III. First-Order Systems of Ordinary Differential Equations
4. Linear Systems: Matrix Exponentials

C. David Levermore
Department of Mathematics
University of Maryland

August 2, 2022

Contents

4. Linear Systems: Matrix Exponentials
- 4.1. Introduction
- 4.2. Matrices and Polynomials
  - 4.2.1. Polynomials of a Matrix
  - 4.2.2. Matrix Key Identity
  - 4.2.3. Cayley-Hamilton Theorem and Characteristic Polynomials
- 4.3. Natural Fundamental Set Method
  - 4.3.1. Basic Steps
  - 4.3.2. Natural Fundamental Sets of Solutions
  - 4.3.3. Computing Matrix Exponentials
- 4.4. Two-by-Two Matrix Exponential Formulas
- 4.5. Natural Fundamental Sets from Green Functions (optional)
  - 4.5.1. The Recipe
  - 4.5.2. Its Justification
- 4.6. General Matrix Exponential Formula (not covered)
  - 4.6.1. The Formula
  - 4.6.2. Its Justification
- 4.7. Hermite Interpolation Method (not covered)
  - 4.7.1. Computing Hermite Interpolants
  - 4.7.2. Application to Matrix Exponentials
  - 4.7.3. Application to Other Functions of Matrices

Exercises on Matrix Exponentials

Navigation to other Chapters

Return to Main Webpage

4. Linear Systems: Matrix Exponentials

4.1. **Introduction.** Consider the vector-valued initial-value problem

$$\mathbf{x}' = \mathbf{A}\mathbf{x}\,, \qquad \mathbf{x}(0) = \mathbf{x}_I\,, \tag{4.1}$$

where the coefficient matrix $\mathbf{A}$ is a *constant*, $n\times n$, real matrix. Let $\mathbf{\Phi}(t)$ be the *natural fundamental matrix* associated with (4.1) for the initial time 0. It satisfies the matrix-valued initial-value problem

$$\mathbf{\Phi}' = \mathbf{A}\mathbf{\Phi}\,, \qquad \mathbf{\Phi}(0) = \mathbf{I}\,, \tag{4.2}$$

where $\mathbf{I}$ is the $n\times n$ identity matrix.

Because $\mathbf{A}$ is constant $\mathbf{\Phi}(t)$ satisfies the properties

$$\begin{aligned} &\text{(i)} \quad \mathbf{\Phi}(t+s) = \mathbf{\Phi}(t)\mathbf{\Phi}(s) \qquad &&\text{for every } t \text{ and } s \text{ in } \mathbb{R}\,,\\ &\text{(ii)} \quad \mathbf{\Phi}(t)\mathbf{\Phi}(-t) = \mathbf{I} &&\text{for every } t \text{ in } \mathbb{R}\,. \end{aligned} \tag{4.3}$$

Property (i) follows because both sides satisfy the matrix-valued initial-value problem

$$\mathbf{\Psi}' = \mathbf{A}\mathbf{\Psi}\,, \qquad \mathbf{\Psi}(0) = \mathbf{\Phi}(s)\,,$$

and therefore are equal. Property (ii) follows by setting $s = -t$ in property (i) and using the fact $\mathbf{\Phi}(0) = \mathbf{I}$.

Properties (i) and (ii) look like the properties satisfied by the real-valued exponential function e^{at}, but they are satisfied by the matrix-valued function $\mathbf{\Phi}(t)$. They motivate the following definition.

Definition 4.1. The *matrix exponential* of $\mathbf{A}$ is the natural fundamental matrix associated with (4.1) for the initial-time 0, which is the solution $\mathbf{\Phi}(t)$ of the matrix-valued initial-value problem (4.2). It is usually denoted either as $e^{t\mathbf{A}}$ or as $\exp(t\mathbf{A})$.

When recast in the $e^{t\mathbf{A}}$ notation, the defining matrix-valued initial-value problem (4.2) becomes

$$\frac{\mathrm{d}}{\mathrm{d}t} e^{t\mathbf{A}} = \mathbf{A} e^{t\mathbf{A}}\,, \qquad e^{t\mathbf{A}}\Big|_{t=0} = \mathbf{I}\,, \tag{4.4}$$

and the properties (4.3) become the rules for matrix exponentials:

$$\begin{aligned} &\text{(i)} \quad e^{(t+s)\mathbf{A}} = e^{t\mathbf{A}} e^{s\mathbf{A}} \qquad &&\text{for every } t \text{ and } s \text{ in } \mathbb{R}\,,\\ &\text{(ii)} \quad e^{t\mathbf{A}} e^{-t\mathbf{A}} = \mathbf{I} &&\text{for every } t \text{ in } \mathbb{R}\,. \end{aligned} \tag{4.5}$$

Remark. These rules are remarkable. The first shows that matrix multiplication commutes for $e^{t\mathbf{A}}$ and $e^{s\mathbf{A}}$. Specifically, for every t and s in $\mathbb{R}$ we have

$$e^{t\mathbf{A}} e^{s\mathbf{A}} = e^{(t+s)\mathbf{A}} = e^{(s+t)\mathbf{A}} = e^{s\mathbf{A}} e^{t\mathbf{A}}\,.$$

This should be surprising because we know that in general $\mathbf{BC} \neq \mathbf{CB}$. The second rule shows that the inverse of $e^{t\mathbf{A}}$ is $e^{-t\mathbf{A}}$. In other words, the inverse of $e^{t\mathbf{A}}$ is computed by simply replacing t with $-t$ in $e^{t\mathbf{A}}$.

Remark. We see from the matrix-valued initial-value problem (4.4) that

$$\frac{\mathrm{d}}{\mathrm{d}t}e^{t\mathbf{A}}\Big|_{t=0} = \mathbf{A}e^{t\mathbf{A}}\Big|_{t=0} = \mathbf{A}\,.$$

This fact gives a quick way to recover $\mathbf{A}$ from $e^{t\mathbf{A}}$.

By differentiating the matrix-valued differential equation in (4.4), we see that

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}e^{t\mathbf{A}} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}}{\mathrm{d}t}e^{t\mathbf{A}}\right) = \frac{\mathrm{d}}{\mathrm{d}t}\big(\mathbf{A}e^{t\mathbf{A}}\big) = \mathbf{A}\frac{\mathrm{d}}{\mathrm{d}t}e^{t\mathbf{A}} = \mathbf{A}^2e^{t\mathbf{A}}\,.$$

By repeated differentiation we see that for every positive integer k we have

$$\frac{\mathrm{d}^k}{\mathrm{d}t^k}e^{t\mathbf{A}} = \mathbf{A}^k e^{t\mathbf{A}}\,. \tag{4.6}$$

By evaluating this at $t = 0$ we obtain

$$\frac{\mathrm{d}^k}{\mathrm{d}t^k}e^{t\mathbf{A}}\Big|_{t=0} = \mathbf{A}^k e^{t\mathbf{A}}\Big|_{t=0} = \mathbf{A}^k\,. \tag{4.7}$$

Remark. Formula (4.7) implies that the Taylor expansion of $e^{t\mathbf{A}}$ about $t = 0$ is

$$e^{t\mathbf{A}} = \sum_{k=0}^{\infty}\frac{1}{k!}\,t^k\mathbf{A}^k = \mathbf{I} + t\mathbf{A} + \tfrac{1}{2}t^2\mathbf{A}^2 + \tfrac{1}{6}t^3\mathbf{A}^3 + \tfrac{1}{24}t^4\mathbf{A}^4 + \cdots\,, \tag{4.8}$$

where we define $\mathbf{A}^0 = \mathbf{I}$. Recall that the Taylor expansion of e^{at} is

$$e^{at} = \sum_{k=0}^{\infty}\frac{1}{k!}\,a^kt^k = 1 + at + \tfrac{1}{2}a^2t^2 + \tfrac{1}{6}a^3t^3 + \tfrac{1}{24}a^4t^4 + \cdots\,.$$

Motivated by the similarity of these expansions, many textbooks define $e^{t\mathbf{A}}$ by the infinite series (4.8). This approach dances around questions about the convergence of the infinite series — questions that are seldom favorites of students. Worse, this approach does not lend itself to finding effecient methods for computing $e^{t\mathbf{A}}$. We will avoid such difficulties because we defined $e^{t\mathbf{A}}$ to be the solution of the matrix-valued initial-value problem (4.4), which is also more central to the material of this course.

Remark. It is important to realize that for every $n \geq 2$ the exponential matrix $e^{t\mathbf{A}}$ is *not* the matrix obtained by exponentiating each entry! This means that

$$\text{if}\quad \mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad\text{then}\quad e^{t\mathbf{A}} \neq \begin{pmatrix} e^{ta_{11}} & e^{ta_{12}} \\ e^{ta_{21}} & e^{ta_{22}} \end{pmatrix}\,.$$

In particular, the initial condition of (4.4) implies that $e^{0\mathbf{A}} = \mathbf{I}$, while the matrix on the right-hand side above has every entry equal to 1 when $t = 0$.

This chapter and the next present methods by which we can compute $e^{t\mathbf{A}}$. Given $e^{t\mathbf{A}}$ then the general methods of Chapter 2 can be applied to find general solutions of homogeneous or nonhomogeneous first-order linear systems, or to solve initial-value problems associated with such systems. For example, a general solution of the homogeneous first-order system $\mathbf{x}' = \mathbf{A}\mathbf{x}$ is given by

$$\mathbf{x}(t) = e^{t\mathbf{A}}\mathbf{c}\,, \qquad \text{where } \mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}.$$

Because the coefficient matrix $\mathbf{A}$ is constant we can use the rules of matrix exponentials (4.5) to simplify many of these methods. For example, the natural fundamental matrix for the inital time t_I becomes simply

$$\Phi(t)\Phi(t_I)^{-1} = e^{t\mathbf{A}}e^{-t_I\mathbf{A}} = e^{(t-t_I)\mathbf{A}}\,,$$

while the Green matrix $\mathbf{G}(t,s)$ becomes simply

$$\mathbf{G}(t,s) = \Phi(t)\Phi(s)^{-1} = e^{t\mathbf{A}}e^{-s\mathbf{A}} = e^{(t-s)\mathbf{A}}\,.$$

Notice that computing the matrix inverse and matrix multiplication is easy; we simply replace the t in $e^{t\mathbf{A}}$ with $-s$ to get $\Phi(s)^{-1}$ and with $t-s$ to get $\Phi(t)\Phi(s)^{-1}$. Therefore the solution of the initial-value problem

$$\mathbf{x}' = \mathbf{A}\mathbf{x} + \mathbf{f}(t)\,, \qquad \mathbf{x}(t_I) = \mathbf{x}^I\,,$$

becomes simply

$$\mathbf{x}(t) = e^{(t-t_I)\mathbf{A}}\,\mathbf{x}^I + \int_{t_I}^{t} e^{(t-s)\mathbf{A}}\,\mathbf{f}(s)\,\mathrm{d}s = e^{(t-t_I)\mathbf{A}}\,\mathbf{x}^I + e^{t\mathbf{A}}\int_{t_I}^{t} e^{-s\mathbf{A}}\,\mathbf{f}(s)\,\mathrm{d}s\,.$$

Given any real $n\times n$ matrix $\mathbf{A}$, there are many ways to compute $e^{\mathbf{A}t}$ that are far easier than evaluating the infinite series (4.8). In the next chapter we will show how to do it using eigen methods. In this chapter we present methods that are generally much faster to carry out when n is not too large. More specifically, we will show how to compute $e^{t\mathbf{A}}$ when $\mathbf{A}$ is an $n\times n$ matrix using natural fundundamental sets and using Hermite interpolation. We will also give explicit formulas for $e^{t\mathbf{A}}$ when $\mathbf{A}$ is a 2×2 matrix.

4.2. **Matrices and Polynomials.** The notion of *matrix polynomials* will play a central role in rest of the course. In this section we introduce them, give their relationship to *matrix exponentials*, and state an important theorem about them. In subsequent sections of this chapter and in the next chapter we will use these tools to construct matrix exponentials.

4.2.1. *Polynomials of a Matrix.* We now define what it means to evaluate a polynomial at a square matrix.

Definition 4.2. Given any m^{th} degree polynomial

$$p(z) = p_0 z^m + p_1 z^{m-1} + \cdots + p_{m-1} z + p_m \,,$$

and any $n\times n$ matrix $\mathbf{A}$, we define the $n\times n$ matrix $p(\mathbf{A})$ by

$$p(\mathbf{A}) = p_0 \mathbf{A}^m + p_1 \mathbf{A}^{m-1} + \cdots + p_{m-1}\mathbf{A} + p_m \mathbf{I} \,. \tag{4.9}$$

Remark. We know what is meant by $\mathbf{A}^k$ for every positive integer k and every square matrix $\mathbf{A}$. Definition 4.2 extends this notion to a notion of polynomials of matrices. Notice that $p(\mathbf{A})$ is a square matrix with the same dimensions as $\mathbf{A}$.

This definition has some natural consequences.

(1) If $p(z)$, $q(z)$, and $r(z)$ are polynomials such that $p(z) = q(z) + r(z)$ then for any square matrix $\mathbf{A}$ we have

$$p(\mathbf{A}) = q(\mathbf{A}) + r(\mathbf{A}) \,. \tag{4.10a}$$

(2) If $p(z)$, $q(z)$, and $r(z)$ are polynomials such that $p(z) = q(z)r(z)$ then for any square matrix $\mathbf{A}$ we have

$$p(\mathbf{A}) = q(\mathbf{A})r(\mathbf{A}) = r(\mathbf{A})q(\mathbf{A}) \,. \tag{4.10b}$$

Remark. Property (4.10b) might have been surprising had you recalled that matrix multiplication generally does not commute — i.e. in general $\mathbf{BC} \neq \mathbf{CB}$. It shows matrix multiplication does commute when both maticies are polynomials of a third matrix — i.e. $\mathbf{BC} = \mathbf{CB}$ when $\mathbf{B} = q(\mathbf{A})$ and $\mathbf{C} = r(\mathbf{A})$ for some polynomials $q(z)$ and $r(z)$ and some square matrix $\mathbf{A}$.

Remark. Property (4.10b) for matrix polynomials is similar to rule (i) for matrix exponentials (4.5). Later we will see the reason for this similarity.

Repeated application of property (4.10b) implies that if a polynomial $p(z)$ can be factored as

$$p(z) = (z - r_1)^{m_1}(z - r_2)^{m_2} \cdots (z - r_k)^{m_k} \,,$$

then for any square matrix $\mathbf{A}$ the matrix $p(\mathbf{A})$ can be factored as

$$p(\mathbf{A}) = (\mathbf{A} - r_1\mathbf{I})^{m_1}(\mathbf{A} - r_2\mathbf{I})^{m_2} \cdots (\mathbf{A} - r_k\mathbf{I})^{m_k} \,.$$

Example. Confirm property (4.10b) for the polynomial $q(z) = z^2 - 5z + 6 = (z - 2)(z - 3)$ and the matrix

$$\mathbf{C} = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix} .$$

Then

$$q(\mathbf{C}) = \mathbf{C}^2 - 5\mathbf{C} + 6\mathbf{I} = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}^2 - 5\begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix} + 6\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} 13 & 12 \\ 12 & 13 \end{pmatrix} - \begin{pmatrix} 15 & 10 \\ 10 & 15 \end{pmatrix} + \begin{pmatrix} 6 & 0 \\ 0 & 6 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix},$$
$$\mathbf{C} - 2\mathbf{I} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \qquad \mathbf{C} - 3\mathbf{I} = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}.$$

Because

$$(\mathbf{C} - 2\mathbf{I})(\mathbf{C} - 3\mathbf{I}) = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix},$$
$$(\mathbf{C} - 3\mathbf{I})(\mathbf{C} - 2\mathbf{I}) = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix},$$

we see that

$$q(\mathbf{C}) = \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix} = (\mathbf{C} - 2\mathbf{I})(\mathbf{C} - 3\mathbf{I}) = (\mathbf{C} - 3\mathbf{I})(\mathbf{C} - 2\mathbf{I}),$$

which confirms property (4.10b). □

4.2.2. *Matrix Key Identity.* Recall that for every polynomial $p(z)$ in the form

$$p(z) = p_0 z^m + p_1 z^{m-1} + \cdots + p_{m-1} z + p_m,$$

where $p_0 \neq 0$, the m^{th}-order linear differential operator

$$p(\mathrm{D}) = p_0 \mathrm{D}^m + p_1 \mathrm{D}^{m-1} + \cdots + p_{m-1}\mathrm{D} + p_m,$$

satisfies the Key Identity, which for every complex number z states

$$p(\mathrm{D})e^{zt} = p(z)e^{zt}.$$

This identity was used to construct a fundamental set of solutions for the homogeneous equation $p(\mathrm{D})y = 0$ from the roots of $p(z)$. It also was used to construct particular solutions to the nonhomogeneous equation $p(\mathrm{D})y = f(t)$ when the forcing is in characteristic form. Here we give a matrix version of the Key Identity which will relate matrix exponentials to matrix polynomials. It will be used to compute the matrix exponential $e^{t\mathbf{A}}$, which is the natural fundamental matrix for the homogeneous system $\mathbf{x}' = \mathbf{A}\mathbf{x}$.

Recall from (4.6) that for every square matrix $\mathbf{A}$ and every positive integer k we have the identity

$$\mathrm{D}^k e^{t\mathbf{A}} = \mathbf{A}^k e^{t\mathbf{A}}, \qquad \text{where } \mathrm{D} = \frac{\mathrm{d}}{\mathrm{d}t}.$$

It follows that

$$\begin{aligned} p(\mathrm{D})e^{t\mathbf{A}} &= \big(p_0\mathrm{D}^m + p_1\mathrm{D}^{m-1} + \cdots + p_{m-1}\mathrm{D} + p_m\big)e^{t\mathbf{A}} \\ &= p_0\mathrm{D}^m e^{t\mathbf{A}} + p_1\mathrm{D}^{m-1}e^{t\mathbf{A}} + \cdots + p_{m-1}\mathrm{D}e^{t\mathbf{A}} + p_m e^{t\mathbf{A}} \\ &= p_0\mathbf{A}^m e^{t\mathbf{A}} + p_1\mathbf{A}^{m-1}e^{t\mathbf{A}} + \cdots + p_{m-1}\mathbf{A}e^{t\mathbf{A}} + p_m e^{t\mathbf{A}} \\ &= \big(p_0\mathbf{A}^m + p_1\mathbf{A}^{m-1} + \cdots + p_{m-1}\mathbf{A} + p_m\mathbf{I}\big)e^{t\mathbf{A}} = p(\mathbf{A})e^{t\mathbf{A}}. \end{aligned}$$

Therefore for any square matrix $\mathbf{A}$ and any polynomial $p(z)$ we have

$$p(\mathrm{D})e^{t\mathbf{A}} = p(\mathbf{A})e^{t\mathbf{A}}\,. \tag{4.11}$$

This is the *Matrix Key Identity.*

The Matrix Key Identity (4.11) will be used to compute $e^{t\mathbf{A}}$. The key will be to show that for every square matrix $\mathbf{A}$ there is a polynomial $p(z)$ such that $p(\mathbf{A}) = \mathbf{0}$. For such a polynomial the Matrix Key Identity shows that

$$p(\mathrm{D})e^{t\mathbf{A}} = \mathbf{0}\,.$$

Therefore each entry of $e^{t\mathbf{A}}$ satifies the equation $p(\mathrm{D})y = 0$. But this is an m^{th}-order homogeneous linear equation with constant coefficients, which we can solve.

4.2.3. *Cayley-Hamiltion Theorem and Characteristic Polynomials.* As mentioned above, for any given square matrix it will be useful to be able to find polynomials with the following property.

Definition 4.3. A polynomial $p(z)$ *annihilates* a square matrix $\mathbf{A}$ if $p(\mathbf{A}) = \mathbf{0}$.

Remark. If a polynomial $q(z)$ annihilates a square matrix $\mathbf{A}$ then so does every polynomial $p(z) = q(z)r(z)$ where $r(z)$ is *any* polynomial because property (4.10b) shows

$$p(\mathbf{A}) = q(\mathbf{A})\,r(\mathbf{A}) = \mathbf{0}\,r(\mathbf{A}) = \mathbf{0}\,.$$

So if one polynomial annihilates a square matrix then many polynomials do.

The *Cayley-Hamilton Theorem* says that every square matrix is annihilated by its so-called *characteristic polynomial.*

Definition 4.3. The *characteristic polynomial* of an $n\times n$ matrix $\mathbf{A}$ is

$$p_{\mathbf{A}}(z) = \det(z\,\mathbf{I} - \mathbf{A})\,. \tag{4.12}$$

This polynomial has degree n and is easy to compute when n is not large.

Remark. Sometimes the characteristic polynomial of $\mathbf{A}$ is defined to be $\det(\mathbf{A} - z\mathbf{I})$. Because $\det(z\mathbf{I} - \mathbf{A}) = (-1)^n \det(\mathbf{A} - z\mathbf{I})$, this alternative coincides with our definition when n is even and is its negative when n is odd. Both conventions are common. We have chosen the convention that makes $p_{\mathbf{A}}(z)$ monic — i.e. that makes the coefficient of z^n equal to 1. What matters most about $p_{\mathbf{A}}(z)$ is its roots and their multiplicity, which are the same for both conventions.

Remark. The roots of $p_{\mathbf{A}}(z)$ are the *eigenvalues* of $\mathbf{A}$. They will play a much bigger role in the next chapter. When $n > 2$ these roots might not be easy to find.

The characteristic polynomial of any 2×2 matrix can be computed quickly. Consider the general 2×2 matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\,.$$

Then by definition (4.12), its characteristic polynomial is

$$\begin{aligned} p_{\mathbf{A}}(z) = \det(z\,\mathbf{I} - \mathbf{A}) = \det\begin{pmatrix} z - a_{11} & -a_{12} \\ -a_{21} & z - a_{22} \end{pmatrix} \\ &= (z - a_{11})(z - a_{22}) - a_{21}a_{12} \\ &= z^2 - (a_{11} + a_{22})z + (a_{11}a_{22} - a_{21}a_{12})\,. \end{aligned}$$

This may be expressed as the formula

$$p_{\mathbf{A}}(z) = z^2 - \mathrm{tr}(\mathbf{A})z + \det(\mathbf{A})\,, \tag{4.13}$$

where $\mathrm{tr}(\mathbf{A}) = a_{11} + a_{22}$ is the *trace* of $\mathbf{A}$.

Remark. Formula (4.13) is handy for computing the characteristic polynomial of a $2{\times}2$ matrix. However, it applies only to $2{\times}2$ matrices! For an $n{\times}n$ matrix with $n > 2$ we can use definition (4.12) to compute its characteristic polynomial. In that case one of the methods from Chapter 3 of Part II can be used to compute the determinant.

Example. Compute the characteristic polynomial of

$$\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}\,.$$

Solution. Because $\mathbf{A}$ is $2{\times}2$, by formula (4.13) its characteristic polynomial is

$$p(z) = z^2 - \mathrm{tr}(\mathbf{A})z + \det(\mathbf{A}) = z^2 - 6z + 5 = (z-1)(z-5)\,.$$

Example. Compute the characteristic polynomial of

$$\mathbf{B} = \begin{pmatrix} -3 & -1 \\ 4 & 1 \end{pmatrix}\,.$$

Solution. Because $\mathbf{B}$ is $2{\times}2$, by formula (4.13) its characteristic polynomial is

$$p(z) = z^2 - \mathrm{tr}(\mathbf{B})z + \det(\mathbf{B}) = z^2 + 2z + 1 = (z+1)^2\,.$$

Example. Compute the characteristic polynomial of

$$\mathbf{C} = \begin{pmatrix} 6 & -5 \\ 5 & -2 \end{pmatrix}\,.$$

Solution. Because $\mathbf{C}$ is $2{\times}2$, by formula (4.13) its characteristic polynomial is

$$p(z) = z^2 - \mathrm{tr}(\mathbf{C})z + \det(\mathbf{C}) = z^2 - 4z + 13 = (z-2)^2 + 3^2\,.$$

We are now ready to state the Cayley-Hamilton Theorem.

Theorem 4.1. (Cayley-Hamilton Theorem) For every $n{\times}n$ matrix $\mathbf{A}$

$$p_{\mathbf{A}}(\mathbf{A}) = \mathbf{0}\,. \tag{4.14}$$

In other words, every square matrix is annihilated by its characteristic polynomial.

Reason. We will not prove this theorem for general $n\times n$ matrices. However, it is easy to verify for 2×2 matrices by a direct calculation. Formulas (4.9) and (4.13) show that

$$\begin{aligned} p_{\mathbf{A}}(\mathbf{A}) &= \mathbf{A}^2 - \mathrm{tr}(\mathbf{A})\mathbf{A} + \det(\mathbf{A})\mathbf{I} \\ &= \mathbf{A}^2 - (a_{11}+a_{22})\mathbf{A} + (a_{11}a_{22}-a_{21}a_{12})\mathbf{I} \\ &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^2 - (a_{11}+a_{22})\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + (a_{11}a_{22}-a_{21}a_{12})\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} a_{11}^2 + a_{12}a_{21} & (a_{11}+a_{22})a_{12} \\ (a_{11}+a_{22})a_{21} & a_{21}a_{12} + a_{22}^2 \end{pmatrix} - \begin{pmatrix} (a_{11}+a_{22})a_{11} & (a_{11}+a_{22})a_{12} \\ (a_{11}+a_{22})a_{21} & (a_{11}+a_{22})a_{22} \end{pmatrix} \\ &\quad + \begin{pmatrix} a_{11}a_{22}-a_{21}a_{12} & 0 \\ 0 & a_{11}a_{22}-a_{21}a_{12} \end{pmatrix} = \mathbf{0}\,, \end{aligned}$$

which verifies (4.14) for 2×2 matrices. □

Remark. Neither Cayley nor Hamilton proved the general theorem stated above. In 1853 Hamilton gave an analog of it in the setting of quaternions, which corresponds to certain 4×4 real matrices. He had introduced quaterions as a setting in which to study geometry. Vectors and matrices were being developed at the same time by others as an alternative setting. In 1858 Cayley stated the result for 3×3 matrices, but gave the proof only for 2×2 matrices with complex entries. His proof was exactly the calculation that we did above. There is a similar proof for 3×3 matrices that is longer. This is certainly the proof discovered by Cayley, but it gives no insight into how to prove the general theorem. In 1878 Frobenius proved the general theorem for $n\times n$ matrices stated above. Perhaps his name should be attached to the theorem too.

4.3. **Natural Fundamental Set Method.** The task of computing $e^{t\mathbf{A}}$ can be reduced to that of computing some powers of $\mathbf{A}$ and the natural fundamental set of solutions for a higher-order differential operator associated with $\mathbf{A}$.

4.3.1. *Basic Steps.* Let $\mathbf{A}$ be any square real matrix. Suppose that we can find a polynomial $p(z)$ of degree m that annihilates $\mathbf{A}$. Then we know from the Matrix Key Identity (4.11) that $e^{t\mathbf{A}}$ satisfies

(4.15a) $$p(\mathrm{D})e^{t\mathbf{A}} = p(\mathbf{A})e^{t\mathbf{A}} = \mathbf{0}\,.$$

This says that each entry of $e^{t\mathbf{A}}$ is a solution of the m^{th}-order scalar homogeneous linear differential equation $p(\mathrm{D})y = 0$. Moreover, we see from (4.7) that each entry of $e^{t\mathbf{A}}$ satisfies initial conditions that can be read off from

(4.15b) $$\mathrm{D}^k e^{t\mathbf{A}}\Big|_{t=0} = \mathbf{A}^k e^{t\mathbf{A}}\Big|_{t=0} = \mathbf{A}^k \qquad \text{for } k = 0, 1, \cdots, m-1\,.$$

Here we understand that $\mathrm{D}^0 = 1$ and that $\mathbf{A}^0 = \mathbf{I}$. We thereby see that the entries of $e^{t\mathbf{A}}$ can be computed by solving the matrix-valued initial-value problem (4.15).

If $y_1(t)$, $y_2(t)$, $\cdots$, $y_m(t)$ is any fundamental set of solutions to the m^{th}-order equation $p(\mathrm{D})y = 0$ then a general solution of equation (4.15a) is

$$e^{t\mathbf{A}} = \mathbf{C}_1 y_1(t) + \mathbf{C}_2 y_2(t) + \cdots + \mathbf{C}_m y_m(t)\,,$$

where $\mathbf{C}_1$, $\mathbf{C}_2$, $\cdots$, $\mathbf{C}_m$ are constant $n\times n$ matrices. We can determine these constant matrices by imposing the initial conditions (4.15b), which is not an easy task in general. However, this becomes easy if we work with $N_0(t)$, $N_1(t)$, $\cdots$, $N_{m-1}(t)$, the natural fundamental set of solutions associated with the m^{th}-order differential operator $p(\mathrm{D})$ and the initial time 0. Recall that the solution of the general initial-value problem

$$p(\mathrm{D})y = 0\,, \qquad y(0) = y_0\,, \quad y'(0) = y_1\,, \quad \cdots \quad y^{(m-1)}(0) = y_{m-1}\,,$$

is expressed in terms of this natural fundamental set as

$$y(t) = N_0(t)y_0 + N_1(t)y_1 + \cdots + N_{m-1}(t)y_{m-1}\,.$$

Therefore the solution of the matrix-valued initial-value problem (4.15) is given by

$$e^{t\mathbf{A}} = N_0(t)\mathbf{I} + N_1(t)\mathbf{A} + \cdots + N_{m-1}(t)\mathbf{A}^{m-1}\,. \tag{4.16}$$

This so-called *natural fundamental set method* for computing $e^{t\mathbf{A}}$ has three steps.

1. Find a polynomial $p(z)$ that annihilates $\mathbf{A}$. Let m be the degree of $p(z)$.
2. Compute $N_0(t)$, $N_1(t)$, $\cdots$, $N_{m-1}(t)$, the natural fundamental set of solutions associated with the m^{th}-order differential operator $p(\mathrm{D})$ and the initial time 0.
3. Compute the matrix exponential $e^{t\mathbf{A}}$ by the formula (4.16).

The Cayley-Hamilton Theorem says that the first step can be done by picking $p(z)$ to be the characteristic polynomial of $\mathbf{A}$. The second step can be done by methods from either Chapter 2 or Chapter 9 of Part II, which we review in the next subsection. Often this is the step that will require the most work. The third step is straightforward.

Remark. The Cayley-Hamilton Theorem insures that for every $n\times n$ matrix $\mathbf{A}$ there is a polynomial $p(z)$ of degree n that annihilates $\mathbf{A}$ — namely, $p_{\mathbf{A}}(z)$, the characteristic polynomial of $\mathbf{A}$ given by definition (4.12). If we can find a polynomial of smaller degree that also annihilates $\mathbf{A}$ then the computation of $e^{t\mathbf{A}}$ will be shortened. However, there is no such polynomial for most matrices that you will see in this course. For example, there is no such polynomial when

- $\mathbf{A}$ is a 2×2 matrix that is not a scalar multiple of $\mathbf{I}$,
- $\mathbf{A}$ is a matrix whose characteristic polynomial has only simple roots.

When $\mathbf{A}$ is an $n\times n$ matrix for some $n > 2$ and its characteristic polynomial $p_{\mathbf{A}}(z)$ has roots that are not simple, we can seek a polynomial $p(z)$ of smaller degree that also annihilates $\mathbf{A}$ guided by the following linear algebra facts.

- Every root of $p_{\mathbf{A}}(z)$ must be a root of any polynomial $p(z)$ that annihilates $\mathbf{A}$, but it might have smaller multiplicity.
- If $\mathbf{A}$ is either symmetric ($\mathbf{A}^{\mathrm{T}} = \mathbf{A}$), skew-symmetric ($\mathbf{A}^{\mathrm{T}} = -\mathbf{A}$), or normal ($\mathbf{A}^{\mathrm{T}}\mathbf{A} = \mathbf{A}\mathbf{A}^{\mathrm{T}}$) then a polynomial with the same roots as $p_{\mathbf{A}}(z)$ but with each root being simple will annihilate $\mathbf{A}$.

Remark. Formula (4.16) shows that $e^{t\mathbf{A}}$ is a time-dependent matrix polynomial of $\mathbf{A}$ whose coefficients are $N_0(t)$, $N_1(t)$, $\cdots$, $N_{m-1}(t)$. Therefore property (4.10b) implies that for any polynomial $q(z)$ we have

$$q(\mathbf{A})\,e^{t\mathbf{A}} = e^{t\mathbf{A}}\,q(\mathbf{A})\,.$$

In particular, this confirms rule (i) for matrix exponentials given in (4.5).

4.3.2. *Natural Fundamental Sets of Solutions.* After we have found a polynomial $p(z)$ of degree $m \leq n$ that annihilates $\mathbf{A}$, the second step of our method is to compute $N_0(t)$, $N_1(t)$, $\cdots$, $N_{m-1}(t)$, the natural fundamental set of solutions associated with the m^{th}-order differential operator $p(\mathrm{D})$ and initial time 0.

In Chapter 2 of Part II we saw how to compute this natural fundamental set by solving the general initial-value problem

$$p(\mathrm{D})y = 0\,, \qquad y(0) = y_0\,, \quad y'(0) = y_1\,, \quad \cdots\,, \quad y^{(m-1)}(0) = y_{m-1}\,. \tag{4.17}$$

Because $p(\mathrm{D})y = 0$ is an m^{th}-order linear equation with constant coefficients, we do this by first factoring $p(z)$ and using our recipe to generate a fundamental set of solutions $Y_1(t)$, $Y_2(t)$, $\cdots$, $Y_m(t)$. We then determine c_1, c_2, $\cdots$, c_m so that the general solution,

$$y(t) = c_1Y_1(t) + c_2Y_2(t) + \cdots + c_mY_m(t)\,,$$

satifies the general initial conditions of (4.17). By grouping the terms that multiply each y_k, we can express the solution of the general initial-value problem (4.17) as

$$y(t) = N_0(t)y_0 + N_1(t)y_1 + \cdots + N_{m-1}(t)y_{m-1}\,.$$

We can then read off $N_0(t)$, $N_1(t)$, $\cdots$, $N_{m-1}(t)$ from this expression.

In Chapter 9 of Part II we saw how to compute this natural fundamental set from the Green function $g(t)$ associated with the m^{th}-order differential operator $p(\mathrm{D})$ and initial time 0. The Green function $g(t)$ can be found either by solving the initial-value problem

$$p(\mathrm{D})g = 0\,, \qquad g(0) = \cdots = g^{(m-2)}(0) = 0\,, \quad g^{(m-1)}(0) = 1\,, \tag{4.18a}$$

or by taking the inverse Laplace transform

$$g(t) = \mathcal{L}^{-1}\left[\frac{1}{p(s)}\right](t)\,. \tag{4.18b}$$

Then if the polynomial $p(z)$ has the form

$$p(z) = z^m + p_1z^{m-1} + \cdots + p_{m-1}z + p_m\,,$$

the natural fundamental set of solutions is given by the recipe

$$\begin{aligned} N_{m-1}(t) &= g(t)\,, \\ N_{m-2}(t) &= N'_{m-1}(t) + p_1g(t)\,, \\ &\ \vdots \\ N_0(t) &= N'_1(t) + p_{m-1}g(t)\,. \end{aligned} \tag{4.18c}$$

An alternative presentation of recipe (4.18c) is given in Section 4.5.

4.3.3. *Computing Matrix Exponentials.* After we have the natural fundamental set $N_0(t)$, $N_1(t)$, $\cdots$, $N_{m-1}(t)$, the third step in our method is to compute $e^{t\mathbf{A}}$ by formula (4.16). If $m = 2$ this only requires doing one matrix addition because only $\mathbf{I}$ and $\mathbf{A}$ appear in formula (4.16). If $m > 2$ this requires computing $\mathbf{A}^k$ up to $k = m - 1$. This can require up to $(m-2)n^3$ multiplications, which grows fast as m and n get large. (Often $m = n$.) However, for small matrices like the ones commonly faced in this course, it is generally efficient.

We now illustrate this method. The first four examples will have $m = n$ with $n = 2$ or $n = 3$. The fifth example will have $n = 4$ and $m = 3$. For each example the second step of computing the natural fundamental set will require the most work. We will show how to do this both by recipe (4.18c) and by solving the general initial-value problem (4.17).

Example. Compute $e^{t\mathbf{A}}$ for

$$\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix} .$$

Solution. Because $\mathbf{A}$ is 2×2, by formula (4.13) its characteristic polynomial is

$$p(z) = z^2 - \mathrm{tr}(\mathbf{A})z + \det(\mathbf{A}) = z^2 - 6z + 5 = (z-5)(z-1) .$$

It has the simple roots 5 and 1, so we will use it.

Below we will show by using recipe (4.18c) and by solving the general initial-value problem (4.17) that the natural fundamental set associated with $p(\mathrm{D})$ is

$$N_0(t) = \frac{5e^t - e^{5t}}{4} , \qquad N_1(t) = \frac{e^{5t} - e^t}{4} .$$

Given this natural fundamental set, formula (4.16) yields

$$\begin{aligned} e^{t\mathbf{A}} &= N_0(t)\mathbf{I} + N_1(t)\mathbf{A} \\ &= \frac{5e^t - e^{5t}}{4} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{e^{5t} - e^t}{4} \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} e^{5t} + e^t & e^{5t} - e^t \\ e^{5t} - e^t & e^{5t} + e^t \end{pmatrix} . \end{aligned}$$

From the Green Function. By the partial fraction identity

$$\frac{1}{p(s)} = \frac{1}{s^2 - 6s + 5} = \frac{1}{(s-5)(s-1)} = \frac{\frac{1}{4}}{s-5} + \frac{-\frac{1}{4}}{s-1} ,$$

a Laplace transform table shows that the Green function $g(t)$ is given by

$$\begin{aligned} g(t) &= \mathcal{L}^{-1}\left[\frac{1}{p(s)}\right](t) = \mathcal{L}^{-1}\left[\frac{1}{s^2 - 6s + 5}\right](t) \\ &= \tfrac{1}{4}\mathcal{L}^{-1}\left[\frac{1}{s-5}\right](t) - \tfrac{1}{4}\mathcal{L}^{-1}\left[\frac{1}{s-1}\right](t) = \tfrac{1}{4}e^{5t} - \tfrac{1}{4}e^t . \end{aligned}$$

Then because $p(s) = s^2 - 6s + 5$, recipe (4.18c) yields the natural fundamental set

$$\begin{aligned} N_1(t) &= g(t) = \tfrac{1}{4}e^{5t} - \tfrac{1}{4}e^t , \\ N_0(t) &= N_1'(t) - 6g(t) = \tfrac{5}{4}e^{5t} - \tfrac{1}{4}e^t - \tfrac{6}{4}e^{5t} + \tfrac{6}{4}e^t \\ &= \tfrac{5}{4}e^t - \tfrac{1}{4}e^{5t} . \end{aligned}$$

From the General Initial-Value Problem. The general initial-value problem (4.17) is

$$y'' - 6y' + 5y = 0\,, \qquad y(0) = y_0\,, \quad y'(0) = y_1\,.$$

A general solution of the differential equation is $y(t) = c_1 e^{5t} + c_2 e^{t}$. Because

$$y'(t) = 5c_1 e^{5t} + c_2 e^{t}\,,$$

the general initial conditions imply

$$\begin{aligned} y_0 &= y(0) = c_1 e^0 + c_2 e^0 = c_1 + c_2\,, \\ y_1 &= y'(0) = 5c_1 e^0 + c_2 e^0 = 5c_1 + c_2\,. \end{aligned}$$

Upon solving this system for c_1 and c_2 we find that

$$c_1 = \frac{y_1 - y_0}{4}\,, \qquad c_2 = \frac{5y_0 - y_1}{4}\,.$$

Therefore the solution of the general initial-value problem is

$$y(t) = \frac{y_1 - y_0}{4}\, e^{5t} + \frac{5y_0 - y_1}{4}\, e^{t} = \frac{5e^{t} - e^{5t}}{4}\, y_0 + \frac{e^{5t} - e^{t}}{4}\, y_1\,.$$

We then read off that the natural fundamental set is

$$N_0(t) = \frac{5e^{t} - e^{5t}}{4}\,, \qquad N_1(t) = \frac{e^{5t} - e^{t}}{4}\,.$$

Example. Compute $e^{t\mathbf{B}}$ for

$$\mathbf{B} = \begin{pmatrix} -3 & -1 \\ 4 & 1 \end{pmatrix}\,.$$

Solution. Because $\mathbf{B}$ is $2{\times}2$, by formula (4.13) its characteristic polynomial is

$$p(z) = z^2 - \mathrm{tr}(\mathbf{B})z + \det(\mathbf{B}) = z^2 + 2z + 1 = (z+1)^2\,.$$

It has the double root -1. Because $\mathbf{B}$ is not a scalar multiple of $\mathbf{I}$, we will use this polynomial.

Below we will show by using recipe (4.18c) and by solving the general initial-value problem (4.17) that the natural fundamental set associated with $p(\mathrm{D})$ is

$$N_0(t) = (1+t)e^{-t}\,, \qquad N_1(t) = te^{-t}\,.$$

Given this natural fundamental set, formula (4.16) yields

$$\begin{aligned} e^{t\mathbf{B}} &= N_0(t)\mathbf{I} + N_1(t)\mathbf{B} \\ &= (1+t)e^{-t}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + te^{-t}\begin{pmatrix} -3 & -1 \\ 4 & 1 \end{pmatrix} = e^{-t}\begin{pmatrix} 1-2t & -t \\ 4t & 1+2t \end{pmatrix}\,. \end{aligned}$$

From the Green Function. A Laplace transform table shows that the Green function $g(t)$ is given by

$$g(t) = \mathcal{L}^{-1}\left[\frac{1}{p(s)}\right](t) = \mathcal{L}^{-1}\left[\frac{1}{(s+1)^2}\right](t) = t\, e^{-t}\,.$$

Then because $p(s) = s^2 + 2s + 1$, recipe (4.18c) yields the natural fundamental set

$$\begin{aligned} N_1(t) &= g(t) = t\,e^{-t}t\,, \\ N_0(t) &= g'(t) + 2g(t) = e^{-t} - t\,e^{-t} + 2t\,e^{-t} = e^{-t} + te^{-t}\,. \end{aligned}$$

From the General Initial-Value Problem. The general initial-value problem (4.17) is

$$y'' + 2y' + y = 0\,, \qquad y(0) = y_0\,, \quad y'(0) = y_1\,.$$

A general solution of the differential equation is $y(t) = c_1 e^{-t} + c_2 te^{-t}$. Because

$$y'(t) = -c_1 e^{-t} - c_2 te^{-t} + c_2 e^{-t}\,,$$

the general initial conditions imply

$$\begin{aligned} y_0 &= y(0) = c_1 e^0 + c_2 0 e^0 = c_1\,, \\ y_1 &= y'(0) = -c_1 e^0 - c_2 0 e^0 + c_2 e^0 = -c_1 + c_2\,. \end{aligned}$$

Upon solving this system for c_1 and c_2 we find that

$$c_1 = y_0\,, \qquad c_2 = y_0 + y_1\,.$$

Therefore the solution of the general initial-value problem is

$$\begin{aligned} y(t) &= y_0 e^{-t} + (y_0 + y_1)te^{-t} \\ &= (1+t)e^{-t}y_0 + te^{-t}y_1\,. \end{aligned}$$

We then read off that the natural fundamental set is

$$N_0(t) = (1+t)e^{-t}\,, \qquad N_1(t) = te^{-t}\,.$$

Example. Compute $e^{t\mathbf{C}}$ for

$$\mathbf{C} = \begin{pmatrix} 6 & -5 \\ 5 & -2 \end{pmatrix}\,.$$

Solution. Because $\mathbf{C}$ is 2×2, by formula (4.13) its characteristic polynomial is

$$p(z) = z^2 - \mathrm{tr}(\mathbf{C})z + \det(\mathbf{C}) = z^2 - 4z + 13 = (z-2)^2 + 3^2\,.$$

It has the simple roots $2 \pm i3$, so we will use it.

Below we will show by using recipe (4.18c) and by solving the general initial-value problem (4.17) that the natural fundamental set associated with $p(\mathrm{D})$ is

$$N_0(t) = e^{2t}\big(\cos(3t) - \tfrac{2}{3}\sin(3t)\big)\,, \qquad N_1(t) = \tfrac{1}{3}e^{2t}\sin(3t)\,.$$

Given this natural fundamental set, formula (4.16) yields

$$\begin{aligned} e^{t\mathbf{C}} &= N_0(t)\mathbf{I} + N_1(t)\mathbf{C} \\ &= e^{2t}\big(\cos(3t) - \tfrac{2}{3}\sin(3t)\big)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + e^{2t}\tfrac{1}{3}\sin(3t)\begin{pmatrix} 6 & -5 \\ 5 & -2 \end{pmatrix} \\ &= e^{2t}\begin{pmatrix} \cos(3t) + \frac{4}{3}\sin(3t) & -\frac{5}{3}\sin(3t) \\ \frac{5}{3}\sin(3t) & \cos(3t) - \frac{4}{3}\sin(3t) \end{pmatrix}\,. \end{aligned}$$

From the Green Function. A Laplace transform table shows that the Green function $g(t)$ is given by

$$\begin{aligned} g(t) = \mathcal{L}^{-1}\left[\frac{1}{p(s)}\right](t) &= \mathcal{L}^{-1}\left[\frac{1}{s^2-4s+13}\right](t) \\ &= \mathcal{L}^{-1}\left[\frac{1}{(s-2)^2+3^2}\right](t) = \tfrac{1}{3}e^{2t}\sin(3t)\,. \end{aligned}$$

Then because $p(s) = s^2 - 4s + 13$, recipe (4.18c) yields the natural fundamental set

$$\begin{aligned} N_1(t) &= g(t) = \tfrac{1}{3}e^{2t}\sin(3t)\,, \\ N_0(t) &= N_1'(t) - 4g(t) = \tfrac{2}{3}e^{2t}\sin(3t) + e^{2t}\cos(3t) - \tfrac{4}{3}e^{2t}\sin(3t) \\ &= e^{2t}\cos(3t) - \tfrac{2}{3}e^{2t}\sin(3t)\,. \end{aligned}$$

From the General Initial-Value Problem. The general initial-value problem (4.17) is

$$y'' - 4y' + 13y = 0\,, \qquad y(0) = y_0\,, \quad y'(0) = y_1\,.$$

A general solution of the differential equation is $y(t) = c_1 e^{2t}\cos(3t) + c_2 e^{2t}\sin(3t)$. Because

$$y'(t) = 2c_1e^{2t}\cos(3t) - 3c_1e^{2t}\sin(3t) + 2c_2e^{2t}\sin(3t) + 3c_2e^{2t}\cos(3t)\,,$$

the general initial conditions imply

$$\begin{aligned} y_0 &= y(0) = c_1e^0\cos(0) + c_2e^0\sin(0) = c_1\,, \\ y_1 &= y'(0) = 2c_1e^0\cos(0) - 3c_1e^0\sin(0) + 2c_2e^0\sin(0) + 3c_2e^0\cos(0) = 2c_1 + 3c_2\,. \end{aligned}$$

Upon solving this system for c_1 and c_2 we find that

$$c_1 = y_0\,, \qquad c_2 = \frac{y_1 - 2y_0}{3}\,.$$

Therefore the solution of the general initial-value problem is

$$\begin{aligned} y(t) &= y_0e^{2t}\cos(3t) + \frac{y_1-2y_0}{3}\,e^{2t}\sin(3t) \\ &= \big(e^{2t}\cos(3t) - \tfrac{2}{3}e^{2t}\sin(3t)\big)\,y_0 + \tfrac{1}{3}e^{2t}\sin(3t)y_1\,. \end{aligned}$$

We then read off that the natural fundamental set is

$$N_0(t) = e^{2t}\big(\cos(3t) - \tfrac{2}{3}\sin(3t)\big)\,, \qquad N_1(t) = e^{2t}\tfrac{1}{3}\sin(3t)\,.$$

Example. Compute $e^{t\mathbf{A}}$ for

$$\mathbf{A} = \begin{pmatrix} 0 & 2 & -1 \\ -2 & 0 & 2 \\ 1 & -2 & 0 \end{pmatrix}.$$

Solution. Because $\mathbf{A}$ is $3{\times}3$, we cannot use formula (4.13) to compute its characteristic polynomial. Rather, by definition (4.12) and the recipe for computing determinants of

3×3 matrices (see Chapter 3 of Part II), the characteristic polynomial of $\mathbf{A}$ is

$$p_{\mathbf{A}}(z) = \det(\mathbf{I}z - \mathbf{A}) = \det\begin{pmatrix} z & -2 & 1 \\ 2 & z & -2 \\ -1 & 2 & z \end{pmatrix} = z^3 + 4 - 4 + 4z + 4z + z$$
$$= z^3 + 9z = z(z^2 + 9)\,.$$

It has the simple roots 0, $\pm i3$, so we will use it.

Below we will show by using recipe (4.18c) and by solving the general initial-value problem (4.17) that the natural fundamental set associated with $p(\mathrm{D})$ is

$$N_0(t) = 1\,, \qquad N_1(t) = \frac{\sin(3t)}{3}\,, \qquad N_2(t) = \frac{1 - \cos(3t)}{9}\,.$$

Given this natural fundamental set, formula (4.16) yields

$$\begin{aligned} e^{t\mathbf{A}} &= N_0(t)\mathbf{I} + N_1(t)\mathbf{A} + N_2(t)\mathbf{A}^2 \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \frac{\sin(3t)}{3}\begin{pmatrix} 0 & 2 & -1 \\ -2 & 0 & 2 \\ 1 & -2 & 0 \end{pmatrix} + \frac{1 - \cos(3t)}{9}\begin{pmatrix} 0 & 2 & -1 \\ -2 & 0 & 2 \\ 1 & -2 & 0 \end{pmatrix}^2 \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \frac{\sin(3t)}{3}\begin{pmatrix} 0 & 2 & -1 \\ -2 & 0 & 2 \\ 1 & -2 & 0 \end{pmatrix} + \frac{1 - \cos(3t)}{9}\begin{pmatrix} -5 & 2 & 4 \\ 2 & -8 & 2 \\ 4 & 2 & -5 \end{pmatrix} \\ &= \begin{pmatrix} \dfrac{4 + 5\cos(3t)}{9} & \dfrac{2 - 2\cos(3t) + 6\sin(3t)}{9} & \dfrac{4 - 4\cos(3t) - 3\sin(3t)}{9} \\ \dfrac{2 - 2\cos(3t) - 6\sin(3t)}{9} & \dfrac{1 + 8\cos(3t)}{9} & \dfrac{2 - 2\cos(3t) + 6\sin(3t)}{9} \\ \dfrac{4 - 4\cos(3t) + 3\sin(3t)}{9} & \dfrac{2 - 2\cos(3t) + 6\sin(3t)}{9} & \dfrac{4 + 5\cos(3t)}{9} \end{pmatrix}. \end{aligned}$$

From the Green Function. We use the partial fraction identity

$$\frac{1}{p(s)} = \frac{1}{s^3 + 9s} = \frac{1}{s(s^2 + 9)} = \frac{\frac{1}{9}}{s} + \frac{-\frac{1}{9}s}{s^2 + 9}\,.$$

A Laplace transform table shows that the Green function $g(t)$ is given by

$$\begin{aligned} g(t) = \mathcal{L}^{-1}\left[\frac{1}{p(s)}\right](t) &= \mathcal{L}^{-1}\left[\frac{1}{s^3 + 9s}\right](t) \\ &= \tfrac{1}{9}\mathcal{L}^{-1}\left[\frac{1}{s}\right](t) - \tfrac{1}{9}\mathcal{L}^{-1}\left[\frac{s}{s^2 + 3^2}\right](t) = \frac{1 - \cos(3t)}{9}\,. \end{aligned}$$

Then because we see $p(s) = s^3 + 9s$ as

$$p(s) = s^3 + 0s^2 + 9s + 0\,,$$

recipe (4.18c) yields the natural fundamental set

$$
\begin{aligned}
N_2(t) &= g(t) = \frac{1-\cos(3t)}{9}\,,\\
N_1(t) &= N_2'(t) + 0g(t) = \frac{\sin(3t)}{3}\,,\\
N_0(t) &= N_1'(t) + 9g(t) = \cos(3t) + 9\frac{1-\cos(3t)}{9} = 1\,.
\end{aligned}
$$

From the General Initial-Value Problem. The general initial-value problem (4.17) is

$$
y''' + 9y' = 0\,, \qquad y(0) = y_0\,, \quad y'(0) = y_1\,, \quad y''(0) = y_2\,.
$$

A general solution of the differential equation is $y(t) = c_1 + c_2\cos(3t) + c_3\sin(3t)$. Because

$$
y'(t) = -3c_2\sin(3t) + 3c_3\cos(3t)\,, \qquad y''(t) = -9c_2\cos(3t) - 9c_3\sin(3t)\,,
$$

the general initial conditions imply

$$
\begin{aligned}
y_0 &= y(0) = c_1 + c_2\cos(0) + c_3\sin(0) = c_1 + c_2\,,\\
y_1 &= y'(0) = -3c_2\sin(0) + 3c_3\cos(0) = 3c_3\,,\\
y_2 &= y''(0) = -9c_2\cos(0) - 9c_3\sin(0) = -9c_2\,,
\end{aligned}
$$

Upon solving this system for c_1, c_2, and c_3 we find that

$$
c_1 = \frac{9y_0 + y_2}{9}\,, \qquad c_2 = -\frac{y_2}{9}\,, \qquad c_3 = \frac{y_1}{3}\,.
$$

Therefore the solution of the general initial-value problem is

$$
\begin{aligned}
y(t) &= \frac{9y_0 + y_2}{9} - \frac{y_2}{9}\cos(3t) + \frac{y_1}{3}\sin(3t)\\
&= y_0 + \frac{\sin(3t)}{3}\,y_1 + \frac{1-\cos(3t)}{9}\,y_2\,.
\end{aligned}
$$

We then read off that the natural fundamental set is

$$
N_0(t) = 1\,, \qquad N_1(t) = \frac{\sin(3t)}{3}\,, \qquad N_2(t) = \frac{1-\cos(3t)}{9}\,.
$$

Remark. In each of the foregoing examples we used the characteristic polynomial for the natural fundamental set method. In the next example, we use a polynomial with a smaller degree than the characteristic polynomial.

Example. Compute $e^{t\mathbf{S}}$ for

$$
\mathbf{S} = \begin{pmatrix} 0 & 1 & 0 & 1\\ 1 & 0 & 1 & 0\\ 0 & 1 & 0 & 1\\ 1 & 0 & 1 & 0 \end{pmatrix}.
$$

Solution. Because $\mathbf{S}$ is 4×4, we cannot use formula (4.13) to compute its characteristic polynomial. We also cannot use the recipes for computing the determinant of a 3×3

matrix. Rather, by definition (4.12) and the Laplace expansion method for computing determinants (see Chapter 3 of Part II), the characteristic polynomial of $\mathbf{S}$ is

$$
\begin{aligned}
p_{\mathbf{S}}(z) = \det(z\,\mathbf{I} - \mathbf{S}) &= \det\begin{pmatrix} z & -1 & 0 & -1 \\ -1 & z & -1 & 0 \\ 0 & -1 & z & -1 \\ -1 & 0 & -1 & z \end{pmatrix} \\
&= z\det\begin{pmatrix} z & -1 & 0 \\ -1 & z & -1 \\ 0 & -1 & z \end{pmatrix} + \det\begin{pmatrix} -1 & 0 & -1 \\ -1 & z & -1 \\ 0 & -1 & z \end{pmatrix} + \det\begin{pmatrix} -1 & 0 & -1 \\ z & -1 & 0 \\ -1 & z & -1 \end{pmatrix} \\
&= z(z^3 - z - z) + (-z^2 - 1 + 1) + (-1 - z^2 + 1) = z^4 - 4z^2 \,.
\end{aligned}
$$

Because $p_{\mathbf{S}}(z) = z^2(z-2)(z+2)$, its roots are 0, 0, 2, and -2. Because 0 is a double root of $p_{\mathbf{S}}(z)$ and because $\mathbf{S}$ is symmetric, we see that $p(z) = z^3 - 4z$ also annihilates $\mathbf{S}$. This is the polynomial that we will use. It has the simple roots 0, 2, and -2.

Below we will show by using recipe (4.18c) and by solving the general initial-value problem (4.17) that the natural fundamental set associated with $p(\mathrm{D})$ is

$$
N_0(t) = 1\,, \qquad N_1(t) = \frac{e^{2t} - e^{-2t}}{4} = \frac{\sinh(2t)}{2}\,,
$$
$$
N_2(t) = \frac{e^{2t} + e^{-2t} - 2}{8} = \frac{\cosh(2t) - 1}{4}\,.
$$

Given this natural fundamental set, formula (4.16) yields

$$
\begin{aligned}
e^{t\mathbf{S}} &= N_0(t)\mathbf{I} + N_1(t)\mathbf{S} + N_2(t)\mathbf{S}^2 \\
&= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \frac{\sinh(2t)}{2}\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} + \frac{\cosh(2t) - 1}{4}\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}^2 \\
&= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \frac{\sinh(2t)}{2}\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} + \frac{\cosh(2t) - 1}{4}\begin{pmatrix} 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \\ 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} 1+\cosh(2t) & \sinh(2t) & \cosh(2t)-1 & \sinh(2t) \\ \sinh(2t) & 1+\cosh(2t) & \sinh(2t) & \cosh(2t)-1 \\ \cosh(2t)-1 & \sinh(2t) & 1+\cosh(2t) & \sinh(2t) \\ \sinh(2t) & \cosh(2t)-1 & \sinh(2t) & 1+\cosh(2t) \end{pmatrix}.
\end{aligned}
$$

From the Green Function. We use the partial fraction identity

$$
\frac{1}{p(s)} = \frac{1}{s^3 - 4s} = \frac{1}{s(s-2)(s+2)} = \frac{\frac{1}{8}}{s-2} + \frac{-\frac{1}{4}}{s} + \frac{\frac{1}{8}}{s+2}\,.
$$

A Laplace transform table shows that the Green function $g(t)$ is given by

$$\begin{aligned} g(t) = \mathcal{L}^{-1}\left[\frac{1}{p(s)}\right](t) &= \mathcal{L}^{-1}\left[\frac{1}{s^3 - 4s}\right](t) \\ &= \tfrac{1}{8}\mathcal{L}^{-1}\left[\frac{1}{s-2}\right](t) - \tfrac{1}{4}\mathcal{L}^{-1}\left[\frac{1}{s}\right](t) + \tfrac{1}{8}\mathcal{L}^{-1}\left[\frac{1}{s+2}\right](t) \\ &= \frac{e^{2t} - 2 + e^{-2t}}{8} = \frac{\cosh(2t) - 1}{4}\,. \end{aligned}$$

Then because we see $p(s) = s^3 - 4s$ as

$$p(s) = s^3 + 0s^2 - 4s + 0\,,$$

recipe (4.18c) yields the natural fundamental set

$$\begin{aligned} N_2(t) &= g(t) = \frac{\cosh(2t) - 1}{4}\,, \\ N_1(t) &= N_2'(t) + 0g(t) = \frac{\sinh(2t)}{2}\,, \\ N_0(t) &= N_1'(t) - 4g(t) = \cosh(2t) - 4\frac{\cosh(2t) - 1}{4} = 1\,. \end{aligned}$$

From the General Initial-Value Problem. The general initial-value problem (4.17) is

$$y''' - 4y' = 0\,, \qquad y(0) = y_0\,, \quad y'(0) = y_1\,, \quad y''(0) = y_2\,.$$

A general solution of the differential equation is $y(t) = c_1 + c_2e^{2t} + c_3e^{-2t}$. Because

$$y'(t) = 2c_2e^{2t} - 2c_3e^{-2t}\,, \qquad y''(t) = 4c_2e^{2t} + 4c_3e^{-2t}\,,$$

the general initial conditions imply that

$$\begin{aligned} y_0 &= y(0) = c_1 + c_2e^0 + c_3e^0 = c_1 + c_2 + c_3\,, \\ y_1 &= y'(0) = 2c_2e^0 - 2c_3e^0 = 2c_2 - 2c_3\,, \\ y_2 &= y''(0) = 4c_2e^0 + 4c_3e^0 = 4c_2 + 4c_3\,, \end{aligned}$$

Upon solving this system for c_1, c_2, and c_3 we find that

$$c_1 = \frac{4y_0 - y_2}{4}\,, \qquad c_2 = \frac{2y_1 + y_2}{8}\,, \qquad c_3 = \frac{-2y_1 + y_2}{8}\,.$$

Therefore the solution of the general initial-value problem is

$$\begin{aligned} y(t) &= \frac{4y_0 - y_2}{4} + \frac{2y_1 + y_2}{8}\,e^{2t} + \frac{-2y_1 + y_2}{8}\,e^{-2t} \\ &= y_0 + \frac{e^{2t} - e^{-2t}}{4}\,y_1 + \frac{e^{2t} + e^{-2t} - 2}{8}\,y_2\,. \end{aligned}$$

We then read off that the natural fundamental set is

$$N_0(t) = 1\,, \qquad N_1(t) = \frac{e^{2t} - e^{-2t}}{4} = \frac{\sinh(2t)}{2}\,,$$

$$N_2(t) = \frac{e^{2t} + e^{-2t} - 2}{8} = \frac{\cosh(2t) - 1}{4}\,.$$

Remark. We saved a lot of work by using $p(z) = z^3 - 4z$ as the annihilating polynomial rather than $p_{\mathbf{S}}(z) = z^4 - 4z^2$ because $p(z)$ has a smaller degree. We had a smaller natural fundamental set of solutions to compute. We did not have to compute $\mathbf{S}^3$. We had fewer terms in formula (4.16) to add up.

4.4. **Two-by-Two Matrix Exponential Formulas.** We will now apply the natural fundamental set method to derive simple formulas for the matrix exponential of any real matrix $\mathbf{A}$ that is annihilated by a quadratic polynomial $p(z) = z^2 + p_1 z + p_2$. This includes every 2×2 real matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

because the Cayley-Hamiltion Theorem says that $\mathbf{A}$ is annihilated by its characteristic polynomial,

$$p(z) = \det(\mathbf{I}z - \mathbf{A}) = z^2 - \mathrm{tr}(\mathbf{A})z + \det(\mathbf{A}).$$

It also includes every $n\times n$ matrix for $n > 2$ such that $p(\mathbf{A}) = \mathbf{0}$ for some quadratic polynomial. Such $n\times n$ matrices are special, but sometimes they arise in applications.

If $\mathbf{A}$ is a real matrix that is annihilated the polynomial $p(z) = z^2 + p_1 z + p_2$ then formula (4.16) reduces to

$$e^{t\mathbf{A}} = N_0(t)\mathbf{I} + N_1(t)\mathbf{A}, \tag{4.19}$$

where $N_0(t)$, $N_1(t)$ is the natural fundamental set of solutions for $t = 0$ of $p(\mathrm{D})y = 0$. The idea of this section is to exploit the fact that the second-order operator $p(\mathrm{D})$ is simple enough to precompute $N_0(t)$ and $N_1(t)$ in general, which when plugged into (4.19) will yield formulas for $e^{t\mathbf{A}}$.

The nature of these matrix exponential formulas will depend on the nature of the roots of $p(z)$. Upon *completing the square* we see that

$$p(z) = z^2 + p_1 z + p_2 = (z-\mu)^2 - \delta,$$

where the mean μ and discriminant δ are given by

$$\mu = -\frac{p_1}{2}, \qquad \delta = \mu^2 - p_2.$$

There are three cases which are distinguished by the sign of δ.

- If $\delta > 0$ then $p(z)$ can be written as the *difference of squares*

  $$p(z) = (z-\mu)^2 - \nu^2,$$

  where $\nu = \sqrt{\delta} > 0$. Then $p(z)$ has the simple real roots $\mu - \nu$ and $\mu + \nu$. The natural fundamental set of solutions of $p(\mathrm{D})$ can be found to be

$$N_0(t) = e^{\mu t}\cosh(\nu t) - \mu e^{\mu t}\frac{\sinh(\nu t)}{\nu}, \qquad N_1(t) = e^{\mu t}\frac{\sinh(\nu t)}{\nu}. \tag{4.20a}$$

- If $\delta < 0$ then $p(z)$ can be written as the *sum of squares*

$$p(z) = (z - \mu)^2 + \nu^2 \,,$$

where $\nu = \sqrt{|\delta|}$. Then $p(z)$ has the complex conjugate roots $\mu - i\nu$ and $\mu + i\nu$. The natural fundamental set of solutions of $p(\mathrm{D})$ can be found to be

$$N_0(t) = e^{\mu t}\cos(\nu t) - \mu e^{\mu t}\frac{\sin(\nu t)}{\nu}\,, \qquad N_1(t) = e^{\mu t}\frac{\sin(\nu t)}{\nu}\,. \tag{4.20b}$$

- If $\delta = 0$ then can be written as the *perfect square*

$$p(z) = (z - \mu)^2 \,.$$

Then $p(z)$ has the double real root μ. The natural fundamental set of solutions of $p(\mathrm{D})$ can be found to be

$$N_0(t) = e^{\mu t} - \mu e^{\mu t} t\,, \qquad N_1(t) = e^{\mu t} t\,. \tag{4.20c}$$

Remark. The details of finding $N_0(t)$ and $N_1(t)$ given by (4.20) are left as exercises.

Remark. Notice that (4.20c) is the limiting case of both (4.20a) and (4.20b) as $\nu \to 0$ because

$$\lim_{\nu\to 0}\cosh(\nu t) = \lim_{\nu\to 0}\cos(\nu t) = 1\,, \qquad \lim_{\nu\to 0}\frac{\sinh(\nu t)}{\nu} = \lim_{\nu\to 0}\frac{\sin(\nu t)}{\nu} = t\,.$$

When the natural fundamental sets (4.20a), (4.20b), and (4.20c) respectively are plugged into formula (4.19) we obtain formulas for $e^{t\mathbf{A}}$. If $p(z)$ is a *difference of squares* then we plug the natural fundamental set (4.20a) into formula (4.19) to obtain

$$\begin{aligned} e^{t\mathbf{A}} = N_0(t)\mathbf{I} + N_1(t)\mathbf{A} &= \left(e^{\mu t}\cosh(\nu t) - \mu e^{\mu t}\frac{\sinh(\nu t)}{\nu}\right)\mathbf{I} + e^{\mu t}\frac{\sinh(\nu t)}{\nu}\,\mathbf{A} \\ &= e^{\mu t}\left[\cosh(\nu t)\mathbf{I} + \frac{\sinh(\nu t)}{\nu}(\mathbf{A} - \mu\mathbf{I})\right]\,. \end{aligned}$$

If $p(z)$ is a *sum of squares* then we plug the natural fundamental set (4.20b) into formula (4.19) to obtain

$$\begin{aligned} e^{t\mathbf{A}} = N_0(t)\mathbf{I} + N_1(t)\mathbf{A} &= \left(e^{\mu t}\cos(\nu t) - \mu e^{\mu t}\frac{\sin(\nu t)}{\nu}\right)\mathbf{I} + e^{\mu t}\frac{\sin(\nu t)}{\nu}\,\mathbf{A} \\ &= e^{\mu t}\left[\cos(\nu t)\mathbf{I} + \frac{\sin(\nu t)}{\nu}(\mathbf{A} - \mu\mathbf{I})\right]\,. \end{aligned}$$

If $p(z)$ is a *perfect square* then we plug the natural fundamental set (4.20c) into formula (4.19) to obtain

$$\begin{aligned} e^{t\mathbf{A}} = N_0(t)\mathbf{I} + N_1(t)\mathbf{A} &= \left(e^{\mu t} - \mu e^{\mu t} t\right)\mathbf{I} + e^{\mu t} t\,\mathbf{A} \\ &= e^{\mu t}\left[\mathbf{I} + t\,(\mathbf{A} - \mu\mathbf{I})\right]\,. \end{aligned}$$

We summarize these results below.

- If $p(z)$ is the *difference of squares*

$$p(z) = (z-\mu)^2 - \nu^2 \quad \text{for some } \nu > 0\,,$$

then

$$e^{t\mathbf{A}} = e^{\mu t}\left[\cosh(\nu t)\mathbf{I} + \frac{\sinh(\nu t)}{\nu}\,(\mathbf{A}-\mu\mathbf{I})\right]. \tag{4.21a}$$

- If $p(z)$ is the *sum of squares*

$$p(z) = (z-\mu)^2 + \nu^2 \quad \text{for some } \nu > 0\,,$$

then

$$e^{t\mathbf{A}} = e^{\mu t}\left[\cos(\nu t)\mathbf{I} + \frac{\sin(\nu t)}{\nu}\,(\mathbf{A}-\mu\mathbf{I})\right]. \tag{4.21b}$$

- If $p(z)$ is the *perfect square*

$$p(z) = (z-\mu)^2\,,$$

then

$$e^{t\mathbf{A}} = e^{\mu t}\big[\mathbf{I} + t\,(\mathbf{A}-\mu\mathbf{I})\big]\,. \tag{4.21c}$$

For almost any matrix that is annihilated by a quadratic polynomial the fastest way to compute its matrix exponential is to apply these formulas. Notice that formulas (4.21a) and (4.21b) are similar — the first uses hyperbolic functions when $p(z)$ has simple real roots, while the second uses trigonometric functions when $p(z)$ has complex conjugate roots. Formula (4.21c) is the limiting case of both (4.21a) and (4.21b) as $\nu \to 0$ because

$$\lim_{\nu\to 0}\cosh(\nu t) = \lim_{\nu\to 0}\cos(\nu t) = 1\,, \qquad \lim_{\nu\to 0}\frac{\sinh(\nu t)}{\nu} = \lim_{\nu\to 0}\frac{\sin(\nu t)}{\nu} = t\,.$$

These relationships make these formulas easier to remember.

Remark. For 2×2 matrices $\mu = \frac{1}{2}\mathrm{tr}(\mathbf{A})$ while $\mathrm{tr}(\mathbf{I}) = 2$. In that case the matrix $\mathbf{A}-\mu\mathbf{I}$ that appears in each of the formulas (4.21) has trace zero. This fact should be used as a check whenever these formulas are used to compute $e^{t\mathbf{A}}$ for 2×2 matrices!

We now illustrate this method on the same 2×2 matrices used in the previous section.

Example. Compute $e^{t\mathbf{A}}$ for

$$\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}.$$

Solution. Because $\mathbf{A}$ is 2×2, by (4.13) its characteristic polynomial is

$$p(z) = z^2 - \mathrm{tr}(\mathbf{A})z + \det(\mathbf{A}) = z^2 - 6z + 5 = (z-3)^2 - 2^2\,.$$

This is a difference of squares with $\mu = 3$ and $\nu = 2$. So by formula (4.21a) we obtain

$$\begin{aligned} e^{t\mathbf{A}} &= e^{3t}\left[\cosh(2t)\mathbf{I} + \frac{\sinh(2t)}{2}(\mathbf{A}-3\mathbf{I})\right] \\ &= e^{3t}\left[\cosh(2t)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\sinh(2t)}{2}\begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}\right] = e^{3t}\begin{pmatrix} \cosh(2t) & \sinh(2t) \\ \sinh(2t) & \cosh(2t) \end{pmatrix}. \end{aligned}$$

Example. Compute $e^{t\mathbf{B}}$ for

$$\mathbf{B} = \begin{pmatrix} -3 & -1 \\ 4 & 1 \end{pmatrix} .$$

Solution. Because $\mathbf{B}$ is 2×2, by (4.13) its characteristic polynomial is

$$p(z) = z^2 - \mathrm{tr}(\mathbf{B})z + \det(\mathbf{B}) = z^2 + 2z + 1 = (z+1)^2 .$$

This is a perfect square with $\mu = -1$. So by formula (4.21c) we obtain

$$\begin{aligned} e^{t\mathbf{B}} &= e^{-t}\big[\mathbf{I} + t\,(\mathbf{B}+\mathbf{I})\big] \\ &= e^{-t}\left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + t\begin{pmatrix} -2 & -1 \\ 4 & 2 \end{pmatrix}\right] = e^{-t}\begin{pmatrix} 1-2t & -t \\ 4t & 1+2t \end{pmatrix} . \end{aligned}$$

Example. Compute $e^{t\mathbf{C}}$ for

$$\mathbf{C} = \begin{pmatrix} 6 & -5 \\ 5 & -2 \end{pmatrix} .$$

Solution. Because $\mathbf{C}$ is 2×2, by (4.13) its characteristic polynomial is

$$p(z) = z^2 - \mathrm{tr}(\mathbf{C})z + \det(\mathbf{C}) = z^2 - 4z + 13 = (z-2)^2 + 3^2 .$$

This is a sum of squares with $\mu = 2$ and $\nu = 3$. So by formula (4.21b) we obtain

$$\begin{aligned} e^{t\mathbf{C}} &= e^{2t}\left[\cos(3t)\mathbf{I} + \frac{\sin(3t)}{3}(\mathbf{C} - 2\mathbf{I})\right] \\ &= e^{2t}\left[\cos(3t)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\sin(3t)}{3}\begin{pmatrix} 4 & -5 \\ 5 & -4 \end{pmatrix}\right] \\ &= e^{2t}\begin{pmatrix} \cos(3t) + \frac{4}{3}\sin(3t) & -\frac{5}{3}\sin(3t) \\ \frac{5}{3}\sin(3t) & \cos(3t) - \frac{4}{3}\sin(3t) \end{pmatrix} . \end{aligned}$$

4.5. **Natural Fundamental Sets from Green Functions.** If $\mathbf{A}$ is an $n\times n$ matrix that is annihilated by a polynomial of degree $m > 2$ and n is not too large then the hardest step in the natural fundamental set method for computing $e^{t\mathbf{A}}$ is often generating the natural fundamental set $N_0(t)$, $N_1(t)$, $\cdots$, $N_{m-1}(t)$. One way to do this is to solve the general initial-value problem (4.17). Here we present the alternative way to do this given by (4.18).

4.5.1. *The Recipe.* Observe that in the two-by-two case each of the natural fundamental sets of solutions given by (4.18) the solutions $N_0(t)$ and $N_1(t)$ are related by

$$N_0(t) = N_1'(t) - 2\mu N_1(t) .$$

This is an instance of a more general fact. For the m^{th}-order equation

$$p(\mathrm{D})y = 0 , \qquad \text{where} \quad p(z) = z^m + p_1 z^{m-1} + \cdots + p_{m-1}z + p_m , \tag{4.22}$$

we can generate its entire natural fundamental set of solutions from the *Green function* $g(t)$ associated with (4.22). Recall that the Green function $g(t)$ satisfies the initial-value problem

$$p(\mathrm{D})g = 0 , \qquad g(0) = g'(0) = \cdots = g^{(m-2)}(0) = 0 , \quad g^{(m-1)}(0) = 1 . \tag{4.23}$$

The natural fundamental set of solutions is then given by the recipe

$$\begin{aligned} N_{m-1}(t) &= g(t)\,, \\ N_{m-2}(t) &= N_{m-1}'(t) + p_1 g(t)\,, \\ &\ \ \vdots \\ N_0(t) &= N_1'(t) + p_{m-1} g(t)\,. \end{aligned} \tag{4.24}$$

This entire set thereby is generated by the solution g of the initial-value problem (4.23).

Remark. It is clear that solving the initial-value problem (4.23) for the Green function $g(t)$ requires less algebra than solving the general initial-value problem (4.19) directly for $N_0(t)$, $N_1(t)$, $\cdots$, $N_{m-1}(t)$. The trade-off is that we now have to compute the $m-1$ derivatives required by recipe (4.24).

Example. Compute the natural fundamental set of solutions to the equation

$$y''' + 9y' = 0\,.$$

Solution. By (4.23) the Green function $g(t)$ satisfies the initial-value problem

$$g''' + 9g' = 0\,, \qquad g(0) = g'(0) = 0\,, \quad g''(0) = 1\,.$$

The characteristic polynomial of this equation is $p(z) = z^3 + 9z$, which has roots 0, $\pm i3$. Therefore we seek a solution in the form

$$g(t) = c_1 + c_2 \cos(3t) + c_3 \sin(3t)\,.$$

Because

$$g'(t) = -3c_2 \sin(3t) + 3c_3 \cos(3t)\,, \qquad g''(t) = -9c_2 \cos(3t) - 9c_3 \sin(3t)\,,$$

the initial conditions for $g(t)$ then yield the algebraic system

$$g(0) = c_1 + c_2 = 0\,, \qquad g'(0) = 3c_3 = 0\,, \qquad g''(0) = -9c_2 = 1\,.$$

Solving this system gives $c_1 = \frac{1}{9}$, $c_2 = -\frac{1}{9}$, and $c_3 = 0$, whereby the Green function is

$$g(t) = \frac{1 - \cos(3t)}{9}\,.$$

Because $p(z) = z^3 + 9z$, we read off from (4.22) that $p_1 = 0$, $p_2 = 9$, and $p_3 = 0$. Then by recipe (4.24) the natural fundamental set of solutions is given by

$$\begin{aligned} N_2(t) &= g(t) = \frac{1 - \cos(3t)}{9}\,, \\ N_1(t) &= N_2'(t) + 0 \cdot g(t) = \frac{\sin(3t)}{3}\,, \\ N_0(t) &= N_1'(t) + 9 \cdot g(t) = \cos(3t) + 9 \cdot \frac{1 - \cos(3t)}{9} = 1\,. \end{aligned}$$

Remark. In Subsection 4.3.3 we computed this natural fundamental set by solving the general initial-value problem (4.17). That calculation should be compared with the one above.

Example. Compute the natural fundamental set of solutions to the equation

$$y''' - 4y' = 0\,.$$

Solution. By (4.23) the Green function $g(t)$ satisfies the initial-value problem

$$g''' - 4g' = 0\,, \qquad g(0) = g'(0) = 0\,, \quad g''(0) = 1\,.$$

The characteristic polynomial of this equation is $p(z) = z^3 - 4z$, which has roots 0, ± 2. Therefore we seek a solution in the form

$$g(t) = c_1 + c_2 e^{2t} + c_3 e^{-2t}\,.$$

Because

$$g'(t) = 2c_2 e^{2t} - 2c_3 e^{-2t}\,, \qquad g''(t) = 4c_2 e^{2t} + 4c_3 e^{-2t}\,,$$

the initial conditions for $g(t)$ then yield the algebraic system

$$g(0) = c_1 + c_2 + c_3 = 0\,, \qquad g'(0) = 2c_2 - 2c_3 = 0\,, \qquad g''(0) = 4c_2 + 4c_3 = 1\,.$$

The solution of this system is $c_1 = -\frac{1}{4}$ and $c_2 = c_3 = \frac{1}{8}$, whereby the Green function is

$$g(t) = -\tfrac{1}{4} + \tfrac{1}{8}e^{2t} + \tfrac{1}{8}e^{-2t} = \tfrac{1}{4}\big(\cosh(2t) - 1\big)\,.$$

Because $p(z) = z^3 - 4z$, we read off from (4.22) that $p_1 = 0$, $p_2 = -4$, and $p_3 = 0$. Then by recipe (4.24) the natural fundamental set of solutions is given by

$$\begin{aligned}
N_2(t) &= g(t) = \frac{\cosh(2t) - 1}{4}\,,\\
N_1(t) &= N_2'(t) + 0 \cdot g(t) = \frac{\sinh(2t)}{2}\,,\\
N_0(t) &= N_1'(t) - 4 \cdot g(t) = \cosh(2t) - 4 \cdot \frac{\cosh(2t) - 1}{4} = 1\,.
\end{aligned}$$

Remark. In Subsection 4.3.3 we computed this natural fundamental set by solving the general initial-value problem (4.15). That calculation should be compared with the one above.

4.5.2. *Its Justification.* The following justification of recipe (4.24) is included for completeness. It does not need to be known in order to use the recipe, so it may be skipped on first reading. However, the recipe itself can be quite useful.

Suppose that the Green function $g(t)$ satisfies the initial-value problem (4.23) while the functions $N_0(t)$, $N_1(t)$, $\cdots$, $N_{m-1}(t)$ are defined by recipe (4.24). We want to show that for each $j = 0, 1, \cdots, m-1$ the function $N_j(t)$ is the solution of

(4.25) $$p(\mathrm{D})y = y^{(m)} + p_1 y^{(m-1)} + p_2 y^{(m-2)} + \cdots + p_{m-1}y' + p_m y = 0\,,$$

that satisfies the initial conditions

(4.26) $$N_j^{(k)}(0) = \delta_{jk} \quad \text{for } k = 0, 1, \cdots, m-1\,.$$

Here δ_{jk} is the Kronecker delta, which is defined by

$$\delta_{jk} = \begin{cases} 1 & \text{when } j = k\,,\\ 0 & \text{when } j \neq k\,. \end{cases}$$

Because $N_{m-1}(t) = g(t)$, we see from (4.23) that it satisfies the differential equation (4.25) and the initial conditions (4.26) for $j = m - 1$. The key step is to show that if $N_j(t)$ satisfies (4.25) and (4.26) for some positive $j < m$ then so does $N_{j-1}(t)$. Once this step is done then we can argue that because $N_{m-1}(t)$ has these properties, so does $N_{m-2}(t)$, which implies so does $N_{m-3}(t)$, and so on down to $N_0(t)$.

We now prove the key step. Suppose that $N_j(t)$ satisfies (4.25) and (4.26) for some positive $j < m$. Because $N_j(t)$ satisfies (4.25), so does $N_j'(t)$. Because $N_{j-1}(t)$ is given by recipe (4.24) as a linear combination of $N_j'(t)$ and $g(t)$, while $N_j'(t)$ and $g(t)$ satisfy (4.25), we see that $N_{j-1}(t)$ also satisfies the differential equation (4.25). The only thing that remains to be checked is that $N_{j-1}(t)$ satifies the initial conditions (4.26).

Because $N_j(t)$ satisfies (4.25) and (4.26), we see that

$$\begin{aligned} 0 = p\big(\mathrm{D}\big)N_j(t)\big|_{t=0} &= N_j^{(m)}(0) + \sum_{k=0}^{m-1} p_{m-k} N_j^{(k)}(0) \\ &= N_j^{(m)}(0) + \sum_{k=0}^{m-1} p_{m-k}\delta_{jk} = N_j^{(m)}(0) + p_{m-j}\,. \end{aligned} \tag{4.27}$$

Because $N_{j-1}(t)$ is given by recipe (4.24) to be $N_{j-1}(t) = N_j'(t) + p_{m-j}g(t)$, we evaluate the $(k-1)^{\text{st}}$ derivative of this relation at $t = 0$ to obtain

$$N_{j-1}^{(k-1)}(0) = N_j^{(k)}(0) + p_{m-j} g^{(k-1)}(0) \quad \text{for } k = 1,\, 2,\, \cdots,\, m\,.$$

Because $g(t)$ satisfies the initial condtions in (4.23), we see from (4.26) that this becomes

$$N_{j-1}^{(k-1)}(0) = \delta_{jk} \quad \text{for } k = 1,\, 2,\, \cdots,\, m-1\,,$$

while we see from (4.27) that for $k = m$ it becomes

$$N_{j-1}^{(m-1)}(0) = N_j^{(m)}(0) + p_{m-j} = 0\,.$$

The initial conditions (4.26) thereby hold for $N_{j-1}(t)$. This completes the proof of the key step, which completes the justification of recipe (4.24). ☐

4.6. **General Matrix Exponential Formula.** We now present a formula for the exponential of any $n\times n$ matrix $\mathbf{A}$. This formula will be the basis for a method presented in Section 4.7 that computes $e^{t\mathbf{A}}$ efficiently when n is not too large. It is also related to the method of generalized eigenvectors that appears in the next chapter and can be used to compute $e^{t\mathbf{A}}$ efficiently when n is large.

4.6.1. *The Formula.* Let be $p(z)$ be any monic polynomial of degree $m \leq n$ that annihilates $\mathbf{A}$. The formula for $e^{t\mathbf{A}}$ will be given explicitly in terms of $\mathbf{A}$ and the roots of $p(z)$. Suppose that $p(z)$ has l roots λ_1, λ_2, $\cdots$, λ_l (possibly complex) with multiplicities m_1, m_2, $\cdots$, m_l respectively. This means that $p(z)$ has the factored form

$$p(z) = \prod_{k=1}^{l} (z - \lambda_k)^{m_k}\,, \tag{4.28}$$

and that $m_1 + m_2 + \cdots + m_l = m$. Here we understand that $\lambda_j \neq \lambda_k$ if $j \neq k$. Then

$$e^{t\mathbf{A}} = \sum_{k=1}^{l} e^{\lambda_k t} \left(\mathbf{I} + t\,(\mathbf{A} - \lambda_k \mathbf{I}) + \cdots + \frac{t^{m_k - 1}}{(m_k - 1)!}\,(\mathbf{A} - \lambda_k \mathbf{I})^{m_k - 1} \right) \mathbf{Q}_k \,, \tag{4.29a}$$

where the $n{\times}n$ matrices $\mathbf{Q}_k$ are defined by

$$\mathbf{Q}_k = \prod_{\substack{j=1 \\ j \neq k}}^{l} \left(\frac{\mathbf{A} - \lambda_j \mathbf{I}}{\lambda_k - \lambda_j} \right)^{m_j} \qquad \text{for every } k = 1,\, 2,\, \cdots,\, l\,. \tag{4.29b}$$

Before we justify formula (4.29), let us consider its structure. It can be recast as $e^{t\mathbf{A}} = h(\mathbf{A})$ where $h(z)$ is the time-dependent polynomial

$$h(z) = \sum_{k=1}^{l} e^{\lambda_k t} \left(1 + t\,(z - \lambda_k) + \cdots + \frac{t^{m_k - 1}}{(m_k - 1)!}\,(z - \lambda_k)^{m_k - 1} \right) q_k(z)\,, \tag{4.30a}$$

with the time-independent polynomials $q_k(z)$ defined by

$$q_k(z) = \prod_{\substack{j=1 \\ j \neq k}}^{l} \left(\frac{z - \lambda_j}{\lambda_k - \lambda_j} \right)^{m_j} \qquad \text{for every } k = 1,\, 2,\, \cdots,\, l\,. \tag{4.30b}$$

Because the degree of each $q_k(z)$ is $m - m_k$, it follows from (4.30a) that the degree of $h(z)$ is at most $m - 1$.

It can be checked that for every $k = 1, 2, \cdots, l$ the polynomial $h(z)$ satisfies the m_k relations

$$h(\lambda_k) = e^{\lambda_k t}\,, \quad h'(\lambda_k) = t\, e^{\lambda_k t}\,, \quad \cdots\,, \quad h^{(m_k - 1)}(\lambda_k) = t^{m_k - 1} e^{\lambda_k t}\,. \tag{4.31}$$

These relations state that the functions $h(z)$ and e^{zt} and their derivatives with respect to z though order $m_k - 1$ agree at $z = \lambda_k$. Because these m_k relations hold for each k, this gives a total of m relations that $h(z)$ satisfies. The theory of *Hermite interpolation* states that these m relations uniquely specify the polynomial $h(z)$ at every time t, which is the called associated *Hermite interpolant* of e^{zt}.

4.6.2. *Its Justification.* We now justify formula (4.29). Let $\mathbf{\Phi}(t)$ denote the right-hand side of (4.29a). We will show that $\mathbf{\Phi}(t)$ satisfies the matrix-valued initial-value problem (4.2), and thereby conclude that $\mathbf{\Phi}(t) = e^{t\mathbf{A}}$ as asserted by formula (4.29).

We first show that $\mathbf{\Phi}(0) = \mathbf{I}$. By setting $t = 0$ in the right-hand side of (4.29a) we get $\mathbf{\Phi}(0) = h(\mathbf{A})$ where the polynomial $h(z)$ is obtained by setting $t = 0$ in formula (4.30a) — namely,

$$h(z) = \sum_{k=1}^{l} q_k(z)\,,$$

where the polynomials $q_k(z)$ are defined by (4.30b). The polynomial $h(z)$ has degree less than m. By (4.31), for every $k = 1, 2, \cdots, l$ this polynomial satisfies the m_k conditions

$$h(\lambda_k) = 1\,, \quad h'(\lambda_k) = 0\,, \quad \cdots\,, \quad h^{(m_k - 1)}(\lambda_k) = 0\,.$$

Because the constant polynomial 1 satisfies these conditions and has degree less than m, we conclude by the uniqueness of the Hermite interpolant that $h(z) = 1$. Therefore $\mathbf{\Phi}(0) = h(\mathbf{A}) = \mathbf{I}$, which is the initial condition of (4.2).

Next we show that $\mathbf{\Phi}(t)$ also satisfies the differential system of (4.2). Because $p(z)$ annihilates $\mathbf{A}$ and has the factored form (4.29), it follows that

$$(\mathbf{A} - \lambda_k \mathbf{I})^{m_k} \mathbf{Q}_k = \mathbf{0} \qquad \text{for every } k = 1,\, 2,\, \cdots,\, l\,.$$

By using this fact we see that

$$\begin{aligned}
\frac{\mathrm{d}\mathbf{\Phi}}{\mathrm{d}t}(t) &= \sum_{k=1}^{l} \lambda_k e^{\lambda_k t} \left(\mathbf{I} + t\,(\mathbf{A} - \lambda_k \mathbf{I}) + \cdots + \frac{t^{m_k-1}}{(m_k-1)!}\,(\mathbf{A} - \lambda_k \mathbf{I})^{m_k-1} \right) \mathbf{Q}_k \\
&\quad + \sum_{k=1}^{l} e^{\lambda_k t} \left((\mathbf{A} - \lambda_k \mathbf{I}) + \cdots + \frac{t^{m_k-2}}{(m_k-2)!}\,(\mathbf{A} - \lambda_k \mathbf{I})^{m_k-1} \right) \mathbf{Q}_k \\
&= \sum_{k=1}^{l} \lambda_k e^{\lambda_k t} \left(\mathbf{I} + t\,(\mathbf{A} - \lambda_k \mathbf{I}) + \cdots + \frac{t^{m_k-1}}{(m_k-1)!}\,(\mathbf{A} - \lambda_k \mathbf{I})^{m_k-1} \right) \mathbf{Q}_k \\
&\quad + \sum_{k=1}^{l} e^{\lambda_k t} (\mathbf{A} - \lambda_k \mathbf{I}) \left(\mathbf{I} + t\,(\mathbf{A} - \lambda_k \mathbf{I}) + \cdots + \frac{t^{m_k-1}}{(m_k-1)!}\,(\mathbf{A} - \lambda_k \mathbf{I})^{m_k-1} \right) \mathbf{Q}_k \\
&= \mathbf{A}\mathbf{\Phi}(t)\,.
\end{aligned}$$

Therefore $\mathbf{\Phi}(t)$ also satisfies the differential system of (4.2). Because $\mathbf{\Phi}(t)$ satisfies the matrix-valued initial-value problem (4.2), we conclude that $\mathbf{\Phi}(t) = e^{t\mathbf{A}}$, thereby proving formula (4.29).

Remark. Naively plugging $\mathbf{A}$ into formula (4.29) is an extremely inefficient way to evaluate $e^{t\mathbf{A}}$. This is because for each $k = 1, \cdots, l$ it requires $m_k - 1$ matrix multiplications to compute $(\mathbf{A} - \lambda_k \mathbf{I})^j$ for $j = 2, \cdots, m_k$, which sums to $m - l$ matrix multiplications, while formula (4.29b) requires $l - 2$ matrix multiplications to compute each $\mathbf{Q}_k$, which sums to $l^2 - 2l$. Naively evaluating formula (4.29) thereby requires at least $l^2 - 3l + m$ matrix multiplications. Because each matrix multiplication requires n^3 multiplications of numbers, naively evaluating formula (4.29) when $l = m = n$ therefore requires $n^5 - 2n^4$ multiplications of numbers. This is 81 multiplications for $n = 3$ and 512 multiplications for $n = 4$ — ouch!

4.7. **Hermite Interpolation Methods.** We can compute $e^{t\mathbf{A}}$ from formula (4.30) more efficiently by adapting methods for evaluating the Hermite interpolant $h(z)$ efficiently. Moreover, these methods can be applied to compute other functions of matrices like high powers.

4.7.1. *Computing Hermite Interpolants.* Given any function $f(z)$ of the complex variable z. An Hermite interpolant $h(z)$ of $f(z)$ is a polynomial of degree at most $m - 1$ that satisfies m conditions. These m conditions are encoded by a list of m points z_1, z_2, $\cdots$, z_m. These are listed with multiplicity and clustered so that

$$\text{if } z_j = z_k \text{ for some } j < k \text{ then } z_j = z_{j+1} = \cdots = z_{k-1} = z_k\,.$$

The interpolation conditions take the form for every $1 \le j \le k \le m$

$$h^{(k-j)}(z_k) = f^{(k-j)}(z_k) \quad \text{if } z_j = z_k \,. \tag{4.32}$$

These relations state that $h(z)$ and $f(z)$ agree at $z = z_k$.

The way to evaluate $h(z)$ efficiently is to first express it in the Newton form

$$h(z) = \sum_{k=0}^{m-1} h_{k+1}\, p_k(z)\,, \tag{4.33a}$$

where the polynomials $p_k(z)$ are defined by

$$p_0(z) = 1\,, \qquad p_k(z) = \prod_{j=1}^{k}(z - z_j) \quad \text{for } k = 1, \cdots, m-1\,. \tag{4.33b}$$

Notice that each $p_k(z)$ is a monic polynomial of degree k. The coefficients h_k are determined by completing a so-called divided-difference table of the form

$$\begin{array}{ccccccccccc}
z_1 & h[z_1] & & & & & & & & & \\
 & & \searrow & & & & & & & & \\
z_2 & h[z_2] & \rightarrow & h[z_1,z_2] & & & & & & & \\
 & & \searrow & & \searrow & & & & & & \\
z_3 & h[z_3] & \rightarrow & h[z_2,z_3] & \rightarrow & h[z_1,z_2,z_3] & & & & & \\
 & & \searrow & & \searrow & & \searrow & & & & \\
\vdots & \vdots & & \vdots & & \vdots & & \ddots & & & \\
 & & \searrow & & \searrow & & \searrow & & \searrow & & \\
z_m & h[z_m] & \rightarrow & h[z_{m-1},z_m] & \rightarrow & h[z_{m-2},z_{m-1},z_m] & \rightarrow & \cdots & \rightarrow & h[z_1,\cdots,z_m]\,, &
\end{array}$$

where the first column of the table is seeded with the entries $h[z_k] = f(z_k)$ and entries of subsequent columns are obtain for every $1 \le j < k \le m$ from the divided-difference formula

$$h[z_j,\cdots,z_k] = \begin{cases} \dfrac{h[z_{j+1},\cdots,z_k] - h[z_j,\cdots,z_{k-1}]}{z_k - z_j}\,, & \text{if } z_j \ne z_k\,; \\[2ex] \dfrac{1}{(k-j)!} f^{(k-j)}(z_k)\,, & \text{if } z_j = z_k\,. \end{cases} \tag{4.33c}$$

Notice that entry $h[z_j,\cdots,z_k]$ depends only on the entry to its left, and the entry above that one. This dependence is indicated by the arrows in the table. After the table is completed, the coefficients h_k in (4.33a) are read off from the top entries of each column as

$$h_k = h[z_1,\cdots,z_k] \qquad \text{for } k = 1, \cdots, m\,. \tag{4.33d}$$

4.7.2. *Application to Matrix Exponentials.* We can now evaluate formula (4.29) by applying recipe (4.33) to the function $f(z) = e^{tz}$. Let $\mathbf{A}$ be an $n{\times}n$ real matrix and $p(z)$ be a polynomial of degree $m \leq n$ that annihilates $\mathbf{A}$. Let $\{z_1, \cdots, z_m\}$ be the roots of $p(z)$ listed with multiplicity and clustered so that

$$\text{if } z_j = z_k \text{ for some } j < k \text{ then } z_j = z_{j+1} = \cdots = z_{k-1} = z_k\,.$$

Compute the divided-difference table whose first column is seeded with $h[z_k] = e^{tz_k}$ and subsequent columns are given by (4.33c). Read off the functions $h_1(t)$, $h_2(t)$, $\cdots$, $h_m(t)$ from the top entries of each column as prescribed by (4.33d). Then set

$$e^{t\mathbf{A}} = h(\mathbf{A}) = \sum_{k=0}^{m-1} h_{k+1}(t)\,\mathbf{P}_k\,, \tag{4.34a}$$

where $\mathbf{P}_k = p_k(\mathbf{A})$ and the polynomials $p_k(z)$ are defined by (4.33b). The matrices $\mathbf{P}_k$ can be computed efficiently by the recipe

$$\begin{aligned}
&\mathbf{P}_0 = \mathbf{I}\,,\\
&\mathbf{P}_1 = \mathbf{A} - z_1\mathbf{I}\,,\\
&\mathbf{P}_k = \mathbf{P}_{k-1}(\mathbf{A} - z_k\mathbf{I}) \quad \text{for } k = 2, \cdots, m-1\,.
\end{aligned} \tag{4.34b}$$

This approach requires only $m{-}2$ matrix multiplications to compute the $\mathbf{P}_k$. Evaluating formula (4.34) when $m = n$ therefore requires $n^4 - 2n^3$ multiplications of numbers. This is 25 for $n = 3$ and 128 for $n = 4$ — much better!

Remark. This method is closely related to the Putzer Method for computing $e^{t\mathbf{A}}$. That method also uses formula (4.34) but rather than computing $h_1(t)$, $h_2(t)$, $\cdots$, $h_m(t)$ by using a divided-difference table, it obtains them by solving the system of differential equations

$$\begin{aligned}
h_1' &= z_1 h_1 & h_1(0) &= 1\,,\\
h_k' &= z_k h_k + h_{k-1} & h_k(0) &= 0 \quad \text{for } k = 2, \cdots, m\,.
\end{aligned}$$

Of course, the resulting $h_0(t)$, $h_0(t)$, $\cdots$, $h_{m-1}(t)$ are the same, but the divided-difference table gets to them faster. Moreover, the divided-difference table easily extends to other functions of matrices, whereas the Putzer Method does not.

For matrices annihilated by a quadratic polynonial, formula (4.34) does not require any matrix multiplications and it recovers formulas (4.21) for matrix exponentials that we derived earlier. There are two cases to consider.

Fact. If $\mathbf{A}$ is annihilated by a quadratic polynonial with a double root λ_1 then

$$e^{t\mathbf{A}} = e^{\lambda_1 t}\mathbf{I} + t\,e^{\lambda_1 t}(\mathbf{A} - \lambda_1\mathbf{I})\,.$$

Reason. Set $z_1 = z_2 = \lambda_1$. Then by (4.34b)

$$\mathbf{P}_0 = \mathbf{I}\,, \qquad \mathbf{P}_1 = \mathbf{A} - \lambda_1\mathbf{I}\,,$$

while by (4.33c) the divided-difference table is simply

$$\begin{array}{cccc}
\lambda_1 & e^{\lambda_1 t} & & \\
 & & \searrow & \\
\lambda_1 & e^{\lambda_1 t} & \rightarrow & t\,e^{\lambda_1 t}
\end{array}$$

We read off from the top entries of the columns that

$$h_1(t) = e^{\lambda_1 t}\,, \qquad h_2(t) = t\, e^{\lambda_1 t}\,,$$

whereby the result follows from (4.34a).

Remark. The above fact recovers formula (4.21c) for the matrix exponential when the quadratic polynonial $p(z)$ has a double root λ_1.

Fact. If $\mathbf{A}$ is annihilated by a quadratic polynonial with distinct roots λ_1 and λ_2 then

$$e^{t\mathbf{A}} = e^{\lambda_1 t}\mathbf{I} + \frac{e^{\lambda_2 t} - e^{\lambda_1 t}}{\lambda_2 - \lambda_1}(\mathbf{A} - \lambda_1 \mathbf{I})\,.$$

Reason. Set $z_1 = \lambda_1$ and $z_2 = \lambda_2$. Then by (4.34b)

$$\mathbf{P}_0 = \mathbf{I}\,, \qquad \mathbf{P}_1 = \mathbf{A} - \lambda_1 \mathbf{I}\,,$$

while by (4.33c) the divided-difference table is simply

$$\begin{array}{ccccc} \lambda_1 & e^{\lambda_1 t} & & \\ & & \searrow & \\ \lambda_2 & e^{\lambda_2 t} & \rightarrow & \dfrac{e^{\lambda_2 t} - e^{\lambda_1 t}}{\lambda_2 - \lambda_1} \end{array}$$

We read off from the top entries of the columns that

$$h_1(t) = e^{\lambda_1 t}\,, \qquad h_2(t) = \frac{e^{\lambda_2 t} - e^{\lambda_1 t}}{\lambda_2 - \lambda_1}\,,$$

whereby the result follows from (4.34a).

Remark. The above fact recovers formulas (4.21a) and (4.21b) for the matrix exponential when the quadratic polynonial $p(z)$ has distinct roots λ_1 and λ_2. For example, if λ_1 and λ_2 are real with $\lambda_1 = \mu - \nu$ and $\lambda_2 = \mu + \nu$ then

$$\begin{aligned} e^{t\mathbf{A}} &= e^{(\mu-\nu)t}\mathbf{I} + \frac{e^{(\mu+\nu)t} - e^{(\mu-\nu)t}}{2\nu}\,(\mathbf{A} - (\mu-\nu)\mathbf{I}) \\ &= \frac{e^{(\mu+\nu)t} + e^{(\mu-\nu)t}}{2}\,\mathbf{I} + \frac{e^{(\mu+\nu)t} - e^{(\mu-\nu)t}}{2\nu}\,(\mathbf{A} - \mu\mathbf{I}) \\ &= e^{\mu t}\left[\cosh(\nu t)\,\mathbf{I} + \frac{\sinh(\nu t)}{\nu}\,(\mathbf{A} - \mu\mathbf{I})\right]\,. \end{aligned}$$

Similarly, if λ_1 and λ_2 are a conjugate pair with $\lambda_1 = \mu - i\nu$ and $\lambda_2 = \mu + i\nu$ then

$$\begin{aligned} e^{t\mathbf{A}} &= e^{(\mu-i\nu)t}\mathbf{I} + \frac{e^{(\mu+i\nu)t} - e^{(\mu-i\nu)t}}{2i\nu}\,(\mathbf{A} - (\mu-i\nu)\mathbf{I}) \\ &= \frac{e^{(\mu+i\nu)t} + e^{(\mu-i\nu)t}}{2}\,\mathbf{I} + \frac{e^{(\mu+i\nu)t} - e^{(\mu-i\nu)t}}{2i\nu}\,(\mathbf{A} - \mu\mathbf{I}) \\ &= e^{\mu t}\left[\cos(\nu t)\,\mathbf{I} + \frac{\sin(\nu t)}{\nu}\,(\mathbf{A} - \mu\mathbf{I})\right]\,. \end{aligned}$$

It is clear that if λ_1 and λ_2 are a conjugate pair then computing $e^{t\mathbf{A}}$ by this approach is usually slower than simply applying formula (4.21b) because of the complex arithmetic that is introduced.

For matrices annihilated by a cubic polynonial formula (4.34) requires only one matrix multiplication. There are three cases to consider.

Fact. If $\mathbf{A}$ is annihilated by a cubic polynonial with a triple root λ_1 then

$$e^{t\mathbf{A}} = e^{\lambda_1 t}\mathbf{I} + t\, e^{\lambda_1 t}(\mathbf{A} - \lambda_1\mathbf{I}) + \tfrac{1}{2}t^2 e^{\lambda_1 t}(\mathbf{A} - \lambda_1\mathbf{I})^2 .$$

Reason. Set $z_1 = z_2 = z_3 = \lambda_1$. Then by (4.34b)

$$\mathbf{P}_0 = \mathbf{I}\,, \qquad \mathbf{P}_1 = \mathbf{A} - \lambda_1\mathbf{I}\,, \qquad \mathbf{P}_2 = (\mathbf{A} - \lambda_1\mathbf{I})^2\,,$$

while by (4.33c) the divided-difference table is simply

$$\begin{array}{llllll}
\lambda_1 & e^{\lambda_1 t} & & & & \\
 & & \searrow & & & \\
\lambda_1 & e^{\lambda_1 t} & \to & t\, e^{\lambda_1 t} & & \\
 & & \searrow & & \searrow & \\
\lambda_1 & e^{\lambda_1 t} & \to & t\, e^{\lambda_1 t} & \to & \tfrac{1}{2}t^2 e^{\lambda_1 t}
\end{array}$$

We read off from the top entries of the columns that

$$h_1(t) = e^{\lambda_1 t}\,, \qquad h_2(t) = t\, e^{\lambda_1 t}\,, \qquad h_3(t) = \tfrac{1}{2}t^2 e^{\lambda_1 t}\,,$$

whereby the result follows from (4.34a).

Fact. If $\mathbf{A}$ is annihilated by a cubic polynonial with a double root λ_1 and a simple root λ_2 then

$$e^{t\mathbf{A}} = e^{\lambda_1 t}\mathbf{I} + t\, e^{\lambda_1 t}(\mathbf{A} - \lambda_1\mathbf{I}) + \frac{1}{\lambda_2 - \lambda_1}\left(\frac{e^{\lambda_2 t} - e^{\lambda_1 t}}{\lambda_2 - \lambda_1} - t\, e^{\lambda_1 t}\right)(\mathbf{A} - \lambda_1\mathbf{I})^2 .$$

Reason. Set $z_1 = z_2 = \lambda_1$ and $z_3 = \lambda_2$. Then by (4.34b)

$$\mathbf{P}_0 = \mathbf{I}\,, \qquad \mathbf{P}_1 = \mathbf{A} - \lambda_1\mathbf{I}\,, \qquad \mathbf{P}_2 = (\mathbf{A} - \lambda_1\mathbf{I})^2\,,$$

while by (4.33c) the divided-difference table is simply

$$\begin{array}{llllll}
\lambda_1 & e^{\lambda_1 t} & & & & \\
 & & \searrow & & & \\
\lambda_1 & e^{\lambda_1 t} & \to & t\, e^{\lambda_1 t} & & \\
 & & \searrow & & \searrow & \\
\lambda_2 & e^{\lambda_2 t} & \to & \dfrac{e^{\lambda_2 t} - e^{\lambda_1 t}}{\lambda_2 - \lambda_1} & \to & \dfrac{1}{\lambda_2 - \lambda_1}\left(\dfrac{e^{\lambda_2 t} - e^{\lambda_1 t}}{\lambda_2 - \lambda_1} - t\, e^{\lambda_1 t}\right)
\end{array}$$

We read off from the top entries of the columns that

$$h_1(t) = e^{\lambda_1 t}\,, \qquad h_2(t) = t\, e^{\lambda_1 t}\,, \qquad h_3(t) = \frac{1}{\lambda_2 - \lambda_1}\left(\frac{e^{\lambda_2 t} - e^{\lambda_1 t}}{\lambda_2 - \lambda_1} - t\, e^{\lambda_1 t}\right),$$

whereby the result follows from (4.34a).

Fact. If $\mathbf{A}$ is annihilated by a cubic polynonial with distinct roots λ_1, λ_2, and λ_3 then

$$\begin{aligned}
e^{t\mathbf{A}} = e^{\lambda_1 t}\mathbf{I} &+ \frac{e^{\lambda_2 t} - e^{\lambda_1 t}}{\lambda_2 - \lambda_1}(\mathbf{A} - \lambda_1\mathbf{I}) \\
&+ \frac{1}{\lambda_3 - \lambda_1}\left(\frac{e^{\lambda_3 t} - e^{\lambda_2 t}}{\lambda_3 - \lambda_2} - \frac{e^{\lambda_2 t} - e^{\lambda_1 t}}{\lambda_2 - \lambda_1}\right)(\mathbf{A} - \lambda_1\mathbf{I})(\mathbf{A} - \lambda_2\mathbf{I})\,.
\end{aligned}$$

Reason. Set $z_1 = \lambda_1$, $z_2 = \lambda_2$, and $z_3 = \lambda_3$. Then by (4.34b)

$$\mathbf{P}_0 = \mathbf{I}\,, \qquad \mathbf{P}_1 = \mathbf{A} - \lambda_1 \mathbf{I}\,, \qquad \mathbf{P}_2 = (\mathbf{A} - \lambda_1 \mathbf{I})(\mathbf{A} - \lambda_2 \mathbf{I})\,,$$

while by (4.33c) the divided-difference table is simply

$$\begin{array}{ccccccc}
\lambda_1 & e^{\lambda_1 t} & & & & \\
 & & \searrow & & & \\
\lambda_2 & e^{\lambda_2 t} & \rightarrow & \dfrac{e^{\lambda_2 t} - e^{\lambda_1 t}}{\lambda_2 - \lambda_1} & & \\
 & & \searrow & & \searrow & \\
\lambda_3 & e^{\lambda_3 t} & \rightarrow & \dfrac{e^{\lambda_3 t} - e^{\lambda_2 t}}{\lambda_3 - \lambda_2} & \rightarrow & \dfrac{1}{\lambda_3 - \lambda_1}\left(\dfrac{e^{\lambda_3 t} - e^{\lambda_2 t}}{\lambda_3 - \lambda_2} - \dfrac{e^{\lambda_2 t} - e^{\lambda_1 t}}{\lambda_2 - \lambda_1}\right)
\end{array}$$

We read off from the top entries of the columns that

$$h_1(t) = e^{\lambda_1 t}\,, \qquad h_2(t) = \frac{e^{\lambda_2 t} - e^{\lambda_1 t}}{\lambda_2 - \lambda_1}\,, \qquad h_3(t) = \frac{1}{\lambda_3 - \lambda_1}\left(\frac{e^{\lambda_3 t} - e^{\lambda_2 t}}{\lambda_3 - \lambda_2} - \frac{e^{\lambda_2 t} - e^{\lambda_1 t}}{\lambda_2 - \lambda_1}\right)\,,$$

whereby the result follows from (4.34a).

Example. Given the fact that $p(z) = z^3 - 4z$ annihilates $\mathbf{A}$, compute $e^{t\mathbf{A}}$ for

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}\,.$$

Solution. Because we are told that $p(z) = z^3 - 4z$ annihilates $\mathbf{A}$, we do not have to compute the characteristic polynomial of $\mathbf{A}$. Because $p(z) = z(z+2)(z-2)$, its roots are -2, 0, 2. Set $\lambda_1 = -2$, $\lambda_2 = 0$, and $\lambda_3 = 2$. Then by (4.34b)

$$\mathbf{P}_0 = \mathbf{I}\,, \qquad \mathbf{P}_1 = \mathbf{A} + 2\mathbf{I}\,, \qquad \mathbf{P}_2 = (\mathbf{A} + 2\mathbf{I})\mathbf{A}\,,$$

while by (4.33c) the divided-difference table is

$$\begin{array}{cccccc}
-2 & e^{-2t} & & & & \\
 & & \searrow & & & \\
0 & 1 & \rightarrow & \dfrac{1 - e^{-2t}}{2} & & \\
 & & \searrow & & \searrow & \\
2 & e^{2t} & \rightarrow & \dfrac{e^{2t} - 1}{2} & \rightarrow & \dfrac{1}{4}\,\dfrac{e^{2t} - 2 + e^{-2t}}{2}
\end{array}$$

Therefore

$$
\begin{aligned}
e^{t\mathbf{A}} &= e^{-2t}\mathbf{I} + \frac{1-e^{-2t}}{2}\,(\mathbf{A}+2\mathbf{I}) + \frac{e^{2t}-2+e^{-2t}}{8}\,(\mathbf{A}^2+2\mathbf{A}) \\
&= \mathbf{I} + \frac{e^{2t}-e^{-2t}}{4}\,\mathbf{A} + \frac{e^{2t}-2+e^{-2t}}{8}\,\mathbf{A}^2 \\
&= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \frac{\sinh(2t)}{2} \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} + \frac{\cosh(2t)-1}{4} \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}^2 \\
&= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \frac{\sinh(2t)}{2} \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} + \frac{\cosh(2t)-1}{4} \begin{pmatrix} 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \\ 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \end{pmatrix} .
\end{aligned}
$$

4.7.3. *Application to Other Functions of Matrices.* We can use the same method to compute other functions of matrices. Let $\mathbf{A}$ be an $n{\times}n$ real matrix and $p(z)$ be a polyinomial of degree $m \leq n$ that annihilates $\mathbf{A}$. Let $\{z_1, \cdots, z_m\}$ be the roots of $p(z)$ listed with multiplicity and clustered so that

$$
\text{if } z_j = z_k \text{ for some } j < k \text{ then } z_j = z_{j+1} = \cdots = z_{k-1} = z_k \,.
$$

Let $f(z)$ be a function that is defined and differentiable at every complex z. (Such functions are called *entire*.) For example, let $f(z) = z^r$ for some positive integer r, where we think of r as large. Compute the divided-difference table whose first column is seeded with $h[z_k] = f(z_k)$ and subsequent columns are given by (4.33c). Read off the coefficients h_1, h_2, $\cdots$, h_m from the top entries of each column as prescribed by (4.33d). Then set

$$
f(\mathbf{A}) = h(\mathbf{A}) = \sum_{k=0}^{m-1} h_{k+1}\,\mathbf{P}_k \,,
$$

where the matices $\mathbf{P}_k$ are computed by (4.34b). In this way we can compute $\mathbf{A}^{1000}$ efficiently when n is not too large.

Exercises on Matrix Exponentials

True or False

(1) (a) $e^{t\mathbf{A}} = \sum_{k=0}^{\infty} \frac{1}{k!} t^k \mathbf{A}^k$

(b) If $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ and a_{12} and a_{21} are nonzero, then $e^{t\mathbf{A}} = \begin{pmatrix} e^{ta_{11}} & e^{ta_{12}} \\ e^{ta_{21}} & e^{ta_{22}} \end{pmatrix}$

(c) For any $m \times m$ matrix, $e^{t\mathbf{A}} = N_0(t)\mathbf{I} + N_1(t)\mathbf{A} + \cdots + N_{m-1}(t)\mathbf{A}^{m-1}$, where $N_0(t), N_1(t), \cdots, N_{m-1}(t)$ is the natural fundamental set of solutions for the m^{th}-order differential equation corresponding to matrix $\mathbf{A}$ (with initial time 0).

(d) The natural fundamental set of solutions $N_0(t), N_1(t), \cdots, N_{m-1}(t)$ can be obtained by solving the m^{th}-order differential equation $p(\mathrm{D})y = 0$, with general initial conditions of $y(0) = y_0, y'(0) = y_1, \cdots, y^{(m-1)}(0) = y_{m-1}$, where $p(z)$ is the characteristic polynomial of $\mathbf{A}$

(e) The solution to the initial value problem, $\mathbf{x}' = \mathbf{A}\mathbf{x}$ with $\mathbf{x}(0) = \mathbf{x}^0$, is given by $\mathbf{x}(t) = e^{t\mathbf{A}}\mathbf{x}^0$.

Solution

True or False (The following refers only to 2×2 matrices)

(2) (a) If the characteristic polynomial corresponding to the matrix $\mathbf{A}$, has simple real roots $\mu \pm \nu$, then $e^{t\mathbf{A}} = e^{\mu t}\left[\cos(\nu t)\mathbf{I} + \frac{\sin(\nu t)}{\nu}(\mathbf{A} - \mu\mathbf{I})\right]$. (Note: $\mu \pm \nu$ is obtained by completing the squares on the characteristic polynomial (ie $(z-\mu)^2 - \delta$) with $\nu = \sqrt{\delta}$)

(b) If the characteristic polynomial corresponding to the matrix $\mathbf{A}$, has double roots μ, then $e^{t\mathbf{A}} = e^{\mu t}\left[\mathbf{I} + t(\mathbf{A} - \mu\mathbf{I})\right]$

(c) If the characteristic polynomial corresponding to the matrix $\mathbf{A}$, has complex conjugate roots $\mu \pm i\nu$, then $e^{t\mathbf{A}} = e^{\mu t}\left[\cosh(\nu t)\mathbf{I} + \frac{\sinh(\nu t)}{\nu}(\mathbf{A} - \mu\mathbf{I})\right]$

Solution

(3) Let $\mathbf{A}$ be a 2×2 matrix. Use the Caley-Hamilton Theorem to derive the formula for the inverse of $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$,

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

Solution

(4) Let $\mathbf{A}$ be a 2×2 matrix. Use the Caley-Hamilton Theorem to prove the following formula for the determinant of $\mathbf{A}$,

$$\det(\mathbf{A}) = \frac{1}{2}\left(\operatorname{tr}(\mathbf{A})^2 - \operatorname{tr}\left(\mathbf{A}^2\right)\right).$$

Solution

(5) Consider the matrix

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

Given that the characteristic polynomial is $p(z) = z^3 - 3z - 2$, use the Caley-Hamilton Theorem to find the inverse of $\mathbf{A}$.

Solution

(6) In the text it was stated that for any constant $n \times n$ matrix $\mathbf{A}$ and any t and s, the following property holds

$$e^{(t+s)\mathbf{A}} = e^{t\mathbf{A}} e^{s\mathbf{A}}.$$

As was outlined in the text, show this by showing that both sides of the equation satisfy the same initial value problem.

Solution

For 7–9 verify that the matrix $\mathbf{\Phi}(t)$ is a matrix exponential $e^{t\mathbf{A}}$ for some $\mathbf{A}$ by checking that $\mathbf{\Phi}(t)$ satisfies the properties of a matrix exponential, i.e. $\mathbf{\Phi}(0) = \mathbf{I}$ and $\mathbf{\Phi}(s)\mathbf{\Phi}(t) = \mathbf{\Phi}(t+s)$ for any t, s. Use this to find the inverse $\mathbf{\Phi}(t)^{-1}$.

(7) $\mathbf{\Phi}(t) = \begin{pmatrix} \cos(2t) & \frac{1}{2}\sin(2t) \\ -2\sin(2t) & \cos(2t) \end{pmatrix}$

Solution

(8) $\mathbf{\Phi}(t) = e^t \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$

Solution

(9) $\mathbf{\Phi}(t) = e^t \begin{pmatrix} \cosh(2t) & 2\sinh(2t) \\ \frac{1}{2}\sinh(2t) & \cosh(2t) \end{pmatrix}$

Solution

For 10- 19 compute $e^{t\mathbf{A}}$.

(10) $\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}$

Short Answer
Solution

(11) $\mathbf{A} = \begin{pmatrix} -1 & -4 \\ 1 & -1 \end{pmatrix}$

Short Answer
Solution

(12) $\mathbf{A} = \begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix}$

Short Answer

Solution

(13) $\mathbf{A} = \begin{pmatrix} 1 & \frac{-1}{2} \\ 2 & -1 \end{pmatrix}$

Short Answer
Solution

(14) $\mathbf{A} = \begin{pmatrix} 1 & -1 \\ 5 & -3 \end{pmatrix}$

Short Answer
Solution

(15) $\mathbf{A} = \begin{pmatrix} 7 & -9 \\ 1 & 1 \end{pmatrix}$

Short Answer
Solution

(16) $\mathbf{A} = \begin{pmatrix} 1 & 5 \\ -1 & 3 \end{pmatrix}$

Short Answer
Solution

(17) $\mathbf{A} = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$

Short Answer
Solution

(18) $\mathbf{A} = \begin{pmatrix} 1 & -4 \\ 4 & -7 \end{pmatrix}$

Short Answer
Solution

(19) $\mathbf{A} = \begin{pmatrix} -3 & 2 & 0 \\ -1 & 0 & 0 \\ -3 & 3 & 1 \end{pmatrix}$

Short Answer
Solution

For 20- 25, find the solution to the initial value problem using $e^{t\mathbf{A}}$.

(20) $\mathbf{x}' = \begin{pmatrix} -2 & 1 \\ -5 & 4 \end{pmatrix} \mathbf{x}, \qquad \mathbf{x}(0) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$

Short Answer
Solution

(21) $\mathbf{x}' = \begin{pmatrix} -\frac{5}{2} & \frac{3}{2} \\ -\frac{3}{2} & \frac{1}{2} \end{pmatrix} \mathbf{x}, \qquad \mathbf{x}(0) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$

Short Answer
Solution

(22) $\mathbf{x}' = \begin{pmatrix} 1 & -5 \\ 1 & -3 \end{pmatrix} \mathbf{x}, \qquad \mathbf{x}(0) = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$

Short Answer
Solution

(23) $\mathbf{x}' = \begin{pmatrix} 2 & \frac{-5}{2} \\ \frac{9}{5} & -1 \end{pmatrix} \mathbf{x}, \qquad \mathbf{x}(0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

Short Answer
Solution

(24) $\mathbf{x}' = \begin{pmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 3 & 6 & 2 \end{pmatrix} \mathbf{x}, \qquad \mathbf{x}(0) = \begin{pmatrix} -1 \\ 3 \\ 1 \end{pmatrix}$

Short Answer
Solution

(25) $\mathbf{x}' = \begin{pmatrix} 1 & -1 & 4 \\ 0 & 2 & -1 \\ 0 & 0 & -1 \end{pmatrix} \mathbf{x}, \qquad \mathbf{x}(0) = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$

Short Answer
Solution

(26) Solve the general initial value problems to find the natural fundamental solutions given in equations (4.18a), (4.18b), (4.18c) in the text.

Solution

(27) Let $\mathbf{A}$ be a 3×3 matrix of the form

$$\mathbf{A} = \begin{pmatrix} 0 & a_{12} & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & 0 \end{pmatrix}.$$

Compute the matrix exponential in terms of $\mathbf{A}$. How does this compare to the series representation of the the matrix exponential? Suppose that $\mathbf{A}$ is more generally given by

$$\mathbf{A} = \begin{pmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & 0 & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & a_{(n-1),n} \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Can you guess the form of the matrix exponential in this case? (see exercise 30 in the supplement on matricies and vectors for a hint on the annihilator for $\mathbf{A}$).

Solution

(28) Derive $\frac{\mathrm{d}}{\mathrm{d}t} e^{t\mathbf{A}} = \mathbf{A} e^{t\mathbf{A}}$ using the series representation of $e^{t\mathbf{A}}$. (You may pull derivatives inside infinite sums without justification).

Solution

(29) Compute the natural fundamental sets of $y'' - 6y' + 5y = 0$ using Green functions. Confirm that these are the same natural sets obtained from Problem 10.

Solution

(30) Compute the natural fundamental sets of $-y''' + 4y'' - 5y' + 2y = 0$ using Green functions. (The characteristic polynomial of this equation has roots $z = 1, 1, 2$.) Confirm that these are the same natural sets obtained from Problem 24.

Solution

Calculate $e^{t\mathbf{A}}$ using Hermite interpolation methods.

(31) $\mathbf{A} = \begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix}$

Solution

(32) $\mathbf{A} = \begin{pmatrix} 1 & -4 \\ 4 & -7 \end{pmatrix}$

Solution

Navigation to Other Chapters

This page may not work with every browser-driven pdf viewer. When it does work then it will enable you to link directly to any chapter. When it does not work then you can link to any chapter through the main webpage.

Ordinary Differential Equations

0. Course Introduction and Overview

I. First-Order Ordinary Differential Equations

1. Introduction to First-Order Equations
2. Linear Equations
3. Separable Equations
4. General Theory
5. Graphical Methods
6. Applications
7. Numerical Methods
8. Second-Order Equations Reducible to First-Order Ones
9. Exact Differential Forms and Integrating Factors
10. Special Equations and Substitution

II. Higher-Order Linear Ordinary Differential Equations

1. Introduction to Higher-Order Linear Equations
2. Homogenous Equations: General Methods and Theory
3. Supplement: Linear Algebraic Systems and Deteminants
4. Homogenous Equations with Constant Coefficients
5. Nonhomogeneous Equations: General Methods and Theory
6. Nonhomogeneous Equations with Constant Coefficients
7. Nonhomogeneous Equations with Variable Coefficients
8. Application: Mechanical Vibrations
9. Laplace Transform Method

III. First-Order Systems of Ordinary Differential Equations

1. Introduction to First-Order Systems
2. Linear Systems: General Methods and Theory
3. Supplement: Matrices and Vectors
4. Linear Systems: Matrix Exponentials
5. Linear Systems: Eigen Methods
6. Linear Systems: Laplace Transform Methods
7. Linear Planar Systems
8. Autonomous Planar Systems: Integral Methods
9. Autonomous Planar Systems: Nonintegral Methods
10. Application: Population Dynamics