ANALYTIC SOLUTIONS TO NONLINEAR ODES VIA SPECTRAL POWER SERIES

ESTELLE BASOR AND REBECCA MORRISON

Abstract. Solutions to most nonlinear ordinary differential equations (ODEs) rely on numerical solvers, but this gives little insight into the nature of the trajectories and is relatively expense to compute. In this paper, we derive analytic solutions to a class of nonlinear, homogeneous ODEs with linear and quadratic terms on the right-hand side. We formulate a power series expansion of each state variable, whose function depends on the eigenvalues of the linearized system, and solve for the coefficients using some linear algebra and combinatorics. Various experiments exhibit quickly decaying coefficients, such that a good approximation to the true solution consists of just a few terms.

Keywords. Method of matching coefficients; Homogeneous, coupled ODEs; Spectral method; Koopman eigenpairs; Reduced models

1. Introduction

Only a few systems of coupled ordinary differential equations (ODEs) admit known closed-form solutions, in particular, linear systems and the 1-dimensional logistic equation. When a closed-form or analytic solution is not available, mathematicians, scientists, students, and engineers are left instead with numerical solutions: These are relatively slow to compute (usually requiring integration), difficult to analyze (how quickly the solution approaches equilibrium, how good are approximations to the true solution, or how the system might change with different inputs), require the introduction of new parameters (grid size or polynomial degree, for example), and require thorough verification tests.

In this work, we provide closed-form solutions for a class of first-order nonlinear ODEs with linear and quadratic terms on the right-hand side (RHS). Such systems are mainstays in diverse fields including the common SEI or SEIR-type models of epidemiology [4]; chemical reaction mechanisms in combustion and other chemical kinetics [12, 2]; and predator-prey or Lotka-Volterra models in mathematical biology [11]. In all of these examples, the rates of change of different state variables—epidemiological/chemical/biological species—depend on some linear growth rates and pairwise (quadratic) interaction rates.

Our solutions describe the behavior of the system for all $t > t_0$ where t_0 is some fixed value, known exactly in the one-dimensional case, and with a bound provided in the higher-dimensional cases. We assume the eigenvalues of the linearized system at the equilibrium are real and negative and independent over the field of

Date: November 6, 2023.

rationals. The method is based on power series expansions along with some linear algebra and a bit of combinatorics; we refer to it as the Spectral Power Series (SPS) method. Compared to existing work, SPS overlaps with approaches using the eigendecomposition of the Koopman operator [6]. On one hand, SPS is so far limited to linear and quadratic RHS, while Koopman operator theory applies more generally. On the other hand, in contrast with Koopman methods, here we generate explicit functions and coefficients comprising the infinite series. We also note the work of [3], in which the two-dimensional Lotka-Volterra equations are reframed with a new Hamiltonian formalism that does admit an analytic solution but replaces the competition terms with other functions from a specified class.

Our motivation to find such solutions came from the study of reduced models. In applications, certain perturbation methods are useful for finding subsets of solutions for complicated models with incomplete data. Our goal was to understand why sometimes such techniques work and why sometimes they fail, but reliance on numerical solutions stalled much progress. More will be said about this topic at the end of the paper.

As a proof of concept, we demonstrate the method for the logistic equation:

$$\dot{x} = rx\left(1 - \frac{x}{k}\right), \tag{1.1}$$

which has the known solution

$$x(t) = \frac{k}{1 + Ae^{-rt}}$$

where $A = (k - x_0)/x_0$, initial condition $x_0 = x(0)$, and equilibrium $x_{eq} = k$. We will recover this solution with a power series expansion of x. That is, first assume that $x(t) = \sum_n \alpha_n^*(g(t))^n$ for some unknown function g. We substitute this into the differential equation:

$$\sum_n \alpha_n^* n g^{n-1}(t) g'(t) = r \sum_n \alpha_n^* g^n(t) - \frac{r}{k}\left(\sum_n \alpha_n^* g^n(t)\right)\left(\sum_m \alpha_m^* g^m(t)\right).$$

We first notice that matching coefficients for $n = 0$ recovers $\alpha_0 = k$, i.e., the constant term is equal to the equilibrium. For $n = 1$, we find $g'/g = -r$, so $\ln g = -rt + \text{const.}$, and thus $g(t) = ce^{-rt}$. Let $\alpha_n = \alpha_n^* c^n$, so that

$$x(t) = \sum_n \alpha_n e^{-rnt}.$$

Higher-order terms yield $\alpha_n = \alpha_1^n / k^{n-1}, n \geq 2$. That is, α_1 is so far left undetermined, and all higher-order coefficients are written in monomials in terms of α_1. So the solution becomes

$$x(t) = k \sum_{n \geq 0} \left(\frac{\alpha_1}{k}\right)^n e^{-rnt}.$$

Given some fixed α_1 this series converges for large enough t and is

$$x(t) = \frac{k^2}{k - \alpha_1 e^{-rt}}.$$

Suppose that $|\alpha_1/k| < 1$. Then this solution is valid for all $t \geq 0$, and α_1 is given by the initial condition:

$$x_0 = \frac{k^2}{k - \alpha_1}$$

which implies $\alpha_1 = k(\frac{x_0 - k}{x_0})$. Finally, we plug this into the solution to find

$$x(t) = \frac{k}{1 + Ae^{-rt}},$$

where A is the same as defined above.

If we parametize α_1 as $k(\frac{x_0 - k}{x_0})$, note the above condition is satisfied when $x_0 > \frac{1}{2}k$. However, even if $0 < x_0 \leq \frac{1}{2}k$, the series solution extends analytically for all $t > 0$. In general the series converges for

$$t > t_0 = \frac{1}{r}(\log|1 - k/x_0|).$$

It is also worth pointing out that

$$\alpha_1 = \lim_{t \to \infty} (x(t) - k)e^{rt},$$

that is, the solution at infinity retains knowledge about initial conditions.

The logistic solution above can of course be found using simply a separation of variables technique. But the point is it can also be easily found using a series. For the logistic equation, the functions e^{-rnt} and coefficients $\alpha_n = \left(\frac{x_0 - k}{x_0}\right)^n$ above correspond exactly to a set of eigenvalues and eigenfunctions, respectively, of the Koopman operator flow [5]. In particular, the choice to parametrize α_1/k as $\frac{x_0 - k}{x_0}$ yields an Koopman expansion.

This same SPS technique works for systems that are analogous to the logistic equation. We show how this is possible in the next sections and investigate the domains of convergence. For these systems as well, the terms of the series overlap with the Koopman expansion found in [7]. Unlike most of the Koopman literature, we show that the series converges and find estimates for the coefficients and for the domains of convergence. Our methods are linear algebra-based and use a constructive recursive algorithm to find the coefficients. We will say more later in the paper about connections to the Koopman expansions.

2. 2×2 SYSTEMS

In this section, we address the 2×2 case in detail. It is a bit easier to track the steps for this case and then in turns out that the higher dimensional case only needs minor modifications in the argument.

We consider 2-dimensional systems of the form

$$\dot{x_1} = b_1 x_1 + a_{11} x_1^2 + a_{12} x_1 x_2$$
$$\dot{x_2} = b_2 x_2 + a_{21} x_1 x_2 + a_{22} x_2^2.$$

Let

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} \\ \\ a_{21} & a_{22} \end{pmatrix}.$$

The equilibrium of the system is found by computing $\boldsymbol{c} = \boldsymbol{A}^{-1}(-\boldsymbol{b})$ where

$$\boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \text{ and } \boldsymbol{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

The eigenvalues of the linearized system near equilibrium are given by the eigenvalues of the matrix

$$\operatorname{diag}(\boldsymbol{c})\boldsymbol{A} = \begin{pmatrix} c_1a_{11} & c_1a_{12} \\ \\ c_2a_{21} & c_2a_{22} \end{pmatrix}.$$

To see this, let $\boldsymbol{x} = \boldsymbol{c} + \boldsymbol{y}$ around equilibrium, so that $\dot{\boldsymbol{x}} = \dot{\boldsymbol{y}}$ and

$$\begin{aligned} \dot{\boldsymbol{y}} &= \operatorname{diag}(\boldsymbol{y} + \boldsymbol{c})\left(\boldsymbol{b} + \boldsymbol{A}(\boldsymbol{c} + \boldsymbol{y})\right) \\ &= \operatorname{diag}(\boldsymbol{c} + \boldsymbol{y})\left(\boldsymbol{A}\boldsymbol{y}\right) \end{aligned}$$

and the linear part is $\dot{\boldsymbol{y}} \approx \operatorname{diag}(\boldsymbol{c})\boldsymbol{A}\boldsymbol{y}$. The eigenvalues, which we will denote by λ_1 and λ_2 are given by

$$\frac{(c_1a_{11} + c_2a_{22}) \pm \sqrt{(c_1a_{11} - c_2a_{22})^2 + 4c_1c_2a_{12}a_{21}}}{2}.$$

We now assume that both eigenvalues are real, negative, and that they are non-rational, and that our solutions near equilibrium are of the form

$$\begin{aligned} x_1(t) &= \sum_{n_1,n_2 \geq 0} \alpha_1^{n_1,n_2} e^{(n_1\lambda_1 + n_2\lambda_2)t} \\ x_2(t) &= \sum_{n_1,n_2 \geq 0} \alpha_2^{n_1,n_2} e^{(n_1\lambda_1 + n_2\lambda_2)t}. \end{aligned}$$

(From here on, we use subscripts for variable or eigenvalue indices, and superscripts for the expansion indices.) Substituting these solutions into the differential equations and equating like terms, we find that as expected $\alpha_1^{0,0} = c_1$, $\alpha_2^{00} = c_2$. For the cases when $n_1 = 1, n_2 = 0$, and $n_1 = 0, n_2 = 1$, the equations leave $\alpha_1^{1,0}$ and $\alpha_1^{0,1}$ undetermined, $\alpha_2^{1,0}$ and $\alpha_2^{0,1}$ multiples of the undetermined coefficients, but yield values for λ_1 and λ_2 that agree with the ones found above. (To see this, match coefficients for $e^{\lambda_1 t}$ or $e^{\lambda_2 t}$ terms. Both result in a quadratic equation for $\lambda_{1,2}$.)

Returning to the coefficients,

$$\alpha_2^{1,0} = \alpha_1^{1,0} a_{21}c_2/(\lambda_1 - a_{22}c_2)$$

and

$$\alpha_2^{0,1} = \alpha_1^{0,1} a_{12}c_2/(\lambda_2 - a_{11}c_1).$$

For all other values we see that $\alpha_1^{n_1,n_2}$ and $\alpha_2^{n_1,n_2}$ must satisfy

$$\begin{aligned} (n_1\lambda_1 \ + n_2\lambda_2 - a_{11}c_1)\alpha_1^{n_1,n_2} - a_{12}c_1\alpha_2^{n_1,n_2} &= a_{11}S_{1,1}^{n_1,n_2} + a_{12}S_{1,2}^{n_1,n_2} \\ -a_{21}c_2\alpha_1^{n_1,n_2} + (n_1\lambda_1 \ + n_2\lambda_2 - a_{22}c_2)\alpha_2^{n_1,n_2} &= a_{21}S_{1,2}^{n_1,n_2} + a_{22}S_{2,2}^{n_1,n_2} \end{aligned} \tag{2.1}$$

where

$$S_{1,1}^{n_1,n_2} = \sum_{(i,j)\neq(0,0),(n_1,n_2)} \alpha_1^{i,j}\alpha_1^{n_1-i,n_2-j}$$
$$S_{1,2}^{n_1,n_2} = \sum_{(i,j)\neq(0,0),(n_1,n_2)} \alpha_1^{i,j}\alpha_2^{n_1-i,n_2-j}$$

and

$$S_{2,2}^{n_1,n_2} = \sum_{(i,j)\neq(0,0),(n_1,n_2)} \alpha_2^{i,j}\alpha_2^{n_1-i,n_2-j}.$$

Now let $\boldsymbol{n} = \langle n_1, n_2\rangle, \boldsymbol{\lambda} = \langle \lambda_1, \lambda_2\rangle$, $\boldsymbol{\alpha^n} = \langle \alpha_1^{n_1,n_2}, \alpha_2^{n_1,n_2}\rangle$,

$$\boldsymbol{S^n} = \begin{pmatrix} S_{1,1}^{n_1,n_2} & S_{1,2}^{n_1,n_2} \\ \\ S_{1,2}^{n_1,n_2} & S_{2,2}^{n_1,n_2} \end{pmatrix},$$

and

$$\boldsymbol{s^n} = \langle \boldsymbol{e}_1 \cdot \boldsymbol{AS_n e}_1, \boldsymbol{e}_2 \cdot \boldsymbol{AS_n e}_2\rangle$$

with $\boldsymbol{e}_1, \boldsymbol{e}_2$ the standard unit basis vectors. In other words, the vector $\boldsymbol{s^n}$ is the diagonal of $\boldsymbol{AS^n}$.

Our system (2.1) is equivalent to

$$(\boldsymbol{n}\cdot\boldsymbol{\lambda})\left(\boldsymbol{I} - \frac{\operatorname{diag}(\boldsymbol{c})\boldsymbol{A}}{\boldsymbol{n}\cdot\boldsymbol{\lambda}}\right)\boldsymbol{\alpha^n} = \boldsymbol{s^n}$$

or

$$\boldsymbol{\alpha^n} = \frac{1}{\boldsymbol{n}\cdot\boldsymbol{\lambda}}\left(\boldsymbol{I} - \frac{\operatorname{diag}(\boldsymbol{c})\boldsymbol{A}}{\boldsymbol{n}\cdot\boldsymbol{\lambda}}\right)^{-1}\boldsymbol{s^n}.$$

Note that the matrix $\left(I - \frac{\operatorname{diag}(\boldsymbol{c})A}{\boldsymbol{n}\cdot\boldsymbol{\lambda}}\right)$ is always invertible if $\boldsymbol{n}$ is not $(1,0)$ or $(0,1)$.

Our goal now is to estimate the terms in the last equation so that we can get a good estimate of $\boldsymbol{\alpha^n}$.

We already assumed that both λ_1 and λ_2 are real, negative, and non-rational. We now also assume that $|\lambda_1| < |\lambda_2|$ without any loss of generality.

Lemma 2.1. *Let $N = n_1 + n_2$. Then*

$$|\boldsymbol{n}\cdot\boldsymbol{\lambda}| \geq |\lambda_1|N.$$

Proof. This follows easily since all the eigenvalues are negative. □

Lemma 2.2. *Let $N = n_1 + n_2$. Suppose not both i and j are zero and suppose for all other i and j with $i + j < N$ there exists a constant K such that*

$$|\alpha_1^{i,j}| \leq \frac{K^{i+j}}{(i+1)(j+1)}, \quad |\alpha_2^{i,j}| \leq \frac{K^{i+j}}{(i+1)(j+1)}.$$

Then

$$|S_{1,1}^{n_1,n_2}| \leq \frac{4K^{n_1+n_2}((\log(n_1+1))+1)((\log(n_2+1)+1)}{(n_1+1)(n_2+1)}.$$

Proof. We have that

$$S_{1,1}^{n_1,n_2} = \sum_{(i,j)\neq(0,0),(n_1,n_2)} \alpha_1^{i,j}\alpha_1^{n_1-i,n_2-j}$$

is bounded by

$$\sum_{(i,j)\neq(0,0),(n_1,n_2)} \frac{K^{i+j}}{(i+1)(j+1)}\frac{K^{n_1-i+n_2-j}}{(n_1-i+1)(n_2-j+1)}$$

or

$$K^{n_1+n_2}\sum_{i=0}^{n_1}\frac{1}{(i+1)(n_1-i+1)}\sum_{j=0}^{n_2}\frac{1}{(j+1)(n_2-j+1)}$$

Separating the denominators into partial fractions and using an estimate for $\log n$ then yields the result. □

Lemma 2.3. *Assuming the conditions in the previous lemma, the same estimates hold for and $S_{1,2}^{n_1,n_2}$ and $S_{2,2}^{n_1,n_2}$.*

Lemma 2.4. *Suppose the coefficients of the two power series*

$$\sum_{n_1,n_2\geq 0}\alpha_1^{n_1,n_2}e^{(n_1\lambda_1+n_2\lambda_2)t}$$

$$\sum_{n_1,n_2\geq 0}\alpha_2^{n_1,n_2}e^{(n_1\lambda_1+n_2\lambda_2)t}$$

satisfy the conditions

$$|\alpha_1^{n_1,n_2}| < \frac{C^{n_1+n_2}}{(n_1+1)(n_2+1)}, \quad |\alpha_2^{n_1,n_2}| < \frac{C^{n_1+n_2}}{(n_1+1)(n_2+1)}$$

for some positive constant C. Then there exists a t_0 such that for all $t > t_0$ the series converge.

Note that once the above is verified, then term by term differentiation of the series yield series that converge uniformly on the same interval and thus the series are the actual solutions.

Proof. The proof is the same for both series, so we consider

$$\sum_{n_1,n_2\geq 0}\alpha_1^{n_1,n_2}e^{(n_1\lambda_1+n_2\lambda_2)t}.$$

This sum is at most

$$\sum_{n_1,n_2\geq 0}\frac{C^{n_1+n_2}}{(n_1+1)(n_2+1)}e^{(n_1\lambda_1+n_2\lambda_2)t}$$

which is bounded by

$$\sum_{n_1\geq 0}C^{n_1}e^{(n_1\lambda_1)t}\sum_{n_2\geq 0}C^{n_2}e^{(n_2\lambda_1)t}$$

Thus if t_0 is chosen so that $|Ce^{\lambda_1 t_0}| < 1$ and $|Ce^{\lambda_2 t_0}| < 1$ we have convergence. □

Our last step is to show how one can guarantee the estimates in the last lemma.

Theorem 2.5. *There exists a constant K such that the coefficients $\alpha_1^{n_1,n_2}$ and $\alpha_2^{n_1,n_2}$ satisfy*

$$|\alpha_1^{n_1,n_2}| < \frac{K^{n_1+n_2}}{(n_1+1)(n_2+1)}, \quad |\alpha_2^{n_1,n_2}| < \frac{K^{n_1+n_2}}{(n_1+1)(n_2+1)}.$$

Proof. Recall that

$$\boldsymbol{\alpha^n} = \frac{1}{\boldsymbol{n}\cdot\boldsymbol{\lambda}}\left(\boldsymbol{I} - \frac{\operatorname{diag}(\boldsymbol{c})\boldsymbol{A}}{\boldsymbol{n}\cdot\boldsymbol{\lambda}}\right)^{-1}\boldsymbol{s^n}$$

We first note that a matrix of the form $\boldsymbol{I}+\boldsymbol{B}$ where $\boldsymbol{B}$ has operator norm $||\boldsymbol{B}|| \leq \delta < 1$ has an inverse of the form

$$\boldsymbol{I}+\boldsymbol{C}$$

where the norm of $\boldsymbol{C}$ is at most $\frac{\delta}{1-\delta}$.

Now choose N_0 such that for $N \geq N_0$ the operator norm

$$\left\|\frac{\operatorname{diag}(\boldsymbol{c})\boldsymbol{A}}{\boldsymbol{n}\cdot\boldsymbol{\lambda}}\right\| \;\leq\; \left\|\frac{\operatorname{diag}(\boldsymbol{c})\boldsymbol{A}}{N\lambda_1}\right\|$$

is at most $1/2$. Thus the operator norm of

$$\left(\boldsymbol{I} - \frac{\operatorname{diag}(\boldsymbol{c})\boldsymbol{A}}{\boldsymbol{n}\cdot\boldsymbol{\lambda}}\right)^{-1}$$

will be at most 2.

Also, there exists an N_1 such that for $N > N_1$,

$$\frac{8(\log(N+1)+1)^2\|\boldsymbol{A}\|}{N|\lambda_1|} < 1/2.$$

Now let $N_2 = \max\{N_0, N_1\}$.

For pair i,j such that $i+j \leq N_2$ we compute the quantity $((i+1)(j+1)|a_{i,j}|)^{\frac{1}{i+j}}$. There are a finite number of these and we define K_1 to be the maximum value of these. Thus for a finite number of such coefficients, we have that

$$|\alpha_1^{i,j}| < \frac{K_1^{i+j}}{(i+1)(j+1)}.$$

We repeat the process for the $\alpha_2^{i,j}$ and then choose the largest: $K = \max\{K_1, K_2\}$. We are finally in a position to use an induction argument.

We let $n_1+n_2 = N_2+1$. For all i,j with $i+j \leq N_2$ we have the satisfied the condition

$$|\alpha_1^{i,j}| < \frac{K^{i+j}}{(i+1)(j+1)}, \quad |\alpha_2^{i,j}| < \frac{K^{i+j}}{(i+1)(j+1)}.$$

Thus by Lemma 2.2, for any n_1+n_2 with $n_1+n_2 = N_2+1$, we have that $S_{i,j}^{n_1,n_2}$ $(i,j=1,2)$ is bounded by

$$\frac{4K^{n_1+n_2}(\log(n_1+1)+1)(\log(n_2+1)+1)}{(n_1+1)(n_2+1)}.$$

This means the operator norm of the matrix $\boldsymbol{S^n}$ is bounded by

$$\frac{8K^{n_1+n_2}((\log(N_2+2))+1)^2}{(n_1+1)(n_2+1)}.$$

Note that the Hilbert-Schmidt norm of $\boldsymbol{S^n}$, $\|\boldsymbol{S^n}\|_2$, is bounded by the same value and thus

$$\|\boldsymbol{s^n}\| \leq \|\boldsymbol{A}\boldsymbol{S^n}\|_2 \leq \|\boldsymbol{A}\|\|\boldsymbol{S^n}\|_2.$$

Putting everything together we have for $\boldsymbol{n} = \langle n_1, n_2\rangle$, and $n_1 + n_2 = N_2 + 1$,

$$\boldsymbol{\alpha^n} = \frac{1}{\boldsymbol{n}\cdot\boldsymbol{\lambda}}\left(I - \frac{\operatorname{diag}(\boldsymbol{c})\boldsymbol{A}}{\boldsymbol{n}\cdot\boldsymbol{\lambda}}\right)^{-1}\boldsymbol{s^n}$$

so that

$$\begin{aligned}\|\boldsymbol{\alpha^n}\| &\leq \frac{1}{(N_2+1)|\lambda_1|)}\left\|\left(I - \frac{\operatorname{diag}(\boldsymbol{c})\boldsymbol{A}}{\boldsymbol{n}\cdot\boldsymbol{\lambda}}\right)^{-1}\right\|\|\boldsymbol{A}\|\|\boldsymbol{S^n}\|_2 \\ &\leq \frac{2\|\boldsymbol{A}\|}{(N_2+1)|\lambda_1|)}\frac{8K^{n_1+n_2}((\log(N_2+2))+1)^2}{(n_1+1)(n_2+1)} \\ &\leq \frac{K^{n_1+n_2}}{(n_1+1)(n_2+1)}\frac{2\|\boldsymbol{A}\|8(\log(N_2+2))+1)^2}{(N_2+1)|\lambda_1|)}\end{aligned}$$

which implies

$$\|\boldsymbol{\alpha^n}\| \leq \frac{K^{n_1+n_2}}{(n_1+1)(n_2+1)}$$

or the desired result. □

One should note that the conditions that we imposed on the coefficients of our series were designed to make the analysis possible. There may be other approaches with less restrictive bounds that show the series converge.

The other thing to note is that the choice of $1/2$ to find the N_0 and N_1 in the above argument can be replaced by $\delta < 1$ for N_0 and $1-\delta$ for N_1 if it makes the choice of N_2 smaller.

As we will see with some of the examples, the coefficients with choices of reasonable constants seem to decay quickly. This is expected because of the following result.

Theorem 2.6. *Given some choice of $\alpha_1^{1,0}$ and $\alpha_1^{0,1}$ all other $\alpha_i^{n_1,n_2}, i = 1, 2$ have the form*

$$c_{\boldsymbol{n}}(\alpha_1^{1,0})^{n_1}(\alpha_1^{0,1})^{n_2}$$

where the constant $c_{\boldsymbol{n}}$ only depends on all the other data in the system.

Proof. This is straightforward to see from the formula for the the coefficients

$$\boldsymbol{\alpha^n} = \frac{1}{\boldsymbol{n}\cdot\boldsymbol{\lambda}}\left(\boldsymbol{I} - \frac{\operatorname{diag}(\boldsymbol{c})\boldsymbol{A}}{\boldsymbol{n}\cdot\boldsymbol{\lambda}}\right)^{-1}\boldsymbol{s^n}$$

and the form of the terms for the $S_{i,j}^{\boldsymbol{n}}$. □

We end this section with an example that estimates the value of K and hence the value of t_0 in our main theorem. Other examples will be given in Section 4.

For this example we let

$$\boldsymbol{A} = \begin{pmatrix} -1 & -0.3 \\ -0.1 & -1 \end{pmatrix}, \ \boldsymbol{b} = \begin{pmatrix} 2.45 \\ 1.7 \end{pmatrix}, \ \boldsymbol{c} = \begin{pmatrix} 2 \\ 1.5 \end{pmatrix},$$

with initial condition $(3, 3)$. The eigenvalues of $\mathrm{diag}(\boldsymbol{c})\boldsymbol{A}$ are ~ -1.359 and ~ -2.140. Using an estimate for the operator norm of $\boldsymbol{A}$ and some help from Mathematica we found that N_2 is of the order of ~ 320. This yielded a K value of less than 5.83 and we find that we have convergence for $t > t_0 = 1.3$.

In addition, for this example, our eigenvalues are not too far apart in value, so we can find appropriate $\alpha_1^{1,0}$ and $\alpha_1^{0,1}$ by using limits at infinity. In fact

$$\alpha_1^{1,0} = \lim_{t\to\infty} (x_1(t) - 2)e^{-\lambda_1 t}$$

and

$$\alpha_1^{0,1} = \lim_{t\to\infty} (x_1(t) - 2 - \alpha_1^{1,0} e^{\lambda_1 t})e^{-\lambda_2 t}.$$

For our example, $\alpha_1^{1,0} \sim -0.45$ and $\alpha_1^{0,1} \sim 0.91$. Just as in the one-dimensional case these values retain information about initial conditions and can be found with enough data. The alternative is to use directly the initial conditions.

With these values, we plot our analytic SPS answer for $x_1(t)$ and $x_2(t)$ verses the image using a numerical solver.

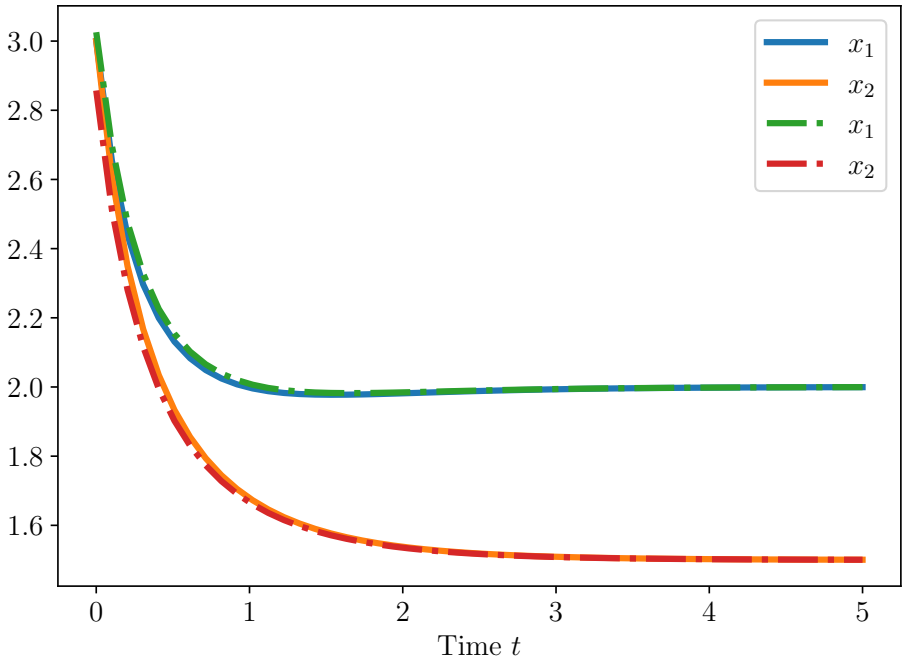


FIGURE 1. SPS (dashed line) and numerical (solid line) solutions, $\boldsymbol{x}(0) = (1, 1)$. Convergence is guaranteed for $t > t_0 = 1.3$.

3. $M \times M$ SYSTEMS

Now consider the $M \times M$ system:

$$\dot{\boldsymbol{x}} = \text{diag}(\boldsymbol{x})(\boldsymbol{b} + \boldsymbol{A}\boldsymbol{x}).$$

Note that existence and uniqueness of the solution is guaranteed, for any $\boldsymbol{x}(0) \in \mathbb{R}^M$, since the RHS is continuous and all its partial derivatives are continuous in $\mathbb{R}^M$ [11]. In what follows we assume that the matrix $\boldsymbol{A}$ is invertible. Equilibrium is given by $\boldsymbol{c} = -\boldsymbol{A}^{-1}\boldsymbol{b}$, and again, the eigenvalues of the linearized system around equilibrium are those of the matrix $\text{diag}(\boldsymbol{c})\boldsymbol{A}$.

Let $\boldsymbol{\lambda}$ be a vector whose entries are the eigenvalues of $\text{diag}(\boldsymbol{c})\boldsymbol{A}$. The order of the eigenvalues does not matter, but we can assume they are in ascending order of magnitude. We also assume that the $M - 1$ dimensional hyperplane defined by $\boldsymbol{z} \cdot \boldsymbol{\lambda} = 0$ has no solutions with $\boldsymbol{z}$ having integer entries. (In other words, the hyperplane does not intersect the integer lattice.) With this assumption we let the solution for each x_i be of the form

$$x_i(t) = \sum_{\boldsymbol{n} \geq 0} \alpha_i^{\boldsymbol{n}} e^{(\boldsymbol{n} \cdot \boldsymbol{\lambda})t}.$$

Setting $\boldsymbol{n} = 0$ yields the system $\boldsymbol{A}\boldsymbol{\alpha}^{\boldsymbol{0}} = -\boldsymbol{b}$, or, $\boldsymbol{\alpha}^{\boldsymbol{0}} = \boldsymbol{A}^{-1}\boldsymbol{b}$, i.e., we recover that the constant terms are equal to equilibrium.

Now let us consider the case when $\boldsymbol{n} = \boldsymbol{e}_i$ where $\boldsymbol{e}_i$ is a standard basis vector with 1 in the ith entry and 0s elsewhere. Then $\boldsymbol{n}\cdot\boldsymbol{\lambda}$ is the eigenvalue λ_i. By again equating coefficients, for these $\boldsymbol{n}$ our system reduces to

$$(\lambda_i \boldsymbol{I} - \operatorname{diag}(\boldsymbol{c})\boldsymbol{A})\,\boldsymbol{\alpha}^{\boldsymbol{n}} = 0,$$

The eigenvalues must be distinct, and so it follows that the matrix

$$\lambda_i \boldsymbol{I} - \operatorname{diag}(\boldsymbol{c})\boldsymbol{A}$$

has rank $M-1$. Thus the vector $\boldsymbol{\alpha}^{\boldsymbol{n}}$ is in the kernel of this matrix; the kernel is one dimensional. One of the coefficients must be nonzero and arbitrary and all others are determined once this value is chosen. Note also that if our initial conditions do not take on equilibrium values and in particular x_1, then in fact $\alpha_1^{\boldsymbol{e}_i}$ must not be zero.

Thus we have shown that for the standard basis vectors, we have M arbitrary values (one for each i). Fixing these values corresponds to choosing initial conditions for the systems.

For other values of $\boldsymbol{n}$, we find the following system holds:

$$(-\operatorname{diag}(\boldsymbol{c})\boldsymbol{A} + (\boldsymbol{n}\cdot\boldsymbol{\lambda})\boldsymbol{I})\,\boldsymbol{\alpha}^{\boldsymbol{n}} = \boldsymbol{s}^{\boldsymbol{n}},$$

where $(\boldsymbol{s}^n)_i = \sum_j a_{i,j} S_{i,j}^{\boldsymbol{n}}$ and

$$S_{i,j}^{\boldsymbol{n}} = \sum_{(i_1,\ldots,i_M)\neq 0,\mathbf{n}} \alpha_i^{i_1,i_2,\ldots,i_M}\alpha_j^{n_1-i_1,n_2-i_2,\ldots,n_M-i_M}.$$

Let's rewrite the system as

$$(\boldsymbol{n}\cdot\boldsymbol{\lambda})\left(I - \frac{\operatorname{diag}(\boldsymbol{c})\boldsymbol{A}}{\boldsymbol{n}\cdot\boldsymbol{\lambda}}\right)\boldsymbol{\alpha}^{\boldsymbol{n}} = \boldsymbol{s}^{\boldsymbol{n}}.$$

Define as before

$$\boldsymbol{S}^{\boldsymbol{n}} = (S_{i,j}^{\boldsymbol{n}}) \quad 1 \le i,j \le M.$$

The matrix $\boldsymbol{S}^{\boldsymbol{n}}$ is symmetric and it follows as in section 2 that $\boldsymbol{s}^{\boldsymbol{n}}$ equals the diagonal entries of $\boldsymbol{A}\boldsymbol{S}^{\boldsymbol{n}}$. Note that the matrix $\boldsymbol{n}\cdot\boldsymbol{\lambda}I - \operatorname{diag}(\boldsymbol{c})\boldsymbol{A}$ is invertible except in the case when $\boldsymbol{n}$ is one of the standard basis vectors. Thus with this set up—except for adjusting for some constants—everything done in the 2×2 case goes through. Summarizing we have

Theorem 3.1. *Suppose that the matrix $\boldsymbol{A}$ is invertible. Let $\boldsymbol{b}$ be some given vector and $\boldsymbol{c} = -\boldsymbol{A}^{-1}\boldsymbol{b}$. Let $\boldsymbol{\lambda}$ be a vector constructed from the eigenvalues of* $\operatorname{diag}(\boldsymbol{c})\boldsymbol{A}$. *We assume the eigenvalues are real and negative and that there are no vector integer solutions of $\boldsymbol{z}\cdot\boldsymbol{\lambda} = 0$ Then there exists a t_0 such that the system*

$$\dot{\boldsymbol{x}} = \operatorname{diag}(\boldsymbol{x})(\boldsymbol{b} + A\boldsymbol{x}).$$

has a solution of the from

$$\boldsymbol{x} = \sum_{\boldsymbol{n}} \boldsymbol{\alpha}^n e^{\boldsymbol{n}\cdot\boldsymbol{\lambda}t}$$

that converges for $t \ge t_0 > 0$.

Proof. The definitions, lemmas and conclusions from the 2×2 case extend to the higher dimensional case without almost any change except in two places. The careful reader will notice that in Lemma 2.2 the conclusion

$$|S_{1,1}^{n_1,n_2}| \leq \frac{4K^{n_1+n_2}((\log(n_1+1)+1)((\log(n_2+1)+1)}{(n_1+1)(n_2+1)}$$

is replaced by

$$|S_{1,1}^{\boldsymbol{n}}| \leq \frac{2^M K^N((\log(n_1+1)+1)((\log(n_2+1)+1)\cdots(\log(n_M+1)+1)}{(n_1+1)(n_2+1)\cdots(n_M+1)},$$

where $N = n_1 + n_2 + \cdots + n_M$.

The other adjustment is that the Hilbert-Schmidt norm of the matrix $\boldsymbol{S^n}$ is then bounded by

$$\frac{2^M M K^N((\log(N_2+2))+1)^M}{(n_1+1)(n_2+1)\cdots(n_M+1)}.$$

□

4. Examples

Here are some examples that illustrate the nature and convergence of the SPS solutions.

4.1. 2×2 system.

Let

$$\boldsymbol{A} = \begin{pmatrix} -2 & -1 \\ -1 & -1 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}.$$

Then $\boldsymbol{c} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. The eigenvalues are given by

$$\lambda_1 = -2 + \sqrt{2} \text{ and } \lambda_2 = -2 - \sqrt{2}.$$

We use the solution from Sec. 2, and let $n_1 \leq 3$ and $n_2 \leq 3$ so that $N = 6$. The coefficients $\boldsymbol{\alpha}$ in terms of $p = \alpha_1^{1,0}$ and $q = \alpha_1^{0,1}$ are given in the matrix below (to the hundredths):

$$\begin{pmatrix} 1 & p & -0.08p^2 & -0.64p^3 \\ q & 2pq & 1.28p^2q & -0.56p^3q \\ 0.93q^2 & 2.82pq^2 & 3.51p^2q^2 & 1.33p^3q^2 \\ 0.85q^3 & 3.46pq^3 & 6.27p^2q^3 & 5.58p^3q^3 \end{pmatrix},$$

where $\alpha_1^{i,j}$ is the ij-th entry, indexed from 0.

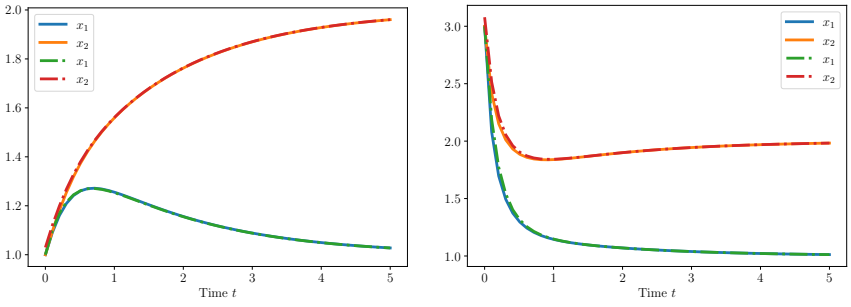
At this point, $\alpha_1^{1,0}$ and $\alpha_1^{0,1}$ are still undetermined: we must incorporate the initial condition. Supposing the series converge at $t = 0$, the sum of all coefficients (for

one variable) must equal its initial value:

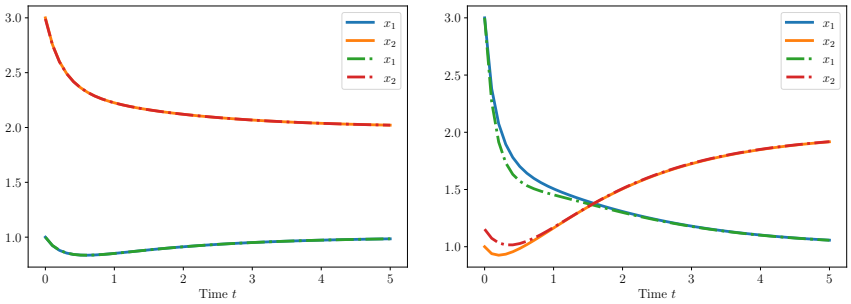
$$\sum_{n_1,n_2} \alpha_1^{n_1,n_2} = x_1(0)$$
$$\sum_{n_1,n_2} \alpha_2^{n_1,n_2} = x_2(0).$$

(Note that if we do not anticipate convergence at $t = 0$, we can use initial values at some other fixed point.)

Fig. 2 shows the solution trajectories for four different initial conditions.



(A) $\boldsymbol{x}(0) = (1, 1)$ (B) $\boldsymbol{x}(0) = (3, 3)$



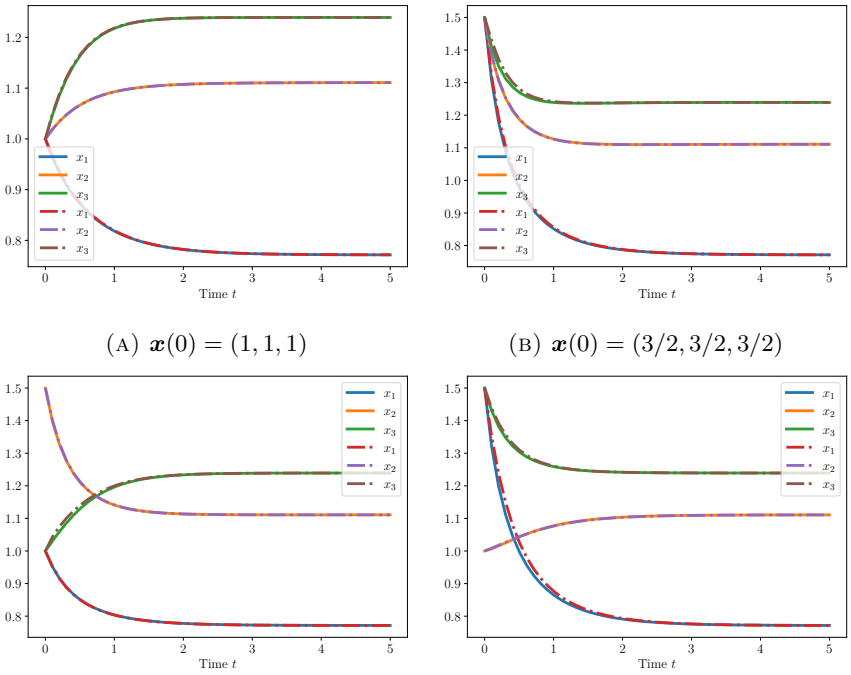
(C) $\boldsymbol{x}(0) = (1, 3)$ (D) $\boldsymbol{x}(0) = (3, 1)$

FIGURE 2. Analytic SPS (dotted line) and numerical (solid line) solution trajectories for the 2×2 example, for four different initial conditions. For (A-C), analytic solutions closely match the numerical solutions for $t > 0$; for (D), the analytic solution converges to the true solution around $t = 2$.

4.2. 3×3 **system.** Let's look at another example, now a 3×3 system. Let the interaction matrix and linear growth rate vector be

$$\boldsymbol{A} = \begin{pmatrix} -2 & -0.3 & -0.1 \\ -0.2 & -2 & -0.1 \\ -0.1 & -0.4 & -2 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 2 \\ 2.5 \\ 3 \end{pmatrix}.$$

Again we let $n_1, n_2, n_3 \leq 3$ so that $N = 9$. Fig. 3 shows numerical and analytic solution trajectories for the three variables, for four different initial conditions.



(A) $\boldsymbol{x}(0) = (1, 1, 1)$ (B) $\boldsymbol{x}(0) = (3/2, 3/2, 3/2)$

(C) $\boldsymbol{x}(0) = (1, 3/2, 1)$ (D) $\boldsymbol{x}(0) = (3/2, 1, 3/2)$

FIGURE 3. Analytic SPS (dotted line) and numerical (solid line) solution trajectories for the 3×3 example, for four different initial conditions. Analytic solutions closely match the numerical solutions for $t > 0$.

5. Remarks about Koopman expansions

In [7] (see displayed equation 69), a general framework is given for the systems we have considered and a Koopman expansion of the form

$$\boldsymbol{x} = \sum_{\boldsymbol{n}} \boldsymbol{\alpha}^{\boldsymbol{n}} e^{\boldsymbol{n}\cdot\boldsymbol{\lambda} t}$$

is produced. Can we see that they are the same?

Let us consider the 2×2 system. Since the SPS coefficients are all generated by products of α_1^{10} and α_1^{01} and these values can be found by either a limiting process or an algebraic method, if the series that appears in the above reference is known to converge then the coefficients must be the same.

But we can say more. Let us make the assumption that α_1^{10} and α_1^{01} are eigenfunctions of the Lie derivative operator for the Koopman flow with eigenvalues λ_1 and λ_2. Recall our system is of the form:

$$\dot{x_1} = f(x_1, x_2) = b_1 x_1 + a_{11} x_1^2 + a_{12} x_1 x_2$$
$$\dot{x_2} = g(x_1, x_2) = b_2 x_2 + a_{21} x_1 x_2 + a_{22} x_2^2.$$

For α_1^{10} and α_1^{01} to be eigenfunctions it follows that

$$\lambda_1 \alpha_1^{10} = f(x_1, x_2)\frac{\partial\, \alpha_1^{10}}{\partial x_1} + g(x_1, x_2)\frac{\partial\, \alpha_1^{10}}{\partial x_2}$$

and

$$\lambda_2 \alpha_1^{01} = f(x_1, x_2)\frac{\partial\, \alpha_1^{01}}{\partial x_1} + g(x_1, x_2)\frac{\partial\, \alpha_1^{01}}{\partial x_2}.$$

Then it is the case that $\alpha_i^{\boldsymbol{n}}$ does align as well with eigenvalue $\boldsymbol{n} \cdot \boldsymbol{\lambda}$. This is easy to see by recalling from Theorem 2.6 that all $\alpha_i^{n_1,n_2}, i = 1, 2$ have the form

$$c_{\boldsymbol{n}}(\alpha_1^{1,0})^{n_1}(\alpha_1^{0,1})^{n_2}$$

where the constant $c_{\boldsymbol{n}}$ only depends on all the other data in the system.

This is analogous to the one-dimensional case where we chose $\frac{x_0-k}{x_0}$ to be our parametrization. The higher-dimensional cases can be analyzed in a similar fashion.

6. Reduced models

In previous work [8, 9], we have seen that reduced models, that only track a subset of $L < M$ variables, can still faithfully represent the dynamics of the complete $M \times M$ system. We lacked a theoretical justification for why this should be the case, but now the analysis here can help explain this phenomenon.

In practice, this type of reduction is very common. In hydrocarbon combustion, a state-of-the-art model may include fifty (or more) chemical species, yielding a set of fifty ODEs [10]. But these models are too expensive for many industrial and academic pursuits; instead reduced mechanisms include only a handful of chemical species, such as just the reactants and products. As another example, consider SEIR-type models of epidemiological outbreaks (which track susceptible, exposed, infected, and recovered populations, or other groups such as hospitalized or quarantined). But of course many other variables (epidemiological species) may contribute to the outbreak dynamics. Typical models of a Zika virus outbreak track sub-populations of humans and mosquitos [1], but omit livestock, non-human primates, and so on. Modelers may not have sufficient data to include these variables, or simply not know which are the most relevant.

Despite such omission of potentially many state variables, still reduced models may capture observed dynamics and even perform well when extrapolated to other scenarios (such as in the future, or for other environmental conditions) [9]. To examine this type of reduced model over $L < M$ state variables, let us first take the *partial* model, using $A_{L\times L}$, the upper left $L \times L$ submatrix of A and $b_{1:L}$, the first L components of b.

Following [8, 9] and as a correction to the partial model, two terms are added for each RHS:

$$\dot{x_i} = b_i x_i + (a_{i,1}x_1 + \cdots + a_{i,L}x_L)x_i + \delta_i x_i + \gamma_i \dot{x}_i.$$

For each i, δ_i and γ_i are calibrated using data from the true (complete) system. For details of the calibration process, see [9].

It is easy to see that $\boldsymbol{\delta}$ corrects for equilibrium and $\boldsymbol{\gamma}$ adjusts the overall transient behavior. Let $(\gamma^*)_i = \frac{1}{1-\gamma_i}$. We can rewrite the system with

$$\hat{\boldsymbol{A}} = \operatorname{diag}(\boldsymbol{\gamma}^*)\boldsymbol{A}_{L\times L}, \qquad \hat{\boldsymbol{b}} = \operatorname{diag}(\boldsymbol{\gamma}^*)(\boldsymbol{b}_{1:L} + \boldsymbol{\delta}).$$

Then the equilibrium of this system is given by

$$\hat{\boldsymbol{c}} = -\boldsymbol{A}^{-1}(\boldsymbol{b}_{1:L} + \boldsymbol{\delta})$$

and the eigenvalues of the new linearized system are the eigenvalues of

$$\operatorname{diag}(\boldsymbol{\gamma}^*)\operatorname{diag}(\hat{\boldsymbol{c}})\boldsymbol{A}_{L\times L}.$$

So we conjecture that a good reduced model sets $\boldsymbol{\delta}$ such that $\hat{\boldsymbol{c}} = \boldsymbol{c}_{1:L}$ (i.e., set $\boldsymbol{\delta} = -(\boldsymbol{A}\boldsymbol{c} + \boldsymbol{b})_{1:L}$) and sets $\boldsymbol{\gamma}$ so that the first L eigenvalues (in decreasing order) match the first L eigenvalues of the $\operatorname{diag}(\boldsymbol{c})\boldsymbol{A}$. Since we have L unknowns (degrees of freedom) in the vector $\boldsymbol{\gamma}$ and we have a target of the first L eigenvalues this should be possible algebraically, at least in some generic sense.

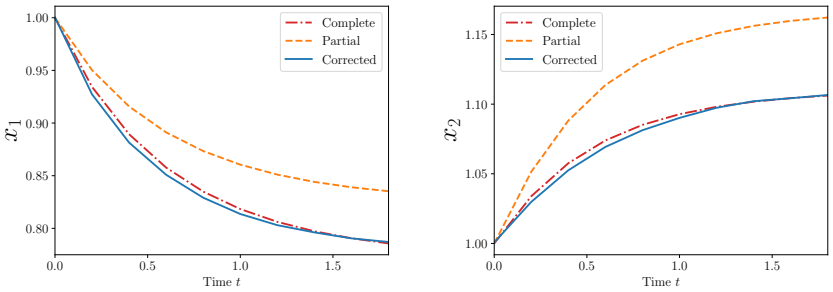
As an example, let's consider the system in section 4.2, and reduce the system to 2 variables:

$$\boldsymbol{A}_{L\times L} = \begin{pmatrix} -2 & -3/10 \\ -2/10 & -2 \end{pmatrix}, \quad \boldsymbol{b}_{1:L} = \begin{pmatrix} 2 \\ 5/2 \end{pmatrix}.$$

In this example, we find that for x_1, $\delta_1 \approx -0.14$ and $\gamma_1 \approx -0.01$, and for x_2, $\delta_2 \approx -0.11$ and $\gamma_2 \approx -0.07$. In fact, this is what happens for this example. With the values of $\boldsymbol{\delta}$ and $\boldsymbol{\gamma}$ given above, we find the following, to the hundredths:

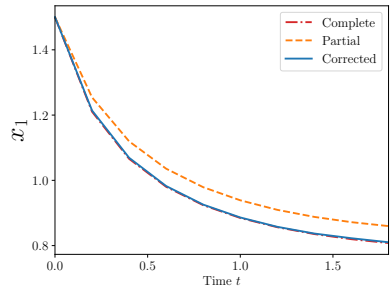
- $c_1 = 0.77$, $\hat{c}_1 = 0.76$
- $c_2 = 1.11$, $\hat{c}_2 = 1.12$
- $\lambda_1 = -1.48$, $\hat{\lambda}_1 = -1.44$
- $\lambda_2 = -2.11$, $\hat{\lambda}_2 = -2.16$.

Finally, note that the four values above of equilibrium and eigenvalues (c_1, c_2, λ_1, λ_2) do not depend on the initial conditions, so a corrected reduced model may be valid over a range of initial conditions, and not just for a single trajectory. This broad domain of validity is critical if the reduced model is meant to be used in other contexts beyond the calibration scenario, which is typically what one asks of a model.

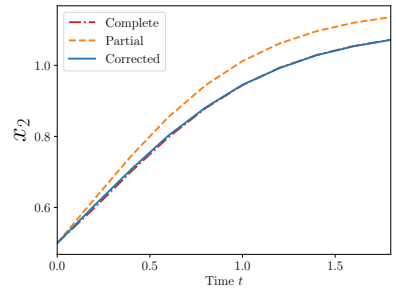


(A) $\boldsymbol{x}(0) = (1, 1, 1)$

(B) $\boldsymbol{x}(0) = (1, 1, 1)$



(C) $\boldsymbol{x}(0) = (3/2, 1/2, 3/2)$



(D) $\boldsymbol{x}(0) = (3/2, 1/2, 3/2)$

FIGURE 4. Numerical solutions for complete, partial, and corrected models. The complete model includes 3 variables, but the corrected model only includes 2 of those. The partial model retains the relevant sub-matrix and sub-vector from the complete model.

7. Conclusion

Analytic solutions for nonlinear sets of differential equations are few and far between. We find analytic solutions to a class of nonlinear coupled ODES with quadratic RHS via power series expansions, and furthermore, we:

(1) Show, for the logistic equation, if *any* power series solution exists, it must be of the exponential form given here;
(2) Derive an explicit construction for the coefficients and how they decay;
(3) Prove these power series solutions converge for $t > t_0$ and give a bound;
(4) Compare the analytic solutions to numerical solutions through various experiments;
(5) Show how, considering the eigenvalues of the linearized system, reduced systems may still capture much of the system dynamics.

We conclude with a couple remarks about these solutions. First, this work reveals the exponential nature of the trajectories, even away from equilibrium. For a given problem and for $t > t_0$, the solutions are infinitely smooth and differentiable, and approach equilibrium exponentially fast. Besides its theoretical interest, this

may inform numerical analysis issues such as convergence, order, and stability, if a numerical solver is still to be used.

Second, numerical experiments typically result in coefficients which decay quickly. Thus, a truncated series even with relatively small N, say $N \approx \gamma M$ where $\gamma = 2$ or 3 can result in very good closed-form approximations to the true solution. If one simply needs a quick estimate of the solution, a truncated series comprised of only a handful terms may do the trick. Third, nothing in this work seems to fundamentally rely on the particular form of the RHS. We propose that many sets of coupled autonomous ODES with polynomial RHS will admit a similar exponential power series solution. Of course, solving for the coefficients may become more complicated (with, e.g., higher-order interactions or complex eigenvalues) but we expect these types of solutions, at least for some semi-bounded time-domain, to exist and converge.

References

[1] Eber Dantas, Michel Tosin, and Americo Cunha Jr. Calibration of a SEIR–SEI epidemic model to describe the Zika virus outbreak in Brazil. *Applied Mathematics and Computation*, 338:249–259, 2018.

[2] James H Espenson. *Chemical kinetics and reaction mechanisms*, volume 2. Citeseer, 1995.

[3] G Kaniadakis. Novel predator-prey model admitting exact analytical solution. *Physical Review E*, 106(4):044401, 2022.

[4] Ellen Kuhl. *Computational Epidemiology*. Springer, 2021.

[5] Bian Li, Yian Ma, J Nathan Kutz, and Xiu Yang. The adaptive spectral Koopman method for dynamical systems. *SIAM Journal on Applied Dynamical Systems*, 22(3):1523–1551, 2023.

[6] Alexandre Mauroy, Y Susuki, and I Mezić. *Koopman operator in systems and control*. Springer, 2020.

[7] Igor Mezić. Spectrum of the Koopman operator, spectral expansions in functional spaces, and state-space geometry. *Journal of Nonlinear Science*, 30:2091—-2145, 2019.

[8] Rebecca E Morrison. Data-driven corrections of partial Lotka–Volterra models. *Entropy*, 22(11):1313, 2020.

[9] Rebecca E Morrison and Americo Cunha. Embedded model discrepancy: A case study of Zika modeling. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 30(5), 2020.

[10] Gregory P. Smith, David M. Golden, Michael Frenklach, Nigel W. Moriarty, Boris Eiteneer, Mikhail Goldenberg, C. Thomas Bowman, Ronald K. Hanson, Soonho Song, Jr. William C. Gardiner, Vitali V. Lissianski, and Zhiwei Qin. GRI-Mech v.3.0.

[11] Steven H Strogatz. *Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering*. CRC press, 2018.

[12] Forman A Williams. *Combustion theory*. CRC Press, 2018.

Estelle Basor
ebasor@aimath.org

Rebecca Morrison
rebeccam@colorado.edu