

# Computer-assisted proofs for radially symmetric solutions of PDEs

István Balázs\*      Jan Bouwe van den Berg†      Julien Courtois‡  
 János Dudás§      Jean-Philippe Lessard¶      Anett Vörös-Kiss||  
 JF Williams\*\*      Xi Yuan Yin††

## Abstract

We introduce a rigorous computer-assisted method to obtain constructive proofs of existence of solutions to nonlinear differential equations. We introduce all main ideas through examples, accessible to undergraduate students, where we consider radially symmetric solutions of partial differential equations. The proofs are obtained by solving for the coefficients of the Taylor series of the solutions in a Banach space of geometrically decaying sequences. The tool that allows us to advance from numerical simulation to mathematical proof is the Banach contraction theorem.

## Keywords

Radially symmetric solutions · PDEs · Computer-assisted proofs  
 Taylor series · Contraction mapping theorem

## Mathematics Subject Classification (2010)

65M99 · 65G40 · 35K57

## 1 Introduction

In the education of mathematics students, we perceive a gap between scientific computing methods for partial differential equations (PDEs) on the one hand, and the mathematically analysis of PDEs on the other. In the last decade powerful techniques to bridge the divide, i.e., to turn numerical calculations into mathematically rigorous statements, have developed

---

\*Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, Szeged, Hungary, H-6720. bal-azs.istvan.88@gmail.com

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

‡Université de Montréal, Département de Mathématiques et de Statistique, Pavillon André-Aisenstadt, 2920 chemin de la Tour, Montreal, QC, H3T 1J4, Canada. courtoisj@dms.umontreal.ca

§Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, Szeged, Hungary, H-6720. dudas88@gmail.com

¶Université Laval, Département de Mathématiques et de Statistique, 1045 avenue de la Médecine, Québec, QC, G1V0A6, Canada. jean-philippe.lessard@mat.ulaval.ca.

||Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, Szeged, Hungary, H-6720. vorosanett89@gmail.com

\*\*Simon Fraser University, Department of Mathematics, 8888 University Drive Burnaby, BC, V5A 1S6, Canada. jfwillia@sfu.ca

††McGill University, Department of Mathematics and Statistics, Burnside Hall, 805 Sherbrooke Street West, Montreal, Quebec, H3A 0B9, Canada. xi.yin@mail.mcgill.ca

rapidly (see e.g. [7, 11, 22, 25, 27, 29]). As computer-assisted mathematical proofs are becoming mainstream, and the link between computational science and mathematics is tightening, it is fitting to introduce these techniques in undergraduate courses.

There are many examples of teaching using the computer for experimentation at all levels and as a rigorous mathematical tool in such areas as number theory and discrete mathematics. But teaching students to use a computer to do functional analysis is rare, if not unheard of. With a clever hybrid approach we can off-load the verification of intricate computationally intensive estimates to the computer to prove existence to infinite dimensional continuum problems near numerical approximations. This permits using classical ideas (such as fixed point methods) to generate results not accessible by any other means. Thus, going beyond simulations, students can get their hands on solutions of nonlinear PDE problems using the computer and still as mathematicians argue about them with all the rigour of a classical pencil and paper proof.

In this paper we present examples of this technique which are accessible to undergraduate students. We believe it is valuable for students to learn about these techniques. First, because it shows them that many simulations can be turned into rigorous proofs. Second, because it teaches them to combine techniques from computational science with areas of mathematics over-frequently thought to be completely separate: pencil and paper proofs and scientific computing.

The examples in this paper are chosen such that the technicalities are minimal (and the associated coding is easily manageable), while solving nontrivial nonlinear PDE problems. Furthermore, these problems require (functional) analysis in infinite dimensions, hence they illustrate the main ideas involved in more advanced involved problems, which require more “technical” machinery, such as approximation theory, Fourier/Chebyshev series, finite elements, etc.

This paper, and in particular the examples discussed, resulted from a summer school for undergraduate students at Simon Fraser University in the summer of 2015. Our experience is that this topic of rigorous analysis using a computer grabs the attention of students. As we will indicate in the examples, there is something there for a wide spectrum of students. With a combination of analysis and coding, it invites them to take a hands-on approach. Moreover, it presents open research problems in applied analysis that are approachable for very junior students. Analysis (of for example PDEs) is about continuous problems, while the computer calculations are inherently discrete. This invites the students to start thinking about problems differently. What can I prove starting from what I can compute? How do I set up a computer program so that it is a proof? What limitations of computers must be overcome to generate an air-tight proof? What happens when I try to implement Newton’s method in infinite dimensions? It shows them the messy difference between textbook examples and real research problems.

We note that one ingredient that is needed in the analysis, is computation using interval arithmetic. We refer to [15, 18, 23] for an introduction to interval arithmetic. We have excellent experiences with the package *Intlab* [17] for *Matlab*, but of course many alternatives are available. The matlab scripts for the examples in this paper are available at [3].

## 2 The central tool

When trying to solve numerically a finite dimensional nonlinear zero-finding for  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , the classical approach is to apply Newton’s algorithm, where one iterates the map

$$T(x) = x - DF(x)^{-1}F(x),$$

until the residue gets small. Often this numerical outcome  $x^{\text{num}}$  is either assumed “good enough” (roughly, in applied mathematics) or “not rigorous” (roughly, in pure mathematics). Our goal is to merge these perspectives by providing a rigorous quantitative statement on how

good this numerical outcome is, i.e., how close it is to a solution of  $F(x) = 0$ . The crucial observation is that Newton's method works so well because the map  $T$  is a contraction with a very small contraction constant on a neighborhood of a zero  $x^{\text{sol}}$  of  $F$ , provided that the Jacobian  $DF(x^{\text{sol}})$  is invertible, i.e.,  $x^{\text{sol}}$  is a non-degenerate zero of  $F$  (the generic situation). This allows us to *prove* that the numerical approximation  $x^{\text{num}}$  is close to the true solution  $x^{\text{sol}}$ .

Rather than going into detail about how to obtain such a computer-assisted proof, we first move to the infinite dimensional setting. After detailing in Theorem 2.1 below the setup in general Banach spaces, we return in Remark 2.2 and Example 2.3 to the (relatively straightforward) application to finite dimensional problems.

Consider now  $F : X \rightarrow X'$  a smooth map between two Banach spaces. One of the main strategies of mathematical analysis is to turn the zero finding problem  $F(x) = 0$  into a fixed point problem. The Newton operator itself is usually impractical because the inverse of the derivative of an infinite dimensional map is hard to work with. Instead, one may *choose* an injective linear map  $A \in L(X', X)$  and study the fixed point problem

$$x = T(x) \stackrel{\text{def}}{=} x - AF(x).$$

The main problem is how to choose  $A$  such that  $T$  is a contraction on some neighborhood of the (unknown) fixed point that we are looking for. In the context of a computer-assisted proof, the approach is the following. First, by your favorite method from scientific computing, choose a finite dimensional 'projection'  $F^{\text{num}}$  of  $F$ , solve it numerically to find  $x^{\text{num}}$ , and reinterpret  $x^{\text{num}}$  as an element of the infinite dimensional space, which we denoted by  $\bar{x}$  (hence  $F^{\text{num}}(x^{\text{num}}) \approx 0$  and  $F(\bar{x}) \approx 0$ ). We expect the solution to be close to  $\bar{x}$ , hence we would like to choose  $A$  an approximate inverse of  $DF(\bar{x})$ .

How to determine when  $A$  is an appropriately accurate approximate inverse? And on which neighborhood of  $\bar{x}$  do we get that  $T$  is a contraction? The following theorem is one way to make this precise. It uses, as an intermediate tool, and approximation  $A^\dagger \in L(X, X')$  of  $DF(\bar{x})$ .

**Theorem 2.1. [Radii polynomial approach in Banach spaces]** *Let  $X$  and  $X'$  be Banach spaces. Denote the norm on  $X$  by  $\|\cdot\|_X$ . Consider bounded linear operators  $A^\dagger \in L(X, X')$  and  $A \in L(X', X)$ . Assume  $F : X \rightarrow X'$  is  $C^1$ , that  $A$  is injective and that*

$$AF : X \rightarrow X. \tag{1}$$

*Consider an approximate solution  $\bar{x}$  of  $F(x) = 0$  (usually obtained using Newton's method on a finite dimensional projection). Let  $Y_0, Z_0, Z_1$ , and  $Z_2$  be positive constants satisfying*

$$\|AF(\bar{x})\|_X \leq Y_0 \tag{2}$$

$$\|I - AA^\dagger\|_{B(X)} \leq Z_0 \tag{3}$$

$$\|A[DF(\bar{x}) - A^\dagger]\|_{B(X)} \leq Z_1, \tag{4}$$

$$\|A[DF(c) - DF(\bar{x})]\|_{B(X)} \leq Z_2 r, \quad \text{for all } \|c - \bar{x}\|_X \leq r, \tag{5}$$

*where  $\|\cdot\|_{B(X)}$  is the induced operator norm for bounded linear operators from  $X$  to itself. Define the radii polynomial*

$$p(r) \stackrel{\text{def}}{=} Z_2 r^2 - (1 - Z_1 - Z_0)r + Y_0. \tag{6}$$

*If there exists  $r_0 > 0$  such that*

$$p(r_0) < 0,$$

*then there exists a unique  $\tilde{x} \in B_{r_0}(\bar{x}) \stackrel{\text{def}}{=} \{x \in X \mid \|x - \bar{x}\|_X \leq r_0\}$  satisfying  $F(\tilde{x}) = 0$ .*

*Proof.* Using the mean value theorem, it is not hard to show that  $T$  maps  $B_{r_0}(\bar{x})$  into itself, and that  $T$  is a contraction on that ball with contraction constant  $\kappa \leq Z_0 + Z_1 + Z_2 r_0 < 1$ . The result then follows directly from the Banach contraction theorem. Although the steps of the proof can be easily filled in by students, we refer to [10] for additional details.  $\square$

This theorem fits in a long tradition of quantitative Newton-Kantorovich type theorems. The bounds are parametrized in terms of  $r$ , so that an appropriate  $r$  does not need to be guessed in advance but rather can be determined as the final step of the proof process. Moreover, it allows us to obtain balls for an interval of values for the radius. Small values of  $r_0$  give us the tightest control on the distance between the solution and the numerical approximation, while large values of  $r_0$  provide us with the best information about the isolation of the solution.

Before we show, in Sections 3 and 4 how to apply Theorem 2.1 in practice in infinite dimensional settings, we first come back to finite dimensional problems, which may also serve as a good initiation for students to start implementing the computer code.

**Remark 2.2.** *In finite dimensions Theorem 2.1 can be implemented very easily and generally, if one has an interval arithmetic package, and if one has explicit formulas for  $D_j F_i$  and  $D_{jk}^2 F_i$  for  $1 \leq i, j, k \leq n$ . We use the  $\infty$ -norm  $|x| = \max_{1 \leq i \leq n} |x_i|$ . Let  $x^{\text{num}}$  be the numerical approximation of a solution (e.g. found using Newton iterations). Let  $A$  be a numerical computed (hence approximate) inverse of the numerically computed  $DF(x^{\text{num}})$ . Now compute with interval arithmetic  $A^\dagger = DF(x^{\text{num}})$ , which implies we may set  $Z_1 = 0$ . Then evaluate, again with interval arithmetic, the following three computable expressions:*

1. the residue

$$Y_0 = \sup \left( \max_{1 \leq i \leq n} |AF(x^{\text{num}})|_i \right),$$

where absolute values are taken component-wise, and  $\sup$  denotes the supremum of the interval obtained.

2. the matrix norm

$$Z_0 = \sup \left( \max_{1 \leq i \leq n} \sum_{1 \leq j \leq n} |(I_n - AA^\dagger)_{ij}| \right).$$

By checking that  $Z_0 < 1$ , one verifies that the hypothesis in Theorem 2.1 that  $A$  is injective.

3. the second derivative estimate (which provides a bound (5) via the mean value theorem)

$$Z_2 = \sup \left( \max_{1 \leq i \leq n} \sum_{1 \leq k, m \leq n} \left| \sum_{1 \leq j \leq n} A_{ij} D_{km}^2 F_j(x^{\text{num}} + [-r_*, r_*]) \right| \right),$$

where  $x^{\text{num}} + [-r_*, r_*]$  is the vector of intervals with components  $[x_k^{\text{num}} - r_*, x_k^{\text{num}} + r_*]$ . Here we choose a loose apriori upper bound  $r_*$  on the value of  $r$ , and we bound the second derivative uniformly on this ball of radius  $r_*$ .

Since  $Y_0$  and  $Z_0$  are usually near machine precision, the quadratic formula then gives a very small  $r_0$  for which  $p(r_0) < 0$ . After checking that  $r_0 \leq r_*$  we can then invoke Theorem 2.1 to prove that a (unique) zero  $x^{\text{sol}}$  of  $F$  lies within distance  $r_0$  of  $x^{\text{num}}$ .

**Example 2.3.** *We consider the circular restricted four body problem (CR4BP), where three bodies (with masses  $(m_1, m_2, m_3)$ , normalized so that  $m_1 + m_2 + m_3 = 1$  and  $m_1 \geq m_2, m_3$ ) move in circular periodic orbits around their center of mass in a triangular configuration that*

is fixed in the co-rotating frame. A fourth massless satellite moves in the effective potential (in this co-rotating frame)

$$\Omega(x, y; m_1, m_2, m_3) \stackrel{\text{def}}{=} \frac{1}{2}(x^2 + y^2) + \sum_{i=1}^3 \frac{m_i}{[(x - x_i)^2 + (y - y_i)^2]^{1/2}}.$$

Here  $(x, y)$  is the position of the satellite in the plane of the triangle, and the fixed positions  $(x_i, y_i)$  of the three bodies can be expressed in terms of their masses:

$$(x_1, y_1) = \left(-\frac{M}{2}, 0\right), \quad (x_2, y_2) = \left(\frac{K_{2,3}}{M}, -\frac{\sqrt{3}m_3}{M}\right), \quad (x_3, y_3) = \left(\frac{K_{3,2}}{M}, \frac{\sqrt{3}m_2}{M}\right),$$

where  $K_{i,j} \stackrel{\text{def}}{=} (m_1 - m_j)m_j + (2m_1 + m_j)m_i$  and  $M \stackrel{\text{def}}{=} 2(m_2^2 + m_2m_3 + m_3^2)^{1/2}$ , see e.g. [4, 6].

The equilibria of the system are given by the critical points of the effective potential  $\Omega$ :

$$F(x, y) \stackrel{\text{def}}{=} \left( x - \sum_{i=1}^3 \frac{m_i(x - x_i)}{[(x - x_i)^2 + (y - y_i)^2]^{3/2}}, y - \sum_{i=1}^3 \frac{m_i(y - y_i)}{[(x - x_i)^2 + (y - y_i)^2]^{3/2}} \right) = 0. \quad (7)$$

It is known that the number of equilibrium points varies from 8 to 10 when the masses are varied (e.g. see [12] and [20]).

Using the general bounds introduced in Remark 2.2 for the finite dimensional case, we applied the radii polynomial approach to prove the existence of several solutions of (7), hence yielding rigorous bounds for relative equilibria of (CR4BP). Let us present a sample result. In case of equal masses  $m_1 = m_2 = m_3 = 1/3$ , the routine `script_equilibria.m`, available at [3], computes (using Newton's method)  $x^{\text{num}} = (-0.467592983336122, 0.809894804400869)$ , and yields the bounds  $Y_0 = 1.775 \times 10^{-15}$ ,  $Z_0 = 1.23 \times 10^{-14}$  and  $Z_2 = 12.6987$  with the choice of  $r_* = 0.02$ . In this case, we obtain  $p(r_0) < 0$  for any  $r_0 \in [1.78 \times 10^{-15}, 0.02]$ , with  $p$  the radii polynomial defined in (6). In Figure 1 one can find several sample results, where each point has been rigorously validated using the computer program.

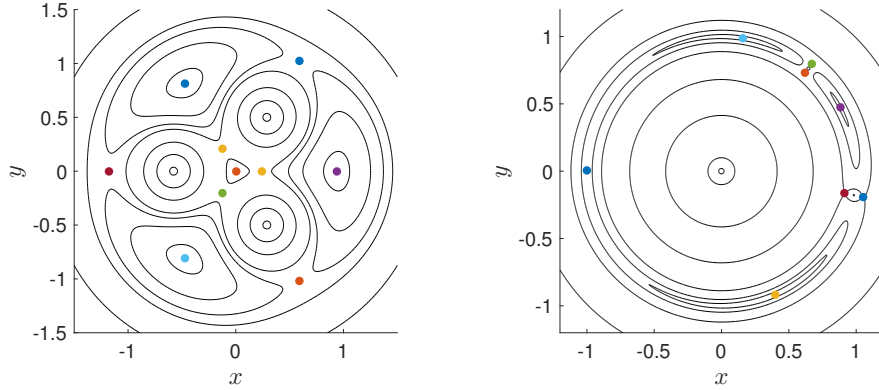


Figure 1: (Left) Ten relative equilibria of (CR4BP) with equal masses. (Right) Eight relative equilibria of (CR4BP) with masses  $m_1 = 0.9987451087$ ,  $m_2 = 0.0010170039$  and  $m_3 = 0.0002378873$ . In both plots, some level sets of the effective potential  $\Omega$  are depicted.

We next present in two relatively simple examples how to apply Theorem 2.1 to infinite dimensional problems.

### 3 Radially symmetric solutions of a nonlinear Laplace-Beltrami operator on the sphere

We consider the partial differential equation

$$\Delta u + \lambda u + u^2 = 0 \quad (8)$$

posed on the sphere  $S^2 \subset \mathbb{R}^3$ , where  $\Delta$  is the Laplace-Beltrami operator on the manifold (this is the natural geometric generalization of the Laplace operator). Here  $\lambda \geq 0$  is a parameter. The PDE (8) describes a classical nonlinear elliptic problem [13], often studied on the unit ball in arbitrary dimension and with a variety of nonlinearities. Here we restrict attention to a quadratic nonlinearity, and we pose the problem on a *sphere*, cf. [8].

Letting  $x = r \cos \phi \sin \theta$ ,  $y = r \sin \phi \sin \theta$  and  $z = r \cos \theta$ , letting  $u(x, y, z) = u(r, \phi, \theta)$ ,

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

We look for solutions of (8) that are radially symmetric (with respect to rotations around the  $z$ -axis) *and* symmetric in the equator. This reduces the PDE to an ODE and leads to the following boundary value problem (BVP) for  $u = u(\theta)$ :

$$\begin{cases} u''(\theta) + \cot(\theta)u'(\theta) + \lambda u(\theta) + u(\theta)^2 = 0, & \text{for } \theta \in (0, \frac{\pi}{2}], \\ u'(0) = 0, \\ u'(\frac{\pi}{2}) = 0. \end{cases} \quad (9)$$

The goal is to prove existence of solutions of (9) via the radii polynomial approach (see Theorem 2.1). The first step in doing so is to introduce a zero finding problem of the form  $F(x) = 0$  on a Banach space.

#### 3.1 The zero finding problem for the Laplace-Beltrami problem

Our approach will be based on Taylor series, and it turns out that it has advantages (see Remark 3.1 below) to work on a domain  $[0, 1]$ . Hence we rescale the independent variable  $\theta = \frac{\pi}{2} \vartheta$ . The algebra is nicer if we also scale the dependent variable, as well as the parameter  $\lambda$ :

$$v(\vartheta) = \frac{\pi^2}{4} u\left(\frac{\pi}{2} \vartheta\right) \quad \text{and} \quad \tilde{\lambda} = \frac{\pi^2}{4} \lambda.$$

The BVP (9) in the new variables becomes

$$\begin{cases} v''(\vartheta) + \frac{\pi}{2} \vartheta \cot\left(\frac{\pi}{2} \vartheta\right) \frac{v'(\vartheta)}{\vartheta} + \tilde{\lambda} v(\vartheta) + v(\vartheta)^2 = 0 & \text{for } \vartheta \in (0, 1], \\ v'(0) = 0, \\ v'(1) = 0. \end{cases} \quad (10)$$

For the sake of presentation, here we have anticipated that it will be convenient to split off a factor  $\vartheta$  in the second term of the differential equation, as it will allow us to deal with smooth functions that are even in  $\vartheta$  only. We search for  $v$  as a power series of  $\vartheta$  around zero:  $v(\vartheta) = \sum_{n=0}^{\infty} a_n \vartheta^n$ . Let us assume for the moment that the radius of convergence of our power

series is larger than 1, then the coefficients are in a Banach space of geometrically decaying coefficients. More precisely, for  $\nu > 1$ , denote

$$\ell_\nu^1 \stackrel{\text{def}}{=} \left\{ a = (a_n)_{n \geq 0} : \|a\|_{1,\nu} \stackrel{\text{def}}{=} \sum_{n \geq 0} |a_n| \nu^n < \infty \right\}.$$

Given  $a, c \in \ell_\nu^1$ , denote by  $a * c$  the Cauchy product given component-wise by

$$(a * c)_n = \sum_{\substack{n_1 + n_2 = n \\ 0 \leq n_1, n_2 \leq n}} a_{n_1} c_{n_2} = \sum_{n_1=0}^n a_{n_1} c_{n-n_1}.$$

An important property of the space  $\ell_\nu^1$  is that it is a Banach algebra under the Cauchy product:

$$\|a * c\|_{1,\nu} \leq \|a\|_{1,\nu} \|c\|_{1,\nu}. \quad (11)$$

It is instructive for students to check this property.

The expansion for the cotangent is

$$\cot(\theta) = \frac{1}{\theta} - 2 \sum_{n=1}^{\infty} \frac{\theta^{2n-1}}{\pi^{2n}} \zeta(2n), \quad \text{where } \zeta(2n) \stackrel{\text{def}}{=} \sum_{k=1}^{\infty} \frac{1}{k^{2n}} \text{ is the Riemann zeta function.}$$

Hence we get

$$\frac{\pi}{2} \vartheta \cot\left(\frac{\pi}{2} \vartheta\right) = 1 - 2 \sum_{n=1}^{\infty} \frac{\zeta(2n)}{2^{2n}} \vartheta^{2n} = \sum_{j=0}^{\infty} b_j \vartheta^j,$$

where the  $b_j$  are the Taylor coefficients defined by

$$b_0 = 1, \quad b_j = -2 \frac{\zeta(j)}{2^j} \quad \text{if } j \geq 1 \text{ is even,} \quad b_j = 0 \quad \text{if } j \text{ is odd.}$$

The decay rate of these coefficients shows that we should restrict attention to  $1 < \nu < 2$ . After expanding all terms in (10) as Taylor series, using the Cauchy product, and equating powers, we arrive at the operator  $F(a) = (F_n(a))_{n \geq 0}$  defined as

$$F_n(a) \stackrel{\text{def}}{=} \begin{cases} a_1 & \text{for } n = 0, \\ \sum_{j=1}^{\infty} j a_j & \text{for } n = 1, \\ n(n-1)a_n + (Ja * b)_n + \tilde{\lambda} a_{n-2} + (a * a)_{n-2} & \text{for } n \geq 2. \end{cases} \quad (12)$$

Here the multiplication operator  $J$  on sequence spaces is defined by

$$(Ja)_j \stackrel{\text{def}}{=} j a_j \quad \text{for } j \geq 0.$$

**Remark 3.1.** *Without the rescaling of the independent variable, the formula for  $F_1$  corresponding to the boundary condition at  $\theta = \frac{\pi}{2}$  would have become  $\sum_{j \geq 1} j a_j (\frac{\pi}{2})^j$ . Evaluating this sum (say up to some finite  $N$ ) is numerically unstable because of the high powers of  $\frac{\pi}{2}$ . Hence the rescaling.*

The following result expresses that to solving  $F(a) = 0$  leads to a solution of (10).

**Lemma 3.2.** *Let  $\nu \in (1, 2)$  and let  $a = (a_n)_{n \geq 0} \in \ell_\nu^1$  be such that  $F_n(a) = 0$  for all  $n \geq 0$ . Then  $v(\vartheta) = \sum_{n \geq 0} a_n \vartheta^n$  is a solution of (10).*

*Proof.* Let  $\nu \in (1, 2)$  and  $a \in \ell_\nu^1$ . Then the power series  $v(\vartheta) = \sum_{n \geq 0} a_n \vartheta^n$  converges uniformly for  $\vartheta \in [0, 1]$ , and similarly for the derivatives of  $v$ . The first two equations  $F_0(a) = 0$  and  $F_1(a) = 0$  imply that  $v$  satisfies the boundary conditions in (10), whereas the remaining equations  $F_n(a) = 0$ ,  $n \geq 2$  imply that  $v$  satisfies the differential equation in (10).  $\square$

Now that we have identified the zero finding problem  $F(a) = 0$  to be solved in the Banach space  $X = \ell_\nu^1$  with  $\nu \in (1, 2)$  to be chosen later, we are ready to apply the radii polynomial approach as introduced in Theorem 2.1. The first ingredient is a numerical approximation  $\bar{a}$  of  $F(a) = 0$ . Given  $N \in \mathbb{N}$  and  $a = (a_n)_{n \geq 0} \in X = \ell_\nu^1$ , denote by  $a^{(N)} = (a_n)_{n=0}^N \in \mathbb{R}^{N+1}$  the finite dimensional projection of  $a$ , and by  $F^{(N)} : \mathbb{R}^{N+1} \rightarrow \mathbb{R}^{N+1}$  the finite dimensional projection of  $F$  defined by

$$F^{(N)}(a^{(N)}) \stackrel{\text{def}}{=} (F_n(a^{(N)}, 0, 0, 0, \dots))_{0 \leq n \leq N}.$$

Assume that a numerical approximation  $\bar{a}^{(N)} \in \mathbb{R}^{N+1}$  has been computed. We abuse slightly the notation by identifying  $\bar{a}^{(N)} \in \mathbb{R}^{N+1}$  with  $\bar{a} = (\bar{a}_0, \bar{a}_1, \dots, \bar{a}_N, 0, 0, 0, \dots) \in X = \ell_\nu^1$ . Denote by  $DF^{(N)}(\bar{a})$  the Jacobian of  $F^{(N)}$  at  $\bar{a}$ . The radii polynomial approach as introduced in Theorem 2.1 requires defining the operators  $A^\dagger$  and  $A$ . Let

$$(A^\dagger h)_n \stackrel{\text{def}}{=} \begin{cases} (DF^{(N)}(\bar{a})h^{(N)})_n & \text{for } 0 \leq n \leq N, \\ n^2 h_n & \text{for } n > N, \end{cases}$$

where the diagonal tail is chosen in view of the dominant term in (12) for large  $n$ . Consider an  $(N+1) \times (N+1)$  matrix  $A^{(N)}$  computed so that  $A^{(N)} \approx DF^{(N)}(\bar{a})^{-1}$ . Define  $A$  as

$$(Ah)_n \stackrel{\text{def}}{=} \begin{cases} (A^{(N)}h^{(N)})_n & \text{for } 0 \leq n \leq N, \\ n^{-2} h_n & \text{for } n > N. \end{cases} \quad (13)$$

One way to visualize the operator  $A$  is as

$$A = \begin{bmatrix} A^{(N)} & \mathbf{0} & \mathbf{0} & \dots \\ \mathbf{0} & \frac{1}{(N+1)^2} & 0 & \dots \\ \mathbf{0} & 0 & \frac{1}{(N+2)^2} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

The following lemma states that condition (1) of Theorem 2.1 holds.

**Lemma 3.3.** *Let  $X = \ell_\nu^1$  with  $\nu \in (1, 2)$ ,  $A$  as in (13) and  $F$  as in (12). Then  $AF : X \rightarrow X$ .*

We leave the proof to the students. We now introduce the formulas for the bounds  $Y_0$ ,  $Z_0$ ,  $Z_1$  and  $Z_2$  satisfying respectively (2), (3), (4) and (5). These bounds are derived in Sections 3.2–3.5. In our experience, students can fill in the details with some effort, as well as grasp the logical structure of the arguments, and sometimes even sharpen the bounds. We first introduce an auxiliary result, used for the  $Z_0$  and  $Z_2$  bounds. We again leave the proof to the students.

**Lemma 3.4.** *Consider a linear operator  $Q : \ell_\nu^1 \rightarrow \ell_\nu^1$  of the form*

$$Q = \begin{bmatrix} Q^{(N)} & & & 0 \\ & q_{N+1} & & \\ & & q_{N+2} & \\ 0 & & & \ddots \end{bmatrix}$$



where  $Q^{(N)} = (Q_{m,n}^{(N)})_{0 \leq m,n \leq N}$  and  $q_n \in \mathbb{R}$ . Assume that  $|q|_\infty = \sup_{n \geq N} |q_n| < \infty$ . Then

$$\|Q\|_{B(\ell_\nu^1)} = \max \left( \max_{0 \leq n \leq N} \frac{1}{\nu^n} \sum_{m=0}^N |Q_{m,n}^{(N)}| \nu^m, |q|_\infty \right).$$

### 3.2 The $Y_0$ bound

We look for a bound  $Y_0$  satisfying  $\|AF(\bar{a})\|_{1,\nu} \leq Y_0$ . We note that since  $\bar{a}_n = 0$  for  $n > N$ , we have  $(\bar{a} * \bar{a})_n = 0$  for all  $n > 2N$ . Hence, recalling the definition of  $A$  in (13),

$$\begin{aligned} \|AF(\bar{a})\|_{1,\nu} &\leq \sum_{n=0}^N \left| (A^{(N)} F^{(N)}(\bar{a}))_n \right| \nu^n + \sum_{n=N+1}^{2N+2} \frac{1}{n^2} |(J\bar{a} * b)_n + (\bar{a} * \bar{a})_n| \nu^n \\ &\quad + \sum_{n=2N+3}^{\infty} \frac{1}{n^2} |(J\bar{a} * b)_n| \nu^n. \end{aligned}$$

When calculating the finite sums in this expression, computing any  $b_l$  involves evaluating the zeta function, which is itself an infinite series. It is a useful exercise for students to bound the tail of this series via an (integral) estimate. Concerning the final term in the expression above, since  $\zeta(s) = \sum_{k=1}^{\infty} k^{-s} \leq \sum_{k=1}^{\infty} k^{-2} = \frac{\pi^2}{6}$  for all  $s \geq 2$ , we have  $|b_l| \leq \frac{4}{3} 2^{-l}$  for all  $l \geq 1$ . Hence

$$\begin{aligned} \sum_{k=2N+3}^{\infty} \frac{1}{k^2} |(J\bar{a} * b)_k| \nu^k &= \sum_{k=2N+3}^{\infty} \left( \frac{1}{k^2} \sum_{j=0}^N |j\bar{a}_j b_{k-j}| \nu^k \right) \\ &\leq \frac{1}{(2N+3)^2} \sum_{j=0}^N \left( |j\bar{a}_j| \nu^j \sum_{k=2N+3}^{\infty} |b_{k-j}| \nu^{k-j} \right) \\ &\leq \frac{1}{(2N+3)^2} \sum_{j=0}^N |j\bar{a}_j| \nu^j \sum_{l=N+3}^{\infty} |b_l| \nu^l \\ &\leq \frac{4\|J\bar{a}\|_{1,\nu}}{3(2N+3)^2} \sum_{l=N+3}^{\infty} \left( \frac{\nu}{2} \right)^l = \frac{4\|J\bar{a}\|_{1,\nu}}{3(2N+3)^2} \left( \frac{\nu}{2} \right)^{N+3} \frac{1}{1 - \frac{\nu}{2}} \stackrel{\text{def}}{=} Y_{\text{tail}}. \end{aligned}$$

We thus set

$$Y_0 \stackrel{\text{def}}{=} \sum_{n=0}^N \left| (A^{(N)} F^{(N)}(\bar{a}))_n \right| \nu^n + \sum_{n=N+1}^{2N+2} \frac{1}{n^2} |(J\bar{a} * b)_n + (\bar{a} * \bar{a})_n| \nu^n + Y_{\text{tail}}. \quad (14)$$

### 3.3 The $Z_0$ bound

Let  $B \stackrel{\text{def}}{=} I - AA^\dagger$ . We remark that the tails of  $A$  and  $A^\dagger$  are exact inverses, hence  $B_{m,n} = 0$  when  $m > N$  or  $n > N$ . Letting

$$Z_0 \stackrel{\text{def}}{=} \max_{0 \leq n \leq N} \frac{1}{\nu^n} \sum_{0 \leq m \leq N} |B_{m,n}| \nu^m, \quad (15)$$

we get from Lemma 3.4 that  $\|I - AA^\dagger\|_{B(\ell_\nu^1)} \leq Z_0$ .

### 3.4 The $Z_1$ bound

We look for a bound  $\|A[DF(\bar{a}) - A^\dagger]\|_{B(\ell_\nu^1)} \leq Z_1$ . Given  $h \in \ell_\nu^1$  with  $\|h\|_{1,\nu} \leq 1$ , we set  $z \stackrel{\text{def}}{=} [DF(\bar{a}) - A^\dagger]h$ . For the finite part ( $0 \leq n \leq N$ ) we see that

$$z_0 = 0, \quad z_1 = \sum_{j \geq N+1} j h_j, \quad \text{and} \quad z_n = 0 \quad \text{for } n = 2, \dots, N.$$

For the tail, i.e.  $n > N$ , we find

$$z_n = n(n-1)h_n + (Jh * b)_n + \tilde{\lambda}h_{n-2} + 2(\bar{a} * h)_{n-2} - n^2h_n.$$

Using the “chain rule” identity  $J(h * b) = (Jh * b) + (h * Jb)$  we obtain

$$z_n = n(\tilde{b} * h)_n - (Jb * h)_n + \tilde{\lambda}h_{n-2} + 2(\bar{a} * h)_{n-2},$$

where  $\tilde{b}$  is the sequence defined by  $\tilde{b}_0 = 0$  and  $\tilde{b}_n = b_n$  for  $n \geq 1$ . Next we estimate

$$\|Az\|_{1,\nu} = \sum_{k=0}^N \left| \sum_{j=0}^N A_{k,j}^{(N)} z_j \right| \nu^k + \sum_{k \geq N+1} \frac{1}{k^2} |z_k| \nu^k. \quad (16)$$

For any  $\|h\|_{1,\nu} \leq 1$ , it is a small exercise to show that  $|z_1| \leq \frac{N+1}{\nu^{N+1}}$  provided  $N+1 \geq (\ln \nu)^{-1}$ . Hence the first term in (16) is bounded by  $\frac{N+1}{\nu^{N+1}} \sum_{k=0}^N |A_{k,1}^{(N)}|$ . For the infinite tail we first estimate

$$\begin{aligned} \|\tilde{b}\|_{1,\nu} &= \sum_{l \geq 2} |b_l| \nu^l \leq \frac{4}{3} \sum_{n=1}^{\infty} \left( \frac{\nu^2}{4} \right)^n = \frac{4\nu^2}{3(4-\nu^2)} \stackrel{\text{def}}{=} C_1, \\ \|Jb\|_{1,\nu} &= \sum_{l \geq 2} l |b_l| \nu^l \leq \frac{8}{3} \sum_{n=1}^{\infty} n \left( \frac{\nu^2}{4} \right)^n = \frac{16\nu^2}{3(4-\nu^2)^2} \stackrel{\text{def}}{=} C_2. \end{aligned}$$

Then, using the Banach algebra property (11), we find, for any  $\|h\|_{1,\nu} \leq 1$ ,

$$\sum_{k \geq N+1} \frac{1}{k^2} |z_k| \nu^k \leq \frac{C_1}{N+1} + \frac{C_2}{(N+1)^2} + \frac{\nu^2}{(N+1)^2} (|\tilde{\lambda}| + 2\|\bar{a}\|_{1,\nu}).$$

Hence, with the requirement that  $N+1 \geq (\ln \nu)^{-1}$ , we set

$$Z_1 \stackrel{\text{def}}{=} \frac{N+1}{\nu^{N+1}} \sum_{k=0}^N |A_{k,1}^{(N)}| + \frac{C_1}{N+1} + \frac{C_2}{(N+1)^2} + \frac{\nu^2}{(N+1)^2} (|\tilde{\lambda}| + 2\|\bar{a}\|_{1,\nu}). \quad (17)$$

### 3.5 The $Z_2$ bound

Let  $c \in B_r(\bar{a})$ , that is  $\|c - \bar{a}\|_{1,\nu} \leq r$ . Given  $h \in B_1(0)$ , note that

$$([DF(c) - DF(\bar{a})]h)_n = \begin{cases} 0 & \text{for } n = 0, 1, \\ 2((c - \bar{a}) * h)_{n-2} & \text{for } n \geq 2, \end{cases}$$

so that

$$\begin{aligned} \|A[DF(c) - DF(\bar{a})]\|_{B(\ell_\nu^1)} &= \sup_{h \in B_1(0)} \|A[DF(c) - DF(\bar{a})]h\|_{1,\nu} \leq 2\nu^2 \sup_{h \in B_1(0)} \|A[(c - \bar{a}) * h]\|_{1,\nu} \\ &\leq 2\nu^2 \sup_{h \in B_1(0)} \|A\|_{B(\ell_\nu^1)} \|c - \bar{a}\|_{1,\nu} \|h\|_{1,\nu} \leq 2\nu^2 \|A\|_{B(\ell_\nu^1)} r. \end{aligned}$$

Using Lemma 3.4, we get that

$$2\nu^2 \|A\|_{B(\ell_\nu^1)} \leq Z_2 \stackrel{\text{def}}{=} 2\nu^2 \max \left( \max_{0 \leq \ell \leq N} \frac{1}{\nu^\ell} \sum_{0 \leq k \leq N} |A_{k,\ell}^{(N)}| \nu^k, \frac{1}{(N+1)^2} \right). \quad (18)$$

### 3.6 Computer-assisted proofs

Combining the bounds (14), (15), (17) and (18), we define the radii polynomial  $p(r)$  as in (6). We prove the existence of three different types of solutions by verifying the hypothesis of Theorem 2.1 with the routine `script_three_proofs_LB.m` available at [3]. The data of each proof can be found in the following table. We find that  $p(r) < 0$  for  $r \in [r_{\min}, r_{\max}]$ .

solution	#1	#2	#3
$\lambda = 4\tilde{\lambda}/\pi^2$	5.67	20	19.961
$\nu$	1.06	1.04	1.058
$N$	250	410	360
$Y_0$	$5.2768 \cdot 10^{-9}$	$3.1552 \cdot 10^{-7}$	$4.4322 \cdot 10^{-7}$
$Z_0$	$5.6260 \cdot 10^{-10}$	$3.6934 \cdot 10^{-10}$	$2.1306 \cdot 10^{-8}$
$Z_1$	$4.1813 \cdot 10^{-1}$	$2.5523 \cdot 10^{-2}$	$1.085 \cdot 10^{-1}$
$Z_2$	$7.6773 \cdot 10^3$	$4.8967 \cdot 10^3$	$3.90306 \cdot 10^6$
$r_{\min}$	$9.07 \cdot 10^{-9}$	$3.2432 \cdot 10^{-7}$	$7.3132 \cdot 10^{-7}$
$r_{\max}$	$7.578 \cdot 10^{-5}$	$1.9868 \cdot 10^{-4}$	$1.5527 \cdot 10^{-6}$

The three solutions can be seen in Figures 2, 3 and 4. One can observe the different qualitative behaviour of these three solutions.

It is not entirely straightforward to find initial guesses for solutions, i.e., starting points for applying Newton's method to the finite truncation  $F^{(N)}$ . To find such approximate solutions, note that the trivial solution  $v = 0$  undergoes transcritical bifurcations at  $\tilde{\lambda} = \pi^2 n(n + \frac{1}{2})$  for  $n \in \mathbb{N}$ . The solution branches that bifurcate at these parameter values can then be followed numerically (using branch following techniques) to other parameter values. The three steps of doing the bifurcation analysis, then finding small nontrivial solutions near the bifurcation point, and finally continuing these using a predictor-corrector algorithm, are all interesting educational avenues for students to pursue.

The main restriction on the Taylor series based approach presented here, is that not all solutions have an analytic extension to the complex ball of radius 1. These solutions, although real analytic on  $[0, 1]$ , cannot be described by a single Taylor series around the origin. Again, this is instructive for students to understand. One way to overcome this is via domain decomposition (i.e. matching together several power series), but we will not pursue that here, and it is also by no means the only option.

## 4 Radially symmetric equilibria of the Swift-Hohenberg equation on the 3D unit ball

We consider the Swift-Hohenberg equation [21] with Dirichlet boundary conditions:

$$\begin{cases} u_t = -(\Delta - 1)^2 u + \lambda u - u^3, & \text{on } D_1, \\ u = \Delta u = 0 & \text{on } \partial D_1. \end{cases} \quad (19)$$

Here  $D_1 \subset \mathbb{R}^3$  is the unit ball, and  $\lambda \in \mathbb{R}$  is a parameter. The parabolic PDE (19) is a popular deterministic model for pattern formation, see e.g. [14]. It has been well studied, analytically in one spatial dimension and predominantly numerically in two spatial dimensions. Here we consider time-independent solutions in three spatial dimensions. Indeed, we will focus on radially symmetric equilibrium solutions of (19). Letting  $v = (\Delta - 1)u$ , these solutions also

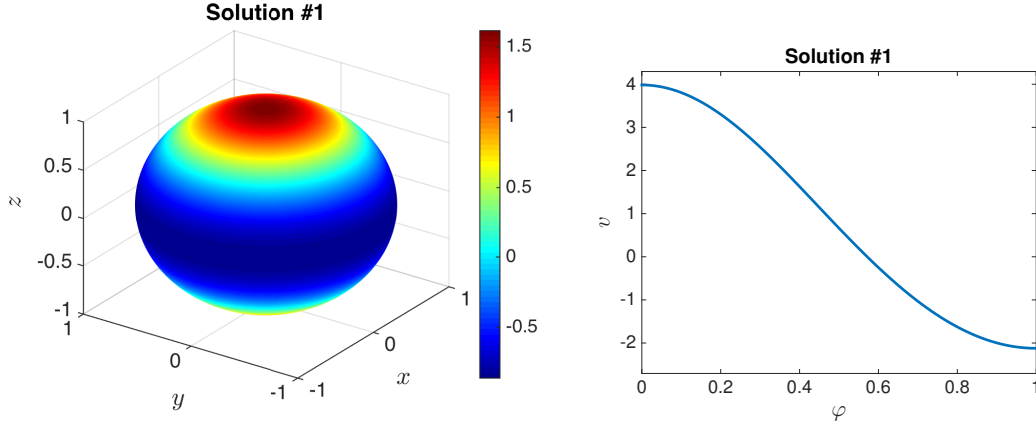


Figure 2: (Left) The first solution of (8) on the unit sphere  $S^2 \subset \mathbb{R}^3$ . (Right) The corresponding (numerical) solution of the BVP (10). Since  $r_{\min} < 10^{-8}$ , the true solution lies with the line-width by Theorem 2.1.

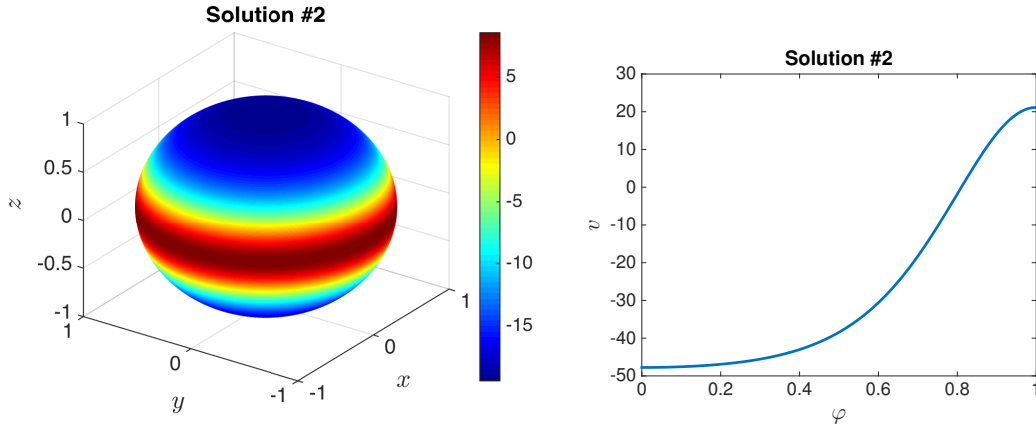


Figure 3: (Left) The second solution of (8) on the unit sphere  $S^2 \subset \mathbb{R}^3$ . (Right) The corresponding (numerical) solution of the BVP (10).

correspond to radially symmetric equilibria of the reaction-diffusion system

$$\begin{cases} u_t = \Delta u - u - v \\ v_t = \Delta v - v - \lambda u + u^3 \end{cases} \quad (20)$$

on the unit ball with Dirichlet boundary conditions. The method works very generally for radially symmetric equilibrium solutions in reaction-diffusion systems (cf. [19]), which are ubiquitous in models in the life sciences. This motivates us to work with the system (20) rather than with the equivalent (at the level of equilibria) scalar equation (19). As an additional benefit, the analysis below illustrates how the method of radii polynomials extends naturally to systems of equations (a fresh challenge for the students, with some new coding hurdles). Looking for radially symmetric equilibria of (20), i.e., time independent solutions of the form

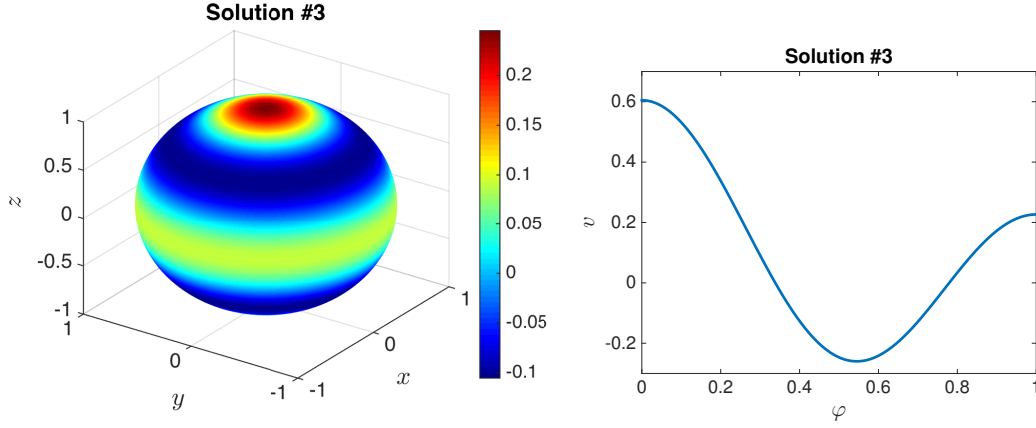


Figure 4: (Left) The third solution of (8) on the unit sphere  $S^2 \subset \mathbb{R}^3$ . (Right) The corresponding (numerical) solution of the BVP (10).

$u(x, y, z) = u(s) = u(\sqrt{x^2 + y^2 + z^2})$ , leads to a coupled systems of ODEs:

$$\begin{cases} u''(s) + \frac{2}{s}u'(s) - u(s) - v(s) = 0 & \text{for } s \in (0, 1], \\ v''(s) + \frac{2}{s}v'(s) - v(s) - \lambda u(s) + u(s)^3 = 0 & \text{for } s \in (0, 1], \\ u'(0) = v'(0) = 0, \\ u(1) = v(1) = 0. \end{cases} \quad (21)$$

We expand the functions  $u$  and  $v$  as power series  $u(s) = \sum_{n=0}^{\infty} a_n s^n$  and  $v(s) = \sum_{n=0}^{\infty} b_n s^n$ . Define the coefficient sequences as  $a = (a_n)_{n \geq 0}$  and  $b = (b_n)_{n \geq 0}$ . Consider the Banach space

$$X = \ell_{\nu}^1 \times \ell_{\nu}^1 = \{x = (a, b) : a, b \in \ell_{\nu}^1\},$$

endowed with the norm  $\|x\|_X = \max\{\|a\|_{1,\nu}, \|b\|_{1,\nu}\}$ . The equations for the Taylor coefficients are  $F_1(x) = ((F_1(x))_n)_{n \geq 0} = 0$  and  $F_2(x) = ((F_2(x))_n)_{n \geq 0} = 0$ , given component-wise by

$$\begin{aligned} (F_1(x))_n &= \begin{cases} a_1 & \text{for } n = 0, \\ \sum_{k \geq 0} a_k & \text{for } n = 1, \\ n(n+1)a_n - a_{n-2} - b_{n-2} & \text{for } n \geq 2, \end{cases} \\ (F_2(x))_n &= \begin{cases} b_1 & \text{for } n = 0, \\ \sum_{k \geq 0} b_k & \text{for } n = 1, \\ n(n+1)b_n - b_{n-2} - \lambda a_{n-2} + (a^3)_{n-2} & \text{for } n \geq 2, \end{cases} \end{aligned}$$

where

$$(a^3)_n = (a * a * a)_n = \sum_{\substack{n_1+n_2+n_3=n \\ n_1, n_2, n_3 \geq 0}} a_{n_1} a_{n_2} a_{n_3} = \sum_{n_1=0}^n a_{n_1} \left( \sum_{n_2=0}^{n-n_1} a_{n_2} a_{n-n_1-n_2} \right).$$

It is not difficult to derive that all odd coefficients will vanish, but we do not exploit that here.

Denoting  $F = (F_1, F_2)$ , the problem is to find  $x = (a, b) \in X = \ell_{\nu}^1 \times \ell_{\nu}^1$  for some  $\nu > 1$  such that  $F(x) = 0$ . To achieve this, we use the radii polynomial approach as introduced in

Theorem 2.1. Given  $N \in \mathbb{N}$ , denote  $x^{(N)} = ((a_n)_{0 \leq n \leq N}, (b_n)_{0 \leq n \leq N}) \in \mathbb{R}^{2N+2}$ . Consider the finite dimensional projection  $F^{(N)} = (F_1^{(N)}, F_2^{(N)}) : \mathbb{R}^{2N+2} \rightarrow \mathbb{R}^{2N+2}$  defined by

$$F_i^{(N)}(x^{(N)}) \stackrel{\text{def}}{=} ((F_i(x^{(N)}))_k)_{0 \leq k \leq N}.$$

Given  $\bar{x} \in \mathbb{R}^{2N+2}$  a numerical approximation of  $F^{(N)}(x) = 0$ , denote by  $DF^{(N)}(\bar{x})$  the Jacobian of  $F^{(N)}$  at  $\bar{x}$ , and let us write it as

$$DF^{(N)}(\bar{x}) = \begin{pmatrix} D_a F_1^{(N)}(\bar{x}) & D_b F_1^{(N)}(\bar{x}) \\ D_a F_2^{(N)}(\bar{x}) & D_b F_2^{(N)}(\bar{x}) \end{pmatrix} \in \mathbb{R}^{(2N+2) \times (2N+2)}.$$

The radii polynomial approach requires defining the operators  $A^\dagger$  and  $A$ . Let

$$A^\dagger = \begin{pmatrix} A_{1,1}^\dagger & A_{1,2}^\dagger \\ A_{2,1}^\dagger & A_{2,2}^\dagger \end{pmatrix}, \quad (22)$$

whose action on an element  $h = (h_1, h_2) \in X$  is defined by  $(A^\dagger h)_i = A_{i,1}^\dagger h_1 + A_{i,2}^\dagger h_2$ , for  $i = 1, 2$ . Here the action of  $A_{i,j}^\dagger$  is defined as

$$(A_{i,1}^\dagger h_1)_n = \begin{cases} (D_a F_i^{(N)}(\bar{x}) h_1^{(N)})_n & \text{for } 0 \leq n \leq N, \\ \delta_{i,1} n(n+1) (h_1)_n & \text{for } n > N, \end{cases}$$

$$(A_{i,2}^\dagger h_2)_n = \begin{cases} (D_b F_i^{(N)}(\bar{x}) h_2^{(N)})_n & \text{for } 0 \leq n \leq N, \\ \delta_{i,2} n(n+1) (h_2)_n & \text{for } n > N, \end{cases}$$

where  $\delta_{i,j}$  is the Kronecker  $\delta$ . Consider now a matrix  $A^{(N)} \in \mathbb{R}^{(2N+2) \times (2N+2)}$  computed so that  $A^{(N)} \approx DF^{(N)}(\bar{x})^{-1}$ . We decompose it into four  $(N+1) \times (N+1)$  blocks:

$$A^{(N)} = \begin{pmatrix} A_{1,1}^{(N)} & A_{1,2}^{(N)} \\ A_{2,1}^{(N)} & A_{2,2}^{(N)} \end{pmatrix}.$$

Thus we define  $A$  as

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}, \quad (23)$$

whose action on an element  $h = (h_1, h_2) \in X$  is defined by  $(Ah)_i = A_{i,1} h_1 + A_{i,2} h_2$ , for  $i = 1, 2$ . The action of  $A_{i,j}$  is defined as

$$(A_{i,j} h_j)_n = \begin{cases} (A_{i,j}^{(N)} h_j^{(N)})_n & \text{for } 0 \leq n \leq N, \\ \delta_{i,j} \frac{1}{n(n+1)} (h_j)_n & \text{for } n > N. \end{cases}$$

Finally, we set  $T(x) = x - AF(x)$ , which indeed maps  $X$  into itself. As in Section 3, the next step is to derive explicit, computable expressions for the bounds (2), (3), (4) and (5).

#### 4.1 The $Y_0$ bound

Observe first that the nonlinear term of  $F_2(\bar{x})$  is the Cauchy product  $(\bar{a} * \bar{a} * \bar{a})_{n-2}$ , which vanishes for  $n \geq 3N+3$ . This implies that  $(F_1(\bar{x}))_n = (F_2(\bar{x}))_n = 0$  for all  $n \geq 3N+3$ . For  $i = 1, 2$ , we set

$$Y_0^{(i)} \stackrel{\text{def}}{=} \sum_{n=0}^N \left| \sum_{j=1}^2 \left( A_{i,j}^{(N)} F_j^{(N)}(\bar{x}) \right)_n \right| \nu^n + \sum_{n=N+1}^{3N+2} \left| \frac{1}{n(n+1)} (F_i(\bar{x}))_n \right| \nu^n,$$

which is a collection of finite sums that can be evaluated with interval arithmetic. We get

$$\| [T(\bar{x}) - \bar{x}]_i \|_{1,\nu} = \| [-AF(\bar{x})]_i \|_{1,\nu} = \left\| \sum_{j=1}^n A_{i,j} F_j(\bar{x}) \right\|_{1,\nu} \leq Y_0^{(i)},$$

and we set

$$Y_0 \stackrel{\text{def}}{=} \max \left( Y_0^{(1)}, Y_0^{(2)} \right). \quad (24)$$

## 4.2 The $Z_0$ bound

We look for a bound of the form  $\|I - AA^\dagger\|_{B(X)} \leq Z_0$ . Recalling the definitions of  $A$  and  $A^\dagger$  given in (23) and (22), let  $B \stackrel{\text{def}}{=} I - AA^\dagger$  the bounded linear operator represented as

$$B = \begin{pmatrix} B_{1,1} & B_{1,2} \\ B_{2,1} & B_{2,2} \end{pmatrix}.$$

We remark that  $(B_{i,j})_{n_1, n_2} = 0$  for any  $i, j = 1, 2$  whenever  $n_1 > N$  or  $n_2 > N$ . Hence we can compute the norms  $\|B_{i,j}\|_{B(\ell_\nu^1)}$  using Lemma 3.4. Given  $h = (h_1, h_2) \in X = \ell_\nu^1 \times \ell_\nu^1$  with  $\|h\|_X = \max(\|h_1\|_{1,\nu}, \|h_2\|_{1,\nu}) \leq 1$ , we get

$$\|(Bh)_i\|_{1,\nu} = \left\| \sum_{j=1}^2 B_{i,j} h_j \right\|_{1,\nu} \leq \sum_{i=1}^2 \|B_{i,j}\|_{B(\ell_\nu^1)}.$$

Hence we define

$$Z_0 \stackrel{\text{def}}{=} \max \left( \|B_{1,1}\|_{B(\ell_\nu^1)} + \|B_{1,2}\|_{B(\ell_\nu^1)}, \|B_{2,1}\|_{B(\ell_\nu^1)} + \|B_{2,2}\|_{B(\ell_\nu^1)} \right). \quad (25)$$

## 4.3 The $Z_1$ bound

Recall that we look for the bound  $\|A[DF(\bar{x}) - A^\dagger]\|_{B(X)} \leq Z_1$ . Given  $h = (h_1, h_2) \in X$  with  $\|h\|_X \leq 1$ , set

$$z \stackrel{\text{def}}{=} [DF(\bar{x}) - A^\dagger]h.$$

Note that for  $j = 1, 2$ ,  $(z_j)_0 = 0$ ,  $(z_j)_1 = \sum_{k \geq N+1} (h_j)_k$ ,  $(z_j)_n = 0$  for  $n = 2, \dots, N$ , and

$$(z_1)_n = -(h_1)_{n-2} - (h_2)_{n-2}, \quad \text{and} \quad (z_2)_n = -(h_2)_{n-2} - \lambda(h_1)_{n-2} + 3(\bar{a} * \bar{a} * h_1)_{n-2},$$

for  $n \geq N+1$ . It is not hard to show that  $|(z_j)_1| \leq \nu^{-(N+1)}$  for all  $\|h\|_X \leq 1$ ,  $j = 1, 2$ , hence

$$\begin{aligned} \|(Az)_1\|_{1,\nu} &\leq \sum_{j=1}^2 \|A_{1,j} z_j\|_{1,\nu} \\ &= \sum_{j=1}^2 \sum_{n=0}^N |(A_{1,j}^{(N)} z_j^{(N)})_n| \nu^n + \sum_{n \geq N+1} \frac{1}{n(n+1)} |(z_1)_n| \nu^n \\ &\leq \frac{1}{\nu^{N+1}} \sum_{j=1}^2 \sum_{n=0}^N |(A_{1,j}^{(N)})_{n,1}| \nu^n + \frac{1}{(N+1)(N+2)} (\|h_1\|_{1,\nu} + \|h_2\|_{1,\nu}) \\ &\leq \frac{1}{\nu^{N+1}} \sum_{j=1}^2 \sum_{n=0}^N |(A_{1,j}^{(N)})_{n,1}| \nu^n + \frac{2}{(N+1)(N+2)} \stackrel{\text{def}}{=} Z_1^{(1)}, \end{aligned}$$

and similarly, now also using the Banach algebra property (11),

$$\|(Az)_2\|_{1,\nu} \leq \frac{1}{\nu^{N+1}} \sum_{j=1}^2 \sum_{n=0}^N |(A_{2,j}^{(N)})_{n,1}| \nu^n + \frac{1}{(N+1)(N+2)} (1 + \lambda + 3\|\bar{a}\|_{1,\nu}^2) \stackrel{\text{def}}{=} Z_1^{(2)}.$$

We thus define

$$Z_1 \stackrel{\text{def}}{=} \max \left( Z_1^{(1)}, Z_1^{(2)} \right). \quad (26)$$

#### 4.4 The $Z_2$ bound

Let  $c = (c_1, c_2) \in B_r(\bar{x})$ , that is  $\|c - \bar{x}\|_X = \max(\|c_1 - \bar{a}\|_{1,\nu}, \|c_2 - \bar{b}\|_{1,\nu}) \leq r$ . Given  $\|h\|_X \leq 1$ , note that  $([DF_1(c) - DF_1(\bar{x})]h)_n = 0$  and that

$$([DF_2(c) - DF_2(\bar{x})]h)_n = \begin{cases} 0 & \text{for } n = 0, 1, \\ 3((c_1 * c_1 - \bar{a} * \bar{a}) * h_1)_{n-2} & \text{for } n \geq 2, \end{cases}$$

so that

$$\begin{aligned} \|A[DF(c) - DF(\bar{x})]\|_{B(X)} &= \sup_{\|h\|_X \leq 1} \|A[DF(c) - DF(\bar{x})]h\|_X \\ &\leq \|A\|_{B(X)} \sup_{\|h\|_X \leq 1} \|[DF(c) - DF(\bar{x})]h\|_X \\ &= 3\nu^2 \|A\|_{B(X)} \sup_{\|h\|_X \leq 1} \|(c_1 - \bar{a}) * (c_1 + \bar{a}) * h_1\|_{1,\nu} \\ &\leq 3\nu^2 \|A\|_{B(X)} \sup_{\|h\|_X \leq 1} \|c_1 - \bar{a}\|_{1,\nu} \|c_1 + \bar{a}\|_{1,\nu} \|h_1\|_{1,\nu} \\ &\leq 3\nu^2 \|A\|_{B(X)} r (\|c_1\|_{1,\nu} + \|\bar{a}\|_{1,\nu}) \\ &\leq 3\nu^2 \|A\|_{B(X)} r (r + 2\|\bar{a}\|_{1,\nu}). \end{aligned}$$

Then, assuming a loose a priori bound  $r \leq 1$  on the radius, we set

$$Z_2 \stackrel{\text{def}}{=} 3\nu^2 \|A\|_{B(X)} (1 + 2\|\bar{a}\|_{1,\nu}), \quad (27)$$

where  $\|A\|_{B(X)}$  is computed using Lemma 3.4, see also (25).

#### 4.5 Computer-assisted proofs

With a numerical continuation algorithm, we continued a branch of solutions of (21) that bifurcates from the zero solution at  $\lambda = (\pi^2 + 1)^2$ . Combining the bounds  $Y_0$ ,  $Z_0$ ,  $Z_1$  and  $Z_2$  given in (24), (25), (26) and (27), respectively, we define the radii polynomial  $p(r)$  as in (6). We prove the existence of six solutions of (21) by verifying the hypotheses of Theorem 2.1 with the routine `script_proofs.SH.m` available at [3]. The following table contains values of the bounds, as well as intervals  $[r_{\min}, r_{\max}]$  on which  $p(r) < 0$ , for sample values of  $\lambda$  ranging from 118.2 to 500.

$\lambda$	$\nu$	$N$	$Y_0$	$Z_1$	$Z_2$	$r_{\min}$	$r_{\max}$
118.2	1.15	39	$8.4414 \cdot 10^{-11}$	0.44499	$1.9642 \cdot 10^6$	$1.5218 \cdot 10^{-10}$	$2.8242 \cdot 10^{-7}$
120	1.1	54	$1.3132 \cdot 10^{-9}$	0.15775	$2.002 \cdot 10^5$	$1.5598 \cdot 10^{-9}$	$4.2056 \cdot 10^{-6}$
250	1.04	114	$4.3386 \cdot 10^{-8}$	0.27014	$4.89752 \cdot 10^5$	$6.2026 \cdot 10^{-8}$	$1.4282 \cdot 10^{-6}$
350	1.03	136	$4.8408 \cdot 10^{-8}$	0.15617	$1.536768 \cdot 10^6$	$6.508 \cdot 10^{-8}$	$4.8401 \cdot 10^{-7}$
450	1.02	164	$4.7337 \cdot 10^{-8}$	0.033858	$3.274572 \cdot 10^6$	$6.2042 \cdot 10^{-8}$	$2.33 \cdot 10^{-7}$
500	1.009	169	$5.6167 \cdot 10^{-8}$	0.06321	$3.73724 \cdot 10^6$	$9.9273 \cdot 10^{-8}$	$1.5139 \cdot 10^{-7}$



We observe from these data that as we increase the parameter  $\lambda$ , the dimension of the projection  $N$  needs to increase while the decay parameter  $\nu$  needs to decrease. This is due to the fact that the Taylor coefficients of the solutions decay slower as  $\lambda$  increases. Moreover, notice that the values of  $r_{\min}$  and  $r_{\max}$  are approaching each other, meaning that the proofs are getting harder and harder to obtain. This suggests that for larger parameter values a single Taylor expansion is not enough. The corresponding solutions can be seen in Figure 5. We plot in Figure 6 the solution  $(u, v)$  of (19) at  $\lambda = 500$ .

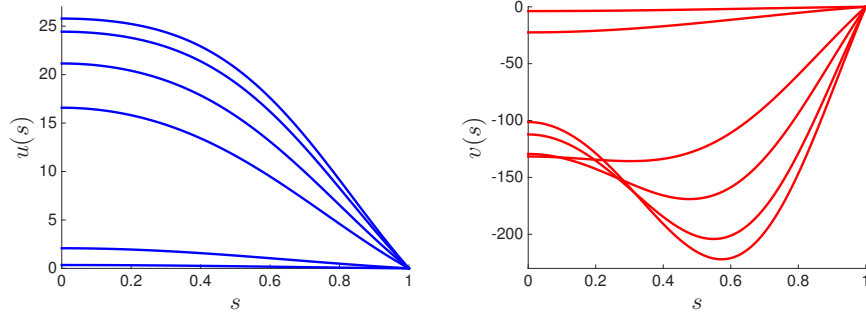


Figure 5: Six solutions of (21) for  $\lambda \in \{118.2, 120, 250, 350, 450, 500\}$ .

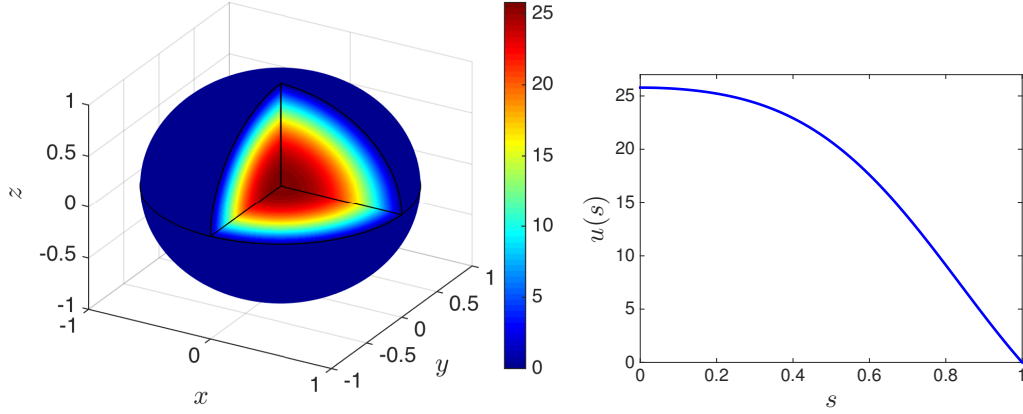


Figure 6: (Left) A stationary solution of the Swift-Hohenberg equation (19) on the unit ball in  $\mathbb{R}^3$  at  $\lambda = 500$ . (Right) The corresponding graph of  $u(s) = u(\sqrt{x^2 + y^2 + z^2})$ .

## 5 Conclusion

We have seen that some nontrivial boundary value problems originating from nonlinear PDEs can be solved in a Taylor series setting, one that requires relatively little technical machinery and is very accessible to students. Cutting of the Taylor series at some finite order and solving the associated finite dimensional algebraic system numerically, leads to an approximate solution, and we have proven that the true solution lies nearby. Indeed, based on Theorem 2.1 and

computable bounds (using interval arithmetic), we can estimate the distance between approximate and true solution rigorously and explicitly. This turns the numerical computation into a mathematical statement about the PDE.

There are limitations to the particular choice of Taylor series as our means of describing a solution, i.e., using monomials as our basis functions. In particular, we have seen that this put limits on the parameter range where we can apply this approach. More generally, there is a large variety of functional analytic setups and numerical algorithms, adapted to the particular problem under study, which fit into the general framework of Theorem 2.1. Successful examples include domain decomposition, Fourier series, Chebyshev series, splines, finite elements, as well as combinations of these. With these tools one is able, using the paradigm illustrated by the two examples in Sections 3 and 4, to solve eigenvalue problems, find periodic and connecting orbits, continue solutions in parameters, and analyze bifurcations. Using this rigorous computer-assisted framework one is able to obtain solutions to systems of ODEs, (time dependent) PDEs, delay differential equations, etc. Due to the nonlinear nature of these problems, you usually simply cannot get your hands on such solutions without the help of a computer. For further reading we refer to [1, 2, 5, 9, 16, 24, 26, 28] and the references therein.

## Acknowledgements

The first, fourth and sixth authors were supported by the Hungarian Scientific Research Fund (NKFIH-OTKA), Grant No. K109782.

## References

- [1] G. Arioli and H. Koch. Computer-assisted methods for the study of stationary solutions in dissipative systems, applied to the Kuramoto-Sivashinski equation. *Arch. Ration. Mech. Anal.*, 197:1033–1051, 2010.
- [2] G. Arioli and H. Koch. Integration of dissipative partial differential equations: a case study. *SIAM J. Appl. Dyn. Syst.*, 9:1119–1133, 2010.
- [3] I. Balázs, J.B. van den Berg, J. Courtois, J. Dudás, J.-P. Lessard, A. Vörös-Kiss, J.F. Williams, and X.Y. Yin. MATLAB code for “Computer-assisted proofs for radially symmetric solutions of PDEs”, 2017. <http://www.math.vu.nl/~janbouwe/code/radialpdes/>.
- [4] A.N. Baltagiannis and K.E. Papadakis. Periodic solutions in the Sun–Jupiter–Trojan Asteroid–Spacecraft system. *Planetary and Space Science*, 75:148–157, 2013.
- [5] B. Breuer, J. Horák, P.J. McKenna, and M. Plum. A computer-assisted existence and multiplicity proof for travelling waves in a nonlinearly supported beam. *J. Diff. Eq.*, 224(1):60–97, 2006.
- [6] J. Burgos-García and M. Gidea. Hill’s approximation in a restricted four-body problem. *Celestial Mech. Dynam. Astronom.*, 122(2):117–141, 2015.
- [7] CAPD: Computer assisted proofs in dynamics, a package for rigorous numerics. <http://capd.ii.uj.edu.pl/>.
- [8] A. Castro and E.M. Fischer. Infinitely many rotationally symmetric solutions to a class of semi-linear Laplace-Beltrami equations on spheres. *Canad. Math. Bull.*, 58(4):723–729, 2015.
- [9] J.-L. Figueras, M. Gameiro, J.-P. Lessard, and R. de la Llave. A framework for the numerical computation and a posteriori verification of invariant objects of evolution equations, 2017. To appear SIAM J. Appl. Dyn. Syst.
- [10] A. Hungria, J.-P. Lessard, and J.D. Mireles James. Rigorous numerics for analytic solutions of differential equations: the radii polynomial approach. *Math. Comp.*, 85(299):1427–1459, 2016.

- [11] H. Koch, A. Schenkel, and P. Wittwer. Computer-assisted proofs in analysis and programming in logic: a case study. *SIAM Rev.*, 38(4):565–604, 1996.
- [12] E.S.G. Leandro. On the central configurations of the planar restricted four-body problem. *J. Diff. Eq.*, 226(1):323–351, 2006.
- [13] P.-L. Lions. On the existence of positive solutions of semilinear elliptic equations. *SIAM Rev.*, 24(4):441–467, 1982.
- [14] S. McCalla and B. Sandstede. Snaking of radial solutions of the multi-dimensional Swift-Hohenberg equation: a numerical study. *Phys. D*, 239(16):1581–1592, 2010.
- [15] R.E. Moore, R.B. Kearfott, and M.J. Cloud. *Introduction to interval analysis*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2009.
- [16] M.T. Nakao. Numerical verification methods for solutions of ordinary and partial differential equations. *Numer. Funct. Anal. Optim.*, 22(3-4):321–356, 2001.
- [17] S.M. Rump. INTLAB - INTerval LABoratory. In Tibor Csendes, editor, *Developments in Reliable Computing*, pages 77–104. Kluwer Academic Publishers, Dordrecht, 1999.
- [18] S.M. Rump. Verification methods: Rigorous results using floating-point arithmetic. *Acta Numerica*, 19:287–449, 2010.
- [19] A. Scheel. Radially symmetric patterns of reaction-diffusion systems. *Mem. Amer. Math. Soc.*, 165(786), 2003.
- [20] C. Simó. Relative equilibrium solutions in the four-body problem. *Celestial Mech.*, 18(2):165–184, 1978.
- [21] J.B. Swift and P.C. Hohenberg. Hydrodynamic fluctuations at the convective instability. *Phys. Rev. A*, 15(1), 1977.
- [22] W. Tucker. A rigorous ODE Solver and Smale’s 14th Problem. *Foundations of Computational Mathematics*, 2(1):53–117–117, 2002.
- [23] W. Tucker. *Validated numerics*. Princeton University Press, Princeton, NJ, 2011. A short introduction to rigorous computations.
- [24] J.B. van den Berg, A. Deschênes, J.-P. Lessard, and J.D. Mireles James. Stationary coexistence of hexagons and rolls via rigorous computations. *SIAM J. Appl. Dyn. Syst.*, 14:942–979, 2015.
- [25] J.B. van den Berg and J.-P. Lessard. Rigorous numerics in dynamics. *Notices Amer. Math. Soc.*, 62(9):1057–1061, 2015.
- [26] J.B. van den Berg, J.-P. Lessard, and K. Mischaikow. Global smooth solution curves using rigorous branch following. *Math. Comp.*, 79(271):1565–1584, 2010.
- [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.
- [28] P. Zgliczyński. Rigorous numerics for dissipative partial differential equations. II. Periodic orbit for the Kuramoto-Sivashinsky PDE — a computer-assisted proof. *Found. Comput. Math.*, 4(2):157–185, 2004.
- [29] P. Zgliczyński and K. Mischaikow. Rigorous numerics for partial differential equations: the Kuramoto-Sivashinsky equation. *Found. Comput. Math.*, 1(3):255–288, 2001.