The primary advantage of the method presented here is its pedagogical simplicity. It is not a practical method of computing e^{At} when A is a large matrix, since the method requires the computation of the first $n-1$ powers of A. This alone requires $n^4 - 2n^3$ multiplications, in general. One should note, however, that Kirchner's formula shares this disadvantage and that in computing n eigenvectors of A, one would do an equivalent amount of work.

References

1. E. A. Coddington and N. Levinson, Theory of Ordinary Differential Equations, McGraw-Hill, New York, 1955.

2. R. B. Kirchner, An explicit formula for e^{At}, this MONTHLY, 74 (1967) 1200–1204.

3. A. D. Ziebur, On determining the structure of A by analyzing e^{At}, SIAM Review, no. 1, 12 (1970) 98–102.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF PITTSBURGH, PITTSBURGH, PA 15213.

EXPLICIT FORMULAS FOR SOLUTIONS OF THE SECOND-ORDER MATRIX DIFFERENTIAL EQUATION $Y'' = AY$

TOM M. APOSTOL

Let A be a given $n \times n$ matrix and let $Y(t)$ denote an $n \times m$ matrix function of a real variable t. It is well known that the first-order matrix differential equation $Y'(t) = AY(t)$, with prescribed initial value $Y(0)$, has the unique solution $Y(t) = e^{tA}Y(0)$ on the interval $(-\infty, +\infty)$. (For example, see [**1**], p. 200.)

This note discusses the second-order matrix equation

$$Y''(t) = AY(t), \tag{1}$$

with prescribed initial values $Y(0)$ and $Y'(0)$. It is easy to show that (1) has a unique solution on $(-\infty, +\infty)$ given by

$$Y(t) = C(t)Y(0) + S(t)Y'(0), \tag{2}$$

where the matrix functions $C(t)$ and $S(t)$ are given by the everywhere-convergent power series

$$C(t) = \sum_{k=0}^{\infty} \frac{t^{2k}A^k}{(2k)!}, \qquad S(t) = \sum_{k=0}^{\infty} \frac{t^{2k+1}A^k}{(2k+1)!}. \tag{3}$$

The solution (2) can be obtained by trying a solution of the form $Y(t) = \sum_{k=0}^{\infty} C_k t^k$ with undetermined matrix coefficients, or by noting that (1) is equivalent to a first-order system of the form

$$Z'(t) = BZ(t), \tag{4}$$

where $Z(t)$ is a $2n \times m$ matrix with block form

$$Z(t) = \begin{bmatrix} Y(t) \\ Y'(t) \end{bmatrix},$$

and B is a $2n \times 2n$ matrix with block form

$$B = \begin{bmatrix} 0 & I \\ A & 0 \end{bmatrix}, \tag{5}$$

where I is the $n \times n$ identity matrix and 0 is the $n \times n$ zero matrix. The differential equation (4) has the solution

$$Z(t) = e^{tB}Z(0), \text{ where } Z(0) = \begin{bmatrix} Y(0) \\ Y'(0) \end{bmatrix}. \tag{6}$$

To calculate e^{tB} we note the powers of B are given by the block forms

$$B^{2k} = \begin{bmatrix} A^k & 0 \\ 0 & A^k \end{bmatrix}, \qquad B^{2k+1} = \begin{bmatrix} 0 & A^k \\ A^{k+1} & 0 \end{bmatrix}.$$

Using these in the exponential series $e^{tB} = \sum_{k=0}^{\infty} t^k B^k / k!$ and substituting in (6) we obtain $Z(t)$, from which we find (2). Incidentally, it is of interest to note that the eigenvalues of B are $\pm\sqrt{\lambda_1}, \cdots, \pm\sqrt{\lambda_n}$, where $\lambda_1, \cdots, \lambda_n$ are the eigenvalues of A.

The Cayley-Hamilton theorem suggests that each of $C(t)$ and $S(t)$ in (3) is a polynomial in A of degree $\leqq n-1$. We shall describe a simple procedure for calculating these polynomials directly without the need of calculating e^{tB} in (6). The method, which is a direct extension of that given by E. J. Putzer [2] for calculating exponential matrices, is much simpler than computing $Z(t)$ in (6) because the matrices in (6) are twice the size of those we shall require.

THEOREM. *Let $\lambda_1, \cdots, \lambda_n$ be the eigenvalues of an $n \times n$ matrix A, and let*

$$P_0(A) = I, \; P_k(A) = \prod_{j=1}^{k}(A - \lambda_j I), \text{ for } k = 1, 2, \cdots, n-1.$$

Then the functions $C(t)$ and $S(t)$ in (3) are given by the formulas

$$C(t) = \sum_{k=0}^{n-1} y_{k+1}(t) P_k(A), \qquad S(t) = \sum_{k=0}^{n-1} z_{k+1}(t) P_k(A), \tag{7}$$

where the functions $y_1, \cdots, y_n$ and $z_1, \cdots, z_n$ are determined recursively by the following triangular systems of second-order linear differential equations with constant coefficients:

$$\begin{cases} y_1''(t) = \lambda_1 y_1(t), & y_1(0) = 1, \quad y_1'(0) = 0, \\ y_{k+1}''(t) = \lambda_{k+1} y_{k+1}(t) + y_k(t), & y_{k+1}(0) = y_{k+1}'(0) = 0, \end{cases} \tag{8}$$

and

$$(9)\qquad \begin{cases} z_1'(t) = \lambda_1 z_1(t), & z_1(0) = 0, \quad z_1'(0) = 1, \\ z_{k+1}''(t) = \lambda_{k+1} z_{k+1}(t) + z_k(t), & z_{k+1}(0) = z_{k+1}'(0) = 0. \end{cases}$$

Proof. A direct extension of the argument given by Putzer in [2] for e^{tA} shows that the function $C(t)$ defined by (7) and (8) satisfies the differential equation $C''(t) = AC(t)$ and the initial conditions $C(0) = I$, $C'(0) = 0$, so $C(t)$ must be the same as the function in (3). Similarly, the function $S(t)$ defined by (7) and (9) satisfies $S''(t) = AS(t)$ with initial values $S(0) = 0$, $S'(0) = I$, so $S(t)$ is the same as the function in (3).

This theorem gives a useful and straightforward procedure for calculating explicitly the functions $C(t)$ and $S(t)$ for every $n \times n$ matrix A. The examples below give the results for all 2×2 matrices.

NOTE. If $\lambda = re^{i\theta}$ where $r \geqq 0$ and $-\pi < \theta \leqq \pi$, we write $\sqrt{\lambda}$ for the square root given by $r^{1/2}e^{i\theta/2}$.

Example 1. If A is a 2×2 matrix with eigenvalues λ, μ, where $\lambda \neq \mu$, then

$$C(t) = \frac{\lambda \cosh\sqrt{\mu}t - \mu\cosh\sqrt{\lambda}t}{\lambda - \mu} I + \frac{\cosh\sqrt{\lambda}t - \cosh\sqrt{\mu}t}{\lambda - \mu} A.$$

$$S(t) = \frac{\sinh\sqrt{\lambda}t}{\lambda\sqrt{\lambda}} A \quad \text{if } \lambda \neq 0,\ \mu = 0.$$

$$S(t) = \frac{1}{\lambda\mu(\lambda - \mu)}\{(\lambda^2\sqrt{\mu}\sinh\sqrt{\mu}t - \mu^2\sqrt{\lambda}\sinh\sqrt{\lambda}t)I$$
$$+ (\mu\sqrt{\lambda}\sinh\sqrt{\lambda}t - \lambda\sqrt{\mu}\sinh\sqrt{\mu}t)A\} \text{ if } \lambda \neq 0, \mu \neq 0.$$

Example 2. If A is a 2×2 matrix with eigenvalues λ, λ, where $\lambda \neq 0$, then

$$C(t) = (\cosh\sqrt{\lambda}t - \tfrac{1}{2}\sqrt{\lambda}t\sinh\sqrt{\lambda}t)I + \left(\frac{t}{2\sqrt{\lambda}}\sinh\sqrt{\lambda}t\right)A.$$

$$S(t) = \left(\frac{3}{2\sqrt{\lambda}}\sinh\sqrt{\lambda}t - \frac{t}{2}\cosh\sqrt{\lambda}t\right)I + \left(\frac{t}{2\lambda}\cosh\sqrt{\lambda}t - \frac{1}{2\lambda\sqrt{\lambda}}\sinh\sqrt{\lambda}t\right)A.$$

Example 3. If A is a 2×2 matrix with eigenvalues 0, 0, then

$$C(t) = I + \frac{t^2}{2}A, \qquad S(t) = tI + \frac{t^3}{6}A.$$

It should be noted that the foregoing method also works for the matrix differential equation of order r,

$$Y^{(r)}(t) = AY(t)$$

with prescribed initial values $Y(0)$, $Y'(0),\cdots,Y^{(r-1)}(0)$. The unique solution on $(-\infty,+\infty)$ is given by the following extension of (2):

$$Y(t) = \sum_{j=1}^{r} C_j(t)Y^{(j-1)}(0),$$

where each matrix function $C_j(t)$ is expressible in the form

$$C_j(t) = \sum_{k=1}^{n} y_{k,j}(t)P_{k-1}(A), \qquad (j = 1,2,\cdots,r)$$

generalizing (7). The scalar functions $y_{k,j}$ are determined recursively by r triangular systems of rth order linear differential equations with constant coefficients, exactly like those in (8) and (9), except that rth derivatives appear on the left and, for the first equation in each system, all the initial conditions are 0 except $y_{1,j}^{(j-1)}(0) = 1$.

References

1. T. M. Apostol, Calculus, Vol. 2, 2nd ed. Xerox, Waltham, Mass., 1969.
2. E. J. Putzer, Avoiding the Jordan canonical form in the discussion of linear systems with constant coefficients, this MONTHLY, 73 (1966) 2–7.

DEPARTMENT OF MATHEMATICS, CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CA 91109.

MATHEMATICAL EDUCATION

EDITED BY SHIRLEY HILL AND PAUL T. MIELKE

Material for this Department should be sent to Shirley Hill, Department of Mathematics, University of Missouri, Kansas City, MO 64110, or to Paul T. Mielke, Department of Mathematics, Wabash College, Crawfordsville, IN 47933.

A PROPOSAL FOR A PROFESSIONAL PROGRAM IN MATHEMATICS

T. I. SEIDMAN

To an overwhelming extent a student majoring in mathematics in the United States receives training predicated on the assumption that he or she is a potential academic mathematician. The standard track assumes a progression through graduate education culminating in a doctorate and followed by a career of teaching and the creation of new mathematics. Such an academic emphasis is, of course, entirely understandable: this is the track which most university faculty members best understand and which, perhaps, they are most likely to encourage for their students—after all, it is probably the track which they themselves followed.