

EFFICIENT SPECTRAL-GALERKIN ALGORITHMS FOR DIRECT SOLUTION OF THE INTEGRATED FORMS OF SECOND-ORDER EQUATIONS USING ULTRASPHERICAL POLYNOMIALS

E. H. DOHA¹ and A. H. BHRAWY^{1, 2}

(Received 1 July 2004)

Abstract

It is well known that spectral methods (tau, Galerkin, collocation) have a condition number of $O(N^4)$ where N is the number of retained modes of polynomial approximations. This paper presents some efficient spectral algorithms, which have a condition number of $O(N^2)$, based on the ultraspherical-Galerkin methods for the integrated forms of second-order elliptic equations in one and two space variables. The key to the efficiency of these algorithms is to construct appropriate base functions, which lead to systems with specially structured matrices that can be efficiently inverted. The complexities of the algorithms are a small multiple of N^{d+1} operations for a d -dimensional domain with $(N - 1)^d$ unknowns, while the convergence rates of the algorithms are exponentials with smooth solutions.

2000 *Mathematics subject classification*: primary 65N35; secondary 65N22, 65F05, 35J05.

Keywords and phrases: spectral-Galerkin method, ultraspherical polynomials, Poisson and Helmholtz equations.

1. Introduction

The problem of approximating solutions of differential equations by spectral methods, known as Galerkin approximations, involves the projection onto the span of some appropriate set of basis functions. The members of the basis functions may automatically satisfy the auxiliary conditions imposed on the problem, such as initial, boundary or more general conditions. Alternatively, these conditions may be imposed as constraints on the expansion coefficients, as in the Lanczos τ -method ([17]).

It is of fundamental importance to realize that the choice of the basis functions is responsible for the superior approximation properties of spectral methods when

¹Department of Mathematics, Faculty of Science, Cairo University, Giza, Egypt; e-mail: eiddoha@frcu.eun.eg.

²Department of Mathematics, Faculty of Science, Beni-Suef University, Beni-Suef, Egypt; e-mail: alibhrawy@yahoo.co.uk.

© Australian Mathematical Society 2007, Serial-fee code 1446-1811/07

compared with the finite difference and finite element methods. The choice of different basis functions leads to different spectral approximations; for instance, trigonometric polynomials for periodic problems, Chebyshev, Legendre, ultraspherical and Jacobi polynomials for non-periodic problems, Laguerre polynomials for problems on the half-line, and Hermite polynomials for problems on the whole line.

Spectral methods provide a computational approach which has achieved substantial popularity over the last four decades. They have gained new popularity in automatic computations for a wide class of physical problems in fluid and heat flow. The principal advantage of the spectral methods lies in their ability to achieve accurate results with substantially fewer degrees of freedom.

This paper aims to develop some efficient spectral algorithms based on the ultraspherical-Galerkin methods (UGM) for elliptic second-order differential equations in one and two space variables but by considering their integrated forms.

Spectral methods (see, for instance, Ben-Yu [1, 2], Canuto *et al.* [4], Doha [5–7], Gottlieb and Orszag [12] and Siyyam and Syam [23]) involve representing the solution to a problem in terms of a truncated series of smooth global functions. They give very accurate approximations for a smooth solution with relatively few degrees of freedom. For Dirichlet problems, Heinrichs [14] uses the Chebyshev polynomials as basis functions. It turns out that for the well-known standard spectral methods (tau, Galerkin, collocation) the condition number is very large and grows as $O(N^4)$, see Orszag [19]. Heinrichs [15] proposes a spectral method based on a subclass of orthogonal ultraspherical polynomials ($\alpha = 3/2$) for solving the Helmholtz equation in two dimensions subject to Dirichlet homogeneous boundary conditions with a symmetric and sparse matrix, whose condition number grows only as $O(N^2)$; and he shows in [16] that certain algebraic spectral multigrid methods can be efficiently used for solving the resulting system. Doha and Abd-Elhameed [10] use ultraspherical-Galerkin approximations to develop a generalization of the improved technique of Heinrichs [14] and Shen [21, 22].

In this paper we are concerned with direct solution techniques for the integrated forms of second-order elliptic equations, using ultraspherical-Galerkin approximations. We present two appropriate bases for the UGM applied to the integrated forms of Helmholtz elliptic equations with various boundary conditions. This leads to discrete systems with specially structured matrices that can be efficiently inverted. We note that two algorithms, namely, the Legendre and Chebyshev-Galerkin approximations, and some other very interesting cases, can be obtained directly as special cases from our proposed ultraspherical-Galerkin approximations.

The remainder of this paper is organized as follows. In Section 2 we give some properties of ultraspherical polynomials, and in Section 3 we discuss two algorithms for solving the integrated forms of Helmholtz equations in one and two dimensions. In Section 4 we consider the integrated forms of the general second-order elliptic

differential equations. In Section 5 we present some numerical results. Finally, some concluding remarks are given in Section 6.

2. Some properties of ultraspherical polynomials

The ultraspherical polynomials $\{C_n^{(\alpha)}(x), n = 0, 1, 2, \dots, \alpha \in (-1/2, \infty)\}$ are a sequence of orthogonal polynomials on the interval $(-1, 1)$ with respect to the weight function $w(x) = (1 - x^2)^{\alpha-1/2}$, that is,

$$\int_{-1}^1 (1 - x^2)^{\alpha-1/2} C_m^{(\alpha)}(x) C_n^{(\alpha)}(x) dx = \begin{cases} 0, & m \neq n, \\ h_n, & m = n, \end{cases} \tag{2.1}$$

where

$$h_n = \frac{\sqrt{\pi} n! \Gamma(\alpha + 1/2)}{(2\alpha)_n (n + \alpha) \Gamma(\alpha)}, \quad (2\alpha)_n = \frac{\Gamma(n + 2\alpha)}{\Gamma(2\alpha)}.$$

They are eigenfunctions of the Sturm-Liouville problem

$$(1 - x^2)\phi''(x) - (2\alpha + 1)x\phi'(x) + n(n + 2\alpha)\phi(x) = 0.$$

For our present purposes, it is convenient to weigh the ultraspherical polynomials so that $C_n^{(\alpha)}(1) = 1, n = 0, 1, 2, \dots$. This is not the usual standardization but has the desirable properties that $C_n^{(0)}(x)$ are identical to the Chebyshev polynomials of the first kind, $C_n^{(1/2)}(x)$ are the Legendre polynomials $L_n(x)$, and $C_n^{(1)}(x)$ is equal to $(1/(n + 1))U_n(x)$, where $U_n(x)$ are the Chebyshev polynomials of the second kind.

In this form the ultraspherical polynomials may be generated by using the recurrence relation $(n + 2\alpha)C_{n+1}^{(\alpha)}(x) = 2(n + \alpha)x C_n^{(\alpha)}(x) - n C_{n-1}^{(\alpha)}(x), n = 1, 2, 3, \dots$, starting from $C_0^{(\alpha)}(x) = 1$ and $C_1^{(\alpha)}(x) = x$, or obtained from Rodrigues' formula

$$C_n^{(\alpha)}(x) = (-1/2)^n \frac{\Gamma(\alpha + 1/2)}{\Gamma(n + \alpha + 1/2)} (1 - x^2)^{1/2-\alpha} \frac{d^n}{dx^n} [(1 - x^2)^{n+\alpha-1/2}].$$

The special values

$$C_n^{(\alpha)}(\pm 1) = (\pm 1)^n \quad \text{and} \quad \frac{d^q C_n^{(\alpha)}(\pm 1)}{dx^q} = (\pm 1)^{n+q} \prod_{k=0}^{q-1} \frac{(n - k)(n + k + 2\alpha)}{(2k + 2\alpha + 1)}$$

will be of important use later.

Let $f(x)$ be an infinitely differentiable function defined on the closed interval $[-1, 1]$. Then we can write $f(x) = \sum_{n=0}^{\infty} b_n^{(q)} C_n^{(\alpha)}(x)$, and for the l th derivative of $f(x)$ in the form

$$f^{(l)}(x) = \sum_{n=0}^{\infty} b_n^{(q-l)} C_n^{(\alpha)}(x), \quad l = 0, 1, \dots, q - 1,$$

and in particular $f^{(q)}(x) = \sum_{n=0}^{\infty} b_n C_n^{(\alpha)}(x)$, $b_n = b_n^{(0)}$, then

$$A_n^{(q)} = 2^{-q} \sum_{j=0}^q (-1)^j \binom{q}{j} \frac{(n - q + 2j + \alpha)\Gamma(n - q + j + \alpha)}{\Gamma(n + j + \alpha + 1)} A_{n-q+2j},$$

where $n \geq q$ for $\alpha \neq 0$, $n > q$ for $\alpha = 0$, and

$$b_n^{(q)} = \frac{2(n + \alpha)\Gamma(n + 2\alpha)}{n!\Gamma(\alpha + 1/2)} A_n^{(q)}, \quad l = 0, 1, \dots, q.$$

If we define the q times repeated integration of $C_n^{(\alpha)}(x)$ by

$$I_n^{(q,\alpha)}(x) = \underbrace{\int \dots \int}_{q \text{ times}} C_n^{(\alpha)}(x) dx \dots dx,$$

then

$$I_n^{(q,\alpha)}(x) = \frac{2^{-q} n!}{\Gamma(n + 2\alpha)} \sum_{j=0}^q (-1)^j \binom{q}{j} \frac{\Gamma(n - j + \alpha)\Gamma(n + q - 2j + 2\alpha)}{(n + q - 2j)!\Gamma(n + q - j + \alpha + 1)} \times (n + q - 2j + \alpha) C_{n+q-2j}^{(\alpha)}(x), \tag{2.2}$$

where $q \geq 0$, $n \geq q + 1$ for $\alpha = 0$; $q \geq 0$, $n \geq q$ for $\alpha \neq 0$. For the proof, see Doha [9]. Another interesting formula is

$$x^l C_k^{(\alpha)}(x) = \frac{k!}{2^l \Gamma(k + 2\alpha)} \sum_{j=0}^l \frac{\Gamma(k + l + 2\alpha - 2j)}{(k + l - 2j)!} a_{lj}(k) C_{k+l-2j}^{(\alpha)}(x) \tag{2.3}$$

for all $k, l \geq 0$, where, for $\alpha \neq 0$,

$$a_{lj}(k) = \begin{cases} \mu(k)a_{l-1,0}(k + 1), & j = 0, \\ \mu(k)a_{l-1,j}(k + 1) + \nu(k)a_{l-1,j-1}(k - 1), & j = 1, \dots, l - 1, l \geq 1, \\ \nu(k)a_{l-1,j-1}(k - 1), & j = l, \end{cases}$$

and $a_{lj}(k) = \binom{l}{j} (k + l - 2j)/k$, for $\alpha = 0$, and where $\mu(k) = (k + 1)/(k + \alpha)$, $\nu(k) = 2 - \mu(k)$, $a_{00}(k) = 1$. For the proof, see Doha [8].

The following five lemmas will be of fundamental importance in what follows.

LEMMA 2.1.

$$(1 - x^2)C_k^{(\alpha)}(x) = -\alpha_k C_{k-2}^{(\alpha)}(x) + (\alpha_k + \beta_k)C_k^{(\alpha)}(x) - \beta_k C_{k+2}^{(\alpha)}(x), \tag{2.4}$$

where

$$\alpha_k = \frac{k(k - 1)}{4(k + \alpha)(k + \alpha - 1)}, \quad \beta_k = \frac{(k + 2\alpha)(k + 2\alpha + 1)}{4(k + \alpha)(k + \alpha + 1)}.$$

PROOF. Set $l = 2$ in relation (2.3) to get

$$x^2 C_k^{(\alpha)}(x) = \alpha_k C_{k-2}^{(\alpha)}(x) + (1 - \alpha_k - \beta_k) C_k^{(\alpha)}(x) + \beta_k C_{k+2}^{(\alpha)}(x),$$

and hence we have relation (2.4). □

LEMMA 2.2.

$$I_n^{(2,\alpha)}(x) = \delta_1(n) C_{n-2}^{(\alpha)}(x) + \delta_2(n) C_n^{(\alpha)}(x) + \delta_3(n) C_{n+2}^{(\alpha)}(x), \tag{2.5}$$

where

$$\begin{aligned} \delta_1(n) &= \begin{cases} 0, & n = 0, 1, \\ \frac{(n-1)_2}{4(n+\alpha-1)_2(n+2\alpha-2)_2}, & n \geq 2, \end{cases} \\ \delta_2(n) &= \begin{cases} \frac{1}{4(n+\alpha+1)}, & n = 0, 1, \\ \frac{-1}{2(n+\alpha-1)(n+\alpha+1)}, & n \geq 2, \end{cases} \\ \delta_3(n) &= \frac{(n+2\alpha)_2}{4(n+1)_2(n+\alpha)_2}, \end{aligned} \tag{2.6}$$

and in particular, for $n = 2$ and $\alpha = 0$

$$\delta_1(2) = -3/16. \tag{2.7}$$

PROOF. By setting $q = 2$ in relation (2.2), we obtain

$$I_n^{(2,\alpha)}(x) = \frac{1}{4} \left[\frac{(n-1)_2}{(n+\alpha-1)_2(n+2\alpha-2)_2} C_{n-2}^{(\alpha)}(x) - \frac{2}{(n+\alpha-1)(n+\alpha+1)} C_n^{(\alpha)}(x) + \frac{(n+2\alpha)_2}{4(n+1)_2(n+\alpha)_2} C_{n+2}^{(\alpha)}(x) \right],$$

where $n \geq 3$ for $\alpha = 0$ and $n \geq 2$ for $\alpha \neq 0$,

$$\begin{aligned} I_0^{(2,\alpha)}(x) &= \frac{1}{4(1+\alpha)} [C_0^{(\alpha)}(x) + (1+2\alpha)C_2^{(\alpha)}(x)], \\ I_1^{(2,\alpha)}(x) &= \frac{1}{12(2+\alpha)} [3C_1^{(\alpha)}(x) + (1+2\alpha)C_3^{(\alpha)}(x)], \\ I_2^{(2,0)}(x) &= -\frac{3}{16}T_0(x) - \frac{1}{6}T_2(x) + \frac{1}{48}T_4(x), \end{aligned}$$

and hence we have relations (2.6) and (2.7). □

LEMMA 2.3.

$$I_n^{(1,\alpha)}(x) = \gamma_1(n)C_{n-1}^{(\alpha)}(x) + \gamma_2(n)C_{n+1}^{(\alpha)}(x),$$

where

$$\gamma_1(n) = \begin{cases} 0, & n = 0, \\ \frac{-n}{2(n+\alpha)(n+2\alpha-1)}, & n \geq 1, \end{cases} \quad \gamma_2(n) = \frac{(n+2\alpha)}{2(n+\alpha)(n+1)}, \quad (2.8)$$

and in particular, for $n = 1$ and $\alpha = 0$,

$$\gamma_1(1) = -1/4. \quad (2.9)$$

PROOF. If we set $q = 1$ in relation (2.2), we get

$$I_n^{(1,\alpha)}(x) = \gamma_1(n)C_{n-1}^{(\alpha)}(x) + \gamma_2(n)C_{n+1}^{(\alpha)}(x),$$

where $n \geq 2$ for $\alpha = 0$, and $n \geq 1$ for $\alpha \neq 0$,

$$I_0^{(1,\alpha)}(x) = C_1^{(\alpha)}(x), \quad \text{and} \quad I_1^{(1,0)}(x) = 1/4[-T_0(x) + T_2(x)],$$

and hence we have relations (2.8) and (2.9). □

LEMMA 2.4. For $|i - j| \leq 2$

$$\left(\iint C_i^{(\alpha)}(x) dx dx, C_j^{(\alpha)}(x) \right)_w = \begin{cases} \delta_1(i)h_{i-2} & i = j + 2, \\ \delta_2(i)h_i & i = j, \\ \delta_3(i)h_{i+2} & i = j - 2, \\ 0, & i + j \text{ odd.} \end{cases}$$

PROOF. From relation (2.5) and making use of the orthogonality relation (2.1), we can show for $i + j$ odd that

$$\left(\iint C_i^{(\alpha)}(x) dx dx, C_j^{(\alpha)}(x) \right)_w = 0,$$

but for $|i - j| \leq 2$, $i + j$ even, and after performing some lengthy manipulations, we get

$$\left(\iint C_i^{(\alpha)} dx dx, C_j^{(\alpha)} \right)_w = \begin{cases} \delta_1(i)h_{i-2} & i = j + 2, \\ \delta_2(i)h_i & i = j, \\ \delta_3(i)h_{i+2} & i = j - 2. \end{cases} \quad \square$$

LEMMA 2.5. For $|i - j| \leq 1$,

$$\left(\int C_i^{(\alpha)}(x) dx, C_j^{(\alpha)}(x) \right)_w = \begin{cases} \gamma_1(i)h_{i-1}, & i = j + 1, \\ \gamma_2(i)h_{i+1}, & i = j - 1, \\ 0 & i + j \text{ even.} \end{cases}$$

PROOF. The proof of this lemma is not difficult, and it can be accomplished by following a similar procedure to that of Lemma 2.4. □

3. Helmholtz equations

We are interested in using the UGM to solve the Helmholtz equation

$$\beta_1 u - \Delta u = g \quad \text{in } \Omega = I^d, \quad u|_{\partial\Omega} = 0, \tag{3.1}$$

where $I = (-1, 1)$ and $d = 1, 2$, but by considering its integrated form.

Let us first introduce some basic notation that will be used in the upcoming sections. We set $S_N = \text{span}\{C_0^{(\alpha)}(x), C_1^{(\alpha)}(x), \dots, C_{N+2}^{(\alpha)}(x)\}$, $V_N = \{v \in S_N : v(\pm 1) = 0\}$, then the standard ultraspherical-Galerkin approximation to (3.1) is to find $u_n \in V_N^d$ such that

$$\beta_1 (u_N, v)_w - (\Delta u_N, v)_w = (g, v)_w \quad \text{for all } v \in V_N^d, \tag{3.2}$$

where $\omega(x) = \prod_{i=1}^d (1 - x_i^2)^{\alpha-1/2}$ and $(u, v)_w = \int_{\Omega} uv\omega dx$ is the scalar product in the weighted space $L^2_{\omega}(\Omega)$. The norm in $L^2_{\omega}(\Omega)$ will be denoted by $\|\cdot\|_{\omega}$.

Let us denote $H^s_{\omega}(\Omega)$ as the weighted Sobolev spaces with the norm $\|v\|_{s,\omega}$. It is well known (see, for example, Canuto *et al.* [4]) that for $\beta_1 \geq 0, s \geq 1$ and $u \in H^s_{\omega}(\Omega)$, that the following optimal error estimate holds:

$$\|u - u_N\|_{\omega} + N \|u - u_N\|_{1,\omega} \leq C(s)N^{-s} \|u\|_{s,\omega}.$$

Although approximation (3.2) achieves the optimal convergence rate, its practical value depends on the choice of a basis for V_N^d . It is essential for the sake of efficiency to choose an appropriate basis for V_N^d such that the resulting linear system is as simple as possible.

However, to the best of our knowledge, the only Chebyshev and Legendre bases available in the literature are

- (1) $V_N = \text{span}\{\phi_2(x), \phi_3(x), \dots, \phi_N(x)\}$ with

$$\phi_k(x) = \begin{cases} T_k(x) - T_0(x), & k \text{ even,} \\ T_k(x) - T_1(x), & k \text{ odd,} \end{cases} \quad \text{or} \quad \phi_k(x) = (1 - x^2)T_{k-2}(x),$$

and

$$\phi_k(x) = \begin{cases} L_k(x) - L_0(x), & k \text{ even,} \\ L_k(x) - L_1(x), & k \text{ odd,} \end{cases} \quad \text{or} \quad \phi_k(x) = (1 - x^2)L_{k-2}(x);$$

see Canuto *et al.* [4], Heinrichs [14], and Gottlib and Orszag [12]. Unfortunately, all these bases lead to linear systems with full matrices even in the simplest case $\beta_1 = 0$.

(2) $V_N = \text{span}\{\phi_0(x), \phi_1(x), \dots, \phi_{N-2}(x)\}$ with $\phi_k(x) = T_k(x) - T_{k+2}(x)$ and $\phi_k(x) = L_k(x) - L_{k+2}(x)$; see Shen [21, 22].

These bases lead to discrete systems with either special matrices that can be efficiently inverted or sparse matrices for variational formulations.

The interested reader is referred to Doha and Abd-Elhameed [10] for efficient spectral-Galerkin algorithms for direct solution of second-order equations using ultraspherical polynomials.

An alternative approach to differentiating solution expansions is to integrate the differential equation q times, where q is the order of the equation. An advantage of this approach is that the general equation in the algebraic system then contains a finite number of terms. Phillips and Karageorghis [20] have followed this approach, see also Fox and Parker [11]. An explicit formula expressing the integration of ultraspherical polynomials in terms of ultraspherical polynomials themselves is given in Doha [9].

3.1. One-dimensional Helmholtz equation It is of fundamental importance to note here that the crucial task in applying the Galerkin-spectral approximations is how to choose an appropriate basis for V_N such that the linear systems resulting from (3.2) and its integrated form are as simple as possible.

In this section, we consider two kinds of bases to numerically solve the one-dimensional Helmholtz equation

$$\beta_1 u(x) - u''(x) = g(x) \quad \text{in } I = (-1, 1), \quad u(\pm 1) = 0,$$

but by considering its integrated form, namely,

$$\beta_1 \iint u(x) dx dx - u(x) = f(x) + d_0 C_0^{(\alpha)}(x) + d_1 C_1^{(\alpha)}(x) \quad \text{in } I = (-1, 1), \tag{3.3}$$

$$u(\pm 1) = 0.$$

(i) The first choice of basis. We choose the basis functions of the expansion

$$\phi_k(x) = (1 - x^2)C_k^{(\alpha)}(x), \quad k = 0, 1, \dots, N - 2, \tag{3.4}$$

which fulfills the boundary conditions of (3.3).

LEMMA 3.1. *We have, for arbitrary constants a_k ,*

$$\sum_{k=0}^{N-2} a_k(1-x^2)C_k^{(\alpha)}(x) = \sum_{k=0}^N e_k C_k^{(\alpha)}(x), \tag{3.5}$$

where $e_k = -\beta_{k-2}a_{k-2} + (\alpha_k + \beta_k)a_k - \alpha_{k+2}a_{k+2}$, and

$$\gamma_k = \frac{k(k-1)}{4(k+\alpha)(k+\alpha-1)}, \quad \beta_k = \frac{(k+2\alpha)(k+2\alpha+1)}{4(k+\alpha)(k+\alpha+1)}.$$

Moreover, if

$$\iint \sum_{k=0}^{N-2} a_k(1-x^2)C_k^{(\alpha)}(x) dx dx = \sum_{k=0}^{N+2} E_k C_k^{(\alpha)}(x), \tag{3.6}$$

then $E_k = e_{k-2}\delta_3(k-2) + e_k\delta_2(k) + e_{k+2}\delta_1(k+2)$, $k = 0, 1, \dots, N+2$.

PROOF. If we write $\sum_{k=0}^{N-2} a_k(1-x^2)C_k^{(\alpha)}(x) = \sum_{k=0}^N e_k C_k^{(\alpha)}(x)$, then it is not difficult to show, by using Lemma 2.1, that $e_k = -\beta_{k-2}a_{k-2} + (\alpha_k + \beta_k)a_k - \alpha_{k+2}a_{k+2}$. Since

$$\iint \left[\sum_{k=0}^{N-2} a_k(1-x^2)C_k^{(\alpha)}(x) \right] dx dx = \sum_{k=0}^N e_k \iint C_k^{(\alpha)}(x) dx dx,$$

then making use of Lemma 2.2 (Formula (2.5)) immediately yields

$$\begin{aligned} & \iint \left[\sum_{k=0}^{N-2} a_k(1-x^2)C_k^{(\alpha)}(x) \right] dx dx \\ &= \sum_{k=0}^N e_k [\delta_1(k)C_{k-2}^{(\alpha)}(x) + \delta_2(k)C_k^{(\alpha)}(x) + \delta_3(k)C_{k+2}^{(\alpha)}(x)]. \end{aligned}$$

Finally, if we write

$$\iint \left[\sum_{k=0}^{N-2} a_k(1-x^2)C_k^{(\alpha)}(x) \right] dx dx = \sum_{k=0}^{N+2} E_k C_k^{(\alpha)}(x),$$

then $E_k = e_{k-2}\delta_3(k-2) + e_k\delta_2(k) + e_{k+2}\delta_1(k+2)$, and this completes the proof of Lemma 3.1. □

THEOREM 3.2. *If $u_N(x) = \sum_{k=0}^{N-2} a_k(1-x^2)C_k^{(\alpha)}(x)$ is the Galerkin approximation to (3.3), then the expansion coefficients $\{a_k, k = 0, 1, \dots, N-2\}$ satisfy the matrix*

system $(\beta_1 A + E)\mathbf{a} = \mathbf{f}^*$, where the nonzero elements of matrices A and E are given by

$$\begin{aligned}
 a_{k,k-2} &= \frac{-(k + 2\alpha - 4)_4}{16(k - 1)_2(k + \alpha - 4)_4}, & k = 4, \dots, N, \\
 a_{k,k} &= \frac{(k + 2\alpha - 2)_2(2k^2 + (2\alpha - 5)k + 2\alpha^2 - \alpha - 3)}{8(k - 1)_2(k + \alpha - 3)_3(k + \alpha + 1)}, & k = 2, \dots, N, \\
 a_{k,k+2} &= \frac{-3(k^2 + 2\alpha k + 2\alpha^2 + \alpha - 4)}{8(k + \alpha - 2)_2(k + \alpha + 1)_2}, & k = 2, \dots, N - 2, \\
 a_{k,k+4} &= \frac{(k + 1)_2(2k^2 + (6\alpha + 5)k + 6\alpha^2 + 9\alpha - 3)}{8(k + \alpha - 1)(k + \alpha + 1)_3(k + 2\alpha)_2}, & k = 2, \dots, N - 4, \\
 a_{k,k+6} &= \frac{-16(k + 1)_4}{(k + \alpha + 1)_4(k + 2\alpha)_2}, & k = 2, \dots, N - 6, \\
 e_{k,k} &= \frac{(k + 2\alpha - 2)_2}{4(k + \alpha - 2)_2}, & k = 2, \dots, N, \\
 e_{k,k+2} &= -\frac{(k^2 + 2\alpha k + 2\alpha^2 - \alpha - 1)}{2(k + \alpha - 1)(k + \alpha + 1)}, & k = 2, \dots, N - 2, \\
 e_{k,k+4} &= \frac{(k + 1)_2}{4(k + \alpha + 1)_2}, & k = 2, \dots, N - 4,
 \end{aligned}$$

and $\mathbf{a} = (a_0, a_1, \dots, a_{N-2})^T$, $\mathbf{f}^* = (f_2^*, f_3^*, \dots, f_N^*)^T$, $f_k^* = (1/h_k)(f, C_k^{(\alpha)}(x))_w$, where h_k is as defined in relation (2.1).

PROOF. The application of the Galerkin method to (3.3) gives

$$\left(\beta_1 \iint u_N dx dx - u_N, \phi_k(x) \right)_w = (f(x) + d_0 C_0^{(\alpha)}(x) + d_1 C_1^{(\alpha)}(x), \phi_k(x))_w, \quad (3.7)$$

where $\phi_k(x)$ is given by (3.4). Substitution of formulae (3.5) and (3.6) into (3.7) yields

$$\begin{aligned}
 &\left(\beta_1 \sum_{j=0}^{N+2} E_j C_j^{(\alpha)}(x) - \sum_{j=0}^N e_j C_j^{(\alpha)}(x), (1 - x^2) C_k^{(\alpha)}(x) \right)_w \\
 &= (f(x) + d_0 + d_1 C_1^{(\alpha)}(x), (1 - x^2) C_k^{(\alpha)}(x))_w. \quad (3.8)
 \end{aligned}$$

The constants of integration d_0 and d_1 would disappear if we take $k \geq 4$ in (3.8), and if we substitute (2.4) into (3.8) and make use of the orthogonality relation (2.1), then we get

$$\begin{aligned}
 &-(\beta_1 E_{k-2} - e_{k-2})\alpha_k h_{k-2} + (\beta_1 E_k - e_k)(\alpha_k + \beta_k)h_k - (\beta_1 E_{k+2} - e_{k+2})\beta_k h_{k+2} \\
 &= -\alpha_k f_{k-2} + (\alpha_k + \beta_k) f_k - \beta_k f_{k+2}, \quad k = 4, 5, \dots, N + 2. \quad (3.9)
 \end{aligned}$$

Now, it is not difficult to show that (3.9) holds for $f_{k-2} = (\beta_1 E_{k-2} - e_{k-2})h_{k-2}$, $k = 4, 5, \dots, N + 2$, or

$$f_k = (\beta_1 E_k - e_k)h_k; \quad k = 2, 3, \dots, N; \tag{3.10}$$

this means that, instead of solving (3.9), we can solve the neater system (3.10), which is equivalent to (3.7). This linear system may be put in the form

$$(\beta_1 E_k - e_k) = f_k^*, \quad k = 2, 3, \dots, N, \tag{3.11}$$

which in turn may be written in the matrix form

$$(\beta_1 A + E)a = f^*, \tag{3.12}$$

and this completes the proof of Theorem 3.2. □

COROLLARY 3.3. *If $\alpha = 0$, then the nonzero elements of A and E in system (3.12) are given by*

$$\begin{aligned} a_{k,k-2} &= \begin{cases} -1/96, & k = 4, \\ \frac{-1}{16(k-1)_2}, & k \geq 5, \end{cases} & a_{k,k} &= \begin{cases} 5/24, & k = 2, \\ 5/192, & k = 3, \\ \frac{(2k+1)}{8(k-1)_3}, & k \geq 4, \end{cases} \\ a_{k,k+2} &= \begin{cases} -5/32, & k = 2, \\ \frac{-3}{8(k-1)(k+1)}, & k \geq 3, \end{cases} & a_{k,k+4} &= \frac{(2k-1)}{8(k-1)_3}, \\ a_{k,k+6} &= \frac{-1}{16(k)_2}, & e_{k,k} &= \begin{cases} 1/2, & k = 2, \\ 1/4, & k \geq 3, \end{cases} \\ e_{k,k+2} &= -1/2, & e_{k,k+4} &= 1/4. \end{aligned}$$

COROLLARY 3.4. *If $\alpha = 1$, then the nonzero elements of A and E in system (3.12) are given by*

$$\begin{aligned} a_{k,k-2} &= \frac{-(k+1)}{16(k-3)(k-1)_2}, & a_{k,k} &= \begin{cases} 21/64, & k = 2, \\ \frac{(k+1)(2k+1)}{8(k-1)^2 k(k+2)}, & k \geq 3, \end{cases} \\ a_{k,k+2} &= \frac{-3(k^2+2k-1)}{8(k-1)_2(k+2)_2}, & a_{k,k+4} &= \frac{(k+1)(2k+3)}{8k(k+2)(k+3)^2}, \\ a_{k,k+6} &= \frac{-(k+1)}{(k+2)_2(k+5)}, & e_{k,k} &= \frac{(k+1)}{4(k-1)}, \\ e_{k,k+2} &= -1/2, & e_{k,k+4} &= \frac{(k+1)}{4(k+3)}. \end{aligned}$$

COROLLARY 3.5. *If $\alpha = 1/2$, then the nonzero elements of A and E in system (3.12) are given by*

$$\begin{aligned}
 a_{k,k-2} &= \frac{-(k-3)_2}{2^4(k-7/2)_4}, & a_{k,k} &= \frac{2(2k^2-4k-3)}{2^3(2k-5/2)_3(2k+3)}, \\
 a_{k,k+2} &= \frac{-6(k^2+k-3)}{2^4(k-3/2)_2(k+3/2)_2}, & a_{k,k+4} &= \frac{(4k^2+16k+9)}{2^3(2k-1)(k+3/2)_3}, \\
 a_{k,k+6} &= \frac{-(k+3)_2}{2^4(k+3/2)_4}, & e_{k,k} &= \frac{(k-1)_2}{2^2(k-3/2)_2}, \\
 e_{k,k+2} &= \frac{-2(k^2+k-1)}{(2k-1)(2k+3)}, & e_{k,k+4} &= \frac{(k+1)_2}{2^2(k+3/2)_2}.
 \end{aligned}$$

In particular, if $\beta_1 = 0, \alpha \in (-1/2, \infty)$, then the system in this case is $Ea = f^*$, where the matrix E is a special upper triangular matrix.

REMARK 3.6. If $\beta_1 \neq 0$ and $\alpha \in (-1/2, \infty)$, then the system (3.12) can be solved by LU factorization.

(ii) The second choice of basis. Here, we choose $\phi_k(x)$ to be of the form

$$\phi_k(x) = C_k^{(\alpha)}(x) - C_{k+2}^{(\alpha)}(x). \tag{3.13}$$

This choice may be considered as a generalization of the bases used by Shen [18, 19].

It is now clear that (3.3) is equivalent, for $k = 0, 1, \dots, N$, to

$$\beta_1 \left(\iint u_N dx dx, \phi_k(x) \right)_w - (u_N, \phi_k(x))_w = (f(x) + d_0 + d_1 C_1^{(\alpha)}(x), \phi_k(x))_w. \tag{3.14}$$

The constants d_0 and d_1 would not appear if we take $k \geq 2$ in (3.14), therefore we get

$$\begin{aligned}
 \beta_1 \left(\iint u_N dx dx, C_k^{(\alpha)}(x) - C_{k+2}^{(\alpha)}(x) \right)_w - (u_N, C_k^{(\alpha)}(x) - C_{k+2}^{(\alpha)}(x))_w \\
 = (f(x), C_k^{(\alpha)}(x) - C_{k+2}^{(\alpha)}(x))_w, \quad k = 2, 3, \dots, N.
 \end{aligned} \tag{3.15}$$

Let us denote $f_k = (f, \phi_k(x))_w, f = (f_2, \dots, f_N)^T, u_N(x) = \sum_{n=2}^N a_{n-2} \phi_{n-2}(x), a = (a_0, \dots, a_{N-2})^T, A = (a_{kj})_{2 \leq k, j \leq N}$ and $B = (b_{kj})_{2 \leq k, j \leq N}$, then (3.15) is equivalent to the matrix equation

$$(\beta_1 A + B)a = f, \tag{3.16}$$

where the nonzero elements of matrices A and B are given explicitly in the following theorem.

THEOREM 3.7. *If we take $\phi_k(x)$ as defined in (3.13), and if we denote*

$$a_{kj} = \left(\iint \phi_{j-2}(x) dx dx, \phi_k(x) \right)_w \quad \text{and} \quad b_{kj} = -(\phi_{j-2}(x), \phi_k(x))_w,$$

then $V_N = \text{span}\{\phi_0(x), \phi_1(x), \dots, \phi_N(x)\}$, and the nonzero elements $(a_{kj}), (b_{kj}), 2 \leq k, j \leq N$, are given by

$$\begin{aligned} a_{k,k-2} &= \frac{-\Gamma(1/2)\Gamma(\alpha + 1/2)(k - 2)!}{4\Gamma(\alpha)(2\alpha)_{k-2}(k + \alpha - 2)_3}, \\ a_{k,k} &= \frac{\Gamma(1/2)\Gamma(\alpha + 1/2)k!(2k^2 + (2\alpha - 3)k + 2\alpha^2 + 3\alpha - 2)}{2\Gamma(\alpha)(2\alpha)_k(k - 1)_2(k + \alpha - 2)(k + \alpha)(k + \alpha + 2)}, \\ a_{k,k+2} &= \frac{-3\Gamma(1/2)\Gamma(\alpha + 1/2)k!(k + \alpha + 1)(2k^2 + 4(\alpha + 1)k + 4\alpha^2 + 10\alpha - 2)}{4\Gamma(\alpha)(2\alpha)_k(k + \alpha - 1)_2(k + \alpha + 2)_2(k + 2\alpha)_2}, \\ a_{k,k+4} &= \frac{\Gamma(1/2)\Gamma(\alpha + 1/2)(k + 2)!(2k^2 + (6\alpha + 11)k + 6\alpha^2 + 21\alpha + 12)}{2\Gamma(\alpha)(2\alpha)_{k+2}(k + \alpha)(k + \alpha + 2)(k + \alpha + 4)(k + 2\alpha + 2)_2}, \\ a_{k,k+6} &= \frac{-\Gamma(1/2)\Gamma(\alpha + 1/2)(k + 4)!}{4\Gamma(\alpha)(2\alpha)_{k+4}(k + \alpha + 2)_3}, \quad b_{kk} = \frac{\Gamma(1/2)\Gamma(\alpha + 1/2)k!}{\Gamma(\alpha)(2\alpha)_k(k + \alpha)}, \\ b_{k,k+2} &= \frac{-\Gamma(1/2)\Gamma(\alpha + 1/2)k!(k + \alpha + 1)(2k^2 + (4\alpha + 4)k + 4\alpha^2 + 6\alpha)}{\Gamma(\alpha)(2\alpha)_k(k + \alpha)(k + \alpha + 2)(k + 2\alpha)_2}, \\ b_{k,k+4} &= \frac{\Gamma(1/2)\Gamma(\alpha + 1/2)(k + 2)!}{\Gamma(\alpha)(2\alpha)_{k+2}(k + \alpha + 2)}. \end{aligned}$$

PROOF. The proof of this theorem is not difficult, but it is rather lengthy, and it can be accomplished by following the same procedure used in proving Theorem 3.2. \square

REMARK 3.8. For $k + j$ odd, $a_{kj} = b_{kj} = 0$. Hence system (3.16) of order $N - 1$ can be decoupled into two separate systems of order $N/2$ and $(N/2 - 1)$, respectively. In this way one needs to solve two systems of order n instead of one of order $2n$, which leads to substantial savings.

COROLLARY 3.9. *If $\alpha = 0$, then the nonzero elements $(a_{kj}), (b_{kj}),$ for $2 \leq k, j \leq N$, are given as follows:*

$$\begin{aligned} a_{k,k-2} &= \frac{-\pi}{8(k - 1)_2}, & a_{k,k} &= \begin{cases} \frac{7\pi}{8(2k - 3)(k + 2)}, & k = 2, 3, \\ \frac{\pi(2k + 1)}{4(k - 1)_2(k + 2)}, & k = 4, \dots, N, \end{cases} \\ a_{k,k+2} &= \frac{-3\pi(k^2 + 2k - 1)}{4(k - 1)_2(k + 2)_2}, & a_{k,k+4} &= \frac{\pi(2k + 3)}{4k(k + 2)_2}, \\ a_{k,k+6} &= \frac{-\pi}{8(k + 2)_2}, & b_{kk} &= \pi/2, \quad b_{k,k+2} = -\pi, \quad b_{k,k+4} = \pi/2. \end{aligned}$$

COROLLARY 3.10. *If $\alpha = 1$, then the nonzero elements $(a_{kj}), (b_{kj})$, for $2 \leq k, j \leq N$, are given as follows:*

$$\begin{aligned}
 a_{k,k-2} &= \frac{-\pi}{8(k-1)^2(k)_2}, & a_{k,k} &= \frac{\pi(2k^2 - k + 3)}{4k(k-1)^2(k+1)^2(k+3)}, \\
 a_{k,k+2} &= \frac{-3\pi(k^2 + 4k + 6)}{4(k-1)_3(k+3)}, & a_{k,k+4} &= \frac{\pi(2k^2 + 17k + 39)}{4(k+1)(k+3)^2(k+4)(k+5)^2}, \\
 a_{k,k+6} &= \frac{-\pi}{8(k+2)(k+4)(k+5)^2}, & b_{kk} &= \frac{\pi}{2(k+1)^2}, \\
 b_{k,k+2} &= \frac{-\pi(k^2 + 4k + 5)}{(k+1)^2(k+3)^2}, & b_{k,k+4} &= \frac{\pi}{2(k+3)^2}.
 \end{aligned}$$

COROLLARY 3.11. *If $\alpha = 1/2$, then the nonzero elements $(a_{kj}), (b_{kj})$, for $2 \leq k, j \leq N$, are given as follows:*

$$\begin{aligned}
 a_{k,k-2} &= \frac{-1}{4(k-3/2)_3}, & a_{k,k} &= \frac{8}{(2k-3)(2k+1)(2k+5)}, \\
 a_{k,k+2} &= \frac{-3(2k+3)}{4(k-1/2)_2(k+5/2)_2}, & a_{k,k+4} &= \frac{8}{(2k+1)(2k+5)(2k+9)}, \\
 a_{k,k+6} &= \frac{-1}{4(k+5/2)_3}, & b_{kk} &= \frac{2}{(2k+1)}, \\
 b_{k,k+2} &= \frac{-4(2k+3)}{(2k+1)(2k+5)}, & b_{k,k+4} &= \frac{2}{(2k+5)}.
 \end{aligned}$$

REMARK 3.12. If $\beta_1 \neq 0, \alpha \in (-1/2, \infty)$, we explicitly form the LU factorization $\beta_1 A + B = LU$.

3.2. Two-dimensional Helmholtz equation Now we consider the Helmholtz equation in two dimensions ($d = 2$)

$$\beta_1 u(x, y) - \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) = g(x, y) \quad \text{in } \Omega = (-1, 1)^2, \quad u|_{\partial\Omega} = 0,$$

but in its integrated form, namely

$$\begin{aligned}
 &\beta_1 \iint \left[\iiint u(x, y) dx dx \right] dy dy - \iint u(x, y) dx dx - \iint u(x, y) dy dy \\
 &= f(x, y) + r(x)C_0^{(\alpha)}(y) + s(x)C_1^{(\alpha)}(y) + t(y)C_0^{(\alpha)}(x) \\
 &+ v(y)C_1^{(\alpha)}(x), \quad \text{in } \Omega, \quad u|_{\partial\Omega} = 0,
 \end{aligned} \tag{3.17}$$

where $r(x), s(x), t(y)$ and $v(y)$ are arbitrary functions in x and y respectively, and

$$f(x, y) = \iint \left[\iiint g(x, y) dx dx \right] dy dy.$$

It is clear that if we take $\phi_k(x)$ as defined in (3.13), then $V_N^2 = \text{span}\{\phi_k(x)\phi_j(y) : k, j = 0, 1, \dots, N\}$. Let us denote $u_N = \sum_{k,j=2}^N u_{k-2,j-2}\phi_{k-2}(x)\phi_{j-2}(y)$, $f_{kj} = (f, \phi_k(x)\phi_j(y))_w$, $U = (u_{kj})_{k,j=0,1,\dots,N-2}$, $F = (f_{kj})_{k,j=2,3,\dots,N}$, then applying the Galerkin method to (3.17) we obtain the matrix equation

$$\beta_1 AUA^T + BUA^T + AUB^T = F, \quad (3.18)$$

where A and B are the matrices defined in Theorem 3.7.

This equation can be solved by the matrix decomposition method described in Buzbee *et al.* [3] and Haidvogel and Zang [13]. Now, let D be the diagonal matrix whose diagonal elements are the eigenvalues of $A^{-1}B$, and let E be the matrix formed by the eigenvectors of $A^{-1}B$. Then $A^{-1}BE = ED$. Applying A^{-1} to (3.17), we obtain $\beta_1 UA^T + A^{-1}BUA^T + UB^T = A^{-1}F$. Setting $U = EV$, the preceding equation becomes $\beta_1 EVA^T + EDVA^T + EVB^T = A^{-1}F$. Now applying E^{-1} to the preceding equation and setting $G = E^{-1}A^{-1}F$, we find

$$\beta_1 VA^T + DVA^T + VB^T = G. \quad (3.19)$$

Let $v_p = (v_{p0}, v_{p1}, \dots, v_{pN-2})^T$, $g_p = (g_{p0}, g_{p1}, \dots, g_{pN-2})^T$ for $p = 0, \dots, N-2$. Then the p th row of (3.18) becomes

$$(\beta_1 + \lambda_p)Av_p + Bv_p = g_p, \quad p = 0, 1, \dots, N-2, \quad (3.20)$$

which is equivalent to $(N-1)$ one-dimensional equations of the form (3.16). In summary, the solution of (3.17) consists of the following four steps:

- (1) Compute the eigenpairs D, E for $A^{-1}B$ and compute E^{-1} .
- (2) Compute $G = E^{-1}A^{-1}F$.
- (3) Obtain V by solving (3.19).
- (4) Set $U = EV$.

It is worth mentioning here that the previous procedure can be applied to (3.17), but with $\phi_k(x)$ as defined in (3.4). In this case we get the matrix equation

$$\beta_1 AUA^T + AUE^T + EUA^T = F,$$

where A and E are the matrices defined in Theorem 3.2. This system has the same method of solution as that of (3.18).

4. Extensions to more general problems in one dimension

The UGM can be applied to more general problems. In this section, we describe several extensions in one dimension.

4.1. Mixed-type boundary conditions When other boundary conditions are prescribed, it is necessary to construct a basis incorporating the boundary conditions. Now consider the equation $\beta_1 u(x) + \beta_2 u'(x) - u''(x) = g(x)$ in $I = (-1, 1)$, and instead, we deal with its integrated form

$$\beta_1 \iint u(x) dx dx + \beta_2 \int u(x) dx - u(x) = f(x) + d_0 + d_1 C_1^{(\alpha)}(x) \quad \text{in } I, \quad (4.1)$$

with the mixed homogeneous boundary conditions

$$\alpha_1 u(1) + \lambda_1 u'(1) = 0, \quad \alpha_2 u(-1) + \lambda_2 u'(-1) = 0.$$

Let $W_N = \{v \in S_N : \alpha_i v(\pm 1) + \lambda_i v'(\pm 1) = 0, i = 1, 2\}$, then the standard ultraspherical-Galerkin approximation to (4.1) is to find $u_N \in W_N$ such that, for all $v \in W_N$,

$$\begin{aligned} &\beta_1 \left(\iint u_N dx dx, v \right)_w + \beta_2 \left(\int u_N dx, v \right)_w - (u_N, v)_w \\ &= (f(x) + d_0 + d_1 C_1^{(\alpha)}(x), v)_w. \end{aligned} \quad (4.2)$$

We can find an appropriate basis for W_N by setting

$$\phi_k(x) = C_k^{(\alpha)}(x) + d_k C_{k+1}^{(\alpha)}(x) + e_k C_{k+2}^{(\alpha)}(x), \quad (4.3)$$

where d_k and e_k are the unique constants such that $\phi_k(x) \in W_N, k = 0, 1, \dots, N - 2$. Therefore $W_N = \text{span}\{\phi_0(x), \phi_1(x), \dots, \phi_N(x)\}$, where

$$d_k = \frac{4(2\alpha + 1)(\alpha_1 \lambda_2 + \alpha_2 \lambda_1)(k + \alpha + 1)}{\text{Den}(k)}, \quad e_k = -\frac{\text{Den}(k - 1)}{\text{Den}(k)}, \quad (4.4)$$

where

$$\begin{aligned} \text{Den}(k) &= 2(2\alpha + 1)^2 \alpha_1 \alpha_2 - 2\lambda_1 \lambda_2 (k + 1)_2 (k + 2\alpha + 1)_2 \\ &\quad - (2\alpha + 1)(\alpha_1 \lambda_2 - \alpha_2 \lambda_1) [(k + 1)(k + 2\alpha + 1) + (k + 2)(k + 2\alpha + 2)], \end{aligned}$$

and $u_N = \sum_{k=2}^N a_{k-2} \phi_{k-2}(x)$.

Now (4.2) is equivalent to the matrix equation

$$(\beta_1 A + \beta_2 B + C)a = f, \quad (4.5)$$

where the nonzero elements of the matrices $A = (a_{kj}), B = (b_{kj})$ and $C = (c_{kj}), 2 \leq k, j \leq N$, are given explicitly in the following theorem.

THEOREM 4.1. *Let $\phi_k(x)$ be as defined in (4.3) and denote*

$$a_{kj} = \left(\iint_w \phi_{j-2}(x) dx dx, \phi_k(x) \right),$$

$$b_{kj} = \left(\int \phi_{j-2}(x) dx, \phi_k(x) \right)_w,$$

$$c_{kj} = (\phi_{j-2}(x), \phi_k(x))_w.$$

Then the nonzero elements of the matrices A , B and C in system (4.5) are given as follows:

$$a_{k,k-2} = e_{k-4}\delta_3(k-2)h_k, \quad a_{k,k-1} = \delta_3(k-2)d_{k-3}h_k + e_{k-3}\delta_3(k-1)d_k h_{k+1},$$

$$a_{k,k} = (\delta_3(k-2) + e_{k-2}\delta_2(k))h_k + d_{k-2}d_k\delta_3(k-1)h_{k+1} + e_{k-2}e_k\delta_3(k)h_{k+2},$$

$$a_{k,k+1} = \delta_2(k)d_{k-1}h_k + (\delta_3(k-1) + e_{k-1}\delta_2(k+1))d_k h_{k+1} + e_k\delta_3(k)d_{k-1}h_{k+2},$$

$$a_{k,k+2} = (\delta_2(k) + e_k\delta_1(k+2))h_k + \delta_2(k+1)d_k^2 h_{k+1} + (\delta_3(k) + e_k\delta_2(k+2))e_k h_{k+2},$$

$$a_{k,k+3} = \delta_1(k+2)d_{k+1}h_k + (\delta_2(k+1) + e_{k+1}\delta_1(k+3))d_k h_{k+1} + e_k\delta_2(k+2)d_{k+1}h_{k+2},$$

$$a_{k,k+4} = \delta_1(k+2)h_k + \delta_1(k+3)d_k d_{k+2}h_{k+1} + (\delta_2(k+2) + e_{k+2}\delta_1(k+4))e_k h_{k+2},$$

$$a_{k,k+5} = \delta_1(k+3)d_k h_{k+1} + e_k\delta_1(k+4)d_{k+3}h_{k+2}, \quad a_{k,k+6} = e_k\delta_1(k+4)h_{k+2},$$

$$b_{k,k-1} = e_{k-3}\gamma_2(k-1)h_k, \quad b_{k,k} = \gamma_2(k-1)d_{k-2}h_k + e_{k-2}\gamma_2(k)d_k h_{k+1},$$

$$b_{k,k+1} = (\gamma_2(k-1) + e_{k-1}\gamma_1(k+1))h_k + \gamma_2(k)d_{k-1}d_k h_{k+1} + \gamma_2(k+1)e_{k-1}e_k h_{k+2},$$

$$b_{k,k+2} = \gamma_1(k+1)d_k h_k + (\gamma_2(k) + e_k\gamma_1(k+2))h_{k+1} + e_k\gamma_2(k+1)d_k h_{k+2},$$

$$b_{k,k+3} = \gamma_1(k+1)h_k + \gamma_1(k+2)d_k d_{k+1}h_{k+1} + (\gamma_2(k+1) + \gamma_1(k+3))e_k h_{k+2},$$

$$b_{k,k+4} = \gamma_1(k+2)d_k h_{k+1} + e_k\gamma_1(k+3)d_{k+2}h_{k+2}, \quad b_{k,k+5} = e_k\gamma_1(k+3)h_{k+2},$$

$$c_{k,k} = c_{k-2,k+2} = e_{k-2}h_k, \quad c_{k,k+2} = h_k + d_k^2 h_{k+1} + e_k^2 h_{k+2},$$

$$c_{k,k+1} = c_{k-1,k+2} = d_{k-1}h_k + d_k e_{k-1}h_{k+1},$$

where $\{\delta_1(k), \delta_2(k), \delta_3(k)\}$, $\{\gamma_1(k), \gamma_2(k)\}$ and $\{d_k, e_k\}$ are given by (2.6)–(2.7), (2.8)–(2.9) and (4.4), respectively, while h_k is defined by (2.1).

PROOF. The proof of this theorem is rather lengthy but not difficult, and it can be accomplished by following the same procedures used in proving Theorem 3.2 and by using Lemma 2.3. □

4.1.1. Dirichlet problem In the special case where $\alpha_1 = \alpha_2 = 1, \lambda_1 = \lambda_2 = 0$ (that is, the homogeneous Dirichlet boundary condition), we have, in particular, $d_k = 0, e_k = -1$, and accordingly the basis functions $\phi_k(x), k = 0, 1, \dots, N - 2$, take the form given in (3.13).

It is of fundamental importance to note here that all the results of Theorem 3.7 can be obtained immediately as a special case from Theorem 4.1.

4.1.2. *Neumann problem* In this case, where $\alpha_1 = \alpha_2 = 0, \lambda_1 = \lambda_2 = 1$, we have $d_k = 0, e_k = -k(k + 2\alpha)/(k + 2)(k + 2\alpha + 2)$ and

$$\phi_k(x) = C_k^{(\alpha)}(x) - \frac{k(k + 2\alpha)}{(k + 2)(k + 2\alpha + 2)} C_{k+2}^{(\alpha)}(x). \tag{4.6}$$

In summary, all the results of such a case are given in the following theorem and its corollaries.

THEOREM 4.2. *Let $\phi_k(x)$ be defined as in (4.6), then the nonzero elements of matrices A, B and C in system (4.5) are given as follows:*

$$a_{k,k-2} = -\frac{\Gamma(1/2)\Gamma(\alpha + 1/2)(k - 3)!(k - 4)(k + 2\alpha - 4)(k + 2\alpha - 1)}{4\Gamma(\alpha)(2\alpha)_k(k + \alpha - 2)_3},$$

$$a_{k,k} = \frac{\Gamma(1/2)\Gamma(\alpha + 1/2)(k - 2)!(k + 2\alpha - 2)}{2\Gamma(\alpha)(2\alpha)_{k+1}(k + 2)(k + \alpha - 2)(k + \alpha)(k + \alpha + 2)(k + 2\alpha + 2)} \\ \times [2k^4 + k^3(6\alpha + 1) + k^2(6\alpha^2 + 9\alpha + 4) + 2k(2\alpha^3 + 13\alpha^2 + 13\alpha - 6) + 4(2\alpha^3 + 7\alpha^2 + 10\alpha + 8)],$$

$$a_{k,k+2} = -\frac{\Gamma(1/2)\Gamma(\alpha + 1/2)k!(k + \alpha + 1)}{2\Gamma(\alpha)(2\alpha)_k(k + 2)(k + \alpha - 1)_2(k + \alpha + 2)_2(k + 2\alpha + 1)(k + 2\alpha + 2)^2} \\ \times [3k^4 + 12k^3(\alpha + 1) + k^2(18\alpha^2 + 55\alpha + 41) + 2k(6\alpha^3 + 37\alpha^2 + 60\alpha + 29) + 8(2\alpha^3 + 11\alpha^2 + 17\alpha + 6)],$$

$$a_{k,k+4} = \frac{\Gamma(1/2)\Gamma(\alpha + 1/2)(k + 1)!(k + \alpha + 1)}{2\Gamma(\alpha)(2\alpha)_{k+2}(k + \alpha)_3(k + \alpha + 4)(k + 2\alpha + 2)_3} \\ \times [2k^4 + 5k^3(2\alpha + 3) + k^2(18\alpha^2 + 63\alpha + 46) + 2k(6\alpha^3 + 35\alpha^2 + 61\alpha + 40) + 4(2\alpha^3 + 15\alpha^2 + 34\alpha + 24)],$$

$$a_{k,k+6} = -\frac{\Gamma(1/2)\Gamma(\alpha + 1/2)k(k + 4)!(k + 2\alpha)}{4\Gamma(\alpha)(2\alpha)_{k+2}(k + 2)(k + \alpha + 2)_3(k + 2\alpha + 2)^2(k + 2\alpha + 3)},$$

$$b_{k,k-1} = -\frac{\Gamma(1/2)\Gamma(\alpha + 1/2)(k - 3)(k - 2)!(k + 2\alpha - 3)}{2\Gamma(\alpha)(2\alpha)_k(k + \alpha - 1)(k + \alpha)},$$

$$b_{k,k+1} = \frac{\Gamma(1/2)\Gamma(\alpha + 1/2)(k - 1)!(k + 2\alpha - 1)}{2\Gamma(\alpha)(2\alpha)_k(k + 2)(k + \alpha - 1)(k + \alpha + 2)(k + 2\alpha)_3} \\ \times [3k^4 + k^3(10\alpha + 7) + 4k^2(3\alpha^2 + 8\alpha + 5) + 4k(2\alpha^3 + 13\alpha^2 + 16\alpha + 2) + 8(2\alpha^3 + 7\alpha^2 + 7\alpha + 2)],$$

$$b_{k,k+3} = -\frac{\Gamma(1/2)\Gamma(\alpha + 1/2)(k + 1)!}{2\Gamma(\alpha)(2\alpha)_{k+1}(k + 2)(k + \alpha)(k + \alpha + 3)(k + 2\alpha + 2)^2(k + 2\alpha + 3)} \\ \times [3k^4 + k^3(14\alpha + 17) + 2k^2(12\alpha^2 + 37\alpha + 25) + 4k(4\alpha^3 + 22\alpha^2 + 36\alpha + 21) + 8(2\alpha^3 + 11\alpha^2 + 18\alpha + 9)],$$

$$\begin{aligned}
 b_{k,k+5} &= \frac{\Gamma(1/2)\Gamma(\alpha + 1/2)(k + 3)k(k + 1)!(k + 2\alpha)}{2\Gamma(\alpha)(2\alpha)_{k+2}(k + \alpha + 2)_2(k + 2\alpha + 2)^2}, \\
 c_{k,k} &= c_{k-2,k+2} = -\frac{\Gamma(1/2)\Gamma(\alpha + 1/2)(k - 2)(k - 1)!(k + 2\alpha - 2)}{\Gamma(\alpha)(2\alpha)_k(k + \alpha)(k + 2\alpha)}, \\
 c_{k,k+2} &= \frac{2\Gamma(1/2)\Gamma(\alpha + 1/2)k!(k + \alpha + 1)(k + 2\alpha)}{\Gamma(\alpha)(2\alpha)_{k+2}(k + 2)(k + \alpha)(k + \alpha + 2)(k + 2\alpha + 2)^2} \\
 &\quad \times [k^4 + 4k^3(\alpha + 1) + k^2(6\alpha^2 + 19\alpha + 12) \\
 &\quad + k(4\alpha^3 + 26\alpha^2 + 38\alpha + 16) + 4(2\alpha^3 + 7\alpha^2 + 7\alpha + 2)].
 \end{aligned}$$

COROLLARY 4.3. *If $\alpha = 0$, then the nonzero elements of matrices A , B and C in system (4.5) are given by*

$$\begin{aligned}
 a_{k,k-2} &= -\frac{\pi(k - 4)^2}{8(k - 2)(k - 2)_3}, & a_{k,k} &= \frac{\pi(2k^4 + k^3 + 4k^2 - 12k + 32)}{4(k - 1)k^2(k + 2)^3}, \\
 a_{k,k+2} &= -\frac{\pi(3k^4 + 12k^3 + 41k^2 + 58k + 48)}{4(k - 1)(k + 2)^4(k + 3)}, \\
 a_{k,k+4} &= \frac{\pi(2k^4 + 15k^3 + 46k^2 + 80k + 96)}{4k(k + 2)^2(k + 3)(k + 4)^2}, \\
 a_{k,k+6} &= -\frac{\pi k^2}{8(k + 2)^3(k + 3)}, & b_{k,k-1} &= -\frac{\pi(k - 3)^2}{4(k - 1)^2k}, \\
 b_{k,k+1} &= \frac{\pi(3k^4 + 7k^3 + 20k^2 + 8k + 16)}{4(k)_3(k + 2)^2}, \\
 b_{k,k+3} &= -\frac{\pi(k + 1)(3k^4 + 17k^3 + 50k^2 + 84k + 72)}{4k(k + 2)^3(k + 3)^2}, & b_{k,k+5} &= \frac{\pi k^2}{4(k + 2)^3}, \\
 c_{k,k} &= c_{k-2,k+2} = -\frac{\pi(k - 2)^2}{2k^2}, & c_{k,k+2} &= \frac{\pi(k^4 + 4k^3 + 12k^2 + 16k + 8)}{(k + 2)^4}.
 \end{aligned}$$

COROLLARY 4.4. *If $\alpha = 1$, then the nonzero elements of matrices A , B and C in system (4.5) are given by*

$$\begin{aligned}
 a_{k,k-2} &= -\frac{\pi(k - 4)}{8(k - 1)_2(k - 1)_3}, & a_{k,k} &= \frac{\pi(2k^4 + 7k^3 + 19k^2 + 44k + 108)}{4(k - 1)^2(k + 1)_2(k + 1)_4}, \\
 a_{k,k+2} &= -\frac{3\pi(k^2 + 4k + 6)(k^2 + 4k + 16)}{4(k)(k + 1)^2(k + 3)^2(k + 4)^3}, \\
 a_{k,k+4} &= \frac{\pi(2k^4 + 25k^3 + 127k^2 + 285k + 300)}{4(k + 1)_6(k + 3)(k + 5)}, & a_{k,k+6} &= -\frac{\pi k}{8(k + 3)_3(k + 4)_2}, \\
 b_{k,k-1} &= -\frac{\pi(k - 3)}{4k^2(k + 1)^2}, & b_{k,k+1} &= \frac{\pi(3k^4 + 17k^3 + 64k^2 + 132k + 144)}{4k^2(k + 2)^2(k + 3)^2(k + 4)}, \\
 b_{k,k+3} &= -\frac{\pi(3k^4 + 31k^3 + 148k^2 + 332k + 320)}{4(k + 1)(k + 2)^2(k + 4)^2(k + 5)}, & b_{k,k+5} &= \frac{\pi k}{4(k + 3)(k + 4)^3},
 \end{aligned}$$

$$c_{k,k} = c_{k-2,k+2} = -\frac{\pi(k-2)}{2(k+1)^2(k+2)}, \quad c_{k,k+2} = \frac{\pi(k^4 + 8k^3 + 37k^2 + 84k + 72)}{(k+1)^2(k+3)^2(k+4)^2}.$$

COROLLARY 4.5. *If $\alpha = 1/2$, then the nonzero elements of matrices A , B and C in system (4.5) are given by*

$$\begin{aligned} a_{k,k-2} &= -\frac{(k-4)_2}{4(k-2)_2(2k-3/2)_3}, & a_{k,k} &= \frac{8(k^4 + 2k^3 + 5k^2 + 4k + 30)}{(k)_4(2k-3)(2k+1)(2k+5)}, \\ a_{k,k+2} &= -\frac{(2k+3)(3k^4 + 18k^3 + 73k^2 + 138k + 140)}{4((k+2)_2)^2(k-1/2)_2(k+5/2)_2}, \\ a_{k,k+4} &= \frac{8(k^4 + 10k^3 + 41k^2 + 80k + 90)}{(k+2)_4(2k+1)(2k+5)(2k+9)}, \\ a_{k,k+6} &= -\frac{(k)_2}{4(k+2)_2(k+5/2)_3}, & b_{k,k-1} &= -\frac{(k-3)_2}{2(k-1)_2(k-1/2)_2}, \\ b_{k,k+1} &= \frac{6(k^2 + 2k + 4)(k^2 + 2k + 5)}{(k+1)_3(k+2)(2k-1)(2k+5)}, \\ b_{k,k+3} &= -\frac{6(k^4 + 8k^3 + 31k^2 + 60k + 56)}{(k+2)_3(k+3)(2k+1)(2k+7)}, \\ b_{k,k+5} &= \frac{(k)_2}{2(k+2)_2(k+5/2)_2}, & c_{k,k} = c_{k-2,k+2} &= -\frac{2(k-2)_2}{(k)_2(2k+1)}, \\ c_{k,k+2} &= \frac{4(2k+3)(k^4 + 6k^3 + 23k^2 + 42k + 30)}{((k+2)_2)^2(2k+1)(2k+5)}. \end{aligned}$$

4.2. Nonhomogeneous boundary conditions In this case, the integrated form of the general second-order differential equation (4.1) is subject to the boundary conditions

$$\alpha_1 u(1) + \lambda_1 u'(1) = \varepsilon_1, \quad \alpha_2 u(-1) + \lambda_2 u'(-1) = \varepsilon_2. \quad (4.7)$$

In such a case we proceed as follows: Set

$$V(x) = u(x) - Ex - F, \quad (4.8)$$

where $E = (\alpha_1 \varepsilon_2 - \alpha_2 \varepsilon_1)/W$, $F = (\varepsilon_1(\lambda_2 - \alpha_2) - \varepsilon_2(\alpha_1 + \lambda_1))/W$ and $W = \alpha_1 \lambda_2 - \alpha_2 \lambda_1 - 2\alpha_1 \alpha_2 \neq 0$. The transformation (4.8) turns the nonhomogeneous boundary conditions (4.7) into the homogeneous boundary conditions

$$\alpha_1 V(1) + \lambda_1 V'(1) = 0, \quad \alpha_2 V(-1) + \lambda_2 V'(-1) = 0. \quad (4.9)$$

Hence it suffices to solve the following modified equation

$$\beta_1 \iint V(x) dx dx + \beta_2 \int V(x) dx - V(x) = f^*(x), \quad (4.10)$$

subject to the homogeneous mixed boundary conditions (4.9), where $V(x)$ is given by (4.8), and

$$\begin{aligned} f^*(x) &= f(x) + d_0 + d_1x - \beta_1 \iint (Ex + F) dx dx - \beta_2 \int (Ex + F) dx + (Ex + F) \\ &= f(x) - \frac{\beta_1(1 + 2\alpha)E}{12(2 + \alpha)} C_3^{(\alpha)}(x) - \frac{(1 + 2\alpha)(\beta_1 F + \beta_2 E)}{4(1 + \alpha)} C_2^{(\alpha)}(x) \\ &\quad + d_3 C_1^{(\alpha)}(x) + d_4 C_0^{(\alpha)}(x). \end{aligned}$$

If we apply the Galerkin method to the modified equation (4.10), we get the equivalent system of equations

$$(\beta_1 A + \beta_2 B + C) \mathbf{a} = \mathbf{f}^*, \tag{4.11}$$

where A, B and C are the matrices defined in Theorem 4.1 and $\mathbf{f}^* = (f_2^*, f_3^*, \dots, f_N^*)$, where

$$f_k^* = \begin{cases} f_2 - \frac{(1 + 2\alpha)}{4(1 + \alpha)} (\beta_1 F + \beta_2 E) h_2, & k = 2, \\ f_3 - \frac{(1 + 2\alpha)}{12(2 + \alpha)} \beta_1 E h_3, & k = 3, \\ f_k, & k \geq 4. \end{cases}$$

REMARK 4.6. In the case of $W = 0$, the previous treatment fails but we can take the approximate solution in the form $u_N(x) = \phi_0(x) + \sum_{k=1}^{N-2} a_k \phi_k(x)$, where the basis $\{\phi_k(x) : k = 0, 1, \dots, N - 2\}$ is chosen so that $\phi_0(x)$ satisfies the nonhomogeneous boundary conditions and so that every $\phi_k(x)$, $k = 1, 2, \dots, N - 2$, satisfies the homogeneous parts of the boundary conditions.

5. Numerical results

In this section we give some numerical results obtained by using the algorithms presented in the previous sections.

It is worth mentioning that the pure spectral-Galerkin method is rarely used in practice, since for a general right-hand side function f one is unable to exactly compute its representation by ultraspherical polynomials. In fact, the so-called pseudospectral method is used to treat the right-hand side; that is, we replace f by its polynomial interpolation over the set of Gauss-Lobatto points.

We consider the following example.

EXAMPLE 1. Consider the one-dimensional nonhomogeneous Helmholtz equation

$$-u'' + b^2 u = e^{ax}, \quad x \in [-1, 1], \quad u(\pm 1) = 0, \tag{5.1}$$

TABLE 1. Maximum pointwise error of $u - u_N$ for $N = 4, 8, 16, 20$.

N	a	b	α	UG1	UG2	a	b	UG1	UG2
4	1	0	3/2	2.111×10^{-3}	2.363×10^{-3}	2	1	2.445×10^{-2}	2.984×10^{-2}
			1	1.808×10^{-3}	2.374×10^{-3}			2.211×10^{-2}	2.736×10^{-2}
			1/2	1.465×10^{-3}	2.374×10^{-3}			1.725×10^{-2}	2.445×10^{-2}
			0	1.075×10^{-3}	1.810×10^{-2}			1.284×10^{-2}	2.110×10^{-2}
			-0.2	8.960×10^{-4}	1.677×10^{-3}			1.093×10^{-2}	1.962×10^{-2}
			-0.45	8.258×10^{-4}	1.502×10^{-3}			8.0465×10^{-3}	1.764×10^{-2}
8	1	0	3/2	9.405×10^{-8}	1.583×10^{-7}	2	1	1.766×10^{-5}	2.925×10^{-5}
			1	6.599×10^{-8}	1.251×10^{-7}			1.245×10^{-5}	2.338×10^{-5}
			1/2	4.171×10^{-8}	9.405×10^{-8}			7.486×10^{-6}	1.766×10^{-5}
			0	2.252×10^{-8}	6.599×10^{-8}			3.542×10^{-6}	1.245×10^{-5}
			-0.2	1.643×10^{-8}	5.373×10^{-8}			3.160×10^{-6}	5.373×10^{-6}
			-0.45	1.621×10^{-8}	4.223×10^{-8}			2.878×10^{-6}	7.691×10^{-6}
16	1	0	3/2	9.884×10^{-16}	5.414×10^{-16}	2	1	3.018×10^{-14}	1.073×10^{-13}
			1	9.662×10^{-16}	5.414×10^{-16}			2.207×10^{-14}	6.853×10^{-14}
			1/2	9.532×10^{-16}	5.356×10^{-16}			1.044×10^{-14}	4.039×10^{-14}
			0	4.796×10^{-16}	8.058×10^{-16}			4.186×10^{-15}	2.212×10^{-14}
			-0.2	9.521×10^{-16}	6.773×10^{-16}			2.983×10^{-15}	1.715×10^{-14}
			-0.45	7.942×10^{-16}	5.641×10^{-16}			2.503×10^{-15}	1.184×10^{-14}
20	1	0	3/2	9.884×10^{-16}	5.414×10^{-16}	2	1	7.126×10^{-16}	7.617×10^{-16}
			1	9.662×10^{-16}	5.414×10^{-16}			7.308×10^{-16}	7.342×10^{-16}
			1/2	9.532×10^{-16}	5.356×10^{-16}			7.192×10^{-16}	7.867×10^{-16}
			0	4.796×10^{-16}	8.058×10^{-16}			7.257×10^{-16}	7.674×10^{-16}
			-0.2	9.521×10^{-16}	6.773×10^{-16}			6.857×10^{-16}	7.512×10^{-16}
			-0.45	7.942×10^{-16}	5.641×10^{-16}			6.922×10^{-16}	7.033×10^{-16}

with an exact smooth solution

$$u(x) = \frac{e^{ax} + e^{-bx} \sinh(a - b) \operatorname{cosech}(2b) - e^{bx} \sinh(a + b) \operatorname{cosech}(2b)}{b^2 - a^2}.$$

Now the integrated form of (5.1) is

$$b^2 \iint u \, dx \, dx - u = \frac{e^{ax}}{a^2} + d_0 C_0^{(\alpha)}(x) + d_1 C_1^{(\alpha)}(x), \quad x \in [-1, 1], \quad u(\pm 1) = 0.$$

We shall compare the ultraspherical-Galerkin approximations obtained, using the first and second choices of bases (UG1) and (UG2), for some different values of the parameter α .

(i) For UG1, $u_N = \sum_{k=0}^{N-2} a_k (1 - x^2) C_k^{(\alpha)}(x)$, where the vector of unknowns \mathbf{a} is the solution of the system

$$(b^2 A + E)\mathbf{a} = \mathbf{f}^*, \tag{5.2}$$

where the nonzero elements of matrices A and E are those given in Theorem 3.2, and

$$f^* = \frac{2\sqrt{\pi}(k + \alpha)\Gamma(k + 2\alpha)}{a^2(2a)^\alpha k! \Gamma(\alpha + 1/2)} I_{k+\alpha}(a), \quad k = 2, 3, \dots, N,$$

TABLE 2. $\text{Cond}(E)$ for $\alpha = 0, 1/2, 1$.

N	α	α_{\min}	α_{\max}	$\text{Cond}(E)$
4	0	2.500×10^{-1}	5.000×10^{-1}	2
8				2
16				2
32				2
64				2
4	0.5	2.669×10^{-1}	6.666×10^{-1}	1.9444
8				2.32143
16				2.49722
32				2.58266
64				2.62483
4	1	2.661×10^{-1}	7.500×10^{-1}	1.800
8				2.3333
16				2.6471
32				2.81818
64				2.90769

where $I_n(\cdot)$ represents the modified Bessel function of the first kind (see Luke [18, page 32]).

(ii) For UG2, $u_N = \sum_{k=0}^{N-2} a_k (C_k^{(\alpha)}(x) - C_{k+2}^{(\alpha)}(x))$,

$$(b^2 A + B)a = f, \tag{5.3}$$

and the nonzero elements of matrices A and B are those given in Theorem 3.7, and

$$f_k = \frac{2\pi\Gamma(2\alpha)}{a^2(2a)^\alpha\Gamma(\alpha)} [I_{k+\alpha}(a) - I_{k+\alpha+2}(a)], \quad k = 2, 3, \dots, N.$$

Table 1 lists the maximum pointwise error of $u - u_N$ using the UGM with various choices of α, a, b and N .

The main source of roundoff errors comes from the matrix decomposition. For UG1, the system resulting from the integrated form of $-u'' = g(x)$ is $Ea = f^*$, where E is an upper triangular matrix whose diagonal elements are

$$e_{kk} = \begin{cases} 1/2, & k = 2, \alpha = 0, \\ 1/4, & k = 3, 4, \dots, N, \alpha = 0, \\ \frac{(k + 2\alpha - 2)_2}{4(k + \alpha - 2)_2}, & k = 2, 3, \dots, N, \alpha \neq 0. \end{cases}$$

Thus we note that the condition number for E behaves like $O(k^0)$ for every value of α . For UG2, the resulting system is $Ba = f$, where B is also an upper triangular matrix with

$$b_{kk} = \begin{cases} \pi/2, & \alpha = 0, \\ \frac{\Gamma(1/2)\Gamma(\alpha + 1/2)k!}{\Gamma(\alpha)(2\alpha)_k(k + \alpha)}, & \alpha \neq 0, \end{cases}$$

and accordingly, its condition number behaves like $O(k^{2\alpha})$ for large values of k and $\alpha > -1/2$.

Tables 2 and 3 illustrate the condition numbers for matrix E in (5.2) and matrix B in (5.3), respectively.

REMARK 5.1. If we add b^2A to matrix E of system (5.2) and add b^2A to matrix B of system (5.3), then we find that the eigenvalues of matrices $D = E + b^2A$ and $C = B + b^2A$ are all real positive. Moreover, the effect of these additions does not significantly change the values of the condition numbers for the two systems. This means that the matrices E and B , which resulted from the integrated form of the highest derivatives of the differential equations under investigation, play the most important role in the propagation of the roundoff errors. The numerical results of Table 4 illustrates this remark.

TABLE 3. $\text{Cond}(B)$ for $\alpha = 0, 1/2, 1$.

N	α	α_{\min}	α_{\max}	$\text{Cond}(B)$	$\text{Cond}(B)/N$	$\text{Cond}(B)/N^2$
4	0	1.571	1.571	1		
8				1		
16				1		
32				1		
64				1		
4	0.5	2.222×10^{-1}	4.000×10^{-1}	1.800	4.500×10^{-1}	
8		1.176×10^{-1}		3.400	4.250×10^{-1}	
16		6.060×10^{-2}		6.600	4.125×10^{-1}	
32		3.676×10^{-2}		1.300×10^1	4.062×10^{-1}	
64		1.550×10^{-2}		2.580×10^1	4.031×10^{-1}	
4	1	6.281×10^{-2}	1.745×10^{-1}	2.777		1.736×10^{-1}
8		1.939×10^{-2}		9.000	1.406×10^{-1}	
16		5.433×10^{-3}		3.211×10^1	1.254×10^{-1}	
32		1.445×10^{-3}		1.210×10^2	1.181×10^{-1}	
64		3.713×10^{-4}		4.694×10^2	1.146×10^{-1}	

6. Concluding remarks

We have presented a systematic way to construct appropriate bases for the UGM applied to the integrated forms of general second-order differential equations in one dimension and, in particular, the integrated forms for Helmholtz equations in one and two dimensions. We have also developed efficient direct solvers, whose complexities are a small multiple of N^{d+1} operations in a d -dimensional domain, with $(N - 1)^d$ unknowns.

TABLE 4. The condition number for the matrix $C = b^2 A + B$ for $\alpha = 0, 1/2, 1$.

N	α	α_{\min}	α_{\max}	Cond(C)	Cond(C)/ N	Cond(C)/ N^2
4	0	1.512	2.414	1.595		
8		1.332	2.425	1.819		
16		1.322	2.425	1.833		
32		1.322	2.425	1.833		
64		1.322	2.425	1.833		
4	0.5	2.237×10^{-1}	5.589×10^{-1}	2.636	6.592×10^{-1}	
8		1.149×10^{-1}	5.590×10^{-1}	5.131	6.414×10^{-1}	
16		5.982×10^{-2}	5.590×10^{-1}	9.861	6.163×10^{-1}	
32		3.055×10^{-2}	5.590×10^{-1}	1.930×10^1	6.033×10^{-1}	
64		1.544×10^{-2}	5.590×10^{-1}	3.819×10^1	5.967×10^{-1}	
4	1	6.377×10^{-2}	2.559×10^{-1}	4.013		2.508×10^{-1}
8		1.922×10^{-2}		1.331×10^1		2.081×10^{-1}
16		5.404×10^{-3}		4.737×10^1		1.850×10^{-1}
32		1.437×10^{-3}		1.780×10^2		1.738×10^{-1}
64		3.711×10^{-4}		6.897×10^2		1.684×10^{-1}

The two algorithms, namely, the Legendre and Chebyshev Galerkin methods can be obtained as special cases from our UGM by taking $\alpha = 1/2$ and $\alpha = 0$, respectively. Although we concentrated on applying our algorithms to solve constant coefficient differential equations, we do claim that such algorithms can be applied to solve not only differential equations with polynomial varying coefficients of any order but also evolution equations.

References

- [1] G. Ben-Yu, "Gengenbauer approximation and its applications to differential equations on the whole line", *J. Math. Anal. Appl.* **226** (1998) 180–206.
- [2] G. Ben-Yu, *Spectral methods and their applications* (World Scientific, London, 1998).
- [3] B. L. Buzbee, G. H. Golub and C. W. Neilson, "On direct methods for solving Poisson's equations", *SIAM J. Numer. Anal.* **7** (1970) 627–656.
- [4] C. Canuto, M. Y. Hussaini, A. Quarteroni and T. A. Zang, *Spectral methods in fluid dynamics* (Springer, New York, 1989).
- [5] E. H. Doha, "An accurate double Chebyshev spectral approximation for Poisson's equation", *Ann. Univ. Sci. Budapest Sect. Comp.* **10** (1990) 243–276.
- [6] E. H. Doha, "An accurate solution of parabolic equations by expansion in ultraspherical polynomials", *J. Comput. Math. Appl.* **19** (1990) 75–88.
- [7] E. H. Doha, "The coefficients of differentiated expansions and derivatives of ultraspherical polynomials", *J. Comput. Math. Appl.* **21** (1991) 115–122.
- [8] E. H. Doha, "The ultraspherical coefficients of the moments of a general-order derivative of an infinitely differentiable function", *J. Comput. Appl. Math.* **89** (1998) 53–72.

- [9] E. H. Doha, "On the coefficients of integrated expansions and integrals of ultraspherical polynomials and their applications for solving differential equations", *J. Comput. Appl. Math.* **139** (2002) 275–298.
- [10] E. H. Doha and W. M. Abd-Elhameed, "Efficient spectral-Galerkin algorithms for direct solution of second-order equations using ultraspherical polynomials", *SIAM J. Sci. Comput.* **24** (2002) 548–571.
- [11] L. Fox and I. B. Parker, *Chebyshev polynomials in numerical analysis* (Oxford University Press, Oxford, 1972).
- [12] D. Gottlieb and S. A. Orszag, *Numerical analysis of spectral methods: theory and applications* (SIAM, Philadelphia, 1977).
- [13] D. B. Haidvogel and T. Zang, "The accurate solution of Poisson's equation by expansion in Chebyshev polynomials", *J. Comput. Phys.* **30** (1979) 167–180.
- [14] W. Heinrichs, "Improved condition number of spectral methods", *Math. Comp.* **53** (1989) 103–119.
- [15] W. Heinrichs, "Spectral methods with sparse matrices", *Numer. Math.* **56** (1989) 25–41.
- [16] W. Heinrichs, "Algebraic spectral multigrid methods", *Comput. Methods Appl. Mech. Engrg.* **80** (1990) 281–286.
- [17] C. Lanczos, *Applied analysis* (Pitman, London, 1957).
- [18] Y. Luke, *The special functions and their approximations, Vol. 1* (Academic Press, New York, 1969).
- [19] S. A. Orszag, "Spectral methods for problems in complex geometrics", *J. Comput. Phys.* **37** (1980) 70–92.
- [20] T. N. Phillips and A. Karageorghis, "On the coefficients of integrated expansions of ultraspherical polynomials", *SIAM J. Numer. Anal.* **27** (1990) 823–830.
- [21] J. Shen, "Efficient spectral-Galerkin method I. Direct solvers of second- and fourth-order equations using Legendre polynomials", *SIAM J. Sci. Comput.* **15** (1994) 1489–1505.
- [22] J. Shen, "Efficient spectral-Galerkin method II. Direct solvers of second- and fourth-order equations using Chebyshev polynomials", *SIAM J. Sci. Comput.* **16** (1995) 74–87.
- [23] H. I. Siyyam and M. I. Syam, "An accurate solution of Poisson equation by the Chebyshev-Tau method", *J. Comput. Appl. Math.* **85** (1997) 1–10.