Rotation invariant patterns for a nonlinear Laplace-Beltrami equation: a Taylor-Chebyshev series approach

Jan Bouwe van den Berg * Gabriel William Duchesne † Jean-Philippe Lessard ‡

January 10, 2022

Abstract

In this paper, we introduce a rigorous computational approach to prove existence of rotation invariant patterns for a nonlinear Laplace-Beltrami equation posed on the 2-sphere. After changing to spherical coordinates, the problem becomes a singular second order boundary value problem (BVP) on the interval $(0, \frac{\pi}{2}]$ with a *removable* singularity at zero. The singularity is removed by solving the equation with Taylor series on $(0, \delta]$ (with δ small) while a Chebyshev series expansion is used to solve the problem on $[\delta, \frac{\pi}{2}]$. The two setups are incorporated in a larger zero-finding problem of the form $F(a) = 0$ with a containing the coefficients of the Taylor and Chebyshev series. The problem $F = 0$ is solved rigorously using a Newton-Kantorovich argument.

Keywords

Rotation invariant patterns · elliptic PDEs on manifolds · computer-assisted proofs
Chebyshev series · Taylor series · contraction mapping theorem

1 Introduction

We consider the semi-linear elliptic partial differential equation (PDE)

$$\Delta u + \lambda u + f(u) = 0 \tag{1}$$

where Δ is the Laplace-Beltrami operator on a given smooth manifold, $f(u)$ is a nonlinearity and λ is a positive parameter. This PDE describes a classical nonlinear elliptic problem [15, 5], which can be studied for a wide range of nonlinearities. It has close connections to questions in differential geometry, as the Yamabe problem is described by a special case of (1), see for example [4]. In this paper, we study the quadratic case $f(u) = u^2$ on the unit sphere $S^2 \subset \mathbb{R}^3$ as considered for instance in [6, 26].

Using the spherical coordinates $(x, y, z) = (\cos\phi \sin\theta, \sin\phi \sin\theta, \cos\theta)$ leads to

$$\Delta u = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta \frac{\partial u}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 u}{\partial\phi^2}, \tag{2}$$

where $u = u(\theta, \phi)$. We look for a specific type of *rotation invariant* solutions, namely solutions of (1) that are radially symmetric around the z-axis (u is constant in ϕ: $u(\theta, \phi) = u(\theta)$) and symmetric in the equator (hence $\frac{\partial u}{\partial\theta}(\frac{\pi}{2}) = 0$). By restricting the PDE to this class of rotation invariant solutions, the second term in the right-hand side of (2) vanishes and the problem is reduced to the following singular second order non-autonomous boundary value problem

$$\begin{cases} u''(\theta) + \cot(\theta)u'(\theta) + \lambda u(\theta) + u(\theta)^2 = 0, & \text{for } \theta \in (0, \frac{\pi}{2}], \\ u'(0) = u'(\frac{\pi}{2}) = 0. \end{cases} \tag{3}$$

*VU Amsterdam, Department of Mathematics, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands. (janbouwe@few.vu.nl)

†McGill University, Department of Mathematics and Statistics, 805 Sherbrooke Street West, Montreal, QC, H3A 0B9, Canada. (gabriel.duchesne@mail.mcgill.ca)

‡McGill University, Department of Mathematics and Statistics, 805 Sherbrooke Street West, Montreal, QC, H3A 0B9, Canada. (jp.lessard@mcgill.ca)

In [2], a few solutions of (3) were proven to exist using the tools of computer-assisted proofs using a single Taylor series expansion in θ about zero. However, this approach has clear limitations: the number of Taylor coefficients needed to represent the solutions on $(0, \pi/2]$ quickly grows (as λ varies) and the radius of convergence becomes smaller than $\pi/2$. For these reasons, a single Taylor series is just not adequate to represent the solutions. In the current paper, we address this issue and extend significantly the results presented in [2]. The idea is still to use a Taylor series expansion of the solution, but now only on a short interval $(0, \delta]$ (with $\delta < \frac{\pi}{2}$ small). Simultaneously, the ODE in (3) is solved on $[\delta, \frac{\pi}{2}]$ with Chebyshev series expansions. The two setups are then incorporated in a larger zero-finding problem of the form $F(a) = 0$ (see Section 2) with a containing the coefficients of the Taylor and Chebyshev series. The problem $F = 0$ incorporates the boundary conditions $u'(0) = u'(\frac{\pi}{2}) = 0$ and a continuity condition at the *interface* $\theta = \delta$. A Newton-Kantorovich theorem (see Theorem 2.1) is then used to demonstrate that exact solutions of $F = 0$ exist close to numerical approximations.

The advantage of this *two-steps* approach is twofold. First, as in [2], the Taylor series expansion of the solution about $\theta = 0$ combined with the boundary condition $u'(0) = 0$ leads to a set-up which gets rid of the (removable) singularity in the term $\cot(\theta)u'(\theta)$. Second, for any $\delta > 0$, the coefficients of the Chebyshev series expansion of the solution of the differential equation (3) on $[\delta, \pi/2]$ has exponential decay, and therefore always converges. That implies that theoretically, this approach is always going to work, as long as the parameter δ is taken small enough so that the Taylor series expansion of the solution converges on $[0, \delta]$. Of course, there are always the standard computational limitations (e.g. the finite dimensional projection size cannot be too large), but there are none theoretically.

Sample results of patterns proven to exist using the approach of the present paper can be found in Figure 1.

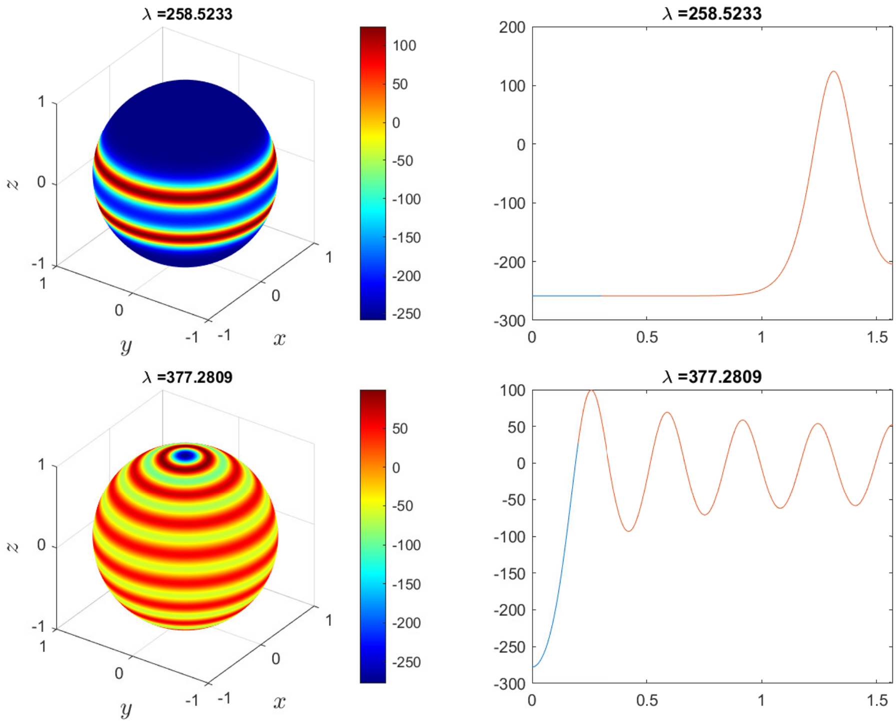

Figure 1: (Left) Two rigorously computed solutions of (1) on the unit sphere $S^2 \subset \mathbb{R}^3$. (Right) The corresponding solution of the BVP (3) with Taylor expansion in blue and Chebyshev expansion in orange.

The results presented in the present paper fall in the domain of computer-assisted proofs (CAPs) for PDEs. Before proceeding any further, it is worth mentioning that CAPs in nonlinear analysis began appearing long before the applications to PDEs. To name a few, in the 1960s, functional analytic methods of computer-assisted proof already existed exemplified by the work of Cesari on Galerkin projections for periodic solutions [7, 8]. In the field of dynamical systems, important open problems were settled with computer-assisted proofs, e.g. the universality of the Feigenbaum constant [12] and the existence of the strange attractor in the Lorenz system [21] (i.e. Smale's 14th problem). Other prominent examples outside dynamics are the proofs of the four-colour theorem [19] and Kepler's densest sphere packing problem [9].

We refer the interested to reader to the expository works [11, 16, 17, 18, 20, 22, 25] and the references therein, for a more complete overview of the field of rigorously verified numerics. Computer-assisted proofs more closely related to the present work concerned existence of radially symmetric solutions to the perturbed Gelfand problem [27], radially symmetric localized solution in a Ginzburg-Landau problem [24] and non-radial solutions for some semilinear elliptic equations on the disk [1].

The paper is organized as follows. In Section 2, we introduce the zero finding problem $F(a) = 0$ whose solutions correspond to solution to the BVP (3). Then, we introduce a Newton-Kantorovich theorem (Theorem 2.1) which is used to prove existence of solutions of $F = 0$. In Section 3, we present the explicit bounds required to apply Theorem 2.1. In Section 4, we introduce a bifurcation analysis to compute branches of solutions (parameterized by λ) bifurcating from the trivial solution $u = 0$. Then we show how to perform a continuation to numerically compute branches of solutions, which can then be validated using Theorem 2.1 using the bounds presented in Section 3.

All the codes necessary to perform the computer-assisted proofs are available at [23].

2 Definition of the zero finding problem

In this section, we introduce a zero finding problem of the form $F(a) = 0$ whose solutions correspond to solution to the BVP (3). As mentioned in Section 1, the second order ODE is solved on $(0, \delta]$ using Taylor series and on $[\delta, \frac{\pi}{2}]$ using Chebyshev series. In order to control numerical instabilities of the coefficients of the Taylor series (whose growth or decay depend on the radius of convergence), we rescale the interval $(0, \delta]$ to $(0, 1]$ in order to obtain a geometric decay of the Taylor coefficients. This stabilizes the numerical computations. Moreover, since Chebyshev series represent functions defined on $[-1, 1]$, we map the interval $[\delta, \frac{\pi}{2}]$ to $[-1, 1]$. Hence, letting $u_0(t) \stackrel{\text{def}}{=} u(\delta t)$ for $t \in (0, 1]$ and $u_1(t) \stackrel{\text{def}}{=} u\left(Kt + \frac{\pi}{4} + \frac{\delta}{2}\right)$ for $t \in [-1, 1]$ with $K \stackrel{\text{def}}{=} \frac{\pi}{4} - \frac{\delta}{2}$, the BVP (3) becomes

$$\begin{cases} u_0''(t) + \delta \cot(\delta t) u_0'(t) + \delta^2(\lambda u_0(t) + u_0(t)^2) = 0, & \text{for } t \in (0, 1], \\ u_1''(t) + K \cot\left(Kt + \frac{\pi}{4} + \frac{\delta}{2}\right) u_1'(t) + K^2(\lambda u_1(t) + u_1(t)^2) = 0, & \text{for } t \in [-1, 1], \\ u_0'(0) = 0, \quad u_1'(1) = 0, \end{cases} \tag{4}$$

together with matching conditions for u_0 at $t = 1$ and u_1 at $t = -1$. In order to use Chebyshev expansions for the u_1 equation in (4) for $t \in [-1, 1]$, we convert the second order equation into a first order system and we use the method introduced in [14]. Denote $u_2(t) \stackrel{\text{def}}{=} u_1'(t)$ and denote the non-autonomous term in the u_1 equation by

$$u_3(t) \stackrel{\text{def}}{=} \cot\left(Kt + \frac{\pi}{4} + \frac{\delta}{2}\right).$$

Note that $u_3(-1) = \cot(\delta)$. Rather than computing directly the Chebyshev coefficients of the expansion of the analytic function $u_3 : [-1, 1] \to \mathbb{R}$, we append a simple polynomial differential equation whose solution is given by u_3, namely $u_3'(t) + K(1 + u_3^2(t)) = 0$. Thus the coupled system of equations (4) is transformed in the *polynomial* (in fact quadratic) system

$$\begin{cases} u_0''(t) + \delta u_0'(t) \cot(\delta t) + \delta^2(\lambda u_0(t) + u_0^2(t)) = 0, & \text{for } t \in (0, 1], \\ u_1'(t) - u_2(t) = 0, & \text{for } t \in [-1, 1], \\ u_2'(t) + K u_2(t) u_3(t) + K^2(\lambda u_1(t) + u_1^2(t)) = 0, & \text{for } t \in [-1, 1], \\ u_3'(t) + K(1 + u_3^2(t)) = 0, & \text{for } t \in [-1, 1], \\ u_0'(0) = 0, \ u_0(1) = u_1(-1), \ u_1'(1) = 0, \ \delta u_2(-1) = K u_0'(1), \ u_3(-1) = \cot(\delta). \end{cases} \tag{5}$$

We remark that the above technique of enlarging a system (with non-polynomial nonlinearities) to make it polynomial is standard (e.g. see [10, 13]).

As mentioned earlier, the idea is to solve the u_0 equation with a Taylor series about 0 while solving the other three differential equations with Chebyshev series. Hence, let

$$u_0(t) \stackrel{\text{def}}{=} \sum_{n=0}^{\infty} (a_0)_n t^n \quad \text{and} \quad u_j(t) \stackrel{\text{def}}{=} (a_j)_0 + 2\sum_{n=1}^{\infty} (a_j)_n T_n(t) \quad (j = 1, 2, 3), \tag{6}$$

where $T_0(t) = 1$, $T_1(t) = t$ and $T_{k+1}(t) = 2tT_k(t) - T_{k-1}(t)$ (for $k \geq 1$) are the Chebyshev polynomials.

To handle the singular term $\cot(\delta t)$ in the u_0 equation, we proceed exactly as in [2]. Note first that

$$\cot(\delta t) = \frac{1}{\delta t} - 2\sum_{n\geq 1} \frac{(\delta t)^{2n-1}}{\pi^{2n}}\zeta(2n), \qquad \zeta(2n) \stackrel{\text{def}}{=} \sum_{k\geq 1} \frac{1}{k^{2n}}$$

where ζ is the Riemann zeta function. Hence, denote

$$\delta t \cot(\delta t) = 1 - 2\sum_{n\geq 1} \left(\frac{\delta t}{\pi}\right)^{2n} \zeta(2n) = \sum_{n\geq 0} b_n t^n,$$

with

$$b_n \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } n = 0, \\ -2\left(\frac{\delta}{\pi}\right)^n \zeta(n) & \text{if } n \geq 1 \text{ is even}, \\ 0 & \text{if } n \text{ is odd}, \end{cases}$$

where the values $\zeta(n)$ can be rigorously estimated using finite sums computations with interval arithmetic and tail estimation using integral estimates (see [2]).

Given $j \in \{0,1,2,3\}$, denote $a_j = \{(a_j)_n\}_{n\in\mathbb{N}}$ and let $a \stackrel{\text{def}}{=} (a_0, a_1, a_2, a_3)$. Denote $b = \{b_n\}_{n\in\mathbb{N}}$ and define the sequence $Ja_0 \stackrel{\text{def}}{=} \{n(a_0)_n\}_{n\in\mathbb{N}}$. Plugging the Taylor series for u_0 in the first ODE of (5), equating powers and assuming that $u_0'(0) = 0$ leads to

$$n(n-1)(a_0)_n + (b * Ja_0)_n + \delta^2(\lambda(a_0)_{n-2} + (a_0 * a_0)_{n-2}) = 0, \quad n \geq 2, \tag{7}$$

where the symbol $*$ denotes the Cauchy product. For example, $(b * Ja_0)_n \stackrel{\text{def}}{=} \sum_{k\geq 0} b_k (Ja_0)_{n-k}$. The conditions $u_0'(0) = 0$ and $u_0(1) = u_1(-1)$ are imposed by requiring that

$$(a_0)_1 = 0 \quad \text{and} \quad \sum_{j=0}^{\infty} (a_0)_j = (a_1)_0 + 2\sum_{j=1}^{\infty} (-1)^j (a_1)_j, \tag{8}$$

respectively, where we used the standard property $T_j(-1) = (-1)^j$ of the Chebyshev polynomials. Combining the conditions (7) and (8) leads to a zero finding problem $F_0 = 0$, where F_0 is defined componentwise by

$$(F_0(a))_n \stackrel{\text{def}}{=} \begin{cases} (a_0)_1 & n = 0, \\ \displaystyle\sum_{j=0}^{\infty} (a_0)_j - (a_1)_0 - 2\sum_{j=1}^{\infty} (-1)^j (a_1)_j, & n = 1, \\ n(n-1)(a_0)_n + (b * Ja_0)_n + \delta^2(\lambda(a_0)_{n-2} + (a_0 * a_0)_{n-2}) & n \geq 2. \end{cases} \tag{9}$$

Next we obtain similar zero finding problems $F_1 = F_2 = F_3 = 0$ resulting from solving the remaining three ODEs of (5) with Chebyshev series. In order to do so, define the tridiagonal operator $\mathcal{T}$ acting on a sequence $c = \{c_n\}_{n\geq 0}$ by

$$(\mathcal{T}c)_n = \begin{cases} 0, & n = 0, \\ -c_{n-1} + c_{n+1}, & n \geq 1 \end{cases} \tag{10}$$

and let

$$\begin{aligned} (\phi_1(a))_n &\stackrel{\text{def}}{=} (a_2)_n, \\ (\phi_2(a))_n &\stackrel{\text{def}}{=} -K(a_2 * a_3)_n - K^2(\lambda(a_1)_n + (a_1 * a_1)_n), \\ (\phi_3(a))_n &\stackrel{\text{def}}{=} -K(\delta_{n,0} + (a_3 * a_3)_n), \end{aligned}$$

where $\delta_{i,j}$ denotes the Kronecker delta and where $*$ denotes the discrete convolution. For example $(a_2 * a_3)_n \stackrel{\text{def}}{=} \sum_{\substack{n_1+n_2=n \\ n_1,n_2\in\mathbb{Z}}} (a_2)_{|n_1|}(a_3)_{|n_2|}$. For the sake of simplicity of the presentation, we use the same symbol $*$ to denote both the Cauchy product and the discrete convolution, and it should be clear from the context which one is which, depending of the type of sequence (Taylor or Chebyshev) the product involves.

Plugging the Chebyshev series expansions of u_1, u_2 and u_3 (see (6)) in the remaining three ODEs of (5) and in the conditions $u_1'(1) = 0$, $\delta u_2(-1) = K u_0'(1)$ and $u_3(-1) = \cot(\delta)$, leads to (e.g. see [14]) the zero finding problems $F_1 = F_2 = F_3 = 0$, where

$$(F_1(a))_n \stackrel{\text{def}}{=} \begin{cases} \sum_{j=1}^{\infty} j^2 (a_1)_j & n = 0, \\ 2n(a_1)_n + (\mathcal{T}\phi_1(a))_n & n \geq 1, \end{cases} \tag{11}$$

$$(F_2(a))_n \stackrel{\text{def}}{=} \begin{cases} \delta\left((a_2)_0 + 2\sum_{j=1}^{\infty}(-1)^j (a_2)_j\right) - K\sum_{j=0}^{\infty} j(a_0)_j & n = 0, \\ 2n(a_2)_n + (\mathcal{T}\phi_2(a))_n & n \geq 1, \end{cases} \tag{12}$$

$$(F_3(a))_n \stackrel{\text{def}}{=} \begin{cases} (a_3)_0 + 2\sum_{j=1}^{\infty}(-1)^j (a_3)_j - \cot(\delta) & n = 0, \\ 2n(a_3)_n + (\mathcal{T}\phi_3(a))_n & n \geq 1. \end{cases} \tag{13}$$

By combining (9), (11), (12) and (13), we obtain the zero finding problem $F(a) = 0$, where $F \stackrel{\text{def}}{=} (F_0, F_1, F_2, F_3)$. We solve this problem using the following Newton-Kantorovich type theorem, whose standard proof is omitted.

Theorem 2.1. *Let X, Y be Banach spaces and assume that $F : X \to Y$ is Fréchet differentiable. Let $\bar{a} \in X$. Consider bounded linear operators $A^\dagger \in B(X,Y)$ and $A \in B(Y,X)$. Observe that*

$$AF \colon X \to X. \tag{14}$$

Assume that A is injective. Let $Y_0, Z_0, Z_1, Z_2 \geq 0$ be bounds satisfying

$$\|AF(\bar{a})\|_X \leq Y_0, \tag{15}$$

$$\|I - AA^\dagger\|_{B(X)} \leq Z_0, \tag{16}$$

$$\|A[DF(\bar{a}) - A^\dagger]\|_{B(X)} \leq Z_1, \tag{17}$$

$$\|A[DF(c) - DF(\bar{a})]\|_{B(X)} \leq Z_2 r, \quad \forall\, c \in B_r(\bar{a}), \tag{18}$$

where $B_r(\bar{a})$ denotes the open ball in X of radius $r > 0$ and centered at $\bar{a}$. Define the radii polynomial by

$$p(r) \stackrel{\text{def}}{=} Z_2 r^2 + (Z_1 + Z_0 - 1)r + Y_0. \tag{19}$$

If there exists $r_0 > 0$ such that

$$p(r_0) < 0, \tag{20}$$

then there exists a unique $\tilde{a} \in B_{r_0}(\bar{a})$ such that $F(\tilde{a}) = 0$.

Typically the choices made to apply Theorem 2.1 are as follows. The Banach space X corresponds to the cartesian product of weighed ℓ^1 sequence spaces of coefficients (in our case of Taylor and Chebyshev coefficients) decaying geometrically to 0. The Banach space Y is similar to X, but incorporates the loss of regularity coming from applying the differential operators in the ODEs to the solutions. The point $\bar{a}$ (the center of the ball) is a numerical approximation of $F = 0$ obtained via applying Newton's method to a finite dimensional reduction. The operator $A^\dagger$ is an approximation of the Fréchet derivative $DF(\bar{a})$ while the operator A is an approximate inverse of $DF(\bar{a})$.

Let us make these choices explicit. To define the Banach space X, we begin by defining weighted ℓ^1 spaces of Taylor and Chebyshev coefficients. For a sequence of weights $\omega \stackrel{\text{def}}{=} (\omega_n)_{n\geq 0}$ with positive entries, and a sequence $c = (c_n)_{n\geq 0}$, denote

$$\|c\|_{1,\omega} \stackrel{\text{def}}{=} \sum_{n\geq 0} |c_n|\omega_n$$

and

$$\ell^1_\omega \stackrel{\text{def}}{=} \{c = (c_n)_{n\geq 0} : \|c\|_{1,\omega} < \infty\}.$$

Given a number $\mu > 1$, define the *Taylor sequence of weights* $\omega_{\mathbf{T}} = \omega_{\mathbf{T}}(\mu)$ component-wise by $(\omega_{\mathbf{T}})_n \stackrel{\text{def}}{=} \mu^n$. Using these weights, $\ell^1_{\omega_{\mathbf{T}}}$ is a Banach algebra under the Cauchy product, that is for all $c_1, c_2 \in \ell^1_{\omega_{\mathbf{T}}}$,

$\|c_1 * c_2\|_{1,\omega_{\mathbf{T}}} \leq \|c_1\|_{1,\omega_{\mathbf{T}}}\|c_2\|_{1,\omega_{\mathbf{T}}}$. Given a number $\nu \geq 1$, define the *Chebyshev sequence of weights* $\omega_{\mathbf{C}} = \omega_{\mathbf{C}}(\nu)$ component-wise by

$$(\omega_{\mathbf{C}})_n = \begin{cases} 1, & n = 0 \\ 2\nu^n, & n \geq 1. \end{cases}$$

Using these weights, $\ell^1_{\omega_{\mathbf{C}}}$ is a Banach algebra under the discrete convolution, that is for all $c_1, c_2 \in \ell^1_{\omega_{\mathbf{C}}}$, $\|c_1 * c_2\|_{1,\omega_{\mathbf{C}}} \leq \|c_1\|_{1,\omega_{\mathbf{C}}}\|c_2\|_{1,\omega_{\mathbf{C}}}$. Letting $X_0 \stackrel{\rm def}{=} \ell^1_{\omega_{\mathbf{T}}}$ and $X_j \stackrel{\rm def}{=} \ell^1_{\omega_{\mathbf{C}}}$ for $j = 1, 2, 3$, set

$$X \stackrel{\rm def}{=} X_0 \oplus X_1 \oplus X_2 \oplus X_3$$

so that $a = (a_0, a_1, a_2, a_3) \in X$. Given $\alpha_j > 0$, $j = 0, 1, 2, 3$ (to be chosen later), the norm in X is given by

$$\|a\|_X = \max\{\alpha_0\|a_0\|_{1,\omega_{\mathbf{T}}}, \alpha_1\|a_1\|_{1,\omega_{\mathbf{C}}}, \alpha_2\|a_2\|_{1,\omega_{\mathbf{C}}}, \alpha_3\|a_3\|_{1,\omega_{\mathbf{C}}}\}.$$

Define the new weights $\tilde{\omega}_{\mathbf{T}}$ and $\tilde{\omega}_{\mathbf{C}}$ component-wise by $(\tilde{\omega}_{\mathbf{T}})_0 = (\tilde{\omega}_{\mathbf{C}})_0 = 1$ and for $n \geq 1$, $(\tilde{\omega}_{\mathbf{T}})_n \stackrel{\rm def}{=} \omega_{\mathbf{T}}/n^2$ and $\tilde{\omega}_{\mathbf{C}} \stackrel{\rm def}{=} \omega_{\mathbf{C}}/n$. Set $Y \stackrel{\rm def}{=} \ell^1_{\tilde{\omega}_{\mathbf{T}}} \oplus \ell^1_{\tilde{\omega}_{\mathbf{C}}} \oplus \ell^1_{\tilde{\omega}_{\mathbf{C}}} \oplus \ell^1_{\tilde{\omega}_{\mathbf{C}}}$. Using the fact that $\ell^1_{\omega_{\mathbf{T}}}$ and $\ell^1_{\omega_{\mathbf{C}}}$ are Banach algebras under the Cauchy product and discrete convolution, respectively, it is a simple to verify that $F : X \to Y$. Having defined the Banach spaces X and Y, we now turn to the question of computing a numerical approximation $\bar{a}$ of $F = 0$. This requires first considering a finite dimensional projection.

Given a number $m \in \mathbb{N}$, and given a vector $c = (c_n)_{n\geq 0} \in \ell^1_\omega$, consider the projection

$$\pi^m : \ell^1_\omega \to \mathbb{R}^{m+1}$$
$$c \mapsto \pi^m c \stackrel{\rm def}{=} (c_n)_{n=0}^m \in \mathbb{R}^{m+1}.$$

Given $c \in \ell^1_\omega$, we sometimes will use the notation $c^{(m)} \stackrel{\rm def}{=} \pi^m c$. Given a Taylor projection M and a Chebyshev projection N, we generalize that projection to get $\Pi^{(M,N)} : X \to \mathbb{R}^{M+1} \times \mathbb{R}^{N+1} \times \mathbb{R}^{N+1} \times \mathbb{R}^{N+1}$ defined by

$$\Pi^{(M,N)}(a) = \Pi^{(M,N)}(a_0, a_1, a_2, a_3) \stackrel{\rm def}{=} (\pi^M a_0, \pi^N a_1, \pi^N a_2, \pi^N a_3) \in \mathbb{R}^{(M+1)+3(N+1)}.$$

Often, given $a \in X$, we denote

$$a^{(M,N)} \stackrel{\rm def}{=} \Pi^{(M,N)} a \in \mathbb{R}^{M+3N+4}.$$

For any weights ω, we define the natural inclusion $\iota^m : \mathbb{R}^{m+1} \hookrightarrow \ell^1_\omega$ as follows. For $c = (c_n)_{n=0}^m \in \mathbb{R}^{m+1}$, we define $\iota^m c \in \ell^1_\omega$ component-wise by

$$(\iota^m c)_n = \begin{cases} c_n, & n = 0, \dots, m \\ 0, & n \geq m. \end{cases}$$

Similarly, let $\iota^{(M,N)} : \mathbb{R}^{M+3N+4} \hookrightarrow X$ be the natural inclusion defined as follows. Given $a = (a_0, a_1, a_2, a_3) \in \mathbb{R}^{M+3N+4}$, we define

$$\iota^{(M,N)} a \stackrel{\rm def}{=} \left(\iota^M a_0, \iota^N a_1, \iota^N a_2, \iota^N a_3\right) \in X.$$

Let the *finite dimensional projection* $F^{(M,N)} : \mathbb{R}^{M+3N+4} \to \mathbb{R}^{M+3N+4}$ of the map F be defined, for $a \in \mathbb{R}^{M+3N+4}$, as

$$F^{(M,N)}(a) = \Pi^{(M,N)} F(\iota^{(M,N)} a). \tag{21}$$

Similarly, we define $F_0^{(M)}(a) = \pi^M F_0(\iota^{(M,N)} a)$, and $F_j^{(N)}$ by $F_j^{(N)}(a) = \pi^N F_j(\iota^{(M,N)} a)$, for $j = 1, 2, 3$. Using that notation, we may write $F^{(M,N)}(a) = \left(F_0^{(M)}(a), F_1^{(N)}(a), F_2^{(N)}(a), F_3^{(N)}(a)\right)$.

Having defined the finite dimensional reduction, we can apply Newton's method to compute $\bar{a} = (\bar{a}_0, \bar{a}_1, \bar{a}_2, \bar{a}_3) \in \mathbb{R}^{M+3N+4}$ such that $F^{(M,N)}(\bar{a}) \approx 0$. For Newton's method to converge, we however need a good initial guess, which in our case is provided via a bifurcation analysis and a continuation method (see Section 4.1).

The next step towards applying Theorem 2.1 is to define the operators $A^\dagger$ and A. Denote by $DF^{(M,N)}(\bar{a})$ the Jacobian of $F^{(M,N)}$ at $\bar{a}$, and let us write it as

$$DF^{(M,N)}(\bar{a}) = \begin{pmatrix} D_{a_0}F_0^{(M)}(\bar{a}) & D_{a_1}F_0^{(M)}(\bar{a}) & D_{a_2}F_0^{(M)}(\bar{a}) & D_{a_3}F_0^{(M)}(\bar{a}) \\ D_{a_0}F_1^{(N)}(\bar{a}) & D_{a_1}F_1^{(N)}(\bar{a}) & D_{a_2}F_1^{(N)}(\bar{a}) & D_{a_3}F_1^{(N)}(\bar{a}) \\ D_{a_0}F_2^{(N)}(\bar{a}) & D_{a_1}F_2^{(N)}(\bar{a}) & D_{a_2}F_2^{(N)}(\bar{a}) & D_{a_3}F_2^{(N)}(\bar{a}) \\ D_{a_0}F_3^{(N)}(\bar{a}) & D_{a_1}F_3^{(N)}(\bar{a}) & D_{a_2}F_3^{(N)}(\bar{a}) & D_{a_3}F_3^{(N)}(\bar{a}) \end{pmatrix} \in M_{M+3N+4}(\mathbb{R}).$$

Using the above notation, let

$$A^\dagger \stackrel{\text{def}}{=} \begin{pmatrix} A^\dagger_{00} & A^\dagger_{01} & A^\dagger_{02} & A^\dagger_{03} \\ A^\dagger_{10} & A^\dagger_{11} & A^\dagger_{12} & A^\dagger_{13} \\ A^\dagger_{20} & A^\dagger_{21} & A^\dagger_{22} & A^\dagger_{23} \\ A^\dagger_{30} & A^\dagger_{31} & A^\dagger_{32} & A^\dagger_{33} \end{pmatrix} \tag{22}$$

where $A^\dagger_{00} : \ell^1_{\omega_\mathbf{T}} \to \ell^1_{\tilde{\omega}_\mathbf{T}}$, $A^\dagger_{0j} : \ell^1_{\omega_\mathbf{C}} \to \ell^1_{\omega_\mathbf{T}}$, $A^\dagger_{i0} : \ell^1_{\omega_\mathbf{T}} \to \ell^1_{\omega_\mathbf{C}}$, $A^\dagger_{ij} : \ell^1_{\omega_\mathbf{C}} \to \ell^1_{\tilde{\omega}_\mathbf{C}}$ for $i, j = 1, 2, 3$, are defined by

$$(A^\dagger_{00} h_0)_n = \begin{cases} \left(D_{a_0} F_0^{(M)}(\bar{a}) h_0^{(M)}\right)_n & \text{for } 0 \le n \le M, \\ n(n-1)(h_0)_n & \text{for } n > M, \end{cases}$$

$$(A^\dagger_{0j} h_j)_n = \begin{cases} \left(D_{a_j} F_0^{(M)}(\bar{a}) h_j^{(N)}\right)_n & \text{for } 0 \le n \le M, \\ 0 & \text{for } n > M, \end{cases} \qquad (\text{for } j = 1, 2, 3),$$

$$(A^\dagger_{i0} h_0)_n = \begin{cases} \left(D_{a_0} F_i^{(N)}(\bar{a}) h_0^{(M)}\right)_n & \text{for } 0 \le n \le N, \\ 0 & \text{for } n > N, \end{cases} \qquad (\text{for } i = 1, 2, 3),$$

$$(A^\dagger_{ij} h_j)_n = \begin{cases} \left(D_{a_j} F_i^{(N)}(\bar{a}) h_j^{(N)}\right)_n & \text{for } 0 \le n \le N, \\ \delta_{i,j} 2n (h_j)_n & \text{for } n > N, \end{cases} \qquad (\text{for } i, j = 1, 2, 3).$$

The action of $A^\dagger$ on an element $h = (h_0, h_1, h_2, h_3) \in X$ is defined by $(A^\dagger h)_i = \sum_{j=0}^3 A^\dagger_{i,j} h_j$, for $i = 0, 1, 2, 3$.

Consider now a matrix $A^{(M,N)} \in M_{M+3N+4}(\mathbb{R})$ computed so that $A^{(M,N)} \approx DF^{(M,N)}(\bar{a})^{-1}$. We decompose it block-wise as

$$A^{(M,N)} = \begin{pmatrix} A_{00}^{(M,M)} & A_{01}^{(M,N)} & A_{02}^{(M,N)} & A_{03}^{(M,N)} \\ A_{10}^{(N,M)} & A_{11}^{(N,N)} & A_{12}^{(N,N)} & A_{13}^{(N,N)} \\ A_{20}^{(N,M)} & A_{21}^{(N,N)} & A_{22}^{(N,N)} & A_{23}^{(N,N)} \\ A_{30}^{(N,M)} & A_{31}^{(N,N)} & A_{32}^{(N,N)} & A_{33}^{(N,N)} \end{pmatrix}.$$

This allows defining the linear operator A as

$$A = \begin{pmatrix} A_{00} & A_{01} & A_{02} & A_{03} \\ A_{10} & A_{11} & A_{12} & A_{13} \\ A_{20} & A_{21} & A_{22} & A_{23} \\ A_{30} & A_{31} & A_{32} & A_{33} \end{pmatrix}, \tag{23}$$

where $A_{00} : \ell^1_{\omega_\mathbf{T}} \to \ell^1_{\omega_\mathbf{T}}$, for $i, j > 0$, $A_{0j} : \ell^1_{\omega_\mathbf{C}} \to \ell^1_{\omega_\mathbf{T}}$, $A_{i0} : \ell^1_{\omega_\mathbf{T}} \to \ell^1_{\omega_\mathbf{C}}$, $A_{ij} : \ell^1_{\omega_\mathbf{C}} \to \ell^1_{\omega_\mathbf{C}}$ are defined by

$$(A_{00} h_0)_n = \begin{cases} \left(A_{00}^{(M,M)} h_0^{(M)}\right)_n & \text{for } 0 \le n \le M, \\ \frac{1}{n(n-1)} (h_0)_n & \text{for } n > M, \end{cases}$$

$$(A_{0j} h_j)_n = \begin{cases} \left(A_{0j}^{(M,N)} h_j^{(N)}\right)_n & \text{for } 0 \le n \le M, \\ 0 & \text{for } n > M, \end{cases} \qquad (\text{for } j = 1, 2, 3),$$

$$(A_{i0} h_0)_n = \begin{cases} \left(A_{i0}^{(N,M)} h_0^{(M)}\right)_n & \text{for } 0 \le n \le N, \\ 0 & \text{for } n > N, \end{cases} \qquad (\text{for } i = 1, 2, 3),$$

$$(A_{ij} h_j)_n = \begin{cases} \left(A_{ij}^{(N,N)} h_j^{(N)}\right)_n & \text{for } 0 \le n \le N, \\ \delta_{i,j} \frac{1}{2n} (h_j)_n & \text{for } n > N, \end{cases} \qquad (\text{for } i, j = 1, 2, 3).$$

The action of A on an element $h = (h_0, h_1, h_2, h_3) \in X$ is defined by $(Ah)_i = \sum_{j=0}^3 A_{i,j} h_j$, for $i = 0, 1, 2, 3$.

Having obtained an approximate solution $\bar{a}$ and the linear operators $A^\dagger$ and A, the next step is to construct the bounds Y_0, Z_0, Z_1 and Z_2 satisfying (15), (16), (17) and (18), respectively.

3 Bounds for the Newton-Kantorovich theorem

In this section, we introduce the formulas for the bounds $Y_0, Z_0, Z_1, Z_2 \geq 0$ satisfying (15), (16), (17) and (18), respectively.

3.1 The bound Y_0

Recalling (15), note that Y_0 satisfies

$$\|AF(\bar{a})\|_X = \max_{0\leq i\leq 3} \alpha_i \sum_{j=0}^{3} \|A_{ij}F_j(\bar{a})\|_{X_i} \leq Y_0$$

For $i, j \geq 1$,

$$\begin{aligned}
\|A_{ij}F_j(\bar{a})\|_{1,\omega_{\mathbf{C}}} &= |(A_{ij}F_j(\bar{a}))_0| + 2\sum_{n=1}^{\infty} |(A_{ij}F_j(\bar{a}))_n|\nu^n \\
&= |(A_{ij}F_j(\bar{a}))_0| + 2\sum_{n=1}^{N} |(A_{ij}F_j(\bar{a}))_n|\nu^n + 2\sum_{n=N+1}^{\infty} |(A_{ij}F_j(\bar{a}))_n|\nu^n \\
&= \|\iota^N A_{ij}^{(N,N)} F_j^{(N)}(\bar{a})\|_{1,\omega_{\mathbf{C}}} + 2\sum_{n=N+1}^{\infty} \frac{\delta_{i,j}}{2n}|(F_i(\bar{a}))_n|\nu^n \\
&\leq \|\iota^N A_{ij}^{(N,N)} F_j^{(N)}(\bar{a})\|_{1,\omega_{\mathbf{C}}} + \delta_{i,j}\sum_{n=N+1}^{2N+1} \frac{1}{n}|(F_i(\bar{a}))_n|\nu^n \stackrel{\text{def}}{=} Y_{i,j}^{(0)}.
\end{aligned}$$

The last inequality follows from the fact that for $j = 1, 2, 3$, the Chebyshev coefficients satisfy $(\bar{a}_j)_n = 0$ for $n > N$, and then $(F_i(\bar{a}))_n = 0$ for all $n > 2N + 1$ since the problem (5) is quadratic and because of the action of $\mathcal{T}$ (being tridiagonal).

Using the same idea, we define for $i = 0$ and $j = 1, 2, 3$,

$$\|A_{0j}F_j(\bar{a})\|_{1,\omega_{\mathbf{T}}} = \|\iota^M A_{0j}^{(M,N)} F_j^{(N)}(\bar{a})\|_{1,\omega_{\mathbf{T}}} \stackrel{\text{def}}{=} Y_{0,j}^{(0)}$$

and for $i = 1, 2, 3$ and $j = 0$,

$$\|A_{i0}F_0(\bar{a})\|_{1,\omega_{\mathbf{C}}} = \|\iota^N A_{i0}^{(N,M)} F_0^{(M)}(\bar{a})\|_{1,\omega_{\mathbf{C}}} \stackrel{\text{def}}{=} Y_{i,0}^{(0)}.$$

For the case $i = j = 0$, since we analytically defined the Taylor coefficients $\{b_n\}_{n\in\mathbb{N}}$ of $\delta t\cot(\delta t)$ in F_0, we have

$$\begin{aligned}
\|A_{00}F_0(\bar{a})\|_{1,\omega_{\mathbf{T}}} &= \sum_{n=0}^{\infty} |(A_{00}F_0(\bar{a}))_n|\mu^n \\
&= \|\iota^M A_{00}^{(M,M)} F_0^{(M)}(\bar{a})\|_{1,\omega_{\mathbf{T}}} + \sum_{n=M+1}^{2M+2} \frac{|(F_0(\bar{a}))_n|}{n(n-1)}\mu^n + \sum_{n=2M+3}^{\infty} \frac{|(b * J\bar{a}_0)_n|}{n(n-1)}\mu^n.
\end{aligned}$$

To bound the tail, we use the estimate

$$|b_n| = \left|2\left(\frac{\delta}{\pi}\right)^n \zeta(k)\right| = 2\left(\frac{\delta}{\pi}\right)^n \sum_{j\geq 1}\frac{1}{j^n} \leq 2\left(\frac{\delta}{\pi}\right)^n \sum_{j\geq 1}\frac{1}{j^2} = \left(\frac{\delta}{\pi}\right)^n \frac{\pi^2}{3}, \qquad \text{for } n \geq 2.$$

From there, we have

$$\begin{aligned}
\sum_{n=2M+3}^{\infty} \frac{|(b*J\bar{a}_0)_n|}{n(n-1)}\mu^n &\leq \frac{1}{(2M+3)(2M+2)} \sum_{n=2M+3}^{\infty} \left| \sum_{j=0}^{M} j(\bar{a}_0)_j b_{n-j} \right| \mu^n \\
&\leq \frac{1}{(2M+3)(2M+2)} \sum_{j=0}^{M} |j(\bar{a}_0)_j|\mu^j \sum_{n=2M+3}^{\infty} |b_{n-j}|\, \mu^{n-j} \\
&\leq \frac{\|\iota^M J\bar{a}_0\|_{1,\omega_{\mathbf{T}}}}{(2M+3)(2M+2)} \sum_{n=M+3}^{\infty} |b_n|\, \mu^n \\
&\leq \frac{\pi^2 \|\iota^M J\bar{a}_0\|_{1,\omega_{\mathbf{T}}}}{3(2M+3)(2M+2)} \sum_{n=M+3}^{\infty} \left(\frac{\delta\mu}{\pi}\right)^n \\
&\leq \left(\frac{\pi^2 \|\iota^M J\bar{a}_0\|_{1,\omega_{\mathbf{T}}}}{3(2M+3)(2M+2)}\right) \left(\frac{1}{1-\frac{\delta\mu}{\pi}}\right) \left(\frac{\delta\mu}{\pi}\right)^{M+3} \stackrel{\text{def}}{=} Y_{tail}^{(0)}.
\end{aligned}$$

The last inequality comes from the fact that $\delta \in (0, \frac{\pi}{2})$ and the fact that we choose $\mu < \frac{\pi}{\delta}$.

We define

$$Y_{0,0}^{(0)} \stackrel{\text{def}}{=} \|\iota^M A_{00}^{(M,M)} F_0^{(M)}(\bar{a})\|_{1,\omega_{\mathbf{T}}} + \sum_{n=M+1}^{2M+2} \frac{1}{n(n-1)} |(F_0(\bar{a}))_n|\mu^n + Y_{tail}^{(0)}.$$

Finally, using the above results, set

$$Y_0 \stackrel{\text{def}}{=} \max_{i\in\{0,1,2,3\}} \left(\alpha_i \sum_{j=0}^{3} Y_{i,j}^{(0)}\right). \tag{24}$$

3.2 The bound Z_0

Recall from (16) that the bound Z_0 satisfies

$$\|I - AA^\dagger\|_{B(X,X)} \leq Z_0.$$

We set

$$\mathcal{E} \stackrel{\text{def}}{=} I - AA^\dagger = \begin{pmatrix} \mathcal{E}_{00} & \mathcal{E}_{01} & \mathcal{E}_{02} & \mathcal{E}_{03} \\ \mathcal{E}_{10} & \mathcal{E}_{11} & \mathcal{E}_{12} & \mathcal{E}_{13} \\ \mathcal{E}_{20} & \mathcal{E}_{21} & \mathcal{E}_{22} & \mathcal{E}_{23} \\ \mathcal{E}_{30} & \mathcal{E}_{31} & \mathcal{E}_{32} & \mathcal{E}_{33} \end{pmatrix},$$

then $\mathcal{E}_{00} : \ell^1_{\omega_{\mathbf{T}}} \to \ell^1_{\omega_{\mathbf{T}}}$, and for $i,j>0$, $\mathcal{E}_{0j} : \ell^1_{\omega_{\mathbf{C}}} \to \ell^1_{\omega_{\mathbf{T}}}$, $\mathcal{E}_{i0} : \ell^1_{\omega_{\mathbf{T}}} \to \ell^1_{\omega_{\mathbf{C}}}$, $\mathcal{E}_{ij} : \ell^1_{\omega_{\mathbf{C}}} \to \ell^1_{\omega_{\mathbf{C}}}$, with

$$\begin{aligned}
(\mathcal{E}_{00}h_0)_n &= \begin{cases} \left(\mathcal{E}_{00}^{(M,M)} h_0^{(M)}\right)_n & \text{for } 0 \leq n \leq M, \\ 0 & \text{for } n > M, \end{cases} \\
(\mathcal{E}_{0j}h_j)_n &= \begin{cases} \left(\mathcal{E}_{0j}^{(M,N)} h_j^{(N)}\right)_n & \text{for } 0 \leq n \leq M, \\ 0 & \text{for } n > M, \end{cases} \qquad (\text{for } j = 1,2,3), \\
(\mathcal{E}_{i0}h_0)_n &= \begin{cases} \left(\mathcal{E}_{i0}^{(N,M)} h_0^{(M)}\right)_n & \text{for } 0 \leq n \leq N, \\ 0 & \text{for } n > N, \end{cases} \qquad (\text{for } i = 1,2,3), \\
(\mathcal{E}_{ij}h_j)_n &= \begin{cases} \left(\mathcal{E}_{ij}^{(N,N)} h_j^{(N)}\right)_n & \text{for } 0 \leq n \leq N, \\ 0 & \text{for } n > N, \end{cases} \qquad (\text{for } i,j = 1,2,3).
\end{aligned}$$

where for instance, the matrix $\mathcal{E}_{00}^{(M,M)} \stackrel{\text{def}}{=} I - A^{(M,M)} D_{a_j} F_0^{(M)}(\bar{a}) \in M_{M+1}(\mathbb{R})$. Denoting the components of $\mathcal{E}_{ij}$ by $(\mathcal{E}_{i,j})_{m,n}$, using standard operator norms computations, one can show that

$$\|I - AA^\dagger\|_{B(X,X)} \leq \max_{0\leq i\leq 3} \left(\alpha_i \sum_{j=0}^{3} \frac{1}{\alpha_j} K_{i,j} \right) \stackrel{\text{def}}{=} Z_0, \tag{25}$$

where

$$K_{i,j} \stackrel{\text{def}}{=} \begin{cases} \max\limits_{0\leq n\leq M} \dfrac{1}{(\omega_{\mathbf{T}})_n} \sum\limits_{m=0}^{M} |(\mathcal{E}_{0,0})_{m,n}|(\omega_{\mathbf{T}})_m, & i=j=0, \\ \max\limits_{0\leq n\leq N} \dfrac{1}{(\omega_{\mathbf{T}})_n} \sum\limits_{m=0}^{M} |(\mathcal{E}_{0,j})_{m,n}|(\omega_{\mathbf{C}})_m, & i=0, j\geq 1, \\ \max\limits_{0\leq n\leq M} \dfrac{1}{(\omega_{\mathbf{C}})_n} \sum\limits_{m=0}^{N} |(\mathcal{E}_{i,0})_{m,n}|(\omega_{\mathbf{T}})_m, & i\geq 1, j=0, \\ \max\limits_{0\leq n\leq N} \dfrac{1}{(\omega_{\mathbf{C}})_n} \sum\limits_{m=0}^{N} |(\mathcal{E}_{i,j})_{m,n}|(\omega_{\mathbf{C}})_m, & i,j\geq 1. \end{cases}$$

3.3 The bound Z_1

Recall from (17) that the bound Z_1 satisfies

$$\|A[DF(\bar{a}) - A^\dagger]\|_{B(X,X)} = \sup_{\|h\|_X=1} \|A[DF(\bar{a}) - A^\dagger]h\|_X \leq Z_1.$$

Before going further in the computation of the bound, we need the identity

$$(b * Jh_0)_n = n(b * h_0)_n - (Jb * h_0)_n,$$

where $Jb = \{nb_n\}_{n\in\mathbb{N}}$, which we will apply to the Taylor coefficients b_n of $\delta t \cot(\delta t)$. We set

$$\|b\|_{1,\omega_{\mathbf{T}}} = \sum_{n=0}^{\infty} |b_n|\mu^n \leq 1 + \frac{\pi^2}{3} \sum_{n=1}^{\infty} \left(\frac{\delta\mu}{\pi}\right)^{2n} = 1 + \frac{(\delta\mu)^2}{3(1-(\frac{\delta\mu}{\pi})^2)} \stackrel{\text{def}}{=} C_1,$$
$$\|Jb\|_{1,\omega_{\mathbf{T}}} = \sum_{n=0}^{\infty} |nb_n|\mu^n \leq \frac{2\pi^2}{3} \sum_{n=1}^{\infty} n \left(\frac{\delta\mu}{\pi}\right)^{2n} = \frac{2(\delta\mu)^2}{3(1-(\frac{\delta\mu}{\pi})^2)^2} \stackrel{\text{def}}{=} C_2.$$

Let $h \in X$, we define

$$(y_{ij})_n \stackrel{\text{def}}{=} ([D_{a_j} F_i(\bar{a}) - A_{ij}^\dagger]h_j)_n \quad \text{for } i,j \in \{0,1,2,3\} \text{ and } n \geq 0.$$

For i,j,n all nonzero we will often rewrite $(y_{ij})_n$ with the help of $(z_{ij})_{n=1}^{N+1}$ and $(\tilde{z}_{ij})_{n=N}^{\infty}$, which are defined (explicitly below) such that

$$(\mathcal{T} z_{ij})_n = (y_{ij})_n, \qquad \text{for } 1 \leq n \leq N \tag{26}$$
$$(\mathcal{T} \tilde{z}_{ij})_n = (y_{ij})_n, \qquad \text{for } n \geq N+1. \tag{27}$$

We choose this notation $(\mathcal{T} z_{ij})_n$ and $(\mathcal{T} \tilde{z}_{ij})_n$ since it will be, at times, easier to derive the bounds if we are working with $A_{ii'} \mathcal{T} z_{i'j}$ instead of $A_{ii'} y_{i'j}$. The distinction between z and $\tilde{z}$ is necessary since the projection π^N and the tridiagonal operator $\mathcal{T}$ do not commute.

Given $c \in \ell^1_{\omega_{\mathbf{T}}}$, we use the notation $c^{I_M} \stackrel{\text{def}}{=} (I - \pi^M)c$. Similarly, given $c \in \ell^1_{\omega_{\mathbf{C}}}$, we use the notation

$c^{I_N} \stackrel{\text{def}}{=} (I-\pi^N)c$. For $i=0$ we find

$$(y_{00})_k = \begin{cases} 0 & \text{for } k=0, \\ \sum_{j=M+1}^{\infty}(h_0)_j & \text{for } k=1, \\ 0 & \text{for } 2\le k\le M, \\ k(b*h_0)_k - (Jb*h_0)_k + \delta^2(\lambda(h_0)_{k-2} + 2(\bar{a}_0*h_0)_{k-2}) & \text{for } k\ge M+1, \end{cases}$$

$$(y_{01})_k = \begin{cases} 0 & \text{for } k=0, \\ 2\sum_{j=N+1}^{\infty}(-1)^{j+1}(h_1)_j & \text{for } k=1, \\ 0 & \text{for } k\ge 2, \end{cases}$$

$$(y_{02})_k = (y_{03})_k = 0, \qquad \text{for all } k,$$

where we have used that $(y_{00})_k = 2\delta^2(\bar{a}_0 * h_0^{I_M})_{k-2} = 0$ for $k=2,\dots,M$.

For $i=1$, we find $(y_{10})_k = 0$ for all $k\ge 0$ and

$$(y_{11})_0 = \sum_{j=N+1}^{\infty} j^2 (h_1)_j,$$

and $(y_{12})_0 = (y_{13})_0 = 0$. For $i=2$, we have

$$(y_{20})_k = \begin{cases} -K\sum_{j=M+1}^{\infty} j(h_0)_j & \text{for } k=0, \\ 0 & \text{for } k\ge 1, \end{cases}$$

and

$$(y_{22})_0 = 2\delta \sum_{j=N+1}^{\infty} (-1)^j (h_2)_j,$$

and $(y_{21})_0 = (y_{23})_0 = 0$. And finally for $i=3$, we have $(y_{30})_k = 0$ for all $k\ge 0$ and

$$(y_{33})_0 = 2\sum_{j=N+1}^{\infty} (-1)^j (h_3)_j,$$

and $(y_{31})_0 = (y_{32})_0 = 0$.

We now turn to $(y_{ij})_k$ for i,j,k all nonzero, which we describe in terms of z and $\tilde{z}$ as expressed in (26) and (27). Since ϕ_1 does not depend on a_1 and a_3, we define z_{11}, and z_{13} as well as $\tilde{z}_{11}$ and $\tilde{z}_{13}$ to vanish. Moreover, we set

$$\begin{aligned} (z_{12})_k &\stackrel{\text{def}}{=} (h_2^{I_N})_k && \text{for } 1\le k\le N+1. \\ (\tilde{z}_{12})_k &\stackrel{\text{def}}{=} (h_2)_k && \text{for } k\ge N. \end{aligned}$$

Similary, defining $\bar{\lambda} \stackrel{\text{def}}{=} (\lambda,0,0,...) \in \ell^1_{\omega_{\mathbf{C}}}$ we set

$$\begin{aligned} (z_{21})_k &\stackrel{\text{def}}{=} -K^2((\bar{\lambda}+2\bar{a}_1)*h_1^{I_N})_k && \text{for } 1\le k\le N+1, \\ (\tilde{z}_{21})_k &\stackrel{\text{def}}{=} -K^2((\bar{\lambda}+2\bar{a}_1)*h_1)_k && \text{for } k\ge N, \\ (z_{22})_k &\stackrel{\text{def}}{=} -K(\bar{a}_3*h_2^{I_N})_k && \text{for } 1\le k\le N+1, \\ (\tilde{z}_{22})_k &\stackrel{\text{def}}{=} -K(\bar{a}_3*h_2)_k && \text{for } k\ge N, \\ (z_{23})_k &\stackrel{\text{def}}{=} -K(\bar{a}_2*h_3^{I_N})_k && \text{for } 1\le k\le N+1, \\ (\tilde{z}_{23})_k &\stackrel{\text{def}}{=} -K(\bar{a}_2*h_3)_k && \text{for } k\ge N. \end{aligned}$$

Furtermore, we define z_{31}, and z_{32} as well as $\tilde{z}_{31}$ and $\tilde{z}_{32}$ to vanish, while setting

$$\begin{aligned} (z_{33})_k &\stackrel{\text{def}}{=} -2K(\bar{a}_3*h_3^{I_N})_k && \text{for } 1\le k\le N+1, \\ (\tilde{z}_{33})_k &\stackrel{\text{def}}{=} -2K(\bar{a}_3*h_3)_k && \text{for } k\ge N. \end{aligned}$$

To take advantage of the notation with z and $\tilde{z}$ we will need the following two results

Lemma 3.1. *Let $\mathcal{T} : \ell^1_{\omega_{\mathbf{C}}} \to \ell^1_{\omega_{\mathbf{C}}}$ be the operator defined in* (10)*, then*

$$\|\mathcal{T}\|_{B(\ell^1_{\omega_{\mathbf{C}}})} \leq 2\nu.$$

Proof. Let $h \in \ell^1_{\omega_{\mathbf{C}}}$ with $\|h\|_{1,\omega_{\mathbf{C}}} = 1$. Since $\nu \geq 1$, note that $1/\nu \leq \nu$. Hence,

$$\begin{aligned}
\|\mathcal{T}h\|_{1,\omega_{\mathbf{C}}} &= 2\sum_{k=1}^{\infty} |(\mathcal{T}h)_k|\nu^k \\
&= 2\sum_{k=1}^{\infty} |-h_{k-1} + h_{k+1}|\nu^k \\
&\leq 2\nu \sum_{k=0}^{\infty} |h_k|\nu^k + \frac{2}{\nu}\sum_{k=2}^{\infty} |h_k|\nu^k \\
&\leq \nu\left(|h_0| + \|h\|_{1,\omega_{\mathbf{C}}}\right) + 2\nu\sum_{k=2}^{\infty} |h_k|\nu^k \\
&= \nu\|h\|_{1,\omega_{\mathbf{C}}} + \nu\left(|h_0| + 2\sum_{k=2}^{\infty} |h_k|\nu^k\right) \\
&\leq \nu\|h\|_{1,\omega_{\mathbf{C}}} + \nu\|h\|_{1,\omega_{\mathbf{C}}} \\
&\leq 2\nu.
\end{aligned}$$

□

Lemma 3.2. *Let $\bar{a} \in \ell^1_{\omega_{\mathbf{C}}}$ be such that $\bar{a}_n = 0$ for $n > N$. Let $h = \{h_n\}_{n\in\mathbb{N}} \in \ell^1_{\omega_{\mathbf{C}}}$ with $\|h\|_{1,\omega_{\mathbf{C}}} \leq 1$. Let $h^{I_N} = (0, ..., 0, h_{N+1}, h_{N+2}, ...)$. We define $\hat{l}^k_{\bar{a}} : \ell^1_{\omega_{\mathbf{C}}} \to \mathbb{R}$ by*

$$\hat{l}^k_{\bar{a}}(h) \stackrel{\text{def}}{=} (\bar{a} * h^{I_N})_k.$$

Let

$$\Psi_k(\bar{a}) \stackrel{\text{def}}{=} \max_{N+1\leq j\leq k+N} \frac{|\bar{a}_{|k-j|}|}{2\nu^j}, \tag{28}$$

then,

$$\left|\hat{l}^k_{\bar{a}}(h)\right| \leq \Psi_k(\bar{a}). \tag{29}$$

Proof. This follows from the identity $\hat{l}^k_{\bar{a}}(h) = \sum_{j>N}(\bar{a}_{|k-j|} + \bar{a}_{|k+j|})h_j$. The details are left to the reader. □

The bound Z_1 will be assembled, using the triangle inequality, from bounds

$$Z^{(1)}_{ij} \geq \sum_{i'=0}^{3} \|A_{ii'}[D_{a_j}F_{i'}(\bar{a}) - A^{\dagger}_{i'j}]h_j\|_{X_i} \qquad \text{with } i,j = 0,1,2,3,$$

uniformly for $\|h_j\|_{X_j} \leq 1$. In what follows we will estimate each term separately:

$$\|A_{ii'}[D_{a_j}F_{i'}(\bar{a}) - A^{\dagger}_{i'j}]h_j\|_{X_i} = \|A_{ii'}y_{i'j}\|_{X_i} \leq \mathcal{A}_{ii'j} \qquad \text{for all } \|h_j\|_{X_j} \leq 1. \tag{30}$$

It is immediate from the vanishing of the corresponding partial derivatives $D_{a_j}F_{i'}$ that

$$\mathcal{A}_{i02} = \mathcal{A}_{i03} = \mathcal{A}_{i10} = \mathcal{A}_{i13} = \mathcal{A}_{i30} = \mathcal{A}_{i31} = \mathcal{A}_{i32} = 0 \qquad \text{for } i = 0,1,2,3.$$

To determine the nonzero terms, let us start with

$$\|A_{00}y_{00}\|_{1,\omega_{\mathbf{T}}} = \sum_{k=0}^{\infty} |(A_{00}y_{00})_k|\,\mu^k = \sum_{k=0}^{M} |(A_{00}y_{00})_k|\,\mu^k + \sum_{k=M+1}^{\infty} \frac{1}{k(k-1)}\,|(y_{00})_k|\,\mu^k.$$

We estimate each term separately. First, recalling that $(y_{00})_j = 0$ for $j = 0, 2, \ldots, M$,

$$\sum_{k=0}^{M} |(A_{00}y_{00})_k|\,\mu^k = \sum_{k=0}^{M} \left|\sum_{n=1}^{M} (A_{00})_{k,n}(y_{00})_n\right| \mu^k = \sum_{k=0}^{M} |(A_{00})_{k,1}(y_{00})_1|\,\mu^k$$
$$\leq \left|\sum_{n'\geq M+1} (h_0)_{n'}\right| \cdot \|(A_{00})_{:,1}\|_{1,\omega_{\mathbf{T}}} \leq \frac{\|(A_{00})_{:,1}\|_{1,\omega_{\mathbf{T}}}}{\mu^{M+1}},$$

where $(A_{ij})_{:,k}$ represents the vector composed of the elements of the $(k+1)^{th}$ column of A_{ij}. Second, recall that $\bar{\lambda} = (\lambda, 0, 0, ...)$ which we now interpret as an element of $\ell^1_{\omega_{\mathbf{T}}}$. Then

$$\sum_{k=M+1}^{\infty} \frac{1}{k(k-1)} |(y_{00})_k|\,\mu^k = \sum_{k\geq M+1} \frac{|k(b*h_0)_k - (Jb*h_0)_k + \delta^2(\lambda(h_0)_{k-2} + 2(\bar{a}_0*h_0)_{k-2}|}{k(k-1)}\mu^k$$
$$\leq \frac{\|b*h_0\|_{1,\omega_{\mathbf{T}}}}{M} + \frac{\|Jb*h_0\|_{1,\omega_{\mathbf{T}}}}{M(M+1)} + \frac{\delta^2\mu^2\|(\bar{\lambda}+2\bar{a}_0)*h_0\|_{1,\omega_{\mathbf{T}}}}{M(M+1)}$$
$$\leq \frac{C_1}{M} + \frac{C_2}{M(M+1)} + \frac{\delta^2\mu^2\|\bar{\lambda}+2\bar{a}_0\|_{1,\omega_{\mathbf{T}}}}{M(M+1)}$$

where C_1 and C_2 are defined in the beginning of the section. We set

$$\mathcal{A}_{000} \stackrel{\text{def}}{=} \frac{\|(A_{00})_{:,1}\|_{1,\omega_{\mathbf{T}}}}{\mu^{M+1}} + \frac{C_1}{M} + \frac{C_2}{M(M+1)} + \frac{\delta^2\mu^2\|\bar{\lambda}+2\bar{a}_0\|_{1,\omega_{\mathbf{T}}}}{M(M+1)}.$$

Using the same method, but simpler since there are no tail terms, we find

$$\|A_{00}y_{01}\|_{1,\omega_{\mathbf{T}}} \leq \sum_{k=0}^{\infty}\sum_{n=0}^{\infty} |(A_{00})_{k,n}||(y_{01})_n|\mu^k \leq 2\sum_{k=0}^{M}\sum_{n'=N+1}^{\infty} |(A_{00})_{k,1}||(h_1)_{n'}|\mu^k \leq \frac{\|(A_{00})_{:,1}\|_{1,\omega_{\mathbf{T}}}}{\nu^{N+1}} \stackrel{\text{def}}{=} \mathcal{A}_{001}.$$

Next we consider

$$\|A_{02}y_{20}\|_{1,\omega_{\mathbf{T}}} \leq \sum_{k=0}^{M}\sum_{n=0}^{N} |(A_{02})_{k,n}||(y_{20})_n|\mu^k \leq K\sum_{k=0}^{M}\sum_{n'=M+1}^{\infty} |(A_{02})_{k,0}|n'|(h_0)_{n'}|\mu^k$$
$$\leq \frac{K(M+1)\|(A_{02})_{:,0}\|_{1,\omega_{\mathbf{T}}}}{\mu^{M+1}} \stackrel{\text{def}}{=} \mathcal{A}_{020}.$$

Here we have assumed that $M+1 \geq (\log\mu)^{-1}$. Analogously, we find, for $i = 1, 2, 3$

$$\mathcal{A}_{i20} \stackrel{\text{def}}{=} \frac{K(M+1)\|(A_{i2})_{:,0}\|_{1,\omega_{\mathbf{C}}}}{\mu^{M+1}} \qquad \mathcal{A}_{i00} \stackrel{\text{def}}{=} \frac{\|(A_{i0})_{:,1}\|_{1,\omega_{\mathbf{C}}}}{\mu^{M+1}} \qquad \mathcal{A}_{i01} \stackrel{\text{def}}{=} \frac{\|(A_{i0})_{:,1}\|_{1,\omega_{\mathbf{C}}}}{\nu^{N+1}}.$$

We now turn attention to the terms $i', j = 1, 2, 3$. As a representative case, we consider $i = i' = j = 2$:

$$\|A_{22}y_{22}\|_{1,\omega_{\mathbf{C}}} = \sum_{k\geq 0}\left(|(A_{22})_{k,0}(y_{22})_0| + \sum_{n=1}^{N} |(A_{22})_{k,n}(\mathcal{T}z_{22})_n| + \sum_{n=N+1}^{\infty} |(A_{22})_{k,n}(\mathcal{T}\tilde{z}_{22})_n|\right)(\omega_{\mathbf{C}})_k.$$

Here we have used that A_{22} is block-diagonal with a finite part $A_{22}^{(N,N)}$ and a diagonal operator in the tail, as well as that $\mathcal{T}$ is tridiagonal. Let us look at each of the three terms individually. The first term is bounded by

$$|(y_{22})_0|\sum_{k=0}^{N} |(A_{22})_{k,0}|(\omega_{\mathbf{C}})_k = 2\delta\left|\sum_{j\geq N+1} (-1)^j (h_2)_j\right| \cdot \|(A_{22})_{:,0}\|_{1,\omega_{\mathbf{C}}}$$
$$\leq 2\delta\sum_{j\geq N+1} |(h_2)_j|\,\frac{2\nu^j}{2\nu^{N+1}} \cdot \|(A_{22})_{:,0}\|_{1,\omega_{\mathbf{C}}}$$
$$\leq \frac{2\delta\|h_2\|_{1,\omega_{\mathbf{C}}}}{2\nu^{N+1}}\|(A_{22})_{:,0}\|_{1,\omega_{\mathbf{C}}} \leq \frac{\delta}{\nu^{N+1}}\|(A_{22})_{:,0}\|_{1,\omega_{\mathbf{C}}}.$$

The third term is bounded by

$$
\begin{aligned}
2\sum_{k\ge N+1}\left(\sum_{n\ge N+1}|(A_{22})_{k,n}(\mathcal{T}\tilde z_{22})_n|\right)\nu^k &= 2\sum_{k\ge N+1}\left(\frac{|(\tilde z_{22})_{k+1}-(\tilde z_{22})_{k-1}|}{2k}\right)\nu^k\\
&\le \frac{2K}{2(N+1)}\sum_{k\ge N+1}\left(|(\bar a_3 * h_2)_{k+1}|+|(\bar a_3 * h_2)_{k-1}|\right)\nu^k\\
&\le \frac{K}{N+1}\left(\frac{1}{2\nu}\sum_{k\ge N+1}|(\bar a_3 * h_2)_{k+1}|2\nu^{k+1}+\frac{\nu}{2}\sum_{k\ge N+1}|(\bar a_3 * h_2)_{k-1}|2\nu^{k-1}\right)\\
&\le \frac{K\left(\frac{1}{\nu}\|\bar a_3 * h_2\|_{1,\omega_{\mathbf{C}}}+\nu\|\bar a_3 * h_2\|_{1,\omega_{\mathbf{C}}}\right)}{2(N+1)}\\
&\le \frac{K\left(\nu+\frac{1}{\nu}\right)\|\bar a_3\|_{1,\omega_{\mathbf{C}}}}{2(N+1)}.
\end{aligned}
$$

The second term is bounded by

$$
\begin{aligned}
\sum_{k=0}^{N}\left|\sum_{n'=1}^{N}\sum_{n=0}^{N+1}(A_{22})_{k,n'}\mathcal{T}_{n',n}(z_{22})_n\right|(\omega_{\mathbf{C}})_k &\le \sum_{k=0}^{N}\sum_{n=0}^{N+1}\left|\sum_{n'=0}^{N}(A_{22})_{k,n'}\mathcal{T}_{n',n}\right||(z_{22})_n|(\omega_{\mathbf{C}})_k\\
&\le \left\|\left|A_{22}^{(N,N)}\mathcal{T}^{(N,N+1)}\right|\left|z_{22}^{(N+1)}\right|\right\|_{1,\omega_{\mathbf{C}}},
\end{aligned}
$$

where we have used that the first row of $\mathcal{T}$ vanishes, absolute values are to be interpreted componentwise, and we have somewhat abused the notation of restricting operators on and elements of $\ell^1_{\omega_{\mathbf{C}}}$ to finite dimensional projections, e.g. $\mathcal{T}^{(N,N+1)}\in M_{N+1,N+2}(\mathbb{R})$. We then use Lemma 29 to estimate $(z_{22})_n = K(\bar a_3 * h_3^{I_N})_n$ and we obtain

$$
\left\|\left|A_{22}^{(N,N)}\mathcal{T}^{(N,N+1)}\right|\left|z_{22}^{(N+1)}\right|\right\|_{1,\omega_{\mathbf{C}}} \le K\||A_{22}\mathcal{T}|_N^{N+1}\Psi(\bar a_3)\|_{1,\omega_{\mathbf{C}}},
$$

where $|A_{22}\mathcal{T}|_N^{N+1}\Psi(\bar a_3)$ should be read as an abbreviation for $|A_{22}^{(N,N)}\mathcal{T}^{(N,N+1)}|\Psi(\bar a_3)^{(N+1)}$. Combining these three parts, we set

$$
\mathcal{A}_{222} \stackrel{\text{def}}{=} \frac{\delta}{\nu^{N+1}}\|(A_{22})_{:,0}\|_{1,\omega_{\mathbf{C}}} + K\||A_{22}\mathcal{T}|_N^{N+1}\Psi(\bar a_3)\|_{1,\omega_{\mathbf{C}}} + \frac{K\left(\nu+\frac{1}{\nu}\right)\|\bar a_3\|_{1,\omega_{\mathbf{C}}}}{2(N+1)}
$$

All other terms can be estimated similarly. We find

$$
\begin{aligned}
\mathcal{A}_{022} &\stackrel{\text{def}}{=} \frac{\delta}{\nu^{N+1}}\|(A_{02})_{:,0}\|_{1,\omega_{\mathbf{T}}} + K\||A_{02}\mathcal{T}|_N^{N+1}\Psi(\bar a_3)\|_{1,\omega_{\mathbf{C}}}\\
\mathcal{A}_{i22} &\stackrel{\text{def}}{=} \frac{\delta}{\nu^{N+1}}\|(A_{i2})_{:,0}\|_{1,\omega_{\mathbf{C}}} + K\||A_{i2}\mathcal{T}|_N^{N+1}\Psi(\bar a_3)\|_{1,\omega_{\mathbf{C}}} \quad \text{for } i=1,3.
\end{aligned}
$$

Furthermore

$$
\mathcal{A}_{011} \stackrel{\text{def}}{=} \frac{(N+1)^2}{2\nu^{N+1}}\|(A_{01})_{:,0}\|_{1,\omega_{\mathbf{T}}}, \qquad \mathcal{A}_{i11} \stackrel{\text{def}}{=} \frac{(N+1)^2}{2\nu^{N+1}}\|(A_{i1})_{:,0}\|_{1,\omega_{\mathbf{C}}}, \quad \text{for } i=1,2,3.
$$

where we have assumed that $N+1\ge 2(\log\nu)^{-1}$. The remaining constants are

$$
\begin{aligned}
\mathcal{A}_{033} &\stackrel{\text{def}}{=} \frac{1}{\nu^{N+1}}\|(A_{03})_{:,0}\|_{1,\omega_{\mathbf{T}}} + 2K\||A_{03}\mathcal{T}|_N^{N+1}\Psi(\bar a_3)\|_{1,\omega_{\mathbf{T}}}\\
\mathcal{A}_{333} &\stackrel{\text{def}}{=} \frac{1}{\nu^{N+1}}\|(A_{33})_{:,0}\|_{1,\omega_{\mathbf{C}}} + 2K\||A_{33}\mathcal{T}|_N^{N+1}\Psi(\bar a_3)\|_{1,\omega_{\mathbf{C}}} + \frac{K\left(\nu+\frac{1}{\nu}\right)\|\bar a_3\|_{1,\omega_{\mathbf{C}}}}{N+1}\\
\mathcal{A}_{i33} &\stackrel{\text{def}}{=} \frac{1}{\nu^{N+1}}\|(A_{i3})_{:,0}\|_{1,\omega_{\mathbf{C}}} + 2K\||A_{i3}\mathcal{T}|_N^{N+1}\Psi(\bar a_3)\|_{1,\omega_{\mathbf{C}}} \quad \text{for } i=1,2,
\end{aligned}
$$

and

$$\begin{aligned}
\mathcal{A}_{021} &\stackrel{\text{def}}{=} K^2 \| |A_{02} \mathcal{T}|_N^{N+1} \Psi(\bar\lambda + 2\bar a_1) \|_{1,\omega_{\mathbf{T}}} \\
\mathcal{A}_{221} &\stackrel{\text{def}}{=} K^2 \| |A_{22} \mathcal{T}|_N^{N+1} \Psi(\bar\lambda + 2\bar a_1) \|_{1,\omega_{\mathbf{C}}} + \frac{K^2 \left(\nu + \frac{1}{\nu}\right) \|\bar\lambda + 2\bar a_1\|_{1,\omega_{\mathbf{C}}}}{2(N+1)} \\
\mathcal{A}_{i21} &\stackrel{\text{def}}{=} K^2 \| |A_{i2} \mathcal{T}|_N^{N+1} \Psi(\bar\lambda + 2\bar a_1) \|_{1,\omega_{\mathbf{C}}} \quad \text{for } i = 1, 3,
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{A}_{023} &\stackrel{\text{def}}{=} K \| |A_{02} \mathcal{T}|_N^{N+1} \Psi(\bar a_2) \|_{1,\omega_{\mathbf{T}}} \\
\mathcal{A}_{223} &\stackrel{\text{def}}{=} K \| |A_{22} \mathcal{T}|_N^{N+1} \Psi(\bar a_2) \|_{1,\omega_{\mathbf{C}}} + \frac{K \left(\nu + \frac{1}{\nu}\right) \|\bar a_2\|_{1,\omega_{\mathbf{C}}}}{2(N+1)} \\
\mathcal{A}_{i23} &\stackrel{\text{def}}{=} K \| |A_{i2} \mathcal{T}|_N^{N+1} \Psi(\bar a_2) \|_{1,\omega_{\mathbf{C}}} \quad \text{for } i = 1, 3,
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{A}_{012} &\stackrel{\text{def}}{=} \frac{1}{2\nu^{N+1}} \|(A_{01})_{:,N}\|_{1,\omega_{\mathbf{T}}} \\
\mathcal{A}_{112} &\stackrel{\text{def}}{=} \frac{1}{2\nu^{N+1}} \|(A_{11})_{:,N}\|_{1,\omega_{\mathbf{C}}} + \frac{\left(\nu + \frac{1}{\nu}\right)}{2(N+1)} \\
\mathcal{A}_{i12} &\stackrel{\text{def}}{=} \frac{1}{2\nu^{N+1}} \|(A_{i1})_{:,N}\|_{1,\omega_{\mathbf{C}}} \quad \text{for } i = 2, 3.
\end{aligned}$$

Having constructed the bounds $\mathcal{A}_{ii'j}$ satisfying (30) we define

$$Z_{ij}^{(1)} \stackrel{\text{def}}{=} \sum_{i'=0}^{3} \mathcal{A}_{ii'j} \qquad \text{for } i, j = 0, 1, 2, 3.$$

Finally, we have

$$Z_1 \stackrel{\text{def}}{=} \max_{i \in \{0,1,2,3\}} \left(\alpha_i \sum_{j=0}^{3} \frac{1}{\alpha_j} Z_{ij}^{(1)} \right). \tag{31}$$

Following the approach proposed in [3], let us briefly describe the choice of the weights $\alpha_0, \ldots, \alpha_3$ in order to minimize the bound Z_1. This is done by using a result from the Perron–Frobenius theorem. We see that the matrix

$$M_{Z_1} \stackrel{\text{def}}{=} \begin{bmatrix} Z_{00}^{(1)} & Z_{01}^{(1)} & Z_{02}^{(1)} & Z_{03}^{(1)} \\ Z_{10}^{(1)} & Z_{11}^{(1)} & Z_{12}^{(1)} & Z_{13}^{(1)} \\ Z_{20}^{(1)} & Z_{21}^{(1)} & Z_{22}^{(1)} & Z_{23}^{(1)} \\ Z_{30}^{(1)} & Z_{31}^{(1)} & Z_{32}^{(1)} & Z_{33}^{(1)} \end{bmatrix} \tag{32}$$

is non-negative. Thus, according to the Perron–Frobenius theorem, M_{Z_1} has a largest real eigenvalue ρ with the corresponding eigenvector v_ρ and by the Collatz–Wielandt formula, if v_ρ is positive, then it is the solution of

$$\max_v \left(\min_{i \in \{0,1,2,3\}} \frac{(M_{Z_1} v)_i}{v_i} \right)$$

over all positive vector $v \in \mathbb{R}^4$. For our problem this means that we simply need to compute the dominant eigenvalue ρ and corresponding eigenvector v_ρ of M_{Z_1} and setting $\alpha_j = 1/(v_\rho)_j$.

Note that the above choice of α allows us to minimize the bound on Z_1, but it will also have a negative impact on the other bounds. Y_0 will be the one that loses the most in the process. So at that point it is a case by case problem where we have to experiment with the α_j to find the perfect balance.

3.4 The bound Z_2

We recall from (18) that Z_2 satisfies

$$\|A[DF(c)-DF(\bar{a})]\|_{B(X,X)} = \sup_{\|h\|_X=1} \|A[DF(c)-DF(\bar{a})]h\|_X \leq Z_2 r$$

for all $c=(c_0,c_1,c_2,c_3)\in \overline{B_r(\bar{a})}$ and all $r>0$. We define $d=(d_0,d_1,d_2,d_3)\in X$ such that $d_i \stackrel{\text{def}}{=} c_i-\bar{a}_i$ and thus $\|d_i\|_{X_i}\leq \frac{r}{\alpha_i}$. We define the operator $\Delta DF : X\to X$, by

$$\Delta DF(d) \stackrel{\text{def}}{=} \begin{bmatrix} B_{00} & B_{01} & B_{02} & B_{03} \\ \mathcal{T}B_{10} & \mathcal{T}B_{11} & \mathcal{T}B_{12} & \mathcal{T}B_{13} \\ \mathcal{T}B_{20} & \mathcal{T}B_{21} & \mathcal{T}B_{22} & \mathcal{T}B_{23} \\ \mathcal{T}B_{30} & \mathcal{T}B_{31} & \mathcal{T}B_{32} & \mathcal{T}B_{33} \end{bmatrix},$$

with $\mathcal{T}$ defined at (10) and where

$$(B_{00}h_0)_k \stackrel{\text{def}}{=} \begin{cases} 0 & \text{for } k=0,1 \\ 2\delta^2(d_0*h_0)_{k-2} & \text{for } k\geq 2, \end{cases}$$
$$(B_{01}h_1)_k \stackrel{\text{def}}{=} (B_{02}h_2)_k \stackrel{\text{def}}{=} (B_{03}h_3)_k \stackrel{\text{def}}{=} 0,$$
$$(B_{10}h_0)_k \stackrel{\text{def}}{=} (B_{11}h_1)_k \stackrel{\text{def}}{=} (B_{12}h_2)_k \stackrel{\text{def}}{=} (B_{13}h_3)_k \stackrel{\text{def}}{=} 0,$$
$$(B_{21}h_1)_k \stackrel{\text{def}}{=} \begin{cases} 0 & \text{for } k=0, \\ -2K^2(d_1*h_1)_k & \text{for } k\geq 1, \end{cases}$$
$$(B_{22}h_2)_k \stackrel{\text{def}}{=} \begin{cases} 0 & \text{for } k=0, \\ -K(d_3*h_2)_k & \text{for } k\geq 1, \end{cases}$$
$$(B_{23}h_3)_k \stackrel{\text{def}}{=} \begin{cases} 0 & \text{for } k=0, \\ -K(d_2*h_3)_k & \text{for } k\geq 1, \end{cases}$$
$$(B_{30}h_0)_k \stackrel{\text{def}}{=} (B_{31}h_1)_k \stackrel{\text{def}}{=} (B_{32}h_2)_k \stackrel{\text{def}}{=} 0,$$
$$(B_{33}h_3)_k \stackrel{\text{def}}{=} \begin{cases} 0 & \text{for } k=0, \\ -2K(d_3*h_3)_k & \text{for } k\geq 1. \end{cases}$$

We notice that $\Delta DF(D)=DF(c)-DF(\bar{a})\in X$, since the tails of $DF(c)$ and$DF(\bar{a})$ cancel each other. We have

$$\|(A\Delta DF(d)h)_i\|_{X_i} \leq \sum_{j=0}^{3} \|A_{ij}(\Delta DF(d)h)_j\|_{X_i} \leq \sum_{j=0}^{3} \|A_{ij}\|_{B(X_j,X_i)}\|(\Delta DF(d)h)_j\|_{X_j}$$

where we can bound $\|A_{ij}\|_{B(X_j,X_i)}$ using similar bounds as the $K_{i,j}$ presented in Section 3.2. Also, we have

$$\|(\Delta DF(d)h)_i\|_{X_i} \leq \|T\|_{B(\ell^1_{\omega_{\mathbf{C}}})}^{(1-\delta_{i,0})} \sum_{j=0}^{3} \|B_{i,j}h_j\|_{X_j},$$

where $\|T\|_{B(\ell^1_{\omega_{\mathbf{C}}})}\leq 2\nu$ by Lemma 3.1.

First, we start by bounding

$$\begin{aligned}\|B_{00}h_0\|_{1,\omega_{\mathbf{T}}} &= \sum_{k=2}^{\infty} |2\delta^2 (d_0 * h_0)_{k-2}|\mu^k \\ &= 2\delta^2\mu^2 \sum_{k=2}^{\infty} |(d_0 * h_0)_{k-2}|\mu^{k-2} \\ &= 2\delta^2\mu^2 \|d_0 * h_0\|_{1,\omega_{\mathbf{T}}} \\ &\le 2\delta^2\mu^2 \|d_0\|_{1,\omega_{\mathbf{T}}} \|h_0\|_{1,\omega_{\mathbf{T}}} \\ &\le 2\delta^2\mu^2 \left(\frac{r}{\alpha_0}\right)\left(\frac{1}{\alpha_0}\right).\end{aligned}$$

Using the same ideas, we compute the remaining bounds

$$\begin{aligned}\|B_{21}h_1\|_{1,\omega_{\mathbf{C}}} &\le 2K^2 \left(\frac{r}{\alpha_1}\right)\left(\frac{1}{\alpha_1}\right), \\ \|B_{22}h_2\|_{1,\omega_{\mathbf{C}}} &\le K \left(\frac{r}{\alpha_3}\right)\left(\frac{1}{\alpha_2}\right), \\ \|B_{23}h_3\|_{1,\omega_{\mathbf{C}}} &\le K \left(\frac{r}{\alpha_2}\right)\left(\frac{1}{\alpha_3}\right), \\ \|B_{33}h_3\|_{1,\omega_{\mathbf{C}}} &\le 2K \left(\frac{r}{\alpha_3}\right)\left(\frac{1}{\alpha_3}\right).\end{aligned}$$

Then, if we set

$$\begin{aligned}Z_{00}^{(2)} &= \frac{2\delta^2\mu^2}{\alpha_0}\|A_{00}\|_{B(\ell^1_{\omega_{\mathbf{T}}})}, \\ Z_{01}^{(2)} &= Z_{02}^{(2)} = Z_{03}^{(2)} = 0, \\ Z_{10}^{(2)} &= Z_{11}^{(2)} = Z_{12}^{(2)} = Z_{13}^{(2)} = 0, \\ Z_{20}^{(2)} &= 0, \\ Z_{21}^{(2)} &= \frac{2K^2}{\alpha_1}\left(\frac{1}{\nu} + 2\nu\right)\|A_{21}\|_{B(\ell^1_{\omega_{\mathbf{C}}})}, \\ Z_{22}^{(2)} &= \frac{K}{\alpha_3}\left(\frac{1}{\nu} + 2\nu\right)\|A_{22}\|_{B(\ell^1_{\omega_{\mathbf{C}}})}, \\ Z_{23}^{(2)} &= \frac{K}{\alpha_2}\left(\frac{1}{\nu} + 2\nu\right)\|A_{23}\|_{B(\ell^1_{\omega_{\mathbf{C}}})}, \\ Z_{30}^{(2)} &= Z_{31}^{(2)} = Z_{32}^{(2)} = 0, \\ Z_{33}^{(2)} &= \frac{2K}{\alpha_3}\left(\frac{1}{\nu} + 2\nu\right)\|A_{33}\|_{B(\ell^1_{\omega_{\mathbf{C}}})},\end{aligned}$$

the bound Z_2 is given by

$$Z_2 \stackrel{\text{def}}{=} \max_{i\in\{0,1,2,3\}} \left(\alpha_i \sum_{j=0}^{3} Z_{ij}^{(2)} \frac{1}{\alpha_j}\right). \tag{33}$$

4 Presentation of the results

In this section, we first introduce in Section 4.1 a bifurcation analysis to compute branches of solutions (parameterized by λ) bifurcating from the trivial solution $u = 0$. Then in Section 4.2, we show how to perform a continuation to numerically compute branches of solutions, which are then validated using Theorem 2.1 using the bounds presented in Section 3.

4.1 Bifurcating from the trivial solution

Recall that (3) describes rotationally invariant symmetric solutions of the elliptic problem $\Delta u+\lambda u+u^2=0$ on the sphere. It is well-known that the eigenfunctions of the Laplacian on the sphere are given by spherical harmonics. To compute asymptotic approximations of the bifurcating solutions, we can perform the following explicit computations. First we note that the substitution $x=\cos\theta$ and $v(x)=v(\cos\theta)=u(\theta)$ transforms (3) into

$$\begin{cases}(1-x^2)v''(x)-2xv'(x)+\lambda v(x)+v(x)^2=0 & \text{for } x\in[0,1],\\ v'(0)=0.\end{cases} \tag{34}$$

We solve the linearized problem

$$(1-x^2)v''(x)-2xv'(x)+\lambda v(x)=0 \tag{35}$$

by plugging in a power series

$$v(x)=\sum_{k=0}^{\infty}a_k x^k,$$

with $a_1=0$ due to the boundary condition in (34). This leads to a recurrence relation

$$a_{k+2}=a_k\frac{k(k+1)-\lambda}{(k+2)(k+1)}, \tag{36}$$

where one may set $a_0=1$ in view of linearity of (35). The power series breaks off after $k=2n$ when

$$\lambda=\lambda_n \stackrel{\text{def}}{=} 2n(2n+1), \qquad \text{for any } n\in\mathbb{N},$$

in which case one finds a polynomial solution

$$v_n(x) \stackrel{\text{def}}{=} \sum_{m=0}^{n}a_{2m}x^{2m}. \tag{37}$$

To determine the (asymptotic) amplitude of the solution near the bifurcation point, we set

$$\lambda=\lambda_n+\varepsilon,$$

and expand $v(x)$ in terms of ε as

$$v(x)=\varepsilon A_n v_n(x)+\varepsilon^2 w(x)+O(\varepsilon^3),$$

where $A_n\in\mathbb{R}$ is unknown at this stage, and so is $w(x)$. By substituting the expressions above into (34) and using that (v_n,λ_n) solves (35) we find at order ε^2 the following equation for w:

$$(1-x^2)w''(x)-2xw'(x)+\lambda_n w(x)=-A_n^2 v_n(x)^2-A_n v_n(x). \tag{38}$$

The linear operator in the left-hand side has a nontrivial kernel. To find the *solvability condition* we multiply Equation (38) by $v_n(x)$ and integrate over $[0,1]$:

$$\int_0^1\big[(1-x^2)w''(x)-2xw'(x)+\lambda_n w(x)\big]v_n(x)dx=-\int_0^1\big[A_n^2v_n^3(x)+A_nv_n(x)^2\big]dx. \tag{39}$$

Integrating by parts in the left-hand side of (39) we find after some manipulation that

$$\int_0^1\big[(1-x^2)w''(x)-2xw'(x)+\lambda_n w(x)\big]v_n(x)dx$$
$$=\int_0^1\big[(1-x^2)v_n''(x)-2xv_n'(x)+\lambda_n v_n(x)\big]w(x)dx=0, \tag{40}$$

since (v_n,λ_n) solves (35). This is not accidental: the linear operator involved is self-adjoint. By combining (39) and (40) we find

$$A_n \stackrel{\text{def}}{=} -\frac{\int_0^1 v_n(x)^2dx}{\int_0^1 v_n(x)^3dx}, \tag{41}$$

so that for small ε the bifurcating solution of (3) at $\lambda=\lambda_n+\varepsilon$ is well approximated by

$$u^{(n)}(\theta) \stackrel{\text{def}}{=} \varepsilon A_n v_n(\cos\theta). \tag{42}$$

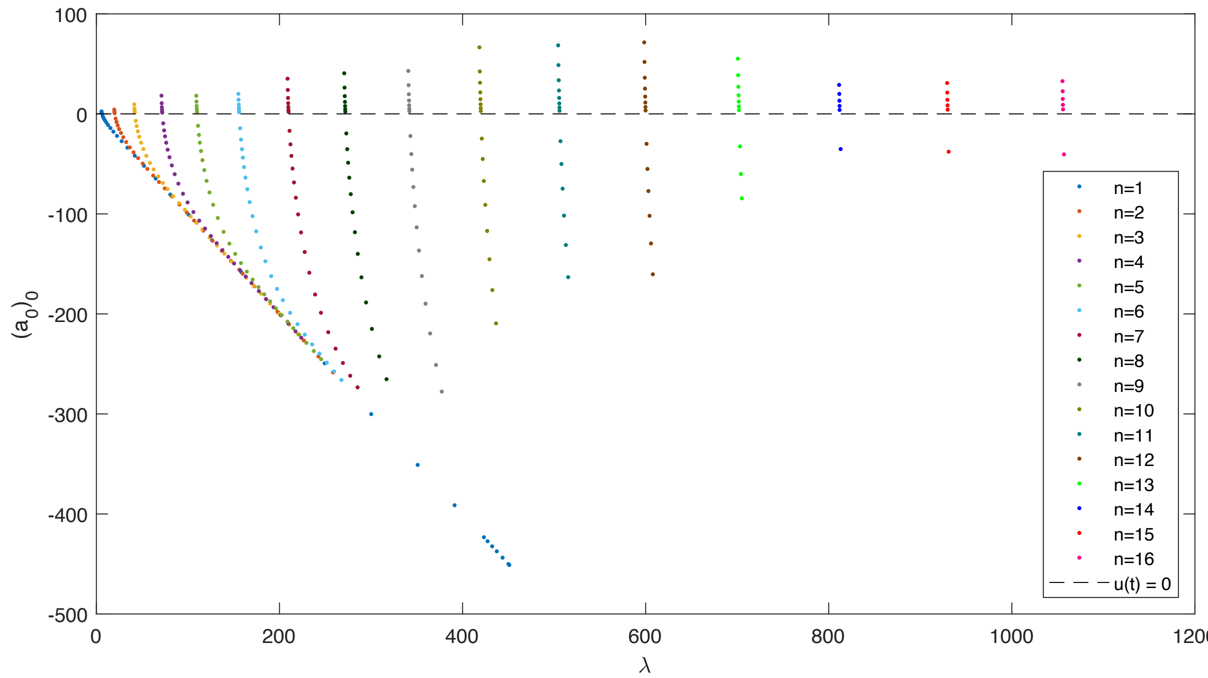

Figure 2: Bifurcation diagram of all proven solutions in this paper.

4.2 Computational details

Using the bifurcation analysis of the previous section, we produce an initial point on a branch of solutions as follows. We fix an integer $n \geq 1$, $\varepsilon > 0$ and let $\lambda = \lambda_n + \varepsilon = 2n(2n+1) + \varepsilon$. Using the explicit recursion formula (36), we compute the sequence $(a_{2m})_{m=0}^{n}$. Then, using $v_n(x)$ defined in (37), we compute A_n in (41). Based on this, Equation (42) provides a function $u^{(n)}(\theta)$ which approximately solves the BVP (3). We extract the Taylor coefficients $a_0^{(0)}$ of $u^{(n)}(\delta t)$ for $t \in (0,1]$, the Chebyshev coefficients $a_1^{(0)}$ of $u^{(n)}\left(Kt + \frac{\pi}{4} + \frac{\delta}{2}\right)$ and the Chebyshev coefficients $a_2^{(0)}$ of $\frac{d}{dt}u^{(n)}\left(Kt + \frac{\pi}{4} + \frac{\delta}{2}\right)$. Also, denote by $a_3^{(0)}$ and the Chebyshev coefficients of $\cot\left(Kt + \frac{\pi}{4} + \frac{\delta}{2}\right)$. Finally, set

$$\hat{a}^{(0)} \stackrel{\text{def}}{=} \left(a_0^{(0)}, a_1^{(0)}, a_2^{(0)}, a_3^{(0)}\right),$$

which is used as an input for Newton's method at $\lambda^{(0)} = 2n(2n+1) + \varepsilon$ to produce a point $\bar{a}^{(0)}$ such that $F\left(\bar{a}^{(0)}, \lambda^{(0)}\right) \approx 0$.

From there, we compute branches of solutions with a standard predictor-corrector continuation algorithm. We computed numerically 16 branches ($n = 1, \ldots, 16$) of solutions (see Figure 2). For $n > 16$, the bounds Z_1 become greater than 1 and we were not able to use Theorem 2.1. For each of these numerically produced solutions, we applied Theorem 2.1 to verify that close to the numerical solution, there exists an exact solution with rigorous error bound.

Since we computed more than 200 solutions, we will only present the details for the two proofs at both ends of the two branches $n = 1$ and $n = 16$. These results have all been obtained using the weights $\mu = 1.1$ and $\nu = 1.2$. Other computational parameters can be found in Table 1 for the four particular solutions, and in the code available at [23] for all other solutions. Both solutions on the $n = 1$ branch are depicted in Figure 3, while the ones on the $n = 16$ branch are shown in Figure 4.

We note that we were able to prove the existence and local uniqueness of solutions on 16 branches, whereas [2] was only able to prove results on 2 branches. In addition, the total numbers of Taylor and Chebyshev coefficients needed for our results is significantly lower than the number of Taylor coefficients in [2] for results with the same parameters. We also mention that choosing the weight α_j as described at the end of Section 3.3 allowed us to have a better control over the the bound Z_1, which needs to be less than one for Theorem 2.1 to be successful. Indeed, choosing the α_j in a way that minimize Z_1 is essential to prove some of the results. However, using the optimal α_j that minimizes Z_1 is not without sacrifice, as it can have detrimental effects on the size of the Y_0 and Z_2 bounds.

n	1		16	
λ	5.6172	451.0710	1055.3153	1057
δ	0.3	0.3	0.1	0.1
M	80	80	90	90
N	80	110	180	180
Y_0	7.6854×10^{-11}	5.5146×10^{-8}	8.1020×10^{-5}	3.1482×10^{-4}
Z_0	1.0868×10^{-8}	1.5306×10^{-10}	2.1409×10^{-4}	4.4249×10^{-5}
Z_1	0.1154	0.21536	0.1597	0.1851
Z_2	66.4698	1.1704	165.8340	108.2655
α	$[2.2299, 2.1083, 1.32, 10^3]$	$[10^3, 9.6736, 1.0054, 10^3]$	$[10^3, 17.376, 1.0017, 10^3]$	$[10^3, 17.3724, 1.0017, 10^3]$
$r_{\min}$	8.6883×10^{-11}	7.028310^{-8}	9.8353×10^{-5}	4.0855×10^{-4}
$r_{\max}$	1.3308×10^{-2}	0.67040	4.9674×10^{-3}	7.1175×10^{-3}

Table 1: Parameter values, computational constants and bounds for the solutions depicted in Figures 3 and 4, where the radii polynomial $p(r) < 0$ for $r \in (r_{\min}, r_{\max})$.

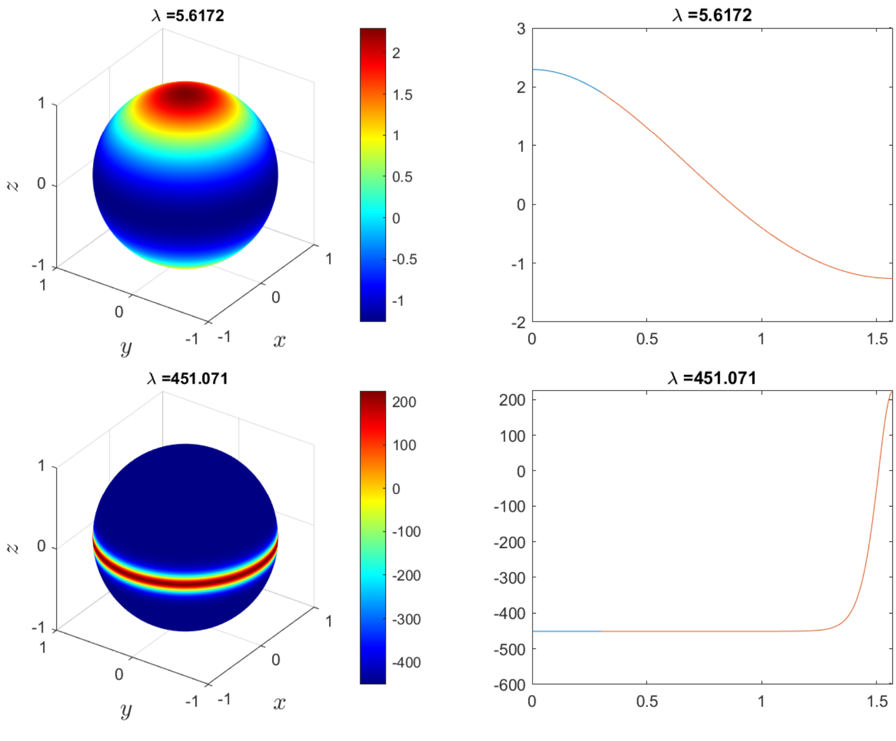

Figure 3: Solutions on the branch bifurcating at $\lambda_1 = 6$. (Left) The solution of (1) on the unit sphere $S^2 \subset \mathbb{R}^3$. (Right) The corresponding numerical solution of the BVP (3) with Taylor expansion in blue and Chebyshev expansion in orange.

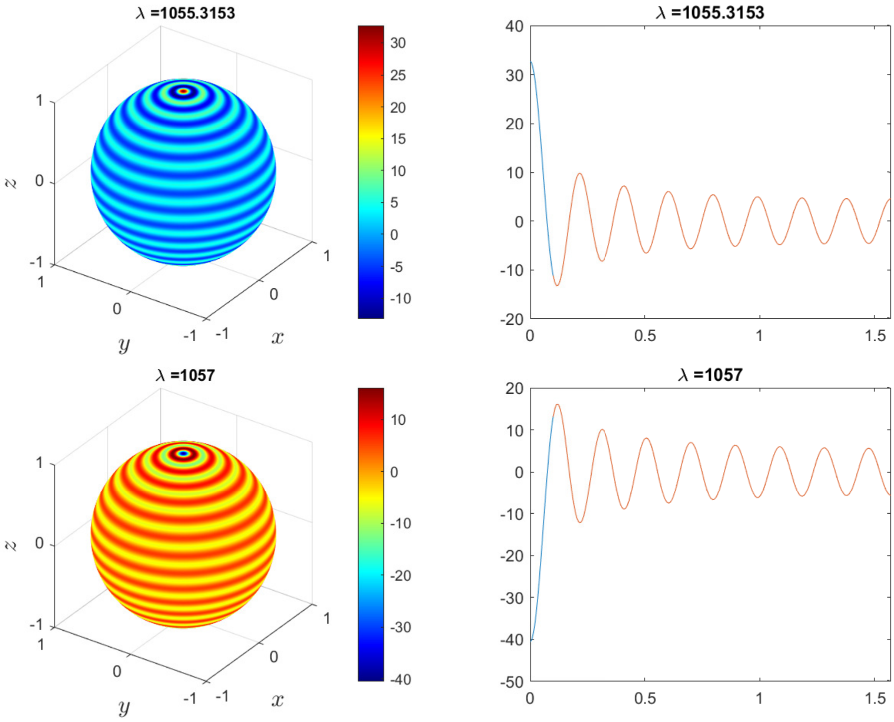

Figure 4: Solutions on the branch bifurcating at $\lambda_{16} = 1056$. (Left) The solution of (1) on the unit sphere $S^2 \subset \mathbb{R}^3$. (Right) The corresponding numerical solution of the BVP (3) with Taylor expansion in blue and Chebyshev expansion in orange.

Finally, looking at the bifurcation diagram of our solutions (see Figure 2), we can see that, for most branches, computing the solutions for $\lambda < \lambda_n$ are harder to prove than those for $\lambda > \lambda_n$. Using a pseudo arc-lenght branch following technique in future studies could be a good idea to prove more solutions for these kinds of problems.

References

[1] Gianni Arioli and Hans Koch. Non-radial solutions for some semilinear elliptic equations on the disk. *Nonlinear Anal.*, 179:294–308, 2019.

[2] István Balázs, Jan Bouwe van den Berg, Julien Courtois, János Dudás, Jean-Philippe Lessard, Anett Vörös-Kiss, J. F. Williams, and Xi Yuan Yin. Computer-assisted proofs for radially symmetric solutions of PDEs. *J. Comput. Dyn.*, 5(1-2):61–80, 2018.

[3] Maxime Breden and Christian Kuehn. Rigorous validation of stochastic transition paths. *J. Math. Pures Appl. (9)*, 131:88–129, 2019.

[4] Simon Brendle and Fernando C. Marques. Recent progress on the Yamabe problem. In *Surveys in geometric analysis and relativity*, volume 20 of *Adv. Lect. Math. (ALM)*, pages 29–47. Int. Press, Somerville, MA, 2011.

[5] Xavier Cabré. Elliptic PDE's in probability and geometry: symmetry and regularity of solutions. *Discrete Contin. Dyn. Syst.*, 20(3):425–457, 2008.

[6] A. Castro and E.M. Fischer. Infinitely many rotationally symmetric solutions to a class of semilinear Laplace-Beltrami equations on spheres. *Canad. Math. Bull.*, 58(4):723–729, 2015.

[7] Lamberto Cesari. Functional analysis and periodic solutions of nonlinear differential equations. *Contributions to Differential Equations*, 1:149–187, 1963.

[8] Lamberto Cesari. Functional analysis and Galerkin's method. *Michigan Math. J.*, 11:385–414, 1964.

[9] Thomas C. Hales. A proof of the Kepler conjecture. *Ann. Math.*, 162(3):1065–1185, 2005.

[10] Donald E. Knuth. *The art of computer programming. Vol. 2.* Addison-Wesley Publishing Co., Reading, Mass., second edition, 1981. Seminumerical algorithms, Addison-Wesley Series in Computer Science and Information Processing.

[11] Hans Koch, Alain Schenkel, and Peter Wittwer. Computer-assisted proofs in analysis and programming in logic: a case study. *SIAM Rev.*, 38(4):565–604, 1996.

[12] Oscar E. Lanford, III. A computer-assisted proof of the Feigenbaum conjectures. *Bull. Amer. Math. Soc. (N.S.)*, 6(3):427–434, 1982.

[13] Jean-Philippe Lessard, J. D. Mireles James, and Julian Ransford. Automatic differentiation for Fourier series and the radii polynomial approach. *Phys. D*, 334:174–186, 2016.

[14] Jean-Philippe Lessard and Christian Reinhardt. Rigorous Numerics for Nonlinear Differential Equations Using Chebyshev Series. *SIAM J. Numer. Anal.*, 52(1):1–22, 2014.

[15] P.-L. Lions. On the existence of positive solutions of semilinear elliptic equations. *SIAM Rev.*, 24(4):441–467, 1982.

[16] Konstantin Mischaikow and J. D. Mireles James. *Encyclopedia of Applied and Computational Mathematics*, chapter Computational Proofs in Dynamics. Springer, 2015.

[17] M. T. Nakao. Numerical verification methods for solutions of ordinary and partial differential equations. *Numer. Funct. Anal. Optim.*, 22(3-4):321–356, 2001.

[18] Michael Plum. Computer-assisted enclosure methods for elliptic differential equations. *Linear Algebra Appl.*, 324(1-3):147–187, 2001.

[19] Neil Robertson, Daniel Sanders, Paul Seymour, and Robin Thomas. The four-colour theorem. *J. Combin. Theory Ser. B*, 70(1):2–44, 1997.

[20] Siegfried M. Rump. Verification methods: rigorous results using floating-point arithmetic. *Acta Numer.*, 19:287–449, 2010.

[21] Warwick Tucker. A rigorous ODE Solver and Smale's 14th Problem. *Found. Comput. Math.*, 2(1):53–117–117, 2002.

[22] Warwick Tucker. *Validated numerics: A short introduction to rigorous computations.* Princeton University Press, Princeton, NJ, 2011.

[23] J. B. van den Berg, G.W. Duchesne, and J.-P. Lessard. MATLAB codes to perform the proofs, 2021. `http://www.math.mcgill.ca/jplessard/ResearchProjects/PatternsSphere/home.html`.

[24] J. B. van den Berg, C. M. Groothedde, and J. F. Williams. Rigorous computation of a radially symmetric localized solution in a Ginzburg-Landau problem. *SIAM J. Appl. Dyn. Syst.*, 14(1):423–447, 2015.

[25] Jan Bouwe van den Berg and Jean-Philippe Lessard. Rigorous numerics in dynamics. *Not. Am. Math. Soc.*, 62(9):1057–1061, 2015.

[26] M. J. Ward. Spots, traps, and patches: asymptotic analysis of localized solutions to some linear and nonlinear diffusive systems. *Nonlinearity*, 31(8):R189–R239, 2018.

[27] N. Yamamoto. A numerical verification method for solutions of boundary value problems with local uniqueness by Banach's fixed-point theorem. *SIAM J. Numer. Anal.*, 35:2004–2013, 1998.