Numer Algor (2019) 81:313–344
https://doi.org/10.1007/s11075-018-0550-y

ORIGINAL PAPER

A spectral method for an elliptic equation with a nonlinear Neumann boundary condition

Kendall Atkinson[1] · **David Chien**[2] · **Olaf Hansen**[2]

Received: 19 February 2018 / Accepted: 14 May 2018 / Published online: 29 May 2018

Abstract Let Ω be an open region in $\mathbb{R}^d$, $d \geq 2$, that is diffeomorphic to $\mathbb{B}^d$. Consider solving $-\Delta u + \gamma u = 0$ on Ω with the Neumann boundary condition $\frac{\partial u}{\partial \mathbf{n}} = b(\cdot, u)$ over $\partial\Omega$. The function b is a nonlinear function of u. The problem is reformulated in a weak form, and then a spectral Galerkin method is used to create a sequence of finite dimensional nonlinear problems. An error analysis shows that under suitable assumptions, the solutions of the finite dimensional problems converge to those of the original problem. To carry out the error analysis, the original problem and the spectral method is converted to a nonlinear integral equation over $H^{1/2}(\Omega)$, and the reformulation is analyzed using tools for solving nonlinear integral equations. Numerical examples are given to illustrate the method. In our error analysis, we assume the existence and local uniqueness of a solution. For the case of three dimensions and a nonlinearity b that is given by the Stefan–Boltzmann law, we will provide an existence proof in the final section.

Keywords Elliptic · Nonlinear · Neumann · Spectral

Mathematics Subject Classification (2010) 65N35

✉ Kendall Atkinson
Kendall-Atkinson@uiowa.edu

David Chien
chien@csusm.edu

Olaf Hansen
ohansen@csusm.edu

[1] The University of Iowa, Iowa City, IA, USA

[2] California State University at San Marcos, San Marcos, CA, USA

1 Introduction

Consider solving the problem

$$-\Delta u(s) + \gamma(s)u(s) = 0, \qquad s \in \Omega, \tag{1.1}$$

$$\frac{\partial u(s)}{\partial n_s} = b(s, u(s)), \qquad s \in \partial\Omega, \tag{1.2}$$

in which the Neumann boundary condition is a function of the solution u. Let Ω be an open region in $\mathbb{R}^d$ that is diffeomorphic to the unit ball $\mathbb{B}^d$ and assume that its boundary $\partial\Omega$ is smooth and homeomorphic to $\mathbb{S}^{d-1}$, $d \geq 2$. The function γ is assumed to be smooth and to satisfy additional conditions, specified below. Let n_s denote the exterior unit normal at $s \in \partial\Omega$. For simplicity, assume $\partial\Omega$ is infinitely differentiable, although that is stronger than needed. The function $b(\cdot, u)$ is generally nonlinear in u, although the linear case is certainly also important. This problem can be given a weak reformulation, much as with finite element methods, and from this, we give a spectral method for its numerical solution. The spectral method is defined using an explicitly given transformation Φ from $\mathbb{B}^d$ onto Ω. An earlier paper using this numerical approach to solve elliptic equations with a Neumann boundary condition is [8].

In Section 2, we derive the variational formulation of the nonlinear boundary value problem and its discrete counterpart, based on trial spaces of multivariate orthogonal polynomials over $\mathbb{B}^d$. The resulting nonlinear equations are solved by using the built in MATLAB function `fsolve` in combination with some bootstrap method. In Section 3, we present two two-dimensional and two three-dimensional numerical examples. For dimension 2, one example has a nonlinear term that is sufficiently bounded such that the corresponding Nemytskii operator (see the definition in Section 4) is continuous on an appropriate Sobolev space. The nonlinearity of the other examples have a Stefan–Boltzmann type nonlinearity. For all given examples, our proposed method shows rapid convergence. Section 4 provides a convergence proof for our method assuming that a solution u^* of the nonlinear boundary problem (1.1)–(1.2) exists, an initial estimate is sufficiently close to u^*, and that the nonlinear term generates a continuous Nemytskii operator. The analysis uses a reformulation of (1.1)–(1.2) as a nonlinear integral equation. The error analysis assumes the nonlinear function $b(\cdot, u)$ is restricted in how it increases as a function of u; but we define and illustrate the numerical method without this limitation. So Section 4 explains the convergence results from Section 3 for the case of the sufficiently bounded nonlinearity. One has to keep in mind that nonlinear problems will have multiple solutions in general, so only local convergence results can be expected. For the case of the Stefan–Boltzmann nonlinearity, we do not have a convergence proof for our method, but in Section 5, we present a proof that the nonlinear variational problem given in Section 2 has at least one positive solution and that our method provides positive solutions to the corresponding discrete nonlinear variational problem.

The theory of the solvability of (1.1)–(1.2) has been studied by a number of researchers. An important approach to this theory is to use the Green's function for solving the linear Neumann problem

$$-\Delta u(s) + \gamma(s)u(s) = 0, \qquad s \in \Omega, \tag{1.3}$$

$$\frac{\partial u(s)}{\partial n_s} = b(s), \qquad s \in \partial\Omega, \tag{1.4}$$

with γ so chosen that there is a unique solution, e.g., assume $\gamma \in C\left(\overline{\Omega}\right)$ and

$$\min_{s\in\overline{\Omega}} \gamma(s) > 0. \tag{1.5}$$

More generally, assume γ is so chosen that the equation (1.3) is strongly elliptic. In the error analysis of our method, the function γ is assumed, for simplicity, to be as smooth as needed. The solution to (1.3)–(1.4) can be written as

$$u(s) = \int_{\partial\Omega} G(s,t)\, b(t)\, dt, \qquad s \in \Omega; \tag{1.6}$$

(see [17, §7.2]). For $\gamma = 0$ (see [27, §9.6; p. 269]). The Green's function G (also called a Neumann function) satisfies

$$G(s,t) = \begin{cases} \mathcal{O}(\log|s-t|), & d = 2, \\ \mathcal{O}\left(|s-t|^{-d+2}\right), & d \geq 3. \end{cases}$$

The operator

$$\left(\widehat{\mathcal{G}}v\right)(s) = \int_{\partial\Omega} G(s,t)\, v(t)\, dt, \qquad s \in \partial\Omega,$$

is a compact operator from $L^2(\partial\Omega) \to L^2(\partial\Omega)$ and from $C(\partial\Omega) \to C(\partial\Omega)$ (see [25, §§7.3, 7.4]). This is also true for other spaces as will be discussed later. With this operator, the solvability of (1.1)–(1.2) can be converted to the solvability of the nonlinear integral equation

$$u(s) = \int_{\partial\Omega} G(s,t)\, b(t, u(t))\, dt, \qquad s \in \partial\Omega, \tag{1.7}$$

over the boundary $\partial\Omega$ for a variety of function spaces, e.g., $C(\partial\Omega)$. Once a solution is found over $\partial\Omega$, the solution over all of Ω follows from using (1.7) with s ranging over all of Ω.

The equation (1.7) can be analyzed as a fixed point problem, e.g., using the Banach fixed point theorem over $C(\partial\Omega)$ [5, p.208] and assuming (i) $b(s,\xi)$ is sufficiently small, and (ii) $b(s,\xi)$ has a sufficiently small Lipschitz constant with respect to ξ. Another approach is to use the theory of positive operators [22]. Using this approach, a very powerful solvability theorem in $C\left(\overline{\Omega}\right)$ is given by Amann [1, p. 47] (also see Zeidler [31, §7.17]). Further results can be obtained using the theory of monotone operators (e.g., see Zeidler [32, §28.3]). Most of these results are in the function spaces $C(\partial\Omega)$ or $L^2(\partial\Omega)$; but we use other Sobolev spaces in Section 4 for the error analysis of our numerical method. Most of these theoretical results require that the function $b(s,\xi)$ be bounded linearly in ξ, eliminating some important cases of interest. Similar restrictions on $b(\cdot,u)$ are also required in the error analysis of our numerical method.

One motivation for studying equations of type (1.1)–(1.2) are heat equations, where the transport in the domain Ω is given by diffusion and on the surface $\partial\Omega$ there is given incoming and radiation is outgoing Stefan–Boltzmann type radiation (see [20]). For this type of nonlinearity, $b(s,\xi)$ behaves like ξ^4. For this problem

(1.1)–(1.2), we give numerical examples in Section 3 and an existence proof in Section 5.

2 The numerical method

Multiplying (1.1) by $v(s)$ and integrating by parts over Ω leads to the weak framework for (1.1)–(1.2): Find $u \in H^1(\Omega)$ such that

$$\mathcal{A}(u, v) = \int_{\partial\Omega} b(t, u(t))\, v(t)\, dt, \qquad \forall v \in H^1(\Omega), \tag{2.1}$$

with the bilinear functional

$$\mathcal{A}(v_1, v_2) = \int_{\Omega} [\nabla v_1(s) \cdot \nabla v_2(s) + \gamma(s) v_1(s) v_2(s)]\, ds. \tag{2.2}$$

The meaning of the Sobolev space $H^m(\Omega)$ (and also $H^{m-\frac{1}{2}}(\partial\Omega)$), with $m \geq 0$ an integer is discussed below in Section 4.

Let $\mathcal{X}_n$ denote the approximation space over Ω, to be based on a transformation Φ of $\mathbb{B}^d$ onto Ω and using Π_n^d, the polynomials of degree $\leq n$ in d variables (see [14]). For the numerical method, find $u_n \in \mathcal{X}_n$ such that

$$\mathcal{A}(u_n, v) = \int_{\partial\Omega} b(t, u_n(t))\, v(t)\, dt, \qquad \forall v \in \mathcal{X}_n. \tag{2.3}$$

2.1 Transformation of Ω

As in earlier papers, assume that a mapping

$$\Phi : \overline{\mathbb{B}}^d \xrightarrow[onto]{1-1} \overline{\Omega}$$

is known, and let $\Psi = \Phi^{-1} : \overline{\Omega} \xrightarrow[onto]{1-1} \overline{\mathbb{B}}^d$ denote the inverse mapping. Let $J(x) \equiv (D\Phi)(x)$, $x \in \overline{\mathbb{B}}^d$, and $K(s) \equiv (D\Psi)(s)$, $s \in \overline{\Omega}$, denote the Jacobian matrix of the transformations Φ and Ψ, respectively. Assume $J(x)$ is nonsingular on $\overline{\mathbb{B}}^d$, and without loss of generality, assume

$$\det J(x) > 0, \qquad x \in \overline{\mathbb{B}}^d.$$

Differentiating the identity $\Psi(\Phi(x)) = x$ over $\mathbb{B}^d$ leads to

$$K(s)\, J(x) = I, \qquad s = \Phi(x), \tag{2.4}$$

$$K(\Phi(x)) = J(x)^{-1}, \qquad x \in \overline{\mathbb{B}}^d. \tag{2.5}$$

Let v denote a general function defined over Ω. For the transformation $s = \Phi(x)$, introduce the notation $\widetilde{v}(x) = v(\Phi(x))$; or equivalently, $v(s) = \widetilde{v}(\Psi(s))$. Consider the derivatives with respect to s of $v(s)$. Let ∇_s denote the gradient with respect to the components of s; and do similarly for ∇_x. Then

$$\nabla_s v(s) = K(s)^{\mathrm{T}} \nabla_x \widetilde{v}(x), \qquad x = \Psi(s), \tag{2.6}$$

$$\nabla_x \widetilde{v}(x) = J(x)^{\mathrm{T}} \nabla_s v(s), \qquad s = \Phi(x), \tag{2.7}$$

with $\nabla_x \widetilde{v}(x)$ the gradient of $\widetilde{v}(x)$ written as a column vector, and analogously for $\nabla_s v(s)$. An investigation of techniques for creating such mappings Φ is given in [7].

The convergence analysis will require comparing norms over Ω and $\mathbb{B}^d$, which the following lemma addresses. For simplicity, assume $\Phi \in C^\infty\left(\overline{\mathbb{B}}^d\right)$, although this can be relaxed. The proof is relatively straightforward.

Lemma 2.1 *Let $m \geq 0$ be an integer. A function $v \in H^m(\Omega)$ if and only if $\widetilde{v} \in H^m\left(\mathbb{B}^d\right)$. Moreover,*

$$c_{1,m}\|v\|_{H^m(\Omega)} \leq \|\widetilde{v}\|_{H^m\left(\mathbb{B}^d\right)} \leq c_{2,m}\|v\|_{H^m(\Omega)}, \qquad v \in H^m(\Omega),\ \widetilde{v} = v \circ \Phi,$$

with constants $c_{1,m}, c_{2,m} > 0$.

2.2 The approximation space

Let Π_n^d denote the polynomials of degree $\leq n$ over $\mathbb{R}^d$. Define

$$\mathcal{X}_n = \left\{\psi \circ \Phi^{-1} |\ \psi \in \Pi_n^d\right\}. \tag{2.8}$$

The following approximation theorem [19, Thm. 4.2] is needed for polynomial approximation over the unit ball $\mathbb{B}^d$.

Theorem 2.2 *(Li and Xu) Let $r \geq 2$ be an integer. Given $v \in H^r\left(\mathbb{B}^d\right)$, there exists a sequence of polynomials $p_n \in \Pi_n^d$ such that*

$$\|v - p_n\|_{H^1\left(\mathbb{B}^d\right)} \leq \varepsilon_{n,r}\|v\|_{H^r\left(\mathbb{B}^d\right)}, \qquad n \geq 1. \tag{2.9}$$

The sequence $\varepsilon_{n,r} = \mathcal{O}\left(n^{-r+1}\right)$ and is independent of v.

Combining Lemma 2.1 with the definition of $\mathcal{X}_n$ in (2.8), the result (2.9) can be extended to the approximation of functions over Ω using $\mathcal{X}_n$.

Let N_n denote the dimension of $\mathcal{X}_n$ and Π_n^d:

$$N \equiv N_n = \binom{n+d}{d}.$$

Let $\{\varphi_1, \ldots, \varphi_N\}$ denote an orthonormal basis for Π_n^d, and let

$$\{\psi_1, \ldots, \psi_N\} = \left\{\varphi_1 \circ \Phi^{-1}, \ldots, \varphi_N \circ \Phi^{-1}\right\}$$

denote a corresponding basis (usually not orthonormal) for $\mathcal{X}_n$. For a discussion of multivariate orthogonal polynomials, see [16] and [30]. A special choice of basis polynomials $\{\varphi_\ell\}$ and their efficient evaluation is discussed in [4].

2.3 The numerical method

For the numerical method (2.3), write

$$u_n(s) = \sum_{\ell=1}^{N_n} \alpha_\ell \psi_\ell(s).$$

Solve for the coefficients $\{\alpha_\ell\}$ for which

$$\begin{aligned}\sum_{\ell=1}^{N_n}\alpha_\ell \int_\Omega \left[\nabla\psi_\ell(s)\cdot\nabla\psi_k(s)+\gamma\,(s)\,\psi_\ell\,(s)\,\psi_k\,(s)\right]\,ds \\ = \int_{\partial\Omega} b\left(s,\sum_{\ell=1}^{N_n}\alpha_\ell\psi_\ell\,(s)\right)\psi_k(s)\,ds, \qquad k=1,\ldots,N_n.\end{aligned} \tag{2.10}$$

The function $\widetilde{u}_n\,(x)\equiv u_n\,(\Phi\,(x)),\, x\in\mathbb{B}^d$ is the equivalent solution considered over $\mathbb{B}^d$,

$$\widetilde{u}_n\,(x)=\sum_{\ell=1}^{N_n}\alpha_\ell\varphi_\ell\,(x)\,.$$

Using the transformation of variables $s=\Phi\,(x)$ in the system (2.10), the coefficients $\{\alpha_\ell|\ell=1,2,\ldots,N_n\}$ are the solutions of

$$\begin{aligned}\sum_{k=1}^{N_n}\alpha_k\int_{\mathbb{B}^d}\left[\sum_{i,j=1}^{d}\widetilde{a}_{i,j}(x)\frac{\partial\varphi_k(x)}{\partial x_j}\frac{\partial\varphi_\ell(x)}{\partial x_i}+\gamma(\Phi\,(x))\varphi_k(x)\varphi_\ell(x)\right]\det J\,(x)\;dx \\ = \int_{\partial\mathbb{B}^d} b\left(x,\sum_{k=1}^{N_n}\alpha_k\varphi_k\,(x)\right)\varphi_\ell\,(x)\;\left|J_{bdy}\,(x)\right|\;dx, \qquad \ell=1,\ldots,N_n\end{aligned} \tag{2.11}$$

with

$$\widetilde{A}\,(x)=J\,(x)^{-1}\,J\,(x)^{-\mathrm{T}}\,.$$

This change of variables is derived in [3] and [8]. The term $\left|J_{bdy}\,(x)\right|$ arises from the transformation of an integral over $\partial\Omega$ to one over $\partial\mathbb{B}^d$, say

$$\int_{\partial\Omega} g\,(s)\,\psi_k(s)\,ds=\int_{\partial\mathbb{B}^d} g\,(\Phi\,(x))\,\varphi_k\,(x)\;\left|J_{bdy}\,(x)\right|\;dx.$$

In one variable ($d=2$), the boundary $\partial\Omega$ is often represented as a mapping

$$\Phi\,(\cos\theta,\sin\theta)\equiv\chi\,(\theta)=(\chi_1\,(\theta)\,,\chi_2\,(\theta))\,, \qquad 0\le\theta\le 2\pi.$$

In that case, $\left|J_{bdy}\,(x)\right|$ is simply $\left|\chi'\,(\theta)\right|$ and the associated integral is

$$\int_0^{2\pi} g\,(\chi\,(\theta))\,\varphi_k\,(\cos\theta,\sin\theta)\,\left|\chi'\,(\theta)\right|\,d\theta$$

The case for $\partial\Omega$ a surface in $\mathbb{R}^3$ is given in the Appendix to this paper.

2.4 Numerical integration

The integrals in (2.11) are over $\mathbb{B}^d$ and $\mathbb{S}^{d-1}=\partial\mathbb{B}^d$, and they need to be evaluated numerically.

2.4.1 The case $d = 2$

For integrals over $\mathbb{B}^2$, use polar coordinates to write

$$\int_{\mathbb{B}^2} q\,(x)\;dx = \int_0^{2\pi}\int_0^1 r\widetilde{q}\,(r,\theta)\;dr\,d\theta.$$

with $\widetilde{q}\,(r,\theta) \equiv q\,(r\cos\theta, r\sin\theta)$. Approximate it by

$$\int_{\mathbb{B}^2} q\,(x)\;dx \approx I_\eta\,(q) \equiv \frac{2\pi}{2\eta+1}\sum_{l=0}^{\eta}\sum_{m=0}^{2\eta}\omega_l r_l \widetilde{q}\left(r_l, \frac{2\pi m}{2\eta+1}\right). \tag{2.12}$$

The formula uses the trapezoidal rule with $2\eta+1$ subdivisions for the integration over $[0, 2\pi]$ in the azimuthal variable θ. The numbers r_l and ω_l denote, respectively, the nodes and weights of the $(\eta+1)$-point Gauss-Legendre quadrature formula on $[0, 1]$. This quadrature over $\mathbb{B}^2$ is exact for all polynomials $q \in \Pi^2_{2\eta}$. For integrals over $\partial\Omega = \mathbb{S}^1$, use the trapezoidal rule.

2.4.2 The case $d = 3$

For integrals over $\mathbb{B}^3$, use spherical coordinates to write

$$\begin{aligned}\int_{\mathbb{B}^3} q\,(x)\;dx &= \int_0^1\int_0^{2\pi}\int_0^{\pi} \widetilde{q}(r,\theta,\phi)\,r^2\sin(\theta)\,d\theta\,d\phi\,dr\\ \widetilde{q}\,(r,\theta,\phi) &= q\,(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)\end{aligned}$$

It is approximated by

$$I_\eta\,(q) \equiv \sum_{i=0}^{2\eta+1}\sum_{j=0}^{\eta}\sum_{k=0}^{\eta}\frac{\pi}{\eta+1}\,\omega_j\,\nu_k\widetilde{q}\left(\frac{\zeta_k+1}{2}, \arccos(\xi_j), \frac{\pi\,i}{2\,(\eta+1)}\right) \tag{2.13}$$

For the ϕ-integration, use the trapezoidal rule with $2\eta+2$ subdivisions, because the function is 2π-periodic in ϕ. For the r-direction, use the transformation and approximation

$$\int_0^1 r^2 v(r)\;dr = \frac{1}{8}\int_{-1}^1 (t+1)^2 v\left(\frac{t+1}{2}\right)\;dt \approx \sum_{k=0}^{\eta}\underbrace{\frac{1}{8}\nu'_k}_{=:\nu_k} v\left(\frac{\zeta_k+1}{2}\right)$$

where the ν'_k and ζ_k are the weights and the nodes of the Gauss quadrature with $\eta+1$ 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 η but this index is omitted here. For the θ-direction, use the transformation

$$\int_0^{\pi}\sin(\theta)v(\theta)\,d\theta = \int_{-1}^1 v(\arccos(z))\,dz \approx \sum_{j=0}^{\eta}\omega_j v(\arccos(\xi_j))$$

where the ω_j and ξ_j are the weights and nodes for the $(\eta+1)$-point Gauss-Legendre quadrature formula on $[-1, 1]$. This quadrature formula is exact for all polynomials $q \in \Pi^3_{2\eta}$.

For the surface $\mathbb{S}^2$, there are many possible quadrature formulas for approximating

$$I(q) = \int_{\mathbb{S}^2} q(x)\, dx.$$

These are discussed at length in the classic text of Stroud [28], and a survey is given in [6, Chap. 5].

A particularly simple and accurate formula is obtained by approximating $I(q)$ when written using spherical coordinates,

$$\begin{aligned} I(q) &= \int_0^{2\pi} \int_0^{\pi} q(\cos\phi \sin\theta, \sin\phi \sin\theta, \cos\theta) \sin\theta \, d\theta \, d\phi \\ &= \int_0^{2\pi} \int_{-1}^{1} q\left(\cos\phi\sqrt{1-z^2}, \sin\phi\sqrt{1-z^2}, z\right) dz \, d\phi. \end{aligned} \tag{2.14}$$

Approximate the ϕ-integration using the trapezoidal rule, and approximate the z-integration using Gauss-Legendre quadrature. More precisely, given $\eta > 1$, apply the trapezoidal rule with 2η subdivisions to the integration in ϕ, and apply Gauss-Legendre quadrature with η nodes to the integration in z over $[-1, 1]$. Let

$$h = \frac{\pi}{\eta}, \qquad \phi_j = jh, \quad j = 0, 1, \ldots, 2\eta.$$

Let $\{z_1, \ldots, z_\eta\}$ and $\{w_1, \ldots, w_\eta\}$ denote the Gauss-Legendre nodes and weights, respectively, over $[-1, 1]$. Then define

$$\begin{aligned} I_\eta(q) &= h \sum_{j=0}^{2\eta-1} \sum_{k=1}^{\eta} w_k q\left(\cos\phi_j\sqrt{1-z_k^2}, \sin\phi_j\sqrt{1-z_k^2}, z_k\right) \\ &= h \sum_{j=0}^{2\eta-1} \sum_{k=1}^{\eta} w_k q\left(\cos\phi_j \sin\theta_k, \sin\phi_j \sin\theta_k, \cos\theta_k\right), \end{aligned} \tag{2.15}$$

where $z_j = \cos\theta_j$, $j = 1, \ldots, \eta$. It is shown in [6, Thm. 5.4] that $I_\eta(q)$ has degree of precision $2\eta - 1$,

$$I_\eta(q) = I(q), \qquad \forall q \in \Pi_{2\eta-1}\left(\mathbb{S}^2\right),$$

with $\Pi_{2\eta-1}\left(\mathbb{S}^2\right)$ denoting all spherical polynomials of degree $\leq 2\eta - 1$.

2.5 Solving the nonlinear system

In our numerical examples, the nonlinear system (2.11) is solved for a small value of the degree n, often with an initial guess associated with $u_n^{(0)} = 0$. As n is increased, the approximate solution from a preceding n is used to generate an initial guess for the new value of n. In some cases, this was combined with a homotopy procedure. This approach to generating an initial guess for the iteration has worked in most cases

we have computed, but will probably not be sufficient in many other cases. Finding an initial guess for nonlinear problems is almost always difficult, and often it requires some knowledge of the solution being sought.

The MATLAB program `fsolve` was used to solve the nonlinear system. In the future, we plan to look at other numerical methods that take advantage of the special structure of (2.11). To estimate the error, we use as a true solution a numerical solution associated with a significantly larger value of n.

3 Numerical examples

Begin with some planar examples. To define the region Ω, begin with the boundary mapping

$$\varphi(\cos\theta, \sin\theta) \equiv (p_3 + p_1\cos\theta + p_2\sin\theta)(\cos\theta, \sin\theta), \quad 0 \le \theta \le 2\pi, \tag{3.1}$$

with $p = (1, 2, 3)$. Using the interpolation/integration method of [7], this is extended to a mapping Φ that is a polynomial of degree 2 in each component. Figure 1 illustrates the mapping, giving the images in $\overline{\Omega}$ of the circles $r = j/20$, $j = 1, \ldots, 20$ and the angular lines $\theta = j\pi/20$, $j = 1, \ldots, 40$. The problem to be solved is

$$-\Delta u(s) + u(s) = 0, \qquad s \in \Omega, \tag{3.2}$$

$$\frac{\partial u(s)}{\partial n_s} = -\tfrac{1}{10}e^{-0.5u(s)^2} + \cos\left(\tfrac{1}{5}s_1 s_2\right), \qquad s \in \partial\Omega. \tag{3.3}$$

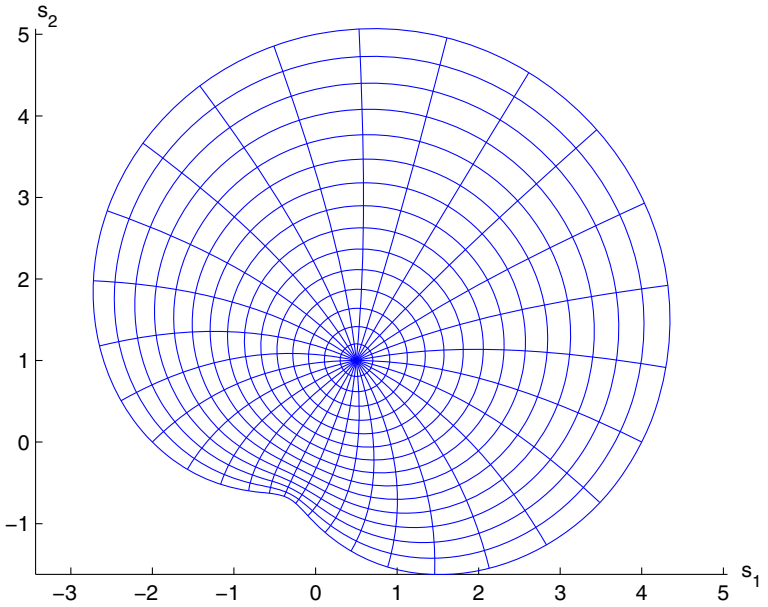

Fig. 1 Illustration of the mapping Φ for the boundary mapping (3.1)

The solution that was found is shown in Fig. 2. The numerical integration of the linear system coefficients in (2.11) were carried out as described in (2.12) with $\eta = 2n$ where n is the degree of the approximating polynomial $\widetilde{u}_n(x)$. The estimated errors are shown in Fig. 3, and the convergence appears to be exponential. Experimentally, the condition numbers of the linear system on the left side of (2.10) are approximately $\mathcal{O}\left(N_n^2\right)$.

3.1 Another planar example

The region Ω is given by means of the mapping Φ:

$$\Phi(x) = \left[x_1 - x_2 + a x_1^2,\ x_1 + x_2\right]^{\mathrm{T}}, \qquad x \in \overline{\mathbb{B}}^2, \tag{3.4}$$

for a given $0 < a < 1$. Figure 4a illustrates the mapping with $a = 0.95$, giving the images in $\overline{\Omega}$ of the circles $r = j/10$, $j = 1, \ldots, 10$ and the angular lines $\theta = j\pi/10$, $j = 1, \ldots, 20$. It can be shown that Φ is a 1-1 mapping on the unit disk $\overline{\mathbb{B}}^2$. Another mapping, denoted here by Φ_{II}, is illustrated in Fig. 4b. It is based on the interpolation/integration method of [7, §3], and each component is a polynomial of degree 2. The mappings Φ and Φ_{II} agree on $\partial\Omega$.

The problem to be solved is

$$-\Delta u(s) + 2u(s) = 0, \qquad s \in \Omega, \tag{3.5}$$

$$\frac{\partial u(s)}{\partial n_s} = -(u(s))^4 + \cos\left(\tfrac{1}{2} s_1 s_2\right), \qquad s \in \partial\Omega. \tag{3.6}$$

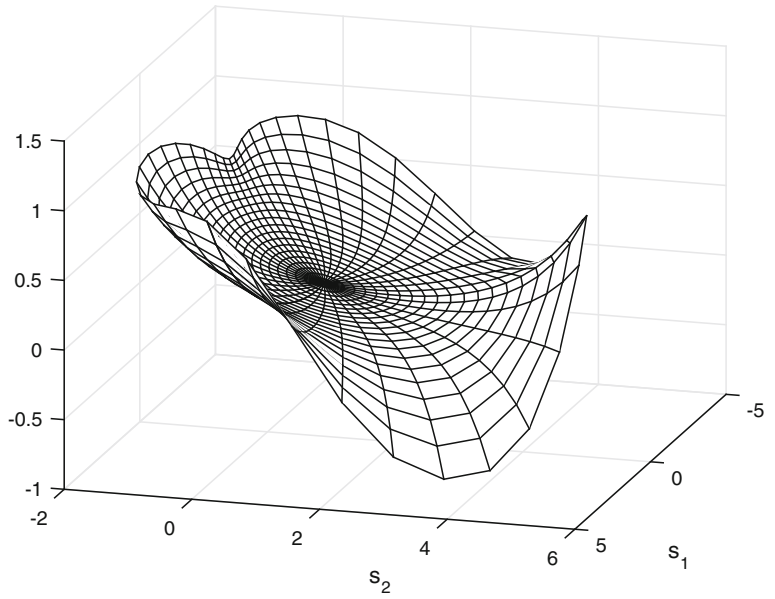

Fig. 2 The solution to (3.2)–(3.3)

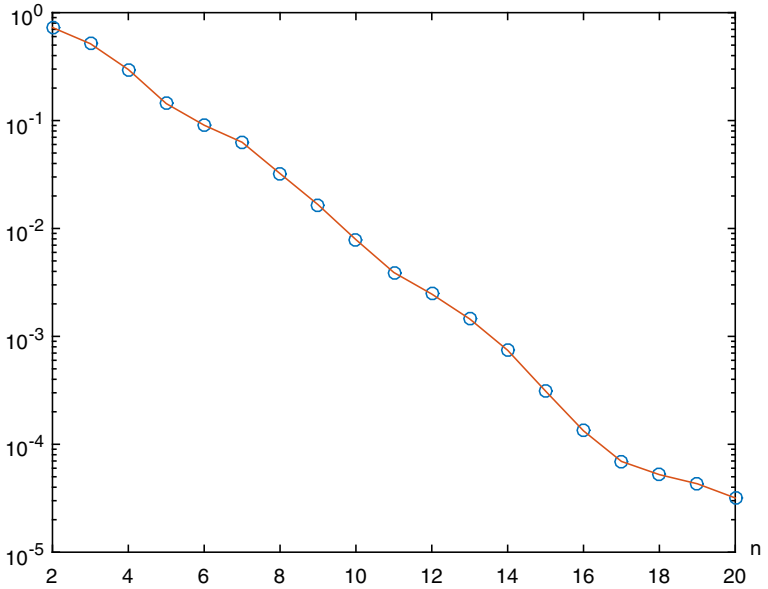

Fig. 3 The error in solving (3.2)–(3.3)

The solution that was found is shown in Fig. 5. The estimated errors when using both the mappings Φ and Φ_{II} are shown in Fig. 6. The right side of (3.6) does not satisfy the error analysis assumptions of (4.27), but the method still works well, converging exponentially.

3.2 Two three-dimensional examples

We will consider two examples in $\mathbb{R}^3$. The first region Ω_1 is given by the mapping

$$\Phi_1(x) = [x_1 - x_2 + ax_1^2, x_1 + x_2, x_3]^T, \qquad x \in \overline{\mathbb{B}}^3,$$

again with $a = 0.95$. This region is a three-dimensional extension of the mapping given in (3.4). Every horizontal (parallel to the xy–plane) cut through Ω_1 will have the same shape as the one shown in Fig. 4. The problem we solve is

$$-\Delta u(s) + 2u(s) = 0, \qquad s \in \Omega_1 \tag{3.7}$$

$$\frac{\partial u(s)}{\partial n_s} = -u^4(s) + \cos\left(\frac{s_1 s_2 s_3}{4}\right), \qquad s \in \partial\Omega_1 \tag{3.8}$$

So, similar to (3.6), the boundary term is given by a Stefan–Boltzmann radiation term and a smooth positive function, describing some incoming energy. The numerical solutions $u_n(s)$ are calculated for $n = 1, \ldots, 25$ and $u_{30}(s)$ is used as a reference solution. Figure 7 shows the approximate surface values, based on u_{30}. Figure 8 shows the the estimated errors. The convergence appears to be exponential, similar to the two-dimensional examples.

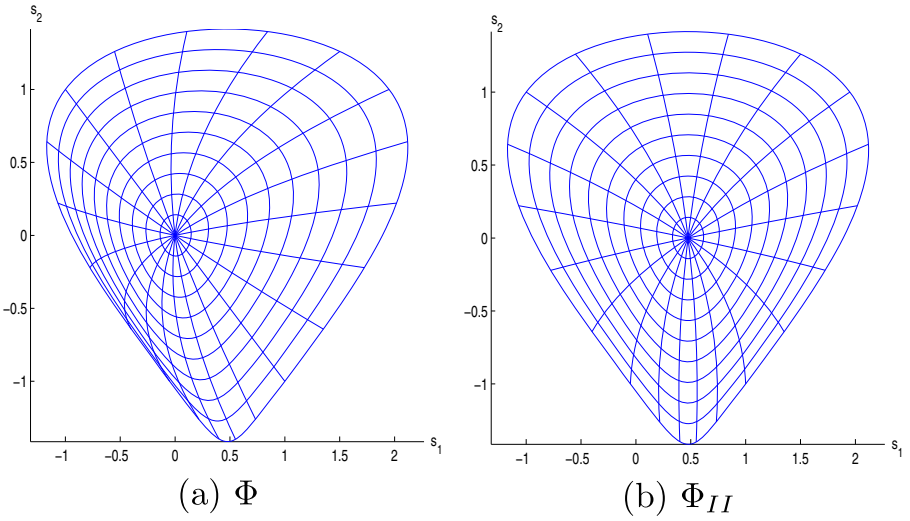

Fig. 4 Illustrations of mappings on $\mathbb{B}^2$ for the region Ω given by (3.4)

For the second example, we use the region Ω_2 defined by

$$\Phi_2(x) = [x_1, 4x_2, 2x_3]^T, \qquad x \in \overline{\mathbb{B}}^3,$$

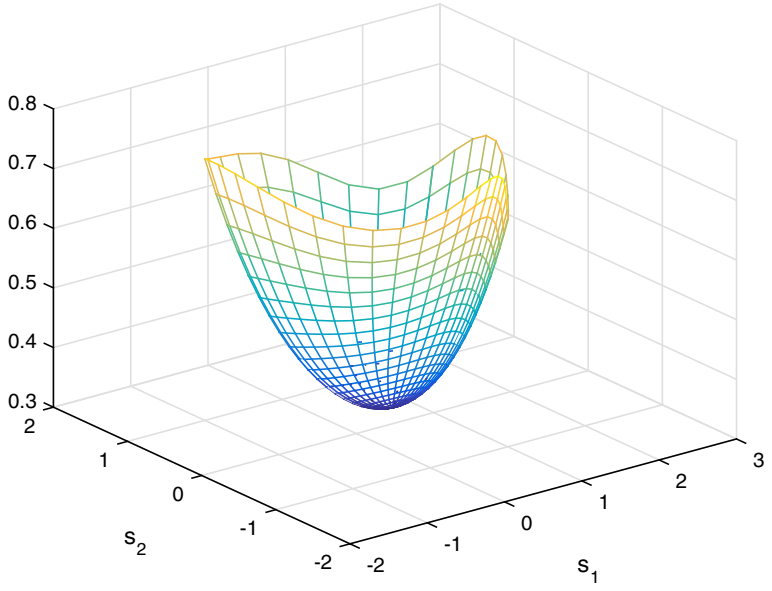

Fig. 5 Solution to (3.5)–(3.6)

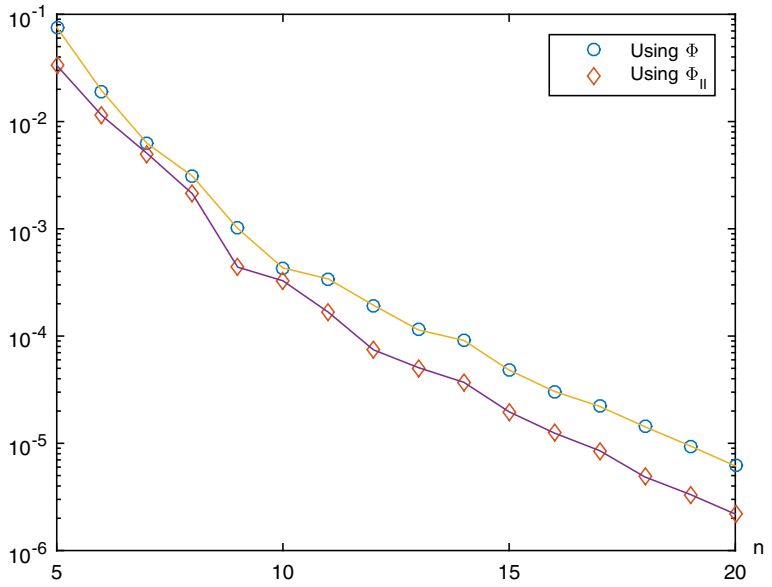

Fig. 6 Errors when solving (3.5)–(3.6) with mappings Φ and Φ_{II}

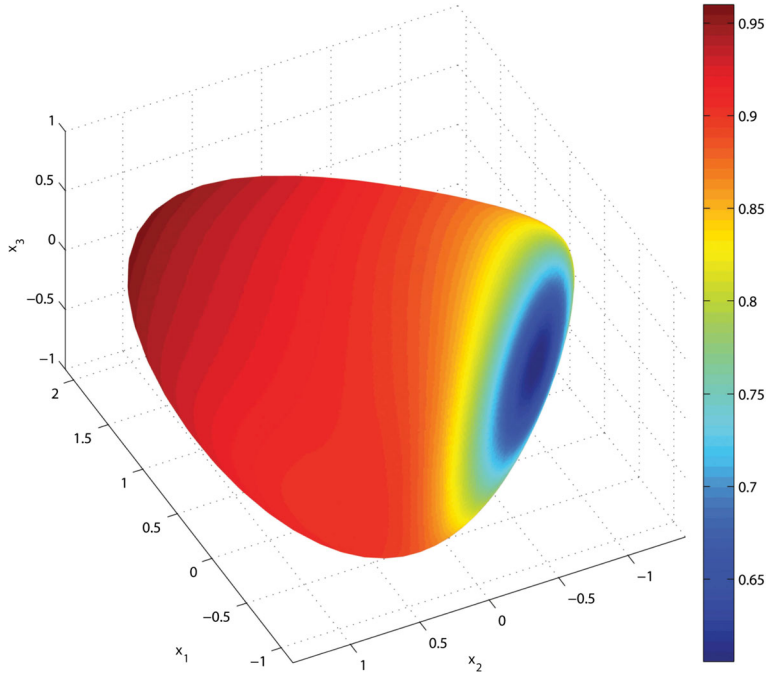

Fig. 7 Surface values of u_{30} for the numerical solution of (3.7)–(3.8)

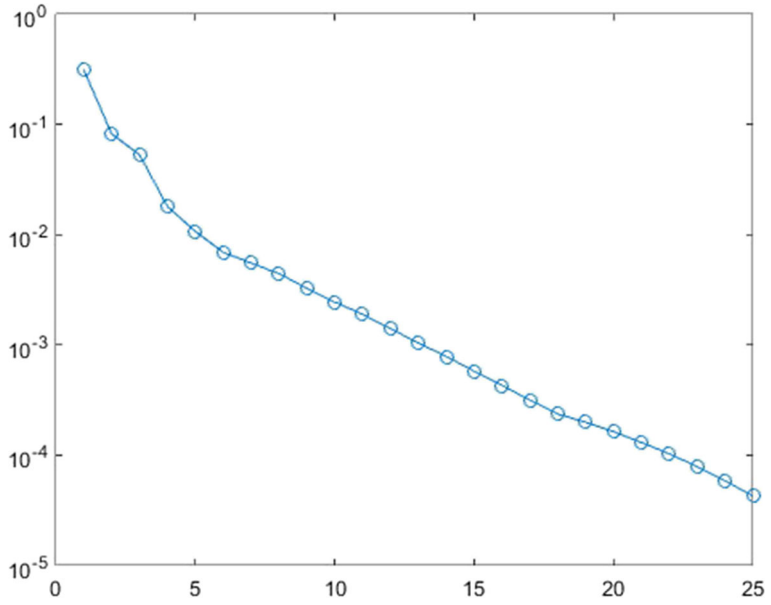

Fig. 8 Estimated errors when solving (3.7)–(3.8)

so the region Ω_2 is an ellipse with semi-axes of length 1, 4, and 2. The equation we solve is given by

$$-\Delta u(s) + 2u(s) = 0, \qquad s \in \Omega_2 \tag{3.9}$$

$$\frac{\partial u(s)}{\partial n_s} = -u^4(s) + \max\{0, (n_s)_{s_1}\}, \qquad s \in \partial\Omega_1 \tag{3.10}$$

The right-hand side is again a Stefan–Boltzmann term plus a source term given by

$$\max\{0, (n_s)_{s_1}\} = \max\{0, n_s \cdot [1, 0, 0]^T\}$$

Here, we take the normal component of an incoming radiation with constant direction $-[1, 0, 0]$ and the maximum ensures that surface side that is in the shadow will not receive a negative energy. This formula is correct, for example, if the region is convex. In the moment where the surface creates shadows on the side of the incoming radiation, these shadows have to be taken into account and the formula would be much more complicated, taking the geometry of $\partial\Omega$ into account. But even for a convex surface, the source term $\max\{0, n_s \cdot [1, 0, 0]^T\}$ is only once continuously differentiable and we expect that the solution $u(s)$ to (3.9)–(3.10) is only three times differentiable. The solution along the surface of Ω_2 is shown in Fig. 9. The speed of

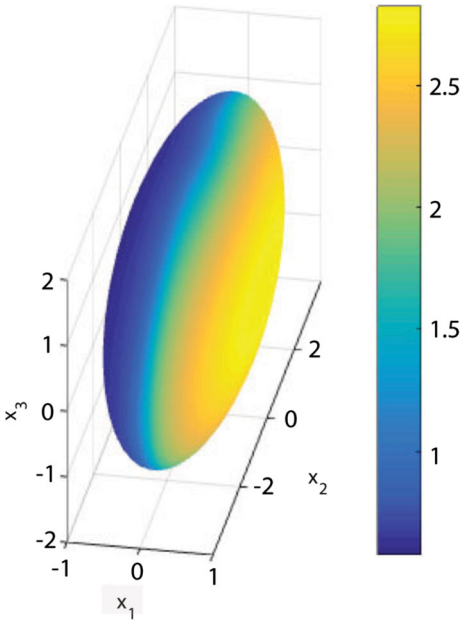

Fig. 9 Surface values of u_{30} for the numerical solution of (3.9)–(3.10)

convergence of our method should be reduced. This is confirmed in Fig. 10 where we use again $u_{30}(s)$ as a reference solution. Figure 10 does not show exponential convergence, but the logarithmic graph indicates a polynomial convergence.

4 Error analysis

The error analysis uses a reformulation of the nonlinear problem (1.1)–(1.2) as a nonlinear integral equation, as discussed earlier in the introduction. The error analysis makes limiting assumptions on the function b, as discussed later in this analysis, e.g., (4.27), while still covering many cases of interest. We begin with a review of notation and results on the linear Neumann boundary value problem.

4.1 The linear Neumann problem

Our numerical method is analyzed in the context of the Sobolev spaces $H^m(\Omega)$ and $H^{m-\frac{1}{2}}(\partial\Omega)$, with $m \geq 0$ an integer. We give a brief review of notation and results, all well-known in the literature. The spaces $H^m(\Omega)$ are the completion of $C^m\left(\overline{\Omega}\right)$ using the norm

$$\|v\|_{H^m(\Omega)} = \sqrt{\sum_{|\mathbf{i}|\leq m} \left\| \frac{\partial^{|\mathbf{i}|} v(s)}{\partial s_1^{i_1} \cdots \partial s_d^{i_d}} \right\|^2_{L^2(\Omega)}}$$

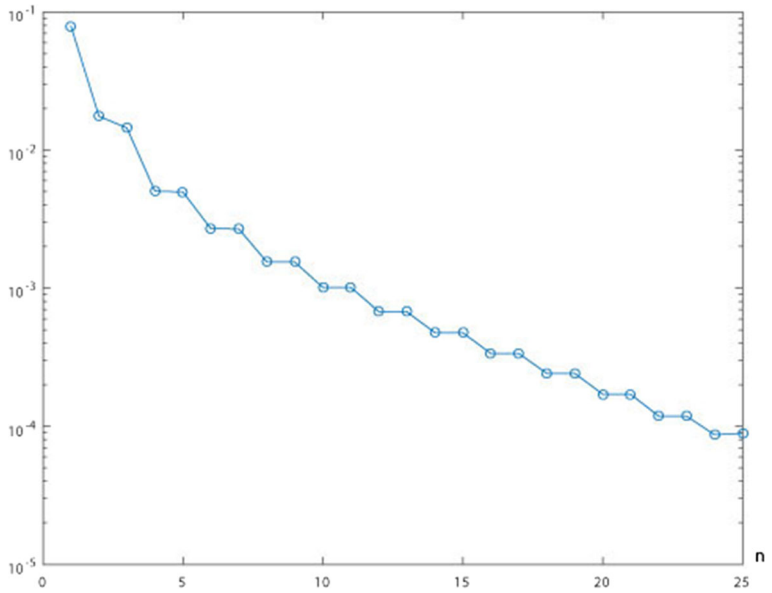

Fig. 10 Estimated errors when solving (3.9)–(3.10)

and $\mathbf{i} = (i_1, \ldots, i_d)$, $|\mathbf{i}| = i_1 + \cdots + i_d$. For $H^{m-\frac{1}{2}}(\partial\Omega)$, begin with $\Omega = \mathbb{B}^d$, noting $\partial\mathbb{B}^d = \mathbb{S}^{d-1}$. For $d = 2$ and $r \in \mathbb{R}$, $v \in H^r\left(\mathbb{S}^1\right)$ if

$$v(s) = \sum_{j=-\infty}^{\infty} a_j e^{ij\theta}, \qquad s = (\cos\theta, \sin\theta) \in \mathbb{S}^1,$$

$$\|v\|_{H^r(\mathbb{S}^1)} \equiv \sqrt{|a_0|^2 + \sum_{\substack{j=-\infty \\ j\neq 0}}^{\infty} |j|^{2r} \left|a_j\right|^2} < \infty.$$

For $d = 3$, begin by considering the spherical harmonics S_n^k of degree n and order k on $\mathbb{S}^2$ (see [5, §7.5.5]). For $r \in \mathbb{R}$, a function $v \in H^r\left(\mathbb{S}^2\right)$ if

$$v(s) = \sum_{n=0}^{\infty} \sum_{k=0}^{2n} a_{n,k} S_n^k(s), \qquad s \in \mathbb{S}^2,$$

$$\|v\|_{H^r(\mathbb{S}^2)} \equiv \sqrt{\sum_{n=0}^{\infty} (2n+1)^{2r} \sum_{k=1}^{2n+1} \left|a_{n,k}\right|^2} < \infty.$$

An analogous definition can be given for $d > 3$.

For the more general region Ω, we have $v \in H^r(\partial\Omega)$ if $\widetilde{v} \in H^r\left(\mathbb{S}^{d-1}\right)$, with $\widetilde{v} = v \circ \Phi$. The norm of v within $H^r(\partial\Omega)$ can be defined in a variety of ways, with

$$\|v\|_{H^r(\partial\Omega)} = \|\widetilde{v}\|_{H^r(\mathbb{S}^{d-1})}$$

being the simplest. In whatever way $\|v\|_{H^r(\partial\Omega)}$ is defined, we assume

$$c_{1,r}\,\|v\|_{H^r(\partial\Omega)} \leq \|\widetilde{v}\|_{H^r(\mathbb{S}^{d-1})} \leq c_{2,r}\,\|v\|_{H^r(\partial\Omega)} \tag{4.1}$$

for constants $c_{1,r}, c_{2,r} > 0$ independent of v.

Recalling the brief discussion of the reformulation (1.7), consider first the linear Neumann problem

$$-\Delta u + \gamma(s)u = 0, \qquad s \in \Omega, \tag{4.2}$$

$$\frac{\partial u(s)}{\partial n_s} = b\,(s)\,, \qquad s \in \partial\Omega. \tag{4.3}$$

with $b \in H^{-1/2}\,(\partial\Omega)$ (see [11, p. 338]). Recalling the definition (2.2) for $\mathcal{A}$, assume $\mathcal{A}$ is *strongly elliptic*:

$$|\mathcal{A}\,(v,v)| \geq c_e\,\|v\|^2_{H^1(\Omega)}\,, \qquad \forall v \in H^1\,(\Omega)\,, \tag{4.4}$$

for some $c_e > 0$. For simplicity, also assume $\gamma \in C^\infty\,(\Omega)$, along with the earlier assumptions following (1.4). It is straightforward to show $\mathcal{A}$ is bounded on $H^1\,(\Omega) \times H^1\,(\Omega)$,

$$|\mathcal{A}\,(v,w)| \leq c_{\mathcal{A}}\,\|v\|_{H^1}\,\|w\|_{H^1}\,, \qquad v, w \in H^1\,(\Omega) \tag{4.5}$$

for some $0 < c_{\mathcal{A}} < \infty$. Introduce the linear functional

$$\ell_b\,(v) \equiv (v,b)_{L^2} = \int_{\partial\Omega} v\,(s)\,b(s)\,ds, \qquad v \in H^{1/2}\,(\partial\Omega) \tag{4.6}$$

which is bounded over $H^{1/2}\,(\partial\Omega)$:

$$\|\ell_b\|_{H^{1/2}(\partial\Omega)\to\mathbb{R}} \leq c\,\|b\|_{H^{-1/2}(\partial\Omega)}$$

for some $c > 0$.

The weak form of the Neumann problem (4.2)–(4.3) is as follows. Let $b \in H^{-1/2}\,(\partial\Omega)$, and then find $u \in H^1\,(\Omega)$ such that

$$\mathcal{A}\,(u,v) = \ell_b(v), \qquad \forall v \in H^1\,(\Omega)\,. \tag{4.7}$$

The Lax-Milgram theorem (cf. [5, §8.3], [12, §2.7], [13, p. 8]) implies the unique existence of $u \in H^1\,(\Omega)$, with

$$\|u\|_{H^1(\Omega)} \leq \frac{1}{c_e}\|\ell_b\|_{H^{-1/2}(\partial\Omega)\to\mathbb{R}}.$$

The solution u can be written

$$u = \mathcal{G}b, \qquad b \in H^{-1/2}\,(\partial\Omega) \mapsto u \in H^1\,(\Omega)\,,$$

with $\mathcal{G} : H^{-1/2}\,(\partial\Omega) \to H^1\,(\Omega)$ a bounded operator (see [10, p. 308]).

More generally,

$$u = \mathcal{G}b, \qquad b \in H^{m-1/2}\,(\partial\Omega) \mapsto u \in H^{m+1}\,(\Omega)\,, \tag{4.8}$$

$m \geq 0$ an integer, and $\mathcal{G}$ is a bounded mapping (see [18, p. 129]). We are usually interested in the case $m = 1$. This formula is more commonly written in an integral form,

$$u\,(s) = \int_{\partial\Omega} G\,(s,t)\,b\,(t)\,dt, \qquad s \in \Omega, \tag{4.9}$$

as earlier in (1.6).This is the Green's integral representation for the solution of (4.2)–(4.3).

In addition, introduce the operator $\widehat{\mathcal{G}}$, a 'restriction' of $\mathcal{G}$:

$$\left(\widehat{\mathcal{G}}v\right)(s) = (\mathcal{G}v)(s), \quad s \in \partial\Omega, \quad v \in H^{-1/2}(\partial\Omega). \tag{4.10}$$

Since $\mathcal{G}v \in H^1(\Omega)$, the trace $(\mathcal{G}v) \in H^{1/2}(\partial\Omega)$, and moreover, the trace operator is a bounded mapping from $H^1(\Omega)$ onto $H^{1/2}(\partial\Omega)$.Thus $\widehat{\mathcal{G}}$ defines a (Neumann to Dirichlet) bounded mapping from $H^{-1/2}(\partial\Omega)$ into $H^{1/2}(\partial\Omega)$. Similarly, $\widehat{\mathcal{G}}$ is a bounded mapping from $H^{m-1/2}(\partial\Omega)$ into $H^{m+1/2}(\partial\Omega)$, $m \geq 1$. Since the unit ball in $H^{m+1/2}(\partial\Omega)$ is compact in $H^{m-1/2}(\partial\Omega)$, it follows that $\widehat{\mathcal{G}}$ is a compact mapping from $H^{m-1/2}(\partial\Omega)$ into $H^{m-1/2}(\partial\Omega)$, $m \geq 1$ an integer. The case of most interest here is $m = 1$.

The numerical solution of (4.2)–(4.3) is as follows. Find $u_n \in \mathcal{X}_n$ such that

$$\mathcal{A}(u_n, v) = \int_{\partial\Omega} b(t)\, v(t)\, dt, \qquad \forall v \in \mathcal{X}_n. \tag{4.11}$$

This has a unique solution $u_n \in \mathcal{X}_n$ by means of the same theory as was used for (4.7). We write

$$u_n(s) = \mathcal{G}_n b(s), \qquad s \in \Omega \tag{4.12}$$

with $\mathcal{G}_n : H^{1/2}(\partial\Omega) \to \mathcal{X}_n \subseteq H^1(\Omega)$. Define $\widehat{\mathcal{G}}_n$ in analogy with $\widehat{\mathcal{G}}$ in (4.10). This was present and analyzed earlier in [8], but a different approach to the error analysis is taken here.

Looking at the approach of Osborn [26, §4], we can show the following.

Theorem 4.1 *Let $b \in H^{1/2}(\partial\Omega)$. Then*

$$\|\mathcal{G}b - \mathcal{G}_n b\|_{H^1(\Omega)} \leq \frac{c}{n} \|b\|_{H^{1/2}(\partial\Omega)}. \tag{4.13}$$

Proof Let $u = \mathcal{G}b$ and note $u \in H^2(\Omega)$ from (4.8) with $m = 1$. Subtracting (4.11) from (4.7), it follows that

$$\mathcal{A}(u - u_n, v) = 0 \qquad \forall v \in \mathcal{X}_n. \tag{4.14}$$

From (4.4),

$$c_e \|u - u_n\|^2_{H^1(\Omega)} \leq \mathcal{A}(u - u_n, u - u_n). \tag{4.15}$$

Using (4.14),

$$\begin{aligned} \mathcal{A}(u - u_n, u - u_n) &= \mathcal{A}(u - u_n, u) \\ &= \mathcal{A}(u - u_n, u - \chi), \qquad \forall \chi \in \mathcal{X}_n. \end{aligned}$$

From (4.5),

$$\mathcal{A}(u - u_n, u - \chi) \leq c_{\mathcal{A}} \|u - u_n\|_{H^1} \|u - \chi\|_{H^1}.$$

Combining this with (4.15) and canceling $\|u - u_n\|_{H^1}$,

$$\|u - u_n\|_{H^1(\Omega)} \leq \frac{c_{\mathcal{A}}}{c_e} \|u - \chi\|_{H^1(\Omega)}, \qquad \forall \chi \in \mathcal{X}_n. \tag{4.16}$$

From the comment following Theorem 2.2 with $r = 2$, this implies

$$\begin{aligned} \|u - u_n\|_{H^1(\Omega)} &\leq \frac{c_1}{n} \|u\|_{H^2(\Omega)} \\ \|\mathcal{G}b - \mathcal{G}_n b\|_{H^1(\Omega)} &\leq \frac{c_1}{n} \|\mathcal{G}b\|_{H^2(\Omega)} \end{aligned} \tag{4.17}$$

for a constant $c > 0$.

In addition,

$$\|\mathcal{G}b\|_{H^2(\Omega)} \leq c_2 \|b\|_{H^{1/2}(\partial\Omega)} .$$

Combining this with (4.17) implies (4.13) for some constant $c > 0$. □

Corollary 4.2

$$\left\|\widehat{\mathcal{G}} - \widehat{\mathcal{G}}_n\right\|_{H^{1/2}(\partial\Omega)\to H^{1/2}(\partial\Omega)} \leq \frac{c}{n} \tag{4.18}$$

Proof From the trace theorem,

$$\left\|\widehat{\mathcal{G}}b - \widehat{\mathcal{G}}_n b\right\|_{H^{1/2}(\partial\Omega)} \leq c \|\mathcal{G}b - \mathcal{G}_n b\|_{H^1(\Omega)} ,$$

for some $c > 0$. Combining this with (4.13),

$$\left\|\widehat{\mathcal{G}}b - \widehat{\mathcal{G}}_n b\right\|_{H^{1/2}(\partial\Omega)} \leq \frac{c}{n} \|b\|_{H^{1/2}(\partial\Omega)} ,$$

proving (4.18) □

4.2 The nonlinear Neumann problem

Consider solving

$$-\Delta u + \gamma(s)u = 0, \qquad s \in \Omega, \tag{4.19}$$

$$\frac{\partial u(s)}{\partial n_s} = b\left(s, u\left(s\right)\right), \qquad s \in \partial\Omega, \tag{4.20}$$

with the nonlinear Neumann boundary condition (4.20). We begin by assuming $b \in C\left(\partial\Omega \times \mathbb{R}\right)$, with added assumptions given as we proceed (e.g., see (4.27) below). Denote the solution being sought by $u^* \in H^2\left(\Omega\right)$, assumed to be unique in some local neighborhood of u^*. An overview of the solvability theory for (4.19)–(4.20) was given earlier in the introduction.

Using the operator $\mathcal{G}$ and the Green's representation (4.9), the original solution u^* satisfies

$$u^*\left(s\right) = \int_{\partial\Omega} G\left(s,t\right) b\left(t, u^*\left(t\right)\right) dt, \qquad s \in \Omega. \tag{4.21}$$

Introduce the "Nemytskii operator" $\mathcal{N}$:

$$\left(\mathcal{N}\left(v\right)\right)\left(t\right) = b\left(t, v\left(t\right)\right), \quad t \in \partial\Omega, \quad v \in L^2\left(\partial\Omega\right).$$

(See [21, Chap. 1, §2], [32, §26.3] for its general properties). The equation (4.21) can be written as

$$u^* = \mathcal{G}\mathcal{N}\left(\text{trace}\left(u^*\right)\right) \tag{4.22}$$

with, as before, trace (u^*) denoting the restriction of u^* to $\partial\Omega$:

$$\text{trace} : H^m(\Omega) \overset{onto}{\longrightarrow} H^{m-1/2}(\partial\Omega), \qquad m \geq 1.$$

Note that knowing trace (u^*) over $\partial\Omega$ determines u^* completely over all of Ω using (4.21).

For the numerical method, following (4.11), find $u_n^* \in \mathcal{X}_n$ such that

$$\mathcal{A}\left(u_n^*, v\right) = \int_{\partial\Omega} b\left(t, u_n^*(t)\right) v(t)\, dt, \qquad \forall v \in \mathcal{X}_n. \tag{4.23}$$

Using (4.12), write

$$u_n^* = \mathcal{G}_n \mathcal{N}\left(\text{trace}\left(u_n^*\right)\right). \tag{4.24}$$

As noted before, $u_n^*(s)$ over Ω is determined from its values over $\partial\Omega$.

The error analysis is similar to that of [9], making use of the equations (4.22) and (4.24). This reformulation of a partial differential equation problem as an integral equation problem follows that of Osborn [26, §4.1]. We begin by analyzing the error $u^* - u_n^*$ over $\partial\Omega$, and then the error over Ω is analyzed.

The error analysis parallels that of [9]. It also was based on the construction of $\mathcal{G}$ and $\mathcal{G}_n$, although with different Green's integral operators. Introduce the notation

$$\begin{aligned} \widehat{u}^* &= \text{trace}\left(u^*\right), \\ \widehat{u}_n^* &= \text{trace}\left(u_n^*\right). \end{aligned}$$

Consider the equations

$$\widehat{u}^* = \widehat{\mathcal{G}}\mathcal{N}\left(\widehat{u}^*\right), \tag{4.25}$$

$$\widehat{u}_n^* = \widehat{\mathcal{G}}_n \mathcal{N}\left(\widehat{u}_n^*\right), \tag{4.26}$$

and consider them with respect to the Hilbert space $H^{1/2}(\partial\Omega)$. Equation (4.25) is a fixed point problem, and as noted earlier following (4.10), $\widehat{\mathcal{G}}$ is a compact operator from $H^{1/2}(\partial\Omega)$ into $H^{1/2}(\partial\Omega)$. This construction converts the numerical method (4.24) to a corresponding method for finding a fixed point for a completely continuous nonlinear integral operator, namely $\widehat{\mathcal{G}}\mathcal{N}$, and this latter numerical method can be analyzed using the results given in [21, Chap. 3] and [2].

It is necessary to make assumptions about the Nemytskii operator $\mathcal{N}$. For an open set $D \subseteq H^{1/2}(\partial\Omega)$ containing the solution $\widehat{u}^*$, assume

$$\begin{gathered} v \in D \implies b(\cdot, v) \in H^{1/2}(\partial\Omega), \\ B \subseteq D \text{ and bounded in } H^{1/2}(\partial\Omega) \Longrightarrow \mathcal{N}(B) \text{ bounded in } H^{1/2}(\partial\Omega) \\ v_n \to v \text{ in } H^{1/2}(\partial\Omega) \implies b(\cdot, v_n) \to b(\cdot, v) \text{ in } H^{1/2}(\partial\Omega). \end{gathered} \tag{4.27}$$

These are somewhat restrictive, limiting the generality of our error analysis. To illustrate this limitation for functions of one variable, if $b(\cdot, v) = v^2$ and if $v \in L^2(0,1)$ then $b(\cdot, v)$ may not belong to $L^2(0,1)$. The function $v(s) \equiv 1/\sqrt[3]{s}$ is in $L^2(0,1)$, whereas $v(s)^2 = 1/\sqrt[3]{s^2}$ does not belong to $L^2(0,1)$. The assumption (4.27) limits the growth behavior of $b(\cdot, v)$ as a function of v. Examples of acceptable cases are $b(s, v) = \exp\left(-v^2\right)$ and $\left(1 + v^2\right)^{-1}$.

An analysis of when (4.27) is true can be based on [23] . Assume the function b can be extended to $b \in C\left(\overline{\Omega} \times \mathbb{R}\right)$. Begin by noting that each function $v \in H^{1/2}(\partial\Omega)$ can be extended to a function $w \in H^1(\Omega)$, and in fact it can be written

$$w(s) = \int_{\partial\Omega} \Gamma(s,t)\, v(t)\, dt, \qquad s \in \Omega, \tag{4.28}$$

a Green's function formula. Such an extension w can be constructed as the solution to the Dirichlet problem

$$\Delta w(s) = 0, \qquad s \in \Omega, \tag{4.29}$$

$$w(s) = v(s), \qquad s \in \partial\Omega. \tag{4.30}$$

There is a unique solution $w(s)$ in $H^1(\Omega)$. Apply the result of Marcel and Mizel [23, Thm. 2.2] to w, examining $b(s, w(s))$. If $b(\cdot, w(\cdot))$ belongs to $H^1(\Omega)$, then $b(s, v(s)) \equiv \operatorname{trace} b(s, w(s))$ belongs to $H^{1/2}(\partial\Omega)$. Their result implies that (4.27) is satisfied if

$$\left|\frac{\partial b(s,v)}{\partial s_j}\right| \le a_1(s) + c\,|v|,$$

$$\left|\frac{\partial b(s,v)}{\partial v}\right| \le a_2(s),$$

for $s \in \Omega$, $v \in \mathbb{R}$, with constant c and functions $a_1, a_2 \in L^2(\Omega)$.

The mapping $\widehat{\mathcal{G}}\mathcal{N}$ is a compact nonlinear operator on an open domain D of the Banach space $\mathcal{X} = H^{1/2}(\partial\Omega)$. This follows from (*i*) the compactness of $\mathcal{G}$ from $H^{1/2}(\partial\Omega)$ to itself, and (*ii*) the boundedness of $\mathcal{N}$ on bounded subsets of $D \subseteq H^{1/2}(\partial\Omega)$.

Let $V \subseteq D$ be a bounded open set containing an isolated fixed point solution $\widehat{u}^*$ of the nonlinear integral equation (4.25). An important concept is that of the "index" of $\widehat{u}^*$ (see Krasnoselskii [21, p. 109]). An important property is for the index of $\widehat{u}^*$ to have a nonzero index. For some intuition as to stability implications of a fixed point having a nonzero index (see [2, Property P5, p. 802]). In essence, if small perturbations of the function $v - \mathcal{G}\mathcal{N}(v)$ leads to small perturbations in the fixed point $\widehat{u}^*$, then the isolated fixed point $\widehat{u}^*$ has a nonzero index.

Theorem 4.3 *Assume the function b satisfies (4.27). Assume the problem (4.25) has a solution $\widehat{u}^*$ that is unique within some open neighborhood $V \subseteq H^{1/2}(\partial\Omega)$ of $\widehat{u}^*$, and further assume that $\widehat{u}^*$ has nonzero index. Then for all sufficiently large n, (4.26) has one or more solutions $\widehat{u}_n^*$ within V, and all such $\widehat{u}_n^*$ converge to $\widehat{u}^*$ as $n \to \infty$.*

Proof This is an application of the methods of [21, Chap. 3, Sec. 3] or [2, Thm. 3]. A sufficient requirement is the norm convergence of $\mathcal{G}_n$ to $\mathcal{G}$, given in (4.18); [2, Thm. 3] uses a weaker form of (4.18). The assumption that a solution $\widehat{u}^*$ exists is linked to an original assumption that (4.19)–(4.20) has a unique solution in some open neighborhood V^* about u^*, $V^* \subseteq H^1(\Omega)$. □

The most standard case of a nonzero index involves a consideration of the Fréchet derivative of $\mathcal{N}$ (see [5, §5.3]). In particular, the linear operator $\mathcal{N}'(v)$ is given by

$$\left(\mathcal{N}'(v)\, w\right)(x) = \left.\frac{\partial b(x,z)}{\partial z}\right|_{z=v(x)} \times w(x) \tag{4.31}$$

Theorem 4.4 *Assume the function b satisfies* (4.27). *Assume the problem* (4.25) *has a solution* $\widehat{u}^*$ *that is unique within some open neighborhood V of* $\widehat{u}^*$*; and further assume that* $I - \widehat{\mathcal{G}}\mathcal{N}'(\widehat{u}^*)$ *is invertible over* $H^{1/2}(\partial\Omega)$. *Then* $\widehat{u}^*$ *has a nonzero index. Moreover, for all sufficiently large n there is a unique solution* $\widehat{u}_n^*$ *to* (4.26) *within V, and* $\widehat{u}_n^*$ *converges to* $\widehat{u}^*$ *with*

$$\begin{aligned}\left\|\widehat{u}^* - \widehat{u}_n^*\right\|_{H^{1/2}(\partial\Omega)} &\leq c\left\|\left(\widehat{\mathcal{G}} - \widehat{\mathcal{G}}_n\right)\mathcal{N}\left(\widehat{u}^*\right)\right\|_{H^{1/2}(\partial\Omega)} \\ &\leq \frac{c}{n}\left\|\mathcal{N}\left(\widehat{u}^*\right)\right\|_{H^{1/2}(\partial\Omega)}\end{aligned} \tag{4.32}$$

Proof Again, this is an immediate application of results in [21, Chap. 3, Sec. 3] or [2, Thm. 4]. □

The assumption that $I - \widehat{\mathcal{G}}\mathcal{N}'(\widehat{u}^*)$ is invertible is comparable to the standard assumption that the solution β for a one-variable fixed point problem

$$x = h(x)$$

satisfies $1 - h'(\beta) \neq 0$. (For further information, see [2, Prop. P4, P5]).

To improve on (4.32), we need to bound $\left\|\left(\widehat{\mathcal{G}} - \widehat{\mathcal{G}}_n\right)\widehat{g}\right\|_{H^{1/2}(\partial\Omega)}$ when given $\widehat{g} \in H^{m+1/2}(\partial\Omega)$ for some integer $m \geq 1$. This can be done by constructing a suitably smooth extension function over Ω. For example, construct an extension function $g \in H^{m+1}(\Omega)$ as the solution of

$$\begin{aligned}\Delta g(s) &= 0, \quad s \in \Omega, \\ g(s) &= \widehat{g}(s), \ s \in \partial\Omega.\end{aligned} \tag{4.33}$$

From Lemma 2.1 and Theorem 2.2, there exists a sequence of polynomials $p_n \in \Pi_n^d$ such that

$$\left\|g - p_n \circ \Phi^{-1}\right\|_{H^1(\Omega)} \leq \delta_{n,m}\|g\|_{H^{m+1}(\Omega)}, \qquad n \geq 1. \tag{4.34}$$

The sequence $\delta_{n,m} = O\left(n^{-m}\right)$ and it is independent of g.Then

$$\|\widehat{g} - \widehat{\chi}_n\|_{H^{1/2}(\partial\Omega)} \leq \varepsilon_{n,m}\|g\|_{H^{m+1}(\Omega)}, \qquad n \geq 1. \tag{4.35}$$

with $\chi_n = p_n \circ \Phi^{-1}$, $\widehat{\chi}_n = \text{trace}(\chi_n)$. The quantity $\|g\|_{H^{m+1}(\Omega)}$ can be bounded in terms of $\|\widehat{g}\|_{H^{m+1/2}(\partial\Omega)}$. To do this, note that the solution to (4.33) can be written as

$$g(s) = \int_{\partial\Omega} \Gamma(s,t)\,\widehat{g}(t)\,dt, \qquad s \in \overline{\Omega},$$

as was done above in (4.28)–(4.30). This is a bounded operator from $H^{m+1/2}(\partial\Omega)$ to $H^{m+1}(\Omega)$,

$$\|g\|_{H^{m+1}(\Omega)} \leq c_m\|\widehat{g}\|_{H^{m+1/2}(\partial\Omega)}$$

for some $c_m > 0$ (see [18, p. 129]). Using this leads to the following:

Corollary 4.5 *Assume* $\mathcal{N}(u^*) \in H^{m+1/2}(\partial\Omega)$, *for some* $m \geq 1$. *Then*

$$\left\|\widehat{u}^* - \widehat{u}_n^*\right\|_{H^{1/2}(\partial\Omega)} \leq \mathcal{O}\left(n^{-m}\right) \left\|\mathcal{N}\left(\widehat{u}^*\right)\right\|_{H^{m+1/2}(\partial\Omega)}. \tag{4.36}$$

4.2.1 The error over Ω

Let u^* and u_n^* denote the solutions of (4.24) and (4.22) over Ω that correspond to $\widehat{u}^*$ and $\widehat{u}_n^*$ respectively. Begin by subtracting (4.24) from (4.22):

$$\begin{aligned} u^* - u_n^* &= \mathcal{G}\mathcal{N}\left(\text{trace}\left(u^*\right)\right) - \mathcal{G}_n\mathcal{N}\left(\text{trace}\left(u_n^*\right)\right) \\ &= \left(\mathcal{G} - \mathcal{G}_n\right)\mathcal{N}\left(\widehat{u}^*\right) + \mathcal{G}_n\left[\mathcal{N}\left(\widehat{u}^*\right) - \mathcal{N}\left(\widehat{u}_n^*\right)\right] \end{aligned} \tag{4.37}$$

This can be used to prove convergence of u_n^* to u^* in $H^1(\Omega)$ by examining each of the two right-hand terms.

Note that (4.13) implies the family of operators $\mathcal{G}_n$ are uniformly bounded from $H^{1/2}(\partial\Omega)$ into $H^1(\partial\Omega)$,

$$\|\mathcal{G}_n\|_{H^{1/2}(\partial\Omega)\to H^1(\Omega)} \leq c_G, \qquad n \geq 1,$$

and thus

$$\begin{aligned} &\left\|\mathcal{G}_n\left[\mathcal{N}\left(\text{trace}\left(u^*\right)\right) - \mathcal{N}\left(\text{trace}\left(u_n^*\right)\right)\right]\right\|_{H^1(\Omega)} \\ &\qquad \leq c_G \left\|\mathcal{N}\left(\text{trace}\left(u^*\right)\right) - \mathcal{N}\left(\text{trace}\left(u_n^*\right)\right)\right\|_{H^{1/2}(\partial\Omega)}. \end{aligned} \tag{4.38}$$

Theorem 4.3 and the assumption (4.27) imply this converges to zero. The first term on the right side of (4.37), $\left(\mathcal{G} - \mathcal{G}_n\right)\mathcal{N}\left(\text{trace}\left(u^*\right)\right)$, also converges to zero, using Theorem 4.1. Thus $u_n^* \to u^*$ in $H^1(\Omega)$.

To improve this for smoother u^* requires further assumptions about $\mathcal{N}$. Begin by assuming $u^* \in H^{m+1}(\Omega)$, $m \geq 1$, corresponding to $\widehat{u}^* \in H^{m+1/2}(\partial\Omega)$. For the derivative operator $\mathcal{N}'(v)$ of (4.31), assume it is a well-defined linear operator over some open neighborhood U of trace (u^*), for every $v \in U \subseteq H^{1/2}(\partial\Omega)$; and further assume $\mathcal{N}'(v)$ is uniformly bounded over U. Then (4.38) implies

$$\begin{aligned} \left\|\mathcal{G}_n\left[\mathcal{N}\left(\text{trace}\left(u^*\right)\right) - \mathcal{N}\left(\text{trace}\left(u_n^*\right)\right)\right]\right\|_{H^1(\Omega)} &\leq c\left\|\text{trace}\left(u^*\right) - \text{trace}\left(u_n^*\right)\right\|_{H^{1/2}(\partial\Omega)} \\ &\leq \mathcal{O}\left(n^{-m}\right), \end{aligned}$$

using (4.36) and (4.27).

For the remaining term in (4.37), recall that $u^* \in H^{m+1}(\partial\Omega)$, and then recall the bound (4.16). Applying Theorem 2.2, we have

$$\left\|\left(\mathcal{G} - \mathcal{G}_n\right)u^*\right\|_{H^1(\Omega)} \leq cn^{-m}\left\|u^*\right\|_{H^{m+1}(\Omega)},$$

for some constant $c > 0$. Together with (4.38) and (4.37), this leads to a bound for $\left\|u^* - u_n^*\right\|_{H^1(\Omega)}$,

$$\left\|u^* - u_n^*\right\|_{H^1(\Omega)} = \mathcal{O}\left(n^{-m}\right), \qquad u^* \in H^{m+1}(\Omega),\ m \geq 1.$$

4.3 A nonhomogeneous boundary value problem

Consider solving the problem

$$-\Delta u(s) + \gamma(s)u(s) = f(s), \qquad s \in \Omega, \tag{4.39}$$

$$\frac{\partial u(s)}{\partial n_s} = b\left(s, u\left(s\right)\right), \qquad s \in \partial\Omega. \tag{4.40}$$

Break this apart into two problems. First solve

$$-\Delta v\left(s\right) + \gamma(s)v\left(s\right) = f\left(s\right), \qquad s \in \Omega, \tag{4.41}$$

$$\frac{\partial v(s)}{\partial n_s} = 0, \qquad s \in \partial\Omega. \tag{4.42}$$

This can be solved approximately using the methods given in [8].

Next, let the solution u be written as

$$u = v + w \tag{4.43}$$

Then w satisfies

$$-\Delta w\left(s\right) + \gamma(s)w\left(s\right) = 0, \qquad s \in \Omega, \tag{4.44}$$

$$\frac{\partial w(s)}{\partial n_s} = b\left(s, v\left(s\right) + w\left(s\right)\right), \qquad s \in \partial\Omega. \tag{4.45}$$

Solve this problem by the method of this paper.

5 An existence theorem for the three-dimensional Stefan–Boltzmann problem

In two of our numerical examples, (3.5)–(3.6) and (3.9)–(3.10), we used the right-hand side of the form u^4 that is motivated by the Stefan–Boltzmann law for radiation, [20]. In the following, we present an existence proof for the weak equation (2.1) for three-dimensional domains if the right-hand side includes a Stefan–Boltzmann type nonlinearity. But as we will note later, the proof for the two-dimensional case follows in a similar way (it requires actually a simpler function space). The proof follows the method used by Delfour, Payre, and Zolésios, [15]. In the following, we will allow a slightly more general second derivative operator than we have used in the previous sections, this accommodates anisotropic heat flows, but does not present any additional technical complications compared to the pure Laplace operator. A solution u of (2.1) is given as the unique minimum of a coercive functional G that we define in this section. (See [32] for the definition of a coercive functional). The proof will also show that this solution u to equation (2.1) is non-negative which corresponds to the interpretation that u represents the absolute temperature.

For a simple connected domain $\Omega \subset \mathbb{R}^3$ with smooth boundary $\partial\Omega$, we consider the boundary value problem

$$\sum_{i,j=1}^{3} \frac{\partial}{\partial s_i}\left(a_{i,j}(s)\frac{\partial}{\partial s_j}u(s)\right) + \gamma(s)u(s) = f(s), \quad s \in \Omega \tag{5.1}$$

$$\frac{\partial u}{\partial n_A}(s) = b(s)u^4(s) + c(s), \quad s \in \partial\Omega. \tag{5.2}$$

This problem describes a diffusive heat flow inside the body Ω with a heat source given by $f(s)$ and on the boundary, we have the emission of energy according

to the Stefan–Boltzmann law and incoming radiation given by $c(s)$. The conormal derivative $\partial u/\partial n_A$ is defined by

$$\frac{\partial u}{\partial n_A}(s) = \sum_{i,j=1}^{3} a_{i,j}(s)\frac{\partial u}{\partial s_j}\cos(\angle(e_i, n))$$

where n is the outward normal at $x \in \partial\Omega$ and e_i is the i. standard unit vector.

We first describe the assumption for the coefficients. We assume that $f, c, \gamma \geq 0$. For the coefficient b we assume strict positivity

$$b(s) \geq b_0 > 0, \qquad s \in \partial\Omega. \tag{5.3}$$

Furthermore, we assume that the matrix $A(s) := [a_{i,j}(s)]_{i,j=1,2,3}$ is symmetric and positive definite

$$\xi^T A(s)\xi \geq m\|\xi\|_2^2, \qquad \xi \in \mathbb{R}^3,\ s \in \Omega \tag{5.4}$$

where $m > 0$ and $\|\cdot\|_2$ denotes the Euclidean norm in $\mathbb{R}^3$. Finally, we assume that the coefficients $b(s)$, $c(s)$, $f(s)$, and $\gamma(s)$, are continuous functions and $a_{i,j}(s)$ is at least on time continuously differentiable on their respective domain.

To derive a weak formulation of the problem (5.1)–(5.2), we assume that we have a $C^2(\overline{\Omega})$ solution $u(s)$ and multiply equation (5.1) by $v(s) \in C^1(\overline{\Omega})$. Integration over Ω and using integrating by parts leads to

$$\int_\Omega \sum_{i,j=1}^{3} a_{i,j}(s)\frac{u(s)}{\partial s_j}\frac{\partial v(s)}{\partial s_i} + \gamma(s)u(s)v(s)\,ds - \int_{\partial\Omega}\frac{u(s)}{\partial n_A}v(s)\,dS$$
$$= \int_\Omega f(s)v(s)\,ds$$

Using equation (5.2) this leads to

$$\int_\Omega \sum_{i,j=1}^{3} a_{i,j}(s)\frac{u(s)}{\partial s_j}\frac{\partial v(s)}{\partial s_i} + \gamma(s)u(s)v(s)ds + \int_{\partial\Omega} b(s)u^4(s)v(s)dS$$
$$= \int_{\partial\Omega} c(s)v(s)\,dS + \int_\Omega f(s)v(s)\,ds \tag{5.5}$$

A weak solution $u \in H^1(\Omega)$ of (5.1)–(5.2) is a function such that (5.5) is true for all $v \in H^1(\Omega)$. The first problem is that in $\mathbb{R}^3$ we only have $H^1(\Omega) \hookrightarrow H^{1/2}(\partial\Omega) \hookrightarrow L^4(\partial\Omega)$, so we would need $v(s) \in L^\infty(\partial\Omega)$ to ensure that the integral

$$\int_{\partial\Omega} b(s)u^4(s)v(s)\,dS$$

exists, but this is not true for all $v \in H^1(\Omega)$. In the following, we will give another derivation of (5.5), based on the minimization of a functional F. Once we have found the right domain of F, the above problem will be solved too.

We start by introducing an appropriate function space:

$$X(\Omega) = \left\{u \in H^1(\Omega) \mid u|_{\partial\Omega} \in L^5(\partial\Omega)\right\} \tag{5.6}$$

This function space is necessary in the three-dimensional case for the functional F, defined further below, to be well defined. In two dimensions, we have $H^1(\Omega) \hookrightarrow H^{1/2}(\partial\Omega) \hookrightarrow L^p(\partial\Omega)$ for all $p \geq 1$. So in the two-dimensional case $X(\Omega) = H^1(\Omega)$. We can turn $X(\Omega)$ into a Banach space by using the norm

$$\|u\|_X := \|u\|_{H^*(\Omega)} + \|u|_{\partial\Omega}\|_{L^5(\partial\Omega)}$$

Here, the first term is given by

$$\|u\|_{H^*(\Omega)} := \left(\int_\Omega \sum_{i=1}^{3} \left(\frac{\partial u}{\partial s_i}(s) \right)^2 ds \right)^{1/2}$$

Note that the first term alone is not a norm on $H^1(\Omega)$. The usual norm on $H^1(\Omega)$ is given by

$$\|u\|_{H^1(\Omega)} := \left(\int_\Omega \sum_{i=1}^{3} \left(\frac{\partial u}{\partial s_i}(s) \right)^2 ds + \int_\Omega u^2(s)\, ds \right)^{1/2}$$

or by an equivalent norm like

$$\left(\int_\Omega \sum_{i=1}^{3} \left(\frac{\partial u}{\partial s_i}(s) \right)^2 ds + \int_{\partial\Omega} u^2(s)\, dS \right)^{1/2}$$

(see [24]). But by Hölder's inequality, we get

$$\int_{\partial\Omega} u^2(s)\, dS \leq \left(\int_{\partial\Omega} 1\, dS \right)^{3/5} \times \left(\int_{\partial\Omega} |u(s)|^5\, dS \right)^{2/5}$$

So there is a constant c (we will use c in the following for all constants, but the value might change from inequality to inequality) such that

$$\|u\|_{L^2(\partial\Omega)} \leq c\|u\|_{L^5(\partial\Omega)}$$

and therefore also

$$\|u\|_{H^1(\Omega)} \leq c\|u\|_X \tag{5.7}$$

The following lemma is proved in [15]

Lemma 5.1 *The normed vector space $X(\Omega)$ is a reflexive Banach space.*

On $X(\Omega)$ we define the functional

$$\begin{aligned} F(u) := \frac{1}{2} \int_\Omega \sum_{i,j=1}^{3} a_{i,j}(s) \frac{\partial u(s)}{\partial s_i} \frac{\partial u(s)}{s_j} + \gamma(s) u^2(s)\, ds \\ + \int_{\partial\Omega} \frac{b(s)}{5} |u(s)|^5\, dS \end{aligned} \tag{5.8}$$

This functional is well defined and continuous on $X(\Omega)$ and because of

$$\begin{aligned} f(t) &= \frac{1}{5}|t|^5, \quad t \in \mathbb{R} \\ f'(t) &= t\,|t|^3, \\ f''(t) &= 4\,|t|^3, \end{aligned}$$

we see that F is twice differentiable on $X(\Omega)$ with

$$\begin{aligned} DF(u)v &= \int_\Omega \sum_{i,j=1}^{3} a_{i,j}(s)\frac{\partial u(s)}{\partial s_i}\frac{\partial v(s)}{s_j} + \gamma(s)u(s)v(s)ds + \\ &\quad + \int_{\partial\Omega} b(s)u(s)|u(s)|^3 v(s)\, dS \qquad (5.9) \\ D^2F(u)(v,v) &= \int_\Omega \sum_{i,j=1}^{3} a_{i,j}(s)\frac{\partial v(s)}{\partial s_i}\frac{\partial v(s)}{s_j} + \gamma(s)v^2(s)ds + \\ &\quad + \int_{\partial\Omega} 4b(s)|u(s)|^3 v^2(s)\, dS \qquad (5.10) \end{aligned}$$

for $v \in X(\Omega)$. This together with the fact that $D^2F(u)(v,v) \geq 0$, for $v \in X(\Omega)$, implies the first result

Lemma 5.2 *The functional $F : X(\Omega) \mapsto \mathbb{R}$, defined in (5.8), is twice differentiable and convex.*

For functions $u, v \in X(\Omega)$ with $DF(u)(v,v) = 0$ we can say more.

Lemma 5.3 *If $u, v \in X(\Omega)$ and $DF(u)(v,v) = 0$ then v, is a constant function, and if $v(s) = k \neq 0$ then $u(s)|_{\partial\Omega} = 0$ for all $s \in \partial\Omega$.*

Proof If $DF(u)(v,v) = 0$ we get

$$\begin{aligned} 0 &= \int_\Omega \sum_{i,j=1}^{3} a_{i,j}(s)\frac{\partial v(s)}{\partial s_i}\frac{\partial v(s)}{s_j} + \gamma(s)v^2(s)\, ds \\ &\quad + \int_{\partial\Omega} 4b(s)|u(s)|^3 v^2(s)\, dS \;\Rightarrow \\ 0 &= \int_\Omega \sum_{i,j=1}^{3} a_{i,j}(s)\frac{\partial v(s)}{\partial s_i}\frac{\partial v(s)}{\partial s_j}\, ds \text{ and} \qquad (5.11) \\ 0 &= \int_{\partial\Omega} 4b(s)|u(s)|^3 v^2(s)\, dS. \qquad (5.12) \end{aligned}$$

From (5.11), we can conclude that $v(s) = k$, $s \in \Omega$, is a constant function, and from (5.12), we can conclude that $u(s)|_{\partial\Omega} = 0$ for $s \in \partial\Omega$, if $k \neq 0$. □

The remaining terms in equation (5.5) are gathered in the linear functional

$$L(v) := \int_\Omega f(s)v(s)\,ds + \int_{\partial\Omega} c(s)v(s)\,dS. \tag{5.13}$$

Because of $X(\Omega) \hookrightarrow L^2(\Omega)$ and $X(\Omega) \hookrightarrow L^2(\partial\Omega)$, the functional L is a linear and continuous functional on $X(\Omega)$. Equation (5.5) can now be reformulated as: Find $u \in X(\Omega)$ such that

$$DF(u) = L \tag{5.14}$$

an equation in $X(\Omega)^*$. Every minimum of the functional

$$G(u) := F(u) - L(u) \tag{5.15}$$

is a solution to (5.14). To show (5.15) has at least one solution, we need one more property. By our assumptions (5.3) and (5.4), we know there is a constant c such that

$$F(u) \geq c(\|u\|^2_{H^*(\Omega)} + \|u\|^5_{L^5(\partial\Omega)})$$

Our next goal is to show that

$$F(u) \geq c_1\|u\|^2_{X(\Omega)} \equiv c_1(\|u\|_{H^*(\Omega)} + \|u\|_{L^5(\partial\Omega)})^2 \tag{5.16}$$

if $\|u\|_{X(\Omega)} \geq 1$, for some suitable constant c_1. This will show that F is a coercive functional on X. We define the function

$$f(x, y) := \frac{x^2 + y^5}{(x + y)^2}$$

on the domain $D := \{(x, y) \mid x, y \geq 0,\ x+y \geq 1\}$. If we can show that $f(x, y) \geq c_1$ for all $(x, y) \in D$, then we found a suitable constant c_1. For this reason let $(\overline{x}, \overline{y}) \in D$ with $\overline{x} + \overline{y} = 1$. Then

$$\begin{aligned} f(t\overline{x}, t\overline{y}) &= \frac{t^2\overline{x}^2 + t^5\overline{y}^5}{t^2(\overline{x} + \overline{y})^2} \\ &= \overline{x}^2 + t^3\overline{y}^5 \\ &\geq \overline{x}^2 + \overline{y}^5 \end{aligned}$$

for $t \geq 1$. So we only need to consider the minimum of the function f on the set $(x, y) \in D, x + y = 1$, which implies $y = 1 - x$. Here, we have

$$\begin{aligned} f(x, 1 - x) &= x^2 + (1 - x)^5 \\ &\geq c_1 > 0 \end{aligned}$$

with some suitable constant c_1, because the polynomial $x^2 + (1 - x)^5$ has obviously no zeros for $x \in [0, 1]$.

For the functional L, we get

$$\begin{aligned} |L(u)| &\leq |\int_\Omega f(x)u(x) : dx| + |\int_{\partial\Omega} c(x)u(x) : dS| \\ &\leq \|f\|_{L^2(\Omega)}\|u\|_{L^2(\Omega)} + \|c\|_{L^2(\partial\Omega)}\|u\|_{L^2(\partial\Omega)} \\ &\leq c_2\|u\|_{X(\Omega)} \end{aligned} \tag{5.17}$$

where we used again the continuous imbeddings $X(\Omega) \hookrightarrow L^2(\Omega)$ and $X(\Omega) \hookrightarrow L^5(\partial\Omega) \hookrightarrow L^2(\partial\Omega)$.

The inequalities (5.16) and (5.17) show now

$$\begin{aligned} G(u) &\geq F(u) - |L(u)| \\ &\geq c_1 \|u\|^2_{X(\Omega)} - c_2 \|u\|_{X(\Omega)} \\ &= c_1 \|u\|^2_{X(\Omega)} \left(1 - \frac{c_2/c_1}{\|u\|_{X(\Omega)}}\right) \end{aligned} \tag{5.18}$$

for all $u \in X(\Omega)$ with $\|u\|_{X(\Omega)} \geq 1$. This implies

$$\lim_{\|u\|_{X(\Omega)} \to \infty} G(u) = \infty \tag{5.19}$$

and proves the next result.

Lemma 5.4 *The functional $G : X(\Omega) \mapsto \mathbb{R}$, defined in (5.15), is coercive.*

Using Lemmata 5.1, 5.2, and 5.4 the Corollary 42.14 in [33] shows

Lemma 5.5 *The functional $G(u)$ (see [5.15]) has at least one minimum and the set of all minimum points is closed, bounded, and convex.*

As a next step we strengthen Lemma 5.5:

Lemma 5.6 *The functional $G(u)$ (see [5.15]) has a unique minimum.*

Proof Assume that the set M of all minimum points of $G(u)$ has two distinct elements $u, v \in M$. Then the function

$$\varphi(t) := G(u + t(v - u)),\ t \in [0, 1]$$

is a constant function according to Lemma 5.5. This implies

$$\begin{aligned} 0 &= \varphi''(0) \\ &= D^2 G(u)(v - u, v - u) \\ &= D^2 F(u)(v - u, v - u) \end{aligned}$$

because the functional L is linear. Using Lemma 5.3, we see that $v = u + k$, where k is a constant. Because of $u \neq v$, we know that $k \neq 0$, so again by Lemma 5.3, we can conclude that $u|_{\partial\Omega} = 0$. But now, we can reverse the roles of u and v and consider the functional

$$\psi(t) := G(v + t(u - v)),\ t \in [0, 1]$$

and the same arguments show $v|_{\partial\Omega} = 0$. But this makes $v = u + k$, with a nonzero constant k impossible. So the assumption that M has two distinct elements leads to a contradiction. This proves the uniqueness. □

Finally, we prove that the unique minimum is a positive function.

Theorem 5.7 *The functional $G(u)$ (see [5.15]) has a unique positive minimum u. The function u is a solution of (5.5).*

Proof Let $u(s)$ be the unique minimum of G, then $u \in H^1(\Omega)$ and therefore $v := |u| \in H^1(\Omega)$ (see [29]). Obviously also $v|_{\partial\Omega} \in L^5(\partial\Omega)$, so $v \in X(\Omega)$. Looking at the definition of F, we see $F(u) = F(v)$, and because the functions f and c in (5.1) and (5.2) are assumed to be non-negative, we also have (see [5.13])

$$L(v) \geq L(u)$$

which implies

$$G(v) \leq G(u)$$

and because u is the unique minimum, we conclude $u = v$.

Because of $u(s) \geq 0$, the function $u(s)$ satisfies

$$\begin{aligned} &\int_\Omega \sum_{i,j=1}^{3} a_{i,j}(s) \frac{u(s)}{\partial s_j} \frac{\partial v(s)}{\partial s_i} + \gamma(s) u(s) v(s)\, ds \\ &= \int_{\partial\Omega} -b(s) u^4(s) v(s) + c(s) v(s)\, dS + \int_\Omega f(s) v(s)\, ds \\ &= 0 \end{aligned}$$

for all $v \in X(\Omega)^*$, and the absolute value which appeared in equation (5.9) is unnecessary. □

The above theorem does not prove the uniqueness of the solution to (2.1), but we know that there is a positive solution. Because we know that the solution minimizes the functional G, one might also try to use minimization methods in combination with spectral methods to approximate the solution u. This will be the subject of further investigations.

Acknowledgments For helpful discussions, we thank Weimin Han and Gerhard Strohmer.

Appendix

Defining surface normals and Jacobian for a general surface. This is well-known in the literature, but we include it for convenience. For notational simplicity in the mapping $\Phi : \overline{\mathbb{B}}^d \xrightarrow[onto]{1-1} \overline{\Omega}$, let x be replaced by (x, y, z), and s be replaced by (s, t, u). Write

$$\begin{aligned} s &= s(x, y, z) \\ t &= t(x, y, z) \\ u &= u(x, y, z) \end{aligned}$$

For derivatives, use the shorthand notation

$$ds_1 = \frac{\partial s(x, y, z)}{\partial x} \equiv \frac{\partial \Phi_1(x, y, z)}{\partial x}, \quad ds_2 = \frac{\partial s(x, y, z)}{\partial y}, \quad ds_3 = \frac{\partial s(x, y, z)}{\partial z}$$

with similar notation for t and u.

For the surface Jacobian $\left|J_{bdy}(x,y,z)\right|$ used in the change of variables expression (2.11),

$$\left|J_{bdy}(x,y,z)\right|^2 = \begin{vmatrix} x & y & z \\ dt_1 & dt_2 & dt_3 \\ du_1 & du_2 & du_3 \end{vmatrix}^2 + \begin{vmatrix} ds_1 & ds_2 & ds_3 \\ x & y & z \\ du_1 & du_2 & du_3 \end{vmatrix}^2 + \begin{vmatrix} ds_1 & ds_2 & ds_3 \\ dt_1 & dt_2 & dt_3 \\ x & y & z \end{vmatrix}^2$$

The normal at $(s,t,u) = \Phi(x,y,z)$, call it $\mathbf{N}(s,t,u)$, is given by

$$\mathbf{N} = \frac{\mathbf{G}}{\|\mathbf{G}\|},$$

$$\mathbf{G} = \begin{bmatrix} (dt_1 du_2 - dt_2 du_1)\, z + (dt_3 du_1 - dt_1 du_3)\, y + (dt_2 du_3 - dt_3 du_2)\, x \\ (du_1 ds_2 - du_2 ds_1)\, z + (du_3 ds_1 - du_1 ds_3)\, y + (du_2 ds_3 - du_3 ds_2)\, x \\ (ds_1 dt_2 - ds_2 dt_1)\, z + (ds_3 dt_1 - ds_1 dt_3)\, y + (ds_2 dt_3 - ds_3 dt_2)\, x \end{bmatrix}.$$

As an example, consider the ellipsoidal mapping

$$\Phi(x,y,z) = (ax, by, cz), \qquad (x,y,z) \in \overline{\mathbb{B}}^3.$$

Then

$$\left|J_{bdy}(x,y,z)\right|^2 = (bcx)^2 + (acy)^2 + (abz)^2, \qquad (x,y,z) \in \mathbb{S}^2.$$

References

1. Amann, H.: Nonlinear elliptic equations with nonlinear boundary conditions. In: Eckhaus, W. (ed.) New Developments in Differential Equations, pp. 43–63. North-Holland (1976)
2. Atkinson, K.: The numerical evaluation of fixed points for completely continuous operators. SIAM J. Num. Anal. **10**, 799–807 (1973)
3. Atkinson, K., Chien, D., Hansen, O.: A spectral method for elliptic equations: the Dirichlet problem. Adv. Comput. Math. **33**, 169–189 (2010)
4. Atkinson, K., Chien, D., Hansen, O.: Evaluating polynomials over the unit disk and the unit ball. Numerical Algorithms **67**, 691–711 (2014)
5. Atkinson, K., Han, W.: Theoretical Numerical Analysis: A Functional Analysis Framework. 3rd edn. Springer, Berlin (2009)
6. Atkinson, K., Han, W.: An Introduction to Spherical Harmonics and Approximations on the Unit Sphere. Springer, Berlin (2012)
7. Atkinson, K., Hansen, O.: Creating domain mappings. Electron. Trans. Numer. Anal. **39**, 202–230 (2012)
8. Atkinson, K., Hansen, O., Chien, D.: A spectral method for elliptic equations: the Neumann problem. Adv. Comput. Math. **34**, 295–317 (2011)
9. Atkinson, K., Chien, D., Hansen, O.: A spectral method for nonlinear elliptic equations. Numerical Algorithms **74**, 797–819 (2017)
10. Aubin, J.-P.: Approximation of Elliptic Boundary-Value Problems. Wiley-Interscience, Hoboken (1972)
11. Aubin, J.-P.: Applied Functional Analysis. 2nd edn. Wiley-Interscience, Hoboken (2000)
12. Brenner, S., Scott, L.R.: The Mathematical Theory of Finite Element Methods. Springer, New York (1994)
13. Ciarlet, P.: The Finite Element Method for Elliptic Problems. North-Holland (1978)
14. Dai, F., Xu, Y.: Approximation Theory and Harmonic Analysis on Spheres and Balls. Springer, Berlin (2013)
15. Delfour, M., Payre, G., Zolésios, J.-P.: Approximation of nonlinear problems associated with radiating bodies in space. SIAM J. Numerical Anal. **24**, 1077–1094 (1987)

16. Dunkl, C., Xu, Y.: Orthogonal Polynomials of Several Variables. Cambridge University Press, Cambridge (2001)
17. Garabedian, P.: Partial Differential Equations. Wiley, Hoboken (1964)
18. Grisvard, P.: Elliptic Problems in Nonsmooth Domains. Pitman Pub (1985)
19. Li, H., Xu, Y.: Spectral approximation on the the unit ball. SIAM J. Num. Anal. **52**, 2647–2675 (2014)
20. Liou, K.-N.: An Introduction to Atmospheric Radiation. Academic Press, Cambridge (1980)
21. Krasnoselskii, M.: Topological Methods in the Theory of Nonlinear Integral Equations. Pergamon press, Oxford (1964)
22. Krasnoselskii, M.: Positive Solutions of Operator Equations. Noordhoff (1964)
23. Marcus, M., Mizel, V.: Absolute continuity on tracks and mappings of Sobolev spaces. Arch. Rat. Mech. Anal. **45**, 294–320 (1972)
24. Maz'ya, V.: Sobolev Spaces. Springer, Berlin (1985)
25. Mikhlin, S.: Mathematical Physics, an Advanced Course. North-Holland (1970)
26. Osborn, J.: Spectral approximation for compact operators. Math. Comput. **29**, 712–725 (1975)
27. Roach, G.: Green's Functions. 2nd edn. Cambridge University Press, Cambridge (1982)
28. Stroud, A.: Approximate Calculation of Multiple Integrals. Prentice-Hall Inc., Eaglewood Cliffs (1971)
29. Treves, F.: Topological Vector Spaces, Distributions, and Kernels. Dover Publications, Mineola (2006)
30. Xu, Y.: Lecture notes on orthogonal polynomials of several variables. In: Advances in the Theory of Special Functions and Orthogonal Polynomials, pp. 135–188. Nova Science Publishers (2004)
31. Zeidler, E.: Nonlinear Functional Analysis and its Applications: I. Springer, Berlin (1986)
32. Zeidler, E.: Nonlinear Functional Analysis and its Applications: II/B. Springer, Berlin (1990)
33. Zeidler, E.: Nonlinear Functional Analysis and its Applications III. Springer, Berlin (1985)