THEOREM 7.10. *If A is an $n \times n$ matrix with all its eigenvalues equal to λ, then we have*

$$e^{tA} = e^{\lambda t}\sum_{k=0}^{n-1}\frac{t^k}{k!}(A - \lambda I)^k. \tag{7.36}$$

Proof. Since the matrices $\lambda t I$ and $t(A - \lambda I)$ commute we have

$$e^{tA} = e^{\lambda t I}e^{t(A-\lambda I)} = (e^{\lambda t}I)\sum_{k=0}^{\infty}\frac{t^k}{k!}(A - \lambda I)^k.$$

The Cayley-Hamilton theorem implies that $(A - \lambda I)^k = 0$ for every $k \geq n$, so the theorem is proved.

THEOREM 7.11. *If A is an $n \times n$ matrix with n distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then we have*

$$e^{tA} = \sum_{k=1}^{n} e^{t\lambda_k}L_k(A),$$

where $L_k(A)$ is a polynomial in A of degree $n - 1$ given by the formula

$$L_k(A) = \prod_{\substack{j=1\\ j\neq k}}^{n}\frac{A - \lambda_j I}{\lambda_k - \lambda_j} \quad \text{for} \quad k = 1, 2, \ldots, n.$$

Note: The polynomials $L_k(A)$ are called *Lagrange interpolation coefficients.*

Proof. We define a matrix function F by the equation

$$F(t) = \sum_{k=1}^{n} e^{t\lambda_k}L_k(A) \tag{7.37}$$

and verify that F satisfies the differential equation $F'(t) = AF(t)$ and the initial condition $F(0) = I$. From (7.37) we see that

$$AF(t) - F'(t) = \sum_{k=1}^{n} e^{t\lambda_k}(A - \lambda_k I)L_k(A).$$

By the Cayley-Hamilton theorem we have $(A - \lambda_k I)L_k(A) = 0$ for each k, so F satisfies the differential equation $F'(t) = AF(t)$.

To complete the proof we need to show that F satisfies the initial condition $F(0) = I$, which becomes

$$\sum_{k=1}^{n} L_k(A) = I. \tag{7.38}$$

A proof of (7.38) is outlined in Exercise 16 of Section 7.15.

The next theorem treats the case when A has two distinct eigenvalues, exactly one of which has multiplicity 1.

THEOREM 7.12. *Let A be an n x n matrix ($n \geq 3$) with two distinct eigenvalues λ and μ, where λ has multiplicity $n - 1$ and μ has multiplicity 1. Then we have*

$$e^{tA} = e^{\lambda t}\sum_{k=0}^{n-2}\frac{t^k}{k!}(A - \lambda I)^k + \left\{\frac{e^{\mu t}}{(\mu - \lambda)^{n-1}} - \frac{e^{\lambda t}}{(\mu - \lambda)^{n-1}}\sum_{k=0}^{n-2}\frac{t^k}{k!}(\mu - \lambda)^k\right\}(A - \lambda I)^{n-1}.$$

Proof. As in the proof of Theorem 7.10 we begin by writing

$$e^{tA} = e^{\lambda t}\sum_{k=0}^{\infty}\frac{t^k}{k!}(A - \lambda I)^k = e^{\lambda t}\sum_{k=0}^{n-2}\frac{t^k}{k!}(A - \lambda I)^k + e^{\lambda t}\sum_{k=n-1}^{\infty}\frac{t^k}{k!}(A - \lambda I)^k$$

$$= e^{\lambda t}\sum_{k=0}^{n-2}\frac{t^k}{k!}(A - \lambda I)^k + e^{\lambda t}\sum_{r=0}^{\infty}\frac{t^{n-1+r}}{(n-1+r)!}(A - \lambda I)^{n-1+r}.$$

Now we evaluate the series over r in closed form by using the Cayley-Hamilton theorem. Since

$$A - \mu I = A - \lambda I - (\mu - \lambda)I$$

we find

$$(A - \lambda I)^{n-1}(A - \mu I) = (A - \lambda I)^n - (\mu - \lambda)(A - \lambda I)^{n-1}.$$

The left member is 0 by the Cayley-Hamilton theorem so

$$(A - \lambda I)^n = (\mu - \lambda)(A - \lambda I)^{n-1}.$$

Using this relation repeatedly we find

$$(A - \lambda I)^{n-1+r} = (\mu - \lambda)^r(A - \lambda I)^{n-1}.$$

Therefore the series over r becomes

$$\sum_{r=0}^{\infty}\frac{t^{n-1+r}}{(n-1+r)!}(\mu - \lambda)^r(A - \lambda I)^{n-1} = \frac{1}{(\mu - \lambda)^{n-1}}\sum_{k=n-1}^{\infty}\frac{t^k}{k!}(\mu - \lambda)^k(A - \lambda I)^{n-1}$$

$$= \frac{1}{(\mu - \lambda)^{n-1}}\left\{e^{t(\mu-\lambda)} - \sum_{k=0}^{n-2}\frac{t^k}{k!}(\mu - \lambda)^k\right\}(A - \lambda I)^{n-1}.$$

This completes the proof.

The explicit formula in Theorem 7.12 can also be deduced by applying Putzer's method, but the details are more complicated.

The explicit formulas in Theorems 7.10, 7.11 and 7.12 cover all matrices of order $n \leq 3$. Since the 3 x 3 case often arises in practice, the formulas in this case are listed below for easy reference.

CASE 1. If a 3 x 3 matrix A has eigenvalues $\lambda, \lambda, \lambda$, then

$$e^{tA} = e^{\lambda t}\{I + t(A - \lambda I) + \tfrac{1}{2}t^2(A - \lambda I)^2\}.$$

CASE 2. If a 3 x 3 matrix A has distinct eigenvalues λ, μ, v, then

$$e^{tA} = e^{\lambda t}\frac{(A-\mu I)(A-\nu I)}{(\lambda-\mu)(\lambda-\nu)} + e^{\mu t}\frac{(A-\lambda I)(A-\nu I)}{(\mu-\lambda)(\mu-\nu)} + e^{\nu t}\frac{(A-\lambda I)(A-\mu I)}{(\nu-\lambda)(\nu-\mu)}.$$

CASE 3. If a 3 x 3 matrix A has eigenvalues λ, λ, μ, with $\lambda \neq \mu$, then

$$e^{tA} = e^{\lambda t}\{I + t(A-\lambda I)\} + \frac{e^{\mu t}-e^{\lambda t}}{(\mu-\lambda)^2}(A-\lambda I)^2 - \frac{te^{\lambda t}}{\mu-\lambda}(A-\lambda I)^2.$$

EXAMPLE. Compute e^{tA} when $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{bmatrix}$.

Solution. The eigenvalues of A are 1, 1, 2, so the formula of Case 3 gives us

(7.39) $$e^{tA} = e^{t}\{I + t(A - Z)\} + (e^{2t} - e^{t})(A - I)^2 - te^{t}(A - I)^2.$$

By collecting powers of A we can also write this as follows,

(7.40) $$e^{tA} = (-2te^{t} + e^{2t})I + \{(3t+2)e^{t} - 2e^{2t}\}A - \{(t+1)e^{t} - e^{2t}\}A^2.$$

At this stage we can calculate $(A - I)^2$ or A^2 and perform the indicated operations in (7.39) or (7.40) to write the result as a 3 x 3 matrix,

$$e^{tA} = \begin{bmatrix} -2te^{t} + e^{2t} & (3t+2)e^{t} - 2e^{2t} & -(t+1)e^{t} + e^{2t} \\ -2(t+1)e^{t} + 2e^{2t} & (3t+5)e^{t} - 4e^{2t} & -(t+2)e^{t} + 2e^{2t} \\ -2(t+2)e^{t} + 4e^{2t} & (3t+8)e^{t} - 8e^{2t} & -(t+4)e^{t} + 4e^{2t} \end{bmatrix}.$$

7.15 Exercises

For each of the matrices in Exercises 1 through 6, express e^{tA} as a polynomial in A.

1. $A = \begin{bmatrix} 5 & -2 \\ 4 & -1 \end{bmatrix}$.
2. $A = \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$.
3. $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}$.
4. $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -11 & -6 \end{bmatrix}$.
5. $A = \begin{bmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{bmatrix}$.
6. $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}$ 3.

7. (a) A 3 x 3 matrix A is known to have all its eigenvalues equal to 1. Prove that

$$e^{tA} = \tfrac{1}{2}e^{\lambda t}\{(\lambda^2t^2 - 2\lambda t + 2)Z + (-2\lambda t^2 + 2t)A + t^2A^2\}.$$

(b) Find a corresponding formula if A is a 4 x 4 matrix with all its eigenvalues equal to λ.

In each of Exercises 8 through 15, solve the system Y' = A Y subject to the given initial condition.

8. $A = \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$, $Y(0) = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$. 9. $A = \begin{bmatrix} -5 & 3 \\ -15 & 7 \end{bmatrix}$, $Y(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

10. $A = \begin{bmatrix} 3 & -1 & 1 \\ & 0 & \\ 1 & -1 & 2 \end{bmatrix}$, $Y(0) = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$. 11. $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$, $Y(0) = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$.

12. $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -11 & -6 \end{bmatrix}$, $Y(0) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. 13. $A = \begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix}$, $Y(0) = \begin{bmatrix} 8 \\ 0 \\ 0 \end{bmatrix}$.

14. $A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$, $Y(0) = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$. 15. $A = \begin{bmatrix} 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 \end{bmatrix}$, $Y(0) = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \end{bmatrix}$.

16. This exercise outlines a proof of Equation (7.38) used in the proof of Theorem 7.11, Let $L_k(\lambda)$ be the polynomial in λ of degree $n - 1$ defined by the equation

$$L_k(\lambda) = \prod_{\substack{j=1 \\ j \neq k}}^{n} \frac{\lambda - \lambda_j}{\lambda_k - \lambda_j},$$

where $\lambda_1, \ldots,$ I, are n distinct scalars.

(a) Prove that

$$L_k(\lambda_i) = \begin{cases} 0 & \text{if } \lambda_i \neq \text{A,,} \\ 1 & \text{if } \lambda_i = \lambda_k. \end{cases}$$

(b) Let $y_1, \ldots, y_n$ be n arbitrary scalars, and let

$$p(\lambda) = \sum_{k=1}^{n} y_k L_k(\lambda).$$

Prove that $p(\lambda)$ is the only polynomial of degree $\leq n - 1$ which satisfies the n equations

$$p(\lambda_k) = y_k \qquad \text{for } k = 1, 2, \ldots, n.$$

(c) Prove that $\sum_{k=1}^{n} L_k(\lambda) = 1$ for every λ, and deduce that for every square matrix A we have

$$\sum_{k=1}^{n} L_k(A) = I,$$

where Z is the identity matrix.

7.16 Nonhomogeneous linear systems with constant coefficients

We consider next the nonhomogeneous initial-value problem

$$Y'(t) = A\,Y(t) + Q(t), \qquad Y(a) = B, \tag{7.41}$$

on an interval J. Here A is an $n \times n$ constant matrix, Q is an n-dimensional vector function (regarded as an $n \times 1$ column matrix) continuous on J, and a is a given point in J. We can obtain an explicit formula for the solution of this problem by the same process used to treat the scalar case.

First we multiply both members of (7.41) by the exponential matrix e^{-tA} and rewrite the differential equation in the form

$$e^{-tA}\{Y'(t) - AY(t)\} = e^{-tA}Q(t). \tag{7.42}$$

The left member of (7.42) is the derivative of the product $e^{-tA}Y(t)$. Therefore, if we integrate both members of (7.42) from a to x, where $x \in J$, we obtain

$$e^{-xA}Y(x) - e^{-aA}Y(a) = \int_a^x e^{-tA}Q(t)\,dt\,.$$

Multiplying by e^{xA} we obtain the explicit formula (7.43) which appears in the following theorem.

THEOREM 7.13. *Let A be an $n \times n$ constant matrix and let Q be an n-dimensional vector function continuous on an interval J. Then the initial-value problem*

$$Y'(t) = A\,Y(t) + Q(t), \qquad Y(a) = B,$$

has a unique solution on J given by the explicit formula

$$Y(x) = e^{(x-a)A}B + e^{xA}\int_a^x e^{-tA}Q(t)\,dt. \tag{7.43}$$

As in the homogeneous case, the difficulty in applying this formula in practice lies in the calculation of the exponential matrices.

Note that the first term, $e^{(x-a)A}B$, is the solution of the homogeneous problem $Y'(t) = A\,Y(t)$, $Y(a) = B$. The second term is the solution of the nonhomogeneous problem

$$Y'(t) = A\,Y(t) + Q(t), \qquad Y(a) = 0.$$

We illustrate Theorem 7.13 with an example.

EXAMPLE. Solve the initial-value problem

$$Y'(t) = AY(t) + Q(t), \qquad Y(0) = B,$$

on the interval (- ∞ , + co), where

$$A = \begin{bmatrix} 2 & -1 & 1 \\ 0 & 3 & -1 \\ 2 & 1 & 3 \end{bmatrix}, \qquad Q(t) = \begin{bmatrix} e^{2t} \\ 0 \\ te^{2t} \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Solution. According to Theorem 7.13, the solution is given by

$$Y(x) = e^{xA}\int_0^x e^{-tA}Q(t)\,dt = \int_0^x e^{(x-t)A}Q(t)\,dt\,. \tag{7.44}$$

The eigenvalues of A are 2, 2, and 4. To calculate e^{xA} we use the formula of Case 3, Section 7.14, to obtain

$$\begin{aligned} e^{xA} &= e^{2x}\{I + x(A - 2I)\} + \tfrac{1}{4}(e^{4x} - e^{2x})(A - 2I)^2 - \tfrac{1}{2}xe^{2x}(A - 2I)^2 \\ &= e^{2x}\{I + x(A - 2I) + \tfrac{1}{4}(e^{2x} - 2x - 1)(A - 2I)^2\}\,. \end{aligned}$$

We can replace x by $x - t$ in this formula to obtain e $^{(x-t)A}$. Therefore the integrand in *(7.44)* is

$$\begin{aligned} e^{(x-t)A}Q(t) &= e^{2(x-t)}\{I + (x - t)(A - 2Z) + \tfrac{1}{4}[e^{2(x-t)} - 2(x - t) - 1](A - 2I)^2\}Q(t) \\ &= e^{2x}\begin{bmatrix} 1 \\ 0 \\ t \end{bmatrix} + (A - 2I)e^{2x}\begin{bmatrix} x - t \\ 0 \\ t(x - t) \end{bmatrix} + \frac{1}{4}(A - 2I)^2e^{2x}\begin{bmatrix} e^{2x}e^{-2t} - 2(x - t) - 1 \\ 0 \\ e^{2x}te^{-2t} - 2t(x - t) - t \end{bmatrix}. \end{aligned}$$

Integrating, we find

$$\begin{aligned} Y(x) = \int_0^x e^{(x-t)A}Q(t)\,dt &= e^{2x}\begin{bmatrix} x \\ 0 \\ \tfrac{1}{2}x^2 \end{bmatrix} + (A - 2I)e^{2x}\begin{bmatrix} \tfrac{1}{2}x^2 \\ 0 \\ \tfrac{1}{6}x^3 \end{bmatrix} \\ &\quad + \frac{1}{4}(A - 2I)^2e^{2x}\begin{bmatrix} \tfrac{1}{2}e^{2x} - \tfrac{1}{2} - x - x^2 \\ 0 \\ \tfrac{1}{4}e^{2x} - \tfrac{1}{4} - \tfrac{1}{2}x - \tfrac{1}{2}x^2 - \tfrac{1}{3}x^3 \end{bmatrix}. \end{aligned}$$

Since we have

$$A - 2I = \begin{bmatrix} 0 & -1 & 1 \\ 0 & 1 & -1 \\ 2 & 1 & 1 \end{bmatrix} \qquad \text{and} \qquad (A - 2I)^2 = \begin{bmatrix} 2 & 0 & 2 \\ -2 & 0 & -2 \\ 2 & 0 & 2 \end{bmatrix},$$

we find

$$\begin{aligned} Y(x) &= e^{2x}\begin{bmatrix} x \\ 0 \\ \tfrac{1}{2}x^2 \end{bmatrix} + e^{2x}\begin{bmatrix} \tfrac{1}{6}x^3 \\ -\tfrac{1}{6}x^3 \\ x^2 + \tfrac{1}{6}x^3 \end{bmatrix} + e^{2x}\begin{bmatrix} \tfrac{3}{8}e^{2x} - \tfrac{3}{8} - \tfrac{3}{4}x - \tfrac{3}{4}x^2 - \tfrac{1}{6}x^3 \\ -\tfrac{3}{8}e^{2x} + \tfrac{3}{8} + \tfrac{3}{4}x + \tfrac{3}{4}x^2 + \tfrac{1}{6}x^3 \\ \tfrac{3}{8}e^{2x} - \tfrac{3}{8} - \tfrac{3}{4}x - \tfrac{3}{4}x^2 - \tfrac{1}{6}x^3 \end{bmatrix} \\ &= e^{2x}\begin{bmatrix} \tfrac{3}{8}e^{2x} - \tfrac{3}{8} + \tfrac{1}{4}x - \tfrac{3}{4}x^2 \\ -\tfrac{3}{8}e^{2x} + \tfrac{3}{8} + \tfrac{3}{4}x + \tfrac{3}{4}x^2 \\ \tfrac{3}{8}e^{2x} - \tfrac{3}{8} - \tfrac{3}{4}x + \tfrac{3}{4}x^2 \end{bmatrix}. \end{aligned}$$

The rows of this matrix are the required functions y_1, y_2, y_3.

7.17 Exercises

1. Let Z be a solution of the nonhomogeneous system

$$Z'(r) = AZ(r) + Q(r),$$

on an interval J with initial value Z(a). Prove that there is only one solution of the nonhomogeneous system

$$Y'(t) = AY(t) + Q(t)$$

on J with initial value Y(a) and that it is given by the formula

$$Y(t) = Z(t) + e^{(t-a)A}\{Y(a) - Z(a)\}.$$

Special methods are often available for determining a particular solution Z(t) which resembles the given function Q(t). Exercises 2, 3, 5, and 7 indicate such methods for Q(t) = C, Q(r) = $e^{\alpha t}C$, $Q(t) = t^m C$, and $Q(t) = (\cos \alpha t)C + (\sin \alpha t)D$, where C and D are constant vectors. If the particular solution Z(t) so obtained does not have the required initial value, we modify Z(t) as indicated in Exercise 1 to obtain another solution $Y(t)$ with the required initial value.

2. (a) Let A be a constant $n \times n$ matrix, B and C constant n-dimensional vectors. Prove that the solution of the system

$$Y'(f) = AY(t) + c, \qquad Y(a) = B,$$

on $(-\infty, +\infty)$ is given by the formula

$$Y(x) = e^{(x-a)A}B + \left(\int_0^{x-a} e^{uA}\,du\right)C.$$

(b) If A is nonsingular, show that the integral in part (a) has the value $\{e^{(x-a)A} - I\}A^{-1}$.
(c) Compute Y(x) explicitly when

$$A = \begin{bmatrix} -1 & 2 \\ -2 & 3 \end{bmatrix}, \quad C = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad B = \begin{bmatrix} b \\ c \end{bmatrix}, \quad a = 0.$$

3. Let A be an $n \times n$ constant matrix, let B and C be n-dimensional constant vectors, and let α be a given scalar.
(a) Prove that the nonhomogeneous system $Z'(t) = AZ(t) + e^{\alpha t}C$ has a solution of the form $Z(t) = e^{\alpha t}B$ if, and only if, $(\alpha I - A)B = C$.
(b) If a is not an eigenvalue of A, prove that the vector B can always be chosen so that the system in (a) has a solution of the form $Z(t) = e^{\alpha t}B$.
(c) If α is not an eigenvalue of A, prove that every solution of the system $Y'(t) = A\,Y(i) + e^{\alpha t}C$ has the form $Y(t) = e^{tA}(Y(0) - B) + e^{\alpha t}B$, where $B = (al - A)^{-1}C$.

4. Use the method suggested by Exercise 3 to find a solution of the nonhomogeneous system $Y'(t) = AY(t) + e^{2t}C$, with

$$A = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \quad Y(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

5. Let A be an n x n constant matrix, let B and C be n-dimensional constant vectors, and let m be a positive integer.
(a) Prove that the nonhomogeneous system Y'(t) = A Y(t) + PC, Y(0) = B, has a particular solution of the form

$$Y(f) = B_0 + tB_1 + t^2B_2 + \cdots + t^mB_m,$$

where $B_0, B_1, \ldots, B_m$ are constant vectors, if and only if

$$C = -\frac{1}{m!}A^{m+1}B$$

Determine the coefficients $B_0, B_1, \ldots, B_m$ for such a solution.
(b) If A is nonsingular, prove that the initial vector B can always be chosen so that the system in (a) has a solution of the specified form.
6. Consider the nonhomogeneous system

$$y_1' = 3y_1 + y_2 + t^3$$
$$y_2' = 2y_1 + 2y_2 + t^3.$$

(a) Find a particular solution of the form Y(f) = $B_0 + tB_1 + t^2B_2 + t^3B_3$.
(b) Find a solution of the system with $y_1(0) = y_2(0) = 1$.
7. Let A be an $n \times n$ constant matrix, let B, C, D be n-dimensional constant vectors, and let α be a given nonzero real number. Prove that the nonhomogeneous system

$$Y'(t) = A\,Y(t) + (\cos \alpha t)C + (\sin \alpha t)D, \qquad Y(0) = B,$$

has a particular solution of the form

$$Y(f) = (\cos \alpha t)E + (\sin \alpha t)F,$$

where E and *Fare* constant vectors, if and only if

$$(A^2 + \alpha^2 I)B = -(AC + \alpha D).$$

Determine E and *Fin* terms of A, B, C for such a solution. Note that if $A^2 + \alpha^2 I$ is nonsingular, the initial vector B can always be chosen so that the system has a solution of the specified form.
8. (a) Find a particular solution of the nonhomogeneous system

$$y_1' = y_1 + 3y_2 + 4\sin 2t$$
$$y_2' = y_1 - y_2.$$

(b) Find a solution of the system with $y_1(0) = y_2(0) = 1$.

In each of Exercises 9 through 12, solve the nonhomogeneous system $Y'(t) = A\,Y(t) + Q(t)$ subject to the given initial condition.

9. $A = \begin{bmatrix} 4 & 1 \\ -2 & 1 \end{bmatrix}, \quad Q(t) = \begin{bmatrix} 0 \\ -2e^t \end{bmatrix}, \quad Y(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$

10. $A = \begin{bmatrix} -5 & -1 \\ 1 & -3 \end{bmatrix}, \quad Q(t) = \begin{bmatrix} e \\ e^{2t} \end{bmatrix}, \quad Y(0) = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$

11. $A = \begin{bmatrix} -5 & -1 \\ 2 & -3 \end{bmatrix}, \quad Q(t) = \begin{bmatrix} 7e^t - 27 \\ -3e^t + 12 \end{bmatrix}, \quad Y(0) = \begin{bmatrix} -\frac{1007}{442} \\ \frac{707}{221} \end{bmatrix}.$

12. $A = \begin{bmatrix} -1 & -1 & 0 \\ 0 & -1 & -1 \\ 0 & 0 & -1 \end{bmatrix}, \quad Q(t) = \begin{bmatrix} t^2 \\ 2t \\ t \end{bmatrix}, \quad Y(0) = \begin{bmatrix} 6 \\ -2 \\ 1 \end{bmatrix}.$

7.18 The general linear system $Y'(t) = P(t)Y(t) + Q(t)$

Theorem 7.13 gives an explicit formula for the solution of the linear system

$$Y'(t) = A\,Y(t) + Q(t), \qquad Y(a) = B,$$

where A is a constant $n \times n$ matrix and $Q(t)$, $Y(t)$ are $n \times 1$ column matrices. We turn now to the more general case

$$Y'(t) = P(t)\,Y(t) + Q(t), \qquad Y(a) = B, \tag{7.45}$$

where the $n \times n$ matrix $P(t)$ is not necessarily constant.

If P and Q are continuous on an open interval J, a general existence-uniqueness theorem which we shall prove in a later section tells us that for each a in J and each initial vector B there is exactly one solution to the initial-value problem (7.45). In this section we use this result to obtain a formula for the solution, generalizing Theorem 7.13.

In the scalar case ($n = 1$) the differential equation (7.45) can be solved as follows. We let $A(x) = \int_a^x P(t)\,dt$, then multiply both members of (7.45) by $e^{-A(t)}$ to rewrite the differential equation in the form

$$e^{-A(t)}\{Y'(t) - P(t)Y(t)\} = e^{-A(t)}Q(t). \tag{7.46}$$

Now the left member is the derivative of the product $e^{-A(t)}\,Y(t)$. Therefore, we can integrate both members from a to x, where a and x are points in J, to obtain

$$e^{-A(x)}Y(x) - e^{-A(a)}Y(a) = \int_a^x e^{-A(t)}Q(t)\,dt.$$

Multiplying by $e^{A(x)}$ we obtain the explicit formula

$$Y(x) = e^{A(x)}e^{-A(a)}Y(a) + e^{A(x)}\int_a^x e^{-A(t)}Q(t)\,dt. \tag{7.47}$$

The only part of this argument that does not apply immediately to matrix functions is the statement that the left-hand member of (7.46) is the derivative of the product $e^{-A(t)}\,Y(t)$. At this stage we used the fact that the derivative of $e^{-A(t)}$ is $-P(t)e^{-A(t)}$. In the scalar case this is a consequence of the following formula for differentiating exponential functions :

$$\text{If} \qquad E(t) = e^{A(t)}, \qquad \text{then} \qquad E'(t) = A'(t)e^{A(t)}.$$

Unfortunately, this differentiation formula is not always true when A is a matrix function. For example, it is false for the 2 x 2 matrix function $A(t) = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$. (See Exercise 7 of Section 7.12.) Therefore a modified argument is needed to extend Equation (7.47) to the matrix case.

Suppose we multiply each member of (7.45) by an unspecified n x n matrix $F(t)$. This gives us the relation

$$F(t)\,Y'(t) = F(t)P(t)\,Y(t) + F(t)Q(t)\,.$$

Now we add $F'(t)Y(t)$ to both members in order to transform the left member to the derivative of the product $F(t)\,Y(t)$. If we do this, the last equation gives us

$$\{F(t)\; r(t)\}' = \{F'(t) + F(t)P(t)\}\;Y(t) + F(t)Q(t)\;.$$

If we can choose the matrix $F(t)$ so that the sum $\{F'(t) + F(t)P(t)\}$ on the right is the zero matrix, the last equation simplifies to

$$\{F(t)\,Y(t)\}' = F(t)Q(t)\,.$$

Integrating this from a to x we obtain

$$F(x)Y(x) - F(a)Y(a) = \int_a^x F(t)Q(t)\,dt\;.$$

If, in addition, the matrix $F(x)$ is nonsingular, we obtain the explicit formula

$$(7.48) \qquad Y(x) = F(x)\text{-'}F(a)\;Y(a) + F(x)^{-1}\int_a^x F(t)Q(t)\;dt\;.$$

This is a generalization of the scalar formula (7.47). The process will work if we can find a n x n matrix function $F(t)$ which satisfies the matrix differential equation

$$F'(t) = -F(t)P(t)$$

and which is nonsingular.

Note that this differential equation is very much like the original differential equation (7.45) with $Q(t) = 0$, except that the unknown function $F(t)$ is a square matrix instead of a column matrix. Also, the unknown function is multiplied on the right by $-P(t)$ instead of on the left by $P(t)$.

We shall prove next that the differential equation for F always has a nonsingular solution. The proof will depend on the following existence theorem for homogeneous linear systems.

THEOREM 7.14. *Assume $A(t)$ is an $n \times n$ matrix function continuous on an open interval J. If $a \in J$ and if B is a given n-dimensional vector, the homogeneous linear system*

$$Y'(t) = A(t)\,Y(t), \qquad Y(a) = B,$$

has an n-dimensional vector solution Y on J.

A proof of Theorem 7.14 appears in Section 7.21. With the help of this theorem we can prove the following.

THEOREM 7.15. *Given an $n \times n$ matrix function P, continuous on an open interval J, and given any point a in J, there exists an $n \times n$ matrix function F which satisfies the matrix differential equation*

$$F'(x) = -F(x)P(x) \tag{7.49}$$

on J with initial value $F(a) = I$. Moreover, $F(x)$ is nonsingular for each x in J.

Proof. Let Y,(x) be a vector solution of the differential equation

$$Y_k'(x) = -P(x)^t Y_k(x)$$

on J with initial vector $Y,(a) = I_k$, where I_k is the kth column of the $n \times n$ identity matrix I. Here $P(x)^t$ denotes the transpose of $P(x)$. Let G(x) be the $n \times n$ matrix whose kth column is Y,(x). Then G satisfies the matrix differential equation

$$G'(x) = -P(x)^t G(x) \tag{7.50}$$

on J with initial value $G(a) = I$. Now take the transpose of each member of (7.50). Since the transpose of a product is the product of transposes in reverse order, we obtain

$$\{G'(x)\}” = -G(x)^t P(x).$$

Also, the transpose of the derivative G' is the derivative of the transpose G^t. Therefore the matrix $F(x) = G(x)^t$ satisfies the differential equation (7.49) with initial value $F(a) = I$.

Now we prove that $F(x)$ is nonsingular by exhibiting its inverse. Let H be the $n \times n$ matrix function whose kth column is the solution of the differential equation

$$Y'(x) = P(x)\,Y(x)$$

with initial vector $Y(a) =$ I,, the kth column of I. Then H satisfies the initial-value problem

$$H'(x) = P(x)H(x), \qquad H(a) = I,$$

on J. The product $F(x)H(x)$ has derivative

$$F(x)H'(x) + F'(x)H(x) = F(x)P(x)H(x) - F(x)P(x)H(x) = O$$

for each x in J. Therefore the product $F(x)H(x)$ is constant, $F(x)H(x) = F(a)H(a) = I$, so H(x) is the inverse of $F(x)$. This completes the proof.

The results of this section are summarized in the following theorem.

THEOREM 7.16. *Given an n x n matrix function P and an n-dimensional vector function Q, both continuous on an open interval J, the solution of the initial-value problem*

$$(7.51) \qquad r'(x) = P(x)\ Y(x) + Q(x), \qquad Y(a) = B,$$

on J is given by the formula

$$(7.52) \qquad Y(x) = F(x)\text{-`}Y(a) + F(x)^{-1} \int_a^x F(t)Q(t)\ dt.$$

The n x n matrix F(x) is the transpose of the matrix whose kth column is the solution of the initial-value problem

$$(7.53) \qquad Y'(x) = -P(x)^t Y(x), \qquad Y(a) = I_k,$$

where I, is the kth column of the identity matrix I.

Although Theorem 7.16 provides an explicit formula for the solution of the general linear system (7.51), the formula is not always a useful one for calculating the solution because of the difficulty involved in determining the matrix function *F.* The determination of F requires the solution of n homogeneous linear systems (7.53). The next section describes a power-series method that is sometimes used to solve homogeneous linear systems.

We remind the reader once more that the proof of Theorem 7.16 was based on Theorem 7.14, the existence theorem for homogeneous linear systems, which we have not yet proved.

7.19 A power-series method for solving homogeneous linear systems

Consider a homogeneous linear system

$$(7.54) \qquad Y'(x) = A(x)\ Y(x)\ , \qquad Y(0) = B,$$

in which the given n x n matrix $A(x)$ has a power-series expansion in x convergent in some open interval containing the origin, say

$$A(x) = A_0 + xA_1 + x^2A_2 + \cdots + x^kA_k + \ldots, \qquad \text{for } |x| < r_1,$$

where the coefficients A_0, A,, A,, . . . are given n x n matrices. Let us try to find a power-series solution of the form

$$Y(x) = B_0 + xB_1 + x^2B_2 + \cdots + x^kB_k + \cdots,$$

with vector coefficients B_0, B_1, B_2, ... Since Y(0) = B_0, the initial condition will be satisfied by taking B_0 = B, the prescribed initial vector. To determine the remaining coefficients we substitute the power series for Y(x) in the differential equation and equate coefficients of like powers of x to obtain the following system of equations:

(7.55) $$B_1 = A_0B, \qquad (k+1)B_{k+1} = \sum_{r=0}^{k} A_rB_{k-r} \qquad \text{for } k = 1, 2, \ldots .$$

These equations can be solved in succession for the vectors B_1, B_2, ... If the resulting power series for Y(x) converges in some interval $|x| < r_2$, then Y(x) will be a solution of the initial-value problem (7.54) in the interval $|x| < r$, where r = min $\{r_1, r_2\}$.

For example, if A(x) is a constant matrix A, then A, $= A$ and $A_k = O$ for $k \geq 1$, so the system of equations in (7.55) becomes

$$B_1 = AB, \qquad (k+1)B_{k+1} = AB_k \qquad \text{for } k \geq 1 .$$

Solving these equations in succession we find

$$B_k = \frac{1}{k!}A^kB \qquad \text{for } k \geq 1 .$$

Therefore the series solution in this case becomes

$$Y(x) = B + \sum_{k=1}^{\infty} \frac{x^k}{k!} A^kB = e^{xA}B .$$

This agrees with the result obtained earlier for homogeneous linear systems with constant coefficients.

7.20 Exercises

1. Let p be a real-valued function and Q an n x 1 matrix function, both continuous on an interval J. Let A be an $n \times n$ constant matrix. Prove that the initial-value problem

$$Y'(x) = p(x)A\,Y(x) + Q(x)\,, \qquad \mathbf{Y(u)} = B,$$

has the solution

$$Y(x) = e^{q(x)A}B + e^{q(x)A}\int_a^x e^{-q(t)A}Q(t)\,dt$$

on J, where $q(x) = \int_a^x p(t)\,dt$.

2. Consider the special case of Exercise 1 in which A is nonsingular, a = 0, $p(x)$ = 2x, and Q(x) = xC, where C is a constant vector. Show that the solution becomes

$$Y(x) = e^{x^2A}(B + \tfrac{1}{2}A^{-1}C) - \tfrac{1}{2}A^{-1}C .$$

3. Let $A(t)$ be an $n \times n$ matrix function and let E(t) = $e^{A(t)}$. Let Q(t), Y(r), and B be n x 1 column matrices. Assume that

$$E'(t) = A'(t)E(t)$$

on an open interval J. If a $\in J$ and if A' and Q are continuous on J, prove that the initial-value problem

$$Y'(t) = A'(t)Y(t) + Q(t), \qquad Y(u) = B,$$

has the following solution on J:

$$Y(x) = e^{A(x)}e^{-A(a)}B + e^{A(x)}\int_a^x e^{-A(t)}Q(t)\,dt.$$

4. Let E(t) = $e^{A(t)}$. This exercise describes examples of matrix functions A(t) for which E'(t) = $A'(t)E(t)$.
(a) Let A(t) = $t^r A$, where A is an $n \times n$ constant matrix and r is a positive integer. Prove that E'(t) = $A'(t)E(t)$ on $(-\infty, \infty)$.
(b) Let A(t) be a polynomial in t with matrix coefficients, say

$$A(t) = \sum_{r=0}^{m} PA,.,$$

where the coefficients commute, $A_rA_s = A_sA_r$ for all r and s. Prove that E'(t) = $A'(t)E(t)$ on $(-\infty, \infty)$.
(c) Solve the homogeneous linear system

$$Y'(r) = (I + tA)Y(t), \qquad Y(0) = B$$

on the interval $(-\infty, \infty)$, where A is an $n \times n$ constant matrix.
5. Assume that the $n \times n$ matrix function $A(x)$ has a power-series expansion convergent for $|x| < r$. Develop a power-series procedure for solving the following homogeneous linear system of second order:

$$Y''(x) = A(x)Y(x), \qquad \text{with } Y(0) = B, \qquad Y'(0) = c.$$

6. Consider the second-order system Y''(x) + $A\,Y(x) = 0$, with Y(0) = B, $Y'(0) = C$, where A is a constant $n \times n$ matrix. Prove that the system has the power-series solution

$$Y(x) = \left(I + \sum_{k=1}^{\infty} \frac{(-1)^k x^{2k} A^k}{(2k)!}\right)B + \left(xI + \sum_{k=1}^{\infty} \frac{(-1)^k x^{2k+1} A^k}{(2k+1)!}\right)C$$

convergent for $-\infty < x < +\infty$.

7.21 Proof of the existence theorem by the method of successive approximations

In this section we prove the existence and uniqueness of a solution for any homogeneous linear system

$$Y'(t) = A(t)\,Y(t), \tag{7.56}$$

where $A(t)$ is an $n \times n$ matrix function, continuous on an open interval J. We shall prove that for any point a in J and any given initial-vector B there exists exactly one solution $Y(t)$ on J satisfying the initial condition $Y(a) = B$.