DYNAMIC SOLUTIONS
OF LINEAR MATRIX DIFFERENTIAL EQUATIONS

BY

JULIO CESAR RUIZ-CLAEYSSEN AND TERESA TSUKAZAN

Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

Abstract. We discuss the $m$th-order linear differential equation with matrix coefficients in terms of a particular matrix solution that enjoys properties similar to the exponential of first-order equations. A new formula for the exponential matrix is established with dynamical solutions related to the generalized Lucas polynomials.

1. Introduction. The study of higher-order linear differential equations with square matrix coefficients is usually left aside for being equivalent to a first-order equation having the companion matrix as its coefficient. The handling of this matrix is not an amenable task for obtaining results proper of higher-order equations which are shadowed or simply unknown. The companion matrix, although sparse, does not preserve any common property that the involved matrix coefficients might share.

In this paper, we shall present a direct study of higher-order equations in terms of a particular square matrix solution, which is characterized by zero initial "displacement" values and a unit initial "impulse" value, and that we shall refer to in the sequel as the dynamical solution. It will be shown that this solution enjoys intrinsic properties not necessarily shared by the complementary basis solutions, and that in a certain way the dynamical solution should do for higher-order equations what the exponential does for first-order equations. In particular, it is a right and left solution, it can be analogous to the scalar case only when the coefficients commute, together its derivatives determine the complementary solutions, and it allows to set up a recursive formula for the powers of the companion matrix.

The above results, once established, are mostly verified by induction. We shall frequently rely on a recurrence lemma involving the power series coefficients of the basis solutions and those of the dynamical solution [9]. Here we shall enunciate it in a more general and clear way. The second-order damped equation is discussed in more detail for its importance in applications, and the dynamical solution is given in terms of the theory of matrix functions developed by Runckel–Pittelkow [14]. In the last section, we discuss a connection between the Lucas polynomials and scalar dynamical solutions which allow us to obtain a new representation for the exponential

Received January 31, 1989.
This work was supported by CNPq.
2. Basic properties of dynamic solutions. We consider the $m$th-order linear differential equation

$$u^{(m)}(t) = A_1 u^{(m-1)}(t) + \cdots + A_{m-1} u'(t) + A_m u(t),$$

(1)

where the coefficients $A_j$ are square matrices. Its general solution can be written as

$$u(t) = C_0(t) u(0) + C_1(t) u'(0) + \cdots + C_{m-1}(t) u^{(m-1)}(0),$$

(2)

where the $C_j(t)$'s are square matrix solutions of (1) satisfying

$$C_j^{(k)}(0) = \delta_{jk} I; \quad k, j = 0, 1, \ldots, m - 1,$$

(3)

where $I$ denotes the identity matrix. The solution $C_{m-1}(t)$ will be denoted by $D(t)$ and referred to as the dynamical solution of (1). The solution (2) is easily obtained by noticing that

$$D(t) = \sum_{k=0}^{\infty} D_k t^k / k!, \quad C_j(t) = \sum_{k=0}^{\infty} C_{k,j} t^k / k!.$$

Recurrence Lemma. For each $j = 0, 1, \ldots, m - 1$, we have that

$$C_{k,j} = \sum_{i=0}^{j} D_{k-j-i} A_{m-i}, \quad k \geq m$$

(6)

where the coefficients $D_k$ of the dynamical solution satisfy the matrix difference equation

$$D_{k+m} = \sum_{j=1}^{m} A_j D_{k+m-j} = \sum_{j=1}^{m} D_{k+m-j} A_j$$

$$D_{m-1} = I, \quad D_0 = D_1 = \cdots = D_{m-2} = 0.$$

Proof. The first formula in (7) follows from direct substitution of $D(t)$ into the given equation (1). The relations (6) are verified by induction [11]. The second summation in (7) is then immediate from (6).
We can now draw two consequences of this lemma. The relation (7) tells us that the dynamical solution is both a right and left solution of the mth-order linear equation, that is

\[ D^{(m)}(t) = \sum_{j=1}^{m} A_j D^{(m-j)}(t) = \sum_{j=1}^{m} D^{(m-j)}(t) A_j. \]  

(8)

We shall see later on that this property is not shared in general by the complementary basis solutions \( C_j(t), j = 0, \ldots, m - 2 \). Let us see how the well-known commuting property with the exponential of the companion matrix

\[ \exp(At)A = A \exp(At) \quad (= (\exp(At))') \]

reflects itself with (8). By performing a block matrix product in \( A \exp(At) \), we shall have that the block diagonal components give rise to the relationships

\[ C_0'(t) = D(t)A_m \]

\[ C_{m-j+1}'(t) = C_{m-j}''(t) + D^{(m-j)}(t) A_j, \quad j = 1, 2, \ldots, m - 1. \]  

(9)

Hence, a backward substitution on \( j \) will take us to (8).

The recurrence lemma suggests the possibility of writing the solutions \( C_j(t) \) in terms of the dynamical solution and the given coefficients \( A_i \). In fact

\[ C_j(t) = D^{(m-j-1)}(t) - \sum_{i=1}^{m-j-1} D^{(m-j-1)}(t) A_i; \quad j = 0, 1, \ldots, m - 1. \]  

(10)

It is enough to verify that such functions are solutions of (1) which satisfy the initial values (3). This was done in [13]. Here we shall prove it by using the relationships (9). For \( j = 0 \) we have

\[ C_0'(t) = D(t)A_m = D^{(m)}(t) - \sum_{j=1}^{m-1} D^{(m-j)}(t) A_j, \]

and therefore (10) follows by integrating both sides between 0 and \( t \). Let us assume that (10) is valid for \( j - 1 < m - 1 \). Then

\[ C_j^{(j+1)}(t) = C_{j-1}^{(j)}(t) + D^{(j)}(t) A_{m-j}, \]

can be written as

\[ C_j^{(j+1)}(t) = D^{(m-1)}(t) - \sum_{i=1}^{m-j-1} D^{(m-i-1)}(t) A_i. \]

We then integrate both sides, \( j \)-times between 0 and \( t \), in order to obtain (10).

**Remark.** When the coefficients \( A_j \) commute among themselves, then they also do with \( D(t) \) and its derivatives. This means that they can appear as right factors in (10), which is the case with certain higher-order evolution equations [2]. We should observe that their natural left position could have certain implications when thinking in numerical spatial discretizations for higher-order distributed linear systems.

**Power of the companion matrix.** We shall now proceed to derive a formula for the powers of the companion matrix \( A \) defined in (5). It is convenient to write

\[ A^k = [A_{ij}^{(k)}], \]
where the components are square matrices. The matrix exponential $e^{At}$ can be expressed operationally as

$$e^{At} = [Id|dt \cdots dt^{m-1}|dt^{m-1}]' [C_0(t)C_1(t) \cdots C_{m-1}(t)].$$

Since the $k$th-derivative of $e^{At}$ at $t = 0$ equals $A^k$ it follows that

$$A^{(k)}_{ij} = C^{(k+i-1)}_{j-1}(0).$$

Consequently, we obtain from the recurrence lemma that

$$A^k = \begin{bmatrix} \sum_{s=0}^{j-1} D_{k+i-j+s-1} A_{m-s} \end{bmatrix}$$

is a $m \times n$ matrix with $n \times n$ components where $n$ is the order of the coefficients.

3. The second-order damped equation. We shall consider now the matrix equation

$$u''(t) = Bu'(t) + Au(t), \quad (12)$$

which has the general solution

$$u(t) = C(t)u(0) + D(t)u'(0).$$

The complementary basis matrix solution $C(t)$ is given by

$$C(t) = D'(t) - D(t)B,$$

as follows from (10). We claim that $C(t)$ is not a left solution of (12) unless the matrices $A$ and $B$ commute. In fact, to have $C'' = C'B + CA$ is equivalent to

$$D'' - D'B = (D'' - D'B)B + (D' - DB)A$$

or

$$D(t)AB = D(t)BA,$$

because $D(t)$ and its derivatives are left solutions of (12). It is clear that the last equality holds only when $A$ and $B$ commute.

Let us now examine how analogous is (12) with a second order scalar equation. When the coefficients commute

$$AB = BA,$$

it is easy to see that $D(t)$ is of the same form as for the case,

$$D(t) = e^{(B/2)t} \sinh \sqrt{\Delta}t/\sqrt{\Delta}, \quad \Delta = \frac{(B^2 + 4A)}{4}$$

where this expression should be understood as the product of two power series, or more appropriately as a matrix function of two commuting variables [15].

A necessary condition for the function $D(t)$ defined by (13) to be a solution of (12) is that it commutes with $B^2 + 4A$. This follows by differentiating (13) and showing that $D'' - BD' - AD = D\Delta - \Delta D$. We claim that (13) cannot be the dynamical solution of (12) unless $A$ and $B$ commute. In fact, if such $D(t)$ were the dynamical solution of (12), then $G(t) = (B^2 + 4A)D(t)$ will also be a solution. Then

$$(B^2 + 4A)D_{k+2} = B(B^2 + 4A)D_{k+1} + A(B^2 + 4A)D_k; \quad k = 0, 1, 2, \ldots$$
reduces to

\[(AB - BA)D_{k+1} = (AB^2 - B^2A)D_k\]

because of (7). The conclusion follows since for \(k = 0\) we must have \(AB - BA = 0\). Therefore, the dynamical solution \(D(t)\) of a linear matrix differential equation could be considered as a mathematical object of its own, the study of which should be pursued.

The following linear differential operators

\[Lu = u'' - Bu' - Au; \quad L^*w = w'' + w'B - wA\]

satisfy

\[\int_0^t wLu ds = \int_0^t (L^*w)u ds + \mathcal{B}(u, w)\bigg|_0^t\]

where

\[\mathcal{B}(u, w) = wu' - w'u - w'u - wBu\]

for column vectors \(u\) and row vectors \(w\). Let \(D^*(t)\) be the dynamical solution of the matrix equation \(v'' = -Bv' + Av\). It will satisfy

\[w'' = -w'B + wA. \quad (14)\]

Let \(C^*(t)\) be the complementary basis solution of \(D^*(t)\). Then

\[\mathcal{B}(D(t), D^*(t)) = 0\]

implies

\[D^*(t)D'(t) = (D^*(t) + D^*(t)B)D(t)\]

or simply

\[D^*(t)D'(t) = C^*(t)D(t).\]

By uniqueness, \(D(t)\) and \(D'(t)\) cannot be zero at the same time nor can either be zero simultaneously with \(C(t)\). Consequently, from the last relation, it follows that \(D(t)\) vanishes whenever \(D^*(t)\) does and vice versa. We shall refer to (14) as the adjoint equation of (12). This equation is the same that we would get by considering the adjoint equation of \(z' = Az\) where \(A\) is the companion matrix associated with (12).

4. Series representation of dynamic solutions. The question of having an explicit series representation for the dynamical solution amounts to solving the matrix difference equation (7) or to taking the appropriate projection of \(e^{At}\) when using the companion matrix \(A\). Since the powers of this matrix have been obtained in terms of the power series coefficients of \(D(t)\), we shall discuss the solution of the difference equation.

For simplicity, we restrict ourselves to the case \(m = 2\). The arguments and formulae could be easily extended to higher-order equations. We consider

\[D_{k+2} = A_1D_{k+1} + A_2D_k\]

\[D_1 = I, \quad D_0 = 0,\]

where \(A_1, A_2\) are arbitrary square matrices of the same order.
After writing out the first two terms $D_k$ that satisfy (15), it is observed that a
register of the involved lower indices for the coefficients, rather than exponents, will
be a convenient way to keep track of a formation law. We claim
\[
D_{2k} = \sum_{i=1}^{k} \sum_{s_1+\cdots+s_{2k-1}=2k-1} A_{s_1} \cdots A_{s_{2k-1}},
\]
\[
D_{2k+1} = \sum_{i=1}^{k} \sum_{s_1+\cdots+s_{2k+1-1}=2k+1} A_{s_1} \cdots A_{s_{2k+1-1}},
\]
where $s_i$'s can take only the values 1, 2.

The above expressions obviously hold when $k = 1$. Let us assume that they are
valid for a certain positive integer $k$. Then splitting the summation for $s_1 = 1$ and
$s_1 = 2$, it will turn out that
\[
D_{2(k+1)} = A_1 \sum_{i=1}^{k} \sum_{t_1+\cdots+t_{2k+1-1}=2k} A_{t_1} \cdots A_{t_{2k+1-1}} + A_2 \sum_{i=1}^{k} \sum_{t_1+\cdots+t_{2k-1}=2k-1} A_{t_1} \cdots A_{t_{2k-1}}
\]
\[= A_1 D_{2k+1} + A_2 D_{2k}, \]
and similarly for $D_{2(k+1)+1}$.

Therefore, the dynamical solution of a second-order matrix differential equation
is given by
\[
D(t) = \sum_{j=1}^{\infty} \sum_{s_1+\cdots+s_{j-1}=j-1}^{(j)} [A_{s_1} \cdots A_{s_{j-1}}] t^{j-i} / (j-i)!
\]
where
\[
(j) = \begin{cases} 
  \frac{j}{2}, & j \text{ even} \\
  \frac{(j+1)}{2}, & j \text{ odd}.
\end{cases}
\]

The above formula furnishes a nontrivial example of a matrix function of two
noncommuting variables. These kinds of functions were considered in the work
of Lappo-Danilevskii [7]. Certainly, this power series is not amenable for drawing
easy analytical or computational conclusions. Therefore, we shall develop in what
follows a more operational approach based in the theory of matrix functions due to
Runckel-Pittelkow [14] and Schwerdtfeger [15].

Let $f(z)$ be analytic in $|z| < K$ where $K > M$ with $M = \max|\lambda_i|$ and $\lambda_i$
the different eigenvalues of a square matrix $S$ of order $N$. Let $c(z) = \sum_{j=0}^{N} c_n z^j$
be the characteristic polynomial of $S$ with $c_0 = 1$, and denote by $d_i$ the coefficients of the
Laurent expansion for $1/c(z)$ around the origin with $|z| > M$. Then
\[
f(S) = \sum_{j=0}^{N-1} \phi_j(f) S^j
\]
where
\[
\phi_j(f) = \sum_{s=j+1}^{N} c_{n-s} \sum_{k=N+j-s}^{\infty} f^{(k)}(0)/k!
\]
\[
= \sum_{s=j}^{N-1} C_{s-j} \sum_{k=j}^{d_{k+n-s}} f^{(k)}(0)/k!
\]
\[
= \left[ f^{(j)}(0)/j! - \sum_{s=0}^{j} c_{N-s} \sum_{k=N+j-s}^{\infty} f^{(k)}(0)/k! \right].
\]

Moreover, the \(d_s\) can be computed recursively by
\[
d_s = -\sum_{k=1}^{N} c_k d_{s-k} \quad \text{for } s \geq N
\]
using \(d_s = 0\) for \(s < N\) and \(d_N = 1\). See [12] for details.

Any of these formulae could be used with \(\exp(St)\) where \(S\) is the companion matrix associated with (12), and then the appropriate projection could be taken, that is \([I0]S^{j\oplus} = D_j\). Thus we have

**Theorem 1.** The solution of the matrix difference equation \(D_{k+1} = BD_{k+1} + AD_k\), \(D_0 = 0\), \(D_1 = I\) is given by
\[
D_k = \sum_{j=0}^{N-1} \phi_j(k)D_j, \quad N = 2n
\]

where
\[
\phi_j(k) = \sum_{s=j+1}^{N} c_{N-s} d_{k+s-j}
\]
\[
= \delta_{jk} - \sum_{s=0}^{j} c_{N-s} d_{k+s-j}.
\]

**Proof.** From (18) it follows that \(S^k = \sum_{j=0}^{N-1} \phi_j(k)S^j\) with those given functions \(\phi_j\) and the conclusion is clear.

Therefore, the dynamical solution of a second-order damped equation can be given by any of the following formulae:
\[
D(t) = \sum_{r=m-1}^{N-1} D_r t^r \sum_{s=r+1}^{N} c_{N-s} \sum_{v=N+r-s}^{\infty} d_{v+s-r} \frac{t^v}{v!}
\]
\[
D(t) = \sum_{r=m-1}^{N-1} D_r \left[ \frac{t^r}{r!} - \sum_{j=0}^{r} c_{N-s} \sum_{v=N+r-s}^{\infty} d_{v+s-r} \frac{t^v}{v!} \right].
\]
5. The Lucas polynomials and dynamical solutions. The relation between a higher-order linear equation and the exponential of the corresponding companion matrix is a well known matter. We wish to discuss the converse, that is, how the exponential of an arbitrary square matrix relates with certain higher-order linear equations. We shall see that there is a relationship that could be attractive from a numerical point of view.

Given a square matrix $A$ of order $N$, we consider the scalar differential equation

$$u^{(N)}(t) = \sum_{j=1}^{N} a_j u^{(N-j)}(t),$$

(22)

associated with the characteristic polynomial

$$P(\lambda) = \det |\lambda I - A| = \lambda^N - \sum_{j=1}^{N} a_j \lambda^{N-j},$$

of the given matrix $A$. From the Cayley–Hamilton theorem we know that any power of $A$ can be written as

$$A^k = \sum_{j=1}^{N} a_j(k) A^{N-j}, \quad k = 1, 2, \ldots$$

for certain scalars $a_j(k)$. Consequently

$$e^{At} = \sum_{j=1}^{N} \beta_{N-j}(t) A^{N-j}$$

where the coefficients

$$\beta_{N-j}(t) = \sum_{k=0}^{\infty} a_j(k) \frac{t^k}{k!}$$

are to be identified. Bakarat and Baumann [1] established that the values $a_j(k)$ could be expressed in terms of the generalized Lucas polynomials

$$d_k = d_k(a_1, \ldots, a_N),$$

where $d_k$ is obtained by solving

$$d_k = \sum_{j=1}^{N} a_j d_{k-j}, \quad k \geq N$$

(23)

$$d_{N-1} = 1, \quad d_0 = d_1 = \cdots = d_{N-2} = 0.$$

Although no compact formula was given for the $a_j$'s, later on, Lavoi [9] showed their relation with the Bell polynomials, while Bruschi and Ricci [3] exhibited a generating function for the generalized Lucas polynomials $d_k$.

For us, it is quite clear that these polynomials are just the derivative values at zero of the dynamical solution $d(t)$ of the scalar equation (22). Thus any formula for the $a_j$'s has to involve the scalar dynamical solution $d(t)$. After trials, if we let $d_k = 0$
for negative integers \( k \) and set \( a_0 = 1 \), we can write Bakarat and Baumann’s formulae in the compact form

\[
\alpha_j(k) = \sum_{i=0}^{N-j} a_{i+1} d_{k-i-1} = \sum_{i=j}^{N} a_i d_{k+j-i-1} = \sum_{i=0}^{j-1} b_i d_{k+j-i-1}
\]

\[
= d_{k+j-1} = \sum_{i=1}^{j-1} a_i d_{k+j-i-1},
\]

where \( b_0 = a_0, b_1 = -a_i \). Thus

\[
\beta_{N-j}(t) = \sum_{k=0}^{\infty} \left( \sum_{i=0}^{j-1} b_i d_{k+j-i-1} \right) t^k / k!,
\]

which are nothing else but the complementary basis solutions \( c_{N-j}(t) \) of equation (22) as it follows from our relation (10). Thus, from (10) we have

**Theorem 2.** For any square matrix \( A \) of order \( N \) having the characteristic polynomial \( P(\lambda) = \sum_{j=0}^{N} b_j \lambda^{N-j}, \) \( b_0 = 1 \), we have

\[
e^{At} = \sum_{j=1}^{N} c_{N-j}(t) A^{N-j} = \sum_{j=1}^{N} \left( \sum_{i=0}^{j-1} b_i d^{(j-i-1)}(t) \right) A^{N-j}, \tag{24}
\]

where \( d(t) \) is the dynamical solution of the scalar equation

\[
\sum_{j=0}^{N} b_j u^{(N-j)}(t) = 0 \tag{25}
\]

and \( c_j(t) \) is the solution of the above equation with initial values \( c_j^{(k)}(0) = \delta_{jk} \) for \( k = 0, 1, \ldots, N - 1 \).

**Corollary.** The dynamical solution of the \( m \)th-order equation

\[
u^{(m)}(t) = \sum_{j=1}^{m} A_j u^{(m-j)}(t), \quad A_j n \times n
\]
is given by

\[
D(t) = \sum_{j=1}^{mn} \sum_{i=0}^{j-1} b_i d^{(j-i-1)}(t) D_{mn-j} \tag{26}
\]

where \( d(t) \) is the scalar dynamical solution of the equation associated with the characteristic polynomial

\[
P(\lambda) = \det \left[ \lambda^m - \sum_{j=1}^{m} A_j \lambda^{m-j} \right] = \sum_{k=0}^{mn} b_k \lambda^{mn-k} \tag{27}
\]

**Proof.** It is immediate from (24) with \( A \) being the block companion matrix of the coefficients \( A_j \).
Remarks. 1. From (26) we conclude that the solution of the matrix difference equation

\[ D_{k+m} = \sum_{j=1}^{m} A_j D_{k+m-j}, \quad D_{m-1} = I, \quad D_0 = D_1 = \cdots = D_{m-2} = 0 \]

is given by

\[ D_k = \sum_{j=1}^{N} \sum_{i=0}^{j-1} b_i d^{(k+j-i-1)}(0)D_{N-j}, \quad N = mn. \]  

(28)

This solution could be considered as an extension of the Cayley–Hamilton theorem for the case of two or more matrix variables.

2. We observe from (27) that the numerical stability of the matrices \( D_k \) can be easily related with that of the coefficients \( d^{(k+j-i-1)}(0) \), i.e., to determine when the characteristic polynomial (27) has its roots within the unit circle \([6, 8, 4]\).

6. Numerical aspects. The formula we have derived for the exponential matrix could be considered from a numerical point of view. We first assume that the coefficients of the characteristic polynomial of the given matrix \( A \) are known. This task could be done with the Faddeva–Frame–Leverrier algorithm [10]

\[ \begin{align*}
kb_k &= - \text{trace} \, Ah_{k-1}(A) \\
h_k(A) &= Ah_{r-1}(A) + b_k I
\end{align*} \]

where \( h_0(A) = I \) and \( h_N(A) = 0 \). Then, we have the alternative of using single or multistep o.d.e.s solved by computing the coefficients \( c_j(t) \) or by numerical inversion by the Laplace transform or by using the matrix methods as follows. Let \( \mathcal{A} \) denote the companion matrix associated with the coefficients \( b_k \). The calculation of the values \( c_j(t) \) could be done by computing \( e^{\mathcal{A}t} \) as it follows from Sec. 2.

Since the companion matrix \( \mathcal{A} \) is sparse, we could compute \( e^{\mathcal{A}t} \) through scaling and squaring, that is

\[ e^{\mathcal{A}t} = (e^{\mathcal{A}/m})^m, \]

where \( e^{\mathcal{A}/m} \) could be computed by either Taylor or Padé approximation. The value for \( m \) can be determined by splitting arguments with the Lie–Trotter formula

\[ e^{B+C} = \lim_{m \to \infty} (e^{B/m} e^{C/m})^m, \]

as it was suggested by M. Gunzburger and D. Gottlieb [10]. More precisely

\[ \| e^{\mathcal{A}t} - e^{B/m} e^{C/m} \| \leq (e^{1+\|C\|})\|C\|/m \]

where

\[ \mathcal{A} = B + C = \begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ c_N & c_{N-1} & \cdots & c_1 \end{bmatrix} \]

and

\[ e^{B/m} = \sum_{j=0}^{N-1} \left( \frac{B}{m} \right)^j \frac{1}{k!}, \quad e^{C/m} = I + \left( \frac{e^{c_1/m} - 1}{c_1} \right) C \]
where $c_k = -b_k$ for each $k$. For a more complete discussion on computational methods for exponential matrices we shall refer to the work of Moler and Van Loan [10].

Acknowledgment. We thank the referee for his comments and suggestions that lead to an improved version of this work.

REFERENCES