RESEARCH Open Access

Check for updates

Computation of solutions to linear difference and differential equations with a prescribed asymptotic behavior

Hendrik Baumann^{1*} ond Thomas Hanschke²

*Correspondence: baumann@math.uni-hannover.de ¹Institute of Analysis, Leibniz-Universität Hannover, 30167 Hannover, Germany Full list of author information is available at the end of the article

Abstract

Linear differential equations usually arise from mathematical modeling of physical experiments and real-world problems. In most applications these equations are linked to initial or boundary conditions. But sometimes the solution under consideration is characterized by its asymptotic behavior, which leads to the question how to infer from the asymptotic growth of a solution to its initial values. In this paper we show that under some mild conditions the initial values of the desired solution can be computed by means of a continuous-time analogue of a modified matrix continued fraction. For numerical applications we develop forward and backward algorithms which behave well in most situations. The topic is closely related to the theory of special functions and its extension to higher-dimensional problems. Our investigations result in a powerful tool for solving some classical mathematical problems. To demonstrate the efficiency of our method we apply it to Poincaré type and Kneser's differential equation.

Keywords: Linear systems of difference and differential equations; Asymptotic behavior of solutions of linear differential equations; Modified matrix-driven Jacobi–Perron algorithm; Poincaré–Perron-type differential equation; Kneser's differential equation

1 Introduction

A huge variety of real-world problems is modelled by means of linear differential equations. Although questions of existence and uniqueness of solutions are easy to answer for linear differential equations, explicit solutions are rare, in particular for differential equations of order > 2, so that numerical procedures are needed. But most of these procedures are restricted to initial and boundary value problems.

However, in some applications we do not have initial or boundary conditions. In probability theory for example, densities of probability distribution functions, stationary measures of stochastic processes or certain functionals defined on diffusion processes are characterized by ordinary differential equations and are often uniquely determined by a single regularity or integrability condition (see [1]).



© The Author(s) 2021. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Another area of possible applications concerns linear differential equations with almost constant coefficients. Following Levinson's seminal paper [2], a relatively rich theory concerning the asymptotic behavior of their solutions has been developed. Here, it is obvious to ask which initial values correspond to which asymptotic behavior, and how to compute these solutions numerically. Since the mathematical theory of differential equations is primarily qualitatively oriented, there is only little literature addressing this problem (see [3]). An exception is the class of second-order linear differential equations, which has come into focus due to Miller's famous algorithm (see [4–8]). This is because it has been turned out that many special functions can be represented as so-called minimal solutions of second-order linear differential equations and their discrete analogues. But when it was recognized that this approach could not directly be transferred to higher order linear difference and differential equations (see [8]), the interest in this subject waned. For discrete systems and scalar differential equations, some of these problems have been investigated by Schäfke [9, 10] and Hanschke [11, 12].

In the sequel it is shown that the initial values corresponding to a specific growth property can be determined by means of a continuous-time analogue of a modified matrix continued fraction. Matrix continued fractions have been originally developed for computing subdominant solutions of linear systems of difference equations (see P. Levrie and A. Bultheel [13]) and stationary measures of discrete state Markov chains with block-band transition matrix (see [14, 15]).

Note that standard numerical methods for solving differential equations are dedicated to solving boundary value problems or initial value problems. In particular, if a numerical method is applied to a differential equation for some function $x:[0,\infty)\to\mathbb{R}$, consistency/convergence statements guarantee that the approximation to x(t) converges to the exact value x(t) as the maximum step size h tends to 0 (note that we only consider linear problems, and hence, consistency implies convergence). Furthermore, the order of consistency/convergence gives some estimates on the speed of convergence. However, for any method, the error will increase with t, in most cases exponentially. This fact entails that the solution of the discretized system can have other asymptotic properties than the exact solution. In so far, finding a solution with prescribed asymptotic behavior, requires a thorough consideration not only of the original equation but also of its corresponding discretization scheme. The alignment of these two systems may be crucial for the success of the approach.

The rest of this paper is organized as follows: After some remarks on 2×2 -block-matrices and their inverses we present our main results. In order to be able to distinguish the asymptotically differently growing solutions of the underlying differential equation we introduce the concept of Σ -subdominant solutions, which is a generalization of the concept introduced in [12]. It turns out that under certain regularity conditions each subspace of Σ -subdominant solutions can be characterized by means of a continuous-time analogue of a modified matrix continued fraction. Since matrix continued fractions represent generalizations of the Jacobi–Perron-algorithm (see [16]), we refer to this correspondence as Jacobi–Perron-characterization. To demonstrate the efficiency of our method we apply it to Poincaré type and Kneser's differential equation. Numerical examples complete our work.

2 Preparation: some notes on block matrices

In this paper, we will frequently use a block partition of some $r \times r$ -matrix

$$A = \begin{pmatrix} B & C \\ D & E \end{pmatrix},$$

where B is a square $p \times p$ -matrix for some p < r. In order to avoid an overload of notation, we introduce $F_1 = \binom{I_p}{0} \in \mathbb{C}^{r \times p}$ and $F_2 = \binom{0}{I_{r-p}} \in \mathbb{C}^{r \times (r-p)}$. Hence, in the above partition, we have $B = F_1^T A F_1$, $C = F_1^T A F_2$,

Sometimes, we have to derive the inverse of a block-partitioned matrix. Fundamental identities are

$$A^{-1} = \begin{pmatrix} B & C \\ D & E \end{pmatrix}^{-1} = \begin{pmatrix} B^{-1} + B^{-1}CSDB^{-1} & -B^{-1}CS \\ -SDB^{-1} & S \end{pmatrix}$$
 (1)

$$= \begin{pmatrix} V & -VCE^{-1} \\ -E^{-1}DV & E^{-1} + E^{-1}DVCE^{-1} \end{pmatrix},$$
 (2)

where $S = (E - DB^{-1}C)^{-1}$ and $V = (B - CE^{-1}D)^{-1}$, provided that A is non-singular and that B or E, respectively, are non-singular square submatrices; see [17, pp. 37–39].

3 Setting, notations, and definitions

We consider solutions to

$$x(t+1) = A(t)x(t), \quad t \in I = \{t_0, t_0 + 1, t_0 + 2, \dots\}$$
 or (3)

$$x'(t) = A(t)x(t), \quad t \in I \setminus t_0 \text{ where } I = [t_0, \infty),$$
 (4)

where $A(t) \in \mathbb{C}^{r \times r}$ is assumed to be invertible for all $t \in I$. In case of the differential equation (4), we look for solutions being differentiable on $I \setminus t_0$ and right-continuous in t_0 . Note that a scalar rth order linear differential equation can always be transformed into a system of the form (4). Therefore, we do not consider scalar rth order linear differential equations separately.

For the solutions of both (3) and (4), we can state some well-known properties:

- The solution space of (3) and (4), respectively, is a linear vector space of dimension r.
- A fundamental system of solutions consists of r linearly independent solutions, say $(x^{(j)}(t))_{t\in I}$ for $j=1,\ldots,r$. By setting $z_{ij}(t)=x_i^{(j)}(t)$, we define matrices $Z(t)=(z_{ij}(t))_{i,j=1}^r$ which are invertible for all $t\in I$ (due to the linear independence of the columns), and the sequence $(Z(t))_{t\in I}$ is a matrix-valued solution of (3) or (4).
- We can directly look for matrix-valued solutions to (3) or (4). If $Z = (Z(t))_{t \in I}$ is a solution with $Z(t) \in \mathbb{C}^{r \times p}$ for some $p \in \mathbb{N}$, then we obtain another solution $\tilde{Z} = (\tilde{Z}(t))_{t \in I}$ with $\tilde{Z}(t) \in \mathbb{C}^{r \times s}$ by setting $\tilde{Z}(t) = Z(t)B$ for all $t \in I$ with some constant matrix $B \in \mathbb{C}^{p \times s}$ for some $s \in \mathbb{N}$. Furthermore, if Z(t) has full rank (which is p for $p \leq r$) for some $t \in I$, then Z(t) has full rank for all $t \in I$.
- In particular, we can look for solutions $Z = (Z(t))_{t \in I}$ with square matrices $Z(t) \in \mathbb{C}^{r \times r}$. Then Z(t) might be invertible, and the above rank statement can be written as follows: If $Z^{-1}(t)$ exists for some $t \in I$ then $Z^{-1}(t)$ exists for all $t \in I$. Furthermore, if Z is a

solution with existing inverse $Z^{-1}(t_0)$, and $\tilde{Z}=(\tilde{Z}(t))$ is another solution, we have $\tilde{Z}(t)=Z(t)Z^{-1}(t_0)\tilde{Z}(t_0)$. For the latter statement, $\tilde{Z}(t)$ is not necessarily a square matrix.

As pointed out in the introduction, often we know that there are vector-valued solutions with a certain asymptotic behavior. These solutions form a p-dimensional subspace (p < r) of the space of all vector-valued solutions. This p-dimensional vector space is uniquely characterized by a matrix-valued solution $P = (P(t))_{t \in I}$ with matrices $P(t) \in \mathbb{C}^{r \times p}$ of full rank p, since then we obtain vector-valued solutions $x = (x(t))_{t \in I}$ by setting $x(t) = P(t)\alpha$ for some $\alpha \in \mathbb{C}^{p \times 1}$.

The characterization of the asymptotic behavior can have different forms:

• In some situations, we know that there is a solution (Z(t)) of square and invertible matrices with an asymptotic representation of the form

$$Z(t) = Y(t) + E(t),$$

where the error-term E(t) is 'neglectible' in comparison to the matrices Y(t) in the sense that $Y^{-1}(t)$ exists for all $t \ge t_1$ (with some $t_1 \in I$) with $\lim_{t \to \infty} ||Y^{-1}(t)E(t)|| = 0$. We will briefly express this assumption as

$$\lim_{t \to \infty} Y^{-1}(t)Z(t) = I. \tag{5}$$

If P(t) contains the first p columns of Z(t), that is, $P(t) = Z(t)F_1$, P(t) has rank p, and hence, the vector space of vector-valued solutions $x = (x(t))_{t \in I}$ with $x(t) = P(t)\alpha$ for some α is p-dimensional. In some sense, the columns of P(t) have an asymptotic expansion given by the first p columns of Y(t). (We have to be careful with this interpretation since the characterization of E(t) as the 'error term' is given in the form $Y^{-1}(t)E(t) \to 0$.)

• Let $P(t) = Z(t)F_1$ contain the first p columns of Z(t), and let $R(t) = Z(t)F_2$ contain the last r - p columns of Z(t). Set $\Sigma^T(t) = F_2^T Y^{-1}(t) \in \mathbb{C}^{(r-p)\times r}$. Under the assumption (5), it becomes obvious that $\Sigma^T(t)R(t)$ converges to I_{r-p} while $\Sigma^T(t)P(t) \in \mathbb{C}^{(r-p)\times p}$ converges to a null matrix of appropriate size. Hence, we have

$$\lim_{t \to \infty} \left(\Sigma^T(t) R(t) \right)^{-1} \left(\Sigma^T(t) P(t) \right) = 0. \tag{6}$$

This may be interpreted as follows: With respect to multiplication with $\Sigma^T(t)$, the solution $(P(t))_{t \in I}$ is *subdominant* to the solution $(R(t))_{t \in I}$, or concisely, (P(t)) is Σ -subdominant. In some sense, P(t) can be interpreted as 'asymptotically normal' to $\Sigma(t)$.

• Note that under the assumption (5), both denominator and numerator in (6) converge for the choice $\Sigma^T(t) = F_2^T Y^{-1}(t)$. On the other hand, there are situations where (6) holds for a specific choice of the family $(\Sigma(t))_{t \in I}$ of matrices with full rank r-p, but neither denominator nor numerator converge for $t \to \infty$. Hence, for Z(t) = (P(t) R(t)), the characterization of P(t) as the first p columns of a solution (Z(t)) of invertible matrices satisfying (5) is a special case of the characterization (6) with a specific choice of $(\Sigma(t))$, whereas the opposite is not true. Hence, the characterization of P(t), being subdominant with respect to some family $(\Sigma(t))$ of matrices, is more general.

• For p = r - 1, $\sigma^T(t) = \Sigma^T(t)$ and R(t) are vectors. Hence (6) may be written as

$$\lim_{t\to\infty} \frac{\langle \sigma^T(t), P_{\bullet,j}(t) \rangle}{\langle \sigma^T(t), R(t) \rangle} = 0, \quad j = 1, \dots, r-1,$$

where $P_{\bullet,i}(t)$ denotes the *j*th column of P(t). This is equivalent to

$$\lim_{t \to \infty} \frac{\langle \sigma^T(t), x(t) \rangle}{\langle \sigma^T(t), \tilde{x}(t) \rangle} = 0 \tag{7}$$

for any solutions x, \tilde{x} of (3) or (4), respectively, with the following assumptions: There is some vector $\alpha \in \mathbb{C}^{p \times 1}$ with $x(t) = P(t)\alpha$ for all $t \in I$, and there is no such vector for \tilde{x} . This is the concept of σ -subdominance which was introduced in [3, 12].

Since the Σ -subdominance is the most general concept for characterizing the asymptotic behavior, we introduce a special notation for the space of all vector-valued solution which can be derived from the matrix-valued Σ -subdominant solution (P(t)): Let $\Sigma = (\Sigma(t))_{t \in I}$ of $r \times (r-p)$ -dimensional matrices with rank r-p, and let $(Z(t))_{t \in I}$ be a solution to (3) or (4), respectively, with the property (6) for $P(t) = Z(t)F_1$ and $R(t) = Z(t)F_2$. Then we define

$$S_{\Sigma} = \{x = (x(t))_{t \in I} : x(t) = P(t)\alpha \text{ for all } t \in I \text{ with some } \alpha \in \mathbb{C}^{p \times 1}\}.$$

The following is about the determination of S_{Σ} .

4 The Jacobi-Perron characterization of Σ -subdominant solutions

In the special case of p = r - 1, (6) and (7) become equivalent. Up to notation, in [12], it was suggested to characterize the subspace S_{σ} as follows:

- For all $\tau \in I$ and for all i = 1, ..., r, define $(\Psi_{ij}(\tau, t))_{i=1}^r$ as a (vector-valued) solution to (3), subject to $\psi_{ij}(\tau, t) = \delta_{ij}$, that is, the sequence of matrices $\Xi(\tau, t) = (\Psi_{ij}(\tau, t))_{i,j=1}^r$ is a matrix-valued solution of (3) with $\Xi(\tau, \tau) = I_r$.
- · In the case of existence, set

$$\eta_i(\tau,\sigma) = \lim_{n \to \infty} \frac{\langle \sigma(t), \Psi_j(\tau,t) \rangle}{\langle \sigma(t), \Psi_r(\tau,t) \rangle}, \quad j = 1, \dots, r-1.$$

• Then it was shown that the limits $\eta_1(\tau, \sigma), \dots, \eta_{r-1}(\tau, \sigma)$ exist and are finite iff there are r linearly independent solutions $x^{(1)}, \dots, x^{(r)}$ to (3) satisfying

$$\lim_{t \to \infty} \frac{\langle \sigma(t), x^{(j)}(t) \rangle}{\langle \sigma(t), x^{(r)}(t) \rangle} = 0, \quad i = 1, \dots, r - 1,$$
(8)

and

$$\begin{vmatrix} x_1^{(1)}(\tau) & \cdots & x_1^{(r-1)}(\tau) \\ x_2^{(1)}(\tau) & \cdots & x_2^{(r-1)}(\tau) \\ \vdots & & \vdots \\ x_{r-1}^{(1)}(\tau) & \cdots & x_{r-1}^{(r-1)}(\tau) \end{vmatrix} \neq 0.$$

$$(9)$$

• If (8) and (9) hold, we have

$$x \in \text{span}\{x^{(1)}, \dots, x^{(r-1)}\}$$
 iff $\sum_{j=1}^{r-1} \eta_j(\tau, \sigma) x_j(\tau) + x_r(\tau) = 0.$ (10)

- (10) characterizes the rth entry of $x(\tau)$ in terms of the other entries. By re-inserting (10) into (3), it is seen that the first r-1 entries of $x(\tau)$ satisfy a system of the form (3) of order r-1.
- In [12], it was demonstrated how this method can be used to reduce the order of the underlying system step by step. In addition, it was shown that the algorithm may be interpreted as a generalization of so-called Jacobi–Perron algorithm; see [16].

In this paper, we deal with the determination of S_{Σ} for arbitrary p < r in only one step. For this purpose, let $(\Sigma(t))_{t \in I}$ be given, define $(\Xi(\tau, t))$ as a solution to (3) or (4) with $\Xi(\tau, \tau) = I_r$, let F_1 , F_2 be defined as above, and, in the case of existence, set

- $\eta(\tau, \Sigma) = \lim_{t \to \infty} (\Sigma^T(t)\Xi(\tau, t)F_2)^{-1}(\Sigma^T(t)\Xi(\tau, t)F_1) \in \mathbb{C}^{(r-p)\times p}$
- $\xi_1(\tau, \Sigma) = \lim_{t \to \infty} \Sigma^T(t) \Xi(\tau, t) F_1 \in \mathbb{C}^{(r-p) \times p}$, and
- $\xi_2(\tau, \Sigma) = \lim_{t \to \infty} \Sigma^T(t) \Xi(\tau, t) F_2 \in \mathbb{C}^{(r-p) \times (r-p)}$.

Obviously, if $\xi_1(\tau, \Sigma)$ and $\xi_2(\tau, \Sigma)$ do exist so does $\eta(\tau, \Sigma) = \xi_2^{-1}(\tau, \Sigma)\xi_1(\tau, \Sigma)$. On the other hand, the existence of $\eta(\tau, \Sigma)$ does not imply the existence of $\xi_i(\tau, \sigma)$ for i = 1, 2. The next result is a straightforward extension of the results in [12].

Theorem 4.1 Let $(\Sigma(t))_{t\in I}$ be a family of matrices $\Sigma(t) \in \mathbb{C}^{r \times (r-p)}$ with rank r-p, and let $F_1, F_2, \Xi(\tau, t), \eta(\tau, \Sigma)$ be defined as above. The limit $\eta(\tau, \Sigma)$ exists if and only if there is a solution $(Z(t))_{t\in I}$ to (3) or (4), respectively, with $\det(Z(t_0)) \neq 0$,

$$\lim_{t \to \infty} \left(\Sigma^T(t) Z(t) F_2 \right)^{-1} \left(\Sigma^T(t) Z(t) F_1 \right) = 0, \quad and$$
 (11)

$$\det(F_1^T Z(\tau) F_1) \neq 0. \tag{12}$$

If these conditions hold, we have $x \in S_{\Sigma}$ *if and only if*

$$\eta(\tau, \Sigma) \cdot F_1^T x(\tau) + F_2^T x(\tau) = 0. \tag{13}$$

Proof In [12], the result was proven with respect to the recurrence relation (3) and p = r - 1. Since our result is more general, we state a detailed proof here. However, the basic principles of the proof remain the same.

Assume first that the limit $\eta(\tau, \Sigma)$ exists. Define Z(t) by

$$Z(t)F_1 = \Xi(\tau, t)F_1 - \Xi(\tau, t)F_2 \cdot \eta(\tau, \Sigma)$$
 and $Z(t)F_2 = \Xi(\tau, t)F_2$.

Then

$$F_1^T Z(\tau) F_1 = F_1^T \Xi(\tau, t) F_1 - F_1^T \Xi(\tau, t) F_2 \cdot \eta(\tau, \Sigma) = I_n - 0 = I_n$$

implying condition (12). Condition (11) follows from

$$\begin{split} &\lim_{t\to\infty} \left(\Sigma^T(t)Z(t)F_2\right)^{-1} \left(\Sigma^T(t)Z(t)F_1\right) \\ &= \lim_{t\to\infty} \left(\Sigma^T(t)\Xi(\tau,t)F_2\right)^{-1} \left(\Sigma^T(t)\Xi(\tau,t)F_1 - \Sigma^T(t)\Xi(\tau,t)F_2 \cdot \eta(\tau,\Sigma)\right) \\ &= \eta(\tau,\Sigma) - \left(\Sigma^T(t)\Xi(\tau,t)F_2\right)^{-1} \Sigma^T(t)\Xi(\tau,t)F_2 \cdot \eta(\tau,\Sigma) = 0. \end{split}$$

Conversely, suppose there is a solution $(Z(t))_{t\in I}$ of invertible matrices satisfying (11) and (12). Then we have $Z(t) = \Xi(\tau,t)Z(t)$, or equivalently $\Xi(\tau,t) = Z(t)Z^{-1}(\tau)$. Let $Z(\tau) = {B \choose DE}$. (12) guarantees that B is invertible so that we are enabled to apply inversion formula (1) with

$$Z^{-1}(\tau) = \begin{pmatrix} B' & C' \\ D' & E' \end{pmatrix} = \begin{pmatrix} B^{-1} + B^{-1}CSDB^{-1} & -B^{-1}CS \\ -SDB^{-1} & S \end{pmatrix}.$$

Note that the definition of S ensures that E' is invertible. Hence, we have $\Xi(\tau,t)F_2 = Z(t)\binom{C'}{E'}$ and $\Xi(\tau,t)F_1 = Z(t)\binom{B'}{D'}$, implying

$$\begin{split} &\lim_{t\to\infty} \left(\Sigma^T(t)\Xi(\tau,t)F_2\right)^{-1}\Sigma^T(t)\Xi(\tau,t)F_1\\ &= \left(\Sigma^T(t)Z(t)F_2E' + \Sigma^T(t)Z(t)F_1C'\right)^{-1}\left(\Sigma^T(t)Z(t)F_2D' + \Sigma^T(t)Z(t)F_1E'\right)\\ &= \lim_{t\to\infty} \left(E' + \left(\Sigma^T(t)Z(t)F_2\right)^{-1}\left(\Sigma^T(t)Z(t)F_1\right)C'\right)^{-1}\\ &\quad \cdot \left(D' + \left(\Sigma^T(t)Z(t)F_2\right)^{-1}\left(\Sigma^T(t)Z(t)F_1\right)E'\right)\\ &= E'^{-1}D' = -DB^{-1}. \end{split}$$

Therefore, $\eta(\tau, \Sigma)$ exists with $\eta(\tau, \Sigma)B + D = 0$, that is,

$$\eta(\tau, \Sigma) \cdot F_1^T Z(\tau) F_1 + F_2^T Z(\nu) F_1 = 0.$$

This concludes the proof since we have $x \in S_{\Sigma}$ if and only if $x(t) = Z(t)F_1\alpha$ for some $\alpha \in \mathbb{C}^{p \times 1}$

We refer to Theorem 4.1 as the *Jacobi–Perron characterization* of Σ -subdominant solutions since the results gives an exact characterization of the *p*-dimensional Σ -subdominant subspace of solutions.

Note that Theorem 4.1 allows one to reduce the order of the differential equation if $\eta(\tau, \Sigma)$ can be computed explicitly: From (3) and (4), respectively, we obtain

$$\left. \begin{array}{l} F_1^T x(\tau+1) \\ F_1^T x'(\tau) \end{array} \right\} = F_1^T A(\tau) F_1 \cdot F_1^T x(\tau) + F_1^T A(\tau) F_2 \cdot F_2^T x(\tau). \tag{14}$$

By inserting (13) into (14), we get a reduced system for the first p entries of $x(t) \in S_{\Sigma}$:

$$\left. \begin{array}{l} F_1^T x(\tau+1) \\ F_1^T x'(\tau) \end{array} \right\} = \left(F_1^T A(\tau) F_1 - F_1^T A(\tau) F_2 \cdot \eta(\tau, \Sigma) \right) F_1^T x(\tau). \tag{15}$$

The other entries of x(t) can be calculated by applying (13) a second time.

Remark 4.1 Let u(t) satisfy an rth order linear difference or differential equation. By setting $x(t) = (u(t), u(t+1), \dots, u(t+r-1))^T$ or $(u(t), u'(t), \dots, u^{(r-1)}(t))^T$, we easily obtain an equation of the form (3) or (4), respectively. In this case, (15) contains all information we are looking for, and another application of (13) for computing $x_{p+1}(t), \dots, x_r(t)$ is not necessary. In particular, for p = 1, we obtain a first order difference or differential equation for u(t).

Remark 4.2 In Theorem 5.2, we will show that the approximate values of $\eta(\tau,t)$ satisfy a continued-fraction-type recursion scheme. For this reason, we refer to $\eta(\tau,t)$ as some kind of matrix-driven continued fraction. Note that a similar construction can be found in [13] although the authors only use the ordinary term of subdominance. The results based on this can be derived from the more general concept by the restriction $Y(t) = I_r$ or $\Sigma^T(t) = F_2^T$. For an overview on extensions of ordinary and generalized continued fractions with a special focus on matrix-driven models consult [18]. With regard to this overview, Theorem 4.1 may be interpreted as a matrix-version of Pincherle-type convergence criteria for ordinary and generalized continued fractions [19–21].

Next, we show that an asymptotic expansion of the form (5) allows to find an even simpler variant of the Jacobi–Perron characterization.

Theorem 4.2 Let $(Z(t))_{t\in I}$ be a fundamental system for (3) or (4), respectively, let $(Y(t))_{t\in I}$ be a family of matrices satisfying (5), and set $\Sigma^T(t) = F_2^T Y^{-1}(t)$ for all $t \in I$. Then $\xi_1(\tau, \Sigma) \in \mathbb{C}^{(r-p)\times p}$ and $\xi_2(\tau, \Sigma) \in \mathbb{C}^{(r-p)\times (r-p)}$ as defined above exist and $x \in \mathcal{S}_{\Sigma}$ if and only if

$$\xi_1(\tau, \Sigma) \cdot F_1^T x(\tau) + \xi_2(\tau, \Sigma) \cdot F_2^T x(\tau) = 0. \tag{16}$$

Proof By definition, we have

$$\xi_i(\tau, \Sigma) = \lim_{t \to \infty} \Sigma^T(t) \Xi(\tau, t) F_i.$$

Due to $Z(t) = \Xi(\tau, t)Z(\tau)$, we can write $\Xi(\tau, t) = Z(t)Z^{-1}(\tau)$. Since $\Sigma^T(t) = F_2^T Y^{-1}(t)$, we conclude that

$$\xi_{i}(\tau, \Sigma) = \lim_{t \to \infty} F_{2}^{T} Y^{-1}(t) Z(t) Z^{-1}(\tau) F_{i} = F_{2}^{T} Z^{-1}(\tau) F_{i}$$

exists. Consequently,

$$Z^{-1}(\tau) = \begin{pmatrix} \bullet & \bullet \\ \xi_1(\tau, \Sigma) & \xi_2(\tau, \Sigma) \end{pmatrix},$$

implying

$$\xi_1(\tau, \Sigma) \cdot F_1^T Z(\tau) F_1 + \xi_2(\tau, \Sigma) \cdot F_2^T Z(\tau) F_1 = 0.$$

Again, this concludes the proof since we have $x \in S_{\Sigma}$ if and only if $x(n) = Z(n)F_1\alpha$ for some $\alpha \in \mathbb{C}^{p \times 1}$.

From (16), we obviously return to (13) by multiplying both sides by $\xi_2^{-1}(\tau, \Sigma)$. Again, it becomes clear that the method (13) still works in situations where the limits $\xi_i(\tau, \Sigma)$ do not exist and therefore (16) cannot be applied.

5 Equivalent backward computation

For the discrete problem (3), the Jacobi–Perron characterization can be exploited in a direct way (we only have to replace the limits by very large choices of t), whereas in the continuous-time case, the numerical calculations must be preceded by a suitable discretization.

Either way, algorithms should be efficient, and it seems natural to compute the values of $\eta(\tau, \Sigma)$ (or the $\xi_i(\tau, \Sigma)$) for all values $\tau \in [t_0, t_1] \cap I$ for an 'interval of interest' $[t_0, t_1]$. For this purpose, we demonstrate that the approximants of $\eta(\tau, \Sigma)$ or $\xi_i(\tau, \Sigma)$ can be computed by different kinds of 'backward procedures'.

5.1 A direct backward computation

In case of the strong asymptotic condition (5), the natural approach is to define $W^{(t)}$ as a solution to (3) or (4), respectively, with $W^{(t)}(t) = Y(t)F_1$, that is, $W^{(t)}(t) \in \mathbb{C}^{r \times p}$ consists of the first p columns of Y(t). For $t \to \infty$, we should expect $W^{(t)}(\tau) \to Z(\tau)$. Indeed, in this situation, we can write

$$W^{(t)}(\tau) = Z(\tau)Z^{-1}(t)W^{(t)}(t) \tag{17}$$

(note that we $W^{(t)}(\tau)$ and $Z(\tau)$ are solutions (3) or (4), respectively, as functions of τ , and $Z^{-1}(t)W^{(t)}(t)$ is constant with respect to τ), and since (5) guarantees that $Z^{-1}(t)W^{(t)}(t) = Z^{-1}(t)Y(t)F_1$ converges to $I_rF_1 = F_1$, the conjecture is true. In particular, in the discrete setting of (3), $W^{(t)}(\tau)$ can be easily computed (numerically) for $\tau \in [t_0, t_1]$, and for $t \gg t_1$, we get good approximations to $P(\tau) = Z(\tau)F_1$.

Next we demonstrate that a similar kind of backward procedure can be used to compute the approximants of $\eta(\tau, \Sigma)$ even in the case that the less restrictive condition (5) holds.

Theorem 5.1 Let $(\Sigma(t))_{t\in I}$ be a family of matrices $\Sigma(t) \in \mathbb{C}^{r \times (r-p)}$ with rank r-p, and let Y(t) be an invertible $r \times r$ -matrix for all $t \in I$ in such a way that $\Sigma^T(t) = F_2^T Y^{-1}(t)$. Furthermore, let $(W^{(t)}(\tau))_{\tau \in I}$ be a solution to (3) or (4), respectively, with $W^{(t)}(t) = Y(t)F_1$. Then we have

$$\left(\Sigma^{T}(t)\Xi(\tau,t)F_{2}\right)^{-1}\left(\Sigma^{T}(t)\Xi(\tau,t)F_{1}\right) = -\left(F_{2}^{T}W^{(t)}(\tau)\right)\left(F_{1}^{T}W^{(t)}(\tau)\right)^{-1}$$
(18)

for the approximants of $\eta(\tau, \Sigma) \in \mathbb{C}^{(r-p)\times p}$.

Proof We write $Y^{-1}(t) = {H^{T}(t) \choose \Sigma^{T}(t)}$. (17) holds for any solution Z with invertible Z(t). In particular, if we choose $Z(t) = \Xi(\tau, t)$, we obtain

$$\begin{split} W^{(t)}(\tau) &= \Xi^{-1}(\tau,t)W^{(t)}(t) = \left(Y^{-1}(t)\Xi(\tau,t)\right)^{-1}F_1 \\ &= \left(\begin{pmatrix} H^T(t) \\ \Sigma^T(t) \end{pmatrix} \left(\Xi(\tau,t)F_1 & \Xi(\tau,t)F_2 \right)\right)^{-1}F_1 \end{split}$$

$$= \begin{pmatrix} H^T(t)\Xi(\tau,t)F_1 & H^T\Xi(\tau,t)F_2 \\ \Sigma^T(t)\Xi(\tau,t)F_1 & \Sigma^T(t)\Xi(\tau,t)F_2 \end{pmatrix}^{-1} F_1$$

$$\stackrel{(1)}{=} \begin{pmatrix} V(\tau,t) \\ -(\Sigma^T(t)\Xi(\tau,t)F_2)^{-1}(\Sigma^T(t)\Xi(\tau,t)F_1)V(\tau,t) \end{pmatrix},$$

where $V(\tau, t)$ is chosen according to (1), that is,

$$(V(\tau,t))^{-1} = H^{T}(t)\Xi(\tau,t)F_{1} - H^{T}\Xi(\tau,t)F_{2}$$
$$\cdot \Sigma^{T}(t)\Xi(\tau,t)F_{1}(\Sigma^{T}(t)\Xi(\tau,t)F_{2})^{-1}(\Sigma^{T}(t)\Xi(\tau,t)F_{1}),$$

which proves our assertion.

Theorem 5.1 suggests the following numerical methods.

Algorithm 5.1 Let (6) hold for (3).

- Find Y(n) with $\Sigma^{T}(n) = F_2^{T} Y^{-1}(n)$, set $W^{(t)}(t) = Y(t)F_1$ for t sufficiently large.
- Then compute $W^{(t)}(n)$ successively for $n = t 1, t 2, ..., t_0$ by backward calculation. For $t \gg n$, the proportion of $F_1^T W^{(t)}(n)$ and $F_2^T W^{(t)}(n)$ is a good approximation to the proportion of $F_1P(n)$ and $F_2P(n)$, where $(P(n))_{\tau \in I}$ is the $\mathbb{C}^{r \times p}$ -valued solution characterizing S_{Σ} .

Algorithm 5.2 Let (6) hold for (3).

- Set $W^{(t)}(t) = Y(t)F_1$ for t sufficiently large.
- Then compute $W^{(t)}(n)$ successively for $n = t 1, t 2, ..., t_0$ by backward calculation. For $t \gg \tau$, $W^{(t)}(n)$ is a good approximation for P(n).

5.2 Continued-fraction type scheme for η

According to Theorem 4.1, the limit $\eta(\tau, \Sigma)$ exists if and only if there is a fundamental system $(Z(t))_{t \in I}$ satisfying (11) and (12). Obviously, condition (11) does not depend on τ . Hence, there will be many situations in which $\eta(\tau, \Sigma)$ exists for all $\tau \in I$.

In this section, we derive a recursion scheme for computing approximations of $\eta(\tau, \Sigma)$ in the discrete-time setting. The advantage of this scheme is that approximations of $\eta(\tau, \Sigma)$ for different values of τ can be computed simultaneously. The special form of the recursive scheme justifies to refer to the values $\eta(\tau)$ as some kind of generalized matrix continued fraction.

In order to derive the recursion scheme for $\eta(\tau,t)$, we define the approximation

$$\eta^{(t)}(\tau, \Sigma) = \left(\Sigma^T(t)\Xi(\tau, t)F_2\right)^{-1} \left(\Sigma^T(t)\Xi(\tau, t)F_1\right)$$
$$= \left(F_2^T Y^{-1}(t)\Xi(\tau, t)F_2\right)^{-1} \left(F_2^T Y^{-1}(t)\Xi(\tau, t)F_1\right).$$

For $t \to \infty$, we have convergence to $\eta(\tau, \Sigma)$ by definition.

Theorem 5.2 With the notation introduced above and with $A(t) = {B(t) \choose D(t)} {C(t) \choose D(t)}$, that is, $B(t) = F_1^T A(t) F_1$, $C(t) = F_1^T A(t) F_2$, ..., we have

$$\eta^{(t)}(t, \Sigma) = -F_2^T Y(t) F_1 \cdot \left(F_1^T Y(t) F_1\right)^{-1}$$

for all $t \in I$ and

$$\eta^{(t)}(\tau,\Sigma) = \left(\eta^{(t)}(\tau+1,\Sigma)C(\tau) + E(\tau)\right)^{-1} \left(\eta^{(t)}(\tau+1,\Sigma)B(\tau) + D(\tau)\right)$$

for all $\tau, t \in I$ with $\tau < t$, provided that these inverses exist.

Proof For $\tau = t$, we have $\Xi(t, t) = I_r$ and hence

$$\eta^{(t)}(t, \Sigma) = (F_2^T Y^{-1}(t) F_2)^{-1} (F_2^T Y^{-1}(t) F_1).$$

Let $Y(t) = \begin{pmatrix} \tilde{B} & \tilde{C} \\ \tilde{D} & \tilde{E} \end{pmatrix}$. According to the inversion formula (2), we have $F_2^T Y^{-1}(t) F_1 = -\tilde{E}^{-1} \tilde{D} \tilde{V}$ and $F_2^T Y^{-1}(t) F_2 = \tilde{E}^{-1} + \tilde{E}^{-1} \tilde{D} \tilde{V} \tilde{C} \tilde{E}^{-1}$ where $\tilde{V} = (\tilde{B} - \tilde{C} \tilde{E}^{-1} \tilde{D})^{-1}$. Using this representation of \tilde{V} , it is easy to show that $(I_r + \tilde{D} \tilde{V} \tilde{C} \tilde{E}^{-1}) \tilde{D} = \tilde{D} \tilde{V} \tilde{B}$, and hence

$$\begin{split} \eta^{(t)}(t) &= - \big(\tilde{E}^{-1} + \tilde{E}^{-1} \tilde{D} \tilde{V} \tilde{C} \tilde{E}^{-1} \big)^{-1} \tilde{E}^{-1} \tilde{D} \tilde{V} = \big(I_r + \tilde{D} \tilde{V} \tilde{C} \tilde{E}^{-1} \big)^{-1} \tilde{D} \tilde{V} \tilde{B} \tilde{B}^{-1} \\ &= - \tilde{D} \tilde{B}^{-1}, \end{split}$$

which is exactly the statement for $\tau = t$.

Now let $\tau < t$ and write $Y^{-1}(t)\Xi(\tau,t) = {\delta_{\tau} \choose \delta_{\tau} \epsilon_{\tau}}$. By construction of $\Xi(\tau,t)$, we have $\Xi(\tau,t) = \Xi(\tau+1,t)A(\tau)$. Hence, we conclude that

$$\begin{split} \eta^{(t)}(\tau) &= \left(F_2^T Y^{-1}(t) \Xi(\tau, t) F_2\right)^{-1} \left(F_2^T Y^{-1}(t) \Xi(\tau, t) F_1\right) \left[= (\epsilon_\tau)^{-1} \delta_\tau \right] \\ &= \left(F_2^T Y^{-1}(t) \Xi(\tau + 1, t) A(\tau) F_2\right)^{-1} \left(F_2^T Y^{-1}(t) \Xi(\tau + 1, t) A(\tau) F_1\right) \\ &= \left(\left(\delta_{\tau+1}, \epsilon_{\tau+1}\right) \begin{pmatrix} C(\tau) \\ E(\tau) \end{pmatrix} \right)^{-1} \left(\left(\delta_{\tau+1}, \epsilon_{\tau+1}\right) \begin{pmatrix} B(\tau) \\ D(\tau) \end{pmatrix} \right) \\ &= \left(\delta_{\tau+1} C(\tau) + \epsilon_{\tau+1} E(\tau)\right)^{-1} \left(\delta_{\tau+1} B(\tau) + \epsilon_{\tau+1} D(\tau)\right) \\ &= \left(\epsilon_{\tau+1}^{-1} \delta_{\tau+1} C(\tau) + E(\tau)\right)^{-1} \left(\epsilon_{\tau+1}^{-1} \delta_{\tau+1} B(\tau) + D(\tau)\right) \\ &= \left(\eta^{(t)}(\tau + 1) C(\tau) + E(\tau)\right)^{-1} \left(\eta^{(t)}(\tau + 1) B(\tau) + D(\tau)\right), \end{split}$$

which completes the proof.

With the notations of Theorem 5.2, we obtain the following

Algorithm 5.3

- Choose t sufficiently large.
- Set $\eta^{(t)}(t,\sigma) = -F_2^T Y(t) F_1 \cdot (F_1^T Y(t) F_1)^{-1}$.
- For $n = t 1, t 2, ..., t_0$, compute

$$\boldsymbol{\eta}^{(t)}(\tau, \boldsymbol{\Sigma}) = \big(\boldsymbol{\eta}^{(t)}(\tau+1, \boldsymbol{\Sigma})C(\tau) + E(\tau)\big)^{-1} \big(\boldsymbol{\eta}^{(t)}(\tau+1, \boldsymbol{\Sigma})B(\tau) + D(\tau)\big).$$

• By means of (15), compute approximate values to $x(t_0 + 1), x(t_0 + 2), \ldots$

6 On the dichotomy of differential equations and their discretization schemes

For the differential equation (4), it is obvious to use Algorithms 5.1 and 5.2, where the numerical calculation of $W^{(t)}(\tau)$ for $\tau < t$ is performed by means of some discretization method. However, this discretization will effect the asymptotic behavior. Therefore,

- in some situations, we will have to replace $\Sigma(t)$ in algorithm 5.1 by a sequence adapted to the discretization method,
- in almost all situations, we will have to replace Y(t) in algorithm 5.2 by a sequence adapted to the discretization method.

Example 6.1 We consider the case r=1, although our results have little meaning in this situation since the space of solutions is one-dimensional and explicitly known. Nevertheless, it becomes clear that the asymptotic behavior of the solutions of x'(t) = A(t)x(t), $t \in I$, and the corresponding discretized system differ. Consider the differential equation $x'(t) = (\lambda + \phi(t))x(t)$, $t \in I = [0, \infty)$, for some constant $\lambda \in \mathbb{R}$ and some function $\phi(t) \in L^1$. Then the exact solution is given by

$$x(t) = e^{\lambda t + \int_0^t \phi(s) \, ds} x(0), \quad t > 0. \tag{19}$$

Now apply Euler's explicit method for computing approximations $x^{[h]}(n)$ to x(nh). Then

$$x^{[h]}(n+1) = (1 + (\lambda + (\phi(nh))h)x(nh), \quad n = 0, 1, 2, \dots,$$
(20)

implying

$$x^{[h]}(n) = \left(\prod_{k=0}^{n-1} \left(1 + \left(\lambda + \phi(kh)\right)h\right)\right) x(0), \quad n = 0, 1, 2, \dots$$
 (21)

From the general theory of discretization methods we know that

$$\lim_{\substack{h \to 0 \\ nh \to t}} x^{[h]}(n) = x(t),\tag{22}$$

since we apply a consistent method to a linear and therefore Lipschitz-continuous problem. Although the limit (22) is a key topic in numerical mathematics, we are interested in a somewhat different property: We want to compare the behavior of the exact solution x(t) for $t \to \infty$ with the behavior of $x^{[h]}(n)$ for $n \to \infty$.

Set $y(t) = e^{\lambda t}$ for $t \ge 0$. From (19), we see that

$$y^{-1}(t)x(t) = e^{\int_0^t \phi(s) \, ds} x(0), \quad t \ge 0.$$

Since $\phi \in L^1$, we observe that there is a solution x(t) with

$$\lim_{t\to\infty}y^{-1}(t)x(t)=1,$$

that is, (5) is satisfied for $x'(t) = (\lambda + \phi(t))x(t)$, $t \ge 0$, with $y(t) = e^{\lambda t}$, $t \ge 0$.

In the sequel, we consider the discretized system. Since $x^{[h]}(n)$ is an approximation to x(nh), we conclude that

$$\lim_{n \to \infty} y^{-1}(nh) x^{[h]}(n) = \lim_{n \to \infty} \frac{(1 + \lambda h)^n}{e^{\lambda h n}} \prod_{k=0}^{n-1} \frac{1 + (\lambda + \phi(kh))h}{1 + \lambda h} x(0)$$

$$= \lim_{n \to \infty} e^{n(\ln(1 + \lambda h) - \lambda h)} \prod_{k=0}^{n-1} \left(1 + \phi(kh) \cdot \frac{h}{1 + \lambda h} \right) x(0).$$

Under appropriate conditions (e.g. monotonicity), $\phi \in L^1$ ensures that $\sum_{k=0}^{\infty} \phi(kh)$ is absolutely convergent, so that the product on the right-hand side converges to a finite value for $n \to \infty$. Observe $\ln(1 + \lambda h) - \lambda h \neq 0$, implying that $y^{-1}(nh)x^{[h]}(n)$ either tends to 0 or ∞ . Hence, (5) is not satisfied for the discretized system with the sequence (y(nh)). Therefore, setting $W^{(N)}(N) = y(Nh)$ for some large N and compute $W^{(N)}(n)$ backwards by use of (20) will not give good approximations to x(nh) for the solution $x(t) \sim e^{\lambda t}$.

The above calculation already tells us what to do: We have to choose $y^{[h]}(n) = (1 + \lambda h)^n$. Then $(y^{[h]}(n))^{-1}x^{[h]}(n)$ converges to some value $\in \mathbb{R} \setminus \{0\}$, that is, it converges to 1 for an appropriate choice of $x^{[h]}(0)$. Hence, we could use Algorithm 5.2 in the following way: Set $W^{(N)}(N) = y^{[h]}(N)$, compute $W^{(N)}(n)$ backwards by means of (20). Then we obtain good approximations to $x^{[h]}(n)$.

Note that usually, there is no need to solve $x'(t) = (\lambda + \phi(t))x(t)$, $t \in I$, numerically, since we have the explicit solution given in (19). However, this behavior is typical also in situations where we have no explicit solution.

Of course, there are better discretization methods (in terms of consistency order, speed of convergence, ...) than the explicit Euler method. However, for any classical discretization method, the discretization error will increase (often exponentially) as $t \to \infty$. Therefore, we can never expect the discretized system to have the same asymptotic behavior as the original differential equation.

For the sake of simplicity, we restrict the discussion in this paper to:

- Single-step methods. The reason is that discretizing (4) with $A(t) \in \mathbb{C}^{r \times r}$ by means of a single-step method will result in a system $x^{[h]}(n+1) = A^{[h]}(n)x^{[h]}(n)$ with $A^{[h]}(n) \in \mathbb{C}^{r \times r}$ whereas multi-step methods will change (increase) the number of dimensions of the discretized system.
- Constant step size h. This is mostly done for the sake of simplicity of the formulas. This means that we approximate x(nh) by $x^{[h]}(n)$ where

$$x^{[h]}(n+1) = x^{[h]}(n) + hC^{[h]}(n)x^{[h]}(n) + hD^{[h]}(n)x^{[h]}(n+1), \quad n = 0, 1, 2, \dots$$

For explicit methods, $D^{[h]}(n) = 0$. With

$$A^{[h]}(n) = (I_r - hD^{[h]}(n))^{-1}(I_r + hC^{[h]}(n)),$$

we obtain $x^{[h]}(n+1) = A^{[h]}(n)x(n)$, $n = 0, 1, 2, \dots$ Popular examples are as follows:

- For the explicit Euler method, we have $C^{[h]}(n) = A(nh)$.
- For the implicit Euler method, we have $C^{[h]}(n) = 0$ and $D^{[h]}(n) = A((n+1)h)$.

• For the (implicit) trapezoidal rule, we have $C^{[h]}(n) = \frac{1}{2}A(nh)$ and $D^{[h]}(n) = \frac{1}{2}A((n+1)h)$.

7 Poincaré-type/Levinson-type systems of linear difference and differential equations

In this section, we consider the case where $A(\infty) := \lim_{t \to \infty} A(t)$ exists. Basic results on the asymptotic behavior of the solutions of linear difference and differential equations with almost constant coefficients are due to Poincaré [22] and Perron [23–25]. A milestone in the analysis of such differential equations is due to Levinson [2]. Therefore, literature refers to these systems as Poincaré-type equations, Poincaré-Perron-type equations, or Levinson-type equations (in particular, in the continuous-time setting). Levinson-type results refer to statements on the asymptotic behavior of the solutions of systems (3) or (4) for which we can write

$$A(t) = A(\infty) + V(t) + R(t),$$

where R(t) is summable/integrable and $V(t) \to 0$ with some additional constraints (e.g. V'(t) should be summable/integrable with some conditions on the eigenvalues of $A(\infty) + V(t)$, the latter conditions are referred to as dichotomy conditions). We refer to the literature (e.g. [2, 26-28]) for results in this setting or similar settings. An exact translation of Levinson's ideas to the discrete-time setting is due to Benzaid and Lutz [29], and therefore, such results are often referred to as Benzaid-Lutz-type results (or sometimes, Levinson-Benzaid-Lutz-type results). A more recent publication in this direction is [30] where further remarks on literature can be found. Note that our focus is not on deriving such results but using these results and finding numerical methods for computing solutions with prescribed asymptotic behavior. Therefore, referring to all literature in this direction is far beyond the scope of this paper, and furthermore, we will only consider the 'simple' situation where V(t) = 0.

7.1 Difference equations

In the discrete-time setting of (3), let $A(t) = A(\infty) + R(t)$ where

- $A(\infty) = BDB^{-1}$ with $D = \operatorname{diag}(\lambda_1, \dots, \lambda_r)$ and $B = (b_1, \dots, b_r)$,
- $\sum_{t=t_0}^{\infty} \|R(t)\| < \infty$ for some $t_0 \in \mathbb{N}$.

By means of [29] or Theorem 8.25 in [31], we conclude that there is a fundamental system of vector-valued solutions to (3), say $x^{(1)}, \ldots, x^{(r)}$, with $x^{(i)}(t) = (b_i + o(1))\lambda_i^t$ for $i = 1, \ldots, r$.

The asymptotic representation of the solutions is equivalent to $\lim_{n\to\infty} z_i \lambda_i^{-n} = b_i$, and in matrix notation, we conclude that there is a solution Z = (Z(n)) with $Z(n) \in \mathbb{C}^{r \times r}$ for $n \in \mathbb{N}_0$ with

$$\lim_{n \to \infty} Z(n) \operatorname{diag}(\lambda_1^{-n}, \dots, \lambda_r^{-n}) = B.$$

Remark 7.1 By setting $Y(t) = B \operatorname{diag}(\lambda_1^t, \dots, \lambda_r^t)$, we find that $Z(t)Y^{-1}(t) \to I$. Unfortunately, this is not equivalent to our strong condition (5) which requires $Y^{-1}(t)Z(t) \to I$. There are some exceptions, e.g. if all eigenvalues of $A(\infty)$ have absolute value 1 since in that situation, both Y(t) and $Y^{-1}(t)$ are bounded and we can argue that

$$Y^{-1}(t)Z(t) = Y^{-1}(t)(Z(t)Y^{-1}(t))Y(t) \rightarrow Y^{-1}(t)IY(t) = I.$$

7.2 Differential equations

Now consider (4) with $A(t) = A(\infty) + R(t)$ where

- $A(\infty) = BDB^{-1}$ with $D = \operatorname{diag}(\lambda_1, \dots, \lambda_r)$ and $B = (b_1, \dots, b_r)$,
- $\int_0^\infty ||R(t)|| dt < \infty$.

By Levinson-type criteria (e.g. [2, 26, 28]), it is guaranteed that there is a fundamental system Z(t) of (4) with

$$\lim_{t \to \infty} Z(t) \operatorname{diag}\left(e^{-\lambda_1 t}, \dots, e^{-\lambda_r t}\right) \to B. \tag{23}$$

Let us consider a special case: Let u(t) satisfy the scalar rth-order linear differential equation

$$u^{(r)}(t) - \sum_{k=0}^{r-1} p_k(t)u^{(k)}(t) = 0, \quad t \in I,$$

with $p_0(t) \neq 0$ for all $t \in I$ and assume that the limits $p_k = \lim_{t \to \infty} p_k(t)$ exist. Then the vector $x(t) = (u(t), u^{(1)}(t), \dots, u^{(r-1)}(t))^T$ satisfies

$$x'(t) = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ p_0(t) & p_1(t) & \cdots & p_{r-2}(t) & p_{r-1}(t) \end{pmatrix} x(t) =: A(t)x(t), \quad t \in I.$$

Then

$$A(\infty) = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & & \\ & & \ddots & \ddots & \\ p_0 & p_1 & \cdots & p_{r-2} & p_{r-1} \end{pmatrix}$$

and the eigenvector corresponding to the eigenvalue λ_j has the form $(1, \lambda_j, \dots, \lambda_j^{r-1})^T$. Hence, the solution associated with this eigenvector satisfies $\lim_{t\to\infty} \frac{u'(t)}{u(t)} = \lambda_j$.

Remark 7.2 Again, setting $Y(t) = B \operatorname{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_r t})$ does not imply that our strong condition (5) is met. Again, there are some exceptions, e.g. for $\operatorname{Re}(\lambda_j) = 0$ for $j = 1, \dots, r$, both Y(t) and $Y^{-1}(t)$ are bounded and we can argue as in the discrete-time setting.

7.3 Discretization

We give some instructions concerning the discretization of Poincaré-type differential equations. As above, we focus on single-step methods with constant step size h. Since these methods rely on the idea of replacing the integral at the right-hand side of

$$x((n+1)h) = x(nh) + \int_{nh}^{(n+1)h} A(t)x(t)$$

by an appropriate summation formula, it becomes apparent that however we calculate, we finally arrive at

$$A^{[h]}(n) = I_r + h(C^{[h]}(n) + D^{[h]}(n)) + o(h) = I_r + hA(nh) + o(h).$$
(24)

Since $A^{[h]}(n)$ tends to some $A^{[h]}(\infty)$ by assumption, we conclude from (24) that $A^{[h]}(\infty) = I_r + hA(\infty)$ has eigenvalues $\mu_j^{[h]}(n) = 1 + h\lambda_j + o(h)$. Applying the Levinson-type theorem for the discrete setting, we note that there is a fundamental matrix $Z^{[h]}(n)$ with

$$\lim_{n \to \infty} Z^{[h]}(n) \operatorname{diag}\left(\left(\mu_1^{[h]}\right)^{-n}, \dots, \left(\mu_r^{[h]}\right)^{-n}\right) \to B^{[h]},\tag{25}$$

provided that R(nh) satisfies the summability condition. For example, if R(t) is (componentwise) monotone, the integrability of R(t) implies summability of R(nh). Due to

$$(I_r + hA(\infty) + o(h))B^{[h]} = B^{[h]} \operatorname{diag}(1 + h\lambda_1 + o(h), \dots, 1 + h\lambda_r + o(h)),$$

we have $\lim_{h\to 0} B^{[h]} = B$. More precisely, for all discretization methods discussed above, $A^{[h]}(\infty)$ may be decomposed into I_r and $A(\infty)$, and hence, the eigenvectors of $A(\infty)$ coincide with the eigenvectors of $A^{[h]}(\infty)$ (but with different eigenvalues), that is, $B^{[h]} = B$.

Hence, if we want to compute solutions generated by the first p columns of the fundamental system Z(t) of the differential equation, a good approximation can be obtained by computing solutions generated by the first p columns of the fundamental system $Z^{[h]}(n)$ of the discretized system. For this purpose, we need to compute the corresponding eigenvalues $\mu_1^{[h]}, \ldots, \mu_r^{[h]}$.

We give some examples. Let $D = \operatorname{diag}(\lambda_1, \dots, \lambda_r)$, then we have $A(\infty)B = BD$ and $B^{-1}A(\infty) = DB^{-1}$.

- For the explicit Euler method, we have $A^{[h]}(\infty) = I + hA(\infty)$. Therefore, we obtain $A^{[h]}B = B + hA(\infty)B = (I + hd)B$, that is, $\mu_i = 1 + h\lambda_i$ for i = 1, ..., r.
- For the implicit Euler method, we have $A^{[h]}(\infty) = (I hA(\infty))^{-1}$, and hence

$$B^{-1}A^{[h]}(\infty) = B^{-1}(I_r - hA(\infty))^{-1} = (B - hA(\infty)B)^{-1}$$
$$= (B - hBD)^{-1} = (I - hD)^{-1}B^{-1},$$

that is, $\mu_j = \frac{1}{1-h\lambda_j}$ for j = 1, ..., r.

• For the trapezoidal rule, we have $A^{[h]}(\infty) = (I - \frac{h}{2}A(\infty))^{-1}(I_r + \frac{h}{2}A(\infty))$, and we obtain

$$\begin{split} B^{-1}A^{[h]}(\infty)B &= \left(B - \frac{h}{2}A(\infty)B\right)^{-1}\left(B + \frac{h}{2}A(\infty)B\right) \\ &= \left(B - \frac{h}{2}DB\right)^{-1}\left(B + \frac{h}{2}DB\right) \\ &= \left(I_r - \frac{h}{2}D\right)^{-1}\left(I + \frac{h}{2}D\right). \end{split}$$

Hence, we have $\mu_j = \frac{1 + \frac{h}{2} \lambda_j}{1 - \frac{h}{2} \lambda_j}$ for $j = 1, \dots, r$.

• For Runge-Kutta-4, we have

$$B^{-1}A^{[h]}(\infty)B$$

$$= B^{-1}\left(I_r + hA(\infty) + \frac{5}{6}h^2A^2(\infty) + \frac{1}{2}h^3A^3(\infty) + \frac{1}{6}h^4A^4(\infty)\right)B$$

$$= I_r + hD + \frac{5}{6}h^2D^2 + \frac{1}{2}h^3D^3 + \frac{1}{6}h^4D^4,$$

that is,
$$\mu_j=1+h\lambda_j+\frac{5}{6}h^2\lambda_j^2+\frac{1}{2}h^3\lambda_j^3+\frac{1}{6}h^4\lambda_j^4$$
.

Remark 7.3 Although each solution of the discretized system corresponds to a solution of the original differential equation, it is worth noting that for fixed step size h, the dominance–subdominance relationship between the solutions may change. As a simple example, consider the continuous-time setting where r = 2, $A(\infty) = B \operatorname{diag}(-1, -100)B^{-1}$ with all entries of B being $\neq 0$. Then, for $\Sigma^T = (0,1)$ (resulting in the classical dominance and subdominance term), the solution corresponding to the eigenvalue -1 dominates over the solution corresponding to the eigenvalue -100. As a simple discretization scheme, choose the explicit Euler method with step size h = 0.1. For the discretized system, the eigenvalues are $1 + h \cdot (-1) = 0.9$ and $1 + h \cdot (-100) = -9$, where the latter eigenvalue corresponds to the original eigenvalue -100. Obviously, for the discrete system, the solutions corresponding to the eigenvalue 0.9 are subdominant.

This problem is well-known in the context of initial value problems: If the initial values are chosen in such a way that the desired solution is the solution corresponding to the eigenvalue -1, the explicit Euler method and many other discretization schemes require a very small step size h since otherwise the solution corresponding to the eigenvalue -100 becomes dominant and will cause massive numerical deviations. This is a well-known effect which belongs to the phenomenon of *stiffness* of differential equations.

In order to avoid the effects of stiffness, it is often recommended to use implicit discretization methods. Indeed, when applying the implicit Euler method in the situation sketched above, we would obtain the new eigenvalues $\frac{1}{1-h\cdot(-1)} = \frac{10}{11}$ and $\frac{1}{1-h\cdot(-100)} = \frac{1}{11}$, that is, the dominance–subdominance relationship of the two solutions is preserved.

In some way, this is not entirely true for all applications of implicit discretization schemes. Think of the situation where $A(\infty)$ admits the eigenvalues +1 and +100. Then the corresponding eigenvalues of the implicit Euler method with step size h=0.1 are $\frac{1}{1-h\cdot 1}=\frac{10}{9}$ and $\frac{1}{1-h\cdot 100}=-\frac{1}{9}$. Hence, the apparently dominant solution (eigenvalue 100) corresponds to the subdominant solution of the discretized system (eigenvalue $-\frac{1}{9}$). Of course, in this situation, the step size h=0.1 is too large for aspects of accuracy.

Anyway, discussing all aspects of stiffness and stability definitions from a numerical point of view is far beyond the scope of this paper since we cannot claim to answer all questions related with them. This remark simply intends to draw attention to the fact that when dealing with solutions of differential equations with prescribed asymptotic behavior numerically, one should carefully study the properties of the resulting discrete system since it is not guaranteed that the dominance–subdominance relationship is preserved during the discretization process.

8 Numerical examples

8.1 A Poincaré-type difference equation

Let D = diag(-2i, 2i, -2, 2),

$$B = \begin{pmatrix} -1+5i & -1-5i & 0 & -2\\ 1+3i & 1-3i & 1 & -4\\ -5+i & -5-i & 0 & 0\\ 10 & 10 & 1 & 1 \end{pmatrix},$$

$$A(\infty) = BDB^{-1} = \frac{1}{27} \begin{pmatrix} 19 & 14 & -83 & -14 \\ 11 & 28 & -193 & -82 \\ 65 & -26 & 23 & 26 \\ -94 & 16 & -118 & -70 \end{pmatrix}.$$

Furthermore, let $A(t) = A(\infty) + R(t)$ for $t \in \mathbb{N}_0$ where $\sum_{t=0}^{\infty} \|R(t)\| < \infty$. The general remarks concerning linear differential equations with almost constant coefficients imply that $Z(t)Y^{-1}(t) \to I$ where

$$Y(t) = Be^{D} = \begin{pmatrix} -1+5i & -1-5i & 0 & -2\\ 1+3i & 1-3i & 1 & -4\\ -5+i & -5-i & 0 & 0\\ 10 & 10 & 1 & 1 \end{pmatrix} \cdot \operatorname{diag}((-2)^{t}, 2^{t}), \quad t = 0, 1, 2, \dots$$

Define $\tilde{Y}(t) = \frac{1}{2^t} Y(t)$. Then $\tilde{Y}(t)$ and $(\tilde{Y}(t))^{-1}$ are bounded. Hence, we obtain

$$\begin{split} Y^{-1}(t)Z(t) &= \left(\tilde{Y}(t)\right)^{-1}\tilde{Y}(t)Y^{-1}(t)Z(t)Y^{-1}(t)Y(t)\left(\tilde{Y}(t)\right)^{-1}\tilde{Y}(t) \\ &= \left(\tilde{Y}(t)\right)^{-1}\cdot 2^t\cdot Z(t)Y^{-1}(t)\cdot \frac{1}{2^t}\cdot \tilde{Y}(t) \\ &= \left(\tilde{Y}(t)\right)^{-1}\left(Z(t)Y^{-1}(t)\right)\tilde{Y}(t) \\ &\to \left(\tilde{Y}(t)\right)^{-1}\cdot I\cdot \tilde{Y}(t) = I, \quad t = 0, 1, 2, \dots, \end{split}$$

that is, our strong assumption (5) is satisfied. For this reason, we may apply all algorithms developed in this paper.

We start with testing our algorithms by setting R(t) = 0. Then $Z(t) = Y(t)B^{-1}Z(0)$ for all solutions Z and all $t \in \mathbb{N}_0$. This means that each linear combination of the columns of Y(t) provides a solution of the underlying differential equation. We slightly modify Y(t):

• Let us change the order of the columns such that $Y(t)F_1$ represents the latter two columns of the original matrices Y(t), that is,

$$Y(t)F_1 = \begin{pmatrix} 0 & -2 \\ 1 & -4 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} \cdot \operatorname{diag}((-2)^t, 2^t).$$

With initialization time t = 100, we apply Algorithm 5.2, that is, we set $W^{(100)}(100) = Y(100)F_1$ and compute $W^{(100)}(n)$ for n = 99, 98, ..., 1, 0. We recognize that all numerically computed values coincide with the true entries of $Y(n)F_1$.

- With the same definition of $Y(t)F_1$ as before, we want to apply Algorithm 5.3, that is, we set $\eta(t) = -\binom{0\ 0}{1\ 1}\binom{0\ -2}{1\ -4}^{-1} = \binom{0\ 0}{-\frac{5}{2}\ 1}$ for some large t (here, t=100), and compute $\eta(n)$ by the continued-fraction-type scheme for $n=t-1,t-2,\ldots,0$. Due to R(t)=0, the exact proportions $\eta(n)$ do not depend on n. Indeed, Algorithm 5.3 reproduces the exact (and constant) matrices $\eta(n)$.
- Next, we are interested in the subspace spanned by the two solutions corresponding to the eigenvalues -2i, 2i. In order to obtain a real-valued solution, we set

$$Y(t)F_1 = \begin{pmatrix} -2 & 10\\ 2 & 6\\ -10 & 2\\ 20 & 0 \end{pmatrix} \operatorname{diag}(2^t, 2^t)$$

for all $t \in \mathbb{N}_0$ which are divisible by four. We apply the backward procedure 5.2 with t = 100. Again, the algorithm reproduces the exact solution.

• Finally, we use $Y(t)F_1$ as defined in the last point, and we set $\eta(100) = -\left(\frac{-10}{20} \frac{-2}{0}\right)\left(\frac{-2}{2} \frac{10}{6}\right)^{-1} = \left(\frac{-2}{15} \frac{3}{-25}\right)$ for applying Algorithm 5.3. It turns out that the algorithm reproduces the exact values (e.g. $\eta(n) = \eta(100)$ for all n which are divisible by four) in this situation, too.

Now let us consider

$$R(t) = \begin{pmatrix} 0 & 2e^{-t} & 0 & -\frac{5}{2t^2+2} \\ 0 & 0 & 0 & 0 \\ \frac{3}{2t^3+2} & 0 & 0 & 0 \\ 0 & 0 & -\frac{3}{2(t+1)\ln(t+2)} & 0 \end{pmatrix}.$$

Obviously, $\sum ||R(t)||$ converges. Suppose we are interested in the solutions corresponding to the real eigenvalues 2, -2. As in the case for R(t) = 0, we set

$$W^{(100)}(100) = \begin{pmatrix} 0 & -2 \\ 1 & -4 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} \cdot \operatorname{diag}((-2)^t, 2^t)$$

and apply Algorithm 5.2 or we set $\eta(100) = \binom{0}{-\frac{5}{2}} \binom{0}{1}$ and apply Algorithm 5.3. In Table 1, we have listed some numerical results. Although we are not in a position to find the exact solution to (3) in the given situation, we remark that the results of both algorithms 'fit together' in the sense that we have $\eta^{(100)}(0) = -(F_2^T W^{(100)}(0))(F_1^T W^{(100)}(0))$.

Since both algorithms lead to the same results, it becomes clear that using the continued-fraction type scheme in Algorithm 5.3 has some advantages: Whereas the growth/decay values of $W^{(100)}(n)$ are clearly influenced by the absolute value of the eigenvalues, this is not true for the values of $\eta^{(t)}(n)$. Hence, computing $\eta^{(t)}(n)$ does not require any 'scaling technique' and Algorithm 5.3 can be applied with much larger t. E.g. for t=1000, we obtain $\eta^{(1000)}(0)=\left(\frac{-1.5504}{1.0773},\frac{0.6340}{-1.3982}\right)$.

Table 1 Poincaré-type linear system of difference equations: Numerical results obtained by Algorithm 5.2 and Algorithm 5.3

n	$W^{(100)}(n)$	$\eta^{(100)}(n)$
100	$\begin{pmatrix} 0 & -2.535 \cdot 10^{30} \\ 1.268 \cdot 10^{30} & -5.071 \cdot 10^{30} \\ 0 & 0 \\ 1.268 \cdot 10^{30} & 1.268 \cdot 10^{30} \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ -2.5 & 1 \end{pmatrix}$
99	$\begin{pmatrix} 4.632 \cdot 10^{25} & -1.268 \cdot 10^{30} \\ -6.339 \cdot 10^{29} & -2.535 \cdot 10^{30} \\ 9.725 \cdot 10^{25} & -9.767 \cdot 10^{25} \\ -6.341 \cdot 10^{29} & 6.341 \cdot 10^{29} \end{pmatrix}$	(-0.0004 0.0002 2.5003 -1.0001)
2	(-0.0937 -14.4523) 6.0451 -15.9751 -1.1255 -9.9137 8.1103 27.0848)	(-0.8767 0.1726) 3.3005 -1.2905)
1	(-0.8344 -3.8417) -3.1130 -31.8064 -1.1025 15.9212 -0.5705 -43.5892)	(-5.8040 1.2016) (8.0617 -2.3442)
0	(-0.4438 3.6676 -4.8386 -3.5192 2.3796 7.9174 -6.2876 -8.8716	(-1.5504 0.6340 1.0772 -1.3983)

8.2 Ordinary subdominance for Poincaré-type differential equations

The computation of subdominant solutions in the ordinary sense is a special case of our approach. By 'ordinary sense' we mean that we consider a scalar rth order linear differential equation, and that we say a subspace $\mathcal S$ to be subdominant if $\frac{u(t)}{v(t)} \to 0$ for every $u \in \mathcal S$ and every solution $v \notin \mathcal S$.

Consider the scalar differential equation

$$u'''(t) = -2u''(t) + u'(t) + 2(1 + \phi(t))u(t), \quad t \ge 0,$$

with some differentiable function $\phi(t)$ converging to 0 monotonically. (The assumption of monotonicity is chosen for the sake of simplicity, it could be replaced by a weaker one.) With the above construction, we have

$$A(t) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2(1 + \phi(t)) & 1 & -2 \end{pmatrix} = A(\infty) + V(t),$$

where the monotonicity ensures that $V'(t) \in L^1$. $A(\infty)$ has the eigenvalues $\lambda_1 = -2$, $\lambda_2 = -1$ and $\lambda_3 = 1$. According to (23), there is a fundamental system of solutions $x^{(1)}$, $x^{(2)}$, $x^{(3)}$ such that the asymptotic growth of $x^{(i)}$ is expressed by λ_i .

First, assume that we want to compute a solution which is a multiple of $x^{(1)}$. Then we set p=1 and $\Sigma^T(t)=F_2^TB^{-1}$. With an arbitrary $\lambda^*\in(\lambda_1,\lambda_2)=(-2,-1)$, we can write

$$\begin{split} & \left(\Sigma^T(t)Z(t)F_2\right)^{-1}\left(\Sigma^T(t)Z(t)F_1\right) \\ & = \left(F_2^TB^{-1}Z(t)\operatorname{diag}\left(e^{-\int_{t_1}^t\lambda_j(s)\,ds}\right)\operatorname{diag}\left(e^{\int_{t_1}^t\lambda_j(s)\,ds}\right)e^{-\lambda^*t}F_2\right)^{-1} \\ & \quad \cdot \left(F_2^TB^{-1}Z(t)\operatorname{diag}\left(e^{-\int_{t_1}^t\lambda_j(s)\,ds}\right)\operatorname{diag}\left(e^{\int_{t_1}^t\lambda_j(s)\,ds}\right)e^{-\lambda^*t}F_1\right) \end{split}$$

$$= (F_2^T (I_r + o(1)) \operatorname{diag}(e^{(\lambda_j - \lambda^*)t + o(t)}) F_2)^{-1} \cdot (F_2^T (I_r + o(1)) \operatorname{diag}(e^{(\lambda_j - \lambda^*)t + o(t)}) F_1).$$

Since $\lambda_j - \lambda^* > 0$ for j = 2, 3 and $\lambda_1 - \lambda^* < 0$, this term obviously converges to 0, that is, the main condition (11) of Theorem 4.1 is satisfied.

For a thorough numerical computation, we have to use some discretization method. The monotonicity of $\phi(t)$ ensures that (V(nh)) converges to 0 in a (componentwise) monotonic sense and therefore, (25) holds. Since $B^{[h]}=B$, we choose $(\Sigma^{[h]}(t))^T=F_2^TB^{-1}$. Then we can easily prove that (11) is satisfied for the discretized system; in the above calculation for the continuous case, we only have to replace $e^{-\lambda^*t}$ by $(\mu^*)^{-n}$ with some $\mu^*\in (1+h\lambda_1,1+h\lambda_2)$.

For our numerical experiment, we set $\phi(t) = \frac{1}{1+t^2}$, h = 0.001 and start our backward procedure (for the discretized system) at $t = 20 = 2 \cdot 10^4 \cdot h$. For the discretization, we consider the explicit Euler method (EE) and the trapezoidal rule (TR). The results are listed in Table 2. We have included the values $\frac{u_1(t)-u_1(t-h)}{hu_1(t-h)}$ as approximations to $\frac{u_1'(t)}{u_1(t)}$. The results clearly indicate that we calculate the solution u_1 for which $\frac{u_1'(t)}{u_1(t)}$ approaches -2 as $t \to \infty$.

Next, assume that we want to compute solutions which are linear combinations of $x^{(1)}$ and $x^{(2)}$. Hence, we set p=2 and $\Sigma^T(t)=F_2^TB^{-1}$. Again, (11) is satisfied not only for the differential equation but also for its discretized version. In addition to the above parameters for p=1, we compute a solution $u_1(t)$ with $u_1(0)=1$ and $u_1'(0)=2$. The results are documented in Table 3, and they clearly indicate that here, $\frac{u_1'(t)}{u_1(t)}$ approaches -1 as $t\to\infty$. This is completely in accordance with our goals.

Table 2 Poincaré-type differential equation: Numerical calculation of the one-dimensional subspace defined by the subdominant solution

t	EE: <i>u</i> ₁ (<i>t</i>)	EE: $\frac{u_1(t)-u_1(t-h)}{hu_1(t-h)}$	TR: <i>u</i> ₁ (<i>t</i>)	TR: $\frac{u_1(t)-u_1(t-h)}{hu_1(t-h)}$
0	1	inf	1	inf
0.001	0.998403	-1.59697	0.998403	-1.59677
0.002	0.996808	-1.59737	0.996809	-1.59717
0.003	0.995216	-1.59776	0.995216	-1.59757
0.004	0.993625	-1.59816	0.993626	-1.59796
0.005	0.992037	-1.59856	0.992038	-1.59836
0.006	0.99045	-1.59896	0.990452	-1.59876
0.007	0.988866	-1.59935	0.988868	-1.59915
0.008	0.987284	-1.59975	0.987286	-1.59955
0.009	0.985705	-1.60015	0.985706	-1.59994
0.01	0.984127	-1.60054	0.984129	-1.60033
0.1	0.85065	-1.63506	0.850673	-1.6347
1	0.174041	-1.85137	0.174195	-1.84999
3	0.00370203	-1.96093	0.00371804	-1.95906
4.99	7.25945e-005	-1.98241	7.31878e-005	-1.98047
4.991	7.24506e-005	-1.98242	7.30429e-005	-1.98047
4.992	7.23069e-005	-1.98243	7.28982e-005	-1.98048
4.993	7.21636e-005	-1.98243	7.27538e-005	-1.98048
4.994	7.20205e-005	-1.98244	7.26097e-005	-1.98049
4.995	7.18778e-005	-1.98244	7.24659e-005	-1.9805
4.996	7.17353e-005	-1.98245	7.23224e-005	-1.9805
4.997	7.15931e-005	-1.98246	7.21792e-005	-1.98051
4.998	7.14511e-005	-1.98246	7.20362e-005	-1.98051
4.999	7.13095e-005	-1.98247	7.18936e-005	-1.98052
5	7.11681e-005	-1.98247	7.17512e-005	-1.98052

t	EE: <i>u</i> ₁ (<i>t</i>)	EE: $\frac{u_1(t)-u_1(t-h)}{hu_1(t-h)}$	TR: <i>u</i> ₁ (<i>t</i>)	TR: $\frac{u_1(t)-u_1(t-h)}{hu_1(t-h)}$
0	1	inf	1	inf
0.001	1.002	2	1.002	1.99533
0.002	1.00399	1.98668	1.00398	1.98206
0.003	1.00597	1.97346	1.00596	1.96888
0.004	1.00794	1.96033	1.00793	1.95579
0.005	1.00991	1.9473	1.00988	1.9428
0.006	1.01186	1.93436	1.01183	1.9299
0.007	1.0138	1.92151	1.01377	1.9171
0.008	1.01574	1.90875	1.0157	1.90438
0.009	1.01767	1.89608	1.01762	1.89175
0.01	1.01958	1.88351	1.01954	1.87921
0.1	1.15754	1.02648	1.15719	1.02459
1	0.891296	-0.844232	0.891478	-0.842916
3	0.110259	-1.09618	0.110554	-1.0952
4.99	0.0129359	-1.05457	0.0129902	-1.05395
4.991	0.0129222	-1.05455	0.0129765	-1.05394
4.992	0.0129086	-1.05453	0.0129628	-1.05392
4.993	0.012895	-1.05452	0.0129491	-1.0539
4.994	0.0128814	-1.0545	0.0129355	-1.05388
4.995	0.0128678	-1.05448	0.0129219	-1.05386
4.996	0.0128542	-1.05446	0.0129082	-1.05384
4.997	0.0128407	-1.05444	0.0128946	-1.05382
4.998	0.0128272	-1.05442	0.012881	-1.0538
4.999	0.0128136	-1.0544	0.0128675	-1.05378
5	0.0128001	-1.05438	0.0128539	-1.05376

Table 3 Poincaré-type differential equation: Computation of $u_1(t)$ for p=2

8.3 Knesers's differential equation

For computing subdominant solutions in the 'ordinary sense', we were allowed to choose $\Sigma^T(t) = (\Sigma^{[h]}(t))^T = F_2^T B^{-1}$ independently of the discretization scheme. Next, we demonstrate that there are situations in which the choice of $\Sigma^{[h]}(t)$ depends on the discretization scheme.

Consider Kneser's differential equation

$$u''(t) + (1 + \phi(t))u(t) = 0, \quad t \ge 0, \tag{26}$$

in which ϕ is continuous and real-valued with $\phi(t) \to 0$ 'sufficiently fast'. Kneser [32, 33] already investigated conditions under which there is a fundamental set of solutions [$u^{(1)}, u^{(2)}$] to equation (26) satisfying

$$u^{(1)}(t) = \sin(t)(1 + o(1))$$
 as $t \to \infty$, (27)

$$u^{(2)}(t) = \cos(t)(1 + o(1))$$
 as $t \to \infty$. (28)

Various conditions on the speed-of-convergence condition on $\phi(t)$ were discussed by Wintner [34]. The algebraic characterization of these solutions in terms of its initial values and its computation for small values of t remained an open problem.

For the sake of simplicity, let us assume $t_0=0$, $\phi(t)\to 0$ monotonically and $\phi(t)\in L^1$. Then we have $\phi((nh))\in \ell^1$ for all h>0 so that we can apply Levinson-type results to the underlying differential equation and Bensaid–Lutz-type results to its discretization scheme. Put $x(t)=\binom{u(t)}{u'(t)}$. Then x(t) satisfies (4) with $A(t)=\binom{0}{-(1+\phi(t))} \stackrel{1}{0}\to A(\infty)=\binom{0}{-1}$ and $A(t)-A(\infty)=\binom{0}{-\phi(t)} \stackrel{0}{0}\in L^1$. In addition, we have $A(\infty)B=BD$ for $B=\binom{1}{i} \stackrel{1}{-i}$ and $D=\binom{1}{i} \stackrel{1}{-i}$ and $D=\binom{1}{i} \stackrel{1}{-i}$ and $D=\binom{1}{i} \stackrel{1}{-i}$ and $D=\binom{1}{i} \stackrel{1}{-i}$

diag(i, -i). Applying (23), we conclude that there is a fundamental system Z(t) with $Z(t)Y^{-1}(t) \rightarrow I$ where

$$Y(t) = B \begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix} = \begin{pmatrix} e^{it} & e^{-it} \\ ie^{it} & -ie^{-it} \end{pmatrix}.$$

With $\tilde{B} = \begin{pmatrix} \frac{1}{2i} & \frac{1}{2} \\ -\frac{1}{2i} & \frac{1}{2} \end{pmatrix}$, set $\tilde{Z}(t) = Z(t)\tilde{B}$ and

$$\tilde{Y}(t) = Y(t)\tilde{B} = \begin{pmatrix} \sin(t) & \cos(t) \\ \cos(t) & -\sin(t) \end{pmatrix},$$

and obtain $\tilde{Z}(t)\tilde{Y}^{-1}(t)\to I$. Note that $\tilde{Y}^{-1}(t)=\tilde{Y}(t)$ and $\sup_{t\geq 0}||\tilde{Y}(t)||<\infty$. Hence,

$$\tilde{Y}^{-1}(t)\tilde{Z}(t) = \tilde{Y}^{-1}(t)(\tilde{Z}(t)\tilde{Y}^{-1}(t))\tilde{Y}(t) \rightarrow I$$
,

and therefore, (5) is satisfied, allowing the naive backward computation for computing the sinus-solution.

However, due to the need of discretization, we have to replace Y(t) by a sequence $(Y^{[h]}(n))$. For the discretized system, we have $A^{[h]}(n) = A^{[h]}(\infty) + R^{[h]}(n)$, and clearly $R^{[h]}(n) \in \ell^1$. As pointed out above, the eigenvalues of $A^{[h]}(\infty)$ depend on the discretization method. For example, for the explicit Euler method, we have the eigenvalues $\mu_j = 1 + h\lambda_j$, that is, $\mu_1 = 1 + ih$ and $\mu_2 = 1 - ih$. Hence, we choose

$$Y^{[h]}(n) = B \begin{pmatrix} \mu_1^n & 0 \\ 0 & \mu_2^n \end{pmatrix}^n \tilde{B}$$

$$= \begin{pmatrix} \text{Im}((1+ih)^n) & \text{Re}((1+ih)^n) \\ \text{Re}((1+ih)^n) & -\text{Im}((1+ih)^n) \end{pmatrix}.$$
(29)

For the trapezoidal rule, we have

$$Y^{[h]}(n) = B \begin{pmatrix} (\frac{1+i\frac{h}{2}}{1-i\frac{h}{2}})^n & 0 \\ 0 & (\frac{1-i\frac{h}{2}}{1+i\frac{h}{2}})^n \end{pmatrix} \tilde{B}$$

$$= B \begin{pmatrix} (\frac{1+ih-\frac{h^2}{4}}{1+\frac{h^2}{4}})^n & 0 \\ 0 & (\frac{1-ih-\frac{h^2}{4}}{1+\frac{h^2}{4}})^n \end{pmatrix} \tilde{B}$$

$$= \begin{pmatrix} \operatorname{Im}((\frac{1+ih-\frac{h^2}{4}}{1+\frac{h^2}{4}})^n) & \operatorname{Re}((\frac{1+ih-\frac{h^2}{4}}{1+\frac{h^2}{4}})^n) \\ \operatorname{Re}((\frac{1+ih-\frac{h^2}{4}}{1+\frac{h^2}{4}})^n) & -\operatorname{Im}((\frac{1+ih-\frac{h^2}{4}}{1+\frac{h^2}{4}})^n) \end{pmatrix}$$
(30)

for $n = 0, 1, 2, \ldots$ In order to assess the accuracy of the backward computation procedure, we apply it to those cases whose solutions we know. For this purpose, we choose

•
$$\phi^{(1)}(t) = 0$$
, implying $u^{(1)}(t) = \sin(t)$ and $u^{(2)}(t) = \cos(t)$ and

•
$$\phi^{(2)}(t) = -\frac{2}{(t+1)^2}$$
, implying
$$u^{(1)}(t) = \sqrt{\frac{1 + (t+1)^2}{(t+1)^2}} \cos(t - \arctan(t+1))$$

$$= \sqrt{\frac{1 + (t+1)^2}{(t+1)^2}} \sin(t - \arctan(t+1) + \frac{\pi}{2}), \quad t \ge 0, \quad \text{and}$$

$$u^{(2)}(t) = -\sqrt{\frac{1 + (t+1)^2}{(t+1)^2}} \sin(t - \arctan(t+1))$$

$$= \sqrt{\frac{1 + (t+1)^2}{(t+1)^2}} \cos(t - \arctan(t+1) + \frac{\pi}{2}), \quad t \ge 0,$$

which can be verified by standard calculations.

The results for $\phi^{(1)}$ can be found in Tables 4 and 5, the corresponding results for $\phi^{(2)}$ are listed in Tables 6 and 7. In all cases, we have chosen $h = 0.001 = 10^{-3}$, and the backward

Table 4 Kneser's differential equation: Computation of Sinus solution for $\phi(t) = 0$

t	u ₁ (t) exact	EE with $u_1(T) = \sin(T)$	Corrected EE	TR with $u_1(T) = \sin(T)$	Corrected TR
0	0	2.24598e-005	-3.58392e-013	0.000833333	1.93341e-013
0.001	0.001	2.91978e-005	0.001	0.00183333	0.001
0.002	0.002	3.59357e-005	0.002	0.00283333	0.002
0.003	0.003	4.26735e-005	0.003	0.00383332	0.003
0.004	0.00399999	4.94114e-005	0.004	0.00483331	0.00399999
0.005	0.00499998	5.61492e-005	0.00499999	0.0058333	0.00499998
0.006	0.00599996	6.28869e-005	0.00599998	0.00683328	0.00599996
0.007	0.00699994	6.96246e-005	0.00699997	0.00783325	0.00699994
0.008	0.00799991	7.63622e-005	0.00799994	0.00883322	0.00799991
0.009	0.00899988	8.30998e-005	0.00899992	0.00983317	0.00899988
0.01	0.00999983	8.98373e-005	0.00999988	0.0108331	0.00999983
0.05	0.0499792	0.000359197	0.0499804	0.0508114	0.0499792
0.1	0.0998334	0.000695052	0.0998384	0.100663	0.0998334
0.15	0.149438	0.00102919	0.149449	0.150262	0.149438
0.2	0.198669	0.00136077	0.198689	0.199486	0.198669
0.25	0.247404	0.00168896	0.247435	0.248211	0.247404

Table 5 Kneser's differential equation: Computation of Cosinus solution for $\phi(t) = 0$

t	u ₂ (t) exact	EE with $u_2(T) = \cos(T)$	Corrected EE	TR with $u_2(T) = \cos(T)$	Corrected TR
0	1	0.00673793	1	1	1
0.001	1	0.0067379	1	0.999998	0.999999
0.002	0.999998	0.00673787	0.999999	0.999996	0.999998
0.003	0.999996	0.00673784	0.999997	0.999993	0.999995
0.004	0.999992	0.0067378	0.999994	0.999988	0.999992
0.005	0.999988	0.00673775	0.99999	0.999983	0.999987
0.006	0.999982	0.00673769	0.999985	0.999977	0.999982
0.007	0.999976	0.00673763	0.999979	0.999969	0.999975
0.008	0.999968	0.00673756	0.999972	0.999961	0.999968
0.009	0.99996	0.00673748	0.999964	0.999952	0.999959
0.01	0.99995	0.0067374	0.999955	0.999941	0.99995
0.05	0.99875	0.00672855	0.998775	0.998708	0.99875
0.1	0.995004	0.00670236	0.995054	0.994921	0.995004
0.15	0.988771	0.00665941	0.988845	0.988646	0.988771
0.2	0.980067	0.00659981	0.980165	0.979901	0.980067
0.25	0.968912	0.00652372	0.969034	0.968706	0.968912

Table 6 Kneser's differential equation: Computation of Sinus solution for $\phi(t) = -1$	$-\frac{2}{(t+1)^2}$
---	----------------------

t	u ₁ (t) exact	EE with $u_1(T) = \sin(T)$	Corrected EE	TR with $u_1(T) = \sin(T)$	Corrected TR
0	1	0.00676985	1.0014	1.00073	0.9999
0.001	1	0.00676982	1.0014	1.00073	0.999901
0.002	1	0.0067698	1.0014	1.00073	0.999902
0.003	1	0.00676978	1.0014	1.00074	0.999905
0.004	1.00001	0.00676978	1.0014	1.00074	0.999908
0.005	1.00001	0.00676977	1.00141	1.00074	0.999913
0.006	1.00002	0.00676978	1.00141	1.00075	0.999919
0.007	1.00002	0.00676979	1.00142	1.00075	0.999925
0.008	1.00003	0.00676981	1.00142	1.00076	0.999933
0.009	1.00004	0.00676983	1.00143	1.00077	0.999941
0.01	1.00005	0.00676987	1.00144	1.00078	0.99995
0.05	1.00117	0.0067762	1.00251	1.00187	1.00107
0.1	1.00438	0.00679644	1.00567	1.00505	1.00429
0.15	1.00924	0.00682785	1.01048	1.00987	1.00915
0.2	1.01539	0.00686808	1.0166	1.01599	1.01531
0.25	1.02253	0.00691504	1.02371	1.0231	1.02246

Table 7 Kneser's differential equation: Computation of Cosinus solution for $\phi(t) = -\frac{2}{(t+1)^2}$

t	$u_2(t)$ exact	EE with $u_2(T) = \cos(T)$	Corrected EE	$TR \ u_2(T) = \cos(T)$	Corrected TR
0	1	0.00672286	1.0011	0.999267	1.0001
0.001	0.999	0.00671611	1.0001	0.998267	0.999101
0.002	0.998002	0.00670937	0.9991	0.997269	0.998102
0.003	0.997004	0.00670264	0.9981	0.996271	0.997105
0.004	0.996008	0.00669591	0.997102	0.995275	0.996108
0.005	0.995012	0.00668919	0.996104	0.994279	0.995113
0.006	0.994018	0.00668247	0.995108	0.993285	0.994118
0.007	0.993024	0.00667576	0.994112	0.992291	0.993125
0.008	0.992032	0.00666906	0.993118	0.991298	0.992132
0.009	0.99104	0.00666237	0.992124	0.990307	0.99114
0.01	0.990049	0.00665568	0.991131	0.989316	0.99015
0.05	0.951151	0.00639307	0.95216	0.950417	0.951252
0.1	0.904247	0.00607641	0.905175	0.90351	0.904347
0.15	0.858825	0.00576978	0.859682	0.858085	0.858926
0.2	0.814509	0.00547062	0.815303	0.813764	0.814611
0.25	0.770989	0.00517685	0.771727	0.770239	0.771092

computation starts at time $T=10^4=10^7 \cdot h$. 'EE' denotes the explicit Euler method, while 'TR' denotes the trapezoidal rule. In order to demonstrate the effect of choosing $Y^{[h]}(n)$ according to the method of discretization, we have computed all results twice: In the first case, the backward computation was initialized with $u_1(T)=\sin(T)$ and $u_2(T)=\cos(T)$, in the 'Corrected EE' and 'Corrected TR', the backward computation was initialized with the corresponding column of $Y^{[h]}(10^7)$ according to (29) and (30), respectively.

The tables clearly indicate that the corrected methods relying on initial values of the discretized problem work much better. In particular, even the explicit Euler method with these initial values yields very good results.

9 Conclusion and further research

As outlined in this paper the concept of Σ -subdominant solutions of linear systems of differential and difference equations, respectively, has a strong impact to a lot of classical mathematical problems. Our aim was to raise awareness of this problem not only from a theoretical but also from a practical point of view. It is therefore to be expected that even

more problems can be subjected to this context. As an example we mention the mathematical representation of regular solutions of complex-valued singular linear systems of differential equations. For the scalar case this problem has been addressed by Pincherle [19] and Perron [35], who linked it to the computation of certain ordinary and generalized continued fractions.

This paper is restricted to linear differential equations. But there is also a connection with nonlinear equations. For example, if $x(t) = {z(t) \choose z'(t)}$ satisfies

$$x'(t) = \begin{pmatrix} g(t) + \frac{f'(t)}{f(t)} & -f(t)h(t) \\ 1 & 0 \end{pmatrix} x(t), \quad t \geq t_0,$$

that is, $z''(t) = (g(t) + \frac{f'(t)}{f(t)})z'(t) - f(t)h(t)z(t)$ for $t \ge t_0$, the function $y(t) = -\frac{z'(t)}{f(t)z(t)}$ solves the Ricatti differential equation

$$y'(t) = f(t)y^{2}(t) + g(t)y(t) + h(t), \quad t \ge t_{0},$$

which means that our methods can also be applied to special cases of Ricatti differential (or difference) equations. But this is not surprising since the continued-fraction-type scheme in Theorem 5.2 reminds one of Ricatti difference equations.

Without further investigations we cannot principally exclude that our algorithm is subject to inherent numerical instability. For example, in the situation of Sect. 7, it is not trivial to find Y(t) or $\Sigma^T(t)$ such that the strong criterion (5) or the weaker condition (6) is met. A trivial exception is given for R(t) = 0 where we know the exact solutions. Numerical computations in this situation reveal that our algorithms might still be susceptible to numerical instabilities. However, a profound study of these effects cannot be performed before it is clarified how to choose $\Sigma^T(t)$ in this situation. Therefore, these questions need further research. In order to solve these problems, it might make sense to consider the adjoint systems of (3) and (4) (some considerations in the scalar case can be found in [3, 11, 35]).

Acknowledgements

We would like to thank the anonymous referees for their valuable comments.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Availability of data and materials

No additional data or material was used to support this study.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

The authors declare that the mathematical results were realized in collaboration with equal responsibility. Both authors read and approved the final manuscript.

Author details

¹Institute of Analysis, Leibniz-Universität Hannover, 30167 Hannover, Germany. ²Simulation Science Center Clausthal-Göttingen, Arnold-Sommerfeld-Str. 6, 38678 Clausthal-Zellerfeld, Germany.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 21 September 2020 Accepted: 5 March 2021 Published online: 19 March 2021

References

- 1. Karlin, S., Taylor, H.M.: A Second Course in Stochastic Processes. Academic Press, San Diego (1981)
- 2. Levinson, N.: The asymptotic nature of solutions of linear systems of differential equations. Duke Math. J. 15, 111–126 (1948)
- 3. Hanschke, T.: Charakterisierung scalar-subdominanter Lösungen von linearen Differenzen- und Differentialgleichungssystemen. Habilitationsschrift, Universität Mainz (1989)
- 4. Miller, K.S.: Linear Difference Equations. Benjamin, New York (1968)
- 5. Wimp, J.: Computation with Recurrence Relations. Pitman, Boston (1984)
- 6. Gautschi, W.: Computational aspects of three-term recurrence relations. SIAM Rev. 9, 24-82 (1967)
- 7. Gautschi, W.: Zur Numerik rekurrenter Relationen. Computing 9, 107–126 (1972)
- 8. Zahar, R.V.M.: A mathematical analysis of Miller's algorithm. Numer. Math. 27, 427–447 (1977)
- 9. Schäfke, F.W.: Lösungstypen von Differenzengleichungen und Summengleichungen in normierten abelschen Gruppen. Math. Z. **88**, 61–104 (1965)
- 10. Schäfke, F.W.: Minimallösungen von Differenzengleichungen in Gruppen: eine verallgemeinerte Kettenbruchmethode. Math. Z. 98, 52–59 (1967)
- 11. Hanschke, T.: Characterization of antidominant solutions of linear differential equations. J. Comput. Appl. Math. 40, 351–361 (1992)
- Hanschke, T.: Ein verallgemeinerter Jacobi–Perron-Algorithmus zur Reduktion linearer Differenzengleichungssysteme. Mh. Math. 126, 287–311 (1998)
- 13. Levrie, P., Bultheel, A.: Matrix continued fractions related to first-order linear recurrence systems. Electron. Trans. Numer. Anal. 4, 46–63 (1996)
- Hanschke, T.: A matrix continued fraction algorithm for the multiserver repeated order queue. Math. Comput. Model. 30, 159–170 (1999)
- 15. Baumann, H., Hanschke, T.: Computation of invariant measures and stationary expectations for Markov chains with block-band transition matrix. J. Appl. Math. 2020, Article ID 4318906 (2020)
- Perron, O.: Über die Konvergenz der Jacobi-Kettenalgorithmen mit komplexen Elementen. Sitzungsber. Akad. München Math.-Phys. 37, 401–482 (1907)
- 17. Zhang, F.: Matrix Theory. Springer, New York (1999)
- 18. Baumann, H.: Generalized continued fractions: a unified definition and a Pringsheim-type convergence criterion. Adv. Differ. Equ. 2019, 406 (2019)
- 19. Pincherle, S.: Sur la génération de systèmes récurrents au moyen d'une équation linéaire différentielle. Acta Math. 16, 341–363 (1892)
- 20. Pincherle, S.: Delle funzioni ipergeometriche e di varie questioni ad esse attinenti. G. Mat. Battaglini 32, 209–291 (1894)
- 21. Cruyssen, P.: Linear difference equations and generalized continued fractions. Computing 22, 269–278 (1979)
- 22. Poincaré, H.: Sur les équations linéaires aux différentielles ordinaires et aux différences finies. Am. J. Math. 7, 203–258 (1885)
- 23. Perron, O.: Über Systeme von linearen Differenzengleichungen erster Ordnung. J. Reine Angew. Math. 147, 36–53 (1917)
- 24. Perron, O.: Über Summengleichungen und Poincarésche Differenzengleichungen. Math. Ann. 84(1), 1–15 (1921)
- 25. Perron, O.: Über Stabilität und asymptotisches Verhalten der Lösungen eines Systems von Differenzengleichungen. J. Reine Angew. Math. **161**, 6–64 (1929)
- 26. Coddington, E.A., Levinson, N.: Theory of Ordinary Differential Equations. McGraw-Hill, New York (1955)
- 27. Hartman, P., Wintner, A.: Asymptotic integrations of linear differential equations. Am. J. Math. 77, 48–86 (1955)
- 28. Devinatz, A.: An asymptotic theorem for systems of linear differential equations. Trans. Am. Math. Soc. **160**, 353–363 (1971)
- Benzaid, Z., Lutz, D.A.: Asymptotic representation of solutions of perturbed systems of linear difference equations. Stud. Appl. Math. 77, 195–221 (1987)
- 30. Ren, G., Shi, Y., Wang, Y.: Asymptotic behavior of solutions of perturbed linear difference systems. Linear Algebra Appl. 395, 283–302 (2005)
- 31. Elaydi, S.: An Introduction to Difference Equations, 3rd edn. Springer, Berlin (2005)
- 32. Kneser, A.: Untersuchung und asymptotische Darstellung der Integrale gewisser linearer Differentialgleichungen bei großen reellen Werten des Arguments. J. Reine Angew. Math. 117, 72–103 (1889)
- Kneser, A.: Untersuchung über die reellen Nullstellen der Integrale linearer Differentialgleichungen. Math. Ann. 42, 409–435 (1893)
- 34. Wintner, A.: Asymptotic integrations of the adiabatic oscillator. Am. J. Math. 69, 251–272 (1947)
- 35. Perron, O.: Über lineare Differenzen- und Differentialgleichungen. Math. Ann. 66, 446–487 (1909)