A Spectral Method for Elliptic Equations: The Dirichlet Problem

Kendall Atkinson
Departments of Mathematics & Computer Science
The University of Iowa

David Chien, Olaf Hansen
Department of Mathematics
California State University San Marcos

February 28, 2009

Abstract

Let Ω be an open, simply connected, and bounded region in $\mathbb{R}^d$, $d \geq 2$, and assume its boundary $\partial\Omega$ is smooth. Consider solving an elliptic partial differential equation $Lu = f$ over Ω with zero Dirichlet boundary values. The problem is converted to an equivalent elliptic problem over the unit ball B; and then a spectral Galerkin method is used to create a convergent sequence of multivariate polynomials u_n of degree $\leq n$ that is convergent to u. The transformation from Ω to B requires a special analytical calculation for its implementation. With sufficiently smooth problem parameters, the method is shown to be rapidly convergent. For $u \in C^\infty\left(\overline{\Omega}\right)$ and assuming $\partial\Omega$ is a C^∞ boundary, the convergence of $\|u - u_n\|_{H^1}$ to zero is faster than any power of $1/n$. Numerical examples in $\mathbb{R}^2$ and $\mathbb{R}^3$ show experimentally an exponential rate of convergence.

1 INTRODUCTION

Consider solving the elliptic partial differential equation

$$Lu(\mathbf{s}) \equiv -\sum_{i,j=1}^{d} \frac{\partial}{\partial s_i}\left(a_{i,j}(\mathbf{s})\frac{\partial u(\mathbf{s})}{\partial s_j}\right) + \gamma(\mathbf{s})u(\mathbf{s}) = f(\mathbf{s}), \qquad \mathbf{s} \in \Omega \subseteq \mathbb{R}^d \tag{1}$$

with the Dirichlet boundary condition

$$u(\mathbf{s}) \equiv 0, \qquad \mathbf{s} \in \partial\Omega \tag{2}$$

Assume $d \geq 2$. Let Ω be an open, simply-connected, and bounded region in $\mathbb{R}^d$, and assume that its boundary $\partial\Omega$ is smooth and sufficiently differentiable.

Similarly, assume the functions $\gamma(\mathbf{s})$, $f(\mathbf{s})$, $a_{i,j}(\mathbf{s})$ are several times continuously differentiable over $\overline{\Omega}$. As usual, assume the matrix $A(\mathbf{s}) = [a_{i,j}(\mathbf{s})]$ is symmetric and satisfies the strong ellipticity condition,

$$\xi^{\mathrm{T}} A(\mathbf{s})\xi \geq c_0 \xi^{\mathrm{T}} \xi, \qquad \mathbf{s} \in \overline{\Omega}, \quad \xi \in \mathbb{R}^d \tag{3}$$

with $c_0 > 0$. Also assume $\gamma(\mathbf{s}) \geq 0$, $\mathbf{s} \in \Omega$.

In §2 we consider the special region $\Omega = B$, the open unit ball in $\mathbb{R}^d$. We define a Galerkin method for (1)-(2) with a special finite-dimensional subspace of polynomials, and we give an error analysis that shows rapid convergence of the method. In §3 we discuss the use of a transformation from a general region Ω to the unit ball B, showing that the transformed equation is again elliptic over B. Implementation issues are discussed in §4 for problems in $\mathbb{R}^2$ and $\mathbb{R}^3$. We conclude in §5 with numerical examples in $\mathbb{R}^2$ and $\mathbb{R}^3$.

The methods of this paper generalize to the equation

$$\begin{aligned} Lu(\mathbf{s}) \equiv &- \sum_{i,j=1}^{d} \frac{\partial}{\partial s_i} \left(a_{i,j}(\mathbf{s}) \frac{\partial u(\mathbf{s})}{\partial s_j} \right) \\ &+ \sum_{j=1}^{d} b_j\left(\mathbf{s}\right) \frac{\partial u(\mathbf{s})}{\partial s_j} + \gamma(\mathbf{s})u(\mathbf{s}) = f(\mathbf{s}), \qquad \mathbf{s} \in \Omega \subseteq \mathbb{R}^d \end{aligned}$$

which contains first order derivative terms, provided the operator L is strongly elliptic. To do so, use the results given in Brenner and Scott [6, §§2.6-2.8], combined with the methods of the present paper. We have chosen to restrict our work to the more standard symmetric problem (1).

There is a rich literature on spectral methods for solving partial differential equations. From the more recent literature, we cite [5], [7], [8], [9] and [19]. Their bibliographies contain references to earlier papers on spectral methods. Our approach is somewhat different than the standard approaches, as we are converting the partial differential equation to an equivalent problem on the unit disk or unit ball, and in the process we are required to work with a more complicated equation. Our approach is reminiscent of the use of conformal mappings for planar problems. Conformal mappings can be used with our approach when working on planar problems, although having a conformal mapping is not necessary. Our approach is also different in that we use multivariable polynomial approximations rather than the more standard use of single variable approximations in each of the spatial variables for the problem.

2 A spectral method on the unit ball

The Dirichlet problem (1)-(2) has the following variational reformulation: Find $u \in H_0^1(\Omega)$ such that

$$\int_\Omega \left[\sum_{i,j=1}^{d} a_{i,j}(\mathbf{s}) \frac{\partial u(\mathbf{s})}{\partial s_j} \frac{\partial v(\mathbf{s})}{\partial s_i} + \gamma(\mathbf{s}) u(\mathbf{s}) v(\mathbf{s}) \right] d\mathbf{s} \\ = \int_\Omega f(\mathbf{s}) v(\mathbf{s})\, d\mathbf{s}, \qquad \forall v \in H_0^1(\Omega) \tag{4}$$

We define a spectral Galerkin method in this section for the special region $\Omega = B$. In §3 we discuss the transformation of (1) from a general $\overline{\Omega}$ to an equivalent equation over the unit ball $\overline{B}$, a transformation that retains the ellipticity of the problem. In the remainder of this section, we replace Ω with B.

Introduce the bilinear form

$$\mathcal{A}(v, w) = \int_B \left[\sum_{i,j=1}^{d} a_{i,j}(\mathbf{x}) \frac{\partial v(\mathbf{x})}{\partial x_j} \frac{\partial w(\mathbf{x})}{\partial x_i} + \gamma(\mathbf{x}) v(\mathbf{x}) w(\mathbf{x}) \right] d\mathbf{x}, \quad v, w \in H_0^1(B) \tag{5}$$

and the bounded linear functional

$$\ell(v) = \int_B f(\mathbf{x})\, v(\mathbf{x})\, d\mathbf{x}, \qquad v \in H_0^1(B)$$

The variational problem (4) can now be written as follows: find $u \in H_0^1(B)$ for which

$$\mathcal{A}(u, v) = \ell(v), \qquad \forall v \in H_0^1(B) \tag{6}$$

It is straightforward to show $\mathcal{A}$ is bounded,

$$|\mathcal{A}(v, w)| \le c_{\mathcal{A}} \|v\|_1 \|w\|_1$$
$$c_{\mathcal{A}} = \max_{\mathbf{x} \in \overline{B}} \|A(\mathbf{x})\|_2 + \|\gamma\|_\infty$$

where $\|\cdot\|_1$ denotes the norm of $H_0^1(\Omega)$ and $\|A(\mathbf{x})\|_2$ the matrix 2-norm of the matrix $A(\mathbf{x})$. In addition, we assume

$$\mathcal{A}(v, v) \ge c_e \|v\|_1^2, \qquad v \in H_0^1(B) \tag{7}$$

This follows generally from (3) and the size of the function $\gamma(\mathbf{x})$ over $\overline{B}$; when $\gamma \equiv 0$, $c_e = c_0$. Under standard assumptions on $\mathcal{A}$, including the strong ellipticity in (7), the Lax-Milgram Theorem implies the existence of a unique solution u to (6) with

$$\|u\|_1 \le \frac{1}{c_e} \|\ell\|$$

with $\|\ell\|$ denoting the operator norm for ℓ regarded as a linear functional on $H_0^1(\Omega)$.

Denote by Π_n the space of polynomials in d variables that are of degree $\leq n$: $p \in \Pi_n$ if it has the form

$$p(\mathbf{x}) = \sum_{|i| \leq n} a_i x_1^{i_1} x_2^{i_2} \ldots x_d^{i_d}$$

with i a multi-integer, $i = (i_1, \ldots, i_d)$, and $|i| = i_1 + \cdots + i_d$. Let $\mathcal{X}_n$ denote our approximation subspace,

$$\mathcal{X}_n = \left\{ \left(1 - \|\mathbf{x}\|_2^2\right) p(\mathbf{x}) \mid p \in \Pi_n \right\} \tag{8}$$

with $\|\mathbf{x}\|_2^2 = x_1^2 + \cdots + x_d^2$. The subspaces Π_n and $\mathcal{X}_n$ have dimension

$$N_n = \binom{n+d}{d}$$

Lemma 1 *Let Δ denote the Laplacian operator in $\mathbb{R}^d$. Then*

$$\Delta : \mathcal{X}_n \xrightarrow[onto]{1-1} \Pi_n \tag{9}$$

For a short proof, see [4].

The Galerkin method for obtaining an approximate solution to (6) is as follows: find $u_n \in \mathcal{X}_n$ for which

$$\mathcal{A}(u_n, v) = \ell(v), \qquad \forall v \in \mathcal{X}_n \tag{10}$$

The Lax-Milgram Theorem (cf. [3, §8.3], [6, §2.7]) implies the existence of u_n for all n. For the error in this Galerkin method, Cea's Lemma (cf. [3, p. 365], [6, p. 62]) implies the convergence of u_n to u, and moreover,

$$\|u - u_n\|_1 \leq \frac{c_{\mathcal{A}}}{c_e} \inf_{v \in \mathcal{X}_n} \|u - v\|_1 \tag{11}$$

It remains to bound the best approximation error on the right side of this inequality. In order to prove the right side of (11) converges to zero, we begin by first considering the case in which $u \in H_0^2(B)$; and later we extend this to $u \in H_0^1(B)$.

Assume $u \in H_0^2(B)$, and define $w = -\Delta u$. Then $w \in L^2(B)$ and u satisfies the boundary value problem

$$\begin{aligned} -\Delta u(P) &= w(P), \qquad P \in B \\ u(P) &= 0, \qquad P \in \partial B \end{aligned}$$

It follows that

$$u(P) = \int_B G(P, Q) w(Q) \, dQ, \qquad P \in \overline{B} \tag{12}$$

For $\mathbb{R}^2$ and $\mathbb{R}^3$, the Green's function is defined as follows.

$$\begin{aligned} d=2: \quad G(P,Q) &= \frac{1}{2\pi}\log\frac{|P-Q|}{|\mathcal{T}(P)-Q|}, \\ d=3: \quad G(P,Q) &= -\frac{1}{4\pi}\left\{\frac{1}{|P-Q|}-\frac{1}{|P|}\frac{1}{|\mathcal{T}(P)-Q|}\right\} \end{aligned} \tag{13}$$

for $P \neq Q$, $Q \in B$, $P \in \overline{B}$. $\mathcal{T}(P)$ denotes the inverse point for P with respect to the unit sphere $S^{d-1} \subseteq \mathbb{R}^d$,

$$\mathcal{T}(r\mathbf{x}) = \frac{1}{r}\mathbf{x}, \qquad 0 < r \leq 1, \quad \mathbf{x} \in S^{d-1}$$

Differentiate (12) to obtain

$$\nabla u\,(P) = \int_B [\nabla_P G(P,Q)]\, w(Q)\, dQ, \qquad P \in \overline{B} \tag{14}$$

Note that $\nabla_P G(P,\cdot)$ is absolutely integrable over $\overline{B}$, for all $P \in \overline{B}$.

Let $w_n \in \Pi_n$ be an approximation of w in $L^2\,(B)$, and let

$$q_n(P) = \int_B G(P,Q) w_n(Q)\, dQ, \qquad P \in \overline{B}$$

We can show $q_n \in \mathcal{X}_n$. This follows from Lemma 1 and noting that the mapping in (12) is the inverse of (9).

Then we have

$$u(P) - q_n(P) = \int_B G(P,Q)\,[w(P) - w_n(Q)]\; dQ, \qquad P \in \overline{B}$$

$$\nabla\,[u\,(P) - q_n(P)] = \int_B [\nabla_P G(P,Q)]\,[w(Q) - w_n(Q)]\; dQ, \qquad P \in \overline{B}$$

The integral operators on the right side are weakly singular compact integral operators on $L^2\,(B)$ to $L^2\,(B)$ [17, Chap. 7, §3]. This implies

$$\|u - q_n\|_1 \leq c\|w - w_n\|_0 \tag{15}$$

By letting w_n be the orthogonal projection of w into Π_n, the right side will go to zero since the polynomials are dense in $L^2\,(B)$. In turn, this implies convergence in the $H_0^1\,(B)$ norm for the right side in (11) provided $u \in H_0^2\,(B)$.

The result

$$\inf_{v\in\mathcal{X}_n} \|u - v\|_1 \to 0 \quad \text{as} \quad n \to \infty, \qquad u \in H_0^2\,(B)$$

can be extended to any $u \in H_0^1\,(B)$. It basically follows from the denseness of $H_0^2\,(B)$ in $H_0^1\,(B)$. Let $u \in H_0^1\,(B)$. We need to find a sequence of polynomials $\{q_n\}$ for which $\|u - q_n\|_1 \to 0$. We know $H_0^2\,(B)$ is dense in $H_0^1\,(B)$. Given any

$k > 0$, choose $u_k \in H_0^2(B)$ with $\|u - u_k\|_1 \leq 1/k$. Then choose a polynomial w_k for which we have the corresponding polynomial q_k satisfying $\|u_k - q_k\|_1 \leq 1/k$, based on (15). [Regarding the earlier notation, q_k need not be of degree $\leq k$.] Then $\|u - q_k\|_1 \leq 2/k$.

To obtain orders of convergence, use (15) and results on best multivariate polynomial approximation over the unit disk. For example, use results of Ragozin [18, Thm 3.4] or Yuan Xu [22]. From [18] we have the following theorem. When applied to our problem with problem parameters that are C^∞, it shows that our numerical method has a convergence rate better than any power of $1/n$. This is sometimes called *spectral convergence*; see [8, p. 10].

Theorem 2 *Assume $u \in C^{k+2}\left(\overline{B}\right)$ for some $k > 0$, and assume $u|_{\partial B} = 0$. Then there is a polynomial $q_n \in \mathcal{X}_n$ for which*

$$\|u - q_n\|_\infty \leq D(k,d)\, n^{-k}\left(n^{-1}\|u\|_{\infty,k+2} + \omega\left(u^{(k+2)}, 1/n\right)\right) \tag{16}$$

In this,

$$\|u\|_{\infty,k+2} = \sum_{|i|\leq k+2} \left\|\partial^i u\right\|_\infty$$

$$\omega(g,\delta) = \sup_{|\mathbf{x}-\mathbf{y}|\leq\delta} |g(\mathbf{x}) - g(\mathbf{y})|$$

$$\omega\left(u^{(k+2)}, \delta\right) = \sum_{|i|=k+2} \omega\left(\partial^i u, \delta\right)$$

This analysis of this section does not include all uniquely solvable Dirichlet problems (1)-(2) for elliptic differential operators, because the differential operator may not satisfy the strong ellipticity condition (7). For example, the Helmholtz equation

$$-\Delta u + \lambda u = f$$

with $\lambda < 0$ need not satisfy (7) even though the Dirichlet problem may be uniquely solvable. Nonetheless, the numerical method (10) still works empirically, as we illustrate later in §5. We are exploring other methods for the error analysis of (10), ones not dependent on the assumption (7).

3 Transformation of the elliptic equation

Consider the differential operator

$$\mathcal{M}v(\mathbf{s}) = -\sum_{i,j=1}^{d} \frac{\partial}{\partial s_i}\left(a_{i,j}(\mathbf{s})\frac{\partial v(\mathbf{s})}{\partial s_j}\right), \qquad \mathbf{s} \in \Omega \subseteq \mathbb{R}^d, \quad v \in C^2\left(\overline{\Omega}\right) \tag{17}$$

which satisfies the ellipticity condition (3) with $c_0 > 0$. The operator $\mathcal{M}$ is said to be elliptic on $H^2(\Omega)$. We want to transform the operator $\mathcal{M}$ to one acting on functions $\widetilde{u} \in C^2(\overline{B})$ with B the unit ball in $\mathbb{R}^d$.

Assume the existence of a continuously differentiable mapping

$$\Phi : \overline{B} \underset{onto}{\xrightarrow{1-1}} \overline{\Omega} \tag{18}$$

and let $\Psi = \Phi^{-1} : \overline{\Omega} \underset{onto}{\xrightarrow{1-1}} \overline{B}$. Let

$$J(\mathbf{x}) \equiv (D\Phi)(\mathbf{x}) = \left[\frac{\partial \varphi_i(\mathbf{x})}{\partial x_j}\right]_{i,j=1}^{d}, \qquad \mathbf{x} \in \overline{B} \subseteq \mathbb{R}^d$$

denote the Jacobian of the transformation. As usual we assume $J(\mathbf{x})$ is nonsingular on $\overline{B}$, and furthermore

$$\min_{\mathbf{x}\in\overline{B}} |\det J(\mathbf{x})| > 0 \tag{19}$$

Similarly, let $K(\mathbf{s}) \equiv (D\Psi)(\mathbf{s})$ denote the Jacobian of Ψ over $\overline{\Omega}$. By differentiating the components of the equation

$$\Psi(\Phi(\mathbf{x})) = \mathbf{x}$$

we obtain

$$K(\Phi(\mathbf{x})) = J^{-1}(\mathbf{x}), \qquad \mathbf{x} \in \overline{B}$$

This general approach is reminiscent of the coordinate transformations in [14, Chap. 2] in which the mapping function is used in generating a mesh on a region Ω.

For $v \in C(\overline{\Omega})$, let

$$\widetilde{v}(\mathbf{x}) = v(\Phi(\mathbf{x})), \qquad \mathbf{x} \in \overline{B} \subseteq \mathbb{R}^d \tag{20}$$

and conversely,

$$v(\mathbf{s}) = \widetilde{v}(\Psi(\mathbf{s})), \qquad \mathbf{s} \in \overline{\Omega} \subseteq \mathbb{R}^d \tag{21}$$

If $\Phi \in C^k(\overline{B})$ and $v \in C^m(\overline{\Omega})$, then $\widetilde{v} \in C^q(\overline{B})$ with $q = \min\{k, m\}$. Thus assumptions about the differentiability of $\widetilde{v}(\mathbf{x})$ can be related back to assumptions on the differentiability of $v(\mathbf{s})$ and $\Phi(\mathbf{x})$. A converse statement can be made as regards $\widetilde{v}$, v, and Ψ in (21).

Let $\Phi = [\varphi_1, \ldots, \varphi_d]^{\mathrm{T}}$, and assume $v \in C^1(\overline{\Omega})$. Then

$$\begin{aligned}
\frac{\partial \widetilde{v}}{\partial x_i} &= \frac{\partial v}{\partial s_1}\frac{\partial \varphi_1(\mathbf{x})}{\partial x_i} + \cdots + \frac{\partial v}{\partial s_d}\frac{\partial \varphi_d(\mathbf{x})}{\partial x_i} \\
&= \left[\frac{\partial \varphi_1(\mathbf{x})}{\partial x_i}, \cdots, \frac{\partial \varphi_d(\mathbf{x})}{\partial x_i}\right] \nabla_{\mathbf{s}} v
\end{aligned}$$

with the gradient $\nabla_{\mathbf{s}} v$ a column vector evaluated at $\mathbf{s} = \Phi(\mathbf{x})$. More concisely,

$$\nabla_{\mathbf{x}} \widetilde{v}(\mathbf{x}) = J(\mathbf{x})^{\mathrm{T}} \nabla_{\mathbf{s}} v(\mathbf{s}), \qquad \mathbf{s} = \Phi(\mathbf{x}) \tag{22}$$

Similarly,

$$\nabla_{\mathbf{s}} v(\mathbf{s}) = K(\mathbf{s})^{\mathrm{T}} \nabla_{\mathbf{x}} \widetilde{v}(\mathbf{x}), \qquad \mathbf{x} = \Psi(\mathbf{s}) \tag{23}$$

Theorem 3 *Assume the transformation $\Phi \in C^2\left(\overline{B}\right)$ and that it satisfies (18) and (19). Assume the functions $a_{i,j} \in C^1\left(\overline{\Omega}\right)$ Then for $\mathbf{s} = \Phi(\mathbf{x})$,*

$$(\mathcal{M}v)(\mathbf{s}) = -\frac{1}{\det(J(\mathbf{x}))}\sum_{i,j=1}^{d}\frac{\partial}{\partial x_i}\left(\det(J(\mathbf{x}))\,\widetilde{a}_{i,j}(\mathbf{x})\frac{\partial \widetilde{v}(\mathbf{x})}{\partial x_j}\right) \tag{24}$$

$$\begin{aligned}\widetilde{A}(\mathbf{x}) &= K(\Phi(\mathbf{x}))\,A(\Phi(\mathbf{x}))K(\Phi(\mathbf{x}))^T \\ &\equiv \left[\widetilde{a}_{i,j}(\mathbf{x})\right]_{i,j=1}^{d}\end{aligned} \tag{25}$$

Proof. *Let $w \in C_0^\infty\left(\overline{\Omega}\right)$. Then*

$$\int_\Omega (\mathcal{M}v)(\mathbf{s})w(\mathbf{s})\,d\mathbf{s} = \int_B (\mathcal{M}v)(\Phi(\mathbf{x}))w(\Phi(\mathbf{x}))\det(J(\mathbf{x}))\,d\mathbf{x} \tag{26}$$

On the other hand, using integration by parts we have

$$\begin{aligned}\int_\Omega (\mathcal{M}v)(\mathbf{s})w(\mathbf{s})\,d\mathbf{s} &= \int_\Omega \sum_{i,j=1}^{d} a_{i,j}(\mathbf{s})\frac{\partial v(\mathbf{s})}{\partial s_j}\frac{\partial w(\mathbf{s})}{\partial s_i}\,d\mathbf{s} \\ &= \int_B \sum_{i,j=1}^{d} a_{i,j}(\Phi(\mathbf{x}))\frac{\partial v(\Phi(\mathbf{x}))}{\partial s_j}\frac{\partial w(\Phi(\mathbf{x}))}{\partial s_i}\det(J(\mathbf{x}))\,d\mathbf{x}\end{aligned} \tag{27}$$

Using (23),

$$\begin{aligned}\sum_{i,j=1}^{d} a_{i,j}(\Phi(\mathbf{x}))\frac{\partial v(\Phi(\mathbf{x}))}{\partial s_j}\frac{\partial w(\Phi(\mathbf{x}))}{\partial s_i} &= \left[\nabla_{\mathbf{s}} w(\Phi(\mathbf{x}))\right]^T A(\Phi(\mathbf{x}))\left[\nabla_{\mathbf{s}} v(\Phi(\mathbf{x}))\right] \\ &= \left[\nabla_{\mathbf{x}}\widetilde{w}(\mathbf{x})\right]^T K(\Phi(\mathbf{x}))A(\Phi(\mathbf{x}))\,K(\Phi(\mathbf{x}))^T\left[\nabla_{\mathbf{x}}\widetilde{v}(\mathbf{x})\right] \\ &= \left[\nabla_{\mathbf{x}}\widetilde{w}(\mathbf{x})\right]^T \widetilde{A}(\mathbf{x})\left[\nabla_{\mathbf{x}}\widetilde{v}(\mathbf{x})\right]\end{aligned}$$

Using this to continue (27),

$$\begin{aligned}\int_\Omega (\mathcal{M}v)(\mathbf{s})w(\mathbf{s})\,d\mathbf{s} &= \int_B \left[\nabla_{\mathbf{x}}\widetilde{w}(\mathbf{x})\right]^T \widetilde{A}(\mathbf{x})\left[\nabla_{\mathbf{x}}\widetilde{v}(\mathbf{x})\right]\det(J(\mathbf{x}))\,d\mathbf{x} \\ &= \int_\Omega \sum_{i,j=1}^{d} \widetilde{a}_{i,j}(\mathbf{s})\frac{\partial \widetilde{v}(\mathbf{x})}{\partial x_j}\frac{\partial \widetilde{w}(\mathbf{x})}{\partial x_i}\det(J(\mathbf{x}))\,d\mathbf{x} \\ &= -\int_\Omega \sum_{i,j=1}^{d}\frac{\partial}{\partial x_i}\left(\det(J(\mathbf{x}))\,\widetilde{a}_{i,j}(\mathbf{x})\frac{\partial \widetilde{v}(\mathbf{x})}{\partial x_j}\right)\widetilde{w}(\mathbf{x})\,d\mathbf{x}\end{aligned} \tag{28}$$

Comparing (26) and (28), and noting that $w \in C_0^\infty\left(\overline{\Omega}\right)$ is arbitrary, we have

$$(\mathcal{M}v)(\Phi(\mathbf{x}))\det(J(\mathbf{x})) = -\sum_{i,j=1}^{d}\frac{\partial}{\partial x_i}\left(\det(J(\mathbf{x}))\,\widetilde{a}_{i,j}(\mathbf{x})\frac{\partial \widetilde{v}(\mathbf{x})}{\partial x_j}\right)$$

which proves (24). ∎

With this transformation, we can solve the Dirichlet problem over a general region Ω by transforming it to an equivalent problem over the unit ball B. We can apply the Galerkin method to (1) by means of the transformation (24). We convert (1) to the equation

$$\begin{aligned}-\sum_{i,j=1}^{d}\frac{\partial}{\partial x_i}\left(\det\left(J(\mathbf{x})\right)\widetilde{a}_{i,j}(\mathbf{x})\frac{\partial\widetilde{v}(\mathbf{x})}{\partial x_j}\right)+\det\left(J(\mathbf{x})\right)\gamma(\Phi\left(\mathbf{x}\right))\widetilde{v}(\mathbf{x})\\ =\det\left(J(\mathbf{x})\right)f\left(\Phi\left(\mathbf{x}\right)\right)\end{aligned}\tag{29}$$

This system is also strongly elliptic.

Theorem 4 *Assume $A(\mathbf{s})$, $\mathbf{s}\in\overline{\Omega}$, satisfies (3); and without loss of generality, assume*

$$\det J(\mathbf{x})>0,\qquad \mathbf{x}\in\overline{B}$$

Recall $\widetilde{A}\left(\mathbf{x}\right)$ as defined by (25). Then $\widetilde{A}\left(\mathbf{x}\right)$ satisfies the strong ellipticity condition

$$\begin{gathered}\xi^T\widetilde{A}(\mathbf{x})\xi\geq\widetilde{c}_0\xi^T\xi,\qquad \mathbf{x}\in\overline{B},\quad \xi\in\mathbb{R}^d\\ \widetilde{c}_0=c_0\lambda_*\equiv c_0\min_{\mathbf{x}\in\overline{B}}\lambda_{\min}(\mathbf{x})\end{gathered}$$

with $\lambda_{\min}(\mathbf{x})$ the smallest eigenvalue of $K(\Phi\left(\mathbf{x}\right))^TK(\Phi\left(\mathbf{x}\right))$ (which equals the reciprocal of the largest eigenvalue of $J\left(\mathbf{x}\right)^T J\left(\mathbf{x}\right)$).
Proof.

$$\begin{aligned}\xi^T\widetilde{A}(\mathbf{x})\xi&=\xi^TKAK^T\xi=\left(K^T\xi\right)^TA\left(K^T\xi\right)\\ &\geq c_0\left(K^T\xi\right)^T\left(K^T\xi\right)=c_0\left\|K^T\xi\right\|_2^2\end{aligned}$$

In addition,

$$\begin{gathered}\left\|K(\Phi\left(\mathbf{x}\right))^T\xi\right\|_2^2\geq\lambda_{\min}(\mathbf{x})\left\|\xi\right\|_2^2\geq\lambda_*\left\|\xi\right\|_2^2\\ \lambda_*=\min_{\mathbf{x}\in\overline{B}}\lambda_{\min}(\mathbf{x})\end{gathered}$$

with $\lambda_{\min}\left(\mathbf{x}\right)$ the smallest eigenvalue of $K(\Phi\left(\mathbf{x}\right))^TK(\Phi\left(\mathbf{x}\right))$; see [2, p. 488].

■

4 Implementation

Consider the implementation of the Galerkin method of §2 for the elliptic problem (6) over the unit ball B. We are to find the function $u_n\in\mathcal{X}_n$ satisfying (10). To do so, we begin by selecting an orthonormal basis for Π_n, denoting it by $\{\varphi_1,\ldots,\varphi_N\}$, with $N\equiv N_n=\dim\Pi_n$. Choosing an orthonormal basis is an attempt to have the linear system in (10) better conditioned. Next, let

$$\psi_i(\mathbf{x})=\left(1-\left\|\mathbf{x}\right\|_2^2\right)\varphi_i(\mathbf{x}),\qquad i=1,\ldots,N_n\tag{30}$$

to form a basis for $\mathcal{X}_n$.

We seek

$$u_n(\mathbf{x}) = \sum_{j=1}^{N} \alpha_j \psi_j(\mathbf{x}) \tag{31}$$

Then (10) becomes

$$\begin{aligned} &\sum_{k=1}^{N_n} \alpha_k \int_B \left[\sum_{i,j=1}^{d} a_{i,j}(\mathbf{x}) \frac{\partial \psi_k(\mathbf{x})}{\partial x_j} \frac{\partial \psi_\ell(\mathbf{x})}{\partial x_i} + \gamma(\mathbf{x}) \psi_k(\mathbf{x}) \psi_\ell(\mathbf{x}) \right] d\mathbf{x} \\ &\qquad = \int_B f(\mathbf{x})\, \psi_\ell(\mathbf{x})\, d\mathbf{x}, \qquad \ell = 1, \ldots, N \end{aligned} \tag{32}$$

We need to calculate the orthonormal polynomials and their first partial derivatives; and we also need to approximate the integrals in the linear system. For an introduction to the topic of multivariate orthogonal polynomials, see Dunkl and Xu [10] and Xu [21]. For multivariate quadrature over the unit ball in $\mathbb{R}^d$, see Stroud [20].

4.1 The planar case

The dimension of Π_n is

$$N_n = \frac{1}{2}(n+1)(n+2) \tag{33}$$

For notation, we replace $\mathbf{x}$ with (x, y). How do we choose the orthonormal basis $\{\varphi_\ell(x, y)\}_{\ell=1}^{N}$ for Π_n? Unlike the situation for the single variable case, there are many possible orthonormal bases over $B = D$, the unit disk in $\mathbb{R}^2$. We have chosen one that is particularly convenient for our computations. These are the "ridge polynomials" introduced by Logan and Shepp [15] for solving an image reconstruction problem. We summarize here the results needed for our work.

Let

$$\mathcal{V}_n = \{P \in \Pi_n : (P, Q) = 0 \quad \forall Q \in \Pi_{n-1}\}$$

the polynomials of degree n that are orthogonal to all elements of Π_{n-1}. Then the dimension of $\mathcal{V}_n$ is $n+1$; moreover,

$$\Pi_n = \mathcal{V}_0 \oplus \mathcal{V}_1 \oplus \cdots \oplus \mathcal{V}_n \tag{34}$$

It is standard to construct orthonormal bases of each $\mathcal{V}_n$ and to then combine them to form an orthonormal basis of Π_n using the latter decomposition. As an orthonormal basis of $\mathcal{V}_n$ we use

$$\varphi_{n,k}(x, y) = \frac{1}{\sqrt{\pi}} U_n(x \cos(kh) + y \sin(kh)), \quad (x, y) \in D, \quad h = \frac{\pi}{n+1} \tag{35}$$

for $k = 0, 1, \ldots, n$. The function U_n is the Chebyshev polynomial of the second kind of degree n:

$$U_n(t) = \frac{\sin(n+1)\theta}{\sin\theta}, \qquad t = \cos\theta, \quad -1 \le t \le 1, \quad n = 0, 1, \ldots \tag{36}$$

The family $\{\varphi_{n,k}\}_{k=0}^{n}$ is an orthonormal basis of $\mathcal{V}_n$. As a basis of Π_n, we order $\{\varphi_{n,k}\}$ lexicographically based on the ordering in (35) and (34):

$$\{\varphi_\ell\}_{\ell=1}^{N} = \{\varphi_{0,0},\, \varphi_{1,0},\, \varphi_{1,1},\, \varphi_{2,0},\, \ldots,\, \varphi_{n,0},\, \ldots, \varphi_{n,n}\}$$

Returning to (30), we define

$$\psi_{n,k}(x,y) = \left(1 - x^2 - y^2\right) \varphi_{n,k}(x,y) \tag{37}$$

To calculate the first order partial derivatives of $\psi_{n,k}(x,y)$, we need $U_n^{'}(t)$. The values of $U_n(t)$ and $U_n^{'}(t)$ are evaluated using the standard triple recursion relations

$$\begin{aligned} U_{n+1}(t) &= 2tU_n(t) - U_{n-1}(t) \\ U_{n+1}^{'}(t) &= 2U_n(t) + 2tU_n^{'}(t) - U_{n-1}^{'}(t) \end{aligned}$$

For the numerical approximation of the integrals in (32), which are over B being the unit disk, we use the formula

$$\int_B g(x,y)\, dx\, dy \approx \sum_{l=0}^{q} \sum_{m=0}^{2q} g\left(r_l, \frac{2\pi\, m}{2q+1}\right) \omega_l \frac{2\pi}{2q+1} r_l \tag{38}$$

Here the numbers ω_l are the weights of the $(q+1)$-point Gauss-Legendre quadrature formula on $[0,1]$. Note that

$$\int_0^1 p(x)dx = \sum_{l=0}^{q} p(r_l)\omega_l,$$

for all single-variable polynomials $p(x)$ with $\deg(p) \le 2q+1$. The formula (38) uses the trapezoidal rule with $2q+1$ subdivisions for the integration over $\overline{B}$ in the azimuthal variable. This quadrature is exact for all polynomials $g \in \Pi_{2q}$. This formula is also the basis of the hyperinterpolation formula discussed in [12].

4.2 The three-dimensional case

In $\mathbb{R}^3$, the dimension of Π_n is

$$N_n = \binom{n+3}{3} = \frac{1}{6}\,(n+1)\,(n+2)\,(n+3)$$

Here we choose orthonormal polynomials on the unit ball as described in [10],

$$\begin{aligned} \varphi_{m,j,\beta}(\mathbf{x}) &= c_{m,j} p_j^{(0,m-2j+\frac{1}{2})}(2\|\mathbf{x}\|^2 - 1) S_{\beta,m-2j}\,(\mathbf{x}) \\ &= c_{m,j} \|\mathbf{x}\|^{m-2j} p_j^{(0,m-2j+\frac{1}{2})}(2\|\mathbf{x}\|^2 - 1) S_{\beta,m-2j}\left(\frac{\mathbf{x}}{\|\mathbf{x}\|}\right), \qquad (39) \\ j &= 0,\ldots,\lfloor m/2 \rfloor, \quad \beta = 0,1,\ldots,2(m-2j), \quad m = 0,1,\ldots,n \end{aligned}$$

Here $c_{m,j} = 2^{\frac{5}{4}+\frac{m}{2}-j}$ is a constant, and $p_j^{(0,m-2j+\frac{1}{2})}$, $j \in \mathbb{N}_0$, are the normalized Jabobi polynomials which are orthonormal on $[-1,1]$ with respect to the inner product

$$(v,w) = \int_{-1}^{1} (1+t)^{m-2j+\frac{1}{2}} v(t) w(t)\, dt,$$

see for example [1], [11]. The functions $S_{\beta,m-2j}$ are spherical harmonic functions, and they are given in spherical coordinates by

$$S_{\beta,k}(\phi,\theta) = \widetilde{c}_{\beta,k} \begin{cases} \cos(\frac{\beta}{2}\phi) T_k^{\frac{\beta}{2}}(\cos\theta), & \beta \text{ even} \\ \sin(\frac{\beta+1}{2}\phi) T_k^{\frac{\beta+1}{2}}(\cos\theta), & \beta \text{ odd} \end{cases}$$

The constant $\widetilde{c}_{\beta,k}$ is chosen in such a way that the functions are orthonormal on the unit sphere S^2 in $\mathbb{R}^3$:

$$\int_{S^2} S_{\beta,k}(\mathbf{x})\, S_{\widetilde{\beta},\widetilde{k}}(\mathbf{x})\, dS = \delta_{\beta,\widetilde{\beta}}\, \delta_{k,\widetilde{k}}$$

The functions T_k^l are the associated Legendre polynomials, see [13], [16]. According to (30) we define the basis for our space of trial functions by

$$\psi_{m,j,\beta}(\mathbf{x}) = (1 - \|\mathbf{x}\|^2) \varphi_{m,j,\beta}(\mathbf{x})$$

and we can order the basis lexicographically. To calculate all of the above functions we can use recursive algorithms similar to the one used for the Chebyshev polynomials. These algorithms also allow the calculation of the derivatives of each of these functions, see [11], [23]

For the numerical approximation of the integrals in (32) we use a quadrature formula for the unit ball B

$$\int_B g(x)\, dx = \int_0^1 \int_0^{2\pi} \int_0^{\pi} \widetilde{g}(r,\theta,\phi)\, r^2 \sin(\phi)\, d\phi\, d\theta\, dr \approx Q_q[g]$$

$$Q_q[g] := \sum_{i=1}^{2q} \sum_{j=1}^{q} \sum_{k=1}^{q} \frac{\pi}{q}\, \omega_j\, \nu_k \widetilde{g}\left(\frac{\zeta_k+1}{2}, \frac{\pi\, i}{2q}, \arccos(\xi_j)\right) \tag{40}$$

Here $\widetilde{g}(r,\theta,\phi) = g(\mathbf{x})$ is the representation of g in spherical coordinates. For the θ integration we use the trapezoidal rule, because the function is $2\pi-$periodic in θ. For the r direction we use the transformation

$$\begin{aligned} \int_0^1 r^2 v(r)\, dr &= \int_{-1}^{1} \left(\frac{t+1}{2}\right)^2 v\left(\frac{t+1}{2}\right) \frac{dt}{2} \\ &= \frac{1}{8} \int_{-1}^{1} (t+1)^2 v\left(\frac{t+1}{2}\right)\, dt \\ &\approx \sum_{k=1}^{q} \underbrace{\frac{1}{8}\nu_k'}_{=:\nu_k} v\left(\frac{\zeta_k+1}{2}\right) \end{aligned}$$

Table 1: Maximum errors in Galerkin solution u_n

n	N_n	$\|u-u_n\|_\infty$	$cond$	n	N_n	$\|u-u_n\|_\infty$	$cond$
2	6	$4.41E-1$	3.42	14	120	$9.95E-6$	141.2
3	10	$4.21E-1$	4.99	15	136	$3.03E-6$	165.8
4	15	$1.70E-1$	9.27	16	153	$8.31E-7$	192.8
5	21	$9.63E-2$	13.6	17	171	$2.09E-7$	222.1
6	28	$4.73E-2$	20.7	18	190	$5.21E-8$	253.8
7	36	$1.88E-2$	28.5	19	210	$1.42E-8$	287.9
8	45	$7.24E-3$	39.0	20	231	$3.53E-9$	324.4
9	55	$2.79E-3$	50.5	21	253	$7.58E-10$	363.4
10	66	$9.58E-4$	64.7	22	276	$1.46E-10$	404.9
11	78	$3.20E-4$	80.4	23	300	$3.36E-11$	448.9
12	91	$9.67E-5$	98.6	24	325	$7.16E-12$	495.4
13	105	$3.01E-5$	118.7	25	351	$1.44E-12$	544.4

where the ν_k' and ζ_k are the weights and the nodes of the Gauss quadrature with q nodes on $[-1,1]$ with respect to the inner product

$$(v,w) = \int_{-1}^{1} (1+t)^2 v(t) w(t)\, dt$$

The weights and nodes also depend on q but we omit this index. For the ϕ direction we use the transformation

$$\begin{aligned}\int_0^\pi \sin(\phi) v(\phi)\, d\phi &= \int_{-1}^{1} v(\arccos(\phi))\, d\phi \\ &\approx \sum_{j=1}^{q} \omega_j v(\arccos(\xi_j))\end{aligned}$$

where the ω_j and ξ_j are the nodes and weights for the Gauss–Legendre quadrature on $[-1,1]$. For more information on this quadrature rule on the unit ball in $\mathbb{R}^3$, see [20].

Finally we need the gradient in Cartesian coordinates to approximate the integral in (32), but the function $\varphi_{m,j,\beta}(x)$ in (39) is given in spherical coordinates. Here we simply use the chain rule, with $\mathbf{x} = (x,y,z)$,

$$\begin{aligned}\frac{\partial}{\partial x} v(r,\theta,\phi) = \frac{\partial}{\partial r} v(r,\theta,\phi)\cos(\theta)\sin(\phi) - \frac{\partial}{\partial \theta} v(r,\theta,\phi)\frac{\sin(\theta)}{r\sin(\phi)} \\ + \frac{\partial}{\partial \phi} v(r,\theta,\phi)\frac{\cos(\theta)\cos(\phi)}{r}\end{aligned}$$

and similarly for $\frac{\partial}{\partial y}$ and $\frac{\partial}{\partial z}$.

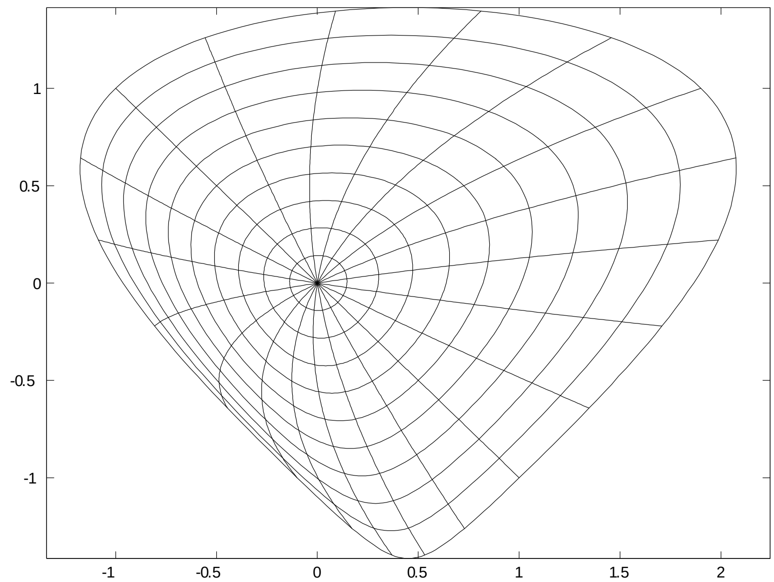

Figure 1: Images of (43), with $a = 0.5$, for lines of constant radius and constant azimuth on the unit disk.

5 Numerical example

Our programs are written in MATLAB and can be obtained from the authors. Our transformations have been so chosen that we can invert explicitly the mapping Φ, to be able to better construct our test examples. This is not needed when applying the method; but it simplified the construction of our test cases. The elliptic equation being solved is

$$Lu(\mathbf{s}) \equiv -\Delta u + \gamma(\mathbf{s})u(\mathbf{s}) = f(\mathbf{s}), \qquad \mathbf{s} \in \Omega \subseteq \mathbb{R}^d \tag{41}$$

which corresponds to choosing $A = I$. Then we need to calculate

$$\begin{aligned} \widetilde{A}(\mathbf{x}) &= K(\Phi(\mathbf{x}))\, K(\Phi(\mathbf{x}))^{\mathrm{T}} \\ K(\Phi(\mathbf{x})) &= J(\mathbf{x})^{-1} \end{aligned} \tag{42}$$

5.1 The planar case

For our variables, we replace $\mathbf{x} \in B$ with (x, y), and we replace $\mathbf{s} \in \Omega$ with (s, t). Define the mapping $\Phi : \overline{B} \to \overline{\Omega}$ by $(s, t) = \Phi(x, y)$,

$$\begin{aligned} s &= x - y + ax^2 \\ t &= x + y \end{aligned} \tag{43}$$

with $0 < a < 1$. It can be shown that Φ is a 1-1 mapping from the unit disk $\overline{B}$. In particular, the inverse mapping $\Psi : \overline{\Omega} \to \overline{B}$ is given by

$$\begin{aligned} x &= \frac{1}{a}\left[-1+\sqrt{1+a\left(s+t\right)}\right] \\ y &= \frac{1}{a}\left[at-\left(-1+\sqrt{1+a\left(s+t\right)}\right)\right] \end{aligned} \tag{44}$$

In Figure 1, we give the images in $\overline{\Omega}$ of the circles $r = j/10$, $j = 1, \ldots, 10$ and the azimuthal lines $\theta = j\pi/10$, $j = 1, \ldots, 20$.

The following information is needed when implementing the transformation from $-\Delta u + \gamma u = f$ on Ω to a new equation on B:

$$D\Phi = J\left(x, y\right) = \begin{pmatrix} 1+2ax & -1 \\ 1 & 1 \end{pmatrix}$$

$$\det\left(J\right) = 2\left(1+ax\right)$$

$$K = \frac{1}{2\left(1+ax\right)}\begin{pmatrix} 1 & 1 \\ -1 & 1+2ax \end{pmatrix}$$

$$\widetilde{A} = KK^T = \frac{1}{2\left(1+ax\right)^2}\begin{pmatrix} 1 & ax \\ ax & 2a^2x^2+2ax+1 \end{pmatrix}$$

$$\det\left(J\right)\widetilde{A} = \frac{1}{1+ax}\begin{pmatrix} 1 & ax \\ ax & 2a^2x^2+2ax+1 \end{pmatrix}$$

The latter are the coefficients for the transformed elliptic operator over B, given in (24).

We give numerical results for solving the equation

$$-\Delta u\left(s,t\right) + e^{s-t}u\left(s,t\right) = f\left(s,t\right), \qquad \left(s,t\right) \in \Omega \tag{45}$$

As a test case, we choose

$$u\left(s,t\right) = \left(1-x^2-y^2\right)\cos\left(\pi s\right) \tag{46}$$

with (x, y) replaced using (44). The solution is pictured in Figure 2. To find $f(s, t)$, we use (45) and (46). We use the domain parameter $a = 0.5$, with Ω pictured in Figure 1.

Numerical results are given in Table 1. The integrations in (32) were performed with (38); and the integration parameter q ranged from 10 to 30. We give the condition numbers of the linear system (32) as produced in Matlab. To calculate the error, we evaluate the the numerical solution and the error on the grid

$$\begin{aligned} \Phi\left(x_{i,j}, y_{i,j}\right) &= \Phi\left(r_i\cos\theta_j, r_i\sin\theta_j\right) \\ \left(r_i, \theta_j\right) &= \left(\frac{i}{10}, \frac{j\pi}{10}\right), \qquad i = 0, 1, \ldots 10; \quad j = 1, \ldots 20 \end{aligned}$$

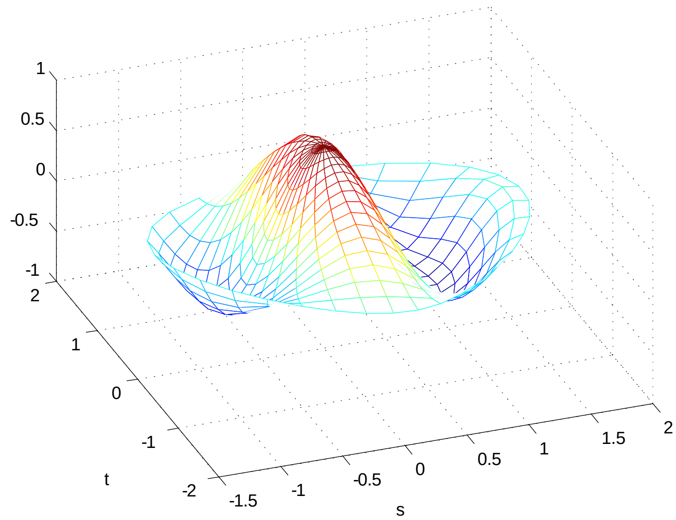

Figure 2: The true solution (46)

The results are shown graphically in Figure 3. The use of a semi-log scale demonstrates the exponential convergence of the method as the degree increases.

To examine experimentally the behaviour of the condition numbers for the linear system (32), we have graphed the condition numbers from Table 1 in Figure 4. Note that we are graphing N_n vs. the condition number of the associated linear system. The graph seems to indicate that the condition number of the system (32) is directly proportional to the order of the system, with the order given in (33).

To check experimentally the behaviour of our method when the strong ellipticity condition (7) is not satisfied, we also solved the equation

$$-\Delta u(s,t) - e^{s-t}u(s,t) = f(s,t), \qquad (s,t) \in \Omega \tag{47}$$

a modification of (45), with the same solution as in (46). The error results are an improvement to those given in Table 1 for (45); for example, $\|u - u_{16}\|_\infty = 1.14 \times 10^{-10}$. The condition numbers are slightly larger; for example, the condition number for the system with $n = 16$ is approximately 283. The condition numbers show the same type of growth with N_n as shown in Figure 4.

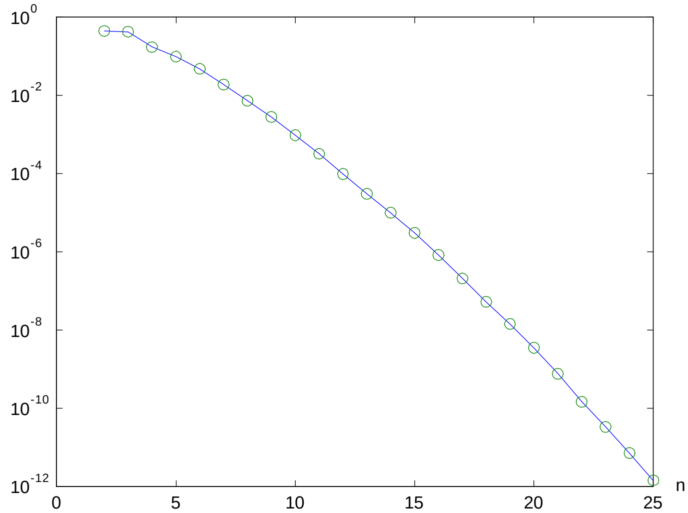

Figure 3: Errors from Table 1

5.2 The three-dimensional case

Here we define the mapping $\Phi : \overline{B} \to \overline{\Omega}$ by $(s,t,u) = \Phi(x,y,z)$,

$$\begin{aligned} s &= x - y + ax^2 \\ t &= x + y \\ u &= 2z + bz^2 \end{aligned} \tag{48}$$

$0 < a, b < 1$, which is an extension of the mapping defined in (43). The inverse mapping $\Psi : \overline{\Omega} \to \overline{B}$ is given by

$$\begin{aligned} x &= \frac{1}{a}\left[-1 + \sqrt{1 + a\,(s+t)}\right] \\ y &= \frac{1}{a}\left[at - \left(-1 + \sqrt{1 + a\,(s+t)}\right)\right] \\ z &= \frac{1}{b}\left[-1 + \sqrt{1 + bu}\right] \end{aligned}$$

In Figure 5 we show the image of the surface of $\overline{B}$ under Φ. As in the planar case, we also need

$$D\Phi(x,y,z) :=: J(x,y,z) = \begin{pmatrix} 1 + 2ax & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 + 2bz \end{pmatrix}$$

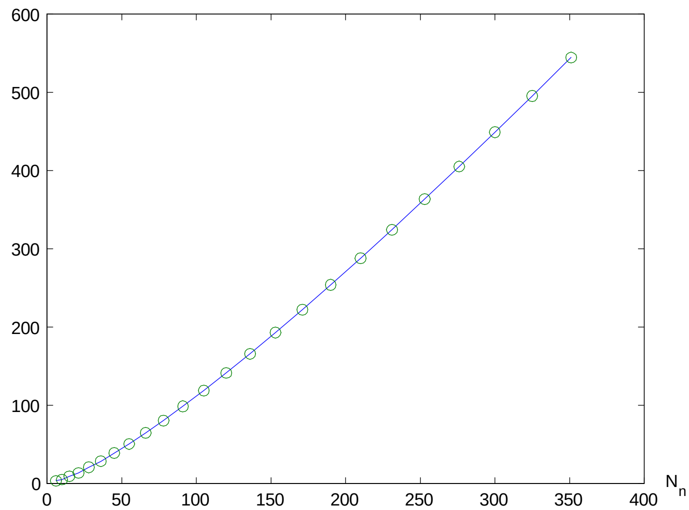

Figure 4: Condition numbers from Table 1

$$\det(J(x,y,z)) = 4(1+ax)(1+bz)$$

and

$$\begin{aligned}
&\det(J(x,y,z))\widetilde{A}(x,y,z)\\
&= \det(J(x,y,z)K(x,y,z)K^T(x,y,z)\\
&= 4(1+ax)(1+bz)\begin{pmatrix} \dfrac{1}{2(1+ax)^2} & \dfrac{ax}{2(1+ax)^2} & 0 \\ \dfrac{ax}{2(1+ax)^2} & \dfrac{1+ax+2a^2x^2}{2(1+ax)^2} & 0 \\ 0 & 0 & \dfrac{1}{4(1+bz)^2} \end{pmatrix}
\end{aligned}$$

Again, these are the coefficients for the second order term for the transformed equation on $\overline{B}$, given in (24). We give numerical results for solving the equation

$$-\Delta v(s,t,u) + e^{s-t}v(s,t,u) = f(s,t,u), \quad (s,t,u) \in \Omega$$

and for our test case we choose

$$v(s,t,u) = \sin\left(\frac{1}{2}(s-t)\right) \cdot (1 - \|\Psi(s,t,u)\|^2)$$

where the second term guarantees the Dirichlet boundary conditions on $\overline{\Omega}$.

Numerical results are given in Table 2.

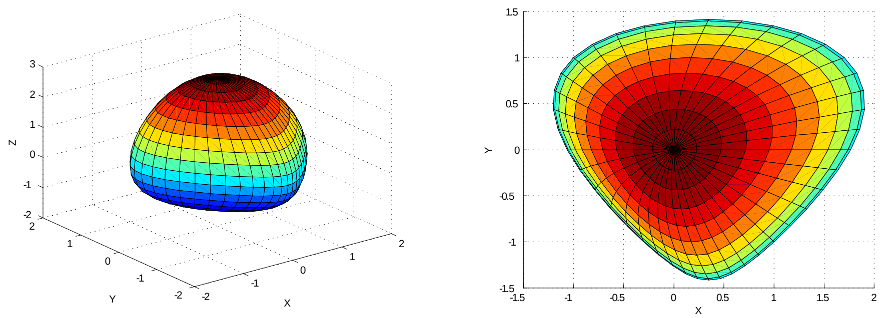

Figure 5: Image of (48) from two different angles, with a=0.7, b=0.9, for lines of constant φ and θ on the sphere.

The integrations in (32) were performed with (40); and the integration parameter q was chosen as $q = n+2$. Numerical experiments indicate that a larger q does not change the results significantly. The condition numbers for the system (32) were again calculated with MATLAB. An estimation for the error in the maximum norm was calculated on the grid given by

$$\begin{pmatrix} x_{i,j,k} \\ y_{i,j,k} \\ z_{i,j,k} \end{pmatrix} = \begin{pmatrix} \frac{i}{21}\sin\left(\frac{k}{21}\pi\right)\cos\left(\frac{2j}{20}\pi\right) \\ \frac{i}{21}\sin\left(\frac{k}{21}\pi\right)\sin\left(\frac{2j}{20}\pi\right) \\ \frac{i}{21}\cos\left(\frac{k}{21}\pi\right) \end{pmatrix}, \quad i,k = 1,\ldots,20, \quad j = 1,\ldots,40.$$

The error for the Galerkin method is shown in Figure 6 and the development of the condition number is shown in Figure 7. Again the numerical experiment seems to indicate an exponential convergence of the method and a linear growth of the condition numbers with respect to the number of degrees of freedom N_n of the linear system (32).

ADDITIONAL REMARKS. We present and study a spectral method for the Neumann problem

$$\begin{aligned} -\Delta u + \gamma(\mathbf{s})u(\mathbf{s}) &= f(\mathbf{s}), \qquad \mathbf{s} \in \Omega \subseteq \mathbb{R}^d \\ \frac{\partial u(\mathbf{s})}{\partial \mathbf{n}_s} &= g(\mathbf{s}), \qquad \mathbf{s} \in \partial\Omega \end{aligned}$$

in a forthcoming paper. We are also investigating the behaviour of the condition number for the linear system (32), attempting to prove that it has size $\mathcal{O}(N_n)$, consistent with the numbers shown in Tables 1 and 2.

Table 2: Maximum errors in Galerkin solution u_n

n	N_n	$\|u-u_n\|_\infty$	$cond$
1	4	$4.98E-1$	1.5
2	10	$1.99E-1$	3.6
3	20	$1.78E-1$	5.7
4	35	$8.22E-2$	11.0
5	56	$2.18E-2$	17.1
6	84	$1.34E-2$	27.1
7	120	$5.95E-3$	39.4
8	165	$1.60E-3$	55.9
9	220	$4.85E-4$	75.8
10	286	$2.56E-4$	100.2
11	364	$1.44E-4$	128.9
12	455	$7.85E-5$	162.4
13	560	$4.19E-5$	200.6
14	680	$2.33E-5$	244.0

Our numerical examples in this section use given

$$\Phi : \overline{B} \xrightarrow[onto]{1-1} \overline{\Omega}, \tag{49}$$

chosen to be nontrivial and illustrative. In general, however, when given a smooth mapping

$$\varphi : \partial B \xrightarrow[onto]{1-1} \partial\Omega$$

it may not be clear as to how to extend φ to Φ over $\overline{B}$. In some cases, there is an obvious choice, as when Ω is an ellipsoid,

$$\Phi(x, y, z) = (ax, by, cz), \qquad (x, y, z) \in \partial B.$$

We are investigating general schemes to produce continuously differentiable extensions Φ of φ, with Φ satisfying (49) and having an easily computable Jacobian $J(\mathbf{x})$ for which

$$\min_{x\in\overline{B}} |\det J(\mathbf{x})| > 0$$

This is the subject of a future paper.

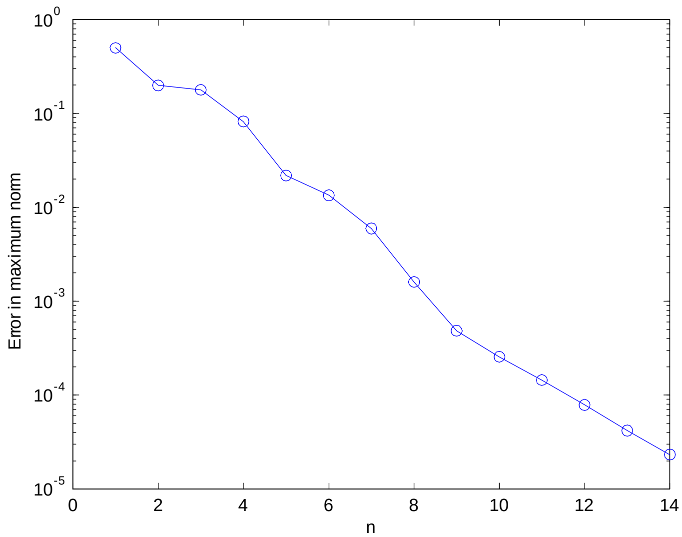

Figure 6: Errors from Table 2

References

[1] M. Abramowitz, I.A. Stegun, *Handbook of Mathematical Functions*, Dover Publications, Inc., New York, 1965.

[2] K. Atkinson. *An Introduction to Numerical Analysis*, 2nd ed., John Wiley, New York, 1989.

[3] K. Atkinson and W. Han. *Theoretical Numerical Analysis*: *A Functional Analysis Framework*, 2nd ed., Springer-Verlag, New York, 2005.

[4] K. Atkinson and O. Hansen. Solving the nonlinear Poisson equation on the unit disk, *Journal of Integral Equations & Applications,* **17** (2005), 223-241.

[5] J. Boyd. *Chebyshev and Fourier Spectral Methods*, 2nd ed., Dover Pub., New York, 2000.

[6] S. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*, 2nd ed., Springer-Verlag, New York, 2002.

[7] C. Canuto, A. Quarteroni, My. Hussaini, and T. Zang, *Spectral Methods in Fluid Mechanics*, Springer-Verlag, 1988.

[8] C. Canuto, A. Quarteroni, My. Hussaini, and T. Zang, *Spectral Methods - Fundamentals in Single Domains*, Springer-Verlag, 2006.

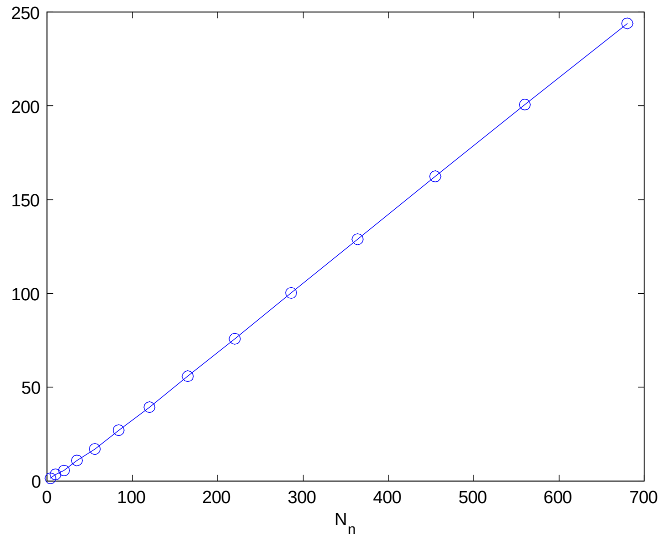

Figure 7: Condition numbers from Table 2

[9] E. Doha and W. Abd-Elhameed. Efficient spectral-Galerkin algorithms for direct solution of second-order equations using ultraspherical polynomials, *SIAM J. Sci. Comput.* **24** (2002), 548-571.

[10] C. Dunkl and Y. Xu. *Orthogonal Polynomials of Several Variables*, Cambridge Univ. Press, Cambridge, 2001.

[11] W. Gautschi, *Orthogonal Polynomials*, Oxford University Press, Oxford, 2004.

[12] O. Hansen, K. Atkinson, and D. Chien. On the norm of the hyperinterpolation operator on the unit disk and its use for the solution of the nonlinear Poisson equation, *IMA J. Numerical Analysis*, to appear.

[13] E.W. Hobson, *The Theory of Spherical and Ellipsoidal Harmonics*, Chelsea Publishing Company, New York, 1965.

[14] V. Liseikin, *Grid Generation Methods*, Springer-Verlag, Berlin, 1999.

[15] B. Logan. and L. Shepp. Optimal reconstruction of a function from its projections, *Duke Mathematical Journal* **42**, (1975), 645–659.

[16] T. M. MacRobert, *Spherical Harmonics*, Dover Publications, Inc., New York, 1948.

[17] S. Mikhlin. *Mathematical Physics, An Advanced Course*, North-Holland, 1970.

[18] D. Ragozin. Constructive polynomial approximation on spheres and projective spaces, *Trans. Amer. Math. Soc.* **162** (1971), 157-170.

[19] J. Shen and L. Wang. Analysis of a spectral-Galerkin approximation to the Helmholtz equation in exterior domains, *SIAM J. Numer. Anal.* **45** (2007), 1954-1978.

[20] A. Stroud. *Approximate Calculation of Multiple Integrals,* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1971.

[21] Yuan Xu. Lecture notes on orthogonal polynomials of several variables, in *Advances in the Theory of Special Functions and Orthogonal Polynomials*, Nova Science Publishers, 2004, 135-188.

[22] Yuan Xu. Analysis on the unit ball and on the simplex, *Electronic Trans. on Numer. Anal.* **25** (2006), 284-301.

[23] S. Zhang, J. Jin, *Computation of Special Functions*, John Wiley & Sons, Inc., New York, 1996.