MATRIX EXPONENTIALS AND INVERSION OF CONFLUENT VANDERMONDE MATRICES

UWE LUTHER* AND KARLA ROST†

Abstract. For a given matrix A we compute the matrix exponential e^{tA} under the assumption that the eigenvalues of A are known, but without determining the eigenvectors. The presented approach exploits the connection between matrix exponentials and confluent Vandermonde matrices V. This approach and the resulting methods are very simple and can be regarded as an alternative to the Jordan canonical form methods. The discussed inversion algorithms for V as well as the matrix representation of V^{-1} are of independent interest also in many other applications.

Key words. matrix exponential, Vandermonde matrix, fast algorithm, inverse

AMS subject classifications. 34A30; 65F05; 15A09; 15A23

1. Introduction

Given a (complex) matrix A of order n, the problem of evaluating its matrix exponential e^{tA} is important in many applications, e.g., in the fields of dynamical systems or control theory.

While the matrix exponential is often represented in terms of an infinite series or by means of the Jordan canonical form our considerations have been inspired by papers like [6] and [5], where alternative methods are discussed. In an elementary way we here develop a representation of e^{tA} which involves only the eigenvalues (but not the eigenvectors), the first $(n-1)$ powers of A, and the inverse of a corresponding confluent Vandermonde matrix V. Such a representation was already given in [4], where an important motivation was to arrange the approach and the proofs simple enough to be taught to beginning students of ordinary differential equations. We here make some slight simplifications by proving such a representation, but we also concentrate our attention to the problem of developing fast recursive algorithms for the inversion of the confluent Vandermonde matrix V in the spirit of [4].

There is a large number of papers dealing with algorithms for nonconfluent and confluent Vandermonde matrices. They mainly utilize the well-known connection of Vandermonde systems with interpolation problems (see, e.g., [3] and the references therein). Moreover, in [9], [8] the displacement structure of V (called there the principle of UV-reduction) is used as a main tool.

In the present paper we want to stay within the framework of ordinary differential equations. Together with some elementary facts of linear algebra, we finally arrive at a first inversion algorithm which requires the computation of the partial fraction decomposition of $(\det(\lambda I - A))^{-1}$. This algorithm is in principle the algorithm developed in a different way in [10]. We here present a second algorithm which can be considered as an improvement of the first one since the preprocessing of the coefficients of the partial fraction decomposition is not needed. Both algorithms are fast, which means that the computational complexity is $O(n^2)$. As far as we know the second algorithm gives a new version for computing V^{-1}. For the sake of simplicity,

*Fakultät für Mathematik, Technische Universität Chemnitz, D-09107 Chemnitz, Germany, uwe.luther@mathematik.tu-chemnitz.de

†Fakultät für Mathematik, Technische Universität Chemnitz, D-09107 Chemnitz, Germany, karla.rost@mathematik.tu-chemnitz.de

let us here roughly explain the main steps of this algorithm for the nonconfluent case $V = (\lambda_i^{j-1})_{i,j=1}^n$:

1. Start with the vector $\mathbf{h}_{n-1} = (1, 1, \cdots, 1)^T \in \mathbb{C}^n$ and do a simple recursion to get vectors $\mathbf{h}_{n-2}, \ldots, \mathbf{h}_0$ and form the matrix $H = (\mathbf{h}_0\ \mathbf{h}_1\ \cdots\ \mathbf{h}_{n-1})$.
2. Multiply the jth row of V with the jth row of H to obtain numbers q_j and form the diagonal matrix $Q = \operatorname{diag}(q_j)_1^n$.
3. Multiply H from the left by the diagonal matrix $P = Q^{-1}$ to obtain V^{-T}.

In the confluent case the diagonal matrix P becomes a block-diagonal matrix with upper triangular Toeplitz blocks (t_{i-j}), $t_\ell = 0$ for $\ell > 0$.

Moreover, we show how the inversion algorithms described above lead to a matrix representation of V^{-1}, the main factor of which is just V^T. The other factors are diagonal matrices, a triangular Hankel matrix $(s_{i+j}), s_\ell = 0$ for $\ell > n$, and a block diagonal matrix with triangular Hankel blocks. For the nonconfluent case such a representation is well known (see [11] or, e.g., [9]). Generalizations of representations of this kind to other classes of matrices such as Cauchy-Vandermonde matrices or generalized Vandermonde matrices can be found in [2] and [7]. Fast inversion algorithms for Vandermonde-like matrices involving orthogonal polynomials are designed in [1].

The paper is organized as follows. In Section 2 we discuss the connection of e^{tA} with confluent Vandermonde matrices V and prove the corresponding representation of e^{tA}. Section 3 is dedicated to recursive inversion algorithms for V. A matrix representation of V^{-1} is presented in Section 4. In Section 5 we demonstrate the steps of our algorithms by means of a matrix A of order 6. Finally, in Section 6 we give some additional remarks concerning alternative possibilities for proving results of Sections 3 and 4, modified representations of e^{tA} in terms of finitely many powers of A, and the determination of analytical functions of A. Moreover, for nonderogatory matrices our results lead to a possibility to compute their (generalized) eigenvectors with the help of products of certain matrices.

2. Connection between matrix exponentials and confluent Vandermonde matrices

Let A be a given $n \times n$ complex matrix and let

$$p(\lambda) = \det(\lambda I - A) = \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0 \tag{2.1}$$

be its characteristic polynomial, $\lambda_1, \lambda_2, \ldots, \lambda_m$ its eigenvalues with the (algebraic) multiplicities $\nu_1, \nu_2, \ldots, \nu_m$, $\sum_{i=1}^m \nu_i = n$. In other words, we associate the polynomial

$$p(\lambda) = (\lambda - \lambda_1)^{\nu_1} \cdot \ldots \cdot (\lambda - \lambda_m)^{\nu_m} \tag{2.2}$$

with the pairs (λ_i, ν_i) $(i = 1, 2, \ldots, m)$. We are going to demonstrate in a very simple way that the computation of e^{tA} can be reduced to the inversion of a corresponding confluent Vandermonde matrix. The only fact we need is well known from a basic course on ordinary differential equations: Each component of the solution $\mathbf{x}(t) = e^{tA}\mathbf{v}$ of the initial value problem

$$\mathbf{x}' = A\mathbf{x}\,, \quad \mathbf{x}(0) = \mathbf{v} \tag{2.3}$$

is a linear combination of the functions

$$e^{\lambda_1 t}, te^{\lambda_1 t}, \ldots, t^{\nu_1 - 1}e^{\lambda_1 t}, \ \ldots\ , e^{\lambda_m t}, te^{\lambda_m t}, \ldots, t^{\nu_m - 1}e^{\lambda_m t}. \tag{2.4}$$

Note that these functions are just n fundamental solutions of the ordinary differential equation

$$y^{(n)}(t) + a_{n-1}\, y^{(n-1)}(t) + \ldots + a_0\, y(t) = 0 \tag{2.5}$$

the constant coefficients of which are given in (2.1). To be more precise, there is an $n \times n$ matrix C such that

$$\mathbf{x}(t) = C\mathbf{e}(t)\,, \tag{2.6}$$

where

$$\mathbf{e}(t) = \left(e^{\lambda_1 t}, te^{\lambda_1 t}, \ldots, t^{\nu_1 - 1}e^{\lambda_1 t},\ \ldots\ , e^{\lambda_m t}, te^{\lambda_m t}, \ldots, t^{\nu_m - 1}e^{\lambda_m t}\right)^T.$$

Now we use $\mathbf{x}(t) = C\mathbf{e}(t)$ as an ansatz and determine the unknown matrix C by comparing the initial values $\mathbf{x}^{(k)}(0) = C\mathbf{e}^{(k)}(0)$ with the given initial values $\mathbf{x}^{(k)}(0) = A^k\mathbf{v}$, $k = 0, 1, \ldots, n-1$:

$$C\mathbf{e}^{(k)}(0) = A^k\mathbf{v}\,, \quad k = 0, 1, \ldots, n-1\,. \tag{2.7}$$

Considering the vectors $\mathbf{e}(0), \mathbf{e}'(0), \ldots, \mathbf{e}^{(n-1)}(0)$ and $\mathbf{v}, A\mathbf{v}, \ldots, A^{n-1}\mathbf{v}$ as columns of the matrices V and $A_{\mathbf{v}}$, respectively,

$$V = \left(\mathbf{e}(0)\ \mathbf{e}'(0)\ \ldots\ \mathbf{e}^{(n-1)}(0)\right), \quad A_{\mathbf{v}} = \left(\mathbf{v}\ A\mathbf{v}\ \ldots\ A^{n-1}\mathbf{v}\right), \tag{2.8}$$

the equalities (2.7) can be rewritten in matrix form

$$CV = A_{\mathbf{v}}\,. \tag{2.9}$$

We state that V has the structure of a confluent Vandermonde matrix,

$$V = \begin{pmatrix} V(\lambda_1, \nu_1) \\ V(\lambda_2, \nu_2) \\ \vdots \\ V(\lambda_m, \nu_m) \end{pmatrix}, \tag{2.10}$$

where

$$V(\lambda, \nu) = \begin{pmatrix} 1 & \lambda & \lambda^2 & \lambda^3 & \ldots & \lambda^{n-1} \\ 0 & 1 & 2\lambda & 3\lambda^2 & \ldots & (n-1)\lambda^{n-2} \\ 0 & 0 & 2 & 6\lambda & \ldots & (n-1)(n-2)\lambda^{n-3} \\ \vdots & & & \ddots & & \vdots \\ 0 & \ldots & 0 & (\nu-1)! & \ldots & (n-1)\cdot\ldots\cdot(n-\nu+1)\lambda^{n-\nu} \end{pmatrix}.$$

It is easy to see that V is nonsingular and together with (2.6) and (2.9) we obtain the following.

Lemma 2.1. *The solution $\mathbf{x}(t)$ of the initial value problem (2.3) is given by*

$$\mathbf{x}(t) = e^{tA}\mathbf{v} = A_{\mathbf{v}}V^{-1}\mathbf{e}(t)\,, \tag{2.11}$$

where V, $A_{\mathbf{v}}$ are defined in (2.8).

Now, taking into account that $A_{\mathbf{v}} V^{-1}\mathbf{e}(t) = \sum_{i=0}^{n-1} y_i(t) A^i \mathbf{v}$, where

$$\big(y_i(t)\big)_{i=0}^{n-1} = V^{-1}\mathbf{e}(t)\,, \tag{2.12}$$

(2.11) leads to the following expression of e^{tA}.

Theorem 2.2. ([4], (1.10) and (1.15)) *The matrix exponential e^{tA} can be represented as*

$$e^{tA} = y_0(t)\, I + y_1(t)\, A + \ldots + y_{n-1}(t)\, A^{n-1}\,, \tag{2.13}$$

where $y_i(t)$, $i = 0, 1, \ldots, n-1$, are defined in (2.12).

Let us define the multiplication of a row vector $(A_1, \ldots, A_n)$ of matrices $A_i \in \mathbb{C}^{n\times n}$ and a column vector $\mathbf{v} = (v_1, \ldots, v_n)^T$ of scalars $v_i \in \mathbb{C}$ by

$$(A_1, A_2, \ldots, A_n)\, \mathbf{v} = v_1 A_1 + v_2 A_2 + \ldots + v_n A_n\,. \tag{2.14}$$

Then (2.12), (2.13) can be rewritten in a more compact form,

$$e^{tA} = \big(I, A, A^2, \ldots, A^{n-1}\big)\, \mathbf{y}(t)\,, \quad \text{where} \quad \mathbf{y}(t) = V^{-1}\mathbf{e}(t)\,. \tag{2.15}$$

Representation (2.15) is already known (see [4] and the references therein). The proof given in [4] is nice and, as promised there, it can be appreciated by students in a course on ordinary differential equations. In this paper the known initial values of e^{tA} are compared with the initial values of the ansatz

$$e^{tA} = (C_1, C_2, \ldots, C_n)\, \mathbf{f}(t), \tag{2.16}$$

where $\mathbf{f}(t)$ is an arbitrary vector of n fundamental solutions of (2.5) and C_i are $n \times n$ matrices. But this requires the application of a formal vector-matrix product defined for matrices A_i, $i = 1, \ldots, n$, and U of order n as follows

$$(A_1, \ldots, A_n)\, U = \left((A_1, \ldots, A_n)\, \mathbf{u}_1, \ldots, (A_1, \ldots, A_n)\, \mathbf{u}_n\right)\,, \tag{2.17}$$

where $\mathbf{u}_i$ denotes the ith column of U.

Possibly, the derivation given above is more natural and easier to understand. In particular, the special choice $\mathbf{f}(t) = \mathbf{e}(t)$ is also a simplification and does not mean any loss of generality. Of course, instead of the vector $\mathbf{e}(t)$, one can use any other vector $\mathbf{f}(t)$ of n fundamental solutions of (2.5). But, denoting by $W_{\mathbf{f}}$ the Wronski matrix of $\mathbf{f}$, $W_{\mathbf{f}} = (\mathbf{f}\ \mathbf{f}'\ \ldots\ \mathbf{f}^{(n-1)})$, the equality (2.11) is replaced by $e^{tA}\mathbf{v} = A_{\mathbf{v}} V_{\mathbf{f}}^{-1}\mathbf{f}(t)\,,$ where $V_{\mathbf{f}} = W_{\mathbf{f}}(0)$. This leads to the same representation (2.15) of e^{tA}, since $V_{\mathbf{f}}^{-1}\mathbf{f}(t)$ does not depend on the choice of $\mathbf{f}$ (compare the proof of Lemma 3.1).

Remark 2.3. *Denoting by $e_k(t)$ the kth component of $\mathbf{e}(t)$ and by $\mathbf{v}_k$ the kth column of V^{-1}, then $V^{-1}\mathbf{e}(t) = \sum_{k=1}^{n} e_k(t)\, \mathbf{v}_k$, and we obtain, as a consequence of (2.15), $e^{tA} = (C_1, C_2, \ldots, C_n)\, \mathbf{e}(t)$, where*

$$C_k = \big(I, A, A^2, \ldots, A^{n-1}\big)\, \mathbf{v}_k\,. \tag{2.18}$$

In the sense of (2.17) this can be written as

$$(C_1, C_2, \ldots, C_n) = \big(I, A, A^2, \ldots, A^{n-1}\big)\, V^{-1}\,. \tag{2.19}$$

Let us summarize: For an $n \times n$ matrix A having the eigenvalues λ_i with the algebraic multiplicities ν_i the matrix exponential e^{tA} is given by (2.15), where V^{-1} is the inverse of the confluent Vandermonde matrix corresponding to the pairs (λ_i, ν_i), $i = 1, \ldots, m$, $\sum \nu_i = n$.

The special structure of the matrix V can be used to compute its inverse V^{-1} in a much more efficient way than e.g. Gaussian eliminations do. Thus we will proceed with designing inversion algorithms the computational complexity of which is $O(n^2)$. Such fast algorithms have been already presented and discussed in a large number of papers (see e.g. [3] and references therein).

On one hand, we want to develop here a new version of such an algorithm, on the other hand, we intend to follow the spirit of the authors of [4], namely to be simple enough for a presentation to students of an elementary course on ordinary differential equations. Thus, we will discuss inversion algorithms exploiting elementary results of Linear Algebra, but, as far as it is possible, all within the framework of ordinary differential equations. Consequently, numerical aspects and criteria such as e.g. stability will be beyond the scope of this paper, but will be discussed in a forthcoming paper.

3. Recursive algorithms for the inversion of V

Hereafter, let $\mathbf{e}_k$ be the kth unit vector and $\mathbf{w}_{k-1}^T$ the kth row of V^{-1}, $\mathbf{w}_{k-1}^T = \mathbf{e}_k^T V^{-1}$. We start with developing a recursion for the rows $\mathbf{w}_i^T$ $(\, i = n-2, \ldots, 0\,)$ of V^{-1} from its last row $\mathbf{w}_{n-1}^T$. As a basis we use the following fact.

Lemma 3.1. *The vector $\mathbf{y}(t) = V^{-1}\mathbf{e}(t)$ is the solution of the initial value problem*

$$\left\{ \begin{array}{l} \mathbf{y}' = B\mathbf{y}, \\ \mathbf{y}(0) = \mathbf{e}_1, \end{array} \right. \qquad \textit{where} \quad B = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ & \ddots & & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & -a_{n-2} \\ 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix} \tag{3.1}$$

and $a_0, \ldots, a_{n-1}$ are the coefficients of the characteristic polynomial (2.1).

Proof. The matrix B is just the companion matrix of $p(\lambda)$. Moreover, in view of $(\, \mathbf{e}_1 \; B\mathbf{e}_1 \; \ldots \; B^{n-1}\mathbf{e}_1 \,) = I_n$, we conclude from (2.11) that

$$e^{tB}\mathbf{e}_1 = V^{-1}\mathbf{e}(t) = \mathbf{y}(t)\,,$$

which completes the proof. $\square$

Corollary 3.2. *The components of $\mathbf{y}(t) = \big(y_k(t)\big)_{k=0}^{n-1}$ can be recurrently determined from the last component $y_{n-1}(t)$:*

$$y_{k-1}(t) = y_k'(t) + a_k y_{n-1}(t)\,, \quad k = n-1, \ldots, 1\,. \tag{3.2}$$

(For $k = 0$ this is also true if we set $y_{-1} = 0$.)

Let us introduce the block-diagonal matrix

$$\widetilde{J} = \operatorname{diag}\big(\widetilde{J}_1, \dots, \widetilde{J}_m\big)\,, \tag{3.3}$$

where $\widetilde{J}_i = \lambda_i$ in case $\nu_i = 1$, otherwise

$$\widetilde{J}_i = \begin{pmatrix} \lambda_i & 1 & & & & \\ & \lambda_i & 2 & & & \\ & & \lambda_i & 3 & & \\ & & & \ddots & \ddots & \\ & & & & \lambda_i & \nu_i - 1 \\ & & & & & \lambda_i \end{pmatrix}.$$

Then, taking into account that $y_k(t) = \mathbf{w}_k^T \mathbf{e}(t)$ and $\mathbf{e}'(t) = \widetilde{J}^T \mathbf{e}(t)$, we obtain the following reformulation of (3.2).

Lemma 3.3. *The rows $\mathbf{w}_0^T, \dots, \mathbf{w}_{n-1}^T$ of V^{-1} satisfy the recursion*

$$\mathbf{w}_{k-1} = \widetilde{J}\mathbf{w}_k + a_k \mathbf{w}_{n-1}\,, \quad k = n-1, \dots, 0\,. \tag{3.4}$$

(For $k = 0$ set $\mathbf{w}_{-1} = \mathbf{0}$.)

Now we are left with the problem how to compute the last row $\mathbf{w}_{n-1}^T$ of V^{-1}. To solve this we decompose $p(\lambda)^{-1}$ into partial fractions

$$\frac{1}{p(\lambda)} = \sum_{i=1}^{m} \sum_{j=1}^{\nu_i} \frac{p_{ij}}{(\lambda - \lambda_i)^j}\,. \tag{3.5}$$

Theorem 3.4. *The last row $\mathbf{w}_{n-1}^T$ of V^{-1} is given by the coefficients of (3.5) as follows*

$$\mathbf{w}_{n-1}^T = \left(\frac{p_{11}}{0!}, \frac{p_{12}}{1!}, \dots, \frac{p_{1\,\nu_1}}{(\nu_1 - 1)!}, \ \dots\ , \frac{p_{m1}}{0!}, \dots, \frac{p_{m\,\nu_m}}{(\nu_m - 1)!} \right)\,. \tag{3.6}$$

Proof. Since the function $\mathbf{y}(t) = (y_i(t))_{i=0}^{n-1}$ satisfies $\mathbf{y}^{(k)}(0) = B^k \mathbf{e}_1 = \mathbf{e}_{k+1}$ for $k = 0, 1, \dots, n-1$ (see Lemma 3.1) we obtain, in particular, that the last component $y_{n-1}(t) = \mathbf{w}_{n-1}^T \mathbf{e}(t)$ is the solution of the initial value problem

$$y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_0 y = 0\,, \tag{3.7}$$

$$y(0) = \cdots = y^{(n-2)}(0) = 0\,,\ y^{(n-1)}(0) = 1\,. \tag{3.8}$$

We consider now the Laplace transform of $y_{n-1}(t)$ defined by

$$\left(\mathcal{L} y_{n-1}\right)(s) = \int_0^\infty e^{-s\,t} y_{n-1}(t)\, dt\,.$$

In view of (3.8), $\mathcal{L}$ applied to (3.7) yields

$$(\mathcal{L} y_{n-1})(s) = \frac{1}{p(s)}$$

which can be decomposed into partial fractions as in (3.5). By applying the back transform we obtain

$$\begin{aligned} y_{n-1}(t) &= \sum_{i=1}^{m} \sum_{j=1}^{\nu_i} \frac{p_{ij}}{(j-1)!}\, t^{j-1} e^{\lambda_i t} \\ &= \Big(\frac{p_{11}}{0!}, \frac{p_{12}}{1!}, \cdots, \frac{p_{1\,\nu_1}}{(\nu_1-1)!}, \;\cdots\;, \frac{p_{m1}}{0!}, \cdots, \frac{p_{m\,\nu_m}}{(\nu_m-1)!} \Big)\, \mathbf{e}(t) \;, \end{aligned}$$

and (3.6) is proved. □

Now we are in the position to propose a first inversion algorithm for V.

Algorithm I:

1) Compute the coefficients p_{ij} of the partial fraction expansion (3.5) and form the vector $\mathbf{w}_{n-1}$ as in (3.6).
2) Compute the remaining rows of V^{-1} via the recursion (3.4).

It is well known that the coefficients in the partial fraction decomposition can be computed by means of an ansatz and the solution of a corresponding linear system of equations. This can be organized in such a way that the computational complexity of Algorithm I is $O(n^2)$.

We propose now a further possibility the advantages of which seem to be that the precomputing of the coefficients p_{ij} is not necessary and that the recursion starts always with a "convenient" vector. To that aim let us adopt some notational convention. Introduce the following block-diagonal matrices with upper triangular Toeplitz blocks,

$$P = \operatorname{diag}(P_k)_{k=1}^m \quad \text{with} \quad P_k = \begin{pmatrix} p_{k\,\nu_k} & p_{k\,\nu_k-1} & \cdots & p_{k1} \\ & p_{k\,\nu_k} & \ddots & \vdots \\ & & \ddots & p_{k\,\nu_k-1} \\ 0 & & & p_{k\,\nu_k} \end{pmatrix},$$

$$J = \operatorname{diag}(J_k)_{k=1}^m \quad \text{with} \quad J_k = \begin{pmatrix} \lambda_k & 1 & & 0 \\ & \ddots & \ddots & \\ & & \lambda_k & 1 \\ 0 & & & \lambda_k \end{pmatrix} \in \mathbb{C}^{\nu_k \times \nu_k},$$

and define the diagonal matrix

$$D = \operatorname{diag}(D_k)_{k=1}^m \quad \text{with} \quad D_k = \operatorname{diag}\left(\frac{1}{0!}, \frac{1}{1!}, \cdots, \frac{1}{(\nu_k-1)!} \right).$$

Obviously, D and P are nonsingular matrices.

It is easy to see that

$$\mathbf{w}_{n-1} = DP\, \mathbf{h}_{n-1} \;, \tag{3.9}$$

where $\mathbf{h}_{n-1}$ is the sum of the unit vectors $\mathbf{e}_{\nu_1}, \mathbf{e}_{\nu_1+\nu_2}, \ldots, \mathbf{e}_n$, i.e.

$$\mathbf{h}_{n-1} = (\underbrace{0, \cdots, 0, 1}_{\nu_1}, \underbrace{0, \cdots, 0, 1}_{\nu_2}, \cdots, \underbrace{0, \cdots, 0, 1}_{\nu_m})^T \;. \tag{3.10}$$

We consider the recursion

$$\mathbf{h}_{k-1} = J\,\mathbf{h}_k + a_k \mathbf{h}_{n-1}\ , \quad k = n-1, \dots, 0\ . \tag{3.11}$$

Lemma 3.5. *The recursion (3.11) produces the rows* $\mathbf{w}_{k-1}^T$ *of* V^{-1} *as follows*

$$\mathbf{w}_{k-1} = DP\,\mathbf{h}_{k-1}\ , \quad k = n, \dots, 1\ . \tag{3.12}$$

(Putting $\mathbf{w}_{-1} = \mathbf{0}$ *the equality (3.12) is also true for* $k = 0$*.)*

Proof. We have, due to (3.11) and (3.9),

$$DP\,\mathbf{h}_{k-1} = DPJ\,\mathbf{h}_k + a_k \mathbf{w}_{n-1}\ , \quad k = n-1, \dots, 0\,. \tag{3.13}$$

Now, it is easy to verify that $\widetilde{J}D = DJ$ and $JP = PJ$. Hence, (3.13) shows that

$$DP\,\mathbf{h}_{k-1} = \widetilde{J}DP\,\mathbf{h}_k + a_k \mathbf{w}_{n-1}\ , \quad k = n-1, \dots, 0\,.$$

We compare this with (3.4) and state that $DP\,\mathbf{h}_{-1} = \mathbf{0}$ implies $\mathbf{h}_{-1} = \mathbf{0}$, which completes the proof. □

The following fact has been already stated in [10].

Corollary 3.6. *The transpose of* V^{-1} *is given by*

$$V^{-T} = DPH\ ,$$

where $H = (\mathbf{h}_0\ \mathbf{h}_1\ \cdots\ \mathbf{h}_{n-1})$.

From Corollary 3.6 it follows that

$$P^{-1} = HV^T D\,.$$

On the other hand we have, obviously,

$$P^{-1} = \operatorname{diag}(Q_k)_{k=1}^m \quad \text{with} \quad Q_k = P_k^{-1}\,.$$

One can easily see that the inverse of a (nonsingular) upper triangular Toeplitz matrix is again an upper triangular Toeplitz matrix, i.e., that Q_k has the form

$$Q_k = \begin{pmatrix} q_{k\,\nu_k} & q_{k\,\nu_k - 1} & \cdots & q_{k1} \\ & q_{k\,\nu_k} & \ddots & \vdots \\ & & \ddots & q_{k\,\nu_k-1} \\ 0 & & & q_{k\,\nu_k} \end{pmatrix}, \tag{3.14}$$

where $\mathbf{q}^{(k)} = (q_{k1}, \dots, q_{k\,\nu_k})^T$ is the last column of P_k^{-1}, i.e. the solution of the equation $P_k \mathbf{q}^{(k)} = \mathbf{e}_{\nu_k}$. Thus, the matrix $Q = P^{-1}$ is completely given by that n elements of $HV^T = QD^{-1}$ which stand in the last columns of the diagonal blocks of HV^T,

$$HV^T = \operatorname{diag}\left[\begin{pmatrix} \ddots & \ddots & \ddots & q_{k1}\,(\nu_k - 1)! \\ & \ddots & \ddots & q_{k2}\,(\nu_k - 1)! \\ & & \ddots & \vdots \\ & & & q_{k\nu_k}\,(\nu_k - 1)! \end{pmatrix}\right]_{k=1}^m . \tag{3.15}$$

With the entries of the vector $\mathbf{q}^{(k)}$ we form the matrix Q_k and compute the solution $\mathbf{p}^{(k)} = (p_{k1}, \ldots, p_{k\nu_k})^T$ of

$$Q_k \mathbf{p}^{(k)} = \mathbf{e}_{\nu_k}, \tag{3.16}$$

which gives the matrix P_k.

Now we propose the following second inversion algorithm for V.

Algorithm II:

1) Compute recurrently the columns $\mathbf{h}_{j-1},\ \ j = 1, \ldots, n$ of H by (3.11).
2) Compute that n elements $q_{kj}(\nu_k - 1)!$ of the product HV^T (see (3.15)) which determine the blocks Q_k (see (3.14)) of the matrix $P^{-1} = \operatorname{diag}(Q_k)_{k=1}^m$.
3) Compute the upper triangular Toeplitz matrices P_k by solving (3.16) and form the matrix $P = \operatorname{diag}(P_k)_{k=1}^m$.
4) Compute $V^{-1} = (DPH)^T$.

In case that the multiplicities ν_k are small compared with n the computational complexity of the algorithm is again $O(n^2)$.

Our next aim is to develop a matrix representation of V^{-1}.

4. Matrix representation of V^{-1}

Let us start with the *nonconfluent case*. Since in this case $\widetilde{J} = \operatorname{diag}(\lambda_i)_1^n$, the recursion (3.4) together with (3.6) leads directly to the following representation

$$V^{-1} = \begin{pmatrix} a_1 & \cdots & a_{n-1} & 1 \\ \vdots & \iddots & \iddots & \\ a_{n-1} & 1 & & \\ 1 & & & 0 \end{pmatrix} V^T \operatorname{diag}\big(p_{k1}\big)_{k=1}^n .$$

(Note that $p_{k1} = p'(\lambda_k)^{-1}$.) This means that the inverse of a Vandermonde matrix V is closely related to its transpose V^T.

Now we are going to show that such a representation is also possible in the *confluent case*. Recall that $p(\lambda) = \sum_{i=0}^n a_i \lambda^i$, where $a_n = 1$. Then the matrix H can be computed as follows.

Lemma 4.1.

$$H = \sum_{i=0}^{n-1} J^i \, \mathbf{h}_{n-1}(a_{i+1}, \cdots, a_n, 0, \cdots, 0) \ . \tag{4.1}$$

Proof. Denote the columns of the matrix on the right hand side of (4.1) by $\widetilde{\mathbf{h}}_k$ $(k = 0, 1, \ldots, n-1)$,

$$\widetilde{\mathbf{h}}_k = \sum_{i=0}^{n-1} J^i \, \mathbf{h}_{n-1}(a_{i+1}, \cdots, a_n, 0, \cdots, 0) \, \mathbf{e}_{k+1} = \sum_{i=0}^{n-k-1} a_{k+i+1} J^i \, \mathbf{h}_{n-1} \ .$$

Then, clearly, $\widetilde{\mathbf{h}}_{n-1} = \mathbf{h}_{n-1}$, and we observe that

$$J \, \widetilde{\mathbf{h}}_k + a_k \widetilde{\mathbf{h}}_{n-1} = \sum_{i=0}^{n-k-1} a_{k+i+1} J^{i+1} \mathbf{h}_{n-1} + a_k \mathbf{h}_{n-1} = \sum_{i=0}^{n-k} a_{k+i} J^i \, \mathbf{h}_{n-1} \ ,$$

which is just $\widetilde{\mathbf{h}}_{k-1}$. Thus, $\widetilde{\mathbf{h}}_k$ satisfies the recursion (3.11) which proves that $\widetilde{\mathbf{h}}_k = \mathbf{h}_k$ for $k = 0, 1, \ldots, n-1$. □

Let Z_{ν_k} be the $\nu_k \times \nu_k$ matrix of counteridentity which has ones on the antidiagonal and zeros elsewhere. Then we obtain the following matrix representation of V^{-1}.

Theorem 4.2.

$$V^{-1} = \begin{pmatrix} a_1 & \cdots & a_{n-1} & 1 \\ \vdots & \iddots & \iddots & \\ a_{n-1} & 1 & & \\ 1 & & & 0 \end{pmatrix} V^T \operatorname{diag}\big(G_k\big)_{k=1}^m , \tag{4.2}$$

where $G_k = D_k P_k Z_{\nu_k} D_k$.

Before we prove this theorem, let us mention that (up to diagonal matrices) the blocks of the last factor of (4.2) are triangular Hankel matrices

$$G_k = D_k \begin{pmatrix} p_{k1} & p_{k2} & \cdots & p_{k\,\nu_k-1} & p_{k\,\nu_k} \\ p_{k2} & & & p_{k\,\nu_k} & \\ \vdots & & \iddots & & \\ p_{k\,\nu_k-1} & p_{k\,\nu_k} & & & \\ p_{k\,\nu_k} & & & & 0 \end{pmatrix} D_k .$$

Proof of Theorem 4.2. Since $P_k Z_{\nu_k} = Z_{\nu_k} P_k^T$, the last factor $\operatorname{diag}\big(G_k\big)_{k=1}^m$ on the right hand side of (4.2) is equal to

$$DZP^TD , \quad \text{where} \quad Z = \operatorname{diag}\big(Z_{\nu_k}\big)_{k=1}^m .$$

The first factor of representation (4.2) can be written as

$$\sum_{i=0}^{n-1} \mathbf{e}_{i+1}\mathbf{a}^{(i)} ,$$

where $\mathbf{a}^{(i)} = (a_{i+1}, \cdots, a_n, 0, \cdots, 0)$. Consequently, (4.2) is equivalent to

$$V^{-T} = DPZDV \sum_{i=0}^{n-1} \mathbf{e}_{i+1}\mathbf{a}^{(i)} .$$

In view of Corollary 3.6 and Lemma 4.1 it remains to show that

$$\sum_{i=0}^{n-1} J^i\, \mathbf{h}_{n-1}\mathbf{a}^{(i)} = ZDV \sum_{i=0}^{n-1} \mathbf{e}_{i+1}\mathbf{a}^{(i)} . \tag{4.3}$$

It is easy to see that

$$J^i = \operatorname{diag}\Big((\lambda_k I_{\nu_k} + S_{\nu_k})^i \Big)_{k=1}^m = \operatorname{diag}\left(\sum_{j=0}^{i} \binom{i}{j} \lambda_k^{i-j} S_{\nu_k}^j \right)_{k=1}^m ,$$

where S_ν denotes the forward shift of order ν,

$$S_\nu = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & 0 & 1 \\ 0 & & & 0 \end{pmatrix} . \tag{4.4}$$

Now, it follows that $J^i\, \mathbf{h}_{n-1} = \left(\mathbf{u}_1^T, \cdots, \mathbf{u}_m^T\right)^T$ with

$$\begin{aligned} \mathbf{u}_k &= \sum_{j=0}^{i} \binom{i}{j} \lambda_k^{i-j} S_{\nu_k}^j \mathbf{e}_{\nu_k} = \sum_{j=0}^{\nu_k - 1} \binom{i}{j} \lambda_k^{i-j} \mathbf{e}_{\nu_k - j} \\ &= \left(\binom{i}{\nu_k - 1} \lambda_k^{i-\nu_k+1}, \binom{i}{\nu_k - 2} \lambda_k^{i-\nu_k+2}, \, \cdots \, , \binom{i}{0} \lambda_k^i \right)^T . \end{aligned}$$

One can easily check that the same vector $\mathbf{u}_k$ can be obtained from the part $V(\lambda_k, \nu_k)$ of V (see (2.10)) by

$$Z_{\nu_k} D_k V(\lambda_k, \nu_k)\, \mathbf{e}_{i+1} \, .$$

This means that

$$J^i\, \mathbf{h}_{n-1} = ZDV \mathbf{e}_{i+1} \, ,$$

and (4.3) is proved. $\square$

5. Example

Let us compute e^{tA} for the following matrix A of order 6,

$$A = \begin{pmatrix} 3.5 & 0.5 & -0.5 & -0.5 & -0.5 & -0.5 \\ 3.5 & 4.5 & 1.5 & -1.5 & -3.5 & -5.5 \\ -1 & -0.5 & 2 & 1.5 & 1 & 0.5 \\ 3 & 1.5 & 0 & 0.5 & -2 & -1.5 \\ -3.5 & -2 & -0.5 & 1 & 4.5 & 2 \\ 5.5 & 3 & 0.5 & -2 & -4.5 & -4 \end{pmatrix} .$$

Its characteristic polynomial $p(\lambda) = \det(\lambda I - A)$ is

$$\begin{aligned} p(\lambda) &= \lambda^6 - 11\lambda^5 + 45\lambda^4 - 77\lambda^3 + 22\lambda^2 + 84\lambda - 72 \\ &= (\lambda - 3)^2 (\lambda - 2)^3 (\lambda + 1) \, . \end{aligned}$$

Thus, the corresponding confluent Vandermonde matrix V is given by

$$\begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \lambda_1^3 & \lambda_1^4 & \lambda_1^5 \\ 0 & 1 & 2\lambda_1 & 3\lambda_1^2 & 4\lambda_1^3 & 5\lambda_1^4 \\ 1 & \lambda_2 & \lambda_2^2 & \lambda_2^3 & \lambda_2^4 & \lambda_2^5 \\ 0 & 1 & 2\lambda_2 & 3\lambda_2^2 & 4\lambda_2^3 & 5\lambda_2^4 \\ 0 & 0 & 2 & 6\lambda_2 & 12\lambda_2^2 & 20\lambda_2^3 \\ 1 & \lambda_3 & \lambda_3^2 & \lambda_3^3 & \lambda_3^4 & \lambda_3^5 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 9 & 27 & 81 & 243 \\ 0 & 1 & 6 & 27 & 108 & 405 \\ 1 & 2 & 4 & 8 & 16 & 32 \\ 0 & 1 & 4 & 12 & 32 & 80 \\ 0 & 0 & 2 & 12 & 48 & 160 \\ 1 & -1 & 1 & -1 & 1 & -1 \end{pmatrix} .$$

Let us compute V^{-1} with the help of the first inversion algorithm. For this aim we need the partial fraction decomposition of $p(\lambda)^{-1}$ which is

$$-\frac{13}{16}\frac{1}{\lambda-3}+\frac{1}{4}\frac{1}{(\lambda-3)^2}+\frac{22}{27}\frac{1}{\lambda-2}+\frac{5}{9}\frac{1}{(\lambda-2)^2}+\frac{1}{3}\frac{1}{(\lambda-2)^3}-\frac{1}{432}\frac{1}{\lambda+1}\,.$$

Due to (3.6) the last row of V^{-1} is given by

$$\mathbf{w}_5^T=\left(-\frac{13}{16},\frac{1}{4},\frac{22}{27},\frac{5}{9},\frac{1}{6},-\frac{1}{432}\right).$$

To compute the other rows $\mathbf{w}_0^T,\cdots,\mathbf{w}_4^T$ of V^{-1} we use (3.4),

$$\begin{aligned}\mathbf{w}_4&=\widetilde{J}\mathbf{w}_5-11\mathbf{w}_5\,,\\ \mathbf{w}_3&=\widetilde{J}\mathbf{w}_4+45\mathbf{w}_5\,,\\ \mathbf{w}_2&=\widetilde{J}\mathbf{w}_3-77\mathbf{w}_5\,,\\ \mathbf{w}_1&=\widetilde{J}\mathbf{w}_2+22\mathbf{w}_5\,,\\ \mathbf{w}_0&=\widetilde{J}\mathbf{w}_1+84\mathbf{w}_5\,,\end{aligned}\qquad \text{where}\quad \widetilde{J}=\begin{pmatrix}3&1&&&&\\&3&&&&\\&&2&1&&\\&&&2&2&\\&&&&2&\\&&&&&-1\end{pmatrix}. \tag{5.1}$$

Starting from below we write $\mathbf{w}_5^T$, $\mathbf{w}_4^T,\cdots,\mathbf{w}_0^T$ as rows of a matrix and obtain

$$V^{-1}=\begin{pmatrix}-\frac{43}{2}&6&\frac{67}{3}&14&6&\frac{1}{6}\\ \frac{69}{4}&-5&-\frac{152}{9}&-\frac{37}{3}&-4&-\frac{13}{36}\\ \frac{103}{8}&-\frac{7}{2}&-\frac{356}{27}&-\frac{76}{9}&-\frac{23}{6}&\frac{67}{216}\\ -\frac{293}{16}&\frac{21}{4}&\frac{166}{9}&\frac{38}{3}&\frac{9}{2}&-\frac{19}{144}\\ \frac{27}{4}&-2&-\frac{61}{9}&-\frac{14}{3}&-\frac{3}{2}&\frac{1}{36}\\ -\frac{13}{16}&\frac{1}{4}&\frac{22}{27}&\frac{5}{9}&\frac{1}{6}&-\frac{1}{432}\end{pmatrix}.$$

Now the matrices C_i in the representation

$$\begin{aligned}e^{tA}&=(C_1,C_2,C_3,C_4,C_5,C_6)\,\mathbf{e}(t)\\ &=e^{3t}C_1+te^{3t}C_2+e^{2t}C_3+te^{2t}C_4+t^2e^{2t}C_5+e^{-t}C_6\end{aligned}$$

can be computed by formula (2.18):

$$C_1 = -\frac{43}{2}\,I + \frac{69}{4}\,A + \frac{103}{8}\,A^2 - \frac{293}{16}\,A^3 + \frac{27}{4}\,A^4 - \frac{13}{16}\,A^5$$

$$= \begin{pmatrix} 1.5 & 1 & 0.5 & 0 & -0.5 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0.5 & 0 & -0.5 & -1 & -0.5 \\ -0.5 & 0 & 0.5 & 1 & 1.5 & 0 \\ 1 & 0.5 & 0 & -0.5 & -1 & -0.5 \end{pmatrix},$$

$$C_2 = 6\,I - 5\,A - \frac{7}{2}\,A^2 + \frac{21}{4}\,A^3 - 2\,A^4 + \frac{1}{4}\,A^5$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1.5 & 1 & 0.5 & 0 & -0.5 & -1 \\ -3 & -2 & -1 & 0 & 1 & 2 \\ 1.5 & 1 & 0.5 & 0 & -0.5 & -1 \end{pmatrix},$$

$$C_3 = \frac{67}{3}\,I - \frac{152}{9}\,A - \frac{356}{27}\,A^2 + \frac{166}{9}\,A^3 - \frac{61}{9}\,A^4 + \frac{22}{27}\,A^5$$

$$= \begin{pmatrix} -0.5 & -1 & -0.5 & 0 & 0.5 & 1 \\ 1 & 1.5 & 0 & -0.5 & -1 & -1.5 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & -0.5 & 0 & 1.5 & 1 & 0.5 \\ 0.5 & 0 & -0.5 & -1 & -0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$C_4 = 14\,I - \frac{37}{3}\,A - \frac{76}{9}\,A^2 + \frac{38}{3}\,A^3 - \frac{14}{3}\,A^4 + \frac{5}{9}\,A^5$$

$$= \begin{pmatrix} 0 & -0.5 & -1 & -0.5 & 0 & 0.5 \\ 0.5 & 1 & 1.5 & 0 & -0.5 & -1 \\ -1 & -0.5 & 0 & 1.5 & 1 & 0.5 \\ 0.5 & 0 & -0.5 & -1 & -0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$C_5 = 6\,I - 4\,A - \frac{23}{6}\,A^2 + \frac{9}{2}\,A^3 - \frac{3}{2}\,A^4 + \frac{1}{6}\,A^5$$

$$= \begin{pmatrix} 0.25 & 0 & -0.25 & -0.5 & -0.25 & 0 \\ -0.5 & 0 & 0.5 & 1 & 0.5 & 0 \\ 0.25 & 0 & -0.25 & -0.5 & -0.25 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$C_6 = \frac{1}{6}\,I - \frac{13}{36}\,A + \frac{67}{216}\,A^2 - \frac{19}{144}\,A^3 + \frac{1}{36}\,A^4 - \frac{1}{432}\,A^5$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -0.5 & 0 & 0.5 & 1 & 1.5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -0.5 & 0 & 0.5 & 1 & 1.5 \end{pmatrix}.$$

In the first step of the above used inversion algorithm for V we have supposed that the coefficients

$$(p_{11}\,,\, p_{12}\,,\, p_{21}\,,\, p_{22}\,,\, p_{23}\,,\, p_{31}) = \left(-\frac{13}{16}\,,\,\frac{1}{4}\,,\,\frac{22}{27}\,,\,\frac{5}{9}\,,\,\frac{1}{3}\,,\,-\frac{1}{432}\right) \tag{5.2}$$

of the partial fraction decomposition of $p(\lambda)^{-1}$ are already computed. If we use the second algorithm of Section 3, then we do not need these coefficients in advance. Let us demonstrate this for our example.

First we have to compute the matrix $H = (\mathbf{h}_0 \ \cdots \ \mathbf{h}_5)$. By (3.10) we have

$$\mathbf{h}_5 = (0,\, 1,\, 0,\, 0,\, 1,\, 1)^T\,.$$

The columns $\mathbf{h}_4, \cdots, \mathbf{h}_0$ are computed by recurrence (3.11) (analogously to (5.1)):

$$\begin{aligned} \mathbf{h}_4 &= J\mathbf{h}_5 - 11\mathbf{h}_5\,, \\ \mathbf{h}_3 &= J\mathbf{h}_4 + 45\,\mathbf{h}_5\,, \\ \mathbf{h}_2 &= J\mathbf{h}_3 - 77\,\mathbf{h}_5\,, \\ \mathbf{h}_1 &= J\mathbf{h}_2 + 22\,\mathbf{h}_5\,, \\ \mathbf{h}_0 &= J\mathbf{h}_1 + 84\,\mathbf{h}_5\,, \end{aligned} \qquad \text{where} \quad J = \begin{pmatrix} 3 & 1 & & & & \\ & 3 & & & & \\ & & 2 & 1 & & \\ & & & 2 & 1 & \\ & & & & 2 & \\ & & & & & -1 \end{pmatrix}.$$

We obtain

$$H = \begin{pmatrix} -8 & 4 & 6 & -5 & 1 & 0 \\ 24 & -20 & -14 & 21 & -8 & 1 \\ 9 & 3 & -5 & 1 & 0 & 0 \\ -18 & 3 & 13 & -7 & 1 & 0 \\ 36 & -24 & -23 & 27 & -9 & 1 \\ -72 & 156 & -134 & 57 & -12 & 1 \end{pmatrix}.$$

Now we compute the last columns of the blocks of HV^T (see (3.15)),

$$HV^T = H \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 1 \\ 3 & 1 & 2 & 1 & 0 & -1 \\ 9 & 6 & 4 & 4 & 2 & 1 \\ 27 & 27 & 8 & 12 & 12 & -1 \\ 81 & 108 & 16 & 32 & 48 & 1 \\ 243 & 405 & 32 & 80 & 160 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} * & 13 & & & & \\ & 4 & & & & \\ & & * & * & 2 & \\ & & & * & -10 & \\ & & & & 6 & \\ & & & & & -432 \end{pmatrix}.$$

From these columns we obtain the matrix $Q = \operatorname{diag}(Q_k)_{k=1}^{3}$ (see (3.14)),

$$Q = \begin{pmatrix} 4 & 13 & & & & \\ & 4 & & & & \\ & & 3 & -5 & 1 & \\ & & & 3 & -5 & \\ & & & & 3 & \\ & & & & & -432 \end{pmatrix}.$$

It is easy to invert this matrix, which yields

$$P = Q^{-1} = \begin{pmatrix} \frac{1}{4} & -\frac{13}{16} & & & & \\ & \frac{1}{4} & & & & \\ & & \frac{1}{3} & \frac{5}{9} & \frac{22}{27} & \\ & & & \frac{1}{3} & \frac{5}{9} & \\ & & & & \frac{1}{3} & \\ & & & & & -\frac{1}{432} \end{pmatrix}.$$

We remark that just the coefficients (5.2) stand in the last columns of the blocks of P. Now we may compute $(DPH)^T$, $D = \operatorname{diag}\big(1,1,1,1,\frac{1}{2},1\big)$, to obtain V^{-1}.

6. Final remarks

Remark 6.1. *The assertion $\mathbf{w}_{k-1} - \widetilde{J}\,\mathbf{w}_k = a_k \mathbf{w}_{n-1}$ $(k = 0, \dots, n-1)$ of Lemma 3.3, written in matrix form*

$$V^{-T} S_n - \widetilde{J}\, V^{-T} = \mathbf{w}_{n-1}(a_0, a_1, \cdots, a_{n-1}), \tag{6.1}$$

where S_n is defined in (4.4), is the so-called displacement equation for V^{-T} (see [8]). This equation may also be concluded from the well known equality

$$B^T V^T = V^T \widetilde{J}\,, \tag{6.2}$$

where B is defined in (3.1). Indeed, (6.2) is equivalent to (6.1), since

$$V^{-T} S_n - \mathbf{w}_{n-1}(a_0, a_1, \cdots, a_{n-1}) = V^{-T} B^T\,.$$

Thus, proving Lemma 3.3 by checking (6.2) is an alternative to the proof given in Section . Moreover, we mention that (6.2) can also be viewed as a direct consequence of Lemma 3.1, since replacing $\mathbf{y}$ in $\mathbf{y}' = B\mathbf{y}$ by its definition $\mathbf{y} = V^{-1}\mathbf{e}(t)$ yields $V^{-1}\widetilde{J}^T\mathbf{e}(t) = BV^{-1}\mathbf{e}(t)$, i.e., $V^{-1}\widetilde{J}^T = BV^{-1}$.

Remark 6.2. *To prove the assertion of Theorem 3.4 one can use, instead of the Laplace transformation, the Laurent series expansion at infinity*

$$\frac{1}{(\lambda-\lambda_i)^j} = \frac{(-1)^{j-1}}{(j-1)!}\left[\frac{1}{\lambda}\sum_{k=0}^{\infty}\left(\frac{\lambda_i}{\lambda}\right)^k\right]^{(j-1)} = \sum_{k=j}^{\infty}\binom{k-1}{j-1}\frac{\lambda_i^{k-j}}{\lambda^k} \qquad (|\lambda|>|\lambda_i|).$$

Indeed, using two representations of $p(\lambda)^{-1}$, namely (3.5) and $\prod_i(\lambda-\lambda_i)^{-\nu_i}$, we obtain its Laurent series expansion in two different forms. Comparing the coefficients of $\lambda^{-1},\lambda^{-2},\cdots,\lambda^{-n}$, one gets n equations which show that the vector on the right hand side of (3.6) multiplied by V from the right yields $\mathbf{e}_n^T$.

Remark 6.3. *The last step leading to the matrix representation (4.2) of V^{-1} was a direct proof of $J^i\,\mathbf{h}_{n-1} = ZDV\mathbf{e}_{i+1}$ $(i=0,\cdots,n-1)$. We mention that this can also be proved by using the well known formula for e^{tJ} applied to $\mathbf{h}_{n-1}$,*

$$e^{tJ}\mathbf{h}_{n-1} = ZD\,\mathbf{e}(t)\,. \tag{6.3}$$

Indeed, we only have to compare (6.3) with the representation

$$e^{tJ}\mathbf{h}_{n-1} = \big(\mathbf{h}_{n-1}\ \ J\,\mathbf{h}_{n-1}\ \ \cdots\ \ J^{n-1}\,\mathbf{h}_{n-1}\big)\,V^{-1}\mathbf{e}(t)$$

from Lemma 2.1 to obtain $\big(\mathbf{h}_{n-1}\ \ J\,\mathbf{h}_{n-1}\ \ \cdots\ \ J^{n-1}\,\mathbf{h}_{n-1}\big) = ZDV$.

Remark 6.4. *In Section 2 we have only used that each component of e^{tA} is a solution of (2.5). This fact is a consequence of the Cayley-Hamilton theorem $p(A)=0$. Hence, if*

$$q(\lambda) = \lambda^N + b_{N-1}\,\lambda^{N-1} + \cdots + b_1\,\lambda + b_0$$

is any other polynomial with $q(A)=0$, for example the minimal polynomial of A satisfying $N\le n$, then the components of e^{tA} are solutions of

$$y^{(N)}(t) + b_{N-1}\,y^{(N-1)}(t) + \cdots + b_0\,y(t) = 0 \tag{6.4}$$

(since the matrix valued function $Y(t)=e^{tA}$ solves this equation). So we obtain, in the same way as in Section 2,

$$e^{tA}\mathbf{v} = \big(\mathbf{v}\ \ A\mathbf{v}\ \ A^2\mathbf{v}\ \ \cdots\ \ A^{N-1}\mathbf{v}\big)\,V_q^{-1}\mathbf{e}_q(t)\,, \tag{6.5}$$

where V_q is the confluent Vandermonde matrix corresponding to the zeros of $q(\lambda)$ and $\mathbf{e}_q(t)$ denotes the vector of the standard fundamental solutions of (6.4). The resulting representation of e^{tA} is

$$e^{tA} = (I,A,A^2,\cdots,A^{N-1})\,\mathbf{y}_q(t)\,, \quad \textit{where} \quad \mathbf{y}_q(t) = V_q^{-1}\mathbf{e}_q(t)\,.$$

Remark 6.5. *Assume for sake of simplicity that all ν_k are equal to ν, where ν is small compared with n. Then using representation (4.2) the matrix-vector multiplication with the inverse of a (confluent) Vandermonde matrix V can be done with $O(n\log^2 n)$ computational complexity. Indeed, utilizing FFT techniques then the complexity of* *the multiplication of an* $n \times n$ *triangular Hankel matrix and a vector is* $O(n \log n)$. *In particular, this leads to a complexity of* $O(n)$ *for the matrix vector multiplication with* $\operatorname{diag}\left(G_k\right)_{k=1}^{m}$. *Now, the rows of the matrix* V^T *can be reordered in such a way that the first* m *rows form a nonconfluent* $m \times n$ *Vandermonde matrix, the* $(km+1)th$ *up to* $((k+1)m)th$ *rows a nonconfluent Vandermonde matrix multiplied by a diagonal matrix,* $k = 1, \ldots, \nu$. *With the ideas of [2] concerning the nonconfluent case this leads to a complexity of* $O(n \log^2 n)$ *to multiply* V^T *with a vector.*

Remark 6.6. *The matrix representation (4.2) of* V^{-1} *suggests a third inversion algorithm for* V*: Let us denote the first and the last factor on the right hand side of (4.2) by* $H(\mathbf{a})$ *and* G*, respectively. Up to diagonal matrices, the blocks of* G *are triangular Hankel matrices. This implies that the inverse* G^{-1} *has the same structure. Thus,* $G^{-1} = VH(\mathbf{a})V^T$ *is completely determined by* n *entries of the product* $VH(\mathbf{a})V^T$. *Using a fast algorithm for the application of* $H(\mathbf{a})$*, these entries can be determined with* $O(n^2 \log n)$ *operations. Now, the resulting matrix can be inverted efficiently to obtain* G.

Remark 6.7. *The exponential of the companion matrix* B *(see (3.1)) is given by*

$$e^{tB} = \Big(\mathbf{y}(t)\ \ \mathbf{y}'(t)\ \ \cdots\ \ \mathbf{y}^{(n-1)}(t)\Big), \quad \mathbf{y}(t) = V^{-1}\mathbf{e}(t)\,. \tag{6.6}$$

This follows from $\mathbf{y}(t) = e^{tB}\mathbf{e}_1$. *Let us assume that* A *is nonderogatory, i.e. that there exists a nonsingular matrix* U *such that* $AU = UB$. *Let* $\mathbf{v}$ *be the first column of* U. *Then*

$$A\mathbf{v} = UB\mathbf{e}_1 = U\mathbf{e}_2,\ \ldots\ , A^{n-1}\mathbf{v} = U\mathbf{e}_n,$$

which means that $U = A_{\mathbf{v}}$*, where* $A_{\mathbf{v}}$ *is defined in (2.8). On the other hand, if there is a* $\mathbf{v}$ *such that* $A_{\mathbf{v}}$ *is nonsingular, then*

$$\begin{aligned} A_{\mathbf{v}}B &= A_{\mathbf{v}}\, S_n^T - A_{\mathbf{v}}\, (a_i)_{i=0}^{n-1}\, \mathbf{e}_n^T \\ &= \big(\, A\mathbf{v}\ \ \cdots\ \ A^{n-1}\mathbf{v}\ \ \ \mathbf{0}\,\big) - \big(\,\mathbf{0}\ \ \cdots\ \ \mathbf{0}\ \ \ p(A)\mathbf{v} - A^n\mathbf{v}\,\big) = AA_{\mathbf{v}}\,, \end{aligned} \tag{6.7}$$

which shows that A *is nonderogatory. To sum up, we can state that* A *is nonderogatory if and only if there exists a* $\mathbf{v}$ *such that* $A_{\mathbf{v}}$ *is invertible. In this case we have*

$$e^{tA} = A_{\mathbf{v}}\Big(\mathbf{y}(t)\ \ \mathbf{y}'(t)\ \ \cdots\ \ \mathbf{y}^{(n-1)}(t)\Big)\, A_{\mathbf{v}}^{-1}\,.$$

Remark 6.8. *If* A *is nonderogatory, i.e. if there is a vector* $\mathbf{v}$ *such that* $A_{\mathbf{v}}$ *is invertible (see Remark 6.7), then the matrices* $A_{\mathbf{v}}$ *and* H^T *(or* V^{-1}*) can be used to compute a linearly independent system of* n *generalized eigenvectors of* A. *Indeed, for the matrix* B *two of such systems are given by the columns of* H^T *(since, by (3.11) and assertion* $\mathbf{h}_{-1} = \mathbf{0}$ *of Lemma 3.5,* $HB^T = JH$*) and the columns of* V^{-1} *(see (6.2)), so that (6.7) shows that* $A_{\mathbf{v}}H^T$ *and* $A_{\mathbf{v}}V^{-1}$ *consist of generalized eigenvectors of* A*, more precisely,*

$$AA_{\mathbf{v}}H^T = A_{\mathbf{v}}H^TJ^T \quad \textit{and} \quad AA_{\mathbf{v}}V^{-1} = A_{\mathbf{v}}V^{-1}\widetilde{J}^T\,.$$

If one wants to transform A into its Jordan canonical form J, then one has only to take into account that $\widetilde{J}^T D^{-1} = D^{-1} J^T$, $J^T Z = ZJ$ (see the proofs of Lemma 3.5 and Theorem 4.2) and, consequently,

$$AA_{\mathbf{v}} H^T Z = A_{\mathbf{v}} H^T ZJ \quad \textit{and} \quad AA_{\mathbf{v}} V^{-1} D^{-1} Z = A_{\mathbf{v}} V^{-1} D^{-1} ZJ\,. \tag{6.8}$$

(This is also true if $A_{\mathbf{v}}$ is singular. But in this case the columns of $A_{\mathbf{v}} H^T Z$ and $A_{\mathbf{v}} V^{-1} D^{-1} Z$, respectively, are not linearly independent.) These equations are simple algebraic facts and surely known. But one can use the matrix representation (4.2) of V^{-1} together with the obvious equality $JP = PJ$ to obtain the following nice reformulation of (6.8),

$$AA_{\mathbf{v}} H(\mathbf{a}) V^T D = A_{\mathbf{v}} H(\mathbf{a}) V^T DJ\,,$$

where $H(\mathbf{a})$ denotes the first factor on the right hand side of (4.2). Now we see that no matrix has to be inverted to obtain generalized eigenvectors of A. One only has to compute the product $A_{\mathbf{v}} H(\mathbf{a}) V^T D$ $(= A_{\mathbf{v}} H^T Z)$.

Remark 6.9. *For $f(z) = e^z$ we have $f(A) = \big(I, A, \cdots, A^{n-1}\big) \cdot f(B)\mathbf{e}_1$ (since $\mathbf{y}(t) = e^{tB}\mathbf{e}_1$). We mention that this is also true for an arbitrary power series*

$$f(z) = \sum_{m=0}^{\infty} \alpha_m (z - z_0)^m$$

with convergence radius $R > 0$ and its corresponding matrix valued function

$$f(A) = \sum_{m=0}^{\infty} \alpha_m (A - z_0 I)^m\,, \quad A \in \mathbb{C}^{n\times n}, \ \ \max_i |\lambda_i(A) - z_0| < R\,.$$

Indeed, if we compare the initial values of both sides of the equation $e^{-z_0 t} e^{tA} = \big(I, A, \cdots, A^{n-1}\big) \cdot e^{-z_0 t} e^{tB} \mathbf{e}_1$, then we obtain

$$(A - z_0 I)^m = \big(I, A, \cdots, A^{n-1}\big) \cdot (B - z_0 I)^m \mathbf{e}_1\,,$$

which yields the assertion. For the application to an vector this means

$$f(A)\mathbf{v} = A_{\mathbf{v}} f(B)\mathbf{e}_1\,.$$

We remark that the vector $(\beta_0, \cdots, \beta_{n-1})^T = f(B)\mathbf{e}_1$ is just the coefficient vector of the Hermite interpolation polynomial $P(z) = \sum_{k=0}^{n-1} \beta_k z^k$ of $f(z)$ with respect to the eigenvalues λ_i of A and their algebraic multiplicities ν_i. This follows from (6.2):

$$\begin{aligned} Vf(B)\mathbf{e}_1 &= Vf(B)V^{-1}V\mathbf{e}_1 = f(VBV^{-1})V\mathbf{e}_1 = f\big(\widetilde{J}^T\big)V\mathbf{e}_1 \\ &= \Big(f(\lambda_1), f'(\lambda_1), \cdots, f^{(\nu_1-1)}(\lambda_1), \ \cdots \ , f(\lambda_m), \cdots, f^{(\nu_m-1)}(\lambda_m)\Big)^T. \end{aligned}$$

REFERENCES

[1] D. Calvetti and L. Reichel, *Fast inversion of Vandermonde-like matrices involving orthogonal polynomials*, BIT, **33** (1993), 473-484.

[2] T. Finck, G. Heinig, and K. Rost, *An inversion formula and fast algorithms for Cauchy-Vandermonde matrices*, Linear Algebra Appl., **183** (1993), 179-192.

[3] G. H. Golub and C. F. van Loan, *Matrix Computations*, Third Edition, The Johns Hopkins U.P., Baltimore, London, 1996.

[4] W. A. Harris, Jr., J. P. Fillmore, and D. R. Smith, *Matrix exponentials - another approach*, SIAM Rev., **43** (2001), 694-706.

[5] I. E. Leonard, *The matrix exponential*, SIAM Rev., **38** (1996), 507-512.

[6] C. Moler and C. F. van Loan, *Nineteen dubious ways to compute the exponential of a matrix*, SIAM Rev., **20** (1978), 801-836.

[7] G. Heinig, *Hermite's formula for vector polynomial interpolation with applications to structured matrices*, Appl. Anal., **70** (1999), 331-345.

[8] G. Heinig, W. Hoppe, and K. Rost, *Structured matrices in interpolation and approximation problems*, Wiss. Zeitschrift der TU Karl-Marx-Stadt, **31**, 2 (1989), 196-202.

[9] G. Heinig and K. Rost, *Algebraic Methods for Toeplitz-like Matrices and Operators*, Akademie-Verlag, Berlin, Birkhäuser Verlag, Basel, 1984.

[10] S.-H. Hou and W.-K. Pang, *Inversion of confluent Vandermonde matrices*, Comp. Math. Appl., **43** (2002), 1539-1547.

[11] J. F. Traub, *Associated polynomials and uniform methods for the solutionof linear problems*, SIAM Rev., **8** (1966), 277-301.