

# A functional analytic approach to validated numerics for eigenvalues of delay equations

Jean-Philippe Lessard <sup>\*</sup>      J.D. Mireles James <sup>†</sup>

June 3, 2019

## Abstract

This work develops validated numerical methods for linear stability analysis at an equilibrium solution of a system of delay differential equations (DDEs). The case of a single constant delay is considered. The method downplays the role of the scalar transcendental characteristic equation in favor of a functional analytic approach exploiting the strengths of numerical linear algebra/techniques of scientific computing. The idea is to consider an equivalent implicitly defined discrete time dynamical system which is projected onto a countable basis of Chebyshev series coefficients. The projected problem reduces to questions about certain sparse infinite matrices, which are well approximated by  $N \times N$  matrices for large enough  $N$ . We develop the appropriate truncation error bounds for the infinite matrices, provide a general numerical implementation which works for any system with one delay, and discuss computer-assisted theorems in a number of example problems.

**Key words.** Delay differential equations, spectral analysis, Chebyshev series, rigorous numerics, computer-assisted proofs

## 1 Introduction

A fundamental problem of numerical linear algebra is to find the eigenvalues and (possibly generalized) eigenvectors of an  $N \times N$  matrix. The literature on the topic is vast, and we refer to [15] for a broad overview. From the perspective of the present work what is important to mention is that a number of researchers have developed self validating numerical algorithms for solving eigenvalue/eigenvector problems. These methods employ fast numerical algorithms, pen and paper analysis, and deliberate control of rounding to provide mathematically rigorous error bounds on approximate eigendata. See the works of [37, 38, 29, 16, 23, 31, 8], and also the survey paper [30] for a thorough review of the literature.

An important line of research is to extend the finite dimensional methods just cited to infinite dimensional problems. Suppose for example that  $\mathcal{X}$  is a Banach space and that  $A: \mathcal{X} \rightarrow \mathcal{X}$  is a bounded linear operator. Numerical analysis of the spectrum of  $A$  presents new challenges, as some truncation is required before  $A$  can be represented on the digital computer. If  $A$  is a compact operator then for large enough  $N \in \mathbb{N}$  there is an  $N \times N$  matrix  $A^N$  approximating  $A$  as well as we like. By studying the eigenvalues of  $A^N$  and bounding the difference between  $A$  and  $A^N$  in an appropriate norm we can, in many cases, obtain mathematically rigorous information about the spectrum of the linear operator  $A$ .

---

<sup>\*</sup>McGill University, Department of Mathematics and Statistics, 805 Sherbrooke Street West, Montreal, QC, H3A 0B9, Canada. [jp.lessard@mcgill.ca](mailto:jp.lessard@mcgill.ca)

<sup>†</sup>Florida Atlantic University, Department of Mathematical Sciences, Science Building, Room 234, 777 Glades Road, Boca Raton, Florida, 33431, USA. [jmirelesjames@fau.edu](mailto:jmirelesjames@fau.edu)

Some works of this kind include the validated numerics for Floquet theory developed in [9], the methods for validated Morse index computations (unstable eigenvalue counts) for infinite dimensional compact maps in [19, 24], similar methods for equilibria of parabolic PDEs posed on compact domains in [26, 1, 28, 35, 25, 36], the validated numerics for stability/instability of traveling waves in [3, 2, 7, 5, 6] the stability analysis for periodic solutions of delay differential equations [18], the computer-assisted proofs of instability for periodic orbits of parabolic partial differential equations found in [14], and the computer-assisted proofs for trapping regions of equilibrium solutions of parabolic PDEs in [11].

The present work develops validated numerical methods for spectral analysis of equilibrium solutions of delay differential equations (DDEs), focusing on systems of scalar equations with a constant delay. Under suitable hypotheses a DDE generates a compact semiflow on a function space, and the problem is inherently infinite dimensional. Yet, as is well known, the equilibrium solutions solve a finite dimensional system of nonlinear equations, and the eigenvalues of the linearized problem are the complex zeros of a scalar analytic characteristic equation. Since the associated eigenfunctions are exponentials, the entire spectral analysis reduces to finding roots of finite dimensional equations.

The present work exploits the observation that parts of the analysis are actually easier when we stay in the infinite dimensional setting. The intuition behind this remark is simple: the infinite dimensional problem is linear, while the transcendental characteristic equation is highly nonlinear. Indeed, even the finite dimensional numerical analysis referred to in the first paragraph rarely passes through the characteristic equation. We argue that a functional analytic/scientific computing perspective is especially well suited to addressing the following problems.

**Problem 1 (approximate eigenvalues):** Given a reasonable approximation of an eigenvalue we iteratively refine via Newton’s method applied to the characteristic equation. This typically results in an approximation good to within a few multiples of machine precision. Moreover, as discussed in Section 2.7, mathematically rigorous a-posteriori error bounds are obtained using a Newton-Kantorovich argument. The hypotheses of the a-posteriori theorem are checked using interval arithmetic. The question remains, how do we find these “reasonable” initial approximations in the first place?

**Problem 2 (eigenvalue exclusion):** Suppose that after some numerical search we locate  $M$  approximate unstable eigenvalues. Assume moreover that we prove the existence of true unstable eigenvalues nearby, as already discussed in the statement of Problem 1. While this procedure provides a lower bound on the number of unstable eigenvalues, we would like to obtain also a sharp upper bound – in fact a validated exact count – on the number of unstable eigenvalues. This count is called the Morse index. This is a delicate problem as it involves ruling out the existence of any unstable eigenvalues *not found* by some search. More generally we would like to be able to count the eigenvalues in the complement of a circle of radius  $r > 0$  in  $\mathbb{C}$ . We refer to this quantity as the  $r$ -generalized Morse index.

One solution to Problem 1 is to perform a random search for approximate zeros in some large enough region of the complex plane. In the present setting something better can be done, as the zeros of the characteristic equation are the eigenvalues of a linear operator. We develop a functional analytic approach to the spectral analysis based on Galerkin projection of a compactified version of the linearized problem. This leads to a matrix whose eigenvalues approximate the compactified spectrum of the linearized DDE. The eigenvalues of the finite matrix are computed using standard methods of numerical linear algebra, and provide the initial guesses used for more refined calculation and validation. The eigenvalues of the compactified operator are related to the zeros of the transcendental characteristic equation through the complex exponential map.

Similarly, since Problem 2 involves counting the zeros of a complex analytic function, one solution is to apply the argument principle of complex analysis. When combined with validated numerical methods for computing line integrals, this provides the desired eigenvalue counts. Unfortunately, as we argue below, an approach based on the argument principle scales poorly with the dimension of the system of DDEs. The functional analytic approach on the other hand leads to a general scheme which is easy to implement for any system of DDEs with a constant delay.

To clarify the goals of the present work, and as motivation for the technical developments to follow, we present two example results obtained using our validated numerical arguments. For  $r > 0$ , let

$$B_r(z_0) = \{z \in \mathbb{C} : |z - z_0| < r\},$$

denote the standard ball of radius  $r$  about  $z_0$  in the complex plane. Here  $|\cdot|$  is the usual complex absolute value.

**Theorem 1.1** (Morse index for Mackey-Glass). *Consider the Mackey-Glass equation with parameter values  $\tau = 2$ ,  $\gamma = 1$ ,  $\beta = 2$ , and  $\rho = 10$ . The constant function  $y(t) = 1$  is an equilibrium solution with Morse index 2. Moreover, let*

$$r = 9.1 \times 10^{-16},$$

and

$$\bar{\lambda}_{1,2} = -1.635336834622171 \pm 1.428179851552561i.$$

The two unstable eigenvalues  $\lambda_u^1, \lambda_u^2 \in \mathbb{C}$  are complex conjugate numbers with

$$|\lambda_u^{1,2} - \bar{\lambda}_{1,2}| < r.$$

(See Section 2.6.1 for the mathematical definition of the Mackey-Glass Equation).

**Theorem 1.2** (Morse index for a delayed van der Pol Equation). *Consider the delayed van der Pol equation with parameter values  $\tau = 2$ ,  $\kappa = -1$ , and  $\varepsilon = 0.15$ . The constant function  $x(t) = 0, y(t) = 0$  is an equilibrium solution with Morse index 2. Moreover, let*

$$r = 4.26 \times 10^{-15},$$

and

$$\bar{\lambda}_{1,2} = -0.61810956461394 \pm 1.84334863710072i$$

The two unstable eigenvalues  $\lambda_u^1, \lambda_u^2$  are complex conjugate numbers with

$$|\lambda_u^{1,2} - \bar{\lambda}_{1,2}| < r.$$

(See Section 2.6.3 for the mathematical definition of the delayed van der Pol Equation)

Other results of this kind are presented in Section 4 using the methods of the present work. Indeed our main result is a computational recipe which applies to any scalar system of DDEs with a constant delay. The remainder of the paper is organized as follows.

In Section 2 we review some background material for abstract dynamical systems defined by an implicit rule, and derive expressions for the linearization at a fixed point. This leads to a generalized eigenvalue problem for the linearized problem. We recall the method of steps for DDEs and see how it fits into the abstract formulation, defining the so-called *step map* which we study throughout the remainder of the paper. We discuss compactness properties of the step map, and give an elementary derivation of its characteristic equation. We relate this equation to the usual characteristic equation for the infinitesimal problem. We recall a simple a-posteriori

theorem which provides validated error bounds for approximate solutions of the characteristic equation, and hence validated eigenvalue bounds. Finally we describe the example systems used in the application sections.

Next, in Section 3 we present some heuristic arguments explaining the potential use of techniques from complex analysis to analyze the spectrum of the linearized step map, and discuss why this analysis is not as straight forward as it first appears. Section 4 presents the main results of the paper, developing the functional analytic approach necessary for studying the spectrum of the step map via numerical linear algebra. We project the problem onto a space of Chebyshev series and see that the linear operators have a sparse representation in this basis. We truncate and compute numerical eigenvalues, and prove a theorem relating  $r$ -generalized Morse index of the numerical matrix to the index of the infinite dimensional problem. We discuss the application of these ideas to a number of problems. Finally in Section 5 we summarize our results and discuss some possible future extensions.

The computer programs which validate the computer assisted theorems presented in this paper are implemented in MATLAB and use the INTLAB library for interval arithmetic [32]. The codes used to produce all the results in the present work are freely available at [21].

## 2 Background

In this section we review some well known facts about delay differential equations. Several of the derivations are included so that the manuscript is more self-contained.

### 2.1 Abstract formulation of the problem

Let  $X, Y$  be Banach spaces and  $T: Y \times X \rightarrow Y$  be a smooth function. Moreover suppose that for each  $(y, x) \in Y \times X$  the Fréchet derivatives with respect to the first and second variables, denoted respectively by  $D_1T(y, x)$  and  $D_2T(y, x)$ , exist and are bounded linear operators. For fixed  $x \in X$  consider the problem of finding a  $y \in Y$  so that

$$T(y, x) = y.$$

We think of  $x$  as a parameter and look for fixed points of the family of fixed point operators  $T_x: Y \rightarrow Y$  defined by

$$T_x(y) = T(y, x).$$

Let  $D \subset X$  have

$$D = \{x \in X : T_x \text{ has a unique fixed point } y \in Y\},$$

and define a mapping  $F: D \subset X \rightarrow Y$  by the correspondence

$$F(x) = y, \quad \text{if and only if} \quad T(y, x) = y \quad (\text{uniquely}).$$

In words  $y = F(x)$  if and only  $y$  is the unique fixed point of  $T_x$  in  $Y$ .

Suppose that  $x_0 \in D$  and let  $y_0 \in Y$  denote the unique fixed point of  $T_{x_0}$  in  $Y$ . Assume that that  $\text{Id} - D_1T(y_0, x_0)$  is an isomorphism of  $Y$ . It follows from the implicit function theorem that  $F$  is defined, continuous, and Fréchet differentiable in a neighborhood of  $x_0$ .

To see this consider the function  $G: Y \times X \rightarrow Y$  defined by

$$G(y, x) = y - T(y, x).$$

Note that  $G(y_0, x_0) = 0$ , and that  $D_1G(y_0, x_0) = \text{Id} - D_1T(y_0, x_0)$  is an isomorphism. Then there is an  $\epsilon > 0$  and a continuous function  $y: B_\epsilon(x_0) \rightarrow Y$  so that  $y(x_0) = y_0$  and

$$G(y(x), x) = 0 \quad \text{for all } x \in B_\epsilon(x_0).$$

It follows that

$$T(y(x), x) = y(x),$$

for all  $x \in B_\epsilon(x_0) \subset X$ . That is, the function  $F$  is locally well defined near  $x_0$  by

$$F(x) = y(x).$$

Moreover, after differentiating the equation  $T(F(x), x) = F(x)$  with respect to  $x$ , we have that

$$DF(x) = [\text{Id} - D_1T(F(x), x)]^{-1} D_2T(F(x), x). \quad (1)$$

## 2.2 Linearization of the abstract problem at a fixed point

Consider the special case when  $X = Y$ , so that  $F: X \rightarrow X$  is a self-map. We are interested in the dynamics generated by  $F$ . In particular we study the linearization at a fixed point. Note that  $x_0 \in X$  is a fixed point of  $F$  if and only if

$$T(x_0, x_0) = x_0.$$

From Equation (1) we have that the derivative of  $F$  at a fixed point  $x_0$  is given by

$$DF(x_0) = [\text{Id} - D_1T(x_0, x_0)]^{-1} D_2T(x_0, x_0),$$

as long as  $\text{Id} - D_1T(x_0, x_0)$  is an isomorphism.

Then  $\lambda \in \mathbb{C}$  is an eigenvalue of  $DF(x_0)$  if and only if there is a non-zero  $\xi \in X$  so that

$$[\text{Id} - D_1T(x_0, x_0)]^{-1} D_2T(x_0, x_0)\xi = \lambda\xi, \quad (2)$$

which is equivalent to the generalized eigenvalue problem

$$M_2\xi = \lambda M_1\xi,$$

where

$$M_2 = D_2T(x_0, x_0), \quad \text{and} \quad M_1 = \text{Id} - D_1T(x_0, x_0). \quad (3)$$

Equations (2) and (3) provide a way to study the spectrum of  $DF(x_0)$  even if  $F$  is only implicitly defined.

## 2.3 The method of steps for DDEs

Let  $f: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  be a smooth function,  $\tau > 0$  a positive constant, and  $x_0(t) \in C^k([-\tau, 0])$  a given smooth function. We say that  $y: C([-\tau, T])$  is a solution of the delay differential equation

$$y'(t) = f(y(t), y(t - \tau)), \quad (4)$$

with history  $x_0(t)$  if  $y(t) = x_0(t)$  for  $t \in [-\tau, 0]$  and  $y(t)$  satisfies Equation (4) for all  $t \in (0, T)$ .

Consider the mapping  $T: C^k([-\tau, 0]) \times C^k([-\tau, 0]) \rightarrow C^k([-\tau, 0])$  defined by

$$T(y, x)(t) = x(0) + \int_{-\tau}^t f(y(s), x(s)) ds. \quad (5)$$

Then we are in precisely the setting of Section 2.1, and we define the map  $F: C^k([-\tau, 0]) \rightarrow C^k([-\tau, 0])$  by the rule that  $F(x) = y$  if and only if  $T(y, x) = y$ .

One checks that if  $x \in C^k([-\tau, 0])$  then  $F(x(t)) = y(t)$  is as differentiable as  $x(t)$  and  $f$  by repeatedly differentiating the formula

$$y(t) = x(0) + \int_{-\tau}^t f(y(s), x(s)) ds.$$

Indeed,  $y(t)$  has one more derivative than the least smooth of  $f$  and  $x$ . It follows that if  $f$  is  $C^\infty$  then  $F: C^k([-\tau, 0]) \rightarrow C^{k+1}([-\tau, 0])$ , so that iterates gain one derivative with every application of  $F$ .

This map  $F$  is called the *step map* for Equation (4), and its iterates are related to solutions of the DDEs by the following Lemma. The elementary proof is found in [20].

**Lemma 2.1** (Orbits of the step map are solutions of the DDE). *Let  $y_0 \in C([-\tau, 0])$  and assume that  $y_1, \dots, y_N \in C([-\tau, 0])$  are the first  $N$  iterates of  $y_0$  under the step map. Then the function  $y: [-\tau, N\tau] \rightarrow \mathbb{R}$  defined by*

$$y(t) = \begin{cases} y_0(t), & t \in [-\tau, 0) \\ y_1(t - \tau), & t \in [0, \tau) \\ y_2(t - 2\tau), & t \in [\tau, 2\tau) \\ \vdots & \\ y_N(t - N\tau), & t \in [(N-1)\tau, N\tau], \end{cases} \quad (6)$$

is a solution of Equation (4) on  $(0, N\tau)$  with initial history  $y_0$ .

## 2.4 Linear stability of constant fixed points of the step map

Now consider the relationship between constant solutions of Equation (4) and fixed points of the step map  $F$ . Indeed, suppose that  $c \in \mathbb{R}^d$  has

$$f(c, c) = 0.$$

We say that the function  $x(t) = c$  is an equilibrium solution of the DDE. Observe that

$$\begin{aligned} T(x(t), x(t)) &= x(0) + \int_{-\tau}^t f(x(s), x(s)) ds \\ &= c + \int_{-\tau}^t f(c, c) ds \\ &= c \\ &= x(t), \end{aligned}$$

so that the constant function  $x(t) = c$  is a fixed point of the map  $F$ . A partial converse holds: one can show that if  $x(t)$  is a fixed point of  $F$  then  $x(t)$  is either constant or is a non-constant function of period  $\tau$  – that is a periodic solution of Equation (4) whose period is in one-to-one resonance with the delay. This latter property is not generic, so that in general fixed points of  $F$  correspond to equilibrium solutions of Equation (4).

Now consider the eigenvalue problem at a fixed point  $x_0(t) = c$ . The eigenvalue problem

$$DF(x_0)\xi(t) = \lambda\xi(t)$$

can be rewritten as

$$[\text{Id} - D_1T(c, c)]^{-1} D_2T(c, c)\xi(t) = \lambda\xi(t),$$

which is equivalent to the generalized eigenvalue problem

$$D_2T(c, c)\xi(t) = \lambda[\text{Id} - D_1T(c, c)]\xi(t).$$

Define the  $d \times d$  matrices

$$K_1 \stackrel{\text{def}}{=} \partial_1 f(c, c) \quad \text{and} \quad K_2 \stackrel{\text{def}}{=} \partial_2 f(c, c), \quad (7)$$

so that we have the eigenvalue problem

$$\lambda\xi(t) - \lambda \int_{-\tau}^t K_1\xi(s) ds = \xi(0) + \int_{-\tau}^t K_2\xi(s) ds. \quad (8)$$

and observe that while the eigenvalue problem involves infinite dimensional integral operators, these operators are completely determined by the entries of the two matrices  $K_1, K_2$ , that is by the partial derivatives of  $f$  at  $c$ .

Observe that if  $(\lambda, \xi)$  is an eigenvalue/eigenvector pair for Equation (8), then  $\xi(t)$  is differentiable for all  $t \in (-\tau, 0)$ . Differentiating Equation (8) with respect to  $t$  gives that  $\xi$  satisfies the constant coefficient linear differential equation

$$\xi'(t) = \left( \partial_1 f(c, c) + \frac{1}{\lambda} \partial_2 f(c, c) \right) \xi(t),$$

or

$$\xi'(t) = \left( K_1 + \frac{1}{\lambda} K_2 \right) \xi(t), \quad (9)$$

subject to the initial condition

$$\lambda\xi(-\tau) = \xi(0). \quad (10)$$

For fixed  $\lambda \in \mathbb{C}$ , Equation (9) is a homogeneous linear system of ordinary differential equations with constant coefficients. Recall that a complex number  $\alpha \in \mathbb{C}$  is an eigenvalue of  $K_1 + \lambda^{-1}K_2$  if and only if  $\alpha$  is a zero of the characteristic equation

$$\det(K_1 + \lambda^{-1}K_2 - \alpha\text{Id}) = 0.$$

For any such  $\alpha$ , a vector  $\eta \in \mathbb{C}^d$  is an eigenvalue of  $K_1 + \lambda^{-1}K_2$  if and only if

$$\left( K_1 + \lambda^{-1}K_2 \right) \eta = \alpha\eta.$$

Given any eigenvalue/eigenvector pair  $(\alpha, \eta)$  of  $K_1 + \lambda^{-1}K_2$ , the function

$$\xi(t) = e^{\alpha t}\eta,$$

is a solution of Equation (9). Then  $\xi(t)$  is in fact real analytic on  $[-\tau, 0]$ .

Imposing the constraint of Equation (10) gives

$$\lambda e^{-\tau\alpha}\eta = \eta,$$

which leads to the scalar constraint

$$\lambda e^{-\tau\alpha} = 1,$$

or

$$\lambda = e^{\tau\alpha}.$$

Solving for  $\alpha$  gives

$$\alpha = \frac{\ln(\lambda)}{\tau},$$

as the relationship connecting  $\lambda$  and  $\alpha$ .

Substituting this expression back into the characteristic equation leads to the transcendental equation

$$\det \left( K_1 + \frac{1}{\lambda} K_2 - \frac{\ln(\lambda)}{\tau} \text{Id} \right) = 0, \quad (11)$$

whose zeros are the eigenvalues of Equation (8).

Equation (11) written in terms of  $\alpha$  is

$$\det (K_1 + e^{-\tau\alpha} K_2 - \alpha \text{Id}) = 0, \quad (12)$$

and for every root  $\alpha$  we obtain an eigenvalue  $\lambda$  by the relationship  $\lambda = e^{\tau\alpha}$ . The dynamical relationship between the two problems is that the solutions of Equation (12) are the usual *infinitesimal eigenvalues* for the ODE on  $C([-\tau, 0])$  generated by Equation (4), while  $\lambda$  solving Equation (11) are the eigenvalues of the time- $\tau$  map on  $C([-\tau, 0])$  generated by the flow. Then it is natural that the two notions are related through the complex logarithm. The advantage of working with the method of steps – that is with the solutions of Equation (11) – is that the spectrum is a compact subset of  $\mathbb{C}$ .

## 2.5 Further remarks on the characteristic equation

Denote the entries of the  $d \times d$  matrices  $K_1 = \partial_1 f(c, c)$  and  $K_2 = \partial_2 f(c, c)$  by

$$K_1 = \begin{pmatrix} k_{11}^1 & \dots & k_{1d}^1 \\ \vdots & \ddots & \vdots \\ k_{d1}^1 & \dots & k_{dd}^1 \end{pmatrix}, \quad \text{and} \quad K_2 = \begin{pmatrix} k_{11}^2 & \dots & k_{1d}^2 \\ \vdots & \ddots & \vdots \\ k_{d1}^2 & \dots & k_{dd}^2 \end{pmatrix},$$

so that the characteristic equation (11) is

$$\det \begin{pmatrix} k_{11}^1 + \frac{1}{\lambda} k_{11}^2 - \frac{\ln(\lambda)}{\tau} & \dots & k_{1d}^1 + \frac{1}{\lambda} k_{1d}^2 \\ \vdots & \ddots & \vdots \\ k_{d1}^1 + \frac{1}{\lambda} k_{d1}^2 & \dots & k_{dd}^1 + \frac{1}{\lambda} k_{dd}^2 - \frac{\ln(\lambda)}{\tau} \end{pmatrix} = 0.$$

Expanding the determinant leads to a polynomial of the form

$$\begin{aligned} p(x, y) &= c_{00} + c_{10}x + c_{01}y + c_{20}x^2 + c_{11}xy + c_{02}y^2 + \dots + c_{0d}x^d + c_{d0}y^d \\ &= \sum_{n=0}^d \sum_{k=0}^n c_{n-k, k} x^{n-k} y^k. \end{aligned}$$

in the variables

$$x = \frac{1}{\lambda}, \quad \text{and} \quad y = \ln(\lambda).$$

Note that when  $|\lambda| < 1$  both  $x$  and  $y$  are large, and taking products and powers introduces numerical instabilities.

When  $d = 1$  Equation (11) reduces to the scalar equation

$$K_1 + \frac{1}{\lambda} K_2 - \frac{\ln(\lambda)}{\tau} = 0,$$

which we rewrite as

$$\ln(\lambda) = \tau K_1 + \frac{\tau}{\lambda} K_2.$$

Exponentiating leads to

$$\lambda = e^{\tau K_1 + \frac{\tau K_2}{\lambda}}.$$

Then in the one-dimensional case we look for complex roots of the function

$$g(z) = z - e^{\tau K_1} e^{\frac{\tau K_2}{z}},$$

to determine the eigenvalues of the equilibrium.

When  $d \geq 2$  the situation is more delicate. To see why we consider explicitly the case of  $d = 2$ , where we must study the equation

$$\begin{aligned} \det \begin{pmatrix} k_{11}^1 + \frac{k_{11}^2}{\lambda} - \frac{\ln(\lambda)}{\tau} & k_{12}^1 + \frac{k_{12}^2}{\lambda} \\ k_{21}^1 + \frac{k_{21}^2}{\lambda} & k_{22}^1 + \frac{k_{22}^2}{\lambda} - \frac{\ln(\lambda)}{\tau} \end{pmatrix} \\ = \left( k_{11}^1 + \frac{k_{11}^2}{\lambda} - \frac{\ln(\lambda)}{\tau} \right) \left( k_{22}^1 + \frac{k_{22}^2}{\lambda} - \frac{\ln(\lambda)}{\tau} \right) - \left( k_{12}^1 + \frac{k_{12}^2}{\lambda} \right) \left( k_{21}^1 + \frac{k_{21}^2}{\lambda} \right) \\ = \frac{\ln(\lambda)^2}{\tau^2} + c_1 \frac{\ln(\lambda)}{\lambda} + c_2 \ln(\lambda) + c_3 \frac{1}{\lambda} + c_4 \frac{1}{\lambda^2} + c_5 \\ = 0, \end{aligned}$$

for some constants  $c_1, c_2, c_3, c_4$ , and  $c_5$  which can be worked out explicitly. Note that the problem is fundamentally different from the one-dimensional case, as there is no obvious way to isolate and remove the logarithmic terms. We can switch back to the exponential form of the characteristic equation given in Equation (12), but then the compactness of the spectrum is lost.

## 2.6 The Example Systems

The following four delay equations are used to illustrate the utility of our validation scheme.

### 2.6.1 The Mackey-Glass Equation

Consider the scalar Mackey-Glass equation [22]

$$y'(t) = f(y(t), y(t - \tau)) = -\gamma y(t) + \beta \frac{y(t - \tau)}{1 + y(t - \tau)^\rho}, \quad \gamma, \beta, \rho > 0 \quad (13)$$

where

$$f(y, x) = -\gamma y + \beta \frac{x}{1 + x^\rho}. \quad (14)$$

This DDE was originally introduced to model the concentration of white blood cells in a subject. We refer to  $\gamma = 1$ ,  $\beta = 2$  and  $\rho = 10$  as the *classic parameter values* for Mackey-Glass.

Note that

$$c_0 = 0, \quad \text{or} \quad c_1 = \left( \frac{\beta}{\gamma} - 1 \right)^{\frac{1}{\rho}},$$

are equilibrium solutions, and at the classic parameter values we see that  $c_1 = 1$ , and moreover that

$$K_1 = \partial_1 f(c, c) = -\gamma \quad \text{and} \quad K_2 = \partial_2 f(c, c) = \beta \frac{1 + (1 - \rho)c^\rho}{(1 + c^\rho)^2}. \quad (15)$$

### 2.6.2 The Cubic Ikeda-Matsumoto Equation

Consider the delay differential equation

$$y'(t) = f(y(t), y(t - \tau)) \stackrel{\text{def}}{=} y(t - \tau) - y(t - \tau)^3, \quad (16)$$

which was considered in [33, 34] as a simple model exhibiting chaotic motion (for instance for the parameter values  $\tau \in [1.538, 1.723]$ ). Remark that Equation (16) can be recovered (via the rescaling  $y(t) \stackrel{\text{def}}{=} \frac{1}{\sqrt{3}} z(t)$ ) as the third order approximation of the DDE  $z'(t) = \sin(z(t - \tau))$  considered by Ikeda and Matsumoto in [17]. For that reason, we call the DDE (16) the *Cubic Ikeda-Matsumoto* equation. There are three steady states given by  $x \equiv c \in \{-1, 0, 1\}$ , and note that

$$K_1 = \partial_1 f(c, c) = 0 \quad \text{and} \quad K_2 = \partial_2 f(c, c) = 1 - 3c^2.$$

### 2.6.3 Delayed van der Pol

Consider the delayed van der Pol delay differential equation (as considered in [27])

$$z''(t) - \varepsilon z'(t)(1 - z(t)^2) + z(t - \tau) - \kappa z(t) = 0,$$

which leads (letting  $y_1(t) = z(t)$ ,  $y_2(t) = z'(t)$  and  $y = (y_1, y_2)$ ) to

$$y'(t) = f(y(t), y(t - \tau)) \stackrel{\text{def}}{=} \begin{pmatrix} y_2(t) \\ \varepsilon y_2(t)(1 - y_1(t)^2) - y_1(t - \tau) + \kappa y_1(t) \end{pmatrix}. \quad (17)$$

Given  $\kappa \neq 1$ , note that  $c = (0, 0)^T$  is the only steady state. Moreover,

$$K_1 = \partial_1 f(c, c) = \begin{pmatrix} 0 & 1 \\ -2\varepsilon y_1(1)y_2(t) + \kappa & \varepsilon(1 - y_1(t)^2) \end{pmatrix} \Big|_{y=c} = \begin{pmatrix} 0 & 1 \\ \kappa & \varepsilon \end{pmatrix} \quad (18)$$

and

$$K_2 = \partial_2 f(c, c) = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}. \quad (19)$$

### 2.6.4 Delayed predator-prey model

Denoting  $y = (y_1, y_2) \in \mathbb{R}^2$ , consider the delayed predator-prey model (as studied in [12])

$$y'(t) = f(y(t), y(t - \tau)) \stackrel{\text{def}}{=} \begin{pmatrix} \tau y_1(t)(r_1 - a y_1(t) - y_2(t - 1)) \\ \tau y_2(t)(-r_2 + y_1(t - 1) - b y_2(t)) \end{pmatrix}, \quad (20)$$

where  $\tau, r_1, r_2 > 0$  and  $a, b \geq 0$ . The model has a unique positive equilibrium solution given by

$$c = \begin{pmatrix} y_1^* \\ y_2^* \end{pmatrix} = \begin{pmatrix} \frac{r_2 + b r_1}{a b + 1} \\ \frac{r_1 - a r_2}{a b + 1} \end{pmatrix}. \quad (21)$$

In this case,

$$\begin{aligned} K_1 = \partial_1 f(c, c) &= \begin{pmatrix} \tau r_1 - 2a\tau y_1(t) - \tau y_2(t - 1) & 0 \\ 0 & -\tau r_2 + \tau y_1(t - 1) - 2b\tau y_2(t) \end{pmatrix} \Big|_{y=c} \\ &= \begin{pmatrix} -\frac{a\tau(r_2 + b r_1)}{a b + 1} & 0 \\ 0 & -\frac{b\tau(r_1 - a r_2)}{a b + 1} \end{pmatrix} = \begin{pmatrix} -a\tau y_1^* & 0 \\ 0 & -b\tau y_2^* \end{pmatrix} \end{aligned}$$

and

$$K_2 = \partial_2 f(c, c) = \begin{pmatrix} 0 & -\tau y_1(t) \\ \tau y_2(t) & 0 \end{pmatrix} \Big|_{y=c} = \begin{pmatrix} 0 & -\tau y_1^* \\ \tau y_2^* & 0 \end{pmatrix}.$$

## 2.7 Validated zero finding for a complex analytic function

Suppose that  $g$  is an analytic function and that we have located an approximate zero  $\bar{z} \in \mathbb{C}$  of  $g$ . We would like to conclude that there is a true zero of  $g$  near  $\bar{z}$ , and the following standard theorem provides numerically verifiable sufficient conditions. We include the elementary proof in the Appendix for the sake of completeness.

**Theorem 2.2.** *Suppose that  $g: B_{r_*}(\bar{z}) \rightarrow \mathbb{C}$  is analytic. Assume that  $g'(\bar{z}) \neq 0$  and let*

$$a = \frac{1}{g'(\bar{z})}.$$

*Suppose that  $Y$ , and  $Z$  are positive constants with*

$$|ag(\bar{z})| \leq Y, \tag{22}$$

*and*

$$|a| \sup_{z \in B_{r_*}(\bar{z})} |g''(z)| \leq Z. \tag{23}$$

*Define the polynomial*

$$p(z) = Zr^2 - r + Y.$$

*Then for any  $0 < r_0 < r_*$  so that*

$$p(r_0) < 0,$$

*there exists a unique  $\tilde{z} \in B_{r_0}(\bar{z})$  so that*

$$g(\tilde{z}) = 0.$$

*Moreover,*

$$g'(\tilde{z}) \neq 0.$$

The utility of the theorem is best seen in examples, so we illustrate the procedure for validating eigenvalue bounds for DDEs. Note that the following examples do not address how initial conditions for the Newton iteration are found, nor do they touch on eigenvalue exclusion. The examples do however show quite successfully that it is very easy to obtain existence and validated error bounds for the eigenvalues, at least in scalar/low dimensional examples.

**Example 1 (eigenvalue validation for a scalar DDE):** Consider for example the Mackey-Glass equation (13) with  $\tau = 2, \gamma = 1, \beta = 2$ , and  $\rho = 10$  at the equilibrium solution  $c = 1$ . Recalling (15), at these parameter values

$$K_1 = -1 \quad \text{and} \quad K_2 = -4$$

and the corresponding characteristic equation is given by

$$g(z) = -1 - \frac{4}{z} - \frac{\ln(z)}{2} = 0.$$

Starting a Newton iteration from  $z_0 = -1$  results in

$$\bar{z} = -1.635336834622171 + 1.428179851552561i,$$

and we can check that

$$|g(\bar{z})| \approx 1.3 \times 10^{-15}.$$

We will now apply Theorem 2.2 with  $\bar{z}$  as our initial data. Using that

$$g'(z) = \frac{8-z}{2z^2}, \quad \text{and} \quad g''(z) = \frac{z-16}{2z^3},$$

we use interval arithmetic in INTLAB to check that

$$a = \frac{1}{g'(\bar{z})} \in B_r(w),$$

where

$$w = 0.269522080830080 - 0.929629614323464i, \quad \text{and} \quad r = 2.09963 \times 10^{-15}.$$

Again using interval arithmetic we check that

$$|ag(\bar{z})| \in [9.015 \times 10^{-16}, 9.016 \times 10^{-16}],$$

and take

$$Y = 9.02 \times 10^{-16},$$

so that  $Y$  satisfies the inequality hypothesized in Equation (22). (This value of  $Y$  simplifies the discussion, however the bounds obtained and stored by the computer are somewhat sharper than this. The interested reader can refer to the MATLAB program referenced at the end of the example). Choosing a ball of radius  $r_* = 0.5$  about  $\bar{z}$  we check, using interval arithmetic that

$$|a| \sup_{z \in B_{r_*}(\bar{z})} |g''(z)| \in [0.4, 3.4].$$

This bound is obtained by evaluating the formula for  $g''$  on the ball about  $\bar{z}$  of radius 2, taking the complex absolute value of the result, and multiplying it by an interval bound on the absolute value of  $a$ . Again, shaper bounds are stored on the computer.

Taking

$$Z = 3.4,$$

insures that  $Z$  satisfies the hypothesis of Equation (23).

The quadratic equation  $p(r) = Zr^2 - r + Y$  has two roots given by

$$r_- = \frac{2Y}{1 + \sqrt{1 - 4ZY}} > 9.1 \times 10^{-16}, \quad \text{and} \quad r_+ = \frac{1 + \sqrt{1 - 4ZY}}{2Z} < 1.19,$$

where the expressions have been evaluated using interval arithmetic. Then for any  $r_- < r_0 < r_+$  we have that  $p(r_0) < 0$ . Since  $r_+ > 0.5 = r_*$  we have that for any  $r_- < r_0 < r_*$ , there is a unique root  $\tilde{z}$  of  $g(z)$  having

$$|\tilde{z} - \bar{z}| < r_0.$$

Since these balls are nested we conclude that there is a true zero  $\hat{z}$  of  $g$  with

$$|\hat{z} - \bar{z}| < 9.1 \times 10^{-16}$$

and that any other zeros of  $g$  are in the complement of the ball  $B_{r_*}(\bar{z})$ .

Observe that since  $K_1$  and  $K_2$  are real, the complex conjugate of  $\hat{z}$  is another zero  $g(z)$ , and we have proven the existence and error bounds claimed in Theorem 1.1. The MATLAB program `script_validateEig_c1_MackeyGlass.m` available at [21] executes the operations described above. To complete the proof of Theorem 1.1 we still have to show that these are the only two unstable eigenvalues. This will be done using the theory of Section 4.

**Example 2 (eigenvalue validation for a system of DDEs):** Consider now the delayed van der Pol equation (17). Recalling (18) and (19), this leads to the characteristic equation

$$\det \left( \begin{pmatrix} 0 & 1 \\ \kappa & \varepsilon \end{pmatrix} + \frac{1}{z} \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} - \frac{\ln(z)}{\tau} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) = 0.$$

$$\det \left( K_1 + \frac{1}{\lambda} K_2 - \frac{\ln(\lambda)}{\tau} \text{Id} \right) = 0,$$

Then for the parameters  $\kappa = -1$ ,  $\varepsilon = 0.15$ , and  $\tau = 2$  we seek complex zeros of the function

$$g(z) = \frac{(\ln(z))^2}{4} - \frac{3 \ln(z)}{40} + \frac{1}{z} + 1.$$

In the MATLAB program `script_validateEig_c1_vanDerPol.m` available at [21], the formulas for  $g'$  and  $g''$  are given. This program performs nearly identical steps as those discussed in Example 1. Indeed, starting from an initial guess of  $z = -0.5$  we obtain an approximate zero of

$$\bar{z} = -0.61810956461394 + 1.84334863710072i,$$

after seven Newton steps. Taking  $r_* = 0.25$  we check that  $Y = 4.26 \times 10^{-15}$  and  $Z = 9$  satisfy the bounds hypothesized in Theorem 2.2. By computing the roots of  $p(r) = Zr^2 - r + Y$  we find that there exists a true zero  $\hat{z}$  of  $g$  having that

$$|\hat{z} - \bar{z}| \leq 4.3 \times 10^{-15},$$

and that any other zeros of  $g$  are in the complement of the ball of radius 0.25 about  $\bar{z}$ . Again, the complex conjugate is also a solution and we have the existence and error bounds for Theorem 1.2.

### 3 Interlude: the argument principle of complex analysis

Suppose we want to count the unstable eigenvalues associated with an equilibrium solution  $c \in \mathbb{R}^d$  of Equation (4). The description of the spectrum of  $DF(c)$  in terms of the zeros of a scalar characteristic equation is at first glance so appealing that it is worth explaining carefully *what we will not do* in our approach, and *why we will not do it*. Recalling that the eigenvalues of  $DF(c)$  are the complex zeros of

$$g(z) = \det \left( K_1 + \frac{1}{z} K_2 - \frac{\ln(z)}{\tau} \text{Id} \right),$$

we make the change of variables

$$\frac{1}{w} \rightarrow z,$$

and define the new function

$$\tilde{g}(w) = \det \left( K_1 + w K_2 + \frac{\ln(w)}{\tau} \text{Id} \right). \quad (24)$$

The zeros of  $\tilde{g}$  inside the unit circle are the desired unstable eigenvalues.

Suppose now that  $\Gamma$  is a simple closed curve in  $\mathbb{C}$  with positive orientation which does not intersect any pole or zero of  $\tilde{g}$ , and that  $\tilde{g}$  is analytic in the open set enclosed by  $\Gamma$  except possibly at a finite number of poles. By the argument principle of complex analysis we have that

$$N_{\text{zeros}} - N_{\text{poles}} = \frac{1}{2\pi i} \int_{\Gamma} \frac{\tilde{g}'(z)}{\tilde{g}(z)} dz. \quad (25)$$

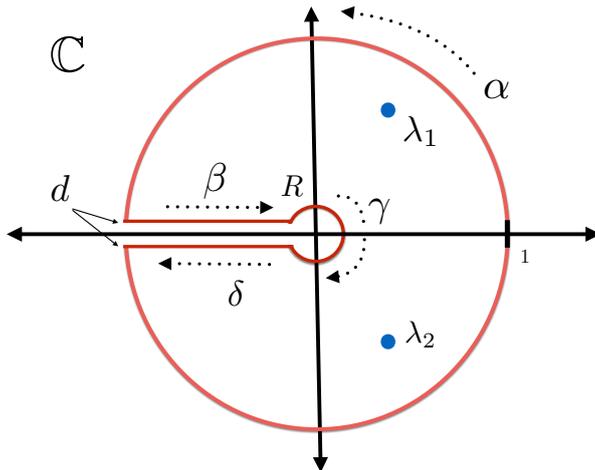


Figure 1: **Validated Morse index by contour integration:** imagine that  $\lambda_1$  and  $\lambda_2$  zeros inside the unit circle for the function  $\tilde{g}$  defined in Equation (24), so that  $\lambda_1^{-1}, \lambda_2^{-1}$  are unstable eigenvalues for the step map  $F$ . To validate the eigenvalue count we choose a branch of the complex logarithm defined on  $\mathbb{C} \setminus (-\infty, 0]$ , and consider the “key hole” contour  $\Gamma = \alpha + \beta + \gamma + \delta$  illustrated in the figure. If  $\tilde{g}$  is analytic inside the region enclosed by  $\Gamma$  then the argument principle counts the number of zeros inside. Supposing that there are no zeros or poles of  $\tilde{g}$  on  $[-1, 0)$ , taking the limit as  $R \rightarrow 0$  and  $d \rightarrow 0$  gives the unstable eigenvalue count, i.e. the Morse index. By choosing other circles we could count the number of stable eigenvalue with modulus larger than some desired bound.

Here  $N_{\text{zeros}}$  is the number of zeros and  $N_{\text{poles}}$  the number of poles enclosed by  $\Gamma$ . By implementing a validated numerical scheme for evaluating the contour integral on the right hand side of Equation (25) one obtains the desired count, that is the Morse index of the equilibrium.

The most significant difficulty in this program comes from the fact that  $\tilde{g}$  and its derivatives involve powers of the complex logarithm  $\ln(z)$ . The function  $\ln(z)$  is not analytic inside the unit circle, and indeed it has an essential singularity rather than a pole at 0. Moreover, no single valued branch of the logarithm can be defined on the punctured disk.

Of course these issues can be resolved satisfactorily using standard arguments from complex analysis. The idea would be to choose a “key hole” contour as in Figure 1. Indeed, since Equation (4) has only finitely many unstable eigenvalues there must be a line segment from the origin to the unit circle in  $\mathbb{C}$  which does not intersect any zero of  $\tilde{g}$ . For the sake of simplicity let us assume that this line segment is the negative real interval  $[-1, 0]$  as drawn in Figure 1.

Assume for example that we have located two zeros  $\lambda_1, \lambda_2$  of  $\tilde{g}$  in the unit disk and that they are not on  $[-1, 0]$ . Taking a semi-circle of radius  $R < \min(|\lambda_1|, |\lambda_2|)$  and removing the strip of width  $0 < d < R$  about  $[-1, 0]$  as in Figure 1, we see that  $\tilde{g}$  (in this example) is analytic inside the curve  $\Gamma = \alpha + \beta + \gamma + \delta$ . If  $\tilde{g}$  has no poles in the unit disk then the argument principle can be used to prove that there are either exactly two zeros in the region enclosed by  $\Gamma$  or, in the case that the contour integral results in a count different from 2, that we have missed some eigenvalues.

The strategy just described does not give the Morse index, as there could be zeros of  $\tilde{g}$  inside the smaller circle of radius  $R$  or along the strip of width  $d$ . Yet by taking the limit as  $R, d \rightarrow 0$  we will obtain the correct index, provided there are no poles or zeros along the limiting contour. Calculations based on interval arithmetic could be used to rule out zeros/poles along the contour and one can write down explicit formulas for the integrals of powers of  $\ln(z)$  around  $\beta, \gamma$ , and  $\delta$  so that the limits can be computed mechanically and incorporated into a computer program.

Even though there is no fundamental obstruction to this approach, it is clear that the inte-

grands involved become increasingly complex as the dimension of the system increases. Writing a general automated code to compute the necessary line integrals over appropriate key hole contours would be a serious programming task. (See Section 5 for some brief comparisons in the one dimensional case). In fact, even expanding the determinants symbolically for the cases of four or five dimensions is cumbersome, so that general purpose solution would need to use either a symbolic manipulation package or to compute validated determinants and their derivatives numerically.

Finally we mention yet another approach, which is to work with Equation (11) instead of the compactified characteristic equation. In this case the unstable eigenvalues are the complex zeros in the right half plane. Equation (11) involves powers of the exponential rather than the logarithm, and its zeros can be counted again via the argument principle. This is not a dramatic improvement over the approach outlined above because one should take a line integral enclosing the entire right half plane. Smaller contours are sufficient if we have explicit bounds on the size of a half circle in the right half plane containing the eigenvalues. However any sharp general estimates will depend in a complicated way on the entries of  $K_1$  and  $K_2$ , and simpler estimates will still require integrating over large semi-circles.

So, while it is possible to perform eigenvalue counts using complex analysis of scalar equations it appears that developing general purpose software for the job is an involved task. In Section 4 we propose an alternative approach which solves the counting problem just discussed but which scales well with respect to the dimension of the problem. The method also provides accurate numerical approximation of the spectrum, and so solves both Problem 1 and 2 from the Introduction. The price of the method proposed in Section 4 is that it abandons the characteristic equation and returns to the functional analytic context from which the DDE came.

## 4 A functional analytic approach to the spectral analysis

Recall from Section 2.4 that any solution  $\xi \in C([-\tau, 0])$  of the eigenvalue problem given in Equation (8) is actually real analytic on  $[-\tau, 0]$ . Then we study the problem in this restricted space. The question now is how should we discretize the space of real analytic functions on  $[-\tau, 0]$ ? One obvious choice is to use power series. This has some disadvantages, as a particular function  $y(t)$ , real analytic on  $[-\tau, 0]$  may require many power series to represent.

This depends on the distance to the nearest pole in the complex plane. More precisely, let  $z_0 \in \mathbb{C}$  denote the nearest pole of  $y$  and suppose that  $\text{dist}([-\tau, 0], z_0) < \tau$ . Then a power series expansion of the form

$$y(t) = \sum_{n=0}^{\infty} a_n(t + \tau)^n,$$

must have radius of convergence smaller than  $\tau$ . And more than one power series is needed to represent  $y$  in all of  $[-\tau, 0]$ , though it can always be done with a finite number of series.

A better choice is to use a Chebyshev series representation. After rescaling  $y$  to the domain  $[-1, 1]$ , recall that the Chebyshev series expansion for  $y : [-1, 1] \rightarrow \mathbb{R}^d$  is

$$h(t) = a_0 + 2 \sum_{n \geq 1} a_n T_n(t), \quad a_n \in \mathbb{R}^d \tag{26}$$

where  $T_0(t) = 1$ ,  $T_1(t) = t$  and  $T_{n+1}(t) = 2tT_n(t) - T_{n-1}(t)$ , for  $n \geq 1$ . It is a result of fundamental importance that if  $y$  is real analytic on  $[-1, 1]$  then the Chebyshev series converges on the largest ellipse with foci at  $(-1, 0)$  and  $(1, 0)$  which does not intersect any poles of  $y$ . This is the so-called *Bernstein ellipse*. Put another way, suppose that  $z_0$  is the nearest pole of  $y$ . Then the semi-minor axis of the Bernstein ellipse is no smaller than  $\rho = |\text{imag}(z_0)|$ . The coefficients  $\{a_n\}_{n=0}^{\infty}$  decay exponentially fast, with rate determined by  $\rho$ . This is a major advantage in the

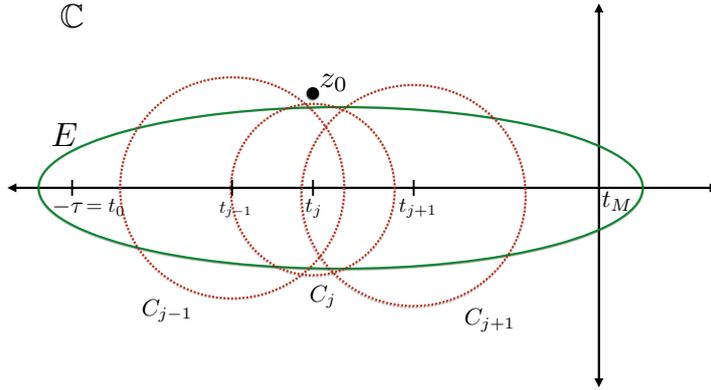


Figure 2: **Representation of real analytic functions:** Illustration of the complex extension of a real analytic function  $y(t)$  defined on  $[-\tau, 0]$  whose nearest complex pole is at  $z_0 \in \mathbb{C}$  with  $\text{dist}([-\tau, 0], z_0) < \tau$ . There is a mesh  $\tau_0 = -\tau, \dots, \tau_M = 0$ , and power series expansions  $y_0(t), \dots, y_M(t)$  so that  $y(t) = y_j(t)$  for any  $t$  where the power series converges. Each power series  $y_j(t)$  is centered at the point  $t_j$  and converges on a disk of radius  $R_j = |t_j - z_0|$ . We refer to these disks as  $C_j$ ,  $0 \leq j \leq M$ . The same function  $y(t)$  can be represented by a Chebyshev series converging absolutely and uniformly on the Bernstein ellipse  $E$  whose semi-minor axis is at least  $\rho = |\text{imag}(z_0)|$ . The fact that one Chebyshev series is always sufficient to represent a real analytic function on  $[-\tau, 0]$  is a significant advantage for our discretization scheme.

discussion to come, hence we adopt the Chebyshev framework from now on. The preceding discussion is recapitulated graphically in Figure 2.

#### 4.1 Banach spaces of infinite sequences and approximation of compact linear operators

Let  $a = \{a_n\}_{n=0}^{\infty}$  be an infinite sequence of complex vectors  $a_n \in \mathbb{C}^d$ . Choose any norm  $\|\cdot\|_d$  in  $\mathbb{C}^d$ . Given a sequence of weights  $\omega = (\omega_n)_{n \geq 0}$  with  $\omega_n > 0$ , define the  $\omega$ -weighted little-ell-one norm

$$\|a\|_{\omega} = \sum_{n=0}^{\infty} \|a_n\|_d \omega_n.$$

The set of all sequences with finite  $\omega$ -weighted norm is a Banach space which we denote by  $\ell_{\omega}^1$ .

**Remark 4.1.** *As any solution of the eigenvalue problem given in Equation (8) is real analytic, its associated sequence of Chebyshev coefficients  $(a_n)_{n \geq 0}$  (i.e. see (26)) has the property that the associate sequence of real numbers  $\{\|a_n\|_d\}_{n \geq 0}$  decays geometrically to 0. The weights  $\omega = \{\omega_n\}_{n \geq 0}$  are therefore chosen to incorporate that property. More precisely, consider a number  $\nu > 1$  and let*

$$\omega_n = \nu^n, \quad n \geq 0.$$

*This choice of weights lead to the Banach space*

$$\ell_{\omega}^1 = \ell_{\nu}^1 = \left\{ a = \{a_n\}_{n \geq 0} : a_n \in \mathbb{C}^d \quad \text{and} \quad \|a\|_{\nu} \stackrel{\text{def}}{=} \sum_{n \geq 0} \|a_n\|_d \nu^n < \infty \right\}.$$

The following result provides a general formula to compute a bound for the norm of bounded linear operators on  $\ell_{\omega}^1$ . Let  $B(\ell_{\omega}^1)$  denote the Banach space of all bounded linear operators from  $\ell_{\omega}^1$  to itself. We have the following proposition.

**Proposition 4.2.** Let  $A = \{a_{m,n}\}_{m,n \geq 0}$  be a bi-infinite sequence with  $a_{m,n} \in M_d(\mathbb{C})$  a  $d \times d$  complex-valued matrix for each  $(m,n) \in \mathbb{N}^2$ . Define the linear mapping  $A$  on  $\ell_\omega^1$  by the formula

$$(Ah)_m = \sum_{n \geq 0} a_{m,n} h_n \in \mathbb{C}^d,$$

for  $h \in \ell_\omega^1$  and  $m \geq 0$ . Then  $A \in B(\ell_\omega^1)$  with

$$\|A\|_{B(\ell_\omega^1)} \leq K_A \stackrel{\text{def}}{=} \sup_{n \geq 0} \frac{1}{\omega_n} \left( \sum_{m \geq 0} \|a_{m,n}\|_d \omega_m \right), \quad (27)$$

where  $\|a_{m,n}\|_d$  denotes the matrix norm of  $a_{m,n} \in M_d(\mathbb{C})$  induced by the norm  $\|\cdot\|_d$ .

*Proof.* Given  $b = \{b_n\}_{n \in \mathbb{N}} \in \ell_\omega^1$ ,

$$\begin{aligned} \|A\|_{B(\ell_\omega^1)} &= \sup_{\|b\|_\omega=1} \|Ab\|_\omega \\ &= \sup_{\|b\|_\omega=1} \sum_{m \geq 0} \left\| \sum_{n \geq 0} a_{m,n} b_n \right\|_d \omega_m \\ &\leq \sup_{\|b\|_\omega=1} \sum_{m \geq 0} \sum_{n \geq 0} \|a_{m,n} \omega_m b_n\|_d \\ &= \sup_{\|b\|_\omega=1} \sum_{n \geq 0} \sum_{m \geq 0} \|a_{m,n} \omega_m b_n\|_d \\ &\leq \sup_{\|b\|_\omega=1} \sum_{n \geq 0} \left( \sum_{m \geq 0} \|a_{m,n}\|_d \omega_m \right) \|b_n\|_d \\ &= \sup_{\|b\|_\omega=1} \sum_{n \geq 0} c_n \|b_n\|_d, \end{aligned}$$

where the third equality follows from Fubini's theorem for infinite sums, and where

$$c_n \stackrel{\text{def}}{=} \sum_{m \geq 0} \|a_{m,n}\|_d \omega_m.$$

By definition of  $K_A$  in (27), observe that

$$K_A = \sup_{n \in \mathbb{N}} \frac{c_n}{\omega_n} \quad \text{and that } c_n \leq K_A \omega_n, \quad \text{for all } n \geq 0.$$

Hence,

$$\|A\|_{B(\ell_\omega^1)} \leq \sup_{\|b\|_\omega=1} \sum_{n \geq 0} c_n \|b_n\|_d \leq K_A \sup_{\|b\|_\omega=1} \sum_{n \geq 0} \|b_n\|_d \omega_n \leq K_A \sup_{\|b\|_\omega=1} \|b\|_\omega = K_A. \quad \square$$

Let  $N \in \mathbb{N}$  and define the projection  $\pi_N, \pi_\infty: \ell_\omega^1 \rightarrow \ell_\omega^1$  by

$$\pi_N(h)_n = \begin{cases} h_n & 0 \leq n \leq N \\ 0 & n \geq N+1 \end{cases}$$

and

$$\pi_\infty(h)_n = \begin{cases} 0 & 0 \leq n \leq N \\ h_n & n \geq N+1. \end{cases}$$

Note that for each  $n \in \mathbb{N}$ ,  $\pi_N(h)_n \in \mathbb{C}^d$ . We note that  $\pi_N(\ell_\omega^1)$  is a finite dimensional complex vector space which we can identify with  $\mathbb{C}^{d(N+1)}$ .

For any  $h \in \ell_\omega^1$  we have that

$$h = \pi_N(h) + \pi_\infty(h),$$

so that

$$\ell_\omega^1 = \pi_N(\ell_\omega^1) \oplus \pi_\infty(\ell_\omega^1).$$

It is clear that  $\pi_n, \pi_\infty$  are bounded linear projection operators. For  $h \in \ell_\omega^1$  we write  $h^N = \pi_N(h)$  and  $h^\infty = \pi_\infty(h)$ . Then  $h = h^N + h^\infty$  and we sometimes identify  $h^N$  with its natural inclusion into  $\mathbb{C}^{d(N+1)}$ , especially when talking about numerics. We think of  $h^\infty$  as “the tail” of the sequence  $h$ .

We are interested in closed linear subspaces of  $C([-\tau, 0])$  isomorphic to  $\ell_\omega^1$ . Suppose that  $X$  is such a subspace, and hence a Banach space in its own right, and let  $\mathcal{I}: X \rightarrow \ell_\omega^1$  denote the isomorphism. Then the map  $T: C([-\tau, 0]) \times C([-\tau, 0]) \rightarrow C([-\tau, 0])$  induces a mapping  $\tilde{T}: \ell_\omega^1 \times \ell_\omega^1 \rightarrow \ell_\omega^1$  by the formula

$$\tilde{T}(u, v) = \mathcal{I} \left[ T \left( \mathcal{I}^{-1}(u), \mathcal{I}^{-1}(v) \right) \right].$$

If  $x(t) = c \in \mathbb{R}^d$  is a constant function with  $T(c, c) = c$  then  $\mathcal{I}(c)$  is a fixed point of  $\tilde{T}$ . Moreover the bounded linear operators  $D_1T(c, c), D_2T(c, c), DF(c): C([-\tau, 0]) \rightarrow C([-\tau, 0])$  induce bounded linear operators on  $\ell_\omega^1$  by similar formulae. In the sequel we suppress the use of the isomorphism  $\mathcal{I}$  and identify these bounded linear operators with the sequence space operators they induce.

Let  $M^N$  be an  $d(N+1) \times d(N+1)$  matrix. Then  $M^N$  induces a compact linear operator  $M: \ell_\omega^1 \rightarrow \ell_\omega^1$  by the formula

$$[Mh]_n = \begin{cases} [M^N h^N]_n & 0 \leq n \leq N \\ 0 & n \geq N+1. \end{cases}$$

We have that

$$\text{spec}(M) = \text{spec}(M^N) \cup \{0\},$$

where 0 is an eigenvalue of infinite multiplicity. That is, the spectrum of the matrix  $M^N$  determines the spectrum of the bounded linear operator  $M$ .

We are interested in the case where  $M^N$  is an approximation of the operator  $DF(x_0)$ , where  $x_0$  is the spectral sequence associated with a fixed point of  $F$ . Recall from Section 2.3 that we do not have explicit access to the mapping  $F$ , which is only implicitly defined through a fixed point operator  $T$ . The next theorem allows us to draw conclusions about the spectrum of  $DF(x_0)$  given knowledge of the spectrum of a good enough approximating matrix  $M^N$ . The meaning of “good enough” has to do with the location of the eigenvalues of  $M^N$ , and also some bounds on the induced operators  $\text{Id} - D_1T(x_0, x_0)$  and  $D_2T(x_0, x_0)$ . The important thing is that the hypotheses of the theorem involve no information about  $F$  or  $DF(x_0)$ . Only the fixed point operator  $T$  and its partial derivatives.

**Theorem 4.3.** *Suppose that  $M: \ell_\omega^1 \rightarrow \ell_\omega^1$  is a compact linear operator of the form*

$$(Mh)_n = \begin{cases} [M^N h^N]_n & 0 \leq n \leq N \\ 0 & n \geq N+1. \end{cases}$$

*Given  $r > 0$  assume that none of the non-zero eigenvalues of  $M^N$  lie on the circle of radius  $r$  in  $\mathbb{C}$ , so that the numbers  $\lambda_j - re^{i\theta}$  are non-zero for each  $\theta \in [0, 2\pi]$ . Assume that  $C_1, C_2, C_3 > 0$  are constants with*

$$\max \left( \sup_{\theta \in [0, 2\pi]} \left\| \left( M^N - re^{i\theta} \text{Id}_N \right)^{-1} \right\|_{B(\ell_\omega^1)}, \frac{1}{r} \right) \leq C_1, \quad (28)$$

$$\left\| (\text{Id} - D_1T(x_0, x_0))^{-1} \right\|_{B(\ell_\omega^1)} \leq C_2,$$

and

$$\left\| (\text{Id} - D_1T(x_0, x_0))M - D_2T(x_0, x_0) \right\|_{B(\ell_\omega^1)} \leq C_3.$$

If  $C_1C_2C_3 < 1$  then  $M$  and  $DF(x_0)$  have the same number of eigenvalues in the complement of the closed disk of radius  $r$ .

*Proof.* Define the homotopy

$$H(s) = (1 - s)M + sDF(x_0),$$

from  $M$  to  $DF(x_0)$ . Clearly then  $H(0) = M$  and  $H(1) = DF(x_0)$ , and  $H(s)$  is continuous for  $s \in [0, 1]$ .  $H(0)$  and  $M$  trivially have the same eigenvalues, and hence the same number of eigenvalues in the complement of the closed disk of radius  $r$ . The argument given in the proof of Lemma 5.3 in [24] shows that  $DF(x_0)$  and  $M$  have a different number of eigenvalues in the complement of the disk if and only if there is a crossing during the homotopy. That is if and only if there is an  $s_0 \in (0, 1]$  and a  $\lambda_0 \in \mathbb{C}$  with  $|\lambda_0| = r$  having that  $\lambda_0$  is an eigenvalue of  $H(s_0)$ .

We show that this cannot happen by showing that  $re^{i\theta}$  is never an eigenvalue of  $H$ . That is, we show that  $H(s) - re^{i\theta}\text{Id}$  is boundedly invertible for all  $s \in [0, 1]$  and  $\theta \in [0, 2\pi]$ . To see this define the family of linear operator

$$N(\theta) = M - re^{i\theta}\text{Id}$$

and note that  $N(\theta)$  is a bounded linear operator. To see this we proceed as follows. For fixed  $\theta \in [0, 1]$  and  $q = (q^N, q^\infty) \in \ell_\omega^1$  we seek an  $h \in \ell_\omega^1$  so that

$$N(\theta)h = q,$$

or equivalently

$$(M^N - re^{i\theta}\text{Id}_N)h^N = q^N,$$

and

$$-re^{i\theta}h^\infty = q^\infty$$

In the tail we have that  $h^\infty = -q^\infty e^{-i\theta}/r$ . In the finite dimensional projection we have that  $(M^N - re^{i\theta}\text{Id}_N)$  is invertible, and hence boundedly invertible, precisely by the assumption that  $M^N$  has no eigenvalues on the circle of radius  $r$ . Then

$$h^N = (M^N - re^{i\theta}\text{Id}_N)^{-1}q^N.$$

Then

$$\left\| N(\theta)^{-1} \right\|_{B(\ell_\omega^1)} \leq \max \left( \sup_{\theta \in [0, 2\pi]} \left\| (M^N - re^{i\theta}\text{Id}_N)^{-1} \right\|_{B(\ell_\omega^1)}, \frac{1}{r} \right) \leq C_1,$$

by the definition of  $C_1$ .

Now we consider the difference

$$\begin{aligned} M - DF(x_0) &= (\text{Id} - D_1T(x_0, x_0))^{-1}(\text{Id} - D_1T(x_0, x_0))(M - DF(x_0)) \\ &= (\text{Id} - D_1T(x_0, x_0))^{-1} [(\text{Id} - D_1T(x_0, x_0))M - (\text{Id} - D_1T(x_0, x_0))DF(x_0)] \\ &= (\text{Id} - D_1T(x_0, x_0))^{-1} [(\text{Id} - D_1T(x_0, x_0))M - D_2T(x_0, x_0)]. \end{aligned}$$

Taking norms gives

$$\begin{aligned} \|M - DF(x_0)\|_{B(\ell_\omega^1)} &\leq \left\| (\text{Id} - D_1T(x_0, x_0))^{-1} [(\text{Id} - D_1T(x_0, x_0))M - D_2T(x_0, x_0)] \right\|_{B(\ell_\omega^1)} \\ &\leq \left\| (\text{Id} - D_1T(x_0, x_0))^{-1} \right\|_{B(\ell_\omega^1)} \|(\text{Id} - D_1T(x_0, x_0))M - D_2T(x_0, x_0)\|_{B(\ell_\omega^1)} \\ &\leq C_2C_3. \end{aligned}$$

Since  $C_1 C_2 C_3 < 1$  we have that

$$\left\| sN^{-1}(\theta)(M - DF(x_0)) \right\|_{B(\ell_\omega^1)} \leq \|N^{-1}(\theta)\|_{B(\ell_\omega^1)} \|M - DF(x_0)\|_{B(\ell_\omega^1)} \leq C_1 C_2 C_3 < 1,$$

for  $s \in [0, 1]$ , hence the operator  $\text{Id} - sN^{-1}(\theta)(M - DF(x_0))$  is boundedly invertible for  $s \in [0, 1]$  with

$$\left\| \left[ \text{Id} - sN^{-1}(\theta)(M - DF(x_0)) \right]^{-1} \right\|_{B(\ell_\omega^1)} \leq \frac{1}{1 - C_1 C_2 C_3},$$

by the Neumann series theorem.

To complete the argument we now consider the homotopy

$$\begin{aligned} H(s) - re^{i\theta}\text{Id} &= M - s(M - DF(x_0)) - re^{i\theta}\text{Id} \\ &= (M - re^{i\theta}\text{Id}) - s(M - DF(x_0)) \\ &= N(\theta) \left( \text{Id} - sN^{-1}(\theta)(M - DF(x_0)) \right). \end{aligned}$$

Since  $N(\theta)$  and  $\text{Id} - sN^{-1}(\theta)(M - DF(x_0))$  are boundedly invertible for  $s \in [0, 1]$ ,  $\theta \in [0, 2\pi]$  we have that

$$\left[ H(s) - re^{i\theta}\text{Id} \right]^{-1} = \left( \text{Id} - sN^{-1}(\theta)(M - DF(x_0)) \right)^{-1} N^{-1}(\theta)$$

with the bound

$$\begin{aligned} \left\| \left[ H(s) - re^{i\theta}\text{Id} \right]^{-1} \right\|_{B(\ell_\omega^1)} &\leq \left\| \left( \text{Id} - sN^{-1}(\theta)(M - DF(x_0)) \right)^{-1} \right\|_{B(\ell_\omega^1)} \left\| N^{-1}(\theta) \right\|_{B(\ell_\omega^1)} \\ &\leq \frac{C_1}{1 - C_1 C_2 C_3}. \end{aligned}$$

Then indeed  $H(s) - e^{i\theta}\text{Id}$  is boundedly invertible for all  $s \in [0, 1]$ ,  $\theta \in [0, 2\pi]$ , and it follows that  $re^{i\theta}$  is never an eigenvalue of  $H(s)$ .  $\square$

**Remark 4.4** (Real systems and complex conjugate eigenvalues). Observe that when  $f: \mathbb{R}^d \rightarrow \mathbb{R}^d$  is real then we are only interested in real equilibrium solutions  $c \in \mathbb{R}^d$  and it follows that the matrices  $K_1 = \partial_1 f(c, c)$  and  $K_2 = \partial_2 f(c, c)$  have real entries. In this case any complex zeros of the determinant is a polynomial with real coefficients and the roots of the characteristic equation occur in complex conjugate pairs. It follows that the spectrum is symmetric about the real axis so the supremum in the definition of  $C_1$  needs only be taken over the interval  $[0, \pi]$ , reducing the computational cost by a factor of 2.

## 4.2 Chebyshev series discretization

Let us map the time interval  $t \in [-\tau, 0]$  to  $\tilde{t} \in [-1, 1]$  via  $\tilde{t} = \frac{2}{\tau}t + 1$ . Set  $\tilde{h}(\tilde{t}) \stackrel{\text{def}}{=} h\left(\frac{\tau}{2}(\tilde{t} - 1)\right) = h(t)$ . Hence, for  $t \in [-\tau, 0]$ ,

$$\begin{aligned} (\text{Id} - D_1 T(c, c)) h(t) &= h(t) - \int_{-\tau}^t K_1 h(s) ds \\ &= \tilde{h}(\tilde{t}) - \frac{\tau}{2} K_1 \int_{-1}^{\tilde{t}} \tilde{h}(\tilde{s}) d\tilde{s}. \end{aligned}$$

For sake of simplicity of the presentation, we simply identify  $h(t)$  and  $\tilde{h}(\tilde{t})$ . Therefore,

$$(\text{Id} - D_1 T(c, c)) h(t) = h(t) - \frac{\tau K_1}{2} \int_{-1}^t h(s) ds.$$

Expand  $h : [-1, 1] \rightarrow \mathbb{R}^d$  with Chebyshev series

$$h(t) = a_0 + 2 \sum_{n \geq 1} a_n T_n(t), \quad a_n \in \mathbb{R}^d.$$

Using that  $\int T_0(s) ds = T_1(s) + \text{const.}$ ,  $\int T_1(s) ds = \frac{T_0(s) + T_2(s)}{4} + \text{const.}$  and  $\int T_n(s) ds = \frac{1}{2} \left( \frac{T_{n+1}(s)}{n+1} - \frac{T_{n-1}(s)}{n-1} \right) + \text{const.}$  for  $n \geq 2$ , yields

$$\int_{-1}^t h(s) ds = \left( a_0 - \frac{a_1}{2} - 2 \sum_{k \geq 2} \frac{(-1)^k}{k^2 - 1} a_k \right) T_0(t) + 2 \sum_{n \geq 1} \frac{1}{2n} (a_{n-1} - a_{n+1}) T_n(t). \quad (29)$$

Hence,

$$\begin{aligned} (\text{Id} - D_1 T(c, c)) h(t) &= h(t) - \frac{\tau K_1}{2} \int_{-1}^t h(s) ds \\ &= a_0 + 2 \sum_{n \geq 1} a_n T_n(t) - \frac{\tau K_1}{2} \left( a_0 - \frac{a_1}{2} - 2 \sum_{k \geq 2} \frac{(-1)^k}{k^2 - 1} a_k \right) T_0(t) \\ &\quad - 2 \frac{\tau K_1}{2} \sum_{n \geq 1} \frac{1}{2n} (a_{n-1} - a_{n+1}) T_n(t) \\ &= \left[ a_0 - \frac{\tau K_1}{2} \left( a_0 - \frac{a_1}{2} - 2 \sum_{k \geq 2} \frac{(-1)^k}{k^2 - 1} a_k \right) \right] T_0(t) \\ &\quad + 2 \sum_{n \geq 1} \left( -\frac{\tau K_1}{4n} a_{n-1} + a_n + \frac{\tau K_1}{4n} a_{n+1} \right) T_n(t), \end{aligned}$$

which has a matrix representation

$$M_1 = \begin{pmatrix} \text{Id}_d - \frac{\tau K_1}{2} & \frac{\tau K_1}{4} & \frac{\tau K_1}{3} & \dots & \frac{\tau K_1 (-1)^n}{n^2 - 1} & \dots \\ -\frac{\tau K_1}{4} & \text{Id}_d & \frac{\tau K_1}{4} & 0 & 0 & \dots \\ 0 & -\frac{\tau K_1}{8} & \text{Id}_d & \frac{\tau K_1}{8} & 0 & \dots \\ 0 & 0 & \ddots & \ddots & \ddots & \dots \\ 0 & 0 & 0 & -\frac{\tau K_1}{4n} & \text{Id}_d & \frac{\tau K_1}{4n} \\ 0 & 0 & 0 & 0 & -\frac{\tau K_1}{4(n+1)} & \text{Id}_d \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Moreover,

$$\begin{aligned} D_2 T(c, c) h(t) &= h(1) + \frac{\tau K_2}{2} \int_{-1}^t h(s) ds \\ &= \left[ a_0 + 2 \sum_{n \geq 1} a_n + \frac{\tau K_2}{2} \left( a_0 - \frac{a_1}{2} - 2 \sum_{k \geq 2} \frac{(-1)^k}{k^2 - 1} a_k \right) \right] T_0(t) \\ &\quad + 2 \sum_{n \geq 1} \left( \frac{\tau K_2}{4n} a_{n-1} - \frac{\tau K_2}{4n} a_{n+1} \right) T_n(t) \end{aligned}$$

which has a matrix representation

$$M_2 = \begin{pmatrix} \text{Id}_d + \frac{\tau K_2}{2} & 2\text{Id}_d - \frac{\tau K_2}{4} & 2\text{Id}_d - \frac{\tau K_2}{3} & \cdots & 2\text{Id}_d - \frac{\tau(-1)^n K_2}{n^2-1} & \cdots \\ \frac{\tau K_2}{4} & 0 & -\frac{\tau K_2}{4} & 0 & 0 & \cdots \\ 0 & \frac{\tau K_2}{8} & 0 & -\frac{\tau K_2}{8} & 0 & \cdots \\ 0 & 0 & \ddots & \ddots & \ddots & \cdots \\ 0 & 0 & 0 & \frac{\tau K_2}{4n} & 0 & -\frac{\tau K_2}{4n} \\ 0 & 0 & 0 & 0 & \frac{\tau K_2}{4(n+1)} & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Truncating to  $N$  modes gives the  $d(N+1) \times d(N+1)$  matrices

$$M_1^N = \begin{pmatrix} \text{Id}_d - \frac{\tau K_1}{2} & \frac{\tau K_1}{4} & \frac{\tau K_1}{3} & \cdots & \frac{\tau(-1)^N K_1}{N^2-1} \\ -\frac{\tau K_1}{4} & \text{Id}_d & \frac{\tau K_1}{4} & 0 & 0 \\ 0 & -\frac{\tau K_1}{8} & \text{Id}_d & \frac{\tau K_1}{8} & 0 \\ 0 & 0 & \ddots & \ddots & \frac{\tau K_1}{4(N-1)} \\ 0 & 0 & 0 & -\frac{\tau K_1}{4N} & \text{Id}_d \end{pmatrix} \quad (30)$$

and

$$M_2^N = \begin{pmatrix} \text{Id}_d + \frac{\tau K_2}{2} & 2\text{Id}_d - \frac{\tau K_2}{4} & 2\text{Id}_d - \frac{\tau K_2}{3} & \cdots & 2\text{Id}_d - \frac{\tau(-1)^N K_2}{N^2-1} \\ \frac{\tau K_2}{4} & 0 & -\frac{\tau K_2}{4} & 0 & 0 \\ 0 & \frac{\tau K_2}{8} & 0 & -\frac{\tau K_2}{8} & 0 \\ 0 & 0 & \ddots & \ddots & -\frac{\tau K_2}{4(N-1)} \\ 0 & 0 & 0 & \frac{\tau K_2}{4N} & 0 \end{pmatrix}. \quad (31)$$

Using (30) and (31), let

$$B^N \stackrel{\text{def}}{=} (M_1^N)^{-1} \quad \text{and} \quad M^N \stackrel{\text{def}}{=} B^N M_2^N = (M_1^N)^{-1} M_2^N. \quad (32)$$

**Lemma 4.5.** *Recalling the definition of  $B^N$  in (32), define the bounded linear operator  $B : \ell_\omega^1 \rightarrow \ell_\omega^1$  by*

$$(Bh)_n = \begin{cases} (B^N h^N)_n & 0 \leq n \leq N \\ h_n & n \geq N+1. \end{cases}$$

For  $n = 0, \dots, N$ , denote by  $B_n^N = (B_{m,n}^N)_{m=0}^N \in \mathbb{R}^{N+1}$  the  $n^{\text{th}}$  column of  $B^N$ . Let  $\check{\omega}$  and  $\hat{\omega}$  two positive numbers satisfying

$$\sup_{n \geq N+2} \frac{\omega_{n-1}}{\omega_n} \leq \check{\omega} \quad \text{and} \quad \sup_{n \geq N+2} \frac{\omega_{n+1}}{\omega_n} \leq \hat{\omega}.$$

Let

$$\rho_n \stackrel{\text{def}}{=} \begin{cases} 0, & n = 0, \dots, N-1 \\ \frac{\tau \|K_1\|_d \omega_{N+1}}{4(N+1) \omega_N}, & n = N \\ \frac{1}{\omega_{N+1}} \sum_{m=0}^N \left\| \frac{\tau(-1)^{N+1}}{(N+1)^2-1} B_{m,0}^N K_1 + \frac{\tau}{4N} B_{m,N}^N K_1 \right\|_d \omega_m + \frac{\tau \|K_1\|_d \omega_{N+2}}{4(N+2) \omega_{N+1}}, & n = N+1 \\ \frac{1}{\omega_{N+2}} \frac{\tau}{(N+2)^2-1} \left( \sum_{m=0}^N \|B_{m,0}^N K_1\|_d \omega_m \right) + \check{\omega} \frac{\tau \|K_1\|_d}{4(N+1)} + \hat{\omega} \frac{\tau \|K_1\|_d}{4(N+3)}, & n = \infty \end{cases} \quad (33)$$

and let

$$\rho \stackrel{\text{def}}{=} \max\{\rho_0, \rho_1, \dots, \rho_{N+1}, \rho_\infty\}. \quad (34)$$

If  $\rho < 1$ , then letting

$$C_2 \stackrel{\text{def}}{=} \frac{\max\{\|B^N\|_{B(\ell_\omega^1)}, 1\}}{1 - \rho} \quad (35)$$

yields (recall that  $M_1$  is the operator representation of  $\text{Id} - D_1T(c, c)$ )

$$\|(\text{Id} - D_1T(c, c))^{-1}\|_{B(\ell_\omega^1)} = \|M_1^{-1}\|_{B(\ell_\omega^1)} \leq C_2.$$

*Proof.* The idea of the proof is to obtain a bound on  $\|M_1^{-1}\|_{B(\ell_\omega^1)}$  by considering the (computable and explicitly representable) approximate inverse  $B$  of  $M_1$ , and apply a Neumann series argument to obtain that bound. Let

$$\Lambda \stackrel{\text{def}}{=} \text{Id} - BM_1.$$

Denote  $\Lambda = \{\Lambda_{m,n}\}_{m,n \geq 0}$  where each  $\Lambda_{m,n}$  is a  $d \times d$  matrix. Recalling (27),

$$\|\Lambda\|_{B(\ell_\omega^1)} \leq K_\Lambda \stackrel{\text{def}}{=} \sup_{n \geq 0} \frac{1}{\omega_n} \left( \sum_{m \geq 0} \|\Lambda_{m,n}\|_{d\omega_m} \right).$$

For any  $n \geq 0$ , denote  $\Lambda_n = (\Lambda_{m,n})_{m \geq 0}$  the  $n^{\text{th}}$  column of  $\Lambda$ . For  $n = 0, \dots, N-1$ ,  $\Lambda_n = 0$ . For  $n = N$ ,

$$\Lambda_N = \begin{pmatrix} \top \\ 0 \\ \perp \\ \frac{\tau K_1}{4(N+1)} \\ 0 \\ 0 \\ \vdots \end{pmatrix},$$

and in this case

$$\frac{1}{\omega_n} \sum_{m \geq 0} \|\Lambda_{m,N}\|_{d\omega_m} = \frac{\tau \|K_1\|_d}{4(N+1)} \frac{\omega_{N+1}}{\omega_N} = \rho_N.$$

For  $n = N+1$ ,

$$\Lambda_{N+1} = \begin{pmatrix} \top \\ -\frac{\tau(-1)^{N+1}}{(N+1)^2-1} \left( B_{m,0}^N K_1 \right)_{m=0}^N - \frac{\tau}{4N} \left( B_{m,N}^N K_1 \right)_{m=0}^N \\ \perp \\ 0 \\ \frac{\tau K_1}{4(N+2)} \\ 0 \\ 0 \\ \vdots \end{pmatrix},$$

and in this case

$$\begin{aligned} \frac{1}{\omega_n} \sum_{m \geq 0} \|\Lambda_{m, N+1}\|_d \omega_m &\leq \frac{1}{\omega_{N+1}} \sum_{m=0}^N \left\| \frac{\tau(-1)^{N+1}}{(N+1)^2 - 1} B_{m,0}^N K_1 + \frac{\tau}{4N} B_{m,N}^N K_1 \right\|_d \omega_m \\ &+ \frac{\tau \|K_1\|_d}{4(N+2)} \frac{\omega_{N+2}}{\omega_{N+1}} = \rho_{N+1}. \end{aligned}$$

For  $n \geq N+2$ ,

$$\Lambda_n = \begin{pmatrix} \top \\ -\frac{\tau(-1)^n}{n^2-1} \left( B_{m,0}^N K_1 \right)_{m=0}^N \\ \perp \\ 0 \\ \vdots \\ 0 \\ -\frac{\tau K_1}{4(n-1)} \\ 0 \\ \frac{\tau K_1}{4(n+1)} \\ 0 \\ 0 \\ \vdots \end{pmatrix},$$

and

$$\frac{1}{\omega_n} \sum_{m \geq 0} \|\Lambda_{m,n}\|_d \omega_m \leq \frac{1}{\omega_n} \frac{\tau}{n^2-1} \left( \sum_{m=0}^N \|B_{m,0}^N K_1\|_d \omega_m \right) + \frac{\omega_{n-1}}{\omega_n} \frac{\tau \|K_1\|_d}{4(n-1)} + \frac{\omega_{n+1}}{\omega_n} \frac{\tau \|K_1\|_d}{4(n+1)}$$

and therefore

$$\begin{aligned} \sup_{n \geq N+2} \frac{1}{\omega_n} \sum_{m \geq 0} \|\Lambda_{m,n}\|_d \omega_m &\leq \frac{1}{\omega_{N+2}} \frac{\tau}{(N+2)^2 - 1} \left( \sum_{m=0}^N \|B_{m,0}^N K_1\|_d \omega_m \right) \\ &+ \check{\omega} \frac{\tau \|K_1\|_d}{4(N+1)} + \hat{\omega} \frac{\tau \|K_1\|_d}{4(N+3)} = \rho_\infty. \end{aligned}$$

Recalling (34), combining formulas from (33) and applying Proposition 4.2 yields

$$\|\text{Id} - BM_1\|_{B(\ell_\omega^1)} = \|\Lambda\|_{B(\ell_\omega^1)} \leq K_\Lambda = \sup_{n \geq 0} \frac{1}{\omega_n} \sum_{m \geq 0} \|\Lambda_{m,n}\|_d \omega_m \leq \max\{\rho_N, \rho_{N+1}, \rho_\infty\} = \rho.$$

Applying a Neumann series argument yields that

$$\left\| (\text{Id} - D_1 T(c, c))^{-1} \right\|_{B(\ell_\omega^1)} = \|M_1^{-1}\|_{B(\ell_\omega^1)} \leq \frac{\|B\|_{B(\ell_\omega^1)}}{1 - \rho} \leq \frac{\max\{\|B^N\|_{B(\ell_\omega^1)}, 1\}}{1 - \rho} = C_2. \quad \square$$

**Lemma 4.6.** *Let*

$$\rho_n \stackrel{\text{def}}{=} \begin{cases} \frac{\tau \|K_1 M_{N,n}^N\|_d \omega_{N+1}}{4(N+1)\omega_n}, & n = 0, \dots, N-1 \\ \frac{\tau \|K_1 M_{N,N}^N + K_2\|_d \omega_{N+1}}{4(N+1)\omega_N}, & n = N \\ \frac{1}{\omega_N} \left( 2 + \frac{\tau \|K_2\|_d}{(N+1)^2 - 1} \right) + \frac{\tau \|K_2\|_d \check{\omega}}{4(N-1)} + \frac{\tau \|K_2\|_d \hat{\omega}}{4(N+1)}, & n = \infty, \end{cases}$$

and let

$$C_3 \stackrel{\text{def}}{=} \max\{\rho_0, \rho_1, \dots, \rho_N, \rho_\infty\}. \quad (36)$$

Then

$$\|(\text{Id} - D_1 T(c, c))M - D_2 T(c, c)\|_{B(\ell_\omega^1)} = \|M_1 M - M_2\|_{B(\ell_\omega^1)} \leq C_3.$$

*Proof.* Denote

$$\Lambda \stackrel{\text{def}}{=} M_1 M - M_2,$$

and note that the finite dimensional block  $\Lambda^N$  of  $\Lambda$  satisfies  $\Lambda^N = M_1^N M^N - M_2^N = 0$ . The proof follows by observing that

$$\Lambda_n = \begin{pmatrix} \top \\ 0 \\ \perp \\ \frac{-\tau K_1 M_{N,n}^N}{4(N+1)} \\ 0 \\ 0 \\ \vdots \end{pmatrix} \text{ for } n = 0, \dots, N-1, \quad \Lambda_N = \begin{pmatrix} \top \\ 0 \\ \perp \\ \frac{-\tau(K_1 M_{N,N}^N + K_2)}{4(N+1)} \\ 0 \\ 0 \\ \vdots \end{pmatrix}$$

and

$$\Lambda_n = \begin{pmatrix} -2\text{Id}_d + \frac{\tau(-1)^n K_2}{n^2 - 1} \\ 0 \\ \vdots \\ 0 \\ \frac{\tau K_2}{4(n-1)} \\ 0 \\ -\frac{\tau K_2}{4(n+1)} \\ 0 \\ 0 \\ \vdots \end{pmatrix} \text{ for } n \geq N+1,$$

and by using Proposition 4.2. □

### 4.3 Applications of the Chebyshev series discretization

In this section, we present applications of the Chebyshev series discretization approach to rigorously compute the number of eigenvalues outside circles of prescribed radii centered at 0 in the complex plane. We apply our approach to Mackey-Glass (13), cubic Ikeda-Matsumoto (16), delayed van der Pol (17) and the predator-prey equation (20). Let us now present a rigorous computational procedure.

After choosing a smooth function  $f: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  and a delay  $\tau > 0$ , assume that  $u_0 \equiv c \in \mathbb{R}^d$  is an equilibrium solution to  $y'(t) = f(y(t), y(t - \tau))$ , that is  $f(c, c) = 0$ . Define the real-valued  $d \times d$  matrices  $K_1$  and  $K_2$  as in (7), that is  $K_1 \stackrel{\text{def}}{=} \partial_1 f(c, c)$  and  $K_2 \stackrel{\text{def}}{=} \partial_2 f(c, c)$ . Given a finite dimensional Chebyshev projection number  $N$ , define the finite dimensional real-valued matrices  $M_1^N$  and  $M_2^N$  given respectively by (30) and (31). Define the real-valued matrices  $B^N$  and  $M^N$  as in (32), that is  $B^N \stackrel{\text{def}}{=} (M_1^N)^{-1}$  and  $M^N \stackrel{\text{def}}{=} (M_1^N)^{-1} M_2^N$ . Choose a sequence of weights  $\omega = (\omega_n)_{n \geq 0}$ . In all our computations, we fix a number  $\nu > 1$  and set  $\omega_n = \nu^n$  (see Remark 4.1). Hence, the Banach space we work with is  $\ell_\omega^1 = \ell_\nu^1$  and represents analytic functions. Use interval arithmetic to compute the constant  $C_2$  satisfying (35) and the constant  $C_3$  satisfying (36). Next, we compute the constant  $C_1$  satisfying (28). Note that since the matrix  $M^N$  is a real-valued matrix, its eigenvalues come in complex conjugate pairs. From Remark 4.4, it is enough to perform the computation of  $C_1$  in (37) over the interval  $[0, \pi]$  instead of  $[0, 2\pi]$  as in (28). Fix a mesh size  $m$  and consider a partition

$$0 = t_1 < t_2 < \cdots < t_{m-1} < t_m = \pi$$

of the interval  $[0, \pi]$ . For a fixed radius value  $r > 0$ , use interval arithmetic to compute  $C_1$  satisfying

$$\max \left( \max_{j=1, \dots, m-1} \sup_{\theta \in [t_j, t_{j+1}]} \left\| (M^N - r e^{i\theta} \text{Id})^{-1} \right\|_{B(\ell_\omega^1)}, \frac{1}{r} \right) \leq C_1. \quad (37)$$

Let  $M: \ell_\omega^1 \rightarrow \ell_\omega^1$  the compact linear operator

$$(Mh)_n = \begin{cases} [M^N h^N]_n & 0 \leq n \leq N \\ 0 & n \geq N + 1. \end{cases}$$

Let

$$C \stackrel{\text{def}}{=} C_1 C_2 C_3. \quad (38)$$

If  $C < 1$  then by Theorem 4.3,  $M$  and  $DF(u_0)$  have the same number of eigenvalues in the complement of the closed disk of radius  $r$ . If  $C > 1$ , then either increase the Chebyshev dimension  $N$ , increase the mesh size  $m$  to compute  $C_1$  or change the decay rate parameter  $\nu$ , recompute the constants  $C_1$ ,  $C_2$  and  $C_3$ , define  $C$  as in (38), and try to verify that  $C < 1$ . The final step is to enclose the eigenvalues of the matrix  $M^N$  (which we do using the approach of [8]) and use that information to obtain a rigorous count for the number of eigenvalues of  $M^N$  outside the circle of radius  $r$ . This count provides the generalized Morse index  $\mu_r(u_0)$ , that is the number of eigenvalues of  $DF(u_0)$  outside the disk  $B_r(0) \subset \mathbb{C}$ . Note that  $\mu_1$  is the standard Morse index, that is the dimension of the unstable manifold of the fixed point  $u_0$ . Using the procedure described above, we proved the following result.

**Theorem 4.7.** *Consider the Mackey-Glass equation (13) at the parameter values  $\tau = 2$ ,  $\gamma = 1$ ,  $\beta = 2$  and  $\rho = 10$ . Denote by  $u_0 \equiv 0$  and  $u_1 \equiv 1$  the two steady states. Then  $\mu_1(u_0) = 1$ ,  $\mu_{0.6}(u_0) = 3$ ,  $\mu_{0.29}(u_0) = 5$ ,  $\mu_{0.2}(u_0) = 7$  and  $\mu_1(u_1) = 2$ ,  $\mu_{0.85}(u_1) = 4$ ,  $\mu_{0.46}(u_1) = 6$ ,  $\mu_{0.341}(u_1) = 8$ . In particular,  $u_0$  has a one-dimensional unstable manifold and  $u_1$  has a two-dimensional unstable manifold.*

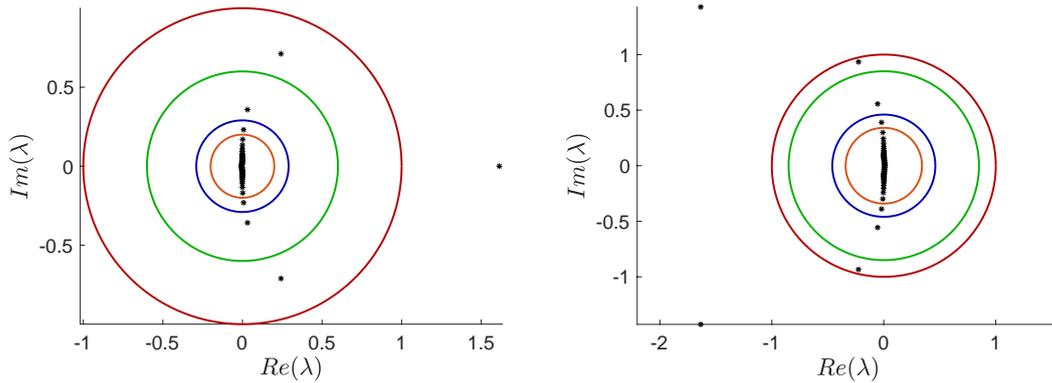


Figure 3: On the left, the spectrum computations for  $DF(u_0)$  and on the right, the spectrum computations for  $DF(u_1)$ . The circles of radii  $r$  used in the computation of the generalized Morse indices in the Mackey-Glass equation (13) at the parameter values  $\tau = 2$ ,  $\gamma = 1$ ,  $\beta = 2$  and  $\rho = 10$  are plotted. On each plot, the unit circle is the largest one and is portrayed in red.

$u_0$	$r$	$N$	$\nu$	$m$	$\mu_r(u_0)$
0	1	32	1.3	10	1
0	0.6	70	1.2	10	3
0	0.29	200	1.15	40	5
0	0.2	600	1.05	60	7
1	1	310	1.1	100	2
1	0.85	130	1.2	30	4
1	0.46	250	1.1	30	6
1	0.341	500	1.05	90	8

Table 1: Parameters used in the proof of Theorem 4.7 to obtain the generalized Morse indices in the Mackey-Glass equation (13) at the parameter values  $\tau = 2$ ,  $\gamma = 1$ ,  $\beta = 2$  and  $\rho = 10$ .

*Proof.* The proof follows by running the program `int_script_compute_spectrum_cheb.m` available at [21]. This MATLAB program requires the use of the interval arithmetic package INTLAB. The data for each proof is available in Table 1. The spectra can be visualized in Figure 3.  $\square$

Similarly, we obtain the following results.

**Theorem 4.8.** *Consider the Ikeda-Matsumoto equation (16) at the parameter value  $\tau = 1.59$ . Denote by  $u_0 \equiv 0$  and  $u_1 \equiv 1$  two steady states. Then  $\mu_1(u_0) = 1$ ,  $\mu_{0.25}(u_0) = 3$  and  $\mu_1(u_1) = 2$ ,  $\mu_{0.31}(u_1) = 4$ . In particular,  $u_0$  has a one-dimensional unstable manifold and  $u_1$  has a two-dimensional unstable manifold.*

**Theorem 4.9.** *Consider the delayed van der Pol equation (17) with parameter values  $\tau = 2$ ,  $\kappa = -1$  and  $\varepsilon = 0.15$ . Denote by  $u_0 \equiv (0, 0)^T$  a steady state. Then  $\mu_1(u_0) = 2$ , that is  $u_0$  has a two-dimensional unstable manifold.*

**Theorem 4.10.** *Consider the delayed predator-prey model (20) with parameter values  $r_1 = 2$ ,  $r_2 = 1$ ,  $a = 1$  and  $b = 1/2$ . Denote by  $u_0 \equiv (y_2, y_2)^T$  the nontrivial equilibrium given in (21). Then  $\mu_1(u_0) = 0$ , that is  $u_0$  is an asymptotically stable steady state.*

$d$	$\tau$	$N$	$\nu$	$m$	$C_1$	$C_2$	$C_3$	$C$	$\mu_1(u_0)$	Elapsed time (in secs)
3	1	5	2.4	15	4.1174	4.0101	0.013124	0.21669	2	1.27
6	1	5	2.4	15	4.3696	4.2823	0.014524	0.27176	2	1.41
12	1	5	2.4	15	4.8091	4.9555	4.9555	0.4729	2	1.63
24	1	5	2.4	15	5.3788	5.9556	0.029061	0.9309	2	2.67

Table 2: Parameters used in the proofs of the higher dimensional examples.

#### 4.4 Higher dimensional examples

In the previous section, we applied our approach to problems with  $d = 1$  (the Mackey-Glass and Ikeda-Matsumoto equations) and  $d = 2$  (the van der Pol and predator-prey equations). We now show that our approach applies to higher dimensional examples.

Given  $d \geq 3$ , consider the  $d \times d$  matrix

$$K_1 = \begin{pmatrix} \ln 3 & 0 & 0 & \cdots & 0 \\ 0 & \ln 2 & 0 & \cdots & 0 \\ 0 & 0 & \ln \frac{1}{2} & 0 & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \ln \frac{1}{d-1} \end{pmatrix}$$

and  $K_2 \in M_d(\mathbb{R})$  a matrix with random positive entries of size at most  $10^{-2}$ . This choice of  $K_1$  and  $K_2$  does not directly come from a specific DDE, but serves instead as a test case to demonstrate the applicability of our approach to higher dimensional examples. Moreover, choosing  $K_2$  with small entries and  $K_1$  as above implies that the first three leading eigenvalues (i.e. the ones with the largest magnitude) of  $M^N$  are roughly given by 3, 2 and  $1/2$ , which does not depend on  $d$ . We present the results in Table 2, where we consider systems with dimensions 3, 6, 12 and 24.

## 5 Conclusions

We have presented a functional analytic approach for validated computation of eigenvalues for DDEs. In addition we have given a numerical implementation of the scheme which applies to any DDE with the form in Equation (4). The problem comes from an infinite dimensional setting but many questions reduce to finite dimensional equations. Our approach moves fluidly between the finite and infinite dimensional settings, exploiting the best strengths of each.

Validated error bounds for eigenvalues follow from a Newton-Kantorovich theorem applied to the scalar characteristic equation. Yet studying truncations of the infinite dimensional operators provides excellent numerical approximation of the eigenvalues. These approximations can be further refined via Newton's method if necessary. Returning to the infinite dimensional setting provides the proper framework for the eigenvalue exclusion problem, giving the best results for generalized Morse indices.

One natural extension of the present work would be to study repeated eigenvalues and generalized eigenvectors. This could be done by looking for  $\lambda$  having that

$$\frac{d}{d\lambda} \det \left( K_1 + \frac{1}{\lambda} K_2 - \frac{\ln(\lambda)}{\tau} \text{Id} \right) = 0,$$

while simultaneously solving Equation (11) in the case of an eigenvalue with multiplicity two. For higher multiplicities we would append higher derivatives. The generalized eigenfunctions of an eigenvalue with multiplicity  $n$  have the form  $e^{\lambda t}, \dots, t^n e^{\lambda t}$ , each times an appropriate generalized eigenvectors of the matrix

$$M = K_1 + \frac{1}{\lambda} K_2.$$

The methods of the present work extend naturally to this problem. Another more ambitious extension would be to develop analogous methods for systems with multiple constant or distributed delays.

In Section 3 we discussed an approach to computing the  $r$ -generalized Morse index of an equilibrium solution via the argument principle of complex analysis, and outlined an implementation. An interesting project would be to complete this outline and compare the results to those of the present work. We hasten to mention that preliminary results in this direction are not encouraging. We have implemented a validated line integrator in the case of the Mackey-Glass equation and while the program works and produces correct results it is no faster, and in most cases slower than the approach proposed in the present work.

Indeed a naive implementation based on a zero-th order approximation of the integrand yields fat interval enclosures which contain two or more integer values, hence are useless for determining the Morse index. Our second attempt at implementing the line integrals exploited high order Taylor expansions of the integrand and adaptively subdivided the interval of integration. This scheme produces a validated enclosure of only a single integer, but still runs slower than the programs discussed above. We have not implemented the contour integrals for any systems of DDEs, thanks to the difficulties in choosing an appropriate contour already discussed in Section 3. Indeed, our implementation of the functional analytic approach usually runs in several seconds or less, and this alone would seem to discourage the approach based on contour integrals.

Another comment is that, while in the present work we have validated eigenvalue bounds using a Newton-Kantorovich argument, it is not clear that this approach is appropriate for higher dimensional problems. To see this consider once again the equation

$$g(z) = \det \left( K_1 + \frac{1}{z} K_2 - \frac{\ln(z)}{\tau} \text{Id} \right) = 0,$$

when  $K_1, K_2$  are  $d \times d$  matrices. If  $d$  is much greater than 3, expanding the determinant symbolically, symbolically computing  $g'$  and managing the resulting formulas are cumbersome tasks.

An alternative is to compute the determinant using validated numerics. While this can be done it is somewhat delicate and has been avoided in the past by many authors. Moreover all validated Newton schemes require rigorous enclosure of the derivative. Using Jacobi's formula we have that

$$\frac{d}{dz} \det(M(z)) = \text{tr}(\text{adj}(M(z))M'(z)),$$

where

$$M(z) = K_1 + \frac{1}{z} K_2 - \frac{\ln(z)}{\tau} \text{Id},$$

so that

$$M'(z) = -\frac{1}{z^2} K_2 - \frac{1}{\tau z} \text{Id}.$$

If second derivatives are desired these can be worked out as well.

While these high level formulas appear to be straight forward, note that computing the adjugate matrix  $\text{adj}(M(z))$  requires further validated computation of the  $d^2$  determinants of the

$d - 1 \times d - 1$  cofactors of  $M(z)$ . This provides a stand out example of a case where developing efficient validated numerical schemes for a single scalar equation presents challenges.

While the problems just describe can certainly be overcome, there is another option which appears to be preferable when  $d$  is large. Beginning with the equation

$$G_1(\xi, \lambda) = M_2\xi - \lambda M_1\xi = 0$$

appending the scalar phase condition

$$G_2(\xi) = \|\xi\|^2 - 1,$$

isolates a single solution. Defining the operator  $G: \ell_\nu^1 \times \mathbb{C} \rightarrow \ell_\nu^1 \times \mathbb{C}$  by

$$G(\xi, \lambda) = \begin{pmatrix} G_1(\xi, \lambda) \\ G_2(\xi) \end{pmatrix},$$

we have an infinite dimensional zero finding problem whose solutions isolate non-degenerate eigenvalue/eigenvector pair for the DDE.

In practice  $G_2$  is replaced with an equivalent, but differentiable, phase condition. Moreover, if the coefficients of  $\xi$  are complex then  $G_2$  can be modified to isolate complex eigenvectors. Given an approximation eigenpair  $(\xi, \lambda)$  an infinite dimensional analog of Theorem 2.2 leads to validated a-posteriori error bounds. See for example Theorem 2.1 in [4] or Theorem 1 in [13]. Implementing these ideas leads to an efficient computer-assisted proof strategy for the eigenvalue problem which goes around the characteristic equation. Similar problems are solved in other infinite dimensional settings in [10, 14, 24]. See also the works mentioned in the Introduction, and the references discussed therein. Extending the arguments of the works just cited for generalized eigenvalue problems leads to validated solutions of eigenvalue problems for DDEs in any dimension without having to work with the complicated formulas discussed above. This is a straight forward exercise but it appears to be the right way forward for higher dimensional problems. In this case the methods of the present work would still provide the initial data needed for numerical zeros of  $G$ , and would still be needed for eigenvalue exclusion.

This is a theme of the present work: that it can be helpful to look at the eigenvalue problem for DDEs through the lens of numerical linear algebra and scientific computing rather than to study very complicated – albeit scalar – nonlinear equations. Sophisticated tools from validated numerics/computer-assisted proofs in analysis are readily applied to the infinite dimensional problem and the resulting numerical implementation is fast, flexible, and reliable.

## 6 Acknowledgments

The first author was supported by an NSERC Discovery Grant. The second author was partially supported by the National Science Foundation grant DMS – 1813501.

## A Proof of Theorem 2.2

Define  $T: \overline{B_{r_*}(\bar{z})} \rightarrow \mathbb{C}$  by

$$T(z) = z - ag(z),$$

and observe that  $z$  is a zero of  $g$  if and only if  $z$  is a fixed point of  $T$ . Note also that

$$T'(z) = 1 - ag'(z).$$

From  $0 < r_0 \leq r_*$  and  $p(r_0) < 0$  it follows that

$$Zr_0^2 + Y < r_0, \tag{39}$$

and dividing by  $r_0$  gives

$$Zr_0 + \frac{Y}{r_0} < 1.$$

Since  $Y, r_0$  are positive it follows that

$$Zr_0 < 1. \quad (40)$$

For  $z \in \overline{B_{r_0}(\bar{z})}$  we obtain the estimate

$$\begin{aligned} |T'(z)| &= |1 - ag'(z)| \\ &\leq |1 - ag'(\bar{z})| + |a||g'(\bar{z}) - g'(z)| \\ &\leq \left|1 - \frac{1}{g'(\bar{z})}g'(\bar{z})\right| + |a| \sup_{z \in \overline{B_{r_0}(\bar{x})}} |g''(z)||z - \bar{z}| \\ &\leq |1 - 1| + |a| \sup_{z \in \overline{B_{r_*}(\bar{x})}} |g''(z)||z - \bar{z}| \\ &\leq Zr_0. \end{aligned} \quad (41)$$

Now for  $z \in \overline{B_{r_0}(\bar{z})}$ , consider

$$\begin{aligned} |T(z) - \bar{z}| &\leq |T(z) - T(\bar{z})| + |T(\bar{z}) - \bar{z}| \\ &\leq \sup_{w \in \overline{B_{r_0}(\bar{z})}} |T'(w)||z - \bar{z}| + |\bar{z} - ag(\bar{z}) - \bar{z}| \\ &\leq (Zr_0)|z - \bar{z}| + |ag(\bar{z})| \\ &\leq Zr_0^2 + Y \\ &< r_0, \end{aligned}$$

by the inequality of Equation (39), and the bound of Equation (41). Then  $T$  maps  $\overline{B_{r_0}(\bar{z})}$  into  $B_{r_0}(\bar{z}) \subset \overline{B_{r_0}(\bar{z})}$ .

Now for  $z_1, z_2 \in \overline{B_{r_0}(\bar{z})}$  consider

$$\begin{aligned} |T(z_1) - T(z_2)| &\leq \sup_{w \in \overline{B_{r_0}(\bar{z})}} |T'(w)||z_1 - z_2| \\ &\leq Zr_0|z_1 - z_2|. \end{aligned}$$

Recalling the inequality of Equation (40), we have that  $T$  is a contraction on  $\overline{B_{r_0}(\bar{x})}$ , and since  $\overline{B_{r_0}(\bar{x})}$  is a complete metric space, it follows from the contraction mapping theorem that there is a unique  $\tilde{z} \in \overline{B_{r_0}(\bar{z})}$  so that  $T(\tilde{z}) = \tilde{z}$ . In fact, since  $T$  maps  $\overline{B_{r_0}(\bar{x})}$  into  $B_{r_0}(\bar{z})$  we have that  $\tilde{z} \in B_{r_0}(\bar{z})$ . Finally, we observe that

$$ag'(\tilde{z}) = 1 - ((1 - ag'(\bar{z})) + ag'(\bar{z}) - ag'(\tilde{z}))$$

where

$$|(1 - ag'(\bar{z})) + ag'(\bar{z}) - ag'(\tilde{z})| \leq |1 - ag'(\bar{z})| + |a||g'(\bar{z}) - g'(\tilde{z})| \leq Zr_0$$

as above, thanks to  $\tilde{z} \in B_{r_0}(\bar{z})$ . Since  $Zr_0 < 1$ , it follows that

$$|ag'(\tilde{z})| \geq |1 - Zr_0| > 0,$$

hence  $ag'(\tilde{z})$  is bounded away from zero. Since  $a \neq 0$  it follows that  $g'(\tilde{z}) \neq 0$ .

## References

- [1] Gianni Arioli and Hans Koch. Non-symmetric low-index solutions for a symmetric boundary value problem. *J. Differential Equations*, 252(1):448–458, 2012.
- [2] Gianni Arioli and Hans Koch. Existence and stability of traveling pulse solutions of the FitzHugh-Nagumo equation. *Nonlinear Anal.*, 113:51–70, 2015.
- [3] Gianni Arioli and Hans Koch. Spectral stability for the wave equation with periodic forcing. *J. Differential Equations*, 265(6):2470–2501, 2018.
- [4] 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.

- [5] Blake Barker. *Numerical proof of stability of roll waves in the small-amplitude limit for inclined thin film flow*. ProQuest LLC, Ann Arbor, MI, 2014. Thesis (Ph.D.)—Indiana University.
- [6] Blake Barker. Numerical proof of stability of roll waves in the small-amplitude limit for inclined thin film flow. *J. Differential Equations*, 257(8):2950–2983, 2014.
- [7] Blake Barker and Kevin Zumbrun. Numerical proof of stability of viscous shock profiles. *Math. Models Methods Appl. Sci.*, 26(13):2451–2469, 2016.
- [8] Roberto Castelli and Jean-Philippe Lessard. A method to rigorously enclose eigenpairs of complex interval matrices. In *Applications of mathematics 2013*, pages 21–31. Acad. Sci. Czech Repub. Inst. Math., Prague, 2013.
- [9] Roberto Castelli and Jean-Philippe Lessard. Rigorous Numerics in Floquet Theory: Computing Stable and Unstable Bundles of Periodic Orbits. *SIAM J. Appl. Dyn. Syst.*, 12(1):204–245, 2013.
- [10] Roberto Castelli and Holger Teismann. Rigorous numerics for NLS: bound states, spectra, and controllability. Preprint, 2013.
- [11] Jacek Cyranka and Thomas Wanner. Computer-assisted proof of heteroclinic connections in the one-dimensional Ohta-Kawasaki Model. *SIAM J. Appl. Dyn. Syst.*, 17(1):694–731, 2018.
- [12] Teresa Faria. Stability and bifurcation for a delayed predator-prey model and the effect of diffusion. *J. Math. Anal. Appl.*, 254(2):433–463, 2001.
- [13] Zbigniew Galias and Piotr Zgliczyński. Infinite-dimensional Krawczyk operator for finding periodic orbits of discrete dynamical systems. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.*, 17(12):4261–4272, 2007.
- [14] Marcio Gameiro and Jean-Philippe Lessard. A posteriori verification of invariant objects of evolution equations: periodic orbits in the Kuramoto-Sivashinsky PDE. *SIAM J. Appl. Dyn. Syst.*, 16(1):687–728, 2017.
- [15] Gene H. Golub and Henk A. van der Vorst. Eigenvalue computation in the 20th century. *J. Comput. Appl. Math.*, 123(1-2):35–65, 2000. Numerical analysis 2000, Vol. III. Linear algebra.
- [16] Milan Hladík, David Daney, and Elias Tsigaridas. Bounds on real eigenvalues and singular values of interval matrices. *SIAM J. Matrix Anal. Appl.*, 31(4):2116–2129, 2009/10.
- [17] Kensuke Ikeda and Kenji Matsumoto. High-dimensional chaotic behavior in systems with time-delayed feedback. *Physica D: Nonlinear Phenomena*, 29(1-2):223–235, December 1987.
- [18] Jonathan Jaquette, Jean-Philippe Lessard, and Konstantin Mischaikow. Stability and uniqueness of slowly oscillating periodic solutions to Wright’s equation. *J. Differential Equations*, 263(11):7263–7286, 2017.
- [19] Hans Koch. On hyperbolicity in the renormalization of near-critical area-preserving maps. *Discrete Contin. Dyn. Syst.*, 36(12):7029–7056, 2016.
- [20] J.-P. Lessard and J. D. Mireles James. An implicit  $C^1$  time-stepping scheme for delay differential equations. In preparation, 2019.

- [21] Jean-Philippe Lessard and J.D. Mireles James. <http://www.math.mcgill.ca/jplessard/ResearchProjects/spectrumDDE/home.html>. MATLAB codes to perform the proofs, 2019.
- [22] M. C. Mackey and L. Glass. Oscillation and Chaos in Physiological Control Systems. *Science*, 197(4300):287–289, July 1977.
- [23] G. Mayer. Result verification for eigenvectors and eigenvalues. In *Topics in validated computations (Oldenburg, 1993)*, volume 5 of *Stud. Comput. Math.*, pages 209–276. North-Holland, Amsterdam, 1994.
- [24] J. D. Mireles James. Fourier-Taylor approximation of unstable manifolds for compact maps: numerical implementation and computer-assisted error bounds. *Found. Comput. Math.*, 17(6):1467–1523, 2017.
- [25] Kaori Nagatou, Michael Plum, and Mitsuhiro T. Nakao. Eigenvalue excluding for perturbed-periodic one-dimensional Schrödinger operators. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 468(2138):545–562, 2012.
- [26] M. T. Nakao, N. Yamamoto, and K. Nagatou. Numerical verifications for eigenvalues of second-order elliptic operators. *Japan J. Indust. Appl. Math.*, 16(3):307–320, 1999.
- [27] Roger D. Nussbaum. Periodic solutions of some nonlinear autonomous functional differential equations. *Ann. Mat. Pura Appl. (4)*, 101:263–306, 1974.
- [28] Christian Reinhardt and J. D. Mireles James. Fourier-Taylor parameterization of unstable manifolds for parabolic partial differential equations: formalism, implementation and rigorous validation. *Indag. Math. (N.S.)*, 30(1):39–80, 2019.
- [29] Siegfried M. Rump. Computational error bounds for multiple or nearly multiple eigenvalues. *Linear Algebra Appl.*, 324(1-3):209–226, 2001. Special issue on linear algebra in self-validating methods.
- [30] Siegfried M. Rump. Verification methods: rigorous results using floating-point arithmetic. *Acta Numer.*, 19:287–449, 2010.
- [31] Siegfried M. Rump and Jens-Peter M. Zemke. On eigenvector bounds. *BIT*, 43(4):823–837, 2003.
- [32] S.M. Rump. INTLAB - INTerval LABoratory. In Tibor Csendes, editor, *Developments in Reliable Computing*, pages 77–104. Kluwer Academic Publishers, Dordrecht, 1999. <http://www.ti3.tu-harburg.de/rump/>.
- [33] J. C. Sprott. A simple chaotic delay differential equation. *Phys. Lett. A*, 366(4-5):397–402, 2007.
- [34] A. Uçar. A prototype model for chaos studies. *Internat. J. Engrg. Sci.*, 40(3):251–258, 2002.
- [35] Yoshitaka Watanabe, Kaori Nagatou, Michael Plum, and Mitsuhiro T. Nakao. Verified computations of eigenvalue enclosures for eigenvalue problems in Hilbert spaces. *SIAM J. Numer. Anal.*, 52(2):975–992, 2014.
- [36] Yoshitaka Watanabe, Michael Plum, and Mitsuhiro T. Nakao. A computer-assisted instability proof for the Orr-Sommerfeld problem with Poiseuille flow. *ZAMM Z. Angew. Math. Mech.*, 89(1):5–18, 2009.

- [37] Tetsuro Yamamoto. Error bounds for computed eigenvalues and eigenvectors. *Numer. Math.*, 34(2):189–199, 1980.
- [38] Tetsuro Yamamoto. Error bounds for computed eigenvalues and eigenvectors. II. *Numer. Math.*, 40(2):201–206, 1982.