# Matrix Exponentials
## Math 246, Spring 2008, Professor David Levermore

The exponential of an $n\times n$ matrix $\mathbf{A}$ is defined to be the solution $\mathbf{\Phi}(t)$ of the matrix-valued initial-value problem

$$\frac{\mathrm{d}\mathbf{\Phi}}{\mathrm{d}t} = \mathbf{A}\mathbf{\Phi}\,, \qquad \mathbf{\Phi}(0) = \mathbf{I}\,, \tag{1}$$

where $\mathbf{I}$ is the $n\times n$ identity matrix. It is commonly denoted as $e^{t\mathbf{A}}$ or as $\exp(t\mathbf{A})$. This is a special fundamental matrix associated with the vector-valued initial-value problem

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

It is easy to check that the solution of this problem 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

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

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 (2).

We can show that $e^{t\mathbf{A}}$ satisfies

$$\begin{aligned}
&\text{(i)} \quad e^{0\mathbf{A}} = \mathbf{I}\,,\\
&\text{(ii)} \quad e^{(t+s)\mathbf{A}} = e^{t\mathbf{A}}e^{s\mathbf{A}} \quad \text{for every } t \text{ and } s \text{ in } \mathbb{R}\,,\\
&\text{(iii)} \quad e^{t\mathbf{A}}e^{-t\mathbf{A}} = \mathbf{I} \quad \text{for every } t \text{ in } \mathbb{R}\,.
\end{aligned}$$

Assertion (i) follows directly from (1). Assertion (ii) follows because both sides satisfy the matrix-valued initial-value problem

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

and are therefore equal. Assertion (iii) follows by setting $s=-t$ in assertion (ii) and using assertion (i).

**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

$$p\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)e^{t\mathbf{A}} = p(\mathbf{A})e^{t\mathbf{A}}\,. \tag{3}$$

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 (2). 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 polynomials are said to *annihilate* $\mathbf{A}$. The Cayley-Hamilton Theorem states that one such polynomial is the *characteristic polynomial* of $\mathbf{A}$, which is given by

$$p_{\mathbf{A}}(z) = \det(\mathbf{I}z - \mathbf{A})\,. \tag{4}$$

This polyinomial has degree $n$. Its roots are the eigenvalues of $\mathbf{A}$. The Cayley-Hamiltion Theorem states that

$$p_{\mathbf{A}}(\mathbf{A}) = 0\,. \tag{5}$$

We will not prove this for general $n\times n$ matrices. However, it is easy to verify for $2\times 2$ matrices by a direct calculation. Consider the general $2\times 2$ 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 (5) for $2\times 2$ 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 (3) 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

$$p\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)y = 0\,. \tag{6}$$

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

$$e^{t\mathbf{A}} = \sum_{j=1}^{m} \mathbf{C}_j y_j(t)\,, \tag{7}$$

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 (7) with respect to $t$ and evaluating them at $t = 0$. The $k^{th}$ derivative of (7) evaluated at $t = 0$ gives

$$\mathbf{A}^k = \sum_{j=1}^{m} \mathbf{C}_j y_j^{(k)}(0)\,. \tag{8}$$

Because $y_1(t)$, $y_1(t)$, $\cdots$, $y_m(t)$ is a fundamental set of solutions to (6), the constant matrices $\mathbf{C}_1$, $\mathbf{C}_2$, $\cdots$, $\mathbf{C}_m$ are determined by (8) for $k = 0,\,1,\,\cdots,\,m-1$.

For example, if $p(z)$ has $m$ simple roots $\lambda_1$, $\lambda_2$, $\cdots$, $\lambda_m$, then one can choose the fundamental set of solutions to (6) 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

$$\mathbf{A}^k = \sum_{j=1}^{m} \mathbf{C}_j \lambda_j^k\,, \quad \text{for } k = 0,\,1,\,\cdots,\,m-1\,. \tag{9}$$

This system may be solved for the constant matrices $\mathbf{C}_1$, $\mathbf{C}_2$, $\cdots$, $\mathbf{C}_m$, and the result placed into (7) 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 (9) 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 (7) 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.** Using the above approach we can easily derive formulas for exponentials of the 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 $\mu$ and discriminant $\delta$ 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 $\delta$.

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

(10) $$e^{t\mathbf{A}} = \mathbf{I}e^{\mu t}\cosh(\nu t) + (\mathbf{A} - \mu\mathbf{I})e^{\mu t}\frac{\sinh(\nu t)}{\nu}\,.$$

- 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

(11) $$e^{t\mathbf{A}} = \mathbf{I}e^{\mu t}\cos(\nu t) + (\mathbf{A} - \mu\mathbf{I})e^{\mu t}\frac{\sin(\nu t)}{\nu}\,.$$

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

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

Notice that (12) is the limiting case of both (10) and (11) 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 $3 \pm 2$. By (10) with $\mu = 3$ and $\nu = 2$ we see that

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

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

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

**Solution.** Because $\mathbf{A}$ is $2{\times}2$, 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 (11) with $\mu = 2$ and $\nu = 3$ we see that

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

**Use of Natural Fundamental Sets.** The *natural fundamental set* of solutions to (6) 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

$$y_j^{(k-1)}(0) = \delta_{jk} \quad \text{for } k = 1, 2, \cdots, m . \tag{13}$$

where $\delta_{jk}$ 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 (6) then (8) 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 (7) becomes

$$e^{t\mathbf{A}} = \sum_{j=1}^{m} \mathbf{A}^{j-1} y_j(t) , \tag{14}$$

If the natural fundamental set of solutions to (6) 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 (13) 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

$$y_1(t) = 1\,, \qquad y_2(t) = \frac{\sin(3t)}{3}\,, \qquad y_3(t) = \frac{1-\cos(3t)}{9}\,. \tag{15}$$

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 (14) yields

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

**Remark.** The above example shows that, once the natural fundamental set of solutions is found for the associated higher-order equation, employing formula (14) is straight forward. It requires only computing $\mathbf{A}^k$ up to $k = m-1$ and some addition. For $m \geq 2$ this requires $(m-2)n^2$ multiplications, which grows fast as $m$ and $n$ gets large. (Often $m = n$.) However, for small systems like the ones you will face in this course, it is generally the fastest method.

The formulas (10), (11), and (12) 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 $\delta$.

- 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

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

- 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

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

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

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

Then by (14), formulas (10), (11), and (12) are abtained by plugging the natural fundamental sets of solutions (16), (17), and (18) respectively into

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

Notice that (18) is the limiting case of both (16) and (17) as $\nu \to 0$.

In each of the natural fundamental sets of solutions given by (16), (17), and (18), 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

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

one can generate its entire natural fundamental set of solutions from the solution $g(t)$ of the single initial-value problem

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

Specifically, the natural fundamental set of solutions is given by the recipe

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

The solution $g(t)$ of the initial-value problem (20) is called the *Green function* of (19).

**Example.** Show that (15) 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 (20) 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 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

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

Then by recipe (21) the natural fundamental set of solutions is given by

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

This is indeed the set given by (15).

**Justification of Recipe (21).** The following justification of recipe (21) 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 (20) that $g(t)$ is a solution of

$$p\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)y = y^{(m)} + \pi_1 y^{(m-1)} + \pi_2 y^{(m-2)} + \cdots + \pi_{m-1}y' + \pi_m y = 0\,, \tag{22}$$

so that all of its derivatives are too. Each $y_j(t)$ defined by (21) must also be a solution of (22) 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 (13) are satisfied.

Because $y_m(t) = g(t)$, we see from (20) that the initial conditions (13) hold for $y_m(t)$. The key step is to show that if the initial conditions (13) 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 (13) 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 (13) hold for $y_{j+1}(t)$. This is the same as

$$y_{j+1}^{(k)}(0) = \delta_{jk} \quad \text{for } k = 0, 1, \cdots, m-1\,. \tag{23}$$

Because $y_{j+1}(t)$ satisfies (22), it follows from the above that

$$\begin{aligned} 0 = p\left(\frac{\mathrm{d}}{\mathrm{d}t}\right) y_{j+1}(t)\bigg|_{t=0} &= y_{j+1}^{(m)}(0) + \sum_{k=0}^{m-1} \pi_{m-k} y_{j+1}^{(k)}(0) \\ &= y_{j+1}^{(m)}(0) + \sum_{k=0}^{m-1} \pi_{m-k} \delta_{jk} = y_{j+1}^{(m)}(0) + \pi_{m-j}\,. \end{aligned} \tag{24}$$

We see from recipe (21) 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 (20), we see from (23) that this becomes

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

while we see from (24) that for $k = m$ it becomes

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

The initial conditions (13) thereby hold for $y_j(t)$. This completes the proof of the key step, which completes the justification of recipe (21).