Electronic Transactions on Numerical Analysis.
Volume 60, pp. 292–326, 2024.
Copyright © 2024, Kent State University.
ISSN 1068–9613.
DOI: 10.1553/etna_vol60s292

A NEW LEGENDRE POLYNOMIAL-BASED APPROACH FOR NON-AUTONOMOUS LINEAR ODES[*]

STEFANO POZZA[†] AND NIEL VAN BUGGENHOUT[‡]

Abstract. We introduce a new method with spectral accuracy to solve linear non-autonomous ordinary differential equations (ODEs) of the kind $\frac{d}{dt}\tilde{u}(t) = \tilde{f}(t)\tilde{u}(t)$, $\tilde{u}(-1) = 1$, with $\tilde{f}(t)$ an analytic function. The method is based on a new analytical expression for the solution $\tilde{u}(t)$ given in terms of a convolution-like operation, the $\star$-product. We prove that, by representing this expression in a finite Legendre polynomial basis, the solution $\tilde{u}(t)$ can be found by solving a matrix problem involving the Fourier coefficients of $\tilde{f}(t)$. An efficient procedure is proposed to approximate the Legendre coefficients of $\tilde{u}(t)$, and the truncation error and convergence are analyzed. We show the effectiveness of the proposed procedure through numerical experiments. Our approach allows for a generalization of the method to solve systems of linear ODEs.

Key words. Legendre polynomials, spectral accuracy, ordinary differential equations

AMS subject classifications. 65F60, 65L05, 35Q41

1. Introduction.

A new numerical approach to solve the ordinary differential equation (ODE)

$$\frac{d}{dt}\tilde{u}(t) = \tilde{f}(t)\tilde{u}(t), \quad \tilde{u}(-1) = 1, \quad t \in [-1, 1]\,, \tag{1.1}$$

where $\tilde{f}(t)$ is a known smooth function, is proposed. Note that any finite interval can be rescaled to $[-1, 1]$. The proposed method computes the Fourier coefficients, in a basis of Legendre polynomials, of the solution $\tilde{u}(t)$ on $[-1, 1]$ by manipulating matrices containing Fourier coefficients related to $\tilde{f}(t)$. The goal of this paper is to introduce, analyze, and numerically verify the new numerical approach. The study we present discusses only the case in which $\tilde{f}$ is a scalar function. Nevertheless, the results obtained in this paper form the necessary basis for extending the approach to the more challenging problem of (large) systems of ODEs. Indeed, our interest in this new approach arises from its significance in developing and analyzing a numerical method for the matrix case. Let $\tilde{A}(t)$ be an $N \times N$ analytic matrix-valued function over the interval $[-1, 1]$. Then, the unique solution of the matrix ODE

$$\frac{d}{dt}\tilde{U}(t) = \tilde{A}(t)\tilde{U}(t), \quad \tilde{U}(-1) = I_N, \quad t \in [-1, 1]\,, \tag{1.2}$$

is also an analytic $N \times N$ matrix-valued function $\tilde{U}(t)$, where I_N is the identity matrix of size $N \times N$. In [6], we demonstrated that a numerical method for solving matrix ODEs, which uses similar concepts as the scalar numerical method proposed here, outperforms the state-of-the-art methods for an important benchmarking problem. The mathematical explanation of these numerical results can be built from the foundation we lay in this paper.

Systems of non-autonomous linear ODEs are ubiquitous in mathematics and related applications. An application of particular interest is nuclear magnetic resonance spectroscopy

[*] Received March 26, 2024. Accepted April 2, 2024. Published online on June 3, 2024. Recommended by Kevin Burrage. This work was supported by Charles University Research programs UNCE/24/SCI/005 and PRIMUS/21/SCI/009 and by the Magica project ANR-20-CE29-0007 funded by the French National Research Agency.

[†] Department of Numerical Mathematics, Charles University, Sokolovská 83, 186 75, Praha 8, Czech Republic (pozza@karlin.mff.cuni.cz).

[‡] Departamento de Matemáticas, Universidad Carlos III de Madrid, Avenida de la Universidad 30, 28911 Leganés, Spain (nvanbugg@math.uc3m.es).

(NMR). In NMR, the given matrix-valued function is of the form $\tilde{A}(t) = -2\pi\imath\tilde{H}(t)$, where $\tilde{H}(t)$ is the Hamiltonian of the system [21, 27]. The Hamiltonian describes the dynamics of the nuclear spins in some sample that is placed in a varying magnetic field. For ℓ spins, the Hamiltonian is of size $2^\ell \times 2^\ell$. So, even for a moderate amount of spins, the Hamiltonian becomes extremely large. Thankfully, this matrix is usually sparse since the dominant interactions are those between neighboring spins.

Our novel numerical approach is based on a recently developed analytical theory [15, 16, 18]. This theory builds on a product of bivariate (matrix-valued) functions, called the $\star$-product, and provides a simple expression for the solution of (1.2). Let $\tilde{F}(t,s)$ be a matrix-valued function analytic[1] in both variables over $[-1,1]$, and let $\Theta(t-s)$ be the Heaviside step function

$$\Theta(t-s) = \begin{cases} 0 & \text{if } t < s, \\ 1 & \text{otherwise.} \end{cases}$$

We denote by $\mathcal{A}_\Theta^N$ the set of all the distributions of the kind $F(t,s) := \tilde{F}(t,s)\Theta(t-s)$ with size $N \times N$. Given $F, G \in \mathcal{A}_\Theta^N$, the $\star$-product [17], denoted by $\star$, is defined as

$$F(t,s) \star G(t,s) := \int_{-1}^{1} F(t,\tau)G(\tau,s)d\tau.$$

From the definition above, by replacing the matrix-matrix product in the integrand with an appropriate one, the $\star$-product can be easily extended to matrix-vector, matrix-scalar, vector-scalar, and scalar-scalar products. The identity under the $\star$-product is the distribution $I_\star(t-s) = \delta(t-s)I_N$, where $\delta(t-s)$ is the Dirac delta distribution [40]. Overall, we have defined an algebraic structure [36] composed of the set of distributions $\mathcal{D}_0^N := \mathcal{A}_\Theta^N \cup \{I_\star\}$, the $\star$-product (interpreted both as a matrix-matrix and a scalar product), and the usual addition. Table 1.1 summarizes the $\star$-product properties and other related definitions.

TABLE 1.1
Operations and related objects in the $\star$-framework.

Operations and related objects	Description
$F(t,s) \star G(t,s)$	matrix-matrix product $\mathcal{D}_0^N \times \mathcal{D}_0^N \to \mathcal{D}_0^N$
$f(t,s) \star G(t,s)$	(left) scalar product $\mathcal{D}_0^1 \times \mathcal{D}_0^N \to \mathcal{D}_0^N$
$G(t,s) \star f(t,s)$	(right) scalar product $\mathcal{D}_0^N \times \mathcal{D}_0^1 \to \mathcal{D}_0^N$
$F + G$	usual addition
$I_\star = \delta(t-s)I_N$	identity (Dirac delta)
$R_\star(F)(t,s) = I_\star + \sum_{k\geq 1} F^{\star k}(t,s)$	$\star$-resolvent

The ODE (1.2) can be formulated in the $\star$-framework in terms of $A(t,s) := \tilde{A}(t)\Theta(t-s)$:

$$\frac{d}{dt}U(t,s) = A(t,s)U(t,s), \quad U(s,s) = I_N, \quad t,s \in [-1,1]\,. \tag{1.3}$$

By Theorem 3.1 in [16], the solution to this ODE can be expressed in the form

$$U(t,s) = \Theta(t-s) \star R_\star(A)(t,s), \tag{1.4}$$

[1] In the previously appeared works on the $\star$-product, the functions are usually assumed to be smooth. Here, we restrict the assumption to analytic for the sake of simplicity since the application we are considering deals with analytic functions.

where $R_\star(\cdot)$ is the $\star$-resolvent, i.e., $R_\star(A)(t,s) = I_\star + \sum_{k\geq 1} A^{\star k}(t,s)$, with $A^{\star k}(t,s)$ the kth $\star$-power of A. Once $U(t,s)$ is known, the solution to the original ODE can be obtained by evaluation in $s=-1$, i.e., $\tilde{U}(t) = U(t,-1)$. The scalar case is retrieved by choosing $A(t,s) = f(t,s) := \tilde{f}(t)\Theta(t-s) \in \mathcal{A}^1_\Theta$; in this case we have $U(t,s) = u(t,s)$, where $u(t,-1) = \tilde{u}(t)$ is the solution of (1.1). In the remainder of this paper we denote $\mathcal{A}^1_\Theta$ by $\mathcal{A}_\Theta$.

Expression (1.4) for the scalar case $u(t,s)$ is the starting point for our numerical approximation method for $\tilde{u}(t)$. This method relies on representing the bivariate distributions by *coefficient matrices* containing their expansion coefficients in a basis of Legendre polynomials. The operations in Table 1.1 can be equivalently written in terms of these coefficient matrices, most notably, the $\star$-product between bivariate distributions is replaced by the usual matrix-matrix multiplication between coefficient matrices. This reformulation leads to an infinite linear system of equations whose solution provides the Legendre coefficients of $\tilde{u}(t)$. A suitable truncation of this system of equations leads to a finite problem that can be solved with techniques from numerical linear algebra, and its solution provides the first Legendre coefficients of $\tilde{u}(t)$. Since the expression (1.4) is also valid for the matrix case, the numerical procedure described above can be generalized to solve matrix ODEs. Such a generalization to a matrix method brings additional computational challenges such as the need for efficient manipulation of large block coefficient matrices and solving systems involving such block matrices, as well as theoretical challenges such as a truncation error analysis in matrix-equation form. This is part of ongoing research.

In this paper, we address these questions for the scalar case, i.e., $N=1$. The results on the scalar case form the building blocks for the matrix case, both the theoretical results and the numerical algorithm. Very preliminary but promising results of the efficiency of the proposed approach for the matrix case can also be found in the master thesis [26]. The next section provides an overview of numerical methods for the matrix case, i.e., systems of ODEs, and motivates our interest in developing a numerical method based on (1.4).

1.1. Methods for systems of ODEs. In the scalar case, the solution to (1.1) is given by the exponential $\tilde{u}(t) = \exp\left(\int_{-1}^t \tilde{f}(\tau)d\tau\right)$. For the matrix case, the solution $\tilde{U}(t)$ to (1.2), for $A(t)$ satisfying $\tilde{A}(\tau_1)\tilde{A}(\tau_2) = \tilde{A}(\tau_2)\tilde{A}(\tau_1)$ for every $\tau_1, \tau_2 \in [-1,1]$, can be expressed simply as a matrix exponential:

$$\tilde{U}(t) = \exp\left(\int_s^t \tilde{A}(\tau)\,\mathrm{d}\tau\right).$$

However, for general $A(t)$, $\tilde{U}(t)$ cannot be written straightforwardly as a matrix exponential involving $\tilde{A}(t)$. There are many approaches, and we will discuss some of them. We distinguish them by the length of the time interval they can handle.

Local methods. Local methods require that the interval is split up into small subintervals, and the ODE is solved, sequentially, on these subintervals. Hence, they are time-stepping methods, such as the Runge–Kutta methods. They can be computationally expensive due to the accumulation of errors during time stepping and the resulting need for very small time steps. Some of these general solvers preserve certain qualitative properties of the solution, e.g., the unitarity of $\tilde{U}(t)$ if $\tilde{A}(t)$ is skew-Hermitian. However, the error accumulates for certain qualitative properties that are not preserved by the method.

Lie-group methods [23] are specifically designed to preserve qualitative (geometric) properties. Those based on the Magnus expansion represent the solution as $\tilde{U}(t) = \exp(\sum_{k=0}^\infty H_k)$, where H_k is a k-fold integral over a polygon involving k commutators of $\tilde{A}(t)$ evaluated at different times. These integrals can be solved efficiently, e.g., by analytic integration for some cases, by quadrature, or by using the Lanczos iteration [22, 23]. An approximation is obtained by truncation $\tilde{U}(t) \approx \exp(\sum_{k=0}^{n} H_k)$. Most common methods of this type have an order of convergence below 10 because higher-order methods become more and more costly. Other Lie-group methods are based on the Fer expansion, which represents the solution as an infinite product of exponentials, $\tilde{U}(t) = \prod_{k=0}^{\infty} \exp(\int_0^t B_k(\tau)d\tau)$, where $B_n(t)$ is essentially also an n-fold integral; for details see [28]. A truncation, $\tilde{U}(t) \approx \prod_{k=0}^{n} \exp(\int_0^t B_k(\tau)d\tau)$, provides an approximation. Since the Magnus and Fer expansion have a limited radius of convergence, the time interval must be split up into small subintervals [10]. Due to the low order of convergence of local methods and the accumulation of errors during time-stepping, they are relatively costly for the accuracy they provide when $\tilde{A}(t)$ is smooth. Another local method is the Cayley method [11], where linear systems of equations have to be solved at each step.

Semi-global methods. Some recent methods exploit the smoothness of $\tilde{A}(t)$ by taking larger time steps. For a time-independent Hamiltonian, a global method (computing an approximation in a single time step) exists, which is based on computing a polynomial approximation of the matrix exponential. However, since the solution for a time-dependent Hamiltonian cannot be simply written in terms of a matrix exponential, this method cannot be easily generalized. Recent generalizations of such polynomial approximation methods rely on a reformulation of (1.2) as an integral equation [29, 38]. By a truncated Chebyshev expansion of the integrand, the integral is discretized, and thus, the solution can be approximated by a fixed-point iteration. In order for the fixed-point iteration to converge, the time step must be restricted, although the time subinterval can be chosen much larger than for local methods. Moreover, on each subinterval there is spectral accuracy thanks to the use of a polynomial approximation. A similar method [39], using other discretizations than Chebyshev polynomials, also requires a restriction of the time interval. Since these methods take large steps but cannot, in general, solve the ODE on the full time interval in a single step, they are often referred to as *semi-global methods*.

Global methods. Two important global methods exist that solve the ODE in a single step. The first is the (t, t')-method [30]. However, it is too expensive in terms of computation and memory [24, 29, 38]. The second is the class of Hamiltonian boundary value methods (HBVM) [1, 7, 8, 9]. These methods are designed as local methods, where on each local subinterval they use an implicit Runge–Kutta-type approximation with a basis of orthogonal polynomials, in particular, the shifted Legendre polynomials. Thanks to the use of the Legendre basis they can increase the degree of the Legendre polynomials and achieve spectral convergence on the whole interval of interest in a single step [1, 8, 9]. This subclass of HBVMs is referred to as spectral HBVMs (SHBVMs). HBVMs expand the right-hand side of the ODE, $f(t)u(t)$, in terms of Legendre polynomials. Since the solution $u(t)$ is unknown, an approximation of the right-hand side must be made, which is achieved by an implicit Runge–Kutta-type approach. Our method is also global and based on a Legendre expansion. However, it takes a completely different approach than HBVM since it expands the bivariate function $f(t)\Theta(t-s)$ in a bivariate Legendre basis. This bivariate function is thus known a-priori, allowing us to compute the results using the Legendre coefficients directly. We will show that equation (1.4) allows us to develop a numerical method that computes the Legendre coefficients of the solution $u(t)$ from the coefficients of the bivariate function.

Due to our use of Legendre polynomials, the interval is required to be finite. In order to obtain a computationally efficient approach, it is paramount to exploit present matrix structures, that is, the structure of $\tilde{A}(t)$ and the structure of the matrix representing the linear system of equations obtained after expansion in a Legendre polynomial basis [33].

Preliminary numerical testing of our proposed method (see the conference proceedings [34, 35]) illustrate that for some simple examples, our method has spectral accuracy (on the whole time interval) and can be implemented efficiently by exploiting matrix structures. The preservation of geometric properties for this novel method is not known; some preliminary numerical results suggest that, if $A(t)$ is skew-Hermitian, then the unitarity of $\tilde{U}(t)$ is not preserved. However, due to the high accuracy of our method, the approximation to $\tilde{U}(t)$ will be unitary up to machine precision or some user-specified precision. Note that since a single polynomial series represents the solution $\tilde{U}(t)$ on the whole time interval, the value of the approximation at any time point can be easily obtained by evaluation of the polynomial series in that point, whereas other methods must rely on interpolation in the time points used during time stepping.

1.2. Outline. Section 2 handles the discretization of the $\star$-operations in the Legendre basis and shows that the problem of solving the scalar version of the ODE (1.3) can be reformulated as an infinite matrix problem with the coefficient matrix of $f(t,s) \in \mathcal{A}_\Theta$. The properties of the coefficient matrix are analyzed in detail in Section 3, and an analytical formula is provided for its entries. This analysis leads to an efficient numerical method to construct the coefficient matrix of $f(t,s)$. Most notably, we prove that coefficient matrices can be approximated by banded matrices. In Section 4, these matrix properties are used to show that the infinite problem can be approximated by a finite problem, and techniques to efficiently solve this finite problem are proposed. Section 5 formulates the finite matrix problem corresponding to approximating $\tilde{u}(t)$ and proposes a numerical procedure to solve it. The effectiveness of this procedure is illustrated by numerical examples.

2. From the $\star$-product to the matrix algebra.

The proposed numerical method for the approximation of $\tilde{u}(t)$ is based on the expansion of the distribution $f(t,s) \in \mathcal{A}_\Theta$ in a basis of orthonormal Legendre polynomials. The distribution $f(t,s)$ can be represented by its *coefficient matrix*, which contains the expansion (Fourier) coefficients of $f(t,s)$. The solution to (1.1), $\tilde{u}(t) = u(t,-1)$, given by (1.4) is obtained by exploiting the connection between the $\star$-product and the usual matrix algebra. Section 2.1 discusses the expansion of functions and distributions in the basis of orthonormal Legendre polynomials and defines the coefficient matrix. In Section 2.2, the connection between the $\star$-product and the matrix algebra is used to reformulate the problem in (1.1) as an infinite matrix problem; see also [31].

2.1. Legendre polynomials. The sequence $\{p_k\}_{k\geq 0}$ of orthonormal Legendre polynomials consists of polynomials p_k of exact degree k that satisfy the orthonormality conditions

$$\int_{-1}^{1} p_k(t)p_\ell(t)dt = \begin{cases} 0 & \text{if } k \neq \ell, \\ 1 & \text{if } k = \ell. \end{cases}$$

For a univariate function $\tilde{f}(t)$, its expansion into the Legendre basis is given by

$$\tilde{f}(t) := \sum_{d=0}^{\infty} \alpha_d p_d(t), \qquad \text{with } \alpha_d = \int_{-1}^{1} \tilde{f}(t)p_d(t)dt.$$

The Fourier coefficients $\{\alpha_d\}_{d\geq 0}$ decay at a rate depending on the smoothness of $\tilde{f}(t)$. Any analytic function over $[-1,1]$ allows an analytic continuation to a Bernstein ellipse $\mathcal{E}_\rho := \{z | z = \frac{\rho}{2}e^{i\theta} + \frac{1}{2\rho}e^{-i\theta}, -\pi \leq \theta \leq \pi\}$ for a $\rho > 1$ small enough. Therefore, for some constant $C > 0$, the Fourier coefficients satisfy

$$|a_d| \leq C\rho^{-d-1}; \tag{2.1}$$

for details we refer to [42, 44]. Moreover, the orthonormal Legendre polynomials can be bounded by

$$|p_d(t)| \leq \sqrt{\frac{2d+1}{2}}, \qquad \text{for } \ t \in [-1,1]\,,$$

and therefore the truncated expansion $\hat{f}_N(t) := \sum_{d=0}^{N} \alpha_d p_d(t)$ has the error, measured in the maximum norm for $t \in [-1,1]$,

$$\|\tilde{f}(t) - \hat{f}_N(t)\|_\infty = \sum_{d=N+1}^{\infty} \alpha_d p_k(t) \leq \sum_{d=N+1}^{\infty} |\alpha_d| \sqrt{\frac{2d+1}{2}}.$$

Hence, if the (decaying) coefficients $\alpha_N, \alpha_{N+1}, \dots$ are smaller than a given threshold, then the truncation $\hat{f}_N(t)$ can provide a good approximation of $\tilde{f}(t)$.

Consider a distribution $f \in \mathcal{A}_\Theta$. Its Legendre expansion is

$$f(t,s) = \sum_{k=0}^{\infty}\sum_{\ell=0}^{\infty} f_{k,\ell}\, p_k(t) p_\ell(s), \qquad \text{for every } t \neq s, t, s \in [-1,1]\,,$$

with Fourier coefficients given by

$$f_{k,\ell} = \int_{-1}^{1}\int_{-1}^{1} f(\tau,\rho) p_k(\tau) p_\ell(\rho) d\rho d\tau. \tag{2.2}$$

We define the *coefficient matrix* F of the distribution $f(t,s)$, which is the infinite matrix composed of the Fourier coefficients (2.2),

$$F := \left[f_{k,\ell}\right]_{k,\ell=0}^{\infty} = \begin{bmatrix} f_{0,0} & f_{0,1} & f_{0,2} & \cdots \\ f_{1,0} & f_{1,1} & f_{1,2} & \cdots \\ f_{2,0} & f_{2,1} & f_{2,2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

The distribution $f(t,s)$ is only piecewise smooth because the Heaviside function $\Theta(t-s)$ introduces discontinuities. Thus, its Fourier coefficients $f_{k,\ell}$ do not decay at a geometric rate. Due to these discontinuities, there is essentially no decay in the coefficients, and the Gibbs phenomenon arises [19]. This means that the reconstruction of $f(t,s)$ over the entire domain $[-1,1] \times [-1,1]$ by using only the coefficients $f_{k,\ell}$ is not possible. For such a reconstruction, there is no convergence at the discontinuities $t = s$, and away from the discontinuities, it converges only linearly to the actual values. There are techniques to resolve the Gibbs phenomenon; see, for example, [13, 19]. In our setting such techniques are not needed since we only need accurate coefficients $f_{k,\ell}$ representing $f(t,s)$ in the Legendre basis, and we will not use these to reconstruct the function values on the entire domain $(t,s) \in [-1,1] \times [-1,1]$ but only on $t \in [-1,1]$ for $s = -1$, i.e., where the function is analytic in t.

2.2. A matrix formulation. The operations of addition and $\star$-multiplication for distributions in $\mathcal{A}_\Theta$ have equivalent operations in the matrix algebra of the associated coefficient matrices, namely the usual matrix addition and matrix-matrix multiplication.

LEMMA 2.1. *Consider $f, g \in \mathcal{A}_\Theta$ and their respective coefficient matrices F, G in the Legendre basis. Then:*

- *$f + g = h \in \mathcal{A}_\Theta$, and its coefficient matrix is $H = F + G$.*
- *$f \star g = h \in \mathcal{A}_\Theta$, and, assuming the matrix product is well defined, its coefficient matrix is $H = FG$.*

Proof. Addition: from the Legendre expansion of f and g it follows that

$$\begin{aligned} h = f + g &= \begin{bmatrix} p_0(t) & p_1(t) & \cdots \end{bmatrix} F \begin{bmatrix} p_0(s) \\ p_1(s) \\ \vdots \end{bmatrix} + \begin{bmatrix} p_0(t) & p_1(t) & \cdots \end{bmatrix} G \begin{bmatrix} p_0(s) \\ p_1(s) \\ \vdots \end{bmatrix} \\ &= \begin{bmatrix} p_0(t) & p_1(t) & \cdots \end{bmatrix} \underbrace{(F+G)}_{=H} \begin{bmatrix} p_0(s) \\ p_1(s) \\ \vdots \end{bmatrix}. \end{aligned}$$

Multiplication: plugging in the double series and using the definition of the $\star$-product provides

$$\begin{aligned} h = f \star g &= \int_{-1}^{1} \begin{bmatrix} p_0(t) & p_1(t) & \cdots \end{bmatrix} F \begin{bmatrix} p_0(\tau) \\ p_1(\tau) \\ \vdots \end{bmatrix} \begin{bmatrix} p_0(\tau) & p_1(\tau) & \cdots \end{bmatrix} G \begin{bmatrix} p_0(s) \\ p_1(s) \\ \vdots \end{bmatrix} d\tau \\ &= \begin{bmatrix} p_0(t) & p_1(t) & \cdots \end{bmatrix} F \left(\int_{-1}^{1} \begin{bmatrix} p_0(\tau)p_0(\tau) & p_0(\tau)p_1(\tau) & \cdots \\ p_1(\tau)p_0(\tau) & p_1(\tau)p_1(\tau) & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} d\tau \right) G \begin{bmatrix} p_0(s) \\ p_1(s) \\ \vdots \end{bmatrix}. \end{aligned}$$

Thanks to the orthonormality of $\{p_k(t)\}_{k\geq 0}$, the matrix in the middle equals the identity matrix

$$\begin{bmatrix} \int_{-1}^{1} p_0(\tau)p_0(\tau)d\tau & \int_{-1}^{1} p_0(\tau)p_1(\tau)d\tau & \cdots \\ \int_{-1}^{1} p_1(\tau)p_0(\tau)d\tau & \int_{-1}^{1} p_1(\tau)p_1(\tau)d\tau & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots \\ 0 & 1 & \ddots \\ \vdots & \ddots & \ddots \end{bmatrix}.$$

Thus, we get

$$h = f \star g = \begin{bmatrix} p_0(t) & p_1(t) & \cdots \end{bmatrix} \underbrace{FG}_{=H} \begin{bmatrix} p_0(s) \\ p_1(s) \\ \vdots \end{bmatrix},$$

that is, the coefficient matrix for h is $H = FG$ under the assumption that this matrix product is well defined. □

In Section 3.3, the infinite matrix product is discussed, and we show that the matrix product between coefficient matrices of distributions $f \in \mathcal{A}_\Theta$ is always well defined. If $f(t,s)$ is bounded for $t, s \in [-1,1]$, then the $\star$-resolvent of $f(t,s)$, which is $R_\star(f) = 1_\star + \sum_{k\geq 1} f^{\star k}$, exists since the series $\sum_{k\geq 1} f^{\star k}$ converges uniformly in $\mathcal{A}_\Theta$ for every $t, s \in [-1,1]$; see [16]. Since

$$R_\star(f) \star (1_\star - f) = \left(1_\star + \sum_{k\geq 1} f^{\star k} \right) \star (1_\star - f) = 1_\star,$$

the $\star$-resolvent is the $\star$-inverse of $(1_\star - f)$, i.e., $R_\star(f) = (1_\star - f)^{-\star}$. Let $g := \sum_{k\geq 1} f^{\star k}$. Then $g \in \mathcal{A}_\Theta$, and hence we can define its coefficient matrix G. Therefore, we have

$$R_\star(f) = 1_\star + \sum_{k\geq 1} f^{\star k} = \phi(t)^T\,(I+G)\,\phi(s), \qquad \text{with } \phi(\tau) := \begin{bmatrix} p_0(\tau) \\ p_1(\tau) \\ \vdots \end{bmatrix},$$

which allows us to derive the following relation between $(I+G)$ and $(I-F)$,

$$\begin{aligned} 1_\star = (R_\star(f) \star (1_\star - f))(t,s) &= \left(\phi(t)^T\,(I+G)\,\phi(s)\right) \star \left(\phi(t)^T\,(I-F)\,\phi(s)\right), \\ &= \phi(t)^T\,(I+G)\,(I-F)\,\phi(s) = \phi(t)^T I\,\phi(s). \end{aligned} \tag{2.3}$$

As a consequence, we have the following result:

LEMMA 2.2. *Consider $f \in \mathcal{A}_\Theta$ and its corresponding coefficient matrix F. If the inverse of the infinite matrix $(I-F)$ exists, then*

$$R_\star(f) = \begin{bmatrix} p_0(t) & p_1(t) & \dots \end{bmatrix} (I-F)^{-1} \begin{bmatrix} p_0(s) \\ p_1(s) \\ \vdots \end{bmatrix}.$$

Proof. Let us define $R_\star(f) := \phi(t)^T(I-F)^{-1}\phi(s)$, i.e., set $(I+G) = (I-F)^{-1}$. Then, by equations (2.3), we get $R_\star(f) \star (1_\star - f) = 1_\star$. □

Combining Lemmas 2.1 and 2.2 allows us to obtain an expression for the Legendre coefficients of $\tilde{u}(t)$ in terms of coefficient matrices. This expression is the matrix counterpart to the expression for $\tilde{u}(t)$ in the $\star$-framework: $\tilde{u}(t) = u(t,s)|_{s=-1} = \Theta(t-s) \star R_\star(f)|_{s=-1}$ (see (1.4)) and is stated in the following theorem:

THEOREM 2.3. *Consider $f \in \mathcal{A}_\Theta$ and its corresponding coefficient matrix F. Let T denote the coefficient matrix of $\Theta(t-s)$, I the identity matrix, and $\{p_k\}_{k\geq 0}$ the sequence of orthonormal Legendre polynomials. Assume that $(I-F)$ is invertible. Then the Legendre coefficients $\{c_k\}_{k\geq 0}$ of the solution $\tilde{u}(t)$ of the ODE (1.1) are given by*

$$\begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \end{bmatrix} = T(I-F)^{-1} \begin{bmatrix} p_0(-1) \\ p_1(-1) \\ p_2(-1) \\ \vdots \end{bmatrix}. \tag{2.4}$$

Based on Theorem 2.3, we can formulate a matrix problem that is equivalent to the problem of solving the ODE (1.1).

PROBLEM 2.1 (Infinite matrix problem). *Given a smooth function $\tilde{f}(t)$, compute the Legendre coefficients $\{c_k\}_{k=0}^\infty$ of the solution $\tilde{u}(t)$ to the ODE (1.1). By (2.4) this corresponds to three matrix problems:*

1. *Construct the infinite coefficient matrix $F = \left[f_{k,\ell}\right]_{k,\ell=0}^\infty$ of Fourier coefficients in the Legendre basis $f_{k,\ell} = \int_{-1}^1 \int_{-1}^1 \tilde{f}(\tau)\Theta(\tau-\rho)p_k(\tau)p_\ell(\rho)d\rho d\tau$.*
2. *Solve the infinite linear system of equations $(I-F)x = \phi(-1)$ for x. The right-hand side is the column vector $\phi(-1) = \left[p_k(-1)\right]_{k=0}^\infty$, and I is the infinite identity matrix.*
3. *Compute the matrix-vector product $Tx = \begin{bmatrix} c_0 & c_1 & c_2 & \cdots \end{bmatrix}^\top$, where T is the coefficient matrix of $\Theta(t-s)$.*

Problem 2.1 is the main problem to solve. In the remainder of this paper, we develop a numerical method to approximate its solution and investigate the conditions under which this approximation is expected to converge.

3. The coefficient matrix and its properties.

Since the coefficient matrix is central to our analysis and to the proposed procedure, we study its structure. In Section 3.1, an analytical expression for the entries of the coefficient matrices is presented, and we show that the entries decay along the diagonals. Section 3.2 proves that the coefficient matrices can be approximated by a banded matrix. These results allow us to show, in Section 3.3, that the matrix-matrix product between two coefficient matrices of distributions in $\mathcal{A}_\Theta$ is well defined; see Lemma 2.1.

3.1. Formula for the Fourier coefficients.

The Fourier coefficients $f_{k,\ell}$ (2.2) of $f(t,s) = \tilde{f}(t)\Theta(t-s) \in \mathcal{A}_\Theta$ are studied by relying on the Legendre expansion of the analytical function $\tilde{f}(t) = \sum_{d=0}^{\infty} \alpha_d p_d(t)$, its fast decaying coefficients $\{\alpha_d\}_{d\geq 0}$, and the coefficient matrices of $p_d(t)\Theta(t-s) \in \mathcal{A}_\Theta$. The Fourier coefficients of $p_d(t)\Theta(t-s)$ can be computed via an analytical formula; see Theorem 3.3. This formula follows from combining the two known properties of Legendre polynomials stated below.

PROPERTY 3.1 (Integral of a Legendre polynomial on a subinterval [37, p. 178]). *Let $p_\ell(t)$ denote the orthonormal Legendre polynomial of degree ℓ. For $\ell = 0$ it holds that*

$$\int_{-1}^{\tau} p_0(\rho)d\rho = \frac{1}{\sqrt{3}}p_1(\tau) + p_0(\tau),$$

and for $\ell > 0$

$$\int_{-1}^{\tau} p_\ell(\rho)d\rho = \frac{1}{\sqrt{2\ell+1}}\left(\frac{1}{\sqrt{2\ell+3}}p_{\ell+1}(\tau) - \frac{1}{\sqrt{2\ell-1}}p_{\ell-1}(\tau)\right).$$

PROPERTY 3.2 (Integral of the triple product of Legendre polynomials [14]). *Let p_ℓ be the orthonormal Legendre polynomial of degree ℓ. Consider integers $a, b, c \geq 0$, and set $s := \frac{a+b+c}{2}$ and $\alpha := |b-c|$. The integral of the product of three orthonormal Legendre polynomials is*

$$\begin{aligned}
\mathcal{F}_{a,b,c} &:= \int_{-1}^{1} p_a(\rho)p_b(\rho)p_c(\rho)d\rho \\
&= \begin{cases} 0 & \textit{if } a+b+c \textit{ odd}, \\ 0 & \textit{if } s < \max(a,b,c), \\ 0 & \textit{if } a < |b-c|, \\ \frac{\sqrt{(2a+1)(2b+1)(2c+1)}}{\sqrt{2}(a+b+c+1)} \binom{2s-2a}{s-a}\binom{2s-2b}{s-b}\binom{2s-2c}{s-c}\binom{2s}{s}^{-1} & \textit{else} \end{cases} \\
&= \begin{cases} 0 & \textit{if } a+b+c \textit{ odd}, \\ 0 & \textit{if } b+c < a \\ 0 & \textit{if } a < \alpha, \\ \frac{\sqrt{(2a+1)(2b+1)(2c+1)}}{2^{(2a+1/2)}} \frac{\prod_{j=1}^{a} \frac{-a+b+c+2j}{-a+b+c+2j-1}}{(a+b+c+1)} \frac{\prod_{j=(\frac{a+\alpha}{2}+1)}^{a+\alpha} j^2}{\prod_{j=1}^{\frac{a-\alpha}{2}} j^2 \prod_{j=(a-\alpha+1)}^{a+\alpha} j} & \textit{else.} \end{cases}
\end{aligned}$$

THEOREM 3.3 (Coefficients of a Legendre polynomial in $\mathcal{A}_\Theta$). *Let $p_d(t), p_k(t), p_\ell(s)$ be the orthonormal Legendre polynomials of degree d, k, ℓ, respectively, and $\mathcal{F}_{a,b,c}$ as in Property 3.2. Then the coefficients $b^{(d)}_{k,\ell}$ of the Legendre expansion of $p_d(t)\Theta(t-s)$ are given, for $\ell = 0$, by*

$$b^{(d)}_{k,0} = \frac{1}{\sqrt{3}}\mathcal{F}_{d,k,1} + \mathcal{F}_{d,k,0},$$

and, for $\ell > 0$, by

$$b^{(d)}_{k,\ell} = \frac{1}{\sqrt{2\ell+1}}\left(\frac{1}{\sqrt{2\ell+3}}\mathcal{F}_{d,k,\ell+1} - \frac{1}{\sqrt{2\ell-1}}\mathcal{F}_{d,k,\ell-1}\right). \tag{3.1}$$

Proof. By orthonormality of the Legendre polynomials, the Fourier coefficients for $\ell > 0$ are given by

$$\begin{aligned}
b^{(d)}_{k,\ell} &:= \int_{-1}^{1}\int_{-1}^{1} p_d(\tau)\Theta(\tau-\rho)p_k(\tau)p_\ell(\rho)d\rho d\tau \\
&= \int_{-1}^{1} p_d(\tau)p_k(\tau)\left(\int_{-1}^{1}\Theta(\tau-\rho)p_\ell(\rho)d\rho\right)d\tau \\
&= \int_{-1}^{1} p_d(\tau)p_k(\tau)\underbrace{\left(\int_{-1}^{\tau}p_\ell(\rho)d\rho\right)}_{\text{Apply Property 3.1}}d\tau \\
&= \frac{1}{\sqrt{2\ell+1}}\left[\frac{1}{\sqrt{2\ell+3}}\int_{-1}^{1}p_d(\tau)p_k(\tau)p_{\ell+1}(\tau)d\tau \right.\\
&\qquad\qquad \left. - \frac{1}{\sqrt{2\ell-1}}\int_{-1}^{1}p_d(\tau)p_k(\tau)p_{\ell-1}(\tau)d\tau\right] \\
&= \frac{1}{\sqrt{2\ell+1}}\left[\frac{1}{\sqrt{2\ell+3}}\mathcal{F}_{d,k,\ell+1} - \frac{1}{\sqrt{2\ell-1}}\mathcal{F}_{d,k,\ell-1}\right].
\end{aligned}$$

For $\ell = 0$ the same derivation holds using the formula for $\ell = 0$ stated in Property 3.1. □

Denote the coefficient matrix of $p_d(t)\Theta(t-s)$ by $B^{(d)} := \left[b^{(d)}_{k,\ell}\right]^{\infty}_{k,\ell=0}$, with $b^{(d)}_{k,\ell}$ as in Theorem 3.3. We will call such a matrix the *Legendre basis matrix of degree d*. Along a diagonal of $B^{(d)}$, the entries decay linearly; this is formally stated in Lemma 3.4.

LEMMA 3.4 (Decay of the Legendre basis coefficients). *For $b^{(d)}_{k,l}$ as in Theorem 3.3, it holds that*

$$\lim_{\substack{k,\ell\to\infty \\ |k-\ell|\ \text{constant}}} |b^{(d)}_{k,\ell}| \sim \mathcal{O}(1/\ell).$$

Proof. In the last equality in the formula in Property 3.2, the last fraction is constant since $\alpha = |b-c|$ is constant and $a = d$ is fixed. Then, it is straightforward to see that there is no decay in the expression of the integral over the triple product,

$$\lim_{\substack{k,\ell\to\infty \\ |k-\ell|\ \text{constant}}} \mathcal{F}_{d,k,\ell} \sim \mathcal{O}(1).$$

Since

$$\lim_{\substack{k,\ell\to\infty \\ |k-\ell| \text{ constant}}} \frac{1}{\sqrt{(2\ell+1)(2\ell-1)}} \sim \mathcal{O}(1/\ell),$$

the statement follows from (3.1). □

The coefficient matrix $F := [f_{k,\ell}]_{k,\ell=0}^{\infty}$ of $f(t,s) = \tilde{f}(t)\Theta(t-s) \in \mathcal{A}_\Theta$ can be written as $F = \sum_{d=0}^{\infty} \alpha_d B^{(d)}$, where $\{\alpha_d\}_{d\geq 0}$ are the Legendre coefficients of $\tilde{f}(t)$. Thus, we can relate properties of $B^{(d)}$ to properties of F. Namely, the fact that along a diagonal of F the entries decay linearly follows from Lemma 3.4 and is illustrated in Example 3.6.

COROLLARY 3.5 (Decay of the expansion coefficients). *Let $\alpha = |k-\ell|$ be constant as k, ℓ go to infinity. Then the coefficients $f_{k,\ell}$ of the Legendre expansion of $f(t,s) \in \mathcal{A}_\Theta$ decay asymptotically at the rate $\frac{1}{\ell}$.*

EXAMPLE 3.6. The polynomial of degree one $\tilde{f}(t) = -\imath\tau(t+1)$, with $\tau > 0$, can be written as a linear combination of $p_0(t)$ and $p_1(t)$, namely $-\imath\tau(t+1) = -2\imath\tau p_0(t) - \sqrt{\frac{2}{3}}\imath\tau p_1(t)$. Thus, its coefficient matrix is $F = -2\imath\tau B^{(0)} - \sqrt{\frac{2}{3}}\imath\tau B^{(1)}$, which is a pentadiagonal matrix. The order of magnitude of the entries of F for $\tau = 4$ are displayed in Figure 3.1. A linear decay is observed in this figure.

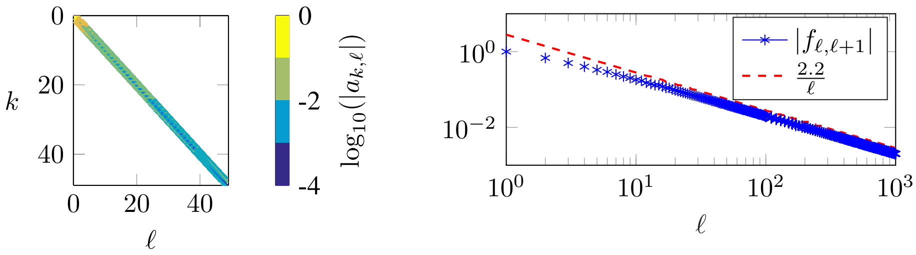


FIG. 3.1. *Left: the order of magnitude of the entries $f_{k,\ell}$ of F, the coefficient matrix of $f(t,s) = [-\imath 4(t+1)]\,\Theta(t-s)$. Right: the magnitude of the entries on the first superdiagonal $|f_{\ell,\ell+1}|$ together with the predicted decay rate $\mathcal{O}(\frac{1}{\ell})$.*

3.2. Banded coefficient matrix. A key property of the coefficient matrices of distributions in $\mathcal{A}_\Theta$ is that they are numerically banded. That is, they can be approximated by a banded matrix for any given threshold, e.g., machine precision. A matrix A is said to be an N-banded matrix or to have bandwidth N, if $a_{k,\ell} = 0$, for $|k-\ell| > N$. In this convention a diagonal matrix is a 0-banded matrix. The following corollary follows from Property 3.2 and Theorem 3.3.

COROLLARY 3.7 (Bandedness of the Legendre basis matrix $B^{(d)}$). *Consider the coefficient matrix $B^{(d)}$ of $p_d(t)\Theta(t-s)$, where $p_d(t)$ is the orthonormal Legendre polynomial of degree d. Then $B^{(d)}$ is a $(d+1)$-banded matrix, i.e.,*

$$b_{k,\ell}^{(d)} = 0, \qquad \textit{for } |k-\ell| > d+1.$$

We would like to truncate the infinite series $F := \sum_{d=0}^{\infty} \alpha_d B^{(d)}$ to a finite series $F^{(N)} := \sum_{d=0}^{N} \alpha_d B^{(d)}$, which is justified if $F^{(N)}$ is in some sense close to F. Closeness will be expressed in terms of the maximum norm $\|F - F^{(N)}\|_\infty$. In order to bound this quantity we state an upper bound for the maximum norm of the Legendre basis matrices $B^{(d)}$ in the next lemma.

LEMMA 3.8. *Consider the Legendre basis matrix $B^{(d)}$. Its maximum norm can be bounded by*

$$\|B^{(d)}\|_\infty \leq 3d + 2.$$

The proof of this lemma is technical and lengthy and is, therefore, postponed to Appendix A. We remark that we have observed, by numerical computations, that the infinity norm $\|B^{(d)}\|_\infty$ can be bounded by a constant. So the bound is too pessimistic. However, it is sufficient to prove the following result.

THEOREM 3.9. *Consider the coefficient matrix F of $\tilde{f}(t)\Theta(t-s) = f(t,s) \in \mathcal{A}_\Theta$. For any given tolerance $\delta_{tol} > 0$, the matrix F can be approximated by an $(N+1)$-banded matrix $F^{(N)}$ satisfying*

$$\|F - F^{(N)}\|_\infty \leq \delta_{tol}.$$

In other words, F is a numerically banded matrix.

Proof. Let $\tilde{f}(t) = \sum_{d=0}^{\infty} \alpha_d p_d(t)$ be the Legendre series of the function $\tilde{f}(t)$, analytic in the Bernstein ellipse $\mathcal{E}_\rho$, $\rho > 1$. Its coefficient matrix is $F = \sum_{d=0}^{\infty} \alpha_d B^{(d)}$. Using the bound (2.1) and Lemma 3.8, we have, for $F^{(N)} := \sum_{d=0}^{N} \alpha_d B^{(d)}$, that

$$\begin{aligned} \|F - F^{(N)}\|_\infty &= \left\| \sum_{d=N+1}^{\infty} \alpha_d B^{(d)} \right\|_\infty \leq \sum_{d=N+1}^{\infty} |\alpha_d| \|B^{(d)}\|_\infty \\ &\leq \sum_{d=N+1}^{\infty} C\rho^{-d-1}(3d+2). \end{aligned} \tag{3.2}$$

Therefore, there exists an N for which $\sum_{d=N+1}^{\infty} C\rho^{-d-1}(3d+2) \leq \delta_{\mathrm{tol}}$. This proves the statement. □

In the proof above, note that the truncated series

$$F^{(N)} = \sum_{d=0}^{N} \alpha_d B^{(d)}$$

defines an $(N+1)$-banded matrix sufficiently close to F for N large enough. The numerical bandedness of F and the bound in equation (3.2) for $F^{(N)}$ are illustrated in the following example.

EXAMPLE 3.10. Consider the function $\tilde{f}(t) = -\imath\omega \sin(\omega t)$, where ω controls the oscillation of the function. This function is not a polynomial, so we cannot expect a banded coefficient matrix; however, it is numerically banded. For $\omega = 1$, Figure 3.2 illustrates (left) the norm $\|F - F^{(N)}\|_\infty$ and the upper bound (3.2) for increasing N and (right) the order of magnitude of the entries of F, i.e., $\log_{10}(|f_{k,\ell}|)$. We see a clear numerical band structure of the matrix F and that the upper bound holds. With the given threshold chosen equal to machine precision $\delta_{\mathrm{tol}} = \epsilon_{\mathrm{mach}}$, the bandwidth is $N = 14$.

For $\omega = 5$, Figure 3.3 displays the same. We notice that $\|F - F^{(N)}\|_\infty$ reaches machine precision at $N = 24$, and again this corresponds to the numerical bandwidth of F. The bandwidth has increased compared to the less oscillatory function with $\omega = 1$ since now we require more Legendre coefficients to represent the function accurately.

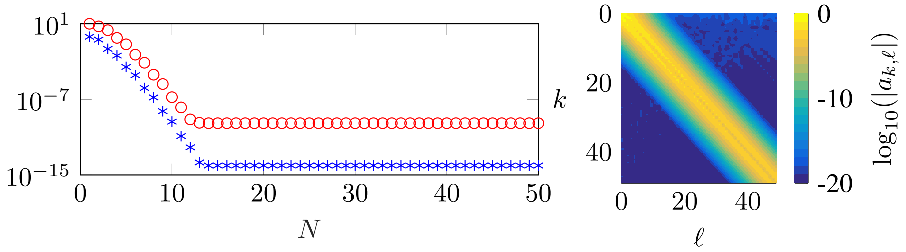


FIG. 3.2. *Coefficient matrix F of $[-\imath\omega\sin(\omega(t+1))]\,\Theta(t-s)$ with $\omega = 1$. Left: maximum norm $\|F - F^{(N)}\|_\infty$ ($*$) and the upper bound $\sum_{d=N+1}^{\infty} |\alpha_d|(3d+2)$ ($\circ$). Right: order of magnitude of the entries $f_{k,\ell}$.*

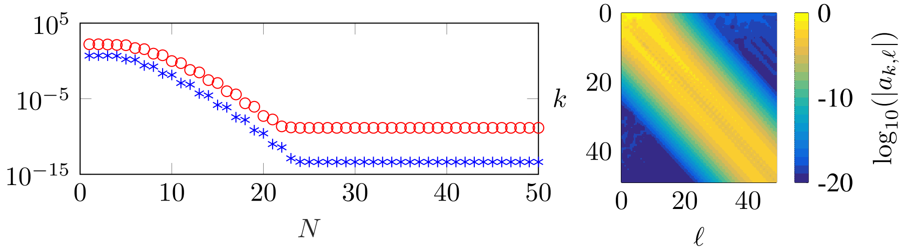


FIG. 3.3. *Coefficient matrix F of $[-\imath\omega\sin(\omega(t+1))]\,\Theta(t-s)$ with $\omega = 5$. Left: maximum norm $\|F - F^{(N)}\|_\infty$ ($*$) and the upper bound $\sum_{d=N+1}^{\infty} |\alpha_d|(3d+2)$ ($\circ$). Right: order of magnitude of the entries $f_{k,\ell}$.*

3.3. Well-defined matrix product. Given the distributions $f, g \in \mathcal{A}_\Theta$ and their coefficient matrices F, G, until now and in particular in Lemma 2.1, we have assumed that the matrix product FG is well defined. Thanks to the properties above, we can show that each element of FG exists, i.e., the matrix product is well defined.

Consider the Legendre series

$$\tilde{f}(t) = \sum_{d=0}^{\infty} \alpha_d p_d(t), \quad \tilde{g}(t) = \sum_{d=0}^{\infty} \beta_d p_d(t)$$

and the related matrix expansions

$$F = \sum_{d=0}^{\infty} \alpha_d B^{(d)}, \quad G = \sum_{d=0}^{\infty} \beta_d B^{(d)}.$$

Using Corollary 3.7 and Lemma 3.8, for $k, \ell = 0, 1, \ldots,$ we get bounds for the (k, ℓ)-entry of these coefficient matrices

$$|f_{k,\ell}| \leq \sum_{d=|k-\ell|-1}^{\infty} |\alpha_d|(3d+2), \qquad |g_{k,\ell}| \leq \sum_{d=|k-\ell|-1}^{\infty} |\beta_d|(3d+2)$$

(by convention, when $k = \ell$, d is set to start from 0). As recalled in equation (2.1), $|\alpha_d| \leq C_f \rho_f^{-d-1}$, with $C_f > 0$, $\rho_f > 1$. Therefore, there exists $K_f > 0$ such that

$$|f_{k,\ell}| \leq C_f \sum_{d=|k-\ell|-1}^{\infty} \frac{(3d+2)}{\rho_f^{d+1}} \leq C_f \rho_f^{-|k-\ell|} \sum_{d=0}^{\infty} \frac{(3d+3|k-\ell|-1)}{\rho_f^{d+2}} \leq K_f \rho_f^{-|k-\ell|}.$$

The same applies to $|g_{k,\ell}|$ for some $K_g > 0$ and $\rho_g > 1$. Note that this shows that F is characterized by an off-diagonal exponential decay. As a consequence, there exist constants $\rho > 1$ and $C > 0$ such that

$$|(FG)_{k,\ell}| = \left|\sum_{j=0}^{\infty} F_{k,j}G_{j,\ell}\right| \leq \sum_{j=0}^{\infty} |F_{k,j}||G_{j,\ell}| \leq \sum_{j=0}^{\infty} C\rho^{-(|j-k|+|j-\ell|)} < \infty,$$

proving that, for every $k, \ell = 0, 1, \dots,$ the product $(FG)_{k,\ell}$ is well defined. Hence, Lemma 2.1 holds for all the distributions in $\mathcal{A}_\Theta$. This allows us to state that there is a correspondence between the $\star$-product algebraic structure and the matrix (sub)algebra of Legendre coefficient matrices. This is summarized in Table 3.1.

TABLE 3.1

The $\star$-operations for distributions $f, g \in \mathcal{A}_\Theta$ and the associated matrix algebra operations for the respective coefficient matrices F, G.

$\star$-framework	matrix algebra
$f(t,s) \star g(t,s)$	FG
$f + g$	$F + G$
$1_\star = \delta(t-s)$	I
$R_\star(f)(t,s) = 1_\star + \sum_{k\geq 1} f^{\star k}(t,s)$	$(I-F)^{-1}$

4. Practical computation in the matrix algebra.

The second matrix problem in Problem 2.1 is to find the solution x to the infinite system of equations

$$(I - F)x = \phi(-1).$$

From Section 3.2 we know that we can accurately represent the infinite numerically banded coefficient matrix F by an infinite banded matrix $F^{(N)}$. In Section 4.1, we will discuss the existence of $(I - F^{(N)})^{-1}$. That is, the banded infinite system of equations

$$(I - F^{(N)})x = \phi(-1)$$

has a unique solution x. This solution can be arbitrarily close to the solution of the original system by choosing N appropriately, and therefore we make a slight abuse of notation by using the same variable x for both these solutions in favor of an easier notation.

To be able to use standard linear algebra techniques we would like to work with a finite system of equations instead of an infinite system. In Section 4.2, we discuss how an accurate approximation $\dot{x}$ to the first entries of x can be obtained by solving the finite system

$$(I_M - F_M^{(N)})\dot{x} = \phi_M(-1),$$

with $\phi_M(-1) = \begin{bmatrix} p_0(-1) & p_1(-1) & \dots & p_{M-1}(-1)\end{bmatrix}^\top$ and $F_M^{(N)}$ the $M \times M$ leading principal submatrix of $F^{(N)}$. In Section 4.3, we elaborate on choosing an appropriate value for M and how the Legendre coefficients can be obtained from $\dot{x}$. Section 4.4 explores further improvements to compute the solution $\dot{x}$ more efficiently based on exploiting hidden matrix structures.

4.1. Resolvent existence and decay phenomenon.

In this section, we deal with the existence of the resolvent $(I - F^{(N)})^{-1}$. Addressing this problem means discussing the invertibility of an operator. More precisely, consider the Hilbert space $\mathcal{H}$ with orthonormal basis $\{\dot{e}_0, \dot{e}_1, \dot{e}_2, \dots\}$. The coefficients $f_{k,\ell}$ of the (banded) matrix $F^{(N)}$ define the operator $\mathcal{R} : \mathcal{H} \to \mathcal{H}$ as follows:

$$\mathcal{R}\, \dot{e}_\ell = \sum_{k=\max\{\ell-N,0\}}^{\ell+N} f_{k,\ell}\, \dot{e}_k. \tag{4.1}$$

Denoting by $\mathcal{H}_M$ the linear span of $\{\dot{e}_0, \dots, \dot{e}_{M-1}\}$ and with $\mathcal{P}_M : \mathcal{H} \to \mathcal{H}_M$ the related orthogonal projection, we can define the finite-dimensional operator $\mathcal{R}_M = \mathcal{P}_M \mathcal{R} \mathcal{P}_M$. The operator $\mathcal{R}_M$ is then represented by the matrix $F_M^{(N)}$. Theorem 3.1 in [25] shows that the operator $\mathcal{R}$ is invertible under the following conditions:

1. $F^{(N)}$ is banded;
2. For every $M = 1, 2, \dots$ and for $j = 1, \dots, (N+1)$, there exist positive constants K_j, L_j, such that

$$\left\| \left(I_M - F_M^{(N)} \right)^{-1} e_{M-j} \right\|_2 \leq K_j, \qquad \left\| \left(I_M - \left(F_M^{(N)} \right)^H \right)^{-1} e_{M-j} \right\|_2 \leq L_j.$$

In the following, we demonstrate that Condition 2 is satisfied under certain assumptions on the *field of values* of the matrices $F_M^{(N)}$, i.e., the convex set in $\mathbb{C}$ defined as

$$W(F_M^{(N)}) := \left\{ v^H F_M^{(N)}\, v,\ \|v\|_2 = 1 \right\}.$$

Under these assumptions, we show that the matrix $(I_M - F_M^{(N)})^{-1}$ is characterized by the so-called *decay phenomenon* (e.g., [3, 4, 5]), i.e., the magnitude of its elements decay exponentially as we move away from the band of $F_M^{(N)}$.

LEMMA 4.1. *Let A be a matrix with bandwidth $N+1$ and so that $W(A)$ is contained in $D(0,r)$, a disk with radius $r < 1$ centered at the origin. Then,*

$$\left| (I - A)^{-1}_{k,\ell} \right| \leq C \mu^{d(k,\ell)}, \qquad d(k,\ell) := \frac{|k-\ell|}{(N+1)},$$

for every $r < \mu < 1$ and a constant C determined by μ.

The proof follows immediately from Theorem 2.3 in [32]; see also [5]. Assume that $W(F_M^{(N)}) \subset D(0,r)$ for a fixed $r < 1$ and for $M \geq M_0$ with M_0 large enough. Then, for every $\ell = 0, 1 \dots, M-1$, we get

$$\begin{aligned} \left\| \left(I_M - F_M^{(N)} \right)^{-1} e_\ell \right\|_2^2 &= \sum_{k=0}^{M} \left| \left(I_M - F_M^{(N)} \right)^{-1}_{k,\ell} \right|^2 \\ &\leq C^2 \sum_{k=0}^{M} \mu^{2d(k,\ell)} \leq C^2 \sum_{k=0}^{M} \tau^{|k-\ell|}, \quad \tau = \mu^{2/(N+1)} < 1, \\ &\leq C^2 \sum_{k=0}^{\infty} \tau^{|k-\ell|} =: K_\ell < \infty. \end{aligned}$$

Note that K_ℓ is independent of M, proving that Condition 2 above holds (a similar argument holds for the Hermitian transposed case). Therefore, by Theorem 3.1 in [25], we proved the following result:

THEOREM 4.2. *Assume that $W(F_M^{(N)}) \subset D(0,r)$, with $r<1$, for every $M > M_0$ with M_0 large enough. Then the operator $\mathcal{R}$ defined in (4.1) is invertible. Moreover, consider the operator equations $\mathcal{R}x = y$ and $\mathcal{R}_M x_M = \mathcal{P}_M y$. If y is in the range of $\mathcal{R}$, then $x_M \to x$ in the norm topology.*

As a final step, we need to determine conditions for the function $\tilde{f}(t)$ so that the coefficient matrix $F_M^{(N)}$ of $\tilde{f}(t)\Theta(t-s)$ satisfies $W(F_M^{(N)}) \subset D(0,r)$ for every M large enough. Since these matrices are usually characterized by a field of values with a disk shape, we estimate it by using the *numerical radius*, which is defined, for a given matrix A, as

$$\nu(A) := \sup\{|\lambda|, \lambda \in W(A)\}.$$

Note that $W(A) \subseteq D(0,\nu(A))$. Moreover, the numerical radius can also be expressed via the formula

$$\nu(A) \leq \max_k \left(\sum_\ell \frac{|a_{k,\ell}| + |a_{\ell,k}|}{2} \right);$$

see, e.g., [20, Corollary 5.2-3]. Unfortunately, obtaining bounds for the numerical radius has proved to be difficult. Therefore, for the moment, we rely on arguments based on numerical observations.

First, consider $B^{(d)}$, the Legendre basis matrix of degree d, and the related truncated matrix $B_M^{(d)}$. For $M = 2000$ and $d = 0, \dots, 500$, we observed numerically that

$$\max_k \left(\sum_\ell \frac{|(B_{2000}^{(d)})_{k,\ell}| + |(B_{2000}^{(d)})_{\ell,k}|}{2} \right) \leq 0.87.$$

Moreover, for all the tested matrices, the maximum was obtained for $k=0$, i.e., the first row. As the magnitude of the elements of $B_M^{(d)}$ tends to decay along the diagonal (see Lemma 3.4), we can bound the numerical radius by

$$\nu(B^{(d)}) \leq 0.87, \quad d = 0, \dots, 500.$$

Finally, by the Legendre expansion of $\tilde{f}(t) = \sum_{d=0}^\infty \alpha_d p_d(t)$ we get the bound

$$\nu(F^{(N)}) \leq \sum_{d=0}^N |\alpha_d| \nu(B^{(d)}).$$

For the reasons given above, we conjecture that as long as

$$\sum_{d=0}^N |\alpha_d| \leq 1.1494, \tag{4.2}$$

we get the inclusion:

$$W(F_M^{(N)}) \subseteq D(0, 0.87), \qquad M = 0, 1, \dots$$

Nevertheless, we have often observed that the solution x_M still converges even when $\nu(F_M^{(N)}) > 1$. The condition in (4.2) is very restrictive; it gives the impression that the techniques presented in this paper are applicable only to slowly oscillating, low amplitude functions $\tilde{f}(t)$. However, in numerical experiments (Section 5) we observe that even when the sum of coefficients is orders of magnitude larger than 1.1494 or $\nu(F_M^{(N)}) > 1$, the matrix $(I - F_M^{(N)})$ is still invertible, and its inverse still has an off-diagonal decay. Hence, both conditions are nondescriptive in our context; they are too pessimistic. More descriptive conditions might be obtained by looking at the pseudospectrum of $F^{(N)}$. Exploring this path is part of ongoing research.

4.2. Truncation error. The finite banded system is obtained by taking the $M \times M$ leading principal submatrix of the matrix $(I - F^{(N)})$:

$$(I_M - F_M^{(N)})\dot{x} = \phi_M(-1).$$

We analyze the error $|\dot{x} - x_M|$, where x_M denotes the vector containing the first M entries of the infinite solution x.

First, we derive an analytical expression for x_M in terms of submatrices of the infinite matrix $(I - F^{(N)})$. Set, for a more compact notation, $A := (I_M - F_M^{(N)}) \in \mathbb{C}^{M\times M}$ and the matrices B, C, D as in Figure 4.1. In this figure, the colored region contains generic nonzeros, white indicates zeros, and arrows are used to emphasize the size of a block or region. Note that $N + 2$ is equal to the bandwidth of the matrix plus one.

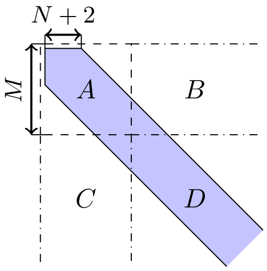


FIG. 4.1. *Block subdivision of an infinite banded matrix $(I - F^{(N)})$. The colored region shows the band of the matrix. The left upper block is $A := (I_M - F_M^{(N)}) \in \mathbb{C}^{M\times M}$; the dimensions of the other blocks follow immediately from this. Open lines indicate that the row and/or column index goes to infinity.*

This block subdivision allows us to rewrite the infinite system of equations as

$$(I - F^{(N)})x = \phi(-1) \Leftrightarrow \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x_M \\ z \end{bmatrix} = \begin{bmatrix} \dot{y} \\ v \end{bmatrix},$$

with $x_M \in \mathbb{C}^M$, $\dot{y} = \phi_M(-1) \in \mathbb{C}^M$, and $v = \begin{bmatrix} p_M(-1) & p_{M+1}(-1) & \dots \end{bmatrix}^\top$. Assume that A and D are invertible. Then the first M entries of the solution are

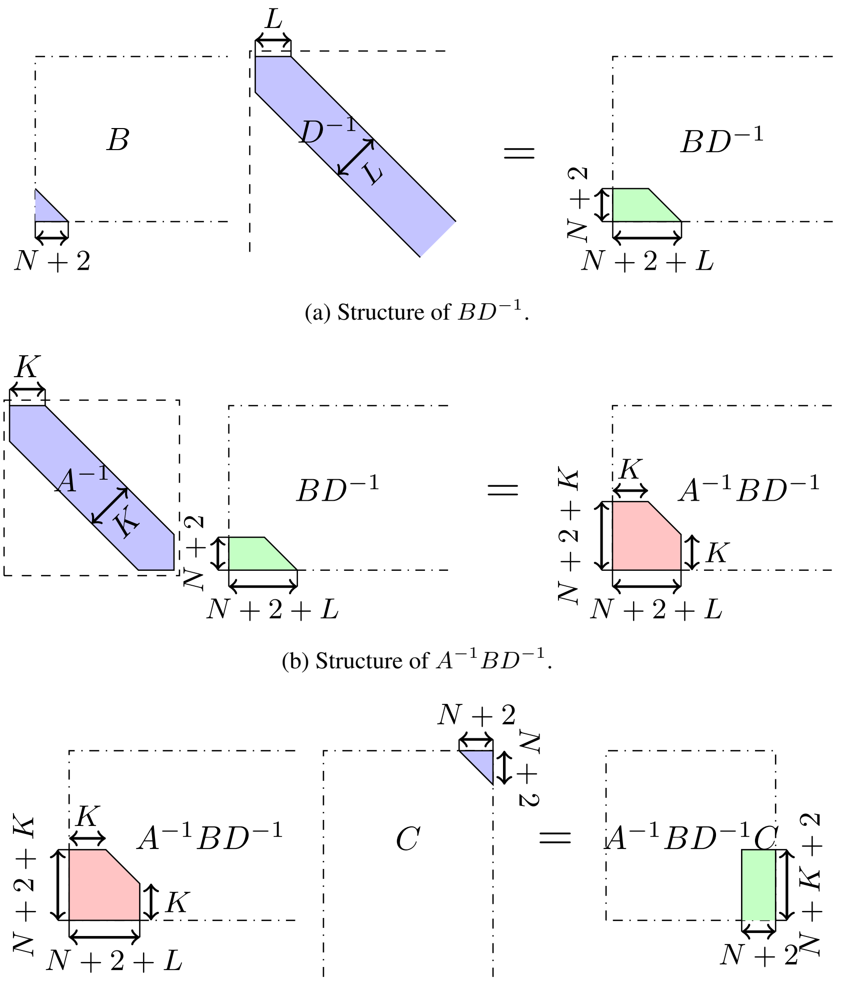
$$x_M = (I_M - A^{-1}BD^{-1}C)^{-1}A^{-1}\dot{y} - (I_M - A^{-1}BD^{-1}C)^{-1}A^{-1}BD^{-1}v.$$

The object under study is the error $|\dot{x} - x_M|$, where $\dot{x} = A^{-1}\dot{y}$, which is given by

$$\left| \left[(I_M - A^{-1}BD^{-1}C)^{-1} - I_M \right] \dot{x} - (I_M - A^{-1}BD^{-1}C)^{-1}A^{-1}BD^{-1}v \right|. \tag{4.3}$$

To study this error we look at the matrix structures of the matrices appearing in equation (4.3). Assume that $A^{-1} = (I_M - F_M^{(N)})^{-1}$ is a numerically banded matrix; see Section 4.1. In the following, matrix entries with magnitude below a given threshold are truncated. As a consequence, the matrix A^{-1} is a K-banded matrix. Since $W(D) \subseteq W(I - F^{(N)})$, if $F^{(N)}$ shows a decay, then, by Lemma 4.1, D^{-1} also shows a decay and can be approximated accurately by an L-banded matrix. The values K and L can be estimated a priori by using spectral information of $F^{(N)}$ using, e.g., Lemma 4.1. In Figure 4.2, the structure of the matrix products $A^{-1}BD^{-1}$ and $A^{-1}BD^{-1}C$ appearing in (4.3) is derived.

The structure of the error is now easily determined from the structure of these matrices. Namely, plug the matrices into the formula $|x_M - \dot{x}|$, and we obtain that (at least) the first



(a) Structure of BD^{-1}.

(b) Structure of $A^{-1}BD^{-1}$.

(c) Structure of $A^{-1}BD^{-1}C$.

FIG. 4.2. *Structure of the matrices appearing in the error analysis of $\dot{x}$. Colored regions indicate generic nonzeros and dashed lines indicate the boundary of the matrix; the lack of a dashed line indicates that the row and/or column index goes to infinity.*

$M-N-K-2$ entries are computed accurately; see Figure 4.3. In practice, thanks to the decay phenomenon, more than $M-N-K-2$ entries might be computed accurately. This is illustrated in Example 4.3.

EXAMPLE 4.3. Consider $\tilde{f}(t) = -\imath \sin(t+1)$. For $M = 50$, Figure 4.4 displays the corresponding matrix $(I_M - F_M)$, its inverse, and the 50×50 leading principal submatrix of D^{-1}. All these matrices are numerically banded, and for $\delta_{\text{tol}} = \epsilon_{\text{mach}}$ they have the bandwidth $N+2 = 16$, $K = 22$, and $L = 16$, respectively.

From the above error analysis, we expect that the first $M-N-K-2$ entries of $\dot{x}$ are close to those of the exact (infinite) solution x. In our example, this would be a number of $50-14-22-2 = 12$ entries. The predicted structure in Figure 4.3 is verified by computing these matrices numerically. This is illustrated in Figure 4.5, where we have chosen to set elements smaller than 10^{-19} to zero (i.e., the color white in the colorbar). The observed

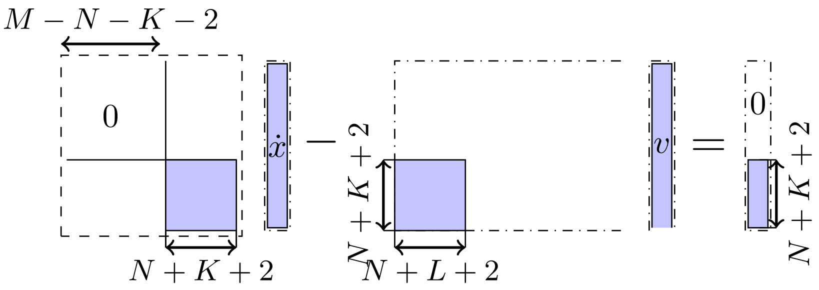


FIG. 4.3. *Structure of the truncation error*
$$|x_M - \dot{x}| = |\left[(I - A^{-1}BD^{-1}C)^{-1} - I\right]\dot{x} - (I - A^{-1}BD^{-1}C)^{-1}A^{-1}BD^{-1}v|.$$
Colored regions indicate generic nonzeros, and dashed lines indicate the boundary of the matrix; the lack of a dashed line indicates that the row and/or column index goes to infinity.

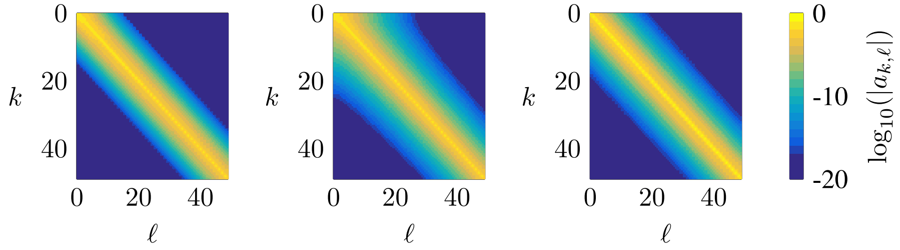


FIG. 4.4. *Order of magnitude of the entries of the matrices* $(I_M - F_M)$, $(I_M - F_M)^{-1}$, *and* D^{-1}, *respectively, for the function* $\tilde{f}(t) = -\imath\sin(t+1)$ *and* $M = 50$. *On the far right the colorbar indicates the colors corresponding to the orders of magnitude.*

matrices adhere to the predicted structure, but thanks to the decay of the entries it does not fill the whole predicted submatrix with large elements. As a consequence, more than the predicted 12 entries of $\dot{x}$ are computed accurately; we observe that 30 entries are computed up to machine precision.

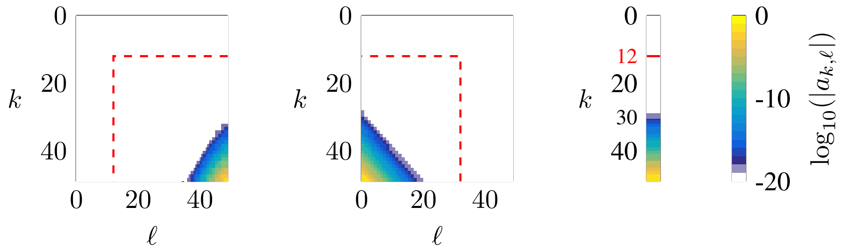


FIG. 4.5. *Order of magnitude of the entries of the matrices* $(I - A^{-1}BD^{-1}C)^{-1} - I$ *and* $(I - A^{-1}BD^{-1})^{-1}A^{-1}BD^{-1}$ *and of the error vector* $|x_M - \dot{x}|$ *for the function* $\tilde{f}(t) = -\imath\sin(t+1)$, *from left to right, respectively. On the far right the colorbar indicates the colors corresponding to the orders of magnitude. Dashed red lines indicate the predicted structure as shown in Figure 4.3.*

The above example numerically validates the error analysis in this section. Thus, the truncation to the leading principal submatrix is both theoretically and numerically justified. We would like to stress that the inverse of $A = (I_M - F_M)$ is not computed explicitly in our proposed procedure, but instead the system of equations $(I_M - F_M^{(N)})\dot{x} = \phi_M(-1)$ is solved to obtain $\dot{x}$. Once $\dot{x}$ is available, the approximate Legendre coefficients are obtained as $\dot{c} = T_M\dot{x}$. Since T_M is a tridiagonal matrix, it follows immediately that at least $M-N-K-3$ coefficients of $\dot{c}$ are computed accurately.

4.3. Finding the accurate Legendre coefficients. From the analysis in the above section two questions arise:

1. What is an optimal choice for M, large enough to accurately compute a sufficient amount of Legendre coefficients and as small as possible, to reduce computational costs?
2. After computing the Legendre coefficients $\dot{c}$, how many should we keep?

For the first question we require some information about the coefficient matrix of the solution. The solution $\tilde{u}(t)$ is analytic inside and on some Bernstein ellipse $\mathcal{E}_\rho$, therefore, from [44], we know that its Legendre coefficients $\{c_k\}_{k\geq 0}$ satisfy

$$|c_k| \leq \frac{(2k+1)\ell(\mathcal{E}_\rho)M_\rho}{\pi\rho^{k+1}(1-\rho^{-2})}, \tag{4.4}$$

where $M_\rho = \max\limits_{z\in\mathcal{E}_\rho}|\tilde{u}(z)|$ and $\ell(\mathcal{E}_\rho)$ is the circumference of $\mathcal{E}_\rho$. Thus, by Theorem 3.9 it follows that the coefficient matrix U of $\tilde{u}(t)\Theta(t-s)$ is a numerically banded matrix.

Let K denote the integer such that $\|U - U^{(K)}\|_\infty \leq \delta_{\text{sol}}$, i.e., U is a numerically $(K+1)$-banded matrix for the requested tolerance. By the truncation error analysis from Section 4.2, it then follows that a Legendre basis of size $M = 2K + N + 4$, with N the numerical bandwidth of F, suffices to compute the $K+1$ first Legendre coefficients of $\tilde{u}(t)$ up to the requested accuracy. If it is known in which Bernstein ellipses the solution $\tilde{u}(t)$ is analytic, then an estimate of K can be obtained. In the numerical examples in Section 5, $f(t)$ are entire functions and so are the corresponding solutions $u(t)$, thus the bound (4.4) holds for all $\rho > 1$. In order to obtain an estimate, we will replace the ellipse $\mathcal{E}_\rho$ in the bound by the circle $\rho e^{i\theta}$. Set $\widehat{M}_\rho := \exp\left(\rho \max\limits_{\theta\in[0,2\pi]}|f(\rho e^{i\theta})|\right)$. Then,

$$|c_k| \leq \frac{(2k+1)\ell(\mathcal{E}_\rho)M_\rho}{\pi\rho^{k+1}(1-\rho^{-2})} \leq \frac{(2k+1)2\pi\rho\widehat{M}_\rho}{\pi\rho^{k+1}(1-\rho^{-2})} \leq \frac{2(2k+1)e^\rho}{\rho^k(1-\rho^{-2})} = \frac{2(2k+1)e^\rho}{1-\rho^{-2}}\rho^{-k}.$$

Using this bound we get, for all $\rho > 1$,

$$\|U - U^{(K)}\|_\infty \leq \sum_{d=K+1}^{\infty}|c_d|\|B^{(d)}\|_\infty \leq \sum_{d=K+1}^{\infty}\frac{2(3d+2)(2d+1)e^\rho}{1-\rho^{-2}}\rho^{-d}.$$

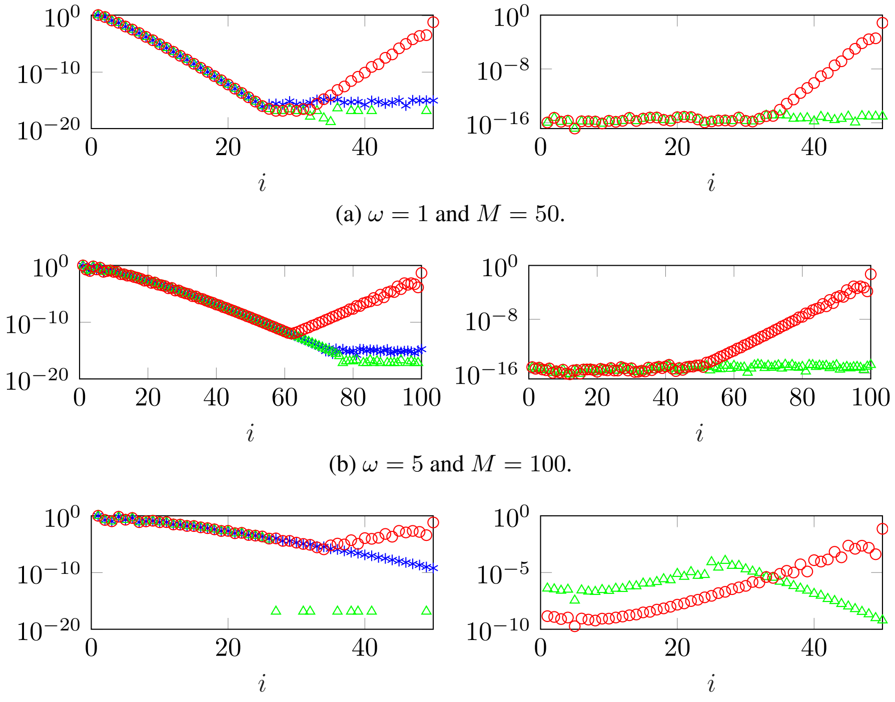
Thus, K can be chosen such that $\sum_{d=K+1}^{\infty}\frac{2(3d+2)(2d+1)e^\rho}{1-\rho^{-2}}\rho^{-d} \leq \delta_{\text{sol}}$, and choosing $M = 2K + N + 4$ guarantees that we can approximate $\tilde{u}(t)$ up to an error of $\mathcal{O}(\delta_{\text{sol}})$. Note that this choice for M is an overestimate of the optimal choice. Since the Legendre coefficients of the solution are computed directly and provide an estimate of the error, they can be used as an automatic way to change M [2].

The second question is answered by a particular truncation combined with a numerical method that chooses the right amount of coefficients automatically. We describe the truncation using an example. Consider the function $\tilde{f}(t) = -\imath\omega\sin(\omega(t+1))$, for which we approximate the solution to the ODE (1.1), that is, $\tilde{u}(t) = \exp\left(-\imath(1-\cos(\omega t+\omega))\right)$. For $\omega = 1$, this function is the one considered in Example 4.3, from which we know that for $M = 50$, computing about 30 Legendre coefficients up to machine precision suffices to represent $\tilde{u}(t)$. In Figure 4.6, we display the Legendre coefficients obtained in the following two ways:

1. Solve $(I - F_M^{(N)})\dot{x} = \phi_M(-1)$ and compute $\dot{c} = T_M\dot{x}$.
2. Solve $(I - \underline{F}_M^{(N)})\underline{\dot{x}} = \phi_M(-1)$ and compute $\underline{\dot{c}} = \underline{T}_M\underline{\dot{x}}$, where the entries of the coefficient matrices are set to zero in the last $N+1$ rows and where $N+1$ is the bandwidth of the matrix. This means that the first $M-N-1$ rows of $\underline{F}_M^{(N)}$ equal those of $F_M^{(N)}$, and the last $N+1$ rows are all zeros. Since T_M has bandwidth equal to one, omitting its last row gives us $\underline{T}_M$.

For $\omega = 1$ and $M = 50$, the Legendre coefficients $\dot{c}$ without truncation lead to the first 30 coefficients being accurately computed, and after this the coefficients increase in amplitude, resulting in a *u*-shaped curve for the coefficients. This *u*-shape does not correspond to the actual Legendre coefficients, and it makes it more difficult to determine how many coefficients one would keep to obtain an accurate approximation of $\tilde{u}(t)$, because using more than 30 coefficients will cause the approximation to deteriorate.

The Legendre coefficients $\underline{\dot{c}}$ with truncation of the coefficient matrices is equally accurate as $\dot{c}$ for the first 30 coefficients and does not have an increase after this; the truncation, in fact, pushes these last coefficients to zero. Thus, $\underline{\dot{c}}$ is easier to use since it has a simpler shape allowing a chopping of the series by, e.g., the procedure in [2]. Moreover, using more coefficients than 30 will not deteriorate the approximation of $\tilde{u}(t)$.



(a) $\omega = 1$ and $M = 50$.

(b) $\omega = 5$ and $M = 100$.

(c) $\omega = 5$ and $M = 50$.

FIG. 4.6. *Legendre coefficients obtained by a system solve for the coefficient matrix* $F_M^{(N)}$ *for* $\tilde{f}(t) = -\imath\omega\sin(\omega(t+1))$. *Left: Legendre coefficients, exact* c ($*$), *approximation* $\dot{c}$ ($\circ$), *approximation with truncation* $\underline{\dot{c}}$ ($\triangle$). *Right: the error of the computed coefficients,* $|c - \dot{c}|$ ($\circ$) *and* $|c - \underline{\dot{c}}|$ ($\triangle$).

In Figure 4.6, we also present the Legendre coefficients and their error for two other choices of parameters. For $\omega = 5$ and $M = 100$, we observe a similar behavior as for $\omega = 1$. For $\omega = 5$ and $M = 50$, M is chosen smaller than the minimal required size to represent $\tilde{u}(t)$ up to machine precision. Now the first coefficients in $\underline{\dot{c}}$ are computed less accurately than $\dot{c}$ due to the truncation. This might be because the last $N = 24$ rows of the coefficient matrix $F_M^{(N)}$ of size $M = 50$ are truncated, thereby discarding too many entries.

4.4. Fast computation of Fourier coefficients. The first subproblem in Problem 2.1 is to construct the coefficient matrix F. We have shown that in fact it suffices to approximate F by the finite banded matrix

$$F_M^{(N)} = \sum_{d=0}^{N} \alpha_d B_M^{(d)}.$$

The efficient construction of the coefficient matrix thus requires the efficient computation of $\{\alpha_d\}_{d\geq 0}$ and $\{B_M^{(d)}\}_{d\geq 0}$. Since $B_M^{(d)}$ are the coefficient matrices of the basis, they must be computed only once for each d and can then be reused for different functions.

First, an efficient algorithm to approximate the Legendre coefficients $\{\alpha_d\}_{d=0}^N$ of $\tilde{f}(t)$ is discussed. We use functions that are available in the MATLAB package `chebfun` [12] . The Legendre coefficients for the smooth function $\tilde{f}(t)$ are given by $\alpha_d = \int_{-1}^{1} \tilde{f}(t) p_d(t) dt$ and will be approximated by $\{\hat{\alpha}_d\}_{d=0}^N$ as follows:

1. Using `chebfun`, we compute the coefficients of the interpolating Chebyshev series $\sum_{k=0}^{N} \hat{c}_d T_d(t) \approx \tilde{f}(t)$. Given a required accuracy, an appropriate truncation value N is chosen automatically [2]. The coefficients $\{\hat{c}_d\}_{d=0}^N$ are obtained with a complexity $\mathcal{O}(N \log(N))$. For details on the error incurred by interpolation instead of by computing the integral $c_d = \int_{-1}^{1} \frac{\tilde{f}(t) T_d(t)}{\sqrt{1-t^2}} dt$, we refer to the book by Trefethen [42, Chapter 4].
2. The Chebyshev coefficients $\{\hat{c}_d\}_{d=0}^N$ can be transformed into Legendre coefficients $\{\hat{\alpha}_d\}_{d=0}^N$ with a complexity $\mathcal{O}(N \log^2(N))$ by using the method proposed by Townsend et al. [41]. In `chebfun`, this method is available under the name `cheb2leg`. This transformation from Chebyshev to Legendre coefficients is expected to have a worst-case error growth of $\mathcal{O}(\sqrt{N} \log(N))$, and for a fast decaying set of coefficients $\{\hat{c}_d\}_{d=0}^N$, Townsend and collaborators have observed numerically that there is no error growth with N.

Thus, at an overall complexity of $\mathcal{O}(N \log^2(N))$ we are able to compute the coefficients $\{\hat{\alpha}_d\}_{d=0}^N$ representing $\tilde{f}(t)$ in the Legendre basis. The coefficient matrix $F_M^{(N)} = \sum_{d=0}^N \alpha_d B_M^{(d)}$ can be accurately approximated by $\hat{F}_M^{(N)} = \sum_{d=0}^N \hat{\alpha}_d B_M^{(d)}$.

Second, we want to compute and store the Legendre basis matrices $B_M^{(d)}$ efficiently. Equation (3.1) provides an expression for the entries $b_{k,\ell}^{(d)}$ in terms of integrals of the triple product of Legendre polynomials $\mathcal{F}_{a,b,c}$. Our approach is based on expressing the matrix $[\mathcal{F}_{a,b,c}]_{b,c=0}^{M-1}$ as

$$[\mathcal{F}_{a,b,c}]_{b,c=0}^{M-1} = \sqrt{2a+1}\,(C_M \circ H_M \circ T_M)\,,$$

with $C_M = \left[\sqrt{(2b+1)}\sqrt{(2c+1)}\right]_{b,c=0}^{M-1}$, H_M a Hankel matrix, T_M a Toeplitz matrix, and $\circ$ denotes the Hadamard product. The Hankel matrix depends on the value $\gamma := b + c$ and is characterized completely by its last row $r_H^\top$ and first column c_H. Let us define a function that generates the entries in H_M,

$$h(a,\gamma) := \frac{1}{(a+\gamma+1)} \left(\prod_{j=1}^{a} \frac{-a+\gamma+2j}{-d+\gamma+2j-1} \right).$$

Then, for $(m+a)$ even, the last row and first column are

$$\begin{aligned} r_H^\top &= \begin{bmatrix} h(a,m) & 0 & h(a,m+2) & 0 & \dots & h(a,2m-2) & 0 & h(a,2m) \end{bmatrix}^\top, \\ c_H &= \begin{bmatrix} \underbrace{0 \ \cdots \ 0}_{a} & h(a,a) & 0 & h(a,a+2) & 0 & \dots & h(a,m-2) & 0 & h(a,m) \end{bmatrix}^\top, \end{aligned}$$

and, for $(m+a)$ odd,

$$\begin{aligned} r_H^\top &= \begin{bmatrix} h(a,m) & 0 & h(a,m+2) & 0 & \dots & 0 & h(a,2m-1) & 0 \end{bmatrix}^\top, \\ c_H &= \begin{bmatrix} \underbrace{0 \ \cdots \ 0}_{a} & h(a,a) & 0 & h(a,a+2) & 0 & \dots & 0 & h(a,m-1) & 0 \end{bmatrix}^\top. \end{aligned}$$

The Toeplitz matrix T_M depends on $\alpha := |b-c|$, and since it is symmetric, it is characterized by its first column c_T. Using the following function generating the entries of T_M,

$$t(a,\alpha) := \frac{1}{2^{(2a+1/2)}} \frac{\prod_{j=(\frac{a+\alpha}{2}+1)}^{a+\alpha} j^2}{\prod_{j=1}^{\frac{a-\alpha}{2}} j^2 \prod_{j=(a-\alpha+1)}^{a+\alpha} j},$$

the first column of the Toeplitz matrix is given, for a odd, by

$$c_T = \begin{bmatrix} t(a,0) & 0 & t(a,2) & 0 & \dots & t(a,a) & \underbrace{0 \ \cdots \ 0}_{m-a} \end{bmatrix}^\top$$

and, for a even, by

$$c_T = \begin{bmatrix} 0 & t(a,1) & 0 & t(a,3) & \dots & 0 & t(a,a) & \underbrace{0 \ \cdots \ 0}_{m-a} \end{bmatrix}^\top.$$

Using this expression, the matrices $[\mathcal{F}_{d,b,c}]_{b,c=0}^{M-1}$, for $d = 0,1,\dots,N$, can be stored by $3(N+1)M + M^2$ numbers instead of $(N+1)M^2$ if each matrix is stored naively. Further reduction of memory cost can be obtained by exploiting the zero structure of the Hankel and Toeplitz matrix.

The Legendre basis matrix of degree d can now be written as

$$B_M^{(d)} = \left[b_{k,l}^{(d)} \right]_{k,l=0}^{M-1} = \sqrt{2d+1} \left(\tilde{C}_M \circ ((\underline{H}_M \circ \underline{T}_M) Z_M) \right), \tag{4.5}$$

where $\tilde{C}_M := \left[\frac{\sqrt{2k+1}}{\sqrt{2\ell+1}} \right]_{k,\ell=0}^{M-1} \in \mathbb{R}^{M \times M}$, $\underline{H}_M$ and $\underline{T}_M$ are, respectively, H_{M+1} and T_{M+1} with the last row removed, and

$$Z_M := \begin{bmatrix} 1 & -1 & & & & \\ 1 & 0 & -1 & & & \\ & 1 & 0 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & 0 & -1 \\ & & & & 1 & 0 \\ & & & & & 1 \end{bmatrix} \in \mathbb{R}^{(M+1) \times M}.$$

The representation of the Legendre basis matrices $B_M^{(d)}$ given by equation (4.5) creates possibilities for the development of efficient methods for storing and computing the coefficient matrix $\hat{F}_M^{(N)}$; see, for example, the procedure proposed in [41] for the positive semidefinite case. Further exploring and implementing efficient memory and computational schemes is subject of ongoing research and is out of the scope of this paper.

5. The proposed numerical procedure. Using the presented results, we can replace the infinite matrix problem formulated in Problem 2.1 by the following finite matrix problem:

PROBLEM 5.1 (Finite matrix problem). *Given a smooth function $\tilde{f}(t)$ and a tolerance δ_{sol}, compute the approximate Legendre coefficients $\{\hat{c}_k\}_{k=0}^{\hat{M}}$ of the solution $\tilde{u}(t)$ to the ODE (1.1) such that they satisfy $\|\tilde{u}(t) - \sum_{d=0}^{\hat{M}} \hat{c}_d p_d(t)\|_\infty \lesssim \delta_{sol}$ on the interval $t \in [-1, 1]$. This corresponds to solving five subproblems:*

1. *Compute the interpolating Legendre coefficients $\{\hat{\alpha}_k\}_{k=0}^{N}$ of $\tilde{f}(t)$ for an appropriate value of N.*
2. *Determine an appropriate value for the size M of the coefficient matrix F such that enough Legendre coefficients, $\hat{M}$, are computed accurately in order to reach the given tolerance; see Section 4.2.*
3. *Construct the finite banded coefficient matrix $\hat{F}_M^{(N)} = \sum_{d=0}^{N} \hat{\alpha}_d B_M^{(d)}$.*
4. *Solve the finite linear system of equations $(I_M - \hat{F}_M^{(N)})\dot{x} = \phi_M(-1)$ for $\dot{x}$. The right-hand side is the column vector $\phi_M(-1) = \big[p_k(-1)\big]_{k=0}^{M-1}$, and I_M is the identity matrix.*
5. *Compute the finite matrix vector product $T_M\dot{x} = \hat{c}$, thereby obtaining the approximate Legendre coefficients $\{\hat{c}_k\}_{k=0}^{M}$.*

The procedure developed in this paper proposes to solve the subproblems in Problem 5.1 as follows:

1. Using `chebfun`, the coefficients $\{\hat{\alpha}_k\}_{k=0}^{N}$ are computed with a complexity of $\mathcal{O}(N\log^2(N))$ as described in Section 4.4. Moreover, for a given tolerance, the value for N is chosen automatically.
2. Section 4.3 gives sufficient condition for choosing a large enough truncation parameter M (which might be an overestimation).
3. Compute the sum $\hat{F}_M^{(N)} = \sum_{d=0}^{N} \hat{\alpha}_d B_M^{(d)}$. An analytical formula for the entries of $B_M^{(d)}$ is stated in Property 3.2. Equation (4.5) provides a formulation for the construction of $B_M^{(d)}$ that is more memory and computationally efficient.
4. Solve the finite system of equations $(I_M - \underline{\hat{F}}_M^{(N)})\underline{\dot{x}} = \phi_M(-1)$, where $\underline{F}_M^{(N)}$ equals $F_M^{(N)}$ with its last $N+1$ rows set equal to zero. See Section 4.2 and Section 4.3 for details. The system is solved in MATLAB by the `backslash` function.
5. Form the product $\underline{T}_M\underline{\dot{x}} = \hat{c}$, where $\underline{T}_M$ equals T_M with its last row set to zero. If requested, the Legendre series can be chopped by applying the procedure proposed by Aurentz and Trefethen [2]; see also Section 4.3.

A MATLAB code implementing this procedure is freely available on the website https://github.com/nielvb/starLegendre. In the following, we present numerical experiments which will confirm the validity of this procedure. The invertibility of $(I_M - F_M^{(N)})$ is studied by its spectral properties. We report the numerical radius $\nu(F_M^{(N)})$ and the upper bound (4.2). However, these quantities are not descriptive of the observed numerical behavior, therefore we also report the pseudospectra of $(I_M - F_M^{(N)})$ which might provide a better description. We denote by $\sigma(A)$ the spectrum of A. Then, for $\epsilon > 0$, the

ϵ-pseudospectrum of $(I_M - F_M^{(N)})$ is defined by

$$\sigma_\epsilon(I_M - F_M^{(N)}) = \left\{ z \in \mathbb{C} | z \in \sigma(I_M - F_M^{(N)} + E) \text{ for some } E \text{ with } \|E\| \le \epsilon \right\}.$$

See, for example, the book by Trefethen and Embree [43] for details.

Since for the scalar ODEs the analytical solution $\tilde{u}(t)$ is available, we can compute the error of our approximation $\hat{u}(t) := \sum_{d=0}^{M} \hat{c}_d p_d(t)$ in the maximum norm

$$\text{err}_\text{f} := \|\tilde{u}(t) - \hat{u}(t)\|_\infty.$$

This error is estimated by evaluating the solution and the approximation in $10M$ equidistant nodes in $t \in [-1, 1]$. Denote by c the vector of the first M Legendre coefficients of $\tilde{u}(t)$. Then the error in the computed Legendre coefficients $\hat{c}$ is quantified by

$$\text{err}_c = \frac{|c - \hat{c}|}{\|c\|_\infty}.$$

Timings are not performed since in this paper we focus on the validity and accuracy of the procedure. For the scalar case studied here we do not expect to outperform state-of-the-art methods. However, thanks to the fact that (1.4) is valid for scalar as well as matrix-valued functions, a similar discretization technique to the one presented here can be used for the matrix case. Proving that the discretization via Legendre polynomials is well defined and the analysis of the truncation error is more challenging for the matrix case; ongoing research aims to develop a similar procedure as the one presented in this paper for the matrix ODE case, which is competitive with the state-of-the-art methods. Understanding the applicability and accuracy of the scalar case is a fundamental step towards developing a competitive procedure for the matrix case.

Because SHBVMs are spectral methods that also use a Legendre basis, we expect that for a similar size M of the basis, both SHBVMs and our proposed one obtain a similar accuracy. For an ODE solved on the time domain $[0, t_\text{end}]$, we will compute an approximation using the (single-step) SHBVM [1] and an approximation using our numerical approach for the same size M of the basis. The obtained approximations are compared to the exact solution at the time point $t = t_\text{end}$.[2] We remark that for the scalar examples below, neither of these methods are the method of choice. The comparison we make only serves to validate our proposed procedure by comparing to a known method.

5.1. Toy problem. The following function is constructed so that we have control over its behavior:

$$\tilde{f}(t) = -\imath \frac{\omega}{\beta} \sin(\omega(t+1)).$$

The parameter ω controls the oscillation of the function, and β controls its amplitude. The solution to the ODE

$$\frac{d}{dt}\tilde{u}(t) = -\imath \frac{\omega}{\beta} \sin(\omega(t+1))\tilde{u}(t), \quad \tilde{u}(-1) = 1, \quad \text{on } t \in [-1, 1]$$

is

$$\tilde{u}(t) = \exp\left(-\frac{\imath}{\beta}(1 - \cos(\omega t + \omega))\right).$$

[2] The code for SHBVM can be found on the webpage:
https://people.dimai.unifi.it/brugnano/LIMbook/software.html.

We report results for three different choices of parameters. The first choice is $\omega = 5$, $\beta = 10$, and we take $M = 100$. The condition (4.2) is satisfied, namely $\sum_{d=0}^{N} |\alpha_d| = 1.0909$, and the numerical radius is $\nu(F_{100}) = 0.2151$. Thus, the matrix $(I_M - F_M^{(N)})$ is nonsingular and subproblem 4 in Problem 5.1 has a unique solution. The approximation has accurate Legendre coefficients, $\max(\text{err}_c) = 1.7828 \cdot 10^{-15}$, and a function error of $\text{err}_{\text{f}} = 1.3345 \cdot 10^{-15}$.

The second choice is $\omega = 5$, $\beta = 1$, for $M = 100$. Condition (4.2) is not satisfied, $\sum_{d=0}^{N} |\alpha_d| = 10.909$, and the numerical radius of the coefficient matrix is $\nu(F_{100}^{(24)}) = 2.151$. Thus the existence of $(I_{100} - F_{100}^{(24)})^{-1}$ cannot be guaranteed by looking at these quantities. Nevertheless, it exists, and the approximation obtained is accurate, $\text{err}_{\text{f}} = 1.8621 \cdot 10^{-15}$ and $\max(\text{err}_c) = 2.5823 \cdot 10^{-15}$.

As a final choice we take a more oscillatory function $\omega = 100$ and $\beta = 1$ and compute an approximation for $M = 1500$. The numerical radius $\nu(F_{1500}^{(148)}) = 45.11$ is much larger than 1, and $\sum_{d=0}^{N} |\alpha_d| = 796.7$ does not satisfy the condition (4.2). The approximation is accurate, $\text{err}_{\text{f}} = 9.9812 \cdot 10^{-14}$ and $\max(\text{err}_c) = 3.6107 \cdot 10^{-14}$. The numerical radius for the second and third choice of parameters does not guarantee the existence of $(I_M - F_M^{(N)})^{-1}$. Therefore, in Figure 5.1, we show the pseudospectra for several levels for $(I_M - F_M^{(N)})$ for these choices with $M = 1500$. They indicate that even for perturbations E that are relatively large in the norm (up to 10^{-5}), the spectrum of $(I_{1500} - F_{1500}^{(N)} + E)$, for both functions, remains contained in a disk centered at $(1, 0)$ with a radius equal to one.

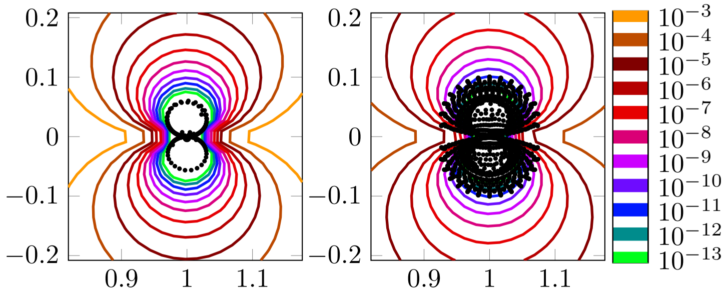


FIG. 5.1. *Spectrum* $(\cdot)$ *and pseudospectra of* $(I - F_M)$ *for the coefficient matrix* F_M *of the function* $\tilde{f}(t) = -\imath \frac{\nu}{\beta} \sin(\nu(t+1))$. *Left:* $\nu = 5$, $\beta = 1$, *and* $M = 1500$. *Right:* $\nu = 100$, $\beta = 1$, *and* $M = 1500$.

Using the SHBVM we compute an approximation $\hat{u}_{\text{HBVM}} \approx \tilde{u}(1)$, and our procedure provides Legendre coefficients for the function $\hat{u}(t)$, which can be evaluated in $t = 1$ such that $\hat{u}(1) \approx \tilde{u}(1)$. Table 5.1, Table 5.2, and Table 5.3 show the error $|\hat{u}_{\text{HBVM}} - \tilde{u}(1)|$ and $|\hat{u}(1) - \tilde{u}(1)|$ for the toy problem with different parameters. Table 5.1 and Table 5.2 show that for a fixed size of the basis, SHBVM is more accurate for lower oscillatory functions, and, as the function becomes more highly oscillatory, our method becomes more accurate; see Table 5.3. This numerical experiment verifies the validity of our proposed numerical method and shows that it is essentially different from SHBVMs.

5.2. A polynomial problem. Consider the following ODE which appeared, e.g., in [39],

$$\frac{d}{d\tau}\tilde{u}(\tau) = -\imath\tau\tilde{u}(\tau), \quad \tilde{u}(0) = 1, \quad \text{on } \tau \in [0, \tau_{\text{end}}]\,.$$

The function $\tilde{f}(\tau) = -\imath\tau$ is a degree-one polynomial, and the solution is $\tilde{u}(\tau) = \exp(-\imath\tau^2)$. Since the Legendre polynomials are defined on $[-1, 1]$, we perform a transformation of the

TABLE 5.1

Accuracy of the SHBVM and our method for approximating $\hat{u}(1)$ for the toy problem $\tilde{f}(t) = -\imath\frac{\omega}{\beta}\sin(\omega(t{+}1))$, with $\omega = 5$ and $\beta = 10$.

M	20	30	40	50	60	70
SHBVM	3.0e-06	3.5e-09	2.4e-12	8.1e-16	8.1e-17	1.8e-16
Our method	1.4e-02	9.0e-03	1.2e-05	3.3e-09	1.1e-12	7.2e-16

TABLE 5.2

Accuracy of the SHBVM and our method for approximating $\hat{u}(1)$ for the toy problem $\tilde{f}(t) = -\imath\frac{\omega}{\beta}\sin(\omega(t{+}1))$, with $\omega = 5$ and $\beta = 1$.

M	20	40	60	80	100
SHBVM	2.5e-02	2.9e-05	5.5e-09	2.9e-13	9.2e-16
Our method	3.3e-01	1.1e-02	9.2e-07	1.8e-11	5.4e-16

ODE, mapping $\tau \in [0, \tau_{\text{end}}]$ onto $t \in [-1, 1]$. This transformation is $t = \frac{2\tau}{\tau_{\text{end}}} - 1$ with its inverse $\tau = (t+1)\frac{\tau_{\text{end}}}{2}$, and this leads to

$$\frac{d}{dt}\tilde{u}(t) = -\imath\left(\frac{\tau_{\text{end}}}{2}\right)^2 (t+1)\tilde{u}(t),$$

with the solution $\tilde{u}(t) = \exp\left(-\frac{\imath}{2}\left(\frac{\tau_{\text{end}}}{2}\right)^2 (t+1)\right)$. The coefficient matrix F of the function $f(t,s) = \tilde{f}(t)\Theta(t-s)$ with $\tilde{f}(t) = -\imath\left(\frac{\tau_{\text{end}}}{2}\right)^2 (t+1)$ is pentadiagonal and is discussed in Example 3.6.

For a fixed size of the coefficient matrix, $M = 1000$, we run our procedure for $\tau_{\text{end}} = 25$ and $\tau_{\text{end}} = 50$. Table 5.4 shows the metrics for these cases; a good approximation is obtained in both cases even though the numerical radius is much larger than one.

In Figure 5.2, the spectrum and pseudospectra for the two choices are shown. The eigenvalues of $(I - F^{(N)}_{1000})$ lie on a circle, and, as τ_{end} increases, the radius of this circle increases; however, the center of the circle also shifts, and the circle does not cross the real line. Thus, $(I - F^{(N)}_{1000})^{-1}$ exists in both cases. The pseudospectra that contain the origin are those associated with large perturbations of the coefficient matrix. In Figure 5.3, the order of magnitude of the entries of the inverse for both cases is presented. It shows that the inverse is characterized by the decay phenomenon.

Tables 5.5 and 5.6 show for both the choices $\tau_{\text{end}} = 25$ and $\tau_{\text{end}} = 50$ the function error err_{f} and the amplitude of the last accurately computed Legendre coefficient, respectively. The last accurately computed Legendre coefficient for both choices is $\hat{c}_{M-1}$. From the tables we can conclude that, for M large enough, an estimate for err_{f} can be obtained by looking only at the Legendre coefficients, which is the expected behavior in function approximation with Legendre polynomials [42].

TABLE 5.3

Accuracy of the SHBVM and our method for approximating $\hat{u}(1)$ for the toy problem $\tilde{f}(t) = -\imath\frac{\omega}{\beta}\sin(\omega(t{+}1))$, with $\omega = 100$ and $\beta = 1$.

M	800	1000	1200	1400	1600
SHBVM	3.6e-05	3.6e-07	5.7e-10	1.6e-11	5.6e-14
Our method	1.6e-06	6.0e-10	1.6e-11	7.0e-14	5.8e-14

TABLE 5.4
Metrics for the function $\tilde{f}(t) = -\imath\left(\frac{\tau_{end}}{2}\right)^2(t+1)$ *and for the approximation to* $\tilde{u}(t)$ *obtained for* $M = 1000$.

τ_{end}	$\lvert\alpha_0\rvert + \lvert\alpha_1\rvert$	$\nu(F_{1000}^{(1)})$	err_c	err_f
25	348.5	147.6	$5.228 \cdot 10^{-14}$	$1.067 \cdot 10^{-13}$
50	1394	590.6	$3.210 \cdot 10^{-13}$	$3.008 \cdot 10^{-13}$

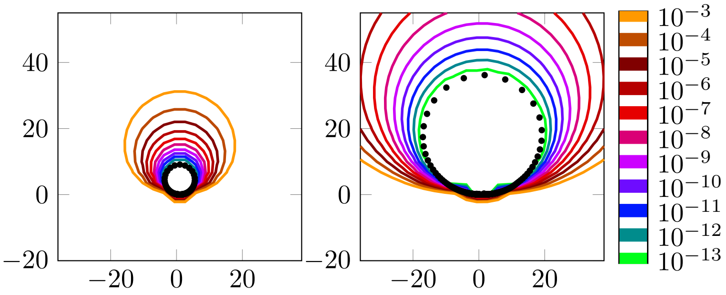


FIG. 5.2. *Spectrum (·) and pseudospectra of* $(I_M - F_M^{(N)})$ *for the coefficient matrix* $F_M^{(N)}$ *of the function* $\tilde{f}(t) = -\imath\left(\frac{\tau_{end}}{2}\right)^2(t+1)$. *Left:* $\tau_{end} = 25$ *and* $M = 1000$. *Right:* $\tau_{end} = 50$ *and* $M = 1000$.

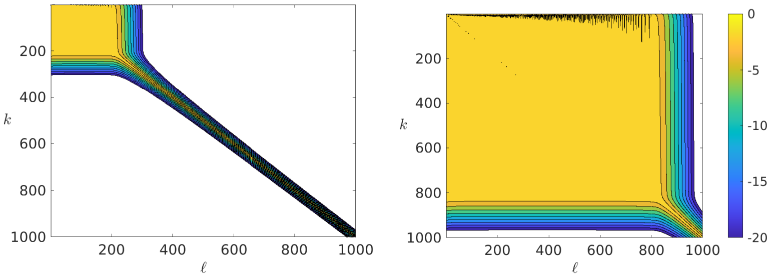


FIG. 5.3. *Order of magnitude of the entries of* $(I_M - F_M^{(N)})^{-1}$ *for* $M = 1000$*, where* $F_M^{(N)}$ *is the coefficient matrix of* $\tilde{f}(t) = -\imath\left(\frac{\tau_{end}}{2}\right)^2(t+1)$. *Left:* $\tau_{end} = 25$. *Right:* $\tau_{end} = 50$.

A comparison of the accuracy of the approximation by SHBVM and our method to $\tilde{u}(\tau_{\text{end}})$ is shown in Table 5.7 and Table 5.8. We observe that our method requires a significantly smaller basis to achieve the same accuracy as the SHBVM.

5.3. NMR-inspired problem. The following experiment is inspired by a problem in nuclear magnetic resonance spectroscopy (NMR), where the matrix ODE

$$\frac{d}{dt}\tilde{A}(t) = \tilde{H}(t)\tilde{A}(t), \quad [0, t_{\text{end}}],$$

governs the dynamics of, e.g., a magic angle spinning experiment. The matrix-valued function $\tilde{H}(t)$ is the Hamiltonian and is of size $2^\ell \times 2^\ell$, where ℓ is the number of spins in the sample [27]. The functions appearing in $\tilde{H}(t)$ are of the form

$$\tilde{f}(t) = -2\imath\pi(\alpha + \beta\cos(2\pi\nu t) + \gamma\cos(4\pi\nu t)), \tag{5.1}$$

where $\alpha \in [-1, 1]$ and $\beta, \gamma \in [100, 5000]$ are typical ranges for these parameters. In a magic angle spinning experiment [21], the sample spins at an angular velocity $\nu \in [5000, 120000]$

TABLE 5.5

Error of the Legendre coefficients and the magnitude of the last accurate Legendre coefficient $\hat{c}_{M-N}$ for increasing M for $\tau_{end} = 25$.

M	$\mathrm{err}_{\mathrm{f}}$	$\lvert\hat{c}_{M-1}\rvert$
200	$1.8 \cdot 10^{0}$	$2.7 \cdot 10^{-2}$
210	$3.3 \cdot 10^{-1}$	$1.3 \cdot 10^{-2}$
220	$1.6 \cdot 10^{-2}$	$1.9 \cdot 10^{-3}$
230	$4.6 \cdot 10^{-4}$	$9.0 \cdot 10^{-5}$
240	$8.0 \cdot 10^{-6}$	$2.2 \cdot 10^{-6}$
250	$8.5 \cdot 10^{-8}$	$2.9 \cdot 10^{-8}$
260	$5.9 \cdot 10^{-10}$	$2.4 \cdot 10^{-10}$
270	$2.8 \cdot 10^{-12}$	$1.2 \cdot 10^{-12}$
280	$9.9 \cdot 10^{-14}$	$2.1 \cdot 10^{-14}$
290	$8.4 \cdot 10^{-14}$	$1.2 \cdot 10^{-14}$
300	$8.7 \cdot 10^{-14}$	$1.2 \cdot 10^{-14}$

TABLE 5.6

Error of the Legendre coefficients and the magnitude of the last accurate Legendre coefficient $\hat{c}_{M-1}$ for increasing M for $\tau_{end} = 50$.

M	$\mathrm{err}_{\mathrm{f}}$	$\lvert\hat{c}_{M-1}\rvert$
830	$8.3 \cdot 10^{-2}$	$2.4 \cdot 10^{-3}$
840	$1.1 \cdot 10^{-2}$	$5.5 \cdot 10^{-4}$
850	$1.1 \cdot 10^{-3}$	$8.4 \cdot 10^{-5}$
860	$9.9 \cdot 10^{-5}$	$9.3 \cdot 10^{-6}$
870	$7.0 \cdot 10^{-6}$	$8.0 \cdot 10^{-7}$
880	$4.0 \cdot 10^{-7}$	$5.4 \cdot 10^{-8}$
890	$1.9 \cdot 10^{-8}$	$3.0 \cdot 10^{-9}$
900	$7.7 \cdot 10^{-10}$	$1.3 \cdot 10^{-10}$
910	$2.6 \cdot 10^{-11}$	$5.0 \cdot 10^{-12}$
920	$9.6 \cdot 10^{-13}$	$1.6 \cdot 10^{-13}$
930	$3.1 \cdot 10^{-13}$	$1.6 \cdot 10^{-14}$

TABLE 5.7

Accuracy of the SHBVM and our method for approximating $\hat{u}(\tau_{\mathrm{end}})$ for the polynomial problem with $\tau_{\mathrm{end}} = 25$.

M	100	200	300	400	500
SHBVM	7.5e-01	1.1e+00	2.0e-01	9.7e-01	2.8e-13
Our method	8.1e-01	1.2e+00	1.4e-14	1.33e-14	1.7e-14

TABLE 5.8

Accuracy of the SHBVM and our method for approximating $\hat{u}(\tau_{\mathrm{end}})$ for the polynomial problem with $\tau_{\mathrm{end}} = 50$.

M	800	1100	1400	1700	2000
SHBVM	9.4e-01	1.0e+00	3.8e-01	8.8e-06	1.4e-14
Our method	1.3e+00	6.8e-14	6.7e-14	9.6e-14	1.2e-13

chosen by the user. The experiment typically runs for about $t_{\mathrm{end}} = 10^{-2}$ seconds. Here, we consider the simpler problem of a scalar ODE

$$\frac{d}{dt}\tilde{u}(t) = \tilde{f}(t)\tilde{u}(t), \quad [0, t_{\mathrm{end}}].$$

We lose the connection to NMR, but studying this problem provides insight into the capabilities of our proposed procedure to tackle the physically relevant matrix case. For $\alpha = 0.05$, $\beta = \gamma = 3450$, $\nu = 5000$, and $M = 1500$, we compute an approximation to $\tilde{u}(t)$. The function approximation error is $\mathrm{err}_{\mathrm{f}} = 1.5994 \cdot 10^{-4}$; increasing M will improve on this error. However, an accuracy of the order 10^{-4} suffices for NMR experiments, where one is limited by the accuracy of the measurements. In Table 5.9, the comparison between SHBVM and our method shows that our method obtains better accuracy in terms of the size of the basis. The spectrum and pseudospectra are displayed in Figure 5.4.

6. Conclusion.

We presented a new approach for the solution of linear non-autonomous scalar ODEs based on the discretization of the $\star$-product by using expansions into series of Legendre polynomials. This approach effectively transforms operations defined on bivariate distributions to operations in a (sub)algebra of infinite matrices. We studied the properties of such matrices and used them to prove the existence of the ODE solution in the infinite matrix algebra. Once the Legendre polynomial series is truncated, the ODE solution is accessible by

TABLE 5.9

Accuracy of the SHBVM and our method for approximating $\hat{u}(t_{\text{end}})$ for the NMR-inspired problem with $\nu = 5000$ and $t_{end} = 10^{-2}$.

M	1100	1200	1300	1400	1500
SHBVM	2.2e-01	1.4e-01	1.2e-02	9.9e-03	7.2e-03
Our method	1.1e-02	2.8e-03	2.2e-03	5.4e-04	8.5e-05

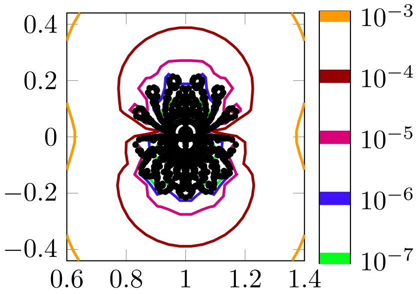


FIG. 5.4. *Spectrum $(\cdot)$ and pseudospectra of $(I - F_M)$ for the coefficient matrix F_M of $\tilde{f}(t)$ (5.1) with parameters $\nu = 5000$, $\alpha = 0.05$, $\beta = \gamma = 3450$, $t_{end} = 10^{-2}$, and $M = 1500$.*

solving a (finite) linear system. We studied the truncation error, proving that obtaining accurate approximations from this finite system is possible. We also presented effective methods to compute the discretization and tested the method on several numerical examples.

The new method was numerically analyzed for the scalar case and validated by comparing its accuracy to an existing method, a spectral Hamiltonian boundary value method. The scalar analysis is a fundamental step toward understanding the more general case of systems of non-autonomous linear ODEs [18, 33]. In fact, the authors are developing a method for this more general case whose analysis and understanding will be built on the crucial results presented here.

Acknowledgments. The second author thanks Marcus Webb for discussion in an early stage of this project. We also thank an anonymous referee for pointing out valuable references.

Appendix A. Proof of Lemma 3.8. In this proof, it is easier to work with the formula based on Legendre polynomials $\dot{p}_k(t)$, which are normalized such that $\dot{p}_k(1) = 1$. Then, setting $s = (a + b + c)/2$, we have

$$\dot{\mathcal{F}}_{a,b,c} := \int_{-1}^{1} \dot{p}_a(\tau)\dot{p}_b(\tau)\dot{p}_c(\tau)d\tau$$
$$= \begin{cases} 0 & \text{if } a+b+c \text{ odd}, \\ 0 & \text{if } s < \max(a,b,c), \\ 0 & \text{if } a < |b-c|, \\ \frac{2}{a+b+c+1}\binom{2(s-a)}{s-a}\binom{2(s-b)}{s-b}\binom{2(s-c)}{s-c}\binom{2s}{s}^{-1} & \text{else.} \end{cases}$$

$$= \begin{cases} 0 & \text{if } a+b+c \text{ odd},\\ 0 & \text{if } b+c<a,\\ 0 & \text{if } a<\alpha:=|b-c|,\\ \frac{1}{2^{(2a-1)}}\frac{1}{(a+b+c+1)}\left(\prod_{j=1}^{a}\frac{-a+b+c+2j}{-a+b+c+2j-1}\right)\frac{\prod_{j=(\frac{a+\alpha}{2}+1)}^{a+\alpha} j^2}{\prod_{j=1}^{\frac{a-\alpha}{2}} j^2 \prod_{j=(a-\alpha+1)}^{a+\alpha} j} & \text{else.}\end{cases}$$

First, we need the following property:

PROPERTY A.1. *The following equality holds for $x \in \mathbb{R}$ and $d = 1, 2, \ldots$:*

$$\frac{\partial}{\partial x}\left(\prod_{j=1}^{d}\frac{2x+2j}{2x+2j-1}\right) = -2\prod_{j=1}^{d}\frac{2x+2j}{2x+2j-1}\sum_{j=1}^{d}\frac{1}{(2x+2j)(2x+2j-1)}.$$

Proof. This follows by induction on d. For $d = 1$:

$$\frac{\partial}{\partial x}\frac{2x+2}{2x+1} = -2\frac{2x+2}{2x+1}\frac{1}{(2x+1)(2x+2)}.$$

Assume the equality holds for d, and consider $d + 1$:

$$\begin{aligned}\frac{\partial}{\partial x}\left(\prod_{j=1}^{d+1}\frac{2x+2j}{2x+2j-1}\right) &= \frac{\partial}{\partial x}\left(\frac{2x+2d+2}{2x+2d+1}\prod_{j=1}^{d}\frac{2x+2j}{2x+2j-1}\right)\\ &= -2\prod_{j=1}^{d+1}\frac{2x+2j}{2x+2j-1}\sum_{j=1}^{d+1}\frac{1}{(2x+2j)(2x+2j-1)}.\end{aligned}$$

This proves the statement. □

LEMMA A.2 (Monotonous decay along the diagonals). *For given integers $d \geq 0$ and $k \leq d$, the following equality is satisfied for $i = 1, 2, \ldots$:*

$$\dot{\mathcal{F}}_{d,k,d-k} > \dot{\mathcal{F}}_{d,k+i,d-k+i} > 0.$$

Proof. For $d = 0$ the integral of the triple product is

$$\dot{\mathcal{F}}_{0,b,c}\int_{-1}^{1}\dot{p}_0(\tau)\dot{p}_b(\tau)\dot{p}_c(\tau)d\tau = \int_{-1}^{1}\dot{p}_b(\tau)\dot{p}_c(\tau)d\tau = \frac{2}{b+c+1},$$

which clearly satisfies $\dot{\mathcal{F}}_{0,0,0} > \dot{\mathcal{F}}_{0,i,i} > 0$. Next, we prove the statement for $d \geq 1$. The elements $\dot{\mathcal{F}}_{d,k+i,d-k+i}$ are clearly positive and nonzero. Set $\alpha = |2k - d|$. Then,

$$\dot{\mathcal{F}}_{d,k+i,d-k+i} = \frac{1}{2^{(2d-1)}}\frac{1}{d+2i+1}\prod_{j=1}^{d}\frac{2i+2j}{2i+2j-1}\frac{\prod_{j=\frac{d+\alpha}{2}+1}^{d+\alpha} j^2}{\prod_{j=1}^{\frac{d-\alpha}{2}} j^2\prod_{j=d-\alpha+1}^{d+\alpha} j}.$$

The term $C(d,\alpha) := \frac{1}{2^{(2d-1)}}\frac{\prod_{j=\frac{d+\alpha}{2}+1}^{d+\alpha} j^2}{\prod_{j=1}^{\frac{d-\alpha}{2}} j^2\prod_{j=d-\alpha+1}^{d+\alpha} j}$ is independent of i. To show that the expression decreases as i increases, we take the derivative with respect to $0 \leq x \in \mathbb{R}$ and use Property A.1:

$$\frac{\partial}{\partial x}\dot{\mathcal{F}}_{d,k+x,d-k+x} = C(d,\alpha)\frac{\partial}{\partial x}\left(\frac{1}{d+2x+1}\prod_{j=1}^{d}\frac{2x+2j}{2x+2j-1}\right)$$

$$= -\frac{2C(d,\alpha)}{d+2x+1}\prod_{j=1}^{d}\frac{2x+2j}{2x+2j-1}\left(\frac{1}{d+2x+1}+\sum_{j=1}^{d}\frac{1}{(2x+2j)(2x+2j-1)}\right)$$
$$< 0.$$

Since $\dot{\mathcal{F}} > 0$ for $i = 0, 1, \ldots$, replacing x with integers $i \geq 0$ proves the statement. □

LEMMA A.3. *For given integers $d \geq 0$ and $k \leq d$,*

$$\dot{\mathcal{F}}_{d,k,d-k} \leq \frac{2}{2d+1}.$$

Proof. In order to prove this lemma it is sufficient to show that

$$\max_{k=0,\ldots,d}\binom{2(d-k)}{d-k}\binom{2k}{k} = \binom{2d}{d}.$$

Note that, because of symmetry, this is equivalent to showing that, for $d/2 \leq k < d$ and $d > 0$, it holds that

$$\binom{2(d-k)}{d-k}\binom{2k}{k} \geq \binom{2(d-k)-2}{d-k-1}\binom{2k-2}{k-1}.$$

The equation above can be reformulated as the quadratic problem

$$4(2d-2k-1)(2k-1) \geq k(d-k).$$

For $d \geq 2$, the inequality above is satisfied for every $d/2 \leq k < d$. The proof is then concluded since the cases $d = 0, 1$ are trivial. □

Proof of Lemma 3.8. The coefficients of $B^{(d)} = \left[b_{k,\ell}^{(d)}\right]_{k,\ell=0}^{\infty}$ are given by the formula

$$b_{k,\ell}^{(d)} = \frac{\sqrt{(2d+1)(2k+1)}}{\sqrt{8}\sqrt{2l+1}}\left(\dot{\mathcal{F}}_{d,k,\ell+1} - \dot{\mathcal{F}}_{d,k,\ell-1}\right).$$

Then, by the definition of the infinity norm, we have

$$\begin{aligned}\|B^{(d)}\|_\infty &= \max_{k\geq 0}\sum_{\ell=0}^{\infty}|b_{k,\ell}^{(d)}| = \frac{\sqrt{2d+1}}{\sqrt{8}}\max_{k\geq 0}\sum_{\ell=0}^{\infty}\left(\frac{\sqrt{2k+1}}{\sqrt{2\ell+1}}\left|\dot{\mathcal{F}}_{d,k,\ell+1} - \dot{\mathcal{F}}_{d,k,\ell-1}\right|\right)\\ &\leq \frac{\sqrt{2d+1}}{\sqrt{8}}\max_{k\geq 0}\sum_{\ell=0}^{\infty}\left(\frac{\sqrt{2k+1}}{\sqrt{2\ell+1}}\left|\dot{\mathcal{F}}_{d,k,\ell+1}\right| + \left|\dot{\mathcal{F}}_{d,k,\ell-1}\right|\right)\\ &= \frac{\sqrt{2d+1}}{\sqrt{8}}\left(\max_{k\geq 0}\sum_{\ell=0}^{\infty}\frac{\sqrt{2k+1}}{\sqrt{2\ell+1}}\left|\dot{\mathcal{F}}_{d,k,\ell+1}\right| + \max_{k\geq 0}\sum_{\ell=0}^{\infty}\frac{\sqrt{2k+1}}{\sqrt{2\ell+1}}\left|\dot{\mathcal{F}}_{d,k,\ell-1}\right|\right).\end{aligned}$$

Consider the first term, and use Lemma A.2, which implies that $\dot{\mathcal{F}}_{d,k+i,d-k+i}$, for any $i \geq 0$, can be bounded from above by $\dot{\mathcal{F}}_{d,k,d-k}$. We can also bound the term $\frac{\sqrt{2k+1}}{\sqrt{2\ell+1}}$:

$$\max_{\substack{0\leq k\leq d\\ \ell=d-k}}\frac{2k+1}{2\ell+1} = \max_{0\leq k\leq d}\left(\frac{2d+2}{2d-(2k-1)}\right) - 1 = \frac{2d+2}{2d-(2d-1)} - 1 = 2d+1.$$

Since $\dot{\mathcal{F}}_{d,k,\ell+1} = 0$, for $|k - \ell - 1| > d$, the infinite sum can be bounded by a finite sum:

$$\max_{k\geq 0} \sum_{\ell=0}^{\infty} \frac{\sqrt{2k+1}}{\sqrt{2\ell+1}} \left|\dot{\mathcal{F}}_{d,k,\ell+1}\right| \leq \sqrt{2d+1} \sum_{k=0}^{d} \dot{\mathcal{F}}_{d,k,d-k}.$$

Since $|\dot{\mathcal{F}}_{d,d,-1}| = |\dot{\mathcal{F}}_{d,d,0}|$, similar arguments lead to the following bound for the second term:

$$\max_{k\geq 0} \sum_{\ell=0}^{\infty} \frac{\sqrt{2k+1}}{\sqrt{2\ell+1}} |\dot{\mathcal{F}}_{d,k,\ell-1}| \leq \sqrt{2d+1}\dot{\mathcal{F}}_{d,d,0} + \sqrt{2d+1} \sum_{k=0}^{d} \dot{\mathcal{F}}_{d,k,d-k}.$$

Hence, we obtain

$$\|B^{(d)}\|_\infty \leq \frac{2d+1}{\sqrt{8}} \left(\dot{\mathcal{F}}_{d,d,0} + 2\sum_{k=0}^{d} \dot{\mathcal{F}}_{d,k,d-k} \right).$$

Now we plug in the formula for $\dot{\mathcal{F}}_{d,k,d-k}$, note that $d+k+d-k = 2d$ and $-d+k+d-k = 0$, and let $\alpha(k) := |2k - d|$. Then,

$$\begin{aligned}\|B^{(d)}\|_\infty &\leq \frac{2d+1}{\sqrt{8}} \left(\dot{\mathcal{F}}_{d,d,0} + 2\sum_{k=0}^{d} \dot{\mathcal{F}}_{d,k,d-k} \right) \leq \frac{2d+1}{\sqrt{8}} \left(1 + 2\sum_{k=0}^{d} \frac{2}{2d+1} \right) \\ &\leq \frac{2d+1}{\sqrt{8}} \left(1 + \frac{4(d+1)}{2d+1} \right) \leq \frac{6d+5}{\sqrt{8}} < 3d+2,\end{aligned}$$

which proves Lemma 3.8. □

REFERENCES

[1] P. Amodio, L. Brugnano, and F. Iavernaro, *Analysis of spectral Hamiltonian boundary value methods (SHBVMs) for the numerical solution of ODE problems*, Numer. Algorithms, 83 (2020), pp. 1489–1508.

[2] J. L. Aurentz and L. N. Trefethen, *Chopping a Chebyshev series*, ACM Trans. Math. Software, 43 (2017), Art. 33, 21 pages.

[3] M. Benzi, *Localization in matrix computations: theory and applications*, in Exploiting Hidden Structure in Matrix Computations: Algorithms and Applications, M. Benzi and V. Simoncini, eds., vol. 2173 of Lecture Notes in Math., Springer, Cham, 2016, pp. 211–317.

[4] ———, *Some uses of the field of values in numerical analysis*, Boll. Unione Mat. Ital., 14 (2021), pp. 159–177.

[5] M. Benzi and P. Boito, *Decay properties for functions of matrices over C^*-algebras*, Linear Algebra Appl., 456 (2014), pp. 174–198.

[6] C. Bonhomme, S. Pozza, and N. Van Buggenhout, *A new fast numerical method for the generalized Rosen-Zener model*, Preprint on arXiv, 2023. https://arxiv.org/abs/2311.04144

[7] L. Brugnano and F. Iavernaro, *Line Integral Methods for Conservative Problems*, CRC Press, Boca Raton, 2016.

[8] L. Brugnano, F. Iavernaro, and D. Trigiante, *A simple framework for the derivation and analysis of effective one-step methods for ODEs*, Appl. Math. Comput., 218 (2012), pp. 8475–8485.

[9] L. Brugnano, J. I. Montijano, and L. Rández, *On the effectiveness of spectral methods for the numerical solution of multi-frequency highly oscillatory Hamiltonian problems*, Numer. Algorithms, 81 (2019), pp. 345–376.

[10] F. Casas, *Sufficient conditions for the convergence of the Magnus expansion*, J. Phys. A, 40 (2007), pp. 15001–15017.

[11] F. Diele, L. Lopez, and R. Peluso, *The Cayley transform in the numerical solution of unitary differential systems*, Adv. Comput. Math., 8 (1998), pp. 317–334.

[12] T. A. Driscoll, N. Hale, and L. N. Trefethen, *Chebfun Guide*, Pafnuty Publications, Oxford, 2014.

[13] A. Gelb and J. Tanner, *Robust reprojection methods for the resolution of the Gibbs phenomenon*, Appl. Comput. Harmon. Anal., 20 (2006), pp. 3–25.

[14] J. GILLIS, J. JEDWAB, AND D. ZEILBERGER, *A combinatorial interpretation of the integral of the product of Legendre polynomials*, SIAM J. Math. Anal., 19 (1988), pp. 1455–1461.

[15] P.-L. GISCARD AND C. BONHOMME, *Dynamics of quantum systems driven by time-varying Hamiltonians: solution for the Bloch-Siegert Hamiltonian and applications to NMR*, Phys. Rev. Res., Art. 023081, 18 pages.

[16] P.-L. GISCARD, K. LUI, S. J. THWAITE, AND D. JAKSCH, *An exact formulation of the time-ordered exponential using path-sums*, J. Math. Phys., 56 (2015), Art. 053503, 18 pages.

[17] P.-L. GISCARD AND S. POZZA, *Lanczos-like algorithm for the time-ordered exponential: the $*$-inverse problem*, Appl. Math., 65 (2020), pp. 807–827.

[18] ———, *A Lanczos-like method for non-autonomous linear ordinary differential equations*, Boll. Unione Mat. Ital., 16 (2023), pp. 81–102.

[19] D. GOTTLIEB AND C.-W. SHU, *On the Gibbs phenomenon and its resolution*, SIAM Rev., 39 (1997), pp. 644–668.

[20] K. E. GUSTAFSON AND D. K. M. RAO, *Numerical Range*, Springer, New York, 1977.

[21] S. HAFNER AND H.-W. SPIESS, *Advanced solid-state NMR spectroscopy of strongly dipolar coupled spins under fast magic angle spinning*, Concepts Magn. Reson., 10 (1998), pp. 99–128.

[22] A. ISERLES, K. KROPIELNICKA, AND P. SINGH, *Magnus-Lanczos methods with simplified commutators for the Schrödinger equation with a time-dependent potential*, SIAM J. Numer. Anal., 56 (2018), pp. 1547–1569.

[23] A. ISERLES, H. Z. MUNTHE-KAAS, S. P. NØRSETT, AND A. ZANNA, *Lie-group methods*, Acta Numer., 9 (2000), pp. 215–365.

[24] K. KORMANN, S. HOLMGREN, AND H. O. KARLSSON, *Accurate time propagation for the Schrödinger equation with an explicitly time-dependent Hamiltonian*, J. Chem. Phys., 128 (2008), Art. 184101, 11 pages

[25] S. H. KULKARNI, R. RADHA, AND K. SARVESH, *Solution of an infinite band matrix equation*, Banach J. Math. Anal., 17 (2023), Art. 14, 28 pages.

[26] L. LAZZARINO, *Numerical Approximation of the Time-Ordered Exponential for Spin Dynamic Simulation*, Master Thesis, Univerzita Karlova, Prague, 2023.

[27] M. H. LEVITT, *Spin Dynamics: Basics of Nuclear Magnetic Resonance*, 2nd ed., Wiley, Chichester, 2008.

[28] E. S. MANANGA, *On the Fer expansion: applications in solid-state nuclear magnetic resonance and physics*, Phys. Rep., 608 (2016), pp. 1–41.

[29] M. NDONG, H. TAL-EZER, R. KOSLOFF, AND C. P. KOCH, *A Chebychev propagator with iterative time ordering for explicitly time-dependent Hamiltonians*, J. Chem Phys., 132 (2010), Art. 064105, 12 pages.

[30] U. PESKIN, R. KOSLOFF, AND N. MOISEYEV, *The solution of the time dependent Schrödinger equation by the (t,t') method: The use of global polynomial propagators for time dependent Hamiltonians*, J. Chem. Phys., 100 (1994), pp. 8849–8855.

[31] S. POZZA, *A new closed-form expression for the solution of ODEs in a ring of distributions and its connection with the matrix algebra*, Linear Multilinear Algebra, (2024), pp. 1–11. Advance online publication https://doi.org/10.1080/03081087.2024.2303058

[32] S. POZZA AND V. SIMONCINI, *Inexact Arnoldi residual estimates and decay properties for functions of non-Hermitian matrices*, BIT Numer. Math., 59 (2019), pp. 969–986.

[33] S. POZZA AND N. VAN BUGGENHOUT, *A new matrix equation expression for the solution of non-autonomous linear systems of ODEs*, Proc. Appl. Math. Mech., 22 (2023), Art. e202200117, 6 pages.

[34] ———, *The $\star$-product approach for linear ODEs: A numerical study of the scalar case*, in Programs and Algorithms of Numerical Mathematics. Proceedings of Seminar, J. Chleboun, P. Kůs, J. Papež, M. Rozložník, K. Segeth, and J. Šístek, eds., Institute of Mathematics CAS, Prague, 2023, pp. 187–198.

[35] ———, *A $\star$-product solver with spectral accuracy for non-autonomous ordinary differential equations*, Proc. Appl. Math. Mech., 23 (2023), Art. e202200050, 6 pages.

[36] M. RYCKEBUSCH, *A Fréchet-Lie group on distributions*, Preprint on arXiv, 2023. https://arxiv.org/abs/2307.09037

[37] G. SANSONE, *Orthogonal Functions*, rev. ed., R. E. Krieger Pub., Huntington, 1977.

[38] I. SCHAEFER, H. TAL-EZER, AND R. KOSLOFF, *Semi-global approach for propagation of the time-dependent Schrödinger equation for time-dependent and nonlinear problems*, J. Comput. Phys., 343 (2017), pp. 368–413.

[39] R. SCHNEIDER, H. GHARIBNEJAD, AND B. I. SCHNEIDER, *ITVOLT: An iterative solver for the time-dependent Schrödinger equation*, Comput. Phys. Commun., 291 (2023), Art. 108780, 13 pages.

[40] L. SCHWARTZ, *Théorie des Distributions*, vol. IX-X of Publications de l'Institut de Mathématique de l'Université de Strasbourg, Hermann, Paris, 1966.

[41] A. TOWNSEND, M. WEBB, AND S. OLVER, *Fast polynomial transforms based on Toeplitz and Hankel matrices*, Math. Comp., 87 (2018), pp. 1913–1934.

[42] L. N. TREFETHEN, *Approximation Theory and Approximation Practice*, SIAM, Philadelphia, 2013.

[43] L. N. TREFETHEN AND M. EMBREE, *Spectra and Pseudospectra*, Princeton University Press, Princeton, 2005.

[44] H. WANG AND S. XIANG, *On the convergence rates of Legendre approximation*, Math. Comp., 81 (2012), pp. 861–877.