Appendix A. Exponential Matrix.

**Definition 1.** For $n \times n$–matrix $A$, the exponential matrix function
\[
e^{tA} = I + \frac{t}{1!}A + \frac{t^2}{2!}A^2 + \cdots = \sum_{k=0}^{\infty} \frac{t^k}{k!}A^k.
\]

**Lemma 1.** $X = X(t) = e^{tA}$ is the unique solution of the Cauchy problem
\[
X' = \frac{dX}{dt} = AX, \quad X(0) = I.
\]

**Proof.** It is easy to see that
\[
\frac{d}{dt} e^{tA} = A + \frac{t}{1!}A^2 + \cdots = Ae^{tA},
\]
so that $X = e^{tA}$ is a solution of (2). Moreover, we can treat $n \times n$–matrix function $X$ as a vector functions with values in $\mathbb{R}^{n^2}$ (or $\mathbb{C}^{n^2}$). We only need to rewrite the matrix equation $X' = AX$ in the vector form $X' = BX$ with a $n^2 \times n^2$–matrix $B$. Then the uniqueness for the Cauchy problem in the vector form implies the uniqueness for the problem (2).

**Definition 2.** If $AV = \lambda V$, for some vector $V \neq 0$, then $\lambda$ is an eigenvalue of $A$, and $V$ is an eigenvector corresponding to $\lambda$.

We have $AV = \lambda V \iff (A - \lambda I)V = 0$. The last system has nontrivial solutions $V \neq 0 \iff |A - \lambda I| = 0$. We introduce the characteristic polynomial of $A$ by the formula $p(\lambda) = |\lambda I - A|$. Now we can conclude:

(i) The eigenvalues of $A$ are roots of the characteristic equation
\[
p(\lambda) = |\lambda I - A| = 0.
\]

(ii) For each eigenvalue $\lambda$, the corresponding eigenvectors $V$ are nontrivial solutions of the system
\[
(A - \lambda I)V = 0.
\]

**Lemma 2.** Let $A$ be a constant $n \times n$ matrix, and let $AV = \lambda V$ for some vector $V$. Then the matrix function $U = U(t) = e^{\lambda t} V$ satisfies $U' = AU$.

**Proof.** $U'(t) = (e^{\lambda t})' V = e^{\lambda t} \lambda V = e^{\lambda t} AV = AU(t)$.

**Lemma 3.** For distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$ of the matrix $A$, the corresponding eigenvectors $V_1, V_2, \cdots, V_m$ are linearly independent.

**Proof.** For $m = 1$, this is trivial:
\[
V_1 \neq 0, \quad c_1 V_1 = 0 \quad \Rightarrow \quad c_1 = 0.
\]
Now suppose this statement is true for some $m = k$, i.e. $V_1, V_2, \cdots, V_k$ are linearly independent. We will prove that it remains true for $m = k + 1$, i.e. the equality
\[
c_1 V_1 + \cdots + c_k V_k + c_{k+1} V_{k+1} = 0
\]
\[A-1\]
holds only in case $c_1 = \ldots = c_k = c_{k+1} = 0$. Multiplying (5) by the matrix $A$ and using the equalities $AV_j = \lambda_j V_j$, we get

$$c_1 \lambda_1 V_1 + \cdots + c_k \lambda_k V_k + c_{k+1} \lambda_{k+1} V_{k+1} = 0.$$ 

Now subtract (5) multiplied by $\lambda_{k+1}$. This gives us

$$c_1 (\lambda_1 - \lambda_{k+1}) V_1 + \cdots + c_k (\lambda_k - \lambda_{k+1}) V_k = 0.$$ 

By our assumption, $V_1, V_2, \ldots, V_k$ are linearly independent. Therefore,

$$c_1 (\lambda_1 - \lambda_{k+1}) = \cdots = c_k (\lambda_k - \lambda_{k+1}) = 0.$$ 

Since all the eigenvalues are distinct, this implies $c_1 = \cdots = c_k = 0$. Now from (5) it follows $c_{k+1} = 0$. This proves our statement for $m = k + 1$, and by induction, it is true for arbitrary $m$. □

Another way of computation $e^{tA}$ is based on the following famous result.

**Theorem 1 (Cayley-Hamilton).** Let $A = (a_{ij})$ be a $n \times n$ matrix, and let $p(\lambda) = |\lambda I - A|$. Then $p(A) = 0$.

**Proof.** The coefficients of $p(\lambda)$ depend on the entries $a_{ij}$, and the entries of the resultant matrix $p(A)$ are polynomials of $a_{ij}$. The statement $p(A) = 0$ means these polynomials are reduced to 0 as a result of algebraic operations, which are based on same rules in both real and complex cases. Thus without loss of generality, we will assume that $A$ is a real matrix.

(i) Suppose the matrix $A$ has $n$ distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. By Lemma 3, the corresponding eigenvectors $V_1, V_2, \ldots, V_n$ are linearly independent and therefore they compose a basis in $\mathbb{R}^n$. We have

$$AV_i = \lambda_i V_i \implies A^k V_i = \lambda_i^k V_i \implies p(A)V_i = p(\lambda_i)V_i = 0$$

for any $i = 1, 2, \ldots, n$. An arbitrary vector $V \in \mathbb{R}^n$ can be represented in the form $V = \sum c_i V_i$. Therefore, we have

$$p(A)V = \sum c_i p(A)V_i = 0 \text{ for all } V \in \mathbb{R}^n \implies p(A) = 0.$$

(ii) For real $\varepsilon$, consider the characteristic polynomial of $\varepsilon A + B$,

$$p_\varepsilon(\lambda) = |\lambda I - \varepsilon A - B|,$$

where $B = \text{diag} \{1, 2, \ldots, n\}$.

Notice that $p_0(\lambda) = (\lambda - 1)(\lambda - 2) \cdots (\lambda - n)$ has different signs at the ends of each of $n$ segments $[k - \frac{1}{2}, k + \frac{1}{2}]$, $k = 1, 2, \ldots, n$. Since $p_\varepsilon(\lambda)$ is continuous with respect to $\varepsilon$, it satisfies the same property for $|\varepsilon| < \varepsilon_0$, provided $\varepsilon_0 > 0$ is small enough. For such $\varepsilon$, the polynomial $p_\varepsilon(\lambda)$ has a root in each of intervals $(k - \frac{1}{2}, k + \frac{1}{2})$, $k = 1, 2, \ldots, n$, thus all $n$ roots are distinct.

We have proved that the matrix $\varepsilon A + B$ have $n$ distinct eigenvalues if $|\varepsilon| < \varepsilon_0$. Multiplying this matrix by $t = \frac{1}{2}$, we obtain that the matrix $A + tB$ has $n$ distinct eigenvalues if $t > t_0 = 1/\varepsilon_0$. In other words, for fixed $t > t_0$, the polynomial $p(\lambda, t) = |\lambda I - A - tB|$ has $n$ distinct roots.

(iii) From (i) it follows $p(A + tB, t) = 0$ for all $t > t_0$. Since the entries of the matrix $p(A + tB, t)$ are polynomials with respect to $t$, we have $p(A + tB, t) = 0$ for all $t$. In particular, taking $t = 0$, we obtain the desired equality $p(A) = p(A, 0) = 0$. □
Let $\lambda_1, \lambda_2, \cdots, \lambda_m$ denote all the distinct eigenvalues of a matrix $A$ (real and complex). Then we can write the characteristic polynomial of $A$ in the form

$$p(\lambda) = |\lambda I - A| = \prod_{k=1}^{m} (\lambda - \lambda_k)^{r_k}, \quad (6)$$

where $r_k$ is the multiplicity of $\lambda_k$. We have

$$\frac{1}{p(\lambda)} = \sum_{k=1}^{m} \frac{r_k}{\prod_{j=k}^{m} (\lambda - \lambda_j)^{r_j}} = \sum_{k=1}^{m} \frac{a_k(\lambda)}{(\lambda - \lambda_k)^{r_k}}, \quad (7)$$

where $a_{kj} = \text{const}$, and $a_k(\lambda)$ is a polynomial of degree $\leq r_k - 1$ for each $k$. Therefore,

$$1 = \sum_{k=1}^{m} a_k(\lambda)p_k(\lambda), \quad I = \sum_{k=1}^{m} a_k(A)p_k(A), \quad (8)$$

where

$$p_k(\lambda) = \frac{p(\lambda)}{(\lambda - \lambda_k)^{r_k}} = \prod_{j\neq k}(\lambda - \lambda_j)^{r_j}. \quad (9)$$

**Theorem 2.** Let $\lambda_1, \lambda_2, \cdots, \lambda_m$ be all the distinct eigenvalues of a matrix $A$ with the characteristic polynomial $p(\lambda)$ in (6). Then

$$e^{tA} = \sum_{k=1}^{m} e^{\lambda_k t}a_k(A)p_k(A)\sum_{j=0}^{r_k-1} \frac{t^j}{j!} (A - \lambda_k I)^j, \quad (10)$$

where $a_k$ are polynomials of degree $\leq r_k - 1$, which are defined from the decomposition (7), and $p_k$ are polynomials in (9).

First we prove a lemma, which allows us to use the equalities

$$e^{tA} = e^{tM}e^{t(A - A)} = e^{M}e^{t(A - \lambda A)}. \quad (11)$$

**Lemma 4.** Let $A$ and $B$ be $n \times n$ matrices satisfying $AB = BA$. Then

$$e^{tA}e^{tB} \equiv e^{t(A+B)}.$$

**Proof of Lemma.** We have

$$AB = BA \implies A^kB = BA^k \implies e^{tA}B = Be^{tA}.$$  

Therefore, the matrix function $X(t) = e^{tA}e^{tB}$ satisfies

$$X' = (e^{tA})'e^{tB} + e^{tA}(e^{tB})' = Ae^{tA}e^{tB} + Be^{tA}e^{tB} = (A + B)X, \quad X(0) = I.$$  

By Lemma 1, we have $X(t) \equiv e^{t(A+B)}$. \hfill \Box
Proof of Theorem 2. Using the equalities (8) and (11) we get

\[ e^{tA} = \sum_{k=1}^{m} a_k(A)p_k(A)e^{tA} = \sum_{k=1}^{m} e^{\lambda_k t} a_k(A)p_k(A)e^{t(A-\lambda_k I)}. \]  

(12)

Since \( p_k(A)(A - \lambda_k I)^{r_k} = p(A) = 0, \)

\[ p_k(A)e^{t(A-\lambda_k I)} = p_k(A)\sum_{j=0}^{\infty} \frac{t^j}{j!}(A - \lambda_k I)^j = p_k(A)\sum_{j=0}^{r_k-1} \frac{t^j}{j!}(A - \lambda_k I)^j. \]

Together with (12), this gives us the equality (10).

Example 1. Consider the matrix \( A = \begin{pmatrix} -1 & 1 & -2 \\ 0 & -1 & 4 \\ 0 & 0 & 1 \end{pmatrix}. \) We have

\[ p(\lambda) = |\lambda I - A| = (\lambda + 1)^2(\lambda - 1) \implies \]

\( \lambda_1 = -1, \quad r_1 = 2, \quad \lambda_2 = 1, \quad r_2 = 1, \quad p_1(\lambda) = \lambda - 1, \quad p_2(\lambda) = (\lambda + 1)^2; \)

\[ \frac{1}{p(\lambda)} = \frac{1}{(\lambda + 1)^2(\lambda - 1)} = \frac{1}{4} \left( \frac{1}{\lambda - 1} - \frac{\lambda + 3}{(\lambda + 1)^2} \right) \implies a_1(\lambda) = -\frac{1}{4}(\lambda + 3), \quad a_2(\lambda) = \frac{1}{4}. \]

By (10), we have

\[ e^{tA} = -\frac{1}{4} e^{-t}(A + 3I)(A - I)[I + t(A + I)] + \frac{1}{4} e^{t}(A + I)^2. \]

Finally,

\[ \frac{1}{4}(A + I)^2 = \frac{1}{4} \begin{pmatrix} 0 & 1 & -2 \\ 0 & 0 & 4 \\ 0 & 0 & 2 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix} \implies \]

\[ -\frac{1}{4}(A + 3I)(A - I) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} \implies \]

\[ e^{tA} = e^{-t} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} + te^{-t} \begin{pmatrix} 0 & 1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + e^t \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix}. \]