SIAM REVIEW
Vol. 43, No. 4, pp. 694–706

© 2001 Society for Industrial and Applied Mathematics

Matrix Exponentials—Another Approach*

William A. Harris, Jr.†
Jay P. Fillmore‡
Donald R. Smith‡

Abstract. The exponential of a matrix and the spectral decomposition of a matrix can be computed knowing nothing more than the eigenvalues of the matrix and the Cayley–Hamilton theorem. The arrangement of the ideas in this paper is simple enough to be taught to beginning students of ODEs.

Key words. matrix exponential, Vandermonde matrix, spectral decomposition

AMS subject classifications. 15-01, 15A21, 34-01

PII. S0036144599362406

Introduction. William Harris (born Dec. 18, 1930, died[1] Jan. 8, 1998) is certainly missed by all who knew him. He was frequently a visiting scholar at UCSD, most recently in the fall of 1997. Smith met Harris in 1975 because of their common interest in ODEs. Fillmore (whose main interest is geometry) did not meet Harris until fall 1997.

Harris possessed a great enthusiasm for ODEs and would frequently exchange insights on various aspects of the subject with colleagues. During his final month at UCSD, Harris gave Fillmore and Smith two short lectures on a nice way of computing the exponential of a matrix. During his last year Harris also lectured and spoke on these ideas at other locations, including Seattle, Rolla, and Los Angeles. Like Harris, we feel that these ideas are simple and elegant and suitable for presentation to undergraduates in an elementary course. In this spirit, we have recorded what Harris told us about computing matrix exponentials, and we present that information here. (See Moler and Van Loan (1978) for 19 other approaches.)

The paper is arranged as follows:

- Section 1 gives a description of the method. We start with the well-known fact that entries of e^{At} satisfy the ODE $c(d/dt)u = 0$, where $c(\lambda)$ is the characteristic polynomial of A, and we proceed to an arrangement of all the information, which is simple and easy to use.
- Section 2 gives an example using the rudiments of *Mathematica*. We include just enough so that a beginning student can do some interesting compu-

*Received by the editors October 6, 1999; accepted for publication (in revised form) June 5, 2001; published electronically October 31, 2001.
http://www.siam.org/journals/sirev/43-4/36240.html

†This author is deceased. Former address: Department of Mathematics, University of Southern California, Los Angeles, CA 90089.
‡Department of Mathematics, University of California at San Diego, La Jolla, CA 92093 (jfillmore@ucsd.edu, drsmith@ucsd.edu).

[1]See *In Memoriam: William Ashton Harris Jr.* (1930–1998), AMS Notices, 45 (1998), p. 684.

tations. Several known results flow unexpectedly without effort from the development of section 1.

- Section 3 gives the spectral decomposition in the case of distinct eigenvalues.
- Section 4 discusses the case of repeated roots. The fact that the generalized (or "confluent") Vandermonde matrix is the change-of-basis matrix, which brings the companion matrix of a characteristic polynomial into Jordan canonical form, was well known early in this century but is now largely forgotten. Here we give a proof using differential equations and, in a remark, a proof that is strictly algebraic. (We leave it to the reader to judge which proof is simpler.)
- Section 5 gives another example, illustrating the sensitivity of the matrix exponential to changes in the data.
- Section 6 comments on other classical as well as recent methods for exponentiating matrices.

1. The Exponential. We define e^{At} as the $n \times n$ solution of the initial value problem

$$\frac{dX}{dt} = AX \quad \text{for} \quad -\infty < t < \infty, \qquad X = I \quad \text{at} \quad t = 0, \tag{1.1}$$

where I is the $n \times n$ identity matrix. Our goal is to compute e^{At}.

We shall often use the symbol $D = d/dt$ to denote differentiation with respect to an independent variable. From (1.2), e^{At} satisfies

$$D^k e^{At} = A^k e^{At}$$

for every nonnegative integer $k = 0, 1, 2, \ldots$ and, by linearity,

$$p(D)e^{At} = p(A)e^{At} \tag{1.2}$$

for every polynomial p. As in (1.2), polynomial operators are applied to a matrix entrywise.

Let c be the characteristic polynomial of A,

$$c(\lambda) = \det(\lambda I - A) = \lambda^n - c_1\lambda^{n-1} - c_2\lambda^{n-2} - \cdots - c_n, \tag{1.3}$$

with coefficients $c_1 = \operatorname{tr} A, \ldots, c_n = (-1)^n \det A$. The Cayley–Hamilton theorem[2] asserts that

$$c(A) = 0 \tag{1.4}$$

in $\mathbb{C}^{n\times n}$. Hence, using the polynomial $p(\lambda) = c(\lambda)$ in (1.2), we have $c(D)\,e^{At} = 0$, so each entry $u(t)$ of e^{At} is a solution of the scalar nth-order linear constant coefficient homogeneous differential equation

$$c(D)\,u(t) = 0. \tag{1.5}$$

[2] A simple proof is based on the formula of Kramer: $(\lambda I - A) \times \text{adjugate}\ (\lambda I - A) = c(\lambda)I$. Viewing this as the product of two polynomials in the commuting indeterminant λ and with matrix coefficients, it is legitimate to set $\lambda = A$. For details, see p. 395 of Broida and Williamson (1989).

The Wronski matrix $W[y;t]$ for n smooth, complex-valued functions $y_1, y_2, \ldots, y_n$ of a real variable is the $n \times n$ matrix

$$W[y;t] := \begin{bmatrix} y_1(t) & y_2(t) & \cdots & y_n(t) \\ y_1'(t) & y_2'(t) & \cdots & y_n'(t) \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{(n-1)}(t) & y_2^{(n-1)}(t) & \cdots & y_n^{(n-1)}(t) \end{bmatrix}. \tag{1.6}$$

If each of the functions $z_1, z_2, \ldots, z_n$ is a (constant coefficient) linear combination of $y_1, y_2, \ldots, y_n$, we have the matrix equation

$$[\, z_1 \quad z_2 \quad \cdots \quad z_n \,] = [\, y_1 \quad y_2 \quad \cdots \quad y_n \,]\, C, \tag{1.7}$$

where C is a complex $n \times n$ matrix with constant entries. This equation relates the first rows of the respective two Wronski matrices $W[z;t]$ and $W[y;t]$, and then, by successive differentiations of (1.7), we have

$$W[z;t] = W[y;t]C. \tag{1.8}$$

Now let $y_1(t), y_2(t), \ldots, y_n(t)$ and $z_1(t), z_2(t), \ldots, z_n(t)$ be any two bases for the n-dimensional vector space (over $\mathbb{C}$) of solutions of the nth-order linear differential equation (1.5). The Wronski matrices are now invertible matrices[3] (see section 8.12 of Apostol (1997)). Relation (1.8) evaluated at $t = 0$ determines the constant matrix $C = W[y;0]^{-1}W[z;0]$, and then (1.8) yields

$$W[z;t]W[z;0]^{-1} = W[y;t]W[y;0]^{-1}. \tag{1.9}$$

The resulting common matrix value appearing on both sides of (1.9) is recognized as the Wronski matrix

$$W[\varphi;t] = \begin{bmatrix} \varphi_1(t) & \varphi_2(t) & \cdots & \varphi_n(t) \\ \varphi_1'(t) & \varphi_2'(t) & \cdots & \varphi_n'(t) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1^{(n-1)}(t) & \varphi_2^{(n-1)}(t) & \cdots & \varphi_n^{(n-1)}(t) \end{bmatrix}$$

of the *principal solutions* $\varphi_1(t), \varphi_2(t), \ldots, \varphi_n(t)$, these latter being the special solutions of (1.5) satisfying the initial conditions contained in the following matrix equation for this Wronski matrix evaluated at $t = 0$:

$$\begin{bmatrix} \varphi_1(0) & \varphi_2(0) & \cdots & \varphi_n(0) \\ \varphi_1'(0) & \varphi_2'(0) & \cdots & \varphi_n'(0) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1^{(n-1)}(0) & \varphi_2^{(n-1)}(0) & \cdots & \varphi_n^{(n-1)}(0) \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

In particular, the unique *principal solutions* for (1.5) may be obtained as

$$[\, \varphi_1(t) \quad \cdots \quad \varphi_n(t) \,] = [\, y_1(t) \quad \cdots \quad y_n(t) \,]\, W[y;0]^{-1} \tag{1.10}$$

in terms of arbitrary fundamental solutions $y_1(t), \ldots, y_n(t)$, where $[\, y_1(t) \quad \cdots \quad y_n(t) \,]$ denotes a $1 \times n$ row matrix.

[3] In an elementary course one may also mention *Abel's result* for such a Wronski matrix $W[y;t]$, according to which the *determinant* of $W[y;t]$ (i.e., the *Wronskian*) is a solution of the first-order differential equation $du/dt = (\mathrm{tr}A)u$.

Each entry of e^{At} is a solution of the nth-order differential equation (1.5) and is hence a unique linear combination (with constant coefficients) of any given n fundamental solutions $y_1(t), \ldots, y_n(t)$. It follows that the array e^{At} can be written as a linear combination

$$e^{At} = y_1(t)F_1 + y_2(t)F_2 + \cdots + y_n(t)F_n = \begin{bmatrix} y_1(t) & y_1(t) & \cdots & y_n(t) \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_n \end{bmatrix} \tag{1.11}$$

for unique $n \times n$ constant matrices $F_i \in \mathbb{C}^{n\times n}$ $(i = 1, 2, \ldots, n)$. The expression

$$\begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_n \end{bmatrix}$$

on the right side of (1.11) is not interpreted as an $n^2 \times n$ block matrix,[4] but rather as an $n \times 1$ array with the ith entry being F_i (for $i = 1, 2, \ldots, n$).

Evaluating successive derivatives of (1.11) at $t = 0$ yields

$$y_1^{(i)}(0)F_1 + \cdots + y_n^{(i)}(0)F_n = A^i \quad \text{for} \quad i = 0, 1, \ldots, n-1,$$

from which we obtain the formal matrix equation

$$\begin{bmatrix} y_1(0) & y_2(0) & \cdots & y_n(0) \\ y_1'(0) & y_2'(0) & \cdots & y_n'(0) \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{(n-1)}(0) & y_2^{(n-1)}(0) & \cdots & y_n^{(n-1)}(0) \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_n \end{bmatrix} = \begin{bmatrix} I \\ A \\ \vdots \\ A^{n-1} \end{bmatrix} \tag{1.12}$$

or, in terms of the Wronski matrix,

$$W[y; 0] \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_n \end{bmatrix} = \begin{bmatrix} I \\ A \\ \vdots \\ A^{n-1} \end{bmatrix}.$$

Hence,

$$\begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_n \end{bmatrix} = W[y; 0]^{-1} \begin{bmatrix} I \\ A \\ \vdots \\ A^{n-1} \end{bmatrix}. \tag{1.13}$$

The F_i in (1.11) are polynomials in A and may be obtained by (1.13).

[4] If V denotes the n-dimensional space of solutions to (1.5), then the space of solutions of $c(D)X = 0$ ($n \times n$ matrices) is the free right module $V \otimes_{\mathbb{C}} \mathbb{C}^{n\times n}$ of rank n over $\mathbb{C}^{n\times n}$. The row on the right side of (1.11) is $\begin{bmatrix} y_1(t)I & \cdots & y_n(t)I \end{bmatrix}$, which is a consequence of the injection $V \longrightarrow V \otimes_{\mathbb{C}} \mathbb{C}^{n\times n}$ given by $y \mapsto y \otimes I$.

Inserting (1.13) into (1.11) yields a formula for e^{At}:

$$e^{At} = [\, y_1(t) \quad y_2(t) \quad \cdots \quad y_n(t) \,] \, W[y;0]^{-1} \begin{bmatrix} I \\ A \\ \vdots \\ A^{n-1} \end{bmatrix} \tag{1.14}$$

or, more simply in terms of the principal solutions (1.10),

$$e^{At} = [\, \varphi_1(t) \quad \varphi_2(t) \quad \cdots \quad \varphi_n(t) \,] \begin{bmatrix} I \\ A \\ \vdots \\ A^{n-1} \end{bmatrix}. \tag{1.15}$$

Remark 1. We thus have two representations for e^{At}: $a(bc)$ from (1.11) and (1.13), and $(ab)c$ from (1.15). These are connected by abc from (1.14). One needs to compute the inverse

$$W[y;0]^{-1}$$

of *only one* $n \times n$ constant matrix. The representation $a(bc)$ uses an arbitrary choice of the n fundamental solutions $y_j(t)$, and the representation $(ab)c$ uses the n unique principal fundamental solutions determined by

$$c(D)\varphi_j(t) = 0 \tag{1.16}$$

subject to the n^2 initial conditions

$$\varphi_j^{(i-1)}(0) = \delta_j^i$$

for $i, j = 1, 2, \ldots, n$ (i.e., the matrix initial condition $W[\varphi;0] = I$).

Remark 2. Beyond their fixed initial conditions, the principal solutions $\varphi_j(t)$ depend only on the characteristic polynomial $c(\lambda) = \det(\lambda I - A)$ and hence remain valid for all A with the same $c(\lambda)$.

Remark 3. One could replace the differential equation (1.2) defining e^{At} and use instead the Maclaurin series $\sum_{k=0}^{\infty} A^k t^k / k!$. Then the fact that (1.15) involves the principal solutions (1.10) for (1.5) is ascertained directly from the Cayley–Hamilton theorem (1.3)–(1.4), according to which A^n is a linear combination of $I, A, A^2, \ldots, A^{n-1}$. In this case the (converging) Maclaurin series implies directly

$$\begin{aligned} e^{At} &= x_1(t)I + x_2(t)A + x_3(t)A^2 + \cdots + x_n(t)A^{n-1} \\ &= [\, x_1(t) \quad x_2(t) \quad \cdots \quad x_n(t) \,] \begin{bmatrix} I \\ A \\ \vdots \\ A^{n-1} \end{bmatrix} \end{aligned} \tag{1.17}$$

for suitable complex-valued functions $x_1(t), \ldots, x_2(t)$. Note that the Maclaurin series and each of its derivatives converge and yield, in particular, at $t = 0$,

$$x_j^{(i-1)}(0) = \delta_j^i \quad \text{for} \quad i, j = 1, 2, \ldots, n.$$

Similarly, the Maclaurin series also implies that e^{At} satisfies the differential equation of (1.2), which, with (1.4) and (1.17), yields

$$c(D)x_j(t) = 0 \quad \text{for} \quad j = 1, 2, \ldots, n.$$

It follows by (1.16) and a standard differential equation uniqueness theorem that the present functions $x_j(t)$ of (1.17) coincide with the previous principal solutions $\varphi_j(t)$. As an aside, we note that some references (e.g., Chapter 14 of Bronson (1989)) factor suitable powers of t from the principal solutions and cast (1.17) in the form

$$e^{At} = \alpha_1(t)I + \alpha_2(t)t\,A + \alpha_3(t)t^2A^2 + \cdots + \alpha_n(t)t^{n-1}A^{n-1},$$

with $x_j(t) = \alpha_j(t)t^{j-1}$. In this case the $\alpha_j(t)$ are not principal solutions.

2. An Example. We illustrate a *Mathematica* computation for the 3×3 matrix

$$a = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 12 & -16 & 7 \end{bmatrix}. \tag{2.1}$$

This matrix[5] a along with the identity matrix are entered into *Mathematica* as

```
a := {{0, 1, 0}, {0, 0, 1}, {12, -16, 7}};

id := IdentityMatrix[3];
```

To `MatrixForm[a]` *Mathematica* responds with

$$\begin{matrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 12 & -16 & 7 \end{matrix}$$

The eigenvalues of `a` are obtained by `Eigenvalues[a]`

$$\{2, 2, 3\}$$

Now, the characteristic polynomial is $c(\lambda) = (\lambda-2)^2(\lambda-3) = \lambda^3 - 7\lambda^2 + 16\lambda - 12$. In this case, fundamental solutions of the characteristic differential equation $x''' - 7\,x'' + 16\,x' - 12\,x = 0$ are simple enough to insert into *Mathematica* "by hand":

```
y[t_] = {Exp[2 * t], t * Exp[2 * t], Exp[3 * t]}
```

The Wronski matrix is computed from `y[t_]` and the derivatives of `y[t_]` as in (1.6):

```
w[t_] = {y[t], D[y[t], {t, 1}], D[y[t], {t, 2}]}
```

Then `MatrixForm[w[t]]` yields

$$\begin{matrix} e^{2\,t} & e^{2\,t}\,t & e^{3\,t} \\ 2\,e^{2\,t} & e^{2\,t} + 2\,e^{2\,t}\,t & 3\,e^{3\,t} \\ 4\,e^{2\,t} & 4\,e^{2\,t} + 4\,e^{2\,t}\,t & 9\,e^{3\,t} \end{matrix} \tag{2.2}$$

[5] One often uses lowercase letters to denote variables in *Mathematica*. The immediate output of a *Mathematica* command is suppressed by following the command with a semicolon.

The principal solutions of $x''' - 7\,x'' + 16\,x' - 12\,x = 0$ are obtained as in (1.10) from

$$\texttt{temp1} = \texttt{y[t].Inverse[w[0]]}$$

(The lowered dot denotes matrix multiplication in *Mathematica*.) The "row vector" of principal solutions is now assigned to `temp1`:

$$\left\{-3\,e^{2\,t} + 4\,e^{3\,t} - 6\,e^{2\,t}\,t, \qquad 4\,e^{2\,t} - 4\,e^{3\,t} + 5\,e^{2\,t}\,t, \qquad -e^{2\,t} + e^{3\,t} - e^{2\,t}\,t\right\}$$

Now, create a "column vector" whose elements are the identity, a, and a^2 as appears on the right side of (1.15):

$$\texttt{ee[u_, x_]} = \{\texttt{u}, \texttt{x}, \texttt{x.x}\}$$

$$\texttt{temp2} = \texttt{ee[id, a]}$$

The indeterminant `u` is needed to obtain the identity matrix by substitution. Then `temp2`, written as nine rows of three, contains

$$\{\{\{1,0,0\},\{0,1,0\},\{0,0,1\}\},\{\{0,1,0\},\{0,0,1\},\{12,-16,7\}\},$$

$$\{\{0,0,1\},\{12,-16,7\},\{84,-100,33\}\}\}$$

Finally, we obtain the matrix exponential e^{At} using (1.15):

$$\texttt{MatrixForm[Simplify[temp1.temp2]]}$$

$$\begin{array}{ccc} e^{2\,t}\,(-3+4\,e^t-6\,t) & e^{2\,t}\,(4-4\,e^t+5\,t) & e^{2\,t}\,(-1+e^t-t) \\ 12\,e^{2\,t}\,(-1+e^t-t) & e^{2\,t}\,(13-12\,e^t+10\,t) & e^{2\,t}\,(-3+3\,e^t-2\,t) \\ 12\,e^{2\,t}\,(-3+3\,e^t-2\,t) & 4\,e^{2\,t}\,(9-9\,e^t+5\,t) & e^{2\,t}\,(-8+9\,e^t-4\,t) \end{array}$$

Mathematica has its own matrix exponential evoked by `Simplify[MatrixForm [MatrixExp[a * t]]]`. `MatrixForm[Simplify[MatrixExp[a * t]]]`. This yields the same answer, as can be checked by subtraction.

3. Spectral Decomposition. As a by-product of the present approach, one gains an elementary development of the spectral decomposition of A. The approach is sufficiently simple that it can be easily presented to undergraduate students. We illustrate the argument for the semisimple case in which A is diagonalizable with distinct eigenvalues. The characteristic polynomial (1.3) factors as

$$c(x) = (x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n) \tag{3.1}$$

with $\lambda_i \neq \lambda_j$ for $i \neq j$, where now we have replaced the variable λ of (1.3) with x.

The functions

$$e^{\lambda_j t} \quad \text{for} \quad j = 1, 2, \ldots, n \tag{3.2}$$

provide n fundamental solutions for (1.5), so these functions may be used for $y_j(t)$ in (1.11) to write

$$e^{At} = e^{\lambda_1 t}E_1 + e^{\lambda_2 t}E_2 + \cdots + e^{\lambda_n t}E_n, \tag{3.3}$$

where in this special case E_j plays the role of the previous constant matrix F_j. Now (1.12) becomes

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \\ \vdots \\ E_n \end{bmatrix} = \begin{bmatrix} I \\ A \\ \vdots \\ A^{n-1} \end{bmatrix}, \tag{3.4}$$

where here the Wronski matrix is a Vandermonde matrix. Solving (3.4) as in (1.13) yields

$$E_j = e_j(A) \quad \text{for} \quad j = 1, 2, \ldots, n, \tag{3.5}$$

where $e_j(x)$ is a polynomial. In fact, Kramer's rule gives that $e_j(x)$ is the solution of

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} = \begin{bmatrix} 1 \\ x \\ \vdots \\ x^{n-1} \end{bmatrix}, \tag{3.6}$$

which is

$$e_j(x) = \frac{(x-\lambda_1)\cdots(x-\lambda_{j-1})(x-\lambda_{j+1})\cdots(x-\lambda_n)}{(\lambda_j-\lambda_1)\cdots(\lambda_j-\lambda_{j-1})(\lambda_j-\lambda_{j+1})\cdots(\lambda_j-\lambda_n)} \tag{3.7}$$

based on the n distinct values $\lambda_1, \lambda_2, \ldots, \lambda_n$. Relation (3.5) with (3.7) is a familiar representation for the polynomial dependence of E_j on A, obtained here by employing the special fundamental solutions (3.2) in (1.11). Note that we have the identity

$$1 = e_1(x) + e_2(x) + \cdots + e_n(x), \tag{3.8}$$

which comes from the top row of (3.6).

Relations (3.1) and (3.7) show that the characteristic polynomial $c(x)$ divides the product

$$e_i(x)e_j(x)$$

for $i \neq j$. Hence (3.5) and the Cayley–Hamilton theorem (1.4) give

$$E_iE_j = 0 \quad (\text{in } \mathbb{C}^{n\times n}) \quad \text{for} \quad i \neq j. \tag{3.9}$$

Analogous to (3.8) it also holds (put $t = 0$ in (3.3)) that

$$I = E_1 + E_2 + \cdots + E_n \quad (\text{in } \mathbb{C}^{n\times n}), \tag{3.10}$$

from which (multiply both sides of (3.10) by E_j and use (3.9))

$$E_j^2 = E_j \quad \text{for} \quad j = 1, 2, \ldots, n.$$

Thus each E_j in (3.3) is a projection. Moreover, with (3.7), $c(x)$ also divides the product

$$(x-\lambda_j)e_j(x),$$

so, again by (3.5) and Cayley–Hamilton,

$$(A - \lambda_j I)E_j = 0 \quad (\text{in } \mathbb{C}^{n\times n}).$$

Hence $AE_j = \lambda_j E_j$ or

$$A(E_j v) = \lambda_j (E_j v) \quad \text{for all} \quad v \in \mathbb{C}^{n\times 1}$$

for each $j = 1, 2, \ldots, n$.

Thus in the present case of n distinct eigenvalues, (3.10) leads to the spectral decomposition

$$A = AE_1 + AE_2 + \cdots + AE_n,$$

where $E_j = e_j(A)$ is a projection onto the eigenspace of A associated with the eigenvalue λ_j.

4. Remarks on the Case of Repeated Roots. We include a few remarks on the case of repeated eigenvalues. Let

$$c(x) = (x - \lambda_1)^{n_1}(x - \lambda_2)^{n_2} \cdots (x - \lambda_r)^{n_r}$$

be the characteristic polynomial of a matrix or of a linear ODE. For each $\rho = 1, 2, \ldots, r$, the characteristic root λ_ρ has multiplicity n_ρ. The degree of the polynomial is $n_1 + n_2 + \cdots + n_r = n$.

A basis of solutions to $c(D)u = 0$ is made up by the functions

$$\frac{1}{j!}\, t^j\, e^{\lambda_\rho t}$$

for $0 \le j \le n_\rho - 1$ and $\rho = 1, 2, \ldots, r$. The resulting Wronski matrix will lead to the *generalized Vandermonde matrix*.

The generalized (or "confluent") Vandermonde matrix T consists of r blocks of size $n_\rho \times n_\rho$ ($\rho = 1, 2, \ldots, r$), each block being of the form (with ρ suppressed in $n = n_\rho$ and in $\lambda = \lambda_\rho$)

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \lambda & 1 & 0 & \cdots & 0 \\ \lambda^2 & 2\lambda & 1 & \cdots & 0 \\ \lambda^3 & 3\lambda^2 & 3\lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda^{n-1} & (n-1)\lambda^{n-2} & \binom{n-1}{2}\lambda^{n-3} & \cdots & 1 \end{bmatrix},$$

and the $\mu\nu$th entry being $\binom{\mu-1}{\nu-1}\lambda^{\mu-\nu}$. It turns out that this matrix T is the change-of-basis matrix, which brings the companion matrix (cf. (1.3))

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ c_n & c_{n-1} & c_{n-2} & \cdots & c_2 & c_1 \end{bmatrix} \tag{4.1}$$

into Jordan canonical form J:

$$AT = TJ. \tag{4.2}$$

This pretty fact can be found (without proof) on p. 60 of Turnbull and Aitken (1961). A short algebraic proof is included at the end of this paper, but here we give a derivation in the spirit of differential equations.

Let φ be the $1 \times (n_1 + n_2 + \cdots + n_r)$ principal matrix (which we write in terms of r $1 \times n_\rho$ blocks, $\rho = 1, 2, \ldots, r$, with only the first block shown)

$$\varphi = \left[e^{\lambda_1 t} \quad t\, e^{\lambda_1 t} \quad \tfrac{t^2}{2!}\, e^{\lambda_1 t} \quad \cdots \quad \tfrac{t^{n_1-1}}{(n_1-1)!}\, e^{\lambda_1 t} \quad (r-1 \text{ additional blocks}) \right].$$

Entrywise differentiation of the matrix $\varphi(t)$ gives directly

$$D\varphi(t) = \varphi(t) \begin{bmatrix} J_1 & & & 0 \\ & J_2 & & \\ & & \ddots & \\ 0 & & & J_r \end{bmatrix}$$

with blocks

$$J_\rho = \begin{bmatrix} \lambda_\rho & 1 & 0 & \cdots & 0 \\ 0 & \lambda_\rho & 1 & \cdots & 0 \\ 0 & 0 & \lambda_\rho & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_\rho \end{bmatrix}$$

for $\rho = 1, 2, \ldots, r$. That is, $D\varphi(t) = \varphi(t)J$, where the constant matrix J on the right is a matrix of Jordan blocks. Now differentiate $D\varphi(t) = \varphi(t)J$ repeatedly to obtain the rows of

$$D \begin{bmatrix} \varphi \\ D\varphi \\ D^2\varphi \\ \vdots \\ D^{n-1}\varphi \end{bmatrix} = \begin{bmatrix} \varphi \\ D\varphi \\ D^2\varphi \\ \vdots \\ D^{n-1}\varphi \end{bmatrix} J$$

or, in terms of the Wronski matrix,

$$DW[\varphi; t] = W[\varphi; t]J. \tag{4.3}$$

Return to $W[\varphi; t]$ and differentiate it again, writing the result differently now as a product with a suitable constant matrix on the left,

$$D \begin{bmatrix} \varphi \\ D\varphi \\ D^2\varphi \\ \vdots \\ D^{n-2}\varphi \\ D^{n-1}\varphi \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ & & & & \ddots & \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ * & & \cdots & \cdots & & * \end{bmatrix} \begin{bmatrix} \varphi \\ D\varphi \\ D^2\varphi \\ \vdots \\ D^{n-2}\varphi \\ D^{n-1}\varphi \end{bmatrix},$$

where we would like to have $D^n\varphi$ for the last entry on the right side. Since every entry of φ, and consequently of $W[\varphi; t]$, is a solution of $c(D)u = 0$, we have $(D^n - c_1\, D^{n-1} - c_2\, D^{n-2} - \cdots - c_n)u = 0$ and so $D^n\, u = (c_1\, D^{n-1} + c_2\, D^{n-2} + \cdots - + c_n)\, u$. Hence

we may fill in $[\,c_n \quad c_{n-1} \quad \cdots \quad c_2 \quad c_1\,]$ for the bottom row of the constant matrix on the right side, from which, with (4.1),

$$DW[\varphi;t] = AW[\varphi;t]. \tag{4.4}$$

Thus, by (4.3)–(4.4), the companion matrix A of $c(x)$ is related to the Jordan matrix J by

$$A\,W[\varphi;t] = W[\varphi;t]\,J. \tag{4.5}$$

But note that the generalized Vandermonde matrix T is just $W[\varphi;0]$, so putting $t=0$ in the previous equation gives the asserted result (4.2):

$$AT = T\,J,$$

which shows that T is the change-of-basis matrix that brings the companion matrix into the Jordan canonical form J.

In the example discussed earlier with companion matrix (2.1) and Wronski matrix (2.2), one can check that

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 12 & -16 & 7 \end{bmatrix} \begin{bmatrix} e^{2t} & te^{2t} & e^{3t} \\ 2e^{2t} & (1+2t)e^{2t} & 3e^{3t} \\ 4e^{2t} & (4+4t)e^{2t} & 9e^{3t} \end{bmatrix} \equiv \begin{bmatrix} e^{2t} & te^{2t} & e^{3t} \\ 2e^{2t} & (1+2t)e^{2t} & 3e^{3t} \\ 4e^{2t} & (4+4t)e^{2t} & 9e^{3t} \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

for all t, in agreement with (4.5).

This change-of-basis property of the generalized Vandermonde matrix can be proved algebraically as follows. Let A be the companion matrix (4.1) of $c(x)$ as before. Repeatedly differentiate the identity

$$A \begin{bmatrix} 1 \\ x \\ x^2 \\ \vdots \\ x^{n-2} \\ x^{n-1} \end{bmatrix} = \begin{bmatrix} 1 \\ x \\ x^2 \\ \vdots \\ x^{n-2} \\ x^{n-1} \end{bmatrix} x - \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} c(x)$$

with respect to the indeterminant x and set $x = \lambda_\rho$. (Higher derivatives of $c(x)$ vanish at λ_ρ since the roots are repeated.) This will yield the $n \times n_\rho$ block of the generalized Vandermonde matrix.

This algebraic proof is tantamount to differentiating solutions to the differential equations with respect to the λ_ρ. Such a procedure can be made rigorous for multiple roots.

5. Another Example. It has been known since the early 1960s with the pioneering work of J. H. Wilkinson (cf. the account in Wilkinson and Reinsch (1971)) that various problems involving repeated (multiple) roots and/or eigenvalues are in general ill conditioned. In particular, in such cases the matrix exponential can be quite sensitive to perturbations in the data. To illustrate this phenomenon we briefly consider a slightly modified version of the *Mathematica* calculation of section 2 for the 3×3 matrix (2.1). We now decrease the $(3,1)$ entry by a small amount, 0.0001, from 12 to 11.9999, so the matrix (2.1) is replaced with

$$B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 11.9999 & -16 & 7 \end{bmatrix}. \tag{5.1}$$

The previous eigenvalues $\{2, 2, 3\}$ of (2.1) are replaced for (5.1) with the (approximate) values

$$\{1.99005, \quad 2.01005, \quad 2.9999\},$$

so the eigenvalues are now distinct.[6] The *Mathematica* calculation of section 2 based on the algorithm of section 1 easily computes for e^{Bt} the following approximate results with entries listed by columns to avoid clutter:

$$\text{col}_1(e^{Bt}) \approx \begin{bmatrix} 4.0017e^{2.9999t} + 298.538e^{1.99005t} - 301.54e^{2.01005t} \\ 12.0047e^{2.9999t} + 594.105e^{1.99005t} - 606.11e^{2.01005t} \\ 36.0129e^{2.9999t} + 1182.3e^{1.99005t} - 1218.31e^{2.01005t} \end{bmatrix},$$

$$\text{col}_2(e^{Bt}) \approx \begin{bmatrix} -4.0017e^{2.9999t} - 248.039e^{1.99005t} + 252.04e^{2.01005t} \\ -12.0047e^{2.9999t} - 493.609e^{1.99005t} + 506.614e^{2.01005t} \\ -36.0129e^{2.9999t} - 982.306e^{1.99005t} + 1018.32e^{2.01005t} \end{bmatrix}, \tag{5.2}$$

$$\text{col}_3(e^{Bt}) \approx \begin{bmatrix} 1.0004e^{2.9999t} + 49.5092e^{1.99005t} - 50.5096e^{2.01005t} \\ 3.0011e^{2.9999t} + 98.5257e^{1.99005t} - 101.527e^{2.01005t} \\ 9.003e^{2.9999t} + 196.071e^{1.99005t} - 204.074e^{2.01005t} \end{bmatrix}.$$

The matrix e^{Bt} can also be computed in this case by the standard elementary method based on diagonalizing the matrix, and the results agree with the present answer (5.2). One sees with (5.2) that a small change in the data from (2.1) to (5.1) results in large changes for the matrix exponential as compared with the results of section 2.

This example (5.1)–(5.2) provides a nice point of contact with the rich subject of numerical linear algebra. The thrust of the present paper is a conceptual understanding of the algorithm described in section 1. We do not intend to provide "the" ultimate algorithm for computing matrix exponentials—that is best left to numerical analysts, who guide us very well through the subtleties of numerical linear algebra, where one can often produce examples for which an algorithm may not be suited in finite precision arithmetic due to extreme sensitivity to changes in data.

6. Concluding Remarks. The classical paper of Moler and Van Loan (1978) treated 19 ways to compute matrix exponentials. The algorithm of Putzer (cf. section 9.12 of Apostol (1997)) can also be used to calculate matrix exponentials in terms of polynomials in A but without the simplicity and elegance of the method discussed here. This can be appreciated by beginning students of differential equations. Others (cf. Ziebur (1970), Fulmer (1975), Leonard (1996), and Elaydi and Harris (1998)) have discussed portions of the method described here, but these references did not develop this method with the ease of use shown to us by Harris.

Acknowledgments. Fillmore and Smith thank Robert Sacker, Louis Grimm, and the referees for their helpful comments on this paper, including the valuable suggestion of one referee to mention the sensitivity of the matrix exponential to changes in the data along with consequent pitfalls when using finite precision arithmetic.

[6] A change of +0.0001 has complex eigenvalues, which leads to a cluttered display, but the effect of either change is comparable.

REFERENCES

T. M. Apostol (1997), *Linear Algebra—A First Course, with Applications to Differential Equations*, Wiley, New York.

J. G. Broida and S. G. Williamson (1989), *Comprehensive Introduction to Linear Algebra*, Addison-Wesley, Reading, MA.

R. Bronson (1989), 2500 *Solved Problems in Differential Equations*, 2nd. ed., Schaum's Solved Problems Series, McGraw-Hill, New York, 1989.

S. N. Elaydi and W. A. Harris, Jr. (1998), *On the computation of* A^n, SIAM Rev., 40, pp. 965–971.

E. P. Fulmer (1975), *Computation of the matrix exponential*, Amer. Math. Monthly, 82, pp. 156–159.

I. E. Leonard (1996), *The matrix exponential*, SIAM Rev., 38, pp. 507–512.

C. Moler and C. Van Loan (1978), *Nineteen dubious ways to compute the exponential of a matrix*, SIAM Rev., 20, pp. 801–836.

H. W. Turnbull and A. C. Aitken (1961), *An Introduction to the Theory of Canonical Matrices*, Dover, New York.

J. H. Wilkinson and C. Reinsch (1971), *Linear Algebra*, Handbook for Automatic Computation 2, Springer-Verlag, Berlin.

A. D. Ziebur (1970), *On determining the structure of A by analyzing* e^{At}, SIAM Rev., 12, pp. 98–102.