Matrix Exponentials
Math 246, Spring 2009, Professor David Levermore

We now consider the homogeneous constant coefficient, vector-valued initial-value problem

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{A}\mathbf{x}\,, \qquad \mathbf{x}(t_I) = \mathbf{x}_I\,, \tag{1}$$

where $\mathbf{A}$ is a constant $n\times n$ real matrix. A special fundamental matrix associated with this problem is the solution $\mathbf{\Phi}(t)$ of the matrix-valued initial-value problem

$$\frac{\mathrm{d}\mathbf{\Phi}}{\mathrm{d}t} = \mathbf{A}\mathbf{\Phi}\,, \qquad \mathbf{\Phi}(0) = \mathbf{I}\,, \tag{2}$$

where $\mathbf{I}$ is the $n\times n$ identity matrix. We assert that $\mathbf{\Phi}(t)$ satisfies

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

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

$$\frac{\mathrm{d}\mathbf{\Psi}}{\mathrm{d}t} = \mathbf{A}\mathbf{\Psi}\,, \qquad \mathbf{\Psi}(0) = \mathbf{\Phi}(s)\,,$$

and are therefore equal. Assertion (ii) follows by setting $s = -t$ in assertion (i) and using the fact $\mathbf{\Phi}(0) = \mathbf{I}$. The fundamental matrix $\mathbf{\Phi}(t)$ is therefore called the exponential of $\mathbf{A}$ and is commonly denoted as either $e^{t\mathbf{A}}$ or $\exp(t\mathbf{A})$. It is easy to check that the solution of the initial-value problem (1) is given by $\mathbf{x}(t) = e^{(t-t_I)\mathbf{A}}\mathbf{x}_I$.

The Taylor expansion of $e^{t\mathbf{A}}$ about $t = 0$ is

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

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

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

Motivated by this fact, the book defines $e^{t\mathbf{A}}$ by the infinite series (3).

Matrix KEY Identity. Given any polynomial $p(z) = \pi_0 z^m + \pi_1 z^{m-1} + \cdots + \pi_{m-1} z + \pi_m$ and any $n\times n$ matrix $\mathbf{A}$ we define the $n\times n$ matrix $p(\mathbf{A})$ by

$$p(\mathbf{A}) = \pi_0\mathbf{A}^m + \pi_1\mathbf{A}^{m-1} + \cdots + \pi_{m-1}\mathbf{A} + \pi_m\mathbf{I}\,.$$

Because for every nonnegative integer k one has

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

it follows from the definition of $p(\mathbf{A})$ that

$$p\left(\frac{\mathrm{d}}{\mathrm{d}t}\right) e^{t\mathbf{A}} = p(\mathbf{A}) e^{t\mathbf{A}}\,. \tag{4}$$

This is the matrix version of the KEY identity. Just as the scalar KEY identity allowed us to construct explicit solutions to higher-order linear differential equations with constant coefficients, the matrix KEY identity allows us to construct explicit solutions to first-order linear differential systems with a constant coefficient matrix.

Computing the Matrix Exponential. Given any $n\times n$ matrix $\mathbf{A}$, there are many ways to compute $e^{\mathbf{A}t}$ that are easier than evaluating the infinite series (3). The book gives a method that is based on computing the eigenvectors and (sometimes) the generalized eigenvectors of the matrix $\mathbf{A}$. This method requires a different approach depending on whether the eigenvalues of the real matrix $\mathbf{A}$ are real, complex conjugate, or have multiplicity greater than one. These approaches are covered in Sections 7.5, 7.6, and 7.8, but these sections do not cover all the possible cases that can arise. Here we will give a different method that covers all possible cases with a single approach. Moreover, this method is generally much faster to carry out than the book's method when n is not too large.

This method begins by identifying a polynomial $p(z)$ such that $p(\mathbf{A})=0$. Such a polynomial is said to *annihilate* $\mathbf{A}$. The Cayley-Hamilton Theorem states that one such polynomial is the *characteristic polynomial* of $\mathbf{A}$, which we define by

$$p_{\mathbf{A}}(z) = \det(\mathbf{I}z - \mathbf{A})\,. \tag{5}$$

This polyinomial has degree n. Because $\det(z\mathbf{I}-\mathbf{A}) = (-1)^n \det(\mathbf{A}-z\mathbf{I})$, this definition of $p_{\mathbf{A}}(z)$ coincides with the book's 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. What matters most about $p_{\mathbf{A}}(z)$ is its roots and their multiplicity, which are the same for both conventions. These roots are called the *eigenvalues* of $\mathbf{A}$.

The Cayley-Hamiltion Theorem states that

$$p_{\mathbf{A}}(\mathbf{A}) = 0\,. \tag{6}$$

We will not prove this for general $n\times n$ matrices. However, it is easy to verify for 2×2 matrices by a direct calculation. Consider the general 2×2 matrix

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

Its characteristic polynomial is

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

where $\operatorname{tr}(\mathbf{A}) = a_{11}+a_{22}$ is the trace of $\mathbf{A}$. Then a direct calculation shows that

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

which verifies (6) for 2×2 matrices.

By the above paragraph, you can always find a polynomial $p(z)$ of degree $m \leq n$ that annihilates $\mathbf{A}$. For this polynomial, we see from the matrix KEY identity (4) that

$$p\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)e^{t\mathbf{A}} = p(\mathbf{A})e^{t\mathbf{A}} = 0\,.$$

This means that each entry of $e^{t\mathbf{A}}$ is a solution of the m^{th}-order scalar homogeneous linear differential equation with constant coefficients

$$p\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)y = 0\,. \tag{7}$$

If $y_1(t)$, $y_1(t)$, $\cdots$, $y_m(t)$ is a fundamental set of solutions to this equation then a general solution of it is

$$y = \sum_{j=1}^{m} c_j y_j(t)\,,$$

where c_1, c_2, $\cdots$, c_m are arbitrary constants. It follows that $e^{t\mathbf{A}}$ must have the form

$$e^{t\mathbf{A}} = \sum_{j=1}^{m} \mathbf{C}_j y_j(t)\,, \tag{8}$$

where $\mathbf{C}_1$, $\mathbf{C}_2$, $\cdots$, $\mathbf{C}_m$ are arbitrary $n{\times}n$ constant matrices.

The constant matrices $\mathbf{C}_1$, $\mathbf{C}_2$, $\cdots$, $\mathbf{C}_m$ may be determined by taking derivatives of (8) with respect to t and evaluating them at $t = 0$. The k^{th} derivative of (8) evaluated at $t = 0$ gives

$$\mathbf{A}^k = \sum_{j=1}^{m} \mathbf{C}_j y_j^{(k)}(0)\,. \tag{9}$$

Because $y_1(t)$, $y_1(t)$, $\cdots$, $y_m(t)$ is a fundamental set of solutions to (7), the constant matrices $\mathbf{C}_1$, $\mathbf{C}_2$, $\cdots$, $\mathbf{C}_m$ are determined by (9) for $k = 0$, 1, $\cdots$, $m-1$.

For example, if $p(z)$ has m simple roots λ_1, λ_2, $\cdots$, λ_m, then one can choose the fundamental set of solutions to (7) given by

$$y_j(t) = e^{\lambda_j t}\,, \quad \text{for } j = 1, 2, \cdots, m\,.$$

Then (8) becomes the system of m linear equations

$$\mathbf{A}^k = \sum_{j=1}^{m} \mathbf{C}_j \lambda_j^k\,, \quad \text{for } k = 0, 1, \cdots, m-1\,. \tag{10}$$

This system may be solved for the constant matrices $\mathbf{C}_1$, $\mathbf{C}_2$, $\cdots$, $\mathbf{C}_m$, and the result placed into (8) to obtain

$$e^{t\mathbf{A}} = \sum_{j=1}^{m} \mathbf{C}_j e^{\lambda_j t}\,.$$

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

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

Solution. Because $\mathbf{A}$ is $2{\times}2$, its characteristic polynomial is

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

Its roots are 1 and 5. System (10) then becomes

$$\mathbf{I} = \mathbf{C}_1 + \mathbf{C}_2\,, \qquad \mathbf{A} = \mathbf{C}_1 + 5\mathbf{C}_2\,.$$

This system can be easily solved to find

$$\mathbf{C}_1 = \tfrac{1}{4}(5\mathbf{I} - \mathbf{A}) = \tfrac{1}{4}\begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix} = \tfrac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\,,$$

$$\mathbf{C}_2 = \tfrac{1}{4}(\mathbf{A} - \mathbf{I}) = \tfrac{1}{4}\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} = \tfrac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\,.$$

Formula (8) then yields

$$e^{t\mathbf{A}} = \mathbf{C}_1 e^t + \mathbf{C}_2 e^{5t} = \tfrac{1}{2}\begin{pmatrix} e^t + e^{5t} & e^{5t} - e^t \\ e^{5t} - e^t & e^t + e^{5t} \end{pmatrix}\,.$$

Exponentials of Two-by-Two Matrices. There are simple formulas for the exponential of a general 2×2 real matrix

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

Because $\mathbf{A}$ is 2×2, its characteristic polynomial is

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

Upon completing the square we see that

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

where the mean μ and discriminant δ are given by

$$\mu = \frac{\operatorname{tr}(\mathbf{A})}{2}\,, \qquad \delta = \frac{\operatorname{tr}(\mathbf{A})^2 - 4\det(\mathbf{A})}{4}\,.$$

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

- If $\delta > 0$ then $p(z)$ has the simple real roots $\mu - \nu$ and $\mu + \nu$ where $\nu = \sqrt{\delta}$. In this case

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

- If $\delta < 0$ then $p(z)$ has the complex conjugate roots $\mu - i\nu$ and $\mu + i\nu$ where $\nu = \sqrt{-\delta}$. In this case

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

- If $\delta = 0$ then $p(z)$ has the double real root μ. In this case

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

For 2×2 matrices you will find it faster to apply these formulas than to set-up and solve for $\mathbf{C}_1$ and $\mathbf{C}_2$. These formulas are eazy to remember. Notice that formulas (11) and (12) 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 (13) is the limiting case of both (11) and (12) as $\nu \to 0$.

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

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

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

$$p(z) = \det(\mathbf{I}z - \mathbf{A}) = z^2 - \operatorname{tr}(\mathbf{A})z + \det(\mathbf{A}) = z^2 - 6z + 5 = (z-3)^2 - 4 .$$

It has the real roots 3 ± 2. By (11) with $\mu = 3$ and $\nu = 2$ we see that

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

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

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

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

$$\begin{aligned} p(z) &= \det(\mathbf{I}z - \mathbf{A}) = z^2 - \operatorname{tr}(\mathbf{A})z + \det(\mathbf{A}) \\ &= z^2 - 4z + 13 = (z-2)^2 + 3^2 . \end{aligned}$$

It has the conjugate roots $2 \pm i3$. By (12) with $\mu = 2$ and $\nu = 3$ we see that

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

Use of Natural Fundamental Sets. The *natural fundamental set* of solutions to (7) are the solutions $y_1(t)$, $y_1(t)$, $\cdots$, $y_m(t)$ such that for each $j = 1, 2, \cdots, m$ the solution $y_j(t)$ satifies the initial conditions

$$y_j^{(k-1)}(0) = \delta_{jk} \quad \text{for } k = 1, 2, \cdots, m . \tag{14}$$

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

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

If $y_1(t)$, $y_1(t)$, $\cdots$, $y_m(t)$ is the natural fundamental set of solutions to (7) then (9) with $k-1$ replacing k becomes

$$\mathbf{A}^{k-1} = \sum_{j=1}^{m} \mathbf{C}_j y_j^{(k-1)}(0) = \sum_{j=1}^{m} \mathbf{C}_j \delta_{jk} = \mathbf{C}_k \quad \text{for } k = 1, 2, \cdots, m .$$

In that case (8) becomes

$$e^{t\mathbf{A}} = \sum_{j=1}^{m} \mathbf{A}^{j-1} y_j(t)\,, \tag{15}$$

If the natural fundamental set of solutions to (7) is either known or easily found then this is the shortest route to computing $e^{t\mathbf{A}}$ when m is not too large.

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

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

Solution. The characteristic polynomial of $\mathbf{A}$ is

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

Its roots are 0, $\pm i3$. The associated higher-order equation is

$$\frac{\mathrm{d}^3 y}{\mathrm{d}t^3} + 9\frac{\mathrm{d}y}{\mathrm{d}t} = 0\,.$$

By (10) its natural fundamental set of solutions $y_1(t)$, $y_2(t)$, and $y_3(t)$ satisfy the initial conditions

$$\begin{aligned} y_1(0) &= 1\,, & y_1'(0) &= 0\,, & y_1''(0) &= 0\,, \\ y_2(0) &= 0\,, & y_2'(0) &= 1\,, & y_2''(0) &= 0\,, \\ y_3(0) &= 0\,, & y_3'(0) &= 0\,, & y_3''(0) &= 1\,. \end{aligned}$$

You can solve these three initial-value problems to find

$$y_1(t) = 1\,, \qquad y_2(t) = \frac{\sin(3t)}{3}\,, \qquad y_3(t) = \frac{1 - \cos(3t)}{9}\,. \tag{16}$$

We will see a more efficient way to find these solutions a bit later, so we will not give any details here. Given these solutions, formula (15) yields

$$\begin{aligned} e^{t\mathbf{A}} &= \mathbf{I}y_1(t) + \mathbf{A}y_2(t) + \mathbf{A}^2 y_3(t) \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 2 & -1 \\ -2 & 0 & 2 \\ 1 & -2 & 0 \end{pmatrix} \frac{\sin(3t)}{3} + \begin{pmatrix} 0 & 2 & -1 \\ -2 & 0 & 2 \\ 1 & -2 & 0 \end{pmatrix}^2 \frac{1 - \cos(3t)}{9} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 2 & -1 \\ -2 & 0 & 2 \\ 1 & -2 & 0 \end{pmatrix} \frac{\sin(3t)}{3} + \begin{pmatrix} -5 & 2 & 4 \\ 2 & -8 & 2 \\ 4 & 2 & -5 \end{pmatrix} \frac{1 - \cos(3t)}{9} \\ &= \begin{pmatrix} \frac{4}{9} + \frac{5}{9}\cos(3t) & \frac{2}{9} - \frac{2}{9}\cos(3t) + \frac{2}{3}\sin(3t) & \frac{4}{9} - \frac{4}{9}\cos(3t) - \frac{1}{3}\sin(3t) \\ \frac{2}{9} - \frac{2}{9}\cos(3t) - \frac{2}{3}\sin(3t) & \frac{1}{9} + \frac{8}{9}\cos(3t) & \frac{2}{9} - \frac{2}{9}\cos(3t) + \frac{2}{3}\sin(3t) \\ \frac{4}{9} - \frac{4}{9}\cos(3t) + \frac{1}{3}\sin(3t) & \frac{2}{9} - \frac{2}{9}\cos(3t) + \frac{2}{3}\sin(3t) & \frac{4}{9} + \frac{5}{9}\cos(3t) \end{pmatrix}. \end{aligned}$$

Three Examples. Formulas (11), (12), and (13) for the exponential of 2×2 matrices can be easily derived using the natural fundamental set of solutions to the equation

$$p\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)y = 0\,, \qquad \text{where} \quad p(z) = (z-\mu)^2 - \delta\,.$$

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

- If $\delta > 0$ then $p(z)$ has the simple real roots $\mu - \nu$ and $\mu + \nu$ where $\nu = \sqrt{\delta}$. In this case the natural fundamental set of solutions is

(17)
$$y_1(t) = e^{\mu t}\cosh(\nu t) - \mu e^{\mu t}\frac{\sinh(\nu t)}{\nu}\,, \qquad y_2(t) = e^{\mu t}\frac{\sinh(\nu t)}{\nu}\,.$$

- If $\delta < 0$ then $p(z)$ has the complex conjugate roots $\mu - i\nu$ and $\mu + i\nu$ where $\nu = \sqrt{-\delta}$. In this case the natural fundamental set of solutions is

(18)
$$y_1(t) = e^{\mu t}\cos(\nu t) - \mu e^{\mu t}\frac{\sin(\nu t)}{\nu}\,, \qquad y_2(t) = e^{\mu t}\frac{\sin(\nu t)}{\nu}\,.$$

- If $\delta = 0$ then $p(z)$ has the double real root μ. In this case the natural fundamental set of solutions is

(19)
$$y_1(t) = e^{\mu t} - \mu e^{\mu t}t\,, \qquad y_2(t) = e^{\mu t}t\,.$$

Then by (15), formulas (11), (12), and (13) are abtained by plugging the natural fundamental sets of solutions (17), (18), and (19) respectively into

$$e^{t\mathbf{A}} = \mathbf{I}y_1(t) + \mathbf{A}y_2(t)\,.$$

Notice that (19) is the limiting case of both (17) and (18) as $\nu \to 0$.

Remark. The above examples shows that, once the natural fundamental set of solutions is found for the associated higher-order equation, employing formula (15) is straight forward. It requires only computing $\mathbf{A}^k$ up to $k = m-1$ and some addition. For $m \geq 2$ this requires $(m-2)n^3$ multiplications, which grows fast as m and n get large. (Often $m = n$.) However, for small systems like the ones you will face in this course, it is generally the fastest method. It will become even faster once you learn how to quickly generate the natural fundamental set of solutions for the associated higher-order equation from its Green function.

Generating Natural Fundamental Sets with Green Functions. In each of the natural fundamental sets of solutions given by (17), (18), and (19), the solutions $y_1(t)$ and $y_2(t)$ are related by

$$y_1(t) = y_2'(t) - 2\mu y_2(t)\,.$$

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

(20)
$$p\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)y = 0\,, \qquad \text{where} \quad p(z) = z^m + \pi_1 z^{m-1} + \cdots + \pi_{m-1}z + \pi_m\,,$$

one can generate its entire natural fundamental set of solutions from the *Green function* $g(t)$ associated with (20). Recall that the Green function $g(t)$ satisfies the initial-value problem

(21)
$$p\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)g = 0\,, \qquad g(0) = g'(0) = \cdots = g^{(m-2)}(0) = 0\,, \quad g^{(m-1)}(0) = 1\,.$$

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

$$
\begin{aligned}
y_m(t) &= g(t)\,, \\
y_{m-1}(t) &= g'(t) + \pi_1 g(t)\,, \\
y_{m-2}(t) &= g''(t) + \pi_1 g'(t) + \pi_2 g(t)\,, \\
&\vdots \\
y_2(t) &= g^{(m-2)}(t) + \pi_1 g^{(m-3)}(t) + \cdots + \pi_{m-3} g'(t) + \pi_{m-2} g(t)\,, \\
y_1(t) &= g^{(m-1)}(t) + \pi_1 g^{(m-2)}(t) + \pi_2 g^{(m-3)}(t) + \cdots + \pi_{m-2} g'(t) + \pi_{m-1} g(t)\,.
\end{aligned}
\tag{22}
$$

This entire set is thereby generated by the solution of the single initial-value problem (21).

Example. Show that (16) is indeed the natural fundamental set of solutions to the equation

$$\frac{\mathrm{d}^3 y}{\mathrm{d}t^3} + 9\frac{\mathrm{d}y}{\mathrm{d}t} = 0\,.$$

Solution. By (21) the Green function $g(t)$ satisfies the initial-value problem

$$\frac{\mathrm{d}^3 g}{\mathrm{d}t^3} + 9\frac{\mathrm{d}g}{\mathrm{d}t} = 0\,, \qquad g(0) = g'(0) = 0\,, \quad g''(0) = 1\,.$$

The characteristic polynomial of this equation is $p(z) = z^3 + 9z$, which has roots 0, $\pm i3$. We therefore 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 inital conditions for $g(t)$ then yield the algebraic system

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

The solution of this system is $c_1 = \frac{1}{9}$, $c_2 = -\frac{1}{9}$, 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 (20) that $\pi_1 = 0$, $\pi_2 = 9$, and $\pi_3 = 0$. Then by recipe (22) the natural fundamental set of solutions is given by

$$
\begin{aligned}
y_3(t) &= g(t) = \frac{1 - \cos(3t)}{9}\,, \\
y_2(t) &= g'(t) + 0 \cdot g(t) = \frac{\sin(3t)}{3}\,, \\
y_1(t) &= g''(t) + 0 \cdot g'(t) + 9 \cdot g(t) = \cos(3t) + 9 \cdot \frac{1 - \cos(3t)}{9} = 1\,.
\end{aligned}
$$

This is indeed the set given by (16).

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 you are told that $p(z) = z^3 - 4z$ annihilates $\mathbf{A}$, you do not have to compute the characteristic polynomial of $\mathbf{A}$. By (21) the Green function $g(t)$ satisfies the initial-value problem

$$\frac{\mathrm{d}^3 g}{\mathrm{d}t^3} - 4\frac{\mathrm{d}g}{\mathrm{d}t} = 0\,, \qquad g(0) = g'(0) = 0\,, \quad g''(0) = 1\,.$$

The characteristic polynomial of this equation is $p(z) = z^3 - 4z$, which has roots 0, ± 2. We therefore 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 inital 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 (20) that $\pi_1 = 0$, $\pi_2 = -4$, and $\pi_3 = 0$. Then by recipe (22) the natural fundamental set of solutions is given by

$$\begin{aligned}
y_3(t) &= g(t) = \frac{\cosh(2t) - 1}{4}\,,\\
y_2(t) &= g'(t) + 0 \cdot g(t) = \frac{\sinh(2t)}{2}\,,\\
y_1(t) &= g''(t) + 0 \cdot g'(t) - 4 \cdot g(t) = \cosh(2t) - 4 \cdot \frac{\cosh(2t) - 1}{4} = 1\,.
\end{aligned}$$

Given these solutions, formula (15) yields

$$\begin{aligned}
e^{t\mathbf{A}} &= \mathbf{I}y_1(t) + \mathbf{A}y_2(t) + \mathbf{A}^2 y_3(t)\\
&= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \frac{\sinh(2t)}{2} + \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}^2 \frac{\cosh(2t) - 1}{4}\\
&= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \frac{\sinh(2t)}{2} + \begin{pmatrix} 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \\ 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \end{pmatrix} \frac{\cosh(2t) - 1}{4}\\
&= \begin{pmatrix} \frac{1}{2} + \frac{1}{2}\cosh(2t) & \frac{1}{2}\sinh(2t) & \frac{1}{2}\cosh(2t) - \frac{1}{2} & \frac{1}{2}\sinh(2t) \\ \frac{1}{2}\sinh(2t) & \frac{1}{2} + \frac{1}{2}\cosh(2t) & \frac{1}{2}\sinh(2t) & \frac{1}{2}\cosh(2t) - \frac{1}{2} \\ \frac{1}{2}\cosh(2t) - \frac{1}{2} & \frac{1}{2}\sinh(2t) & \frac{1}{2} + \frac{1}{2}\cosh(2t) & \frac{1}{2}\sinh(2t) \\ \frac{1}{2}\sinh(2t) & \frac{1}{2}\cosh(2t) - \frac{1}{2} & \frac{1}{2}\sinh(2t) & \frac{1}{2} + \frac{1}{2}\cosh(2t) \end{pmatrix}.
\end{aligned}$$

Remark. The polynomial $p(z) = z^3 - 4z$ used in the above example is not the characteristic polynomial of $\mathbf{A}$. With some effort you can check that $p_{\mathbf{A}}(z) = z^4 - 4z^2$. We saved quite a bit of work in computing $e^{t\mathbf{A}}$ by using $p(z) = z^3 - 4z$ as the annihilating polynomial rather than $p_{\mathbf{A}}(z) = z^4 - 4z^2$ because it has a lower degree. Given the characteristic polynomial $p_{\mathbf{A}}(z)$ of a matrix $\mathbf{A}$, every annihilating polynomial $p(z)$ will have the same roots as $p_{\mathbf{A}}(z)$, but these roots might have lower multiplicity. If $p(z)$ has simple roots then it has the lowest degree possible for an annihilating polynomial of $\mathbf{A}$. If $p_{\mathbf{A}}(z)$ has roots that are not simple then it pays to find an annihilating polynomial of lower degree. If $\mathbf{A}$ is either symmetric ($\mathbf{A}^T = \mathbf{A}$) or skew-symmetric ($\mathbf{A}^T = -\mathbf{A}$) then it has an annihilating polynomial with simple roots. In the example above, because $\mathbf{A}$ is symmetric and $p_{\mathbf{A}}(z)$ has roots -2, 0, 0, and 2, we know that $p(z) = (z+2)z(z-2) = z^3 - 4z$ is an annihilating polynomial. Because each root of $p(z)$ is simple, it has the lowest degree possible for an annihilating polynomial of $\mathbf{A}$.

Justification of Recipe (22). The following justification of recipe (22) is included for completeness. It was not covered in lecture and you do not need to know this argument. However, you should find the recipe itself quite useful.

We see from (21) that $g(t)$ is a solution of

$$p\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)y = y^{(m)} + \pi_1 y^{(m-1)} + \pi_2 y^{(m-2)} + \cdots + \pi_{m-1} y' + \pi_m y = 0\,, \tag{23}$$

so that all of its derivatives are too. Each $y_j(t)$ defined by (22) must also be a solution of (23) because it is a linear combination of $g(t)$ and its derivatives. The only thing that remains to be checked is that the initial conditions (14) are satisfied.

Because $y_m(t) = g(t)$, we see from (21) that the initial conditions (14) hold for $y_m(t)$. The key step is to show that if the initial conditions (14) hold for $y_{j+1}(t)$ for some $j < m$ then they hold for $y_j(t)$. Once this is done then we can argue that because the initial conditions (14) hold for $y_m(t)$, they also hold for $y_{m-1}(t)$, which implies they also hold for $y_{m-2}(t)$, which implies they also hold for $y_{m-3}(t)$, and so on down to $y_1(t)$.

We now prove the key step. We suppose that for some $j < m$ the initial conditions (14) hold for $y_{j+1}(t)$. This is the same as

$$y_{j+1}^{(k)}(0) = \delta_{jk} \quad \text{for } k = 0,\, 1,\, \cdots,\, m-1\,. \tag{24}$$

Because $y_{j+1}(t)$ satisfies (23), it follows from the above that

$$\begin{aligned} 0 = p\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)y_{j+1}(t)\bigg|_{t=0} &= y_{j+1}^{(m)}(0) + \sum_{k=0}^{m-1} \pi_{m-k} y_{j+1}^{(k)}(0) \\ &= y_{j+1}^{(m)}(0) + \sum_{k=0}^{m-1} \pi_{m-k}\delta_{jk} = y_{j+1}^{(m)}(0) + \pi_{m-j}\,. \end{aligned} \tag{25}$$

We see from recipe (22) that $y_j(t)$ is related to $y_{j+1}(t)$ by $y_j(t) = y'_{j+1}(t) + \pi_{m-j} g(t)$. We evaluate the $(k-1)^{st}$ derivative of this relation at $t = 0$ to obtain

$$y_j^{(k-1)}(0) = y_{j+1}^{(k)}(0) + \pi_{m-j} g^{(k-1)}(0) \quad \text{for } k = 1,\, 2,\, \cdots,\, m\,.$$

Because $g(t)$ satisfies the initial condtions in (21), we see from (24) that this becomes

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

while we see from (25) that for $k = m$ it becomes

$$y_j^{(m-1)}(0) = y_{j+1}^{(m)}(0) + \pi_{m-j} = 0\,.$$

The initial conditions (14) thereby hold for $y_j(t)$. This completes the proof of the key step, which completes the justification of recipe (22).