. The operator $\mathcal{L}_\alpha^\varrho$ is selfadjoint and*

$$\mathcal{L}_\alpha^\varrho\, p = \mu_{m,|\alpha|}^\varrho\, p \quad \text{for all } p \in \mathbb{R}_m^\alpha, \tag{3.4}$$

where

$$\mu_{m,|\alpha|}^\varrho = \frac{\Gamma(m+1)}{\Gamma(m-\varrho+1)} \frac{\Gamma(m+\varrho+|\alpha|+2)}{\Gamma(m+|\alpha|+2)}.$$

Proof. Owing to the linearity of $\mathcal{L}_\alpha^\varrho$ and Lemma 3.1, it suffices to prove that

$$\mathcal{L}_\alpha^\varrho x^k y^l = \mu_{k+l,|\alpha|}^\varrho x^k y^l \pmod{\mathbb{P}_{k+l-1}}, \quad k,l \ge 0. \tag{3.5}$$

Now we prove (3.5) by induction. Although the proof for $|n| = 1$ under the barycentric coordinates is given in [5], for the sake of completeness, we provide a brief proof below. In fact, we get by a mechanical calculation that

$$\begin{aligned}
\mathcal{L}_\alpha^1(x^k y^l) &= -\chi^{-\alpha}\partial_x\big(x(1-x-y)\chi^\alpha \partial_x(x^k y^l)\big) - \chi^{-\alpha}\partial_y\big(y(1-x-y)\chi^\alpha \partial_y(x^k y^l)\big) \\
&\quad - \chi^{-\alpha}(\partial_y - \partial_x)\big(xy\chi^\alpha(\partial_y - \partial_x)(x^k y^l)\big) \\
&= -kx^{k-1}y^l\big((\alpha_1+k) - (\alpha_1+\alpha_3+k+1)x - (\alpha_1+k)y\big) \\
&\quad - lx^k y^{l-1}\big(\alpha_2+l) - (\alpha_2+\alpha_3+l+1)y - (\alpha_2+l)x\big) \\
&\quad - x^{k-1}y^{l-1}\big(l(\alpha_2+l)x^2 + k(\alpha_1+k)y^2 \\
&\qquad\qquad - (k(\alpha_2+l+1) + l(k+\alpha_1+1))xy\big) \\
&\equiv (k+l)(k+l+2+|\alpha|)x^k y^l \pmod{\mathbb{P}_{k+l-1}}.
\end{aligned}$$

Now assume (3.5) holds for $|n| = \varrho$. Then by a straightforward calculation (mod $\mathbb{P}_{k+l-1}$), we have

$$\begin{aligned}
&\chi^{-\alpha}\sum_{|n|=\varrho}\frac{(-1)^\varrho \varrho!}{n_1!n_2!n_3!}D^n\partial_x\Big(x(1-x-y)\chi^{n^*+\alpha}D^n\partial_x(x^k y^l)\Big) \\
&= k\chi^{-\alpha}\sum_{|n|=\varrho}\frac{(-1)^\varrho \varrho!}{n_1!n_2!n_3!}D^n\Big(\chi^{n^*+\alpha}\big(x(1-x-y)\partial_x + (n_1+n_3+\alpha_1+1)(1-x-y) \\
&\qquad - (n_1+n_2+\alpha_3+1)x\big)D^n(x^{k-1}y^l)\Big) \\
&= k(k-1)\chi^{-\alpha}\sum_{|n|=\varrho}\frac{(-1)^\varrho \varrho!}{n_1!n_2!n_3!}D^n\Big(x(1-x-y)\chi^{n^*+\alpha}D^n(x^{k-2}y^l)\Big) \\
&\quad + k(n_1+n_3+\alpha_1+1)\chi^{-\alpha}\sum_{|n|=\varrho}\frac{(-1)^\varrho \varrho!}{n_1!n_2!n_3!}D^n\Big(\chi^{n^*+\alpha}(1-x-y)D^n(x^{k-1}y^l)\Big) \\
&\quad - k(n_1+n_2+\alpha_3+1)\chi^{-\alpha}\sum_{|n|=\varrho}\frac{(-1)^\varrho \varrho!}{n_1!n_2!n_3!}D^n\Big(\chi^{n^*+\alpha}xD^n(x^{k-1}y^l)\Big) \\
&\equiv k\big((k-1)\mu_{k+l-2,|\alpha|+2}^\varrho + (n_1+n_3+\alpha_1+1)\mu_{k+l-1,|\alpha|+1}^\varrho\big)x^{k-1}y^l(1-x-y) \\
&\quad - k(n_1+n_2+\alpha_3+1)\mu_{k+l-1,|\alpha|+1}^\varrho x^k y^l.
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\chi^{-\alpha}\sum_{|n|=\varrho}\frac{(-1)^{\varrho}\varrho!}{n_1!n_2!n_3!}D^n\Big(\partial_y y(1-x-y)\chi^{n^*+\alpha}D^n\partial_y(x^k y^l)\Big)\\
&\equiv l\big((l-1)\mu^{\varrho}_{k+l-2,|\alpha|+2}+(n_2+n_3+\alpha_2+1)\mu^{\varrho}_{k+l-1,|\alpha|+1}\big)x^k y^{l-1}(1-x-y)\\
&\quad -l(n_1+n_2+\alpha_3+1)\mu^{\varrho}_{k+l-1,|\alpha|+1}x^k y^l
\end{aligned}$$

and

$$\begin{aligned}
&\chi^{-\alpha}\sum_{|n|=\varrho}\frac{(-1)^{\varrho}\varrho!}{n_1!n_2!n_3!}D^n\Big((\partial_y-\partial_x)xy\chi^{n^*+\alpha}D^n(\partial_y-\partial_x)(x^k y^l)\Big)\\
&\quad\equiv l\big((l-1)\mu^{\varrho}_{k+l-2,|\alpha|+2}+(n_2+n_3+\alpha_2+1)\mu^{\varrho}_{k+l-1,|\alpha|+1}\big)x^{k+1}y^{l-1}\\
&\qquad +k\big((k-1)\mu^{\varrho}_{k+l-2,|\alpha|+2}+(n_1+n_3+\alpha_1+1)\mu^{\varrho}_{k+l-1,|\alpha|+1}\big)x^{k-1}y^{l+1}\\
&\qquad -\big(2kl\mu^{\varrho}_{k+l-2,|\alpha|+2}+(k(n_2+n_3+\alpha_2+1)\\
&\qquad\qquad +l(n_1+n_3+\alpha_1+1))\mu^{\varrho}_{k+l-1,|\alpha|+1}\big)x^k y^l.
\end{aligned}$$

Finally we deduce from the above arguments and Lemma 3.2 that

$$\begin{aligned}
&(-1)^{\varrho+1}\chi^{-\alpha}\sum_{|n|=\varrho+1}\frac{(\varrho+1)!}{n_1!n_2!n_3!}D^n\chi^{n^*+\alpha}D^n\chi^n\\
&\quad\equiv (k+l)\big[(k+l-1)\mu^{\varrho}_{k+l-2,|\alpha|+2}+(2\varrho+|\alpha|+3)\mu^{\varrho}_{k+l-1,|\alpha|+1}\big]x^k y^l\\
&\quad=\mu^{\varrho+1}_{k+l,|\alpha|}x^k y^l,
\end{aligned}$$

which completes the proof. □

3.2. **Sturm-Liouville operator for generalized orthogonal polynomials.** We first define the space of generalized orthogonal polynomials on the triangle T for any $\alpha\in\aleph^3$,

$$\mathbb{P}^{\alpha}_m=\mathbb{P}_m\cap L^2_{\chi^\alpha}(T)=\{\mathcal{J}^{\alpha}_l(x,y):\ l_1\ge\hat\alpha_1+\hat\alpha_2,\ l_2\ge\hat\alpha_3,\ l_1+l_2\le m\}. \tag{3.6}$$

It is obvious that $\mathbb{P}^{\alpha}_m=\mathbb{P}_m$ if $\alpha\in\aleph^3_+$, and $(-1)^{|n|}D^n(\chi^{n^*+\alpha}D^n)$ maps $\mathbb{P}^{\alpha}_m$ into $\mathbb{P}^{\alpha}_m$, noting that any function $f\in\mathbb{P}^{\alpha}_m$ will vanish on the corresponding boundary/boundaries if $\alpha\in\aleph^3\backslash\aleph^3_+$. Thus following the same way as in Lemma 3.1, one can prove that $(-1)^{|n|}D^n(\chi^{n^*+\alpha}D^n)$ is selfadjoint and positive definite with respect to χ^α in $\mathbb{P}^{\alpha}_m$. This, in return, shows that $(-1)^{|n|}D^n(\chi^{n^*+\alpha}D^n)$ maps $\mathbb{R}^{\alpha}_m$ into $\mathbb{R}^{\alpha}_m$.

Since $\{\mathcal{J}^{\alpha}_l(x,y)\}_{l_1\ge\hat\alpha_1+\hat\alpha_2,\ l_2\ge\hat\alpha_3}$ form a basis of $L^2_{\chi^\alpha}(T)$, we have the following lemma.

Lemma 3.3. *Let $\alpha\in\aleph^3$ and $n\in\mathbb{N}^3_0$. The differential operator*

$$(-1)^{|n|}\chi^{-\alpha}D^n\chi^{n^*+\alpha}D^n$$

is selfadjoint and positive definite with respect to the inner product $(\cdot,\cdot)_{\chi^\alpha}$. It maps $\mathbb{P}^{\alpha}_m$ into $\mathbb{P}^{\alpha}_m$ and $\mathbb{R}^{\alpha}_m$ into $\mathbb{R}^{\alpha}_m$.

By using a similar induction argument as in the proof of Theorem 3.1, but acting on $\chi^{l+\hat\alpha}(x,y)$ instead of $x^{l_1}y^{l_2}$, we can prove the following result.

Theorem 3.2. *Let $\alpha \in \aleph^3$ and $\varrho \in \mathbb{Z}^+$. The operator $\mathcal{L}_\alpha^\varrho$ is selfadjoint and*

$$\mathcal{L}_\alpha^\varrho\, p = \mu_{m,|\alpha|}^\varrho\, p \quad \text{for all } p \in \mathbb{R}_m^\alpha, \tag{3.7}$$

where $m \geq |\hat{\alpha}|$ and $\mu_{m,\alpha}^\varrho$ is the same as in Theorem 3.1.

3.3. **Polynomial approximations in Jacobi-weighted Sobolev space.** For any $M \in \mathbb{N}_0$, define the orthogonal projection $P_M^\alpha : L^2_{\chi^\alpha}(T) \to \mathbb{P}_M^\alpha$ such that for any $u \in L^2_{\chi^\alpha}(T)$,

$$(P_M^\alpha u - u, v)_{\chi^\alpha} = 0, \quad \forall v \in \mathbb{P}_M^\alpha.$$

The main purpose of this subsection is to obtain an optimal error estimate of the orthogonal projection. We resort to the space decomposition for our purpose. Note that

$$\mathbb{P}_M^\alpha = \bigoplus_{m=|\hat{\alpha}|}^{M} \mathbb{R}_m^\alpha.$$

We now define the orthogonal projection $R_m^\alpha : L^2_{\chi^\alpha}(T) \to \mathbb{R}_m^\alpha$ by

$$(R_m^\alpha u - u, v)_{\chi^\alpha} = 0, \quad \forall v \in \mathbb{R}_m^\alpha.$$

For notational simplicity, we denote $u_m = R_m^\alpha u$. Then we have the following identities:

$$u = \sum_{m \geq |\hat{\alpha}|} u_m, \qquad P_M^\alpha u = \sum_{m \leq M} u_m.$$

For integer $\varrho \in \mathbb{N}_0$, we define the Jacobi-weight Sobolev space $B_\alpha^\varrho(T)$ and its corresponding norm and semi-norm,

$$B_\alpha^\varrho(T) = \{u : \ u \text{ is measurable and } \|u\|_{B_\alpha^\varrho} < \infty\},$$
$$\|u\|_{B_\alpha^\varrho}^2 = \sum_{\rho=0}^{\varrho} |u|_{B_\alpha^\rho}^2, \qquad |u|_{B_\alpha^\rho}^2 = \sum_{|n|=\rho} \frac{\rho!}{n_1!n_2!n_3!} \|D^n u\|_{\chi^{n^*+\alpha}}^2.$$

We start with the following useful lemma.

Lemma 3.4. *Let $n \in \mathbb{N}_0^3$, $\alpha \in \aleph^3$ and $\alpha + n^* \in \aleph^3$. Then D^n maps $\mathbb{R}_m^\alpha$ onto $\mathbb{R}_{m-|n|}^{\alpha+n^*}$. Moreover, the differential operator D^n and the projection operator commute in the sense that*

$$D^n P_M^\alpha = P_{M-|n|}^{\alpha+n^*} D^n. \tag{3.8}$$

Proof. It is easy to see from (A.7) that $\partial_x \mathcal{J}_l^\alpha$ can be expressed as a linear combination of $\mathcal{J}_{k_1,k_2}^{\alpha_1+1,\alpha_2,\alpha_3+1}$ with $|k| = |l| - 1$, which states that the differential operator ∂_x maps $\mathbb{R}_m^\alpha$ into $\mathbb{R}_{m-1}^{\alpha_1+1,\alpha_2,\alpha_3+1}$. On the other hand, we get from (A.7) that

$$\mathcal{J}_{l_1,l_2}^{\alpha_1+1,\alpha_2,\alpha_3+1}(x,y) = -\frac{2l_1+\alpha_1+\alpha_2+3}{(l_1+\alpha_2+1)(|l|+l_1+\alpha_1+\alpha_2+3)} \partial_x \mathcal{J}_{l_1+1,l_2}^{\alpha_1,\alpha_2,\alpha_3}(x,y)$$
$$- \frac{(l_1+\alpha_1+\alpha_2+2)(|l|+l_1+|\alpha|+4)}{(l_1+\alpha_2+1)(|l|+l_1+\alpha_1+\alpha_2+3)} \mathcal{J}_{l_1+1,l_2-1}^{\alpha_1+1,\alpha_2,\alpha_3+1}(x,y).$$

Noting that $\mathcal{J}^{\alpha_1+1,\alpha_2,\alpha_3+1}_{|l|-\widehat{(\alpha_3+1)}+1,\widehat{(\alpha_3+1)}-1}=0$ and $(l_1+\alpha_2+1)(|l|+l_1+\alpha_1+\alpha_2+3)>0$ for any $l_1\geq\widehat{(\alpha_1+1)}+\hat{\alpha}_2$ and $l_2\geq\widehat{(\alpha_3+1)}$, we further get that

$$\mathcal{J}^{\alpha_1+1,\alpha_2,\alpha_3+1}_{l_1,l_2}(x,y)=\sum_{i=\widehat{(\alpha_3+1)}}^{l_2}c_i\partial_x\mathcal{J}^{\alpha_1,\alpha_2,\alpha_3}_{|l|-i+1,i}(x,y)\qquad\text{for certain } c_i\text{'s}.$$

This shows that $\partial_x:\mathbb{R}^{\alpha}_m\to\mathbb{R}^{\alpha_1+1,\alpha_2,\alpha_3+1}_{m-1}$ is surjective, i.e., $\partial_x\mathbb{R}^{\alpha}_m=\mathbb{R}^{\alpha_1+1,\alpha_2,\alpha_3+1}_{m-1}$. Similar arguments lead to $\partial_y\mathbb{R}^{\alpha}_m=\mathbb{R}^{\alpha_1,\alpha_2+1,\alpha_3+1}_{m-1}$ and $(\partial_y-\partial_x)\mathbb{R}^{\alpha}_m=\mathbb{R}^{\alpha_1+1,\alpha_2+1,\alpha_3}_{m-1}$. Therefore, we deduce that $D^n\mathbb{R}^{\alpha}_m=\mathbb{R}^{\alpha+n^*}_{m-|n|}$, i.e., D^n maps $\mathbb{R}^{\alpha}_m$ onto $\mathbb{R}^{\alpha+n^*}_{m-|n|}$.

Now, for any sufficiently smooth function u,

$$D^nP^{\alpha}_Mu=\sum_{m=|\hat{\alpha}|}^{M}D^nR^{\alpha}_mu.$$

On the other hand, for any $v\in\mathbb{R}^{\alpha+n^*}_{m-|n|}$, there exists $v^\star\in\mathbb{R}^{\alpha}_m$ such that $D^nv^\star=v$. Correspondingly,

$$\begin{aligned}(D^nR^{\alpha}_mu-D^nu,v)_{\chi^{\alpha+n^*}}&=(D^nR^{\alpha}_mu-D^nu,D^nv^\star)_{\chi^{\alpha+n^*}}\\&=(R^{\alpha}_mu-u,(-1)^{|n|}\chi^{-\alpha}D^n(\chi^{\alpha+n^*}D^nv^\star))_{\chi^{\alpha}}=0.\end{aligned}$$

This indicates that $D^nR^{\alpha}_mu=R^{\alpha+n^*}_{m-|n|}D^nu$, which in return leads to (3.8). □

The following lemma is an analogue of the results for Fourier series, and it plays a key role in our error estimates.

Lemma 3.5. *Let $\sigma\in\mathbb{N}_0$ and $\alpha\in\aleph^3$. If $u\in B^{\sigma}_{\alpha}(T)$, then*

$$|D^nu|^2_{B^{\varrho}_{\alpha+n^*}}=\sum_{m\geq|\hat{\alpha}|}\mu^{\varrho}_{m-|n|,|\alpha+n^*|}\,\|D^nu_m\|^2_{\chi^{\alpha+n^*}},\quad 0\leq\varrho\leq\sigma-|n|,\ n\in\mathbb{N}^3_0.\tag{3.9}$$

In particular,

$$|u|^2_{B^{\varrho}_{\alpha}}=\sum_{m\geq|\hat{\alpha}|}\mu^{\varrho}_{m,|\alpha|}\|u_m\|^2_{\chi^{\alpha}},\quad 0\leq\varrho\leq\sigma.\tag{3.10}$$

Proof. Since (3.10) is just a special case of (3.9) with $n=(0,0,0)$, we only need to prove (3.9). By Theorem 3.2 and Lemma 3.4, we have

$$\begin{aligned}|D^nu|^2_{B^{\varrho}_{\alpha+n^*}}&=\sum_{|k|=\varrho}\frac{\varrho!}{k_1!k_2!k_3!}(D^{n+k}\sum_{m\geq|\hat{\alpha}|}u_m,D^{n+k}\sum_{j\geq|\hat{\alpha}|}u_j)_{\chi^{n^*+k^*+\alpha}}\\&=\sum_{m\geq|\hat{\alpha}|}\sum_{j\geq|\hat{\alpha}|}\sum_{|k|=\varrho}\frac{\varrho!}{k_1!k_2!k_3!}(D^{n+k}u_m,D^{n+k}u_j)_{\chi^{n^*+k^*+\alpha}}\\&=\sum_{m\geq|\hat{\alpha}|}\sum_{j\geq|\hat{\alpha}|}(\mathcal{L}^{\varrho}_{\alpha+n^*}D^nu_m,D^nu_j)_{\chi^{\alpha+n^*}}\\&=\sum_{m\geq|\hat{\alpha}|}\sum_{j\geq|\hat{\alpha}|}\mu^{\varrho}_{m-|n|,|\alpha+n^*|}\,(D^nu_m,D^nu_j)_{\chi^{\alpha+n^*}}\\&=\sum_{m\geq|\hat{\alpha}|}\mu^{\varrho}_{m-|n|,|\alpha+n^*|}\,(D^nu_m,D^nu_m)_{\chi^{\alpha+n^*}}.\end{aligned}$$

This ends the proof. □

We are now in position to prove one of the main results in this paper.

Theorem 3.3. *Let $\sigma \geq 0$ and $\alpha \in \aleph^3$. If $u \in B^\sigma_\alpha(T)$, then*

$$\begin{aligned} \left|D^n(P^\alpha_M u - u)\right|_{B^{\varrho-|n|}_{\alpha+n^*}} &\leq \sqrt{\frac{\mu^{\varrho-|n|}_{M-|n|+1,|\alpha+n^*|}}{\mu^{\sigma-|n|}_{M-|n|+1,|\alpha+n^*|}}}\ \left|D^n(P^\alpha_M u - u)\right|_{B^{\sigma-|n|}_{\alpha+n^*}} \\ &\lesssim M^{\varrho-\sigma}\left|D^n u\right|_{B^{\sigma-|n|}_{\alpha+n^*}}, \quad 0 \leq |n| \leq \varrho \leq \sigma,\ n \in \mathbb{N}^3_0. \end{aligned} \tag{3.11}$$

In particular,

$$\left|P^\alpha_M u - u\right|_{B^\varrho_\alpha} \leq \sqrt{\frac{\mu^\varrho_{M+1,|\alpha|}}{\mu^\sigma_{M+1,\alpha}}}\ \left|P^\alpha_M u - u\right|_{B^\sigma_\alpha} \lesssim M^{\varrho-\sigma}|u|_{B^\sigma_\alpha}, \quad 0 \leq \varrho \leq \sigma. \tag{3.12}$$

Proof. It suffices to prove (3.11). We deduce from (3.10) in Lemma 3.5 that

$$\begin{aligned} \left|D^n(P^\alpha_M u - u)\right|^2_{B^{\varrho-|n|}_{\alpha+n^*}} &= \sum_{m>M} \mu^{\varrho-|n|}_{m-|n|,|\alpha+n^*|}\left\|D^n u_m\right\|^2_{\chi^{\alpha+n^*}} \\ &\leq \frac{\mu^{\varrho-|n|}_{M-|n|+1,|\alpha+n^*|}}{\mu^{\sigma-|n|}_{M-|n|+1,|\alpha+n^*|}} \sum_{m>M} \mu^{\sigma-|n|}_{m-|n|,|\alpha+n^*|}\left\|D^n u_m\right\|^2_{\chi^{\alpha+n^*}} \\ &= \frac{\mu^{\varrho-|n|}_{M-|n|+1,|\alpha+n^*|}}{\mu^{\sigma-|n|}_{M-|n|+1,|\alpha+n^*|}}\left|D^n(P^\alpha_M u - u)\right|^2_{B^{\sigma-|n|}_{\alpha+n^*}}, \end{aligned}$$

which gives the first inequality in (3.12). The second inequality is an immediate consequence of the following two estimates:

$$\frac{\mu^{\varrho-|n|}_{M-|n|+1,|\alpha+n^*|}}{\mu^{\sigma-|n|}_{M-|n|+1,|\alpha+n^*|}} = \frac{\Gamma(M-\sigma+2)\Gamma(M+\varrho+|\alpha|+3)}{\Gamma(M-\varrho+2)\Gamma(M+\sigma+|\alpha|+3)} \lesssim M^{2\varrho-2\sigma}$$

and

$$\begin{aligned} \sum_{m>M} \mu^{\sigma-|n|}_{m-|n|,|\alpha+n^*|}\left\|D^n u_m\right\|^2_{\chi^{\alpha+n^*}} &\leq \sum_{m\geq|\hat{\alpha}|} \mu^{\sigma-|n|}_{m-|n|,|\alpha+n^*|}\left\|D^n u_m\right\|^2_{\chi^{\alpha+n^*}} \\ &= \left|D^n u\right|^2_{B^{\sigma-|n|}_{\alpha+n^*}}. \end{aligned}$$

This ends the proof. □

Remark 3.1. Braess and Schwab [5] proved an elegant result for the d-dimensional simplex for $\alpha \in \aleph^d_+$ by using the properties of the Sturm-Liouville operator. Consequently, their results are expressed in the norms associated with the Sturm-Liouville operator. Our results here, while restricted to a two-dimensional triangle, are valid for all $\alpha \in \aleph^3$ and are expressed in semi-norms expressed directly by the derivatives.

3.4. **Error estimates for the H^1_0-orthogonal projection.** We now consider the H^1_0-orthogonal projection which is essential for the analysis of the Galerkin approximation of elliptic equations.

Denote $X^0_M = H^1_0(T) \cap \mathbb{P}_M$ and define the H^1_0-orthogonal projection $\Pi^{1,0}_M : H^1_0(T) \to X^0_M$ by

$$\left(\nabla(\Pi^{1,0}_M u - u),\ \nabla\varphi\right) = 0, \quad \forall \varphi \in X^0_M. \tag{3.13}$$

Another main result of this paper is the following:

Theorem 3.4. *For any $u \in H_0^1(T) \cap H^r(T)$ with $r \geq 1$,*

$$\left\|\Pi_M^{1,0} u - u\right\|_s \lesssim M^{s-r} |u|_{\tilde{B}^r_{0,0,0}}, \quad s = 0, 1, \tag{3.14}$$

where $|u|^2_{\tilde{B}^1_{0,0,0}} := \|\nabla u\|^2$, and for $r \geq 2$,

$$\begin{aligned} |u|^2_{\tilde{B}^r_{0,0,0}} := {} & |\partial_x u|^2_{B^{r-1}_{0,0,0}} + |\partial_y u|^2_{B^{r-1}_{0,0,0}} + |\partial_x \partial_y u|^2_{B^{r-2}_{0,0,1}} \\ & + |\partial_y(\partial_y - \partial_x) u|^2_{B^{r-2}_{0,1,0}} + |\partial_x(\partial_y - \partial_x) u|^2_{B^{r-2}_{1,0,0}}. \end{aligned} \tag{3.15}$$

Proof. We first prove (3.14) with $s = 1$. By the projection theorem,

$$\left\|\nabla(\Pi_M^{1,0} u - u)\right\| \leq \left\|\nabla(\varphi - u)\right\|, \quad \forall\, \varphi \in X_M^0. \tag{3.16}$$

Taking $\varphi = 0$, we get that

$$\left\|\nabla(\Pi_M^{1,0} u - u)\right\| \leq \|\nabla u\|, \tag{3.17}$$

which leads to (3.14) with $s = r = 1$.

Now let $r \geq 2$. For any $u \in H_0^1(T) \cap H^r(T)$, one obtains by using the Hardy inequality (cf. for instance A.14 in [6]) that $u \in B^1_{-1,-1,-1}(T)$. So we write

$$u = \sum_{l_1 > l_2 \geq 1} \frac{2(2l_1 - 1)(2|l| - 1)}{(l_1 - 1)(|l| + l_1 - 1)} \hat{u}_l \mathcal{J}_l^{-1,-1,-1}.$$

Then by (A.14), we have

$$\begin{aligned} \partial_x(u - P_M^{-1,-1,-1} u) &= \sum_{|l| \geq M+1} \frac{2(2l_1 - 1)(2|l| - 1)}{(l_1 - 1)(|l| + l_1 - 1)} \hat{u}_l \partial_x \mathcal{J}_l^{-1,-1,-1} \\ &= \sum_{|l| \geq M+1} \hat{u}_l \big((l_2 - 1)\mathcal{J}^{0,0,0}_{l_1,l_2-2} - (2l_1 - 1)\mathcal{J}^{0,0,0}_{l_1-1,l_2-1} - (|l| + l_1 - 2)\mathcal{J}^{0,0,0}_{l_1-2,l_2} \\ &\quad - (|l| + l_1)\mathcal{J}^{0,0,0}_{l_1,l_2-1} - (2l_1 - 1)\mathcal{J}^{0,0,0}_{l_1-1,l_2} + (l_2 + 1)\mathcal{J}^{0,0,0}_{l_1-2,l_2+1}\big) \\ &= \sum_{|l| \geq M} \big((l_2 + 1)\hat{u}_{l_1,l_2+2} - (2l_1 + 1)\hat{u}_{l_1+1,l_2+1} - (|l| + l_1 + 2)\hat{u}_{l_1+2,l_2} \\ &\quad - (|l| + l_1 + 1)\hat{u}_{l_1,l_2+1} - (2l_1 + 1)\hat{u}_{l_1+1,l_2} + l_2 \hat{u}_{l_1+2,l_2-1}\big)\mathcal{J}_l^{0,0,0} \\ &\quad + \sum_{|l| = M-1} \big((l_2 + 1)\hat{u}_{l_1,l_2+2} - (2l_1 + 1)\hat{u}_{l_1+1,l_2+1} - (|l| + l_1 + 2)\hat{u}_{l_1+2,l_2}\big)\mathcal{J}_l^{0,0,0}. \end{aligned} \tag{3.18}$$

On the other hand, one gets from (A.5) that

$$\big(\partial_y \mathcal{J}_l^{-1,-1,-1}\big)(x, y) = (-1)^{l_1} \big(\partial_x \mathcal{J}_l^{-1,-1,-1}\big)(y, x).$$

Thus we obtain from (3.18) that

$$\begin{aligned} \partial_y(u - P_M^{-1,-1,-1} u) &= \sum_{|l| \geq M} \big((l_2 + 1)\hat{u}_{l_1,l_2+2} + (2l_1 + 1)\hat{u}_{l_1+1,l_2+1} - (|l| + l_1 + 2)\hat{u}_{l_1+2,l_2} \\ &\quad - (|l| + l_1 + 1)\hat{u}_{l_1,l_2+1} + (2l_1 + 1)\hat{u}_{l_1+1,l_2} + l_2 \hat{u}_{l_1+2,l_2-1}\big)\mathcal{J}_l^{0,0,0} \\ &\quad + \sum_{|l| = M-1} \big((l_2 + 1)\hat{u}_{l_1,l_2+2} + (2l_1 + 1)\hat{u}_{l_1+1,l_2+1} - (|l| + l_1 + 2)\hat{u}_{l_1+2,l_2}\big)\mathcal{J}_l^{0,0,0}. \end{aligned} \tag{3.19}$$

Note that for $M = 2$, we have $u - P_M^{-1,-1,-1}u = u$ and all the terms within $\sum_{|l|=M-1}$ in (3.18)-(3.19) vanish.

Next, a direct calculation leads to

$$
\begin{aligned}
\big\|\nabla(u - P_M^{-1,-1,-1}u)\big\|^2 &= \big\|(\partial_x u - P_{M-1}^{0,0,0}\partial_x u)\big\|^2 + \big\|(\partial_y u - P_{M-1}^{0,0,0}\partial_y u)\big\|^2 \\
&\quad + \sum_{|l|=M-1} \big((l_2+1)\hat{u}_{l_1,l_2+2} - (2l_1+1)\hat{u}_{l_1+1,l_2+1} \\
&\qquad - (|l|+l_1+2)\hat{u}_{l_1+2,l_2}\big)^2 \big\|\mathcal{J}_l^{0,0,0}\big\|^2 \\
&\quad + \sum_{|l|=M-1} \big((l_2+1)\hat{u}_{l_1,l_2+2} + (2l_1+1)\hat{u}_{l_1+1,l_2+1} \\
&\qquad - (|l|+l_1+2)\hat{u}_{l_1+2,l_2}\big)^2 \big\|\mathcal{J}_l^{0,0,0}\big\|^2 \\
&= \big\|(\partial_x u - P_{M-1}^{0,0,0}\partial_x u)\big\|^2 + \big\|(\partial_y u - P_{M-1}^{0,0,0}\partial_y u)\big\|^2 \\
&\quad + \sum_{|l|=M-1} \frac{2l_1+1}{|l|+1}\hat{u}^2_{l_1+1,l_2+1} \\
&\quad + \sum_{|l|=M-1} \frac{1}{(2l_1+1)(|l|+1)}\big((l_2+1)\hat{u}_{l_1,l_2+2} - (|l|+l_1+2)\hat{u}_{l_1+2,l_2}\big)^2 \qquad (3.20) \\
&= \big\|(\partial_x u - P_{M-1}^{0,0,0}\partial_x u)\big\|^2 + \big\|(\partial_y u - P_{M-1}^{0,0,0}\partial_y u)\big\|^2 \\
&\quad + \sum_{|l|=M-1} \frac{(l_2+1)(|l|+l_1+2)}{(2l_1+1)(|l|+1)}\big(\hat{u}_{l_1,l_2+2} - \hat{u}_{l_1+2,l_2}\big)^2 \\
&\quad + \sum_{|l|=M-1} \Big(\frac{|l|+l_1+2}{|l|+1}\hat{u}^2_{l_1+2,l_2} + \frac{2l_1+1}{|l|+1}\hat{u}^2_{l_1+1,l_2+1} - \frac{l_2+1}{|l|+1}\hat{u}^2_{l_1,l_2+2}\Big) \\
&= \big\|(\partial_x u - P_{M-1}^{0,0,0}\partial_x u)\big\|^2 + \big\|(\partial_y u - P_{M-1}^{0,0,0}\partial_y u)\big\|^2 \\
&\quad + \sum_{|l|=M-1} \frac{4l_1+2}{|l|+1}\hat{u}^2_{l_1+1,l_2+1} \\
&\quad + \sum_{|l|=M-1} \frac{(l_2+1)(|l|+l_1+2)}{(2l_1+1)(|l|+1)}\big(\hat{u}_{l_1,l_2+2} - \hat{u}_{l_1+2,l_2}\big)^2.
\end{aligned}
$$

We now bound the second and the third terms in the last equation.

By using (A.10) and (2.16), we derive that

$$
\begin{aligned}
&\big\|\partial_x(\partial_y - \partial_x)(u - P_M^{-1,-1,-1}u)\big\|^2_{\chi^{1,0,0}} + \big\|\partial_y(\partial_y - \partial_x)(u - P_M^{-1,-1,-1}u)\big\|^2_{\chi^{0,1,0}} \qquad (3.21) \\
&= \Big(u - P_M^{-1,-1,-1}u,\, \big(D^{1,0,1}\chi^{1,0,0}D^{1,0,1} + D^{0,1,1}\chi^{0,1,0}D^{0,1,1}\big)(u - P_M^{-1,-1,-1}u)\Big) \\
&= \sum_{|l|\geq M+1} l_1(l_1-1)l_2(|l|+l_1-1) \times \Big\|\frac{2(2l_1-1)(2|l|-1)}{(l_1-1)(|l|+l_1-1)}\hat{u}_l J_l^{-1,-1,-1}\Big\|^2_{\chi^{-1,-1,-1}} \\
&= \sum_{|l|\geq M-1} 4(2l_1+1)(2|l|+3)\hat{u}^2_{l_1+1,l_2+1} \geq 2M(2M+1)\sum_{|l|=M-1}\frac{4l_1+2}{|l|+1}\hat{u}^2_{l_1+1,l_2+1}.
\end{aligned}
$$

Furthermore, we derive from (A.13) and (2.16) that

$$
\begin{aligned}
&\big\|\partial_x\partial_y(u-P_M^{-1,-1,-1}u)\big\|^2_{\chi^{0,0,1}}\\
&\quad=\Big\|\sum_{|l|\geq M+1}(2|l|-1)\hat{u}_l\big((|l|+l_1)\mathcal{J}^{0,0,1}_{l_1,l_2-2}-(|l|+l_1-2)\mathcal{J}^{0,0,1}_{l_1-2,l_2}\big)\Big\|^2_{\chi^{0,0,1}}\\
&\quad=\Big\|\sum_{|l|\geq M-1}(2|l|+3)(|l|+l_1+2)(\hat{u}_{l_1,l_2+2}-\hat{u}_{l_1+2,l_2})\mathcal{J}^{0,0,1}_{l_1,l_2}\Big\|^2_{\chi^{0,0,1}}\\
&\quad=\sum_{|l|\geq M-1}\frac{(l_2+1)(2|l|+3)(|l|+l_1+2)}{2l_1+1}(\hat{u}_{l_1,l_2+2}-\hat{u}_{l_1+2,l_2})^2\\
&\quad\geq M(2M+1)\sum_{|l|=M-1}\frac{(l_2+1)(|l|+l_1+2)}{(2l_1+1)(|l|+1)}\big(\hat{u}_{l_1,l_2+2}-\hat{u}_{l_1+2,l_2}\big)^2.
\end{aligned}
\tag{3.22}
$$

Combining (3.20)-(3.22) and using Theorem 3.3, we deduce that

$$
\begin{aligned}
\big\|\nabla(u-P_M^{-1,-1,-1}u)\big\|^2&\leq\big\|(I-P_{M-1}^{0,0,0})\partial_x u\big\|^2+\big\|(I-P_{M-1}^{0,0,0})\partial_y u\big\|^2\\
&\quad+M^{-2}\Big(\big\|\partial_x\partial_y(u-P_M^{-1,-1,-1}u)\big\|^2_{\chi^{0,0,1}}+\big\|\partial_x(\partial_y-\partial_x)(u-P_M^{-1,-1,-1}u)\big\|^2_{\chi^{1,0,0}}\\
&\quad+\|\partial_y(\partial_y-\partial_x)(u-P_M^{-1,-1,-1}u)\|^2_{\chi^{0,1,0}}\Big)\\
&\lesssim M^{2-2r}|\partial_x u|^2_{B^{r-1}_{0,0,0}}+M^{2-2r}|\partial_y u|^2_{B^{r-1}_{0,0,0}}+M^{2-2r}|\partial_x\partial_y u|^2_{B^{r-2}_{0,0,1}}\\
&\quad+M^{2-2r}|\partial_y(\partial_y-\partial_x)u|^2_{B^{r-2}_{0,1,0}}+M^{2-2r}|\partial_x(\partial_y-\partial_x)u|^2_{B^{r-2}_{1,0,0}}\\
&=M^{2-2r}|u|^2_{\tilde{B}^r_{0,0,0}}.
\end{aligned}
$$

Since $P_M^{-1,-1,-1}u\in X_M^0$, taking $\varphi=P_M^{-1,-1,-1}u$ in (3.16), we obtain

$$
\big\|\nabla(u-\Pi_M^0 u)\big\|\leq\big\|\nabla(u-P_M^{-1,-1,-1}u)\big\|\leq M^{1-r}|u|_{\tilde{B}^r_{0,0,0}},
$$

which is (3.14) with $s=1$.

The case with $s=0$ can be proved by using a duality argument. Let $e_M=\Pi_M^{1,0}u-u$ and consider the auxiliary problem:

$$
\text{find } w\in H_0^1(T) \text{ such that } \big(\nabla\phi,\nabla w\big)=(\phi,e_M),\quad\forall\phi\in H_0^1(T),
\tag{3.23}
$$

which admits a unique solution $w\in H_0^1(T)$ with the regularity

$$
\|w\|_2\lesssim\|e_M\|.
$$

Taking $\phi=e_M$ in (3.23), we have from (3.13) and (3.14) with $s=1$ that

$$
\begin{aligned}
\|e_M\|^2=(\nabla e_M,\nabla w)&=(\nabla e_M,\nabla(w-\Pi_M^{1,0}w))\\
&\leq\|\nabla e_M\|\|\nabla(w-\Pi_M^{1,0}w)\|\\
&\lesssim\|\nabla e_M\|M^{-1}\|w\|_2\\
&\lesssim M^{-1}\|\nabla e_M\|\|e_M\|,
\end{aligned}
$$

which implies $\|e_M\|\lesssim M^{-1}\|\nabla e_M\|$. By using (3.14) with $s=1$, we finally get the estimate (3.14) with $s=0$. □

4. Application to second and fourth order equations

4.1. **Modified Helmholtz equation.** Consider the modified Helmholtz equation:

$$-\Delta u+\gamma u=f,\quad \text{in } T;\quad u|_{\partial T}=0,\quad \gamma\geq 0. \tag{4.1}$$

Given $f\in H^{-1}(T)$, the weak formulation for (4.1) is: find $u\in H_0^1(T)$ such that

$$a(u,v):=\big(\nabla u,\nabla v\big)+\gamma(u,v)=(f,v),\quad \forall v\in H_0^1(T). \tag{4.2}$$

It is clear that $X_M^0=\mathbb{P}_M^{-1,-1,-1}$. Hence,

$$X_M^0=\text{span}\big\{\phi_l:\ l_1\geq 2,\ l_2\geq 1,\ l_1+l_2\leq M\big\}$$

with

$$\phi_l(x,y)=\frac{2(2l_1-1)(2|l|-1)}{(l_1-1)(|l|+l_1-1)}\mathcal{J}_l^{-1,-1,-1}(x,y). \tag{4.3}$$

The spectral-Galerkin approximation for (4.2) is: find $u_M\in X_M^0$ such that

$$a(u_M,\varphi)=(f,\varphi),\quad \forall\varphi\in X_M^0. \tag{4.4}$$

Thanks to the Lax-Milgram lemma, (4.2) (resp. (4.4)) admits a unique solution satisfying

$$\|\nabla u\|+\gamma\|u\|\lesssim\|f\|_{-1}\quad(\text{resp. } \|\nabla u_M\|+\gamma\|u_M\|\lesssim\|f\|_{-1}). \tag{4.5}$$

Setting

$$a_{l;k}=(\phi_l,\phi_k),\qquad b_{l;k}=(\partial_x\phi_l,\partial_x\phi_k)+(\partial_y\phi_l,\partial_y\phi_k),$$
$$u_M=\sum_{l_1=2}^{M-1}\sum_{l_2=1}^{M-l_1}\hat{u}_l\phi_l,\qquad \hat{f}_l=(f,\phi_l),$$

and

$$A_{l_1,k_1}=\big(a_{l_1,l_2;k_1,k_2}\big)_{l_1+1\leq|l|\leq M,\,k_1+1\leq|k|\leq M},\qquad A=\big(A_{l_1,k_1}\big)_{2\leq l_1,k_1\leq M-1},$$
$$B_{l_1,k_1}=\big(b_{l_1,l_2;k_1,k_2}\big)_{l_1+1\leq|l|\leq M,\,k_1+1\leq|k|\leq M},\qquad B=\big(B_{l_1,k_1}\big)_{2\leq l_1,k_1\leq M-1},$$
$$\hat{u}^{l_1}=\big(\hat{u}_{l_1,1},\hat{u}_{l_1,2},\cdots,\hat{u}_{l_1,M-l_1}\big),\qquad \hat{u}=\big(\hat{u}^2,\hat{u}^3,\cdots,\hat{u}^{M-1}\big)^{\text{tr}},$$
$$\hat{f}^{l_1}=\big(\hat{f}_{l_1,1},\hat{f}_{l_1,2},\cdots,\hat{f}_{l_1,M-l_1}\big),\qquad \hat{f}=\big(\hat{f}^2,\hat{f}^3,\cdots,\hat{f}^{M-1}\big)^{\text{tr}},$$

the linear system associated with (4.4) becomes

$$(B+\gamma A)\hat{u}=\hat{f}. \tag{4.6}$$

It is easy to see from (A.14), (A.15) and (A.5) that

$$a_{l;k}=0,\qquad \text{if } l_1-k_1\notin\big\{0,\pm2\big\}\ \text{ or } |k|-|l|\notin\big\{0,\pm1,\pm2,\pm3\big\},$$
$$b_{l;k}=0,\qquad \text{if } l_1-k_1\notin\big\{0,\pm2\big\}\ \text{ or } |k|-|l|\notin\big\{0,\pm1\big\}.$$

So $(B+\gamma A)$ is a block penta-diagonal matrix, whose entries are all hepta-diagonal submatrices (cf. Figure 4.1). Moreover, the system can be split into two block tri-diagonal linear systems with respect to even/odd l_1,k_1. The non-zero entries of the matrices A and B can be determined exactly by using the properties of $\mathcal{J}_l^{-1,-1,-1}(x,y)$ in the Appendix.

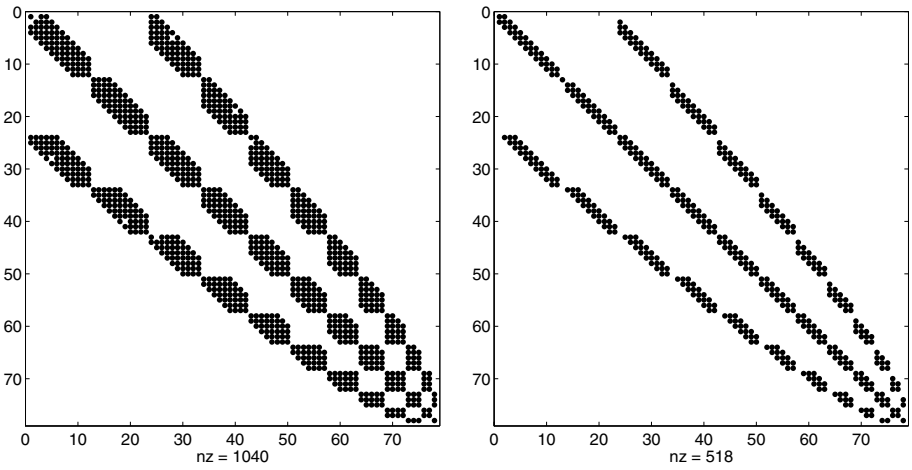

FIGURE 4.1. Structure of the mass matrix A (left) and the stiff matrix B (right) with $M = 14$: nz is the number of non-zero elements.

Theorem 4.1. *Let u and u_M be the solutions of (4.2) and (4.4), respectively. If $u \in H^r(T)$, $r \geq 1$, then we have*

$$\|u_M - u\|_s \lesssim M^{s-r} |u|_{\tilde{B}^r_{0,0,0}}, \quad 0 \leq s \leq 1. \tag{4.7}$$

Proof. One derives from (4.2) and (4.4) that

$$\|u_M - u\|_1 \lesssim \inf_{v \in X^0_M} \|v - u\|_1 .$$

Then, applying Theorem 3.4 to the above, we obtain (4.7) immediately with $\mu = 1$. The result for $\mu = 0$ can then be derived by using a standard duality argument as in the proof of Theorem 3.4. □

4.2. **Biharmonic equation.** Consider the following biharmonic equation:

$$\Delta^2 u = f, \quad \text{in } T; \quad u|_{\partial T} = 0, \ \frac{\partial u}{\partial \Gamma}\Big|_{\partial T} = 0. \tag{4.8}$$

Given $f \in H^{-2}(T)$, the weak formulation for (4.8): find $u \in H^2_0(T)$ such that

$$(\Delta u, \Delta v) = (f, v), \quad \forall v \in H^2_0(T). \tag{4.9}$$

The coercivity of the above bilinear form can be directly obtained from the equivalence norm/semi-norm in $H^2_0(T)$ and the formula

$$\iint_T (\Delta u)^2 dxdy = \iint_T \left[\left(\partial_x^2 u\right)^2 + 2 \left(\partial_x \partial_y u\right)^2 + \left(\partial_y^2 u\right)^2 \right] dxdy.$$

Thus for any $f \in H^{-2}(T)$, problem (4.8) has a unique solution in $H^2_0(T)$, which satisfies

$$\|u\|_2 \lesssim \|f\|_{-2}. \tag{4.10}$$

Define the approximation space $Y^0_M := H^2_0(T) \cap \mathbb{P}_M$. Then, we have

$$Y^0_M = \mathbb{P}^{-2,-2,-2}_M = \operatorname{span}\{\psi_l : \ l_1 \geq 4, \, l_2 \geq 2, \ l_1 + l_2 \leq M\} \tag{4.11}$$

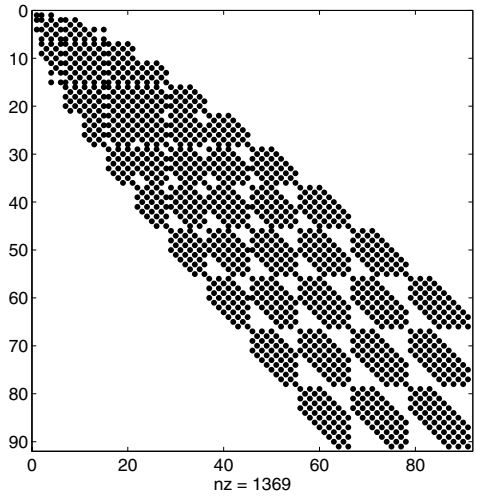

FIGURE 4.2. Structure of the matrix D (with $M = 18$): nz is the number of non-zero elements.

with

$$\psi_l(x,y) = \frac{2\mathcal{J}_l^{-2,-2,-2}(x,y)}{(l_1-2)(l_1-3)(|l|+l_1-3)(|l|+l_1-4)}, \quad l_1 \geq 4,\ l_2 \geq 2. \tag{4.12}$$

The spectral-Galerkin approximation to (4.9) is: to find $u_M \in Y_M^0$ such that

$$(\Delta u_M, \Delta\varphi) = (f,\varphi), \quad \forall \varphi \in Y_M^0, \tag{4.13}$$

which has a unique solution satisfying (4.10) with u_M in place of u.

Setting

$$d_{l;k} = (\Delta\psi_l, \Delta\psi_k), \qquad u_M = \sum_{l_1=4}^{M-2}\sum_{l_2=2}^{M-l_1} \hat{u}_l\psi_l, \qquad \hat{f}_l = (f,\psi_l),$$

and

$$\begin{aligned} &D_{|l|,|k|} = \big(d_{l;k}\big)_{4\leq l_1\leq |l|-2,\, 4\leq k_1\leq |k|-2}, && D = \big(D_{|l|,|k|}\big)_{6\leq |l|,|k|\leq M},\\ &\hat{u}^{|l|} = \big(\hat{u}_{4,|l|-4}, \hat{u}_{5,|l|-5}, \cdots, \hat{u}_{|l|-2,2}\big), && \hat{u} = \big(\hat{u}^6, \hat{u}^7, \cdots, \hat{u}^M\big)^{\mathrm{tr}},\\ &\hat{f}^{|l|} = \big(\hat{f}_{4,|l|-4}, \hat{f}_{5,|l|-5}, \cdots, \hat{f}_{|l|-2,2}\big), && \hat{f} = \big(\hat{f}^6, \hat{f}^7, \cdots, \hat{f}^M\big)^{\mathrm{tr}}, \end{aligned}$$

the linear system associated with (4.13) becomes

$$D\hat{u} = \hat{f}. \tag{4.14}$$

It is easy to see from (A.16) that

$$d_{l;k} = 0, \qquad \text{if } \ l_1 - k_1 \notin \big\{0, \pm 2, \pm 4\big\} \ \text{ or } |k| - |l| \notin \big\{0, \pm 1, \pm 2\big\}.$$

Hence, $(B + \gamma A)$ is a block penta-diagonal matrix, whose entries are all ennea-diagonal submatrices (cf. Figure 4.2). Moreover, the system can be split into two block penta-diagonal linear systems with respect to even/odd l_1, k_1. The non-zero entries of the matrices D can be exactly determined from the properties of $\mathcal{J}_l^{-2,-2,-2}(x,y)$ in the Appendix.

We now present the error analysis for the scheme (4.13).

Theorem 4.2. *Let u and u_M be the solutions of* (4.9) *and* (4.13)*, respectively. If $\partial_x^2 u \in B^{r-2}_{0,-2,0}(T)$ and $\partial_y^2 u \in B^{r-2}_{-2,0,0}(T)$ with $r \geq 2$, then we have*

$$\|u_M - u\|_2 \lesssim M^{2-r}\Big(\big|\partial_x^2 u\big|_{B^{r-2}_{0,-2,0}} + \big|\partial_y^2 u\big|_{B^{r-2}_{-2,0,0}} \Big). \tag{4.15}$$

Furthermore, if $u \in B^{r+1}_{0,0,0}(T)$, then

$$\|u_M - u\|_s \lesssim M^{s-r}\|u\|_{B^{r+1}_{0,0,0}}, \quad 0 \leq s \leq 2. \tag{4.16}$$

Proof. One derives immediately from (4.9) and (4.13) that

$$\|u_M - u\|_2 \lesssim \inf_{v \in Y_M^0} |v - u|_2.$$

Then applying (3.11) of Theorem 3.3 with $\alpha = (-2,-2,-2)$, $\varrho = 2$ to the above, we obtain that

$$\begin{aligned}
\|u_M - u\|_2 &\lesssim |P_M^{-2,-2,-2}u - u|_2 \\
&\lesssim \|\partial_x^2(P_M^{-2,-2,-2}u - u)\| + \|\partial_y^2(P_M^{-2,-2,-2}u - u)\| \\
&\lesssim M^{2-r}\Big(\big|\partial_x^2 u\big|_{B^{r-2}_{0,-2,0}} + \big|\partial_y^2 u\big|_{B^{r-2}_{-2,0,0}} \Big).
\end{aligned}$$

Now, for $u \in B^{r+1}_{0,0,0}(T)$, we have

$$\begin{aligned}
&\big|\partial_x^2 u\big|_{B^{r-2}_{0,-2,0}(T)} + \big|\partial_y^2 u\big|_{B^{r-2}_{-2,0,0}(T)} \\
&\quad\lesssim \sum_{n_1+n_2+n_3=r-2} \big\|\partial_x^{n_1+2}\partial_y^{n_2}(\partial_y - \partial_x)^{n_3} u\big\|_{\chi^{n_1+n_3,n_2+n_3-2,n_1+n_2}} \\
&\quad+ \sum_{n_1+n_2+n_3=r-2} \big\|\partial_x^{n_1}\partial_y^{n_2+2}(\partial_y - \partial_x)^{n_3} u\big\|_{\chi^{n_1+n_3-2,n_2+n_3,n_1+n_2}} \\
&\quad\lesssim |u|_{B^r_{0,0,0}} + \big\|\partial_x^r u\big\|_{\chi^{0,-2,0}} + \big\|\partial_y^r u\big\|_{\chi^{-2,0,0}} \\
&\quad\text{(by the Hardy inequality (cf. for instance A.14 in [6])} \\
&\quad\lesssim |u|_{B^r_{0,0,0}} + \big\|\partial_x^r \partial_y u\big\| + \big\|\partial_x \partial_y^r u\big\| \lesssim \|u\|_{B^{r+1}_{0,0,0}}.
\end{aligned}$$

This implies (4.16) for $s = 2$. The result for $s = 0$ can then be derived by using a standard duality argument, and the general case $0 \leq s \leq 2$ can be concluded by space interpolation. □

Remark 4.1. We note that the above results are derived using only Theorem 3.3. While the estimate in (4.15) has optimal convergence rate, it requires that $\partial_x^2 u$ and $\partial_y^2 u$ satisfy certain additional boundary conditions due to the weight functions with negative index in the norm; on the other hand, this requirement is removed in the estimate (4.16) but the convergence rate becomes suboptimal.

To derive optimal estimates as in Theorem 4.1 for the scheme (4.13), one needs to study the projection operator $P_M^{-2,-2,-2}$, as we did in Theorem 3.4 for $P_M^{-1,-1,-1}$. However, this process becomes extremely tedious and technical, so we decided not to perform it.

5. Numerical results and concluding remarks

To illustrate the convergence behaviors of the spectral-Galerkin method on the triangle, we report below some numerical results. An important purpose of these simple tests is to check the correctness of various properties derived in the Appendix which are used in determining the non-zero entries of the stiffness and mass matrices.

Example 1. We consider the L^2-orthogonal projection to the following two functions:

$$u(x,y) = \exp(x+y) \qquad \text{and} \qquad u(x,y) = \cos(x-y).$$

Since the functions are analytic on T, Theorem 3.3 indicates that the errors will decay faster than any algebraic order. In Figure 5.1, the maximum errors and L^2-error of the L^2 projection are shown. The computation is performed in Maple with an accuracy of 50 decimal digits.

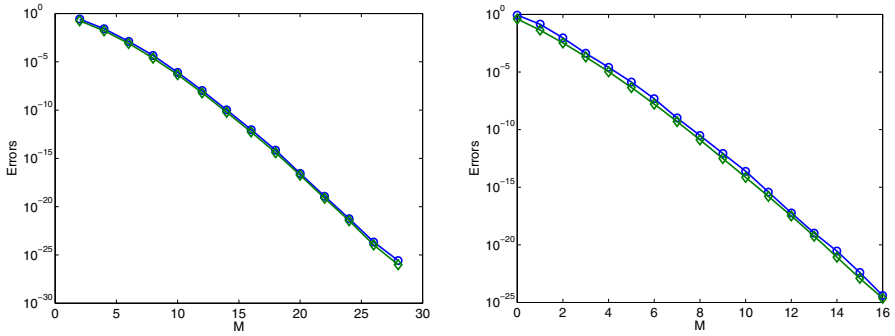

Figure 5.1. Maximum pointwise errors (marked by 'o') and L^2-errors (marked by '$\diamond$') of the L^2 projection (with $\alpha = (0,0,0)$) against various degrees M. Left: function $y = \cos(x-y)$; Right: function $y = \exp(x+y)$.

Example 2. We consider the modified Helmholtz equation (4.1) with the exact solution

$$u(x,y) = xy(e^{x+y} - e).$$

Example 3. We consider the biharmonic equation (4.8) with the exact solution

$$u(x,y) = \sin^2(\pi x)\sin^2(\pi y)\sin^2(\pi(1-x-y)).$$

Computations for Examples 2 and 3 are performed in Matlab and the results are shown in Figures 5.2 and 5.3, respectively. For all three examples, we observe super-geometric convergence rates, which are typical for spectral approximations to analytic functions.

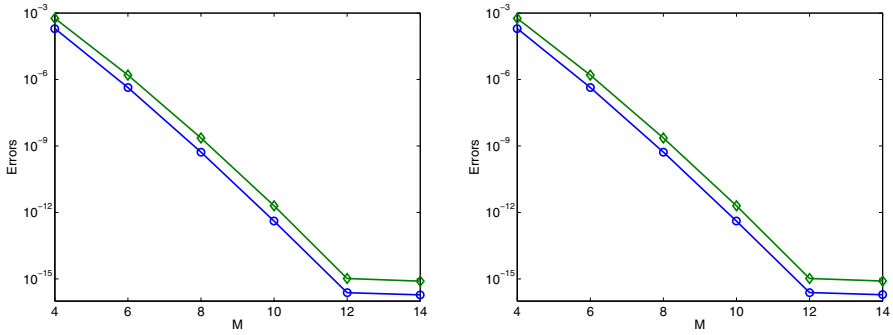

FIGURE 5.2. Maximum pointwise errors (marked by '$\diamond$') and L^2-errors (marked by 'o') against various degrees M for Example 2. Left: function $\gamma = 0$; Right: function $\gamma = 1$.

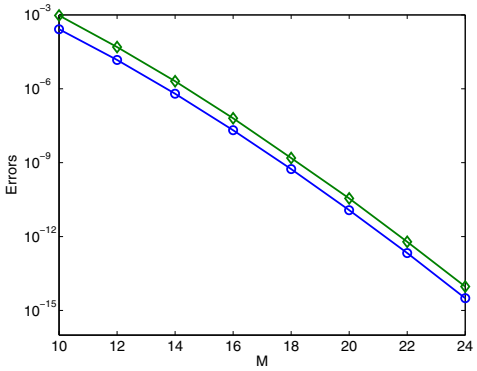

FIGURE 5.3. Maximum pointwise errors (marked by '$\diamond$') and L^2-errors (marked by 'o') against various degrees M for Example 3.

Concluding remarks. We studied in this paper the spectral approximation by orthogonal polynomials on the triangle. By introducing the generalized Koornwinder polynomials and using the properties of the Sturm-Liouville operator on the triangle, we derived optimal error estimates for both the L^2- and H_0^1-orthogonal polynomial projections on the triangle. We then applied these results to derive error estimates for the spectral-Galerkin method for second-order and fourth-order equations on the triangle.

Furthermore, the generalized Koornwinder polynomials with suitable parameters serve as natural basis functions for the spectral-Galerkin method and lead to sparse linear systems that can be efficiently solved.

While we only studied the polynomial approximations on the triangle, it is clear that the method used here can be generalized to the approximations on the tetrahedron. Therefore, the techniques and results developed in this paper will be very useful for studying spectral and spectral-element approximations of PDEs in triangular domains.

Appendix A. Properties of the generalized Koornwinder polynomials

Since the generalized Koornwinder polynomials are defined in terms of the generalized Jacobi polynomials, we will first present some properties of the generalized Jacobi polynomials.

Lemma A.1. *Let $\alpha_1, \alpha_2 \in \aleph$. Then for any $m, n \geq \hat{\alpha}_1 + \hat{\alpha}_2$,*

$$J_n^{\alpha_1,\alpha_2}(\zeta) = (-1)^n J_n^{\alpha_2,\alpha_1}(-\zeta), \tag{A.1}$$

$$(J_m^{\alpha_1,\alpha_2}, J_n^{\alpha_1,\alpha_2})_{\omega^{\alpha_1,\alpha_2},I} = h_n^{\alpha_1,\alpha_2}\delta_{m,n}, \tag{A.2}$$

$$J_n^{\alpha_1,\alpha_2}(\zeta) = \sum_{k=0}^{n} \binom{n+\alpha_1}{n-k}\binom{n+\alpha_2}{k}\Big(\frac{\zeta-1}{2}\Big)^k\Big(\frac{\zeta+1}{2}\Big)^{n-k}, \tag{A.3}$$

$$J_n^{\alpha_1,\alpha_2}(\zeta) = \sum_{k=0}^{n} \binom{n+\alpha_1}{n-k}\binom{n+k+\alpha_1+\alpha_2}{k}\Big(\frac{\zeta-1}{2}\Big)^k, \tag{A.4}$$

where $h_n^{\alpha_1,\alpha_2}$ is defined as in (2.9).

Proof. (A.1) is obvious from the definition, while (A.2) is an immediate consequence of (2.8), (2.9) and (2.10).

We now assume $\alpha_1 \in \aleph_-$, $\alpha_2 \in \aleph_-$. Noting that (A.3) is valid for $\alpha_1 \in \aleph_+$, $\alpha_2 \in \aleph_+$, we derive that

$$\begin{aligned}
J_n^{\alpha_1,\alpha_2}(\zeta) &= \Big(\frac{\zeta-1}{2}\Big)^{-\alpha_1}\Big(\frac{\zeta+1}{2}\Big)^{-\alpha_2} J_{n+\alpha_1+\alpha_2}^{-\alpha_1,-\alpha_2}(\zeta)\\
&= \sum_{k=0}^{n+\alpha_1+\alpha_2}\binom{n+\alpha_2}{n+\alpha_1+\alpha_2-k}\binom{n+\alpha_1}{k}\Big(\frac{\zeta-1}{2}\Big)^{k-\alpha_1}\Big(\frac{\zeta+1}{2}\Big)^{n+\alpha_1-k}\\
&= \sum_{k=-\alpha_1}^{n+\alpha_2}\binom{n+\alpha_2}{n+\alpha_2-k}\binom{n+\alpha_1}{k+\alpha_1}\Big(\frac{\zeta-1}{2}\Big)^{k}\Big(\frac{\zeta+1}{2}\Big)^{n-k}\\
&= \sum_{k=0}^{n}\binom{n+\alpha_2}{k}\binom{n+\alpha_1}{n-k}\Big(\frac{\zeta-1}{2}\Big)^{k}\Big(\frac{\zeta+1}{2}\Big)^{n-k}.
\end{aligned}$$

Similarly, one can prove (A.3) for $\alpha_1 \in \aleph_+$, $\alpha_2 \in \aleph_-$ and for $\alpha_1 \in \aleph_-$, $\alpha_2 \in \aleph_+$.

Using the Chu-Vandermonde sum, one can show that

$$\begin{aligned}
J_n^{\alpha_1,\alpha_2}(\zeta) &= \sum_{k=0}^{n}\sum_{l=0}^{n-k}\binom{n+\alpha_1}{n-k-l}\binom{n+\alpha_2}{k}\binom{k+l+\alpha_1}{l}\Big(\frac{\zeta-1}{2}\Big)^{k+l}\\
&= \sum_{j=0}^{n}\binom{n+\alpha_1}{n-j}\Big(\frac{\zeta-1}{2}\Big)^{j}\sum_{k=0}^{j}\binom{n+\alpha_2}{k}\binom{j+\alpha_1}{j-k}\\
&= \sum_{j=0}^{n}\binom{n+\alpha_1}{n-j}\binom{n+j+\alpha_1+\alpha_2}{j}\Big(\frac{\zeta-1}{2}\Big)^{j}.
\end{aligned}$$

This concludes the proof. □

We now derive some properties of the generalized Koornwinder polynomials.

Lemma A.2. *Let $\alpha \in \aleph^3$. Then for any $l \in \mathbb{N}_0^2$ with $l_1 \geq \hat{\alpha}_1 + \hat{\alpha}_2$ and $l_2 \geq \hat{\alpha}_3$,*

$$\mathcal{J}_l^{\alpha}(x,y) = (-1)^{l_1}\mathcal{J}_l^{\alpha_2,\alpha_1,\alpha_3}(y,x), \tag{A.5}$$

$$\mathcal{J}_l^\alpha(x,y)=\sum_{k_1=0}^{l_1}\sum_{k_2=0}^{l_2} c_l^\alpha(k)\,y^{k_1}(-x-y)^{l_1-k_1+k_2}, \tag{A.6}$$

$$\begin{aligned}\partial_x\mathcal{J}_l^\alpha(x,y)=&-\frac{(l_1+\alpha_2)(|l|+l_1+\alpha_1+\alpha_2+1)}{2l_1+\alpha_1+\alpha_2+1}\mathcal{J}_{l_1-1,l_2}^{\alpha_1+1,\alpha_2,\alpha_3+1}(x,y)\\&-\frac{(l_1+\alpha_1+\alpha_2+1)(|l|+l_1+|\alpha|+2)}{2l_1+\alpha_1+\alpha_2+1}\mathcal{J}_{l_1,l_2-1}^{\alpha_1+1,\alpha_2,\alpha_3+1}(x,y),\end{aligned} \tag{A.7}$$

$$\begin{aligned}\partial_y\mathcal{J}_l^\alpha(x,y)=&\frac{(l_1+\alpha_1)(|l|+l_1+\alpha_1+\alpha_2+1)}{2l_1+\alpha_1+\alpha_2+1}\mathcal{J}_{l_1-1,l_2}^{\alpha_1,\alpha_2+1,\alpha_3+1}(x,y)\\&-\frac{(l_1+\alpha_1+\alpha_2+1)(|l|+l_1+|\alpha|+2)}{2l_1+\alpha_1+\alpha_2+1}\mathcal{J}_{l_1,l_2-1}^{\alpha_1,\alpha_2+1,\alpha_3+1}(x,y),\end{aligned} \tag{A.8}$$

$$(\partial_y-\partial_x)\mathcal{J}_l^\alpha(x,y)=(l_1+\alpha_1+\alpha_2+1)\mathcal{J}_{l_1-1,l_2}^{\alpha_1+1,\alpha_2+1,\alpha_3}(x,y), \tag{A.9}$$

$$\begin{aligned}&\chi^{-\alpha}\partial_y(\partial_y-\partial_x)\left(\chi^{\alpha_1+1,\alpha_2+2,\alpha_3+1}\partial_y(\partial_y-\partial_x)\mathcal{J}_l^\alpha(x,y)\right)\\&\quad+\chi^{-\alpha}\partial_x(\partial_y-\partial_x)\left(\chi^{\alpha_1+2,\alpha_2+1,\alpha_3+1}\partial_x(\partial_y-\partial_x)\mathcal{J}_l^\alpha(x,y)\right)\\&\quad=l_1(l_1+\alpha_1+\alpha_2+1)\big(l_2(l_2+2l_1+|\alpha|+2)\\&\qquad\qquad+(\alpha_3+1)(l_1-1)\big)\mathcal{J}_l^\alpha(x,y),\end{aligned} \tag{A.10}$$

where $c_l^\alpha(k)=\binom{l_1+\alpha_2}{l_1-k_1}\binom{l_1+k_1+\alpha_1+\alpha_2}{k_1}\binom{l_2+2l_1+\alpha_1+\alpha_2+1}{l_2-k_2}\binom{l_2+k_2+2l_1+\alpha_1+\alpha_2+\alpha_3+1}{k_2}$.

Proof. One can readily derive (A.5) from (2.17) and (A.1).

Also from (2.17) and (A.1), we have

$$\mathcal{J}_l^\alpha(x,y)=(-y-x)^{l_1}J_{l_1}^{\alpha_2,\alpha_1}\left(\frac{x-y}{y+x}\right)J_{l_2}^{2l_1+\alpha_1+\alpha_2+1,\alpha_3}(1-2x-2y).$$

Then (A.6) is an immediate consequence of (A.4). By a direct calculation,

$$\begin{aligned}(l_1-k_1+k_2)c_{l_1,l_2}^{\alpha_1,\alpha_2,\alpha_3}(k_1,k_2)=&\frac{(l_1+\alpha_2)(|l|+l_1+\alpha_1+\alpha_2+1)}{2l_1+\alpha_1+\alpha_2+1}c_{l_1-1,l_2}^{\alpha_1+1,\alpha_2,\alpha_3+1}(k_1,k_2)\\&+\frac{(l_1+\alpha_1+\alpha_2+1)(|l|+l_1+|\alpha|+2)}{2l_1+\alpha_1+\alpha_2+1}c_{l_1,l_2-1}^{\alpha_1+1,\alpha_2,\alpha_3+1}(k_1,k_2-1).\end{aligned}$$

Thus we derive (A.7) from (A.6) by comparing coefficients of $y^{k_1}(-x-y)^{l_1-k_1-k_2-1}$ for both the left-hand and right-hand sides. Similar arguments lead to (A.8) and (A.9).

Under the collapsed transform (2.13),

$$\partial_y-\partial_x=\frac{4}{1-\eta}\partial_\xi.$$

Thus by the definition (2.17) of the generalized Koornwinder polynomials and the Sturm-Liouville eigenequation for the generalized Jacobi polynomials (cf. A.6 in [22]), we have

$$\begin{aligned}&-x^{-\alpha_1}y^{-\alpha_2}(\partial_y-\partial_x)\left(x^{\alpha_1+1}y^{\alpha_2+1}(\partial_y-\partial_x)\mathcal{J}_l^\alpha(x,y)\right)\\&\quad=-\omega^{-\alpha_1,-\alpha_2}(\xi)\partial_\xi\left(\omega^{\alpha_1+1,\alpha_2+1}(\xi)\partial_\xi J_{l_1}^{\alpha_1,\alpha_2}(\xi)\right)\times\left(\frac{1-\eta}{2}\right)^{l_1}J_{l_2}^{2l_1+\alpha_1+\alpha_2+1,\alpha_3}(\eta)\\&\quad=l_1(l_1+\alpha_1+\alpha_2+1)J_{l_1}^{\alpha_1,\alpha_2}(\xi)\times\left(\frac{1-\eta}{2}\right)^{l_1}J_{l_2}^{2l_1+\alpha_1+\alpha_2+1,\alpha_3}(\eta)\\&\quad=l_1(l_1+\alpha_1+\alpha_2+1)\mathcal{J}_l^\alpha(x,y).\end{aligned} \tag{A.11}$$

Subtracting (A.11) from (3.7) with $\varrho = 1$, we get

$$\begin{aligned}(\text{A.12})\qquad &-\chi^{-\alpha}\partial_y\left(\chi^{\alpha_1,\alpha_2+1,\alpha_3+1}\partial_y\mathcal{J}_l^\alpha(x,y)\right)-\chi^{-\alpha}\partial_x\left(\chi^{\alpha_1+1,\alpha_2,\alpha_3+1}\partial_x\mathcal{J}_l^\alpha(x,y)\right)\\ &\qquad=\left(l_2(|l|+l_1+|\alpha|+2)+l_1(\alpha_3+1)\right)\mathcal{J}_l^\alpha(x,y).\end{aligned}$$

Thus by (A.9),

$$\begin{aligned}&\chi^{-\alpha}\partial_y(\partial_y-\partial_x)\left(\chi^{\alpha_1+1,\alpha_2+2,\alpha_3+1}\partial_y(\partial_y-\partial_x)\mathcal{J}_l^\alpha(x,y)\right)\\ &\quad+\chi^{-\alpha}\partial_x(\partial_y-\partial_x)\left(\chi^{\alpha_1+2,\alpha_2+1,\alpha_3+1}\partial_x(\partial_y-\partial_x)\mathcal{J}_l^\alpha(x,y)\right)\\ &=(l_1+\alpha_1+\alpha_2+1)\chi^{-\alpha}(\partial_y-\partial_x)\Big(\partial_y\big(\chi^{\alpha_1+1,\alpha_2+2,\alpha_3+1}\partial_y\mathcal{J}_{l_1-1,l_2}^{\alpha_1+1,\alpha_2+1,\alpha_3}(x,y)\big)\\ &\quad+\partial_x\big(\chi^{\alpha_1+1,\alpha_2+2,\alpha_3+1}\partial_x\mathcal{J}_{l_1-1,l_2}^{\alpha_1+1,\alpha_2+1,\alpha_3}(x,y)\big)\Big)\\ &=-(l_1+\alpha_1+\alpha_2+1)\big(l_2(l_2+2l_1+|\alpha|+2)+(\alpha_3+1)(l_1-1)\big)\\ &\quad\times\chi^{-\alpha}(\partial_y-\partial_x)\big(\chi^{\alpha_1+1,\alpha_2+1,\alpha_3}\mathcal{J}_{l_1-1,l_2}^{\alpha_1+1,\alpha_2+1,\alpha_3}(x,y)\big)\\ &=l_1(l_1+\alpha_1+\alpha_2+1)\big(l_2(l_2+2l_1+|\alpha|+2)+(\alpha_3+1)(l_1-1)\big)\mathcal{J}_l^\alpha(x,y).\end{aligned}$$

This concludes the proof. □

We now derive some useful properties for the two specific families of the generalized Koornwinder polynomials with $\alpha=(-1,-1,-1)$ and $\alpha=(-2,-2,-2)$.

Lemma A.3. *For any admissible $l\in\mathbb{N}_0^2$, the following identities hold:*

$$\begin{aligned}(\text{A.13})\qquad &\frac{2(2l_1-1)}{(l_1-1)(|l|+l_1-1)}\partial_x\partial_y\mathcal{J}_{l_1,l_2}^{-1,-1,-1}(x,y)\\ &\quad=(|l|+l_1)\mathcal{J}_{l_1,l_2-2}^{0,0,1}(x,y)-(|l|+l_1-2)\mathcal{J}_{l_1-2,l_2}^{0,0,1}(x,y),\end{aligned}$$

$$\begin{aligned}(\text{A.14})\qquad &\frac{2(2l_1-1)(2|l|-1)}{(l_1-1)(|l|+l_1-1)}\partial_x\mathcal{J}_{l_1,l_2}^{-1,-1,-1}(x,y)\\ &\quad=-(|l|+l_1)\mathcal{J}_{l_1,l_2-1}^{0,0,0}(x,y)-(2l_1-1)\mathcal{J}_{l_1-1,l_2}^{0,0,0}(x,y)\\ &\qquad+(l_2+1)\mathcal{J}_{l_1-2,l_2+1}^{0,0,0}(x,y)+(l_2-1)\mathcal{J}_{l_1,l_2-2}^{0,0,0}(x,y)\\ &\qquad-(2l_1-1)\mathcal{J}_{l_1-1,l_2-1}^{0,0,0}(x,y)-(|l|+l_1-2)\mathcal{J}_{l_1-2,l_2}^{0,0,0}(x,y),\end{aligned}$$

$$\begin{aligned}(\text{A.15})\qquad &\frac{4(2l_1-1)(2|l|-1)}{(l_1-1)(|l|+l_1-1)}\mathcal{J}_{l_1,l_2}^{-1,-1,-1}(x,y)\\ &=\frac{(|l|+l_1)(|l|+l_1+1)}{|l|(2|l|+1)}\mathcal{J}_{l_1,l_2}^{0,0,0}(x,y)-\frac{(l_2+1)(l_2+2)}{|l|(2|l|+1)}\mathcal{J}_{l_1-2,l_2+2}^{0,0,0}(x,y)\\ &\quad-\frac{(|l|+l_1)(|l|-3l_1-1)}{(2|l|+1)(|l|-1)}\mathcal{J}_{l_1,l_2-1}^{0,0,0}(x,y)+\frac{(|l|+3l_1-4)(l_2+1)}{(2|l|+1)(|l|-1)}\mathcal{J}_{l_1-2,l_2+1}^{0,0,0}(x,y)\\ &\quad-\frac{(l_2-1)(|l|+3l_1)}{|l|(2|l|-3)}\mathcal{J}_{l_1,l_2-2}^{0,0,0}(x,y)+\frac{(|l|+l_1-2)(|l|-3l_1+3)}{|l|(2|l|-3)}\mathcal{J}_{l_1-2,l_2}^{0,0,0}(x,y)\\ &\quad+\frac{(l_2-1)(l_2-2)}{(2|l|-3)(|l|-1)}\mathcal{J}_{l_1,l_2-3}^{0,0,0}(x,y)-\frac{(|l|+l_1-3)(|l|+l_1-2)}{(2|l|-3)(|l|-1)}\mathcal{J}_{l_1-2,l_2-1}^{0,0,0}(x,y),\end{aligned}$$

and

$$
\begin{aligned}
&\frac{2}{(l_1-2)(l_1-3)(|l|+l_1-3)(|l|+l_1-4)}\Delta\mathcal{J}^{-2,-2,-2}_{l_1,l_2}(x,y)\\
&\quad=\frac{(|l|+l_1-2)(|l|+l_1-1)}{2(2l_1-1)(2l_1-3)(2|l|-3)(|l|-2)}\mathcal{J}^{0,0,0}_{l_1,l_2-2}(x,y)\\
&\qquad-\frac{2(l_2-2)(|l|+l_1-2)}{(2|l|-5)(2|l|-3)(2l_1-3)(2l_1-1)}\mathcal{J}^{0,0,0}_{l_1,l_2-3}(x,y)\\
&\qquad+\frac{(l_2-3)(l_2-2)}{2(2l_1-1)(2l_1-3)(|l|-2)(2|l|-5)}\mathcal{J}^{0,0,0}_{l_1,l_2-4}(x,y)\\
&\qquad+\frac{3l_1^2-9l_1+2|l|-|l|^2+3}{(2l_1-5)(2l_1-1)(2|l|-3)(|l|-2)}\mathcal{J}^{0,0,0}_{l_1-2,l_2}(x,y)\\
&\qquad+\frac{12l_1^2-36l_1+30-16|l|+4|l|^2}{(2|l|-5)(2|l|-3)(2l_1-5)(2l_1-1)}\mathcal{J}^{0,0,0}_{l_1-2,l_2-1}(x,y)\\
&\qquad+\frac{3l_1^2-9l_1-5-|l|^2+6|l|}{(2l_1-1)(2l_1-5)(|l|-2)(2|l|-5)}\mathcal{J}^{0,0,0}_{l_1-2,l_2-2}(x,y)\\
&\qquad+\frac{(l_2+1)(2+l_2)}{2(2l_1-3)(2l_1-5)(|l|-2)(2|l|-3)}\mathcal{J}^{0,0,0}_{l_1-4,l_2+2}(x,y)\\
&\qquad-\frac{2(l_2+1)(|l|+l_1-5)}{(2|l|-5)(2|l|-3)(2l_1-3)(2l_1-5)}\mathcal{J}^{0,0,0}_{l_1-4,l_2+1}(x,y)\\
&\qquad+\frac{(|l|+l_1-5)(|l|+l_1-6)}{2(2l_1-3)(2l_1-5)(|l|-2)(2|l|-5)}\mathcal{J}^{0,0,0}_{l_1-4,l_2}(x,y).
\end{aligned}
\tag{A.16}
$$

Proof. The identities (A.13)-(A.16) can be verified by using the expansion (A.6) and then comparing the coefficients of $y^{k_1}(-x-y)^{k_2}$ for both the left-hand and the right-hand sides. We leave the details to the interested reader. □

References

[1] M. Abramowitz and I. A. Stegun, editors. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, 1972. MR1225604 (94b:00012)

[2] P. Appell. Sur des polynômes de deux variables analogues aux polynômes de Jacobi. *Arch. Math. Phys.*, 66:238–245, 1881.

[3] P. Appell and J. Kampeé de Fériet. *Fonctions Hypergéométriques et Hypersphériques: Polynômes d'Hermite*. Gauthier-Villars, Paris, 1926.

[4] I. Babuška and Manil Suri. The optimal convergence rate of the p-version of the finite element method. *SIAM J. Numer. Anal.*, 24(4):750–776, 1987. MR899702 (88k:65102)

[5] Dietrich Braess and Christoph Schwab. Approximation on simplices with respect to weighted Sobolev norms. *J. Approx. Theory*, 103(2):329–337, 2000. MR1749969 (2001f:41038)

[6] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang. *Spectral Methods: Fundamentals in Single Domains*. Scientific Computation. Springer-Verlag, Berlin, 2006. MR2223552 (2007c:65001)

[7] Moshe Dubiner. Spectral methods on triangles and other domains. *J. Sci. Comput.*, 6(4):345–390, 1991. MR1154903 (92k:76061)

[8] Charles F. Dunkl and Yuan Xu. *Orthogonal Polynomials of Several Variables*. Cambridge University Press, 2001. MR1827871 (2002m:33001)

[9] D. Funaro. *Polynomial Approximations of Differential Equations*. Springer-Verlag, 1992. MR1176949 (94c:65078)

[10] A. Grundmann and H. M. Möller. Invariant integration formulas for the $n-$simplex by combinatorial methods. *SIAM J. Numer. Anal.*, 15:282–290, 1978. MR488881 (81e:41045)

[11] Ben-Yu Guo, Jie Shen, and Li-Lian Wang. Optimal spectral-Galerkin methods using generalized Jacobi polynomials. *J. Sci. Comput.*, 27(1-3):305–322, 2006. MR2285783 (2008f:65233)

[12] Ben-Yu Guo, Jie Shen, and Li-Lian Wang. Generalized Jacobi polynomials/functions and applications to spectral methods. *Appl. Numer. Math.*, 59(5):1011–1028, 2009.
[13] Benyu Guo and Li-Lian Wang. Error analysis of spectral method on a triangle. *Adv. Comput. Math.*, 26:473–496, 2007. MR2291668 (2007k:41008)
[14] Wilhelm Heinrichs. Spectral collocation on triangular elements. *J. Comput. Phys.*, 145(2):743–757, 1998. MR1645013 (99d:65332)
[15] J. Hesthaven and D. Gottlieb. Stable spectral methods for conservation laws on triangles with unstructured grids. *Comput. Methods Appl. Mech. Eng.*, 175:361–381, 1999. MR1702193 (2000e:65096)
[16] George Em Karniadakis and Spencer J. Sherwin. *Spectral/hp element methods for computational fluid dynamics*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, second edition, 2005. MR2165335 (2006j:65001)
[17] Tom Koornwinder. Two-variable analogues of the classical orthogonal polynomials. In *Theory and application of special functions (Proc. Advanced Sem., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1975)*, pages 435–495. Math. Res. Center, Univ. Wisconsin, Publ. No. 35. Academic Press, New York, 1975. MR0402146 (53:5967)
[18] Heping Ma and Weiwei Sun. A Legendre-Petrov-Galerkin method and Chebyshev-collocation method for the third-order differential equations. *SIAM J. Numer. Anal.*, 38(5):1425–1438, 2000. MR1812518 (2001m:65130)
[19] S. A. Orszag. Spectral methods for complex geometries. *J. Comput. Phys.*, 37:70–92, 1980. MR584322 (83e:65188)
[20] R. G. Owens. Spectral approximations on the triangle. *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.*, 454(1971):857–872, 1998. MR1631583 (99i:33028)
[21] C. Schwab. *p- and hp- Finite Element Methods: Theory and Applications in Solid and Fluid Mechanics*. Numerical Mathematics and Scientific Computation. The Clarendon Press, Oxford University Press, New York, 1998. MR1695813 (2000d:65003)
[22] Jie Shen, Li-Lian Wang, and Huiyuan Li. A triangular spectral element method using fully tensorial rational basis functions. *SIAM J. Numer. Anal.*, 47(3):1619–1650, 2009.
[23] S. J. Sherwin and G. E. Karniadakis. A new triangular and tetrahedral basis for high-order finite element methods. *Int. J. Numer. Methods Engrg.*, 38:3775–3802, 1995. MR1362003 (96h:65140)
[24] S. J. Sherwin and G. E. Karniadakis. A triangular spectral element method; applications to the incompressible Navier-Stokes equations. *Comput. Methods Appl. Mech. Engrg.*, 123(1-4):189–229, 1995. MR1339373 (96b:76069)
[25] A. H. Stroud. Integration formulas and orthogonal polynomials for two variables. *SIAM J. Numer. Anal.*, 1969. MR0261788 (41:6400)

Institute of Software, Chinese Academy of Sciences, Beijing 100190, People's Republic of China

Department of Mathematics, Purdue University, West Lafayette, Indiana, 47907