Suppose an $n \times n$ matrix $A$ has $\lambda = 3$ as an eigenvalue: then $(3I_n - A)v_0 = 0$ for the corresponding eigenvector $v_0$.

So: plugging $A$ into the polynomial $3-x$ yields a matrix annihilating $v_0$. Is there a polynomial into which plugging $A$ gives a matrix annihilating all vectors? — i.e., the zero matrix?

Consider $M_n$, as a vector space over $\IR$ of dimension $n^2$.

Clearly $1, A, A^2, \ldots, A^{n^2}$ are dependent (there are $n^2+1$ of them) \[ \Rightarrow \exists \alpha_0, \ldots, \alpha_{n^2} \in \IR \text{ s.t. } \sum_{k=0}^{n^2} \alpha_k A^k = 0. \]

So indeed, writing $q(x) = \sum_{k=0}^{n^2} \alpha_k x^k$, $q(A) = 0$. But maybe we can do better, i.e. succeed with a polynomial of smaller degree?

If $A$ is diagonalizable with eigenvalues $\lambda_1, \ldots, \lambda_n$, then $(\lambda_1 I - A)(\lambda_2 I - A) \cdots (\lambda_n I - A) v = 0$ for every $v$, by writing $v = \sum_{m=1}^{n^2} \alpha_m v_m$ as a linear combination of eigenvectors.

(Use the fact that the $(\lambda_i I - A)$ all commute.) So then writing $p(x) = (\lambda_1 - x) \cdots (\lambda_n - x)$, we get $p(A) = 0$.

But this $p$ is the characteristic polynomial!

* Small point: constants $c$ become $cI_n (= c \cdot A^0)$
Perhaps plugging $A$ into its characteristic polynomial always gives zero? We should be careful — diagonalizable matrices are not representative of the ‘general’ case. Also, the first idea that comes to mind for proving such a thing — $\det(A\mathbf{I}-A) = \det(0) = 0$ — is wrong: Substituting in $A$ before you take the determinant is not the same as doing so afterwards! On the other hand, here’s a nondiagonalizable $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, with characteristic poly. $P_A(\lambda) := \det\begin{pmatrix} \lambda & -1 & 0 \\ 0 & \lambda & -1 \\ 0 & 0 & \lambda \end{pmatrix} = \lambda^3$, and indeed $A^3 = 0$. So perhaps the following isn’t surprising after all:

**Theorem (Cayley–Hamilton):** Writing $P_A(\lambda) := \det(\lambda\mathbf{I}_n-A)$, we have $P_A(A) = 0$.

**Proof:** We work with matrices whose entries are polynomials in $\lambda$. So we can’t divide by $\lambda$, or polynomials in $\lambda$, like the determinant $\det$ such a matrix. But $(i,j)_{th}$ entry of the matrix product $M \cdot \text{adj}(M) = \left[ \sum_{k=1}^{n} m_{ik} \cdot (-1)^{i+j} \det M_{jk} \right]_{i,j}$ is the determinant of the matrix obtained by replacing the $j$th row of $M$ by the $i$th row. So get $\det M$ if $i=j$, 0 if $i\neq j$. 

$$M \cdot \text{adj}(M) = \left[ \begin{array}{ccc} \det(M) & 0 & \ldots \\ \vdots & \vdots & \ddots \\ 0 & \ldots & \det(M) \end{array} \right] = \det(M) \mathbf{I}_n$$
is true for such matrices. Applying this to $M = \lambda \mathbb{I} - A$,

$$(\lambda \mathbb{I} - A) \cdot \text{adj}(\lambda \mathbb{I} - A) = \text{det}(\lambda \mathbb{I} - A) \mathbb{I} = P_A(\lambda) \mathbb{I}.$$ 

Now decompose $\text{adj}(\lambda \mathbb{I} - A) = \sum \lambda^k S_k$, and write 

$P_A(\lambda) = \sum_{j=0}^{n} a_j \lambda^j$. We have

$$(\lambda \mathbb{I} - A) \sum_{k=0}^{n} \lambda^k S_k = \sum_{j=0}^{n} a_j \lambda^j \mathbb{I}.$$ 

$$\Rightarrow$$ 

$$-A S_0 + \lambda (S_0 - A S_1) + \lambda^2 (S_1 - A S_2) + \ldots + \lambda^{n-1} (S_{n-2} - A S_{n-1}) + \lambda^n S_{n-1}$$

$$= a_0 \mathbb{I} + \lambda (a_1 \mathbb{I}) + \lambda^2 (a_2 \mathbb{I}) + \ldots + \lambda^{n-1} (a_{n-2} \mathbb{I}) + \lambda^n (a_{n-1} \mathbb{I})$$

$$\Rightarrow a_0 \mathbb{I} = -A S_0, \quad a_1 \mathbb{I} = S_0 - A S_1, \quad a_2 \mathbb{I} = S_1 - A S_2,$$

$$\ldots, \quad a_{n-1} \mathbb{I} = S_{n-2} - A S_{n-1}, \quad a_n \mathbb{I} = S_{n-1}.$$ 

$$\Rightarrow P_A(A) = \sum_{j=0}^{n} a_j A^j = a_0 \mathbb{I} + A(a_1 \mathbb{I}) + A^2(a_2 \mathbb{I}) + \ldots + A^{n-1}(a_{n-1} \mathbb{I}) + A^n(a_n \mathbb{I})$$

$$= -A S_0 + A (S_0 - A S_1) + A^2 (S_1 - A S_2) + \ldots$$

$$+ A^{n-1} (S_{n-2} - A S_{n-1}) + A^n S_{n-1}$$

$$= -A S_0 + A S_0 - A^2 S_2 + A^3 S_3 - A^4 S_4 + \ldots + A^{n-1} S_{n-2} - A^n S_{n-1} + A^n S_n,$$

$$= 0.$$ 

\[ \blacksquare \]

*We can decompose any matrix with entries polynomials in $\lambda$ into powers of $\lambda$: for instance, $(\lambda^2 + 2 \lambda + 1)(\lambda^2 + 1) = (0 \ 0) + (1 \ 0) \lambda + (1 \ 1) \lambda^2.$*
So, you may be wondering: exactly what does this have to do with computing $e^{tA}$, our topic from yesterday?

Suppose $A$ is $n \times n$, and has one eigenvalue $\lambda$ (with multiplicity $n$). The characteristic polynomial is $p_A(\lambda) = (\lambda - \lambda)^n$, so by Cayley-Hamilton $0 = p_A(A) = (A - \lambda I)^n \Rightarrow (A - \lambda I)^k = 0$ for $k \geq n$

$\Rightarrow e^{tA} = e^{\lambda t} I + t(A - \lambda I) \xrightarrow{\text{commute}} e^{\lambda t} I e^{t(A - \lambda I)}$

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

\[ \begin{align*}
\text{Ex/} & \quad \text{Suppose we want to solve} \\
& \begin{cases} 
\frac{d^2}{dt^2} f(t) - 6 \frac{d}{dt} f(t) + 12 f(t) - 8 f(t) = 0 \\
\end{cases} \\
& \begin{cases} 
\begin{align*} 
& f(0) = 1, \quad f'(0) = 0, \quad f''(0) = -1. 
\end{align*} \\
\end{cases} \\
& \text{Writing} \\
& \begin{cases} 
& x_1(t) = f(t), \\
& x_2(t) = f'(t), \\
& x_3(t) = f''(t) \\
& \frac{d}{dt} x_1(t) = x_2(t) \\
& \frac{d}{dt} x_2(t) = x_3(t) \\
& \frac{d}{dt} x_3(t) = 8 x_1(t) - 12 x_2(t) + 6 x_3(t) \\
& \begin{align*} 
& x(0) = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \\
\end{align*} \\
\end{cases} \\
\end{align*} \]
which is the same as
\[ \begin{cases} \dot{x}'(t) = Ax'(t) \\ \dot{x}(0) = \ddot{b} \end{cases} \]
where \( \ddot{b} = \left( \frac{1}{4} \right) \) and \( A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 8 & 12 & 6 \end{pmatrix} \). The characteristic polynomial is \( P_A(\lambda) = \det \begin{pmatrix} \lambda & -1 & 0 \\ 0 & \lambda & -1 \\ -8 & 12 & \lambda - 6 \end{pmatrix} = \lambda^3 - 6\lambda^2 + 12\lambda - 8 = (\lambda - 2)^3 \)

same as that of the original DE (1), as it must be.
(Can you prove that?) So the above calculation gives
\[ e^{\tau A} = e^{2\tau} \sum_{k=0}^{\infty} \frac{\tau^k}{k!} (A - 2I)^k \]
\[ = e^{2\tau} I + e^{2\tau} \begin{pmatrix} -2 & 1 & 0 \\ 0 & -2 & 1 \\ 8 & 12 & 4 \end{pmatrix} + \frac{e^{2\tau}}{2} \begin{pmatrix} 4 & -4 & 1 \\ 0 & -8 & 2 \\ 16 & -16 & 4 \end{pmatrix}. \]
The solution is \( \ddot{x}(t) = e^{\tau A} \ddot{b} = e^{2\tau} \left( \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} \right) + e^{2\tau} \left( \begin{pmatrix} -2 \\ 1 \\ 8 \end{pmatrix} \right) + \frac{e^{2\tau}}{2} \left( \begin{pmatrix} 3 \\ -4 \\ 12 \end{pmatrix} \right), \]
which is to say \( f(t) \) (\( = \ddot{x}(t) \)) = \( e^{2\tau} - 2e^{2\tau} + \frac{e^{2\tau}}{2} + e^{2\tau}. \)
(Do you understand why?)

Here is how to do the other "non-diagonalizable" possibility for 3x3 matrices: \( \lambda_1 = \mu, \lambda_2 = \lambda_3 = \gamma \). We have
\[ P_A(\lambda) = (\lambda - \mu)(\lambda - \gamma)^2 \Rightarrow 0 = (A - \mu I)(A - \gamma I)^2 \]
\[ \Rightarrow (A - \gamma I)^k = [(A - \mu I) + (\mu - \gamma) I]^{k-2} (A - \gamma I)^2 \]
\[ (A - \gamma I)^2 = (\mu - \gamma)^{k-2} (A - \gamma I)^2 \]
\[
\Rightarrow e^{tA} = e^{2tI + t(A-\lambda I)} = e^{2t} \sum_{k=0}^{\infty} \frac{t^k}{k!} (A-\lambda I)^k
\]

\[
= e^{2t}I + te^{2t}(A-\lambda I) + e^{2t} \sum_{k=2}^{\infty} \frac{t^k}{k!} (\lambda-\lambda)^{k-2}(A-\lambda I)^2
\]

\[
= e^{2t}I + te^{2t}(A-\lambda I) + \frac{e^{2t}}{(\lambda-\lambda)^2} \left\{ e^{(\lambda-\lambda)t} - 1 - (\lambda-\lambda)t \right\} (A-\lambda I)^2
\]

a formula which looks more complicated than it is.

Notice that the functions of \( t \) that appear in \( e^{tA} \), hence in any solution \( e^{tA} \mathbf{b} \) to \( \dot{x}(t) = A \dot{x}(t) \), are simply \( e^{2t} \), \( te^{2t} \), and \( e^{2t} \).