Problem A.31

Functions of matrices are typically defined by their Taylor series expansions. For example,

$$e^{\mathsf{M}} \equiv \mathsf{I} + \mathsf{M} + \frac{1}{2}\mathsf{M}^2 + \frac{1}{3!}\mathsf{M}^3 + \cdots . \tag{A.99}$$

(a) Find $\exp(\mathsf{M})$, if

$$\text{(i)}\ \ \mathsf{M} = \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}; \quad \text{(ii)}\ \ \mathsf{M} = \begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix}.$$

(b) Show that if M is diagonalizable, then

$$\det\left(e^{\mathsf{M}}\right) = e^{\mathrm{Tr}(\mathsf{M})}. \tag{A.100}$$

Comment: This is actually true even if M is *not* diagonalizable, but it's harder to prove in the general case.

(c) Show that if the matrices M and N *commute*, then

$$e^{\mathsf{M}+\mathsf{N}} = e^{\mathsf{M}}e^{\mathsf{N}}. \tag{A.101}$$

Prove (with the simplest counterexample you can think up) that Equation A.101 is *not* true, in general, for *non*-commuting matrices.[21]

(d) If H is hermitian, show that $e^{i\mathsf{H}}$ is unitary.

Solution

Part (a)

Suppose that

$$\mathsf{M} = \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}.$$

By definition, the exponential of this matrix is

$$e^{\mathsf{M}} = \mathsf{I} + \mathsf{M} + \frac{1}{2}\mathsf{M}^2 + \frac{1}{3!}\mathsf{M}^3 + \cdots .$$

The aim, then, is to find the higher powers of M.

$$\mathsf{M}^2 = \mathsf{M}\mathsf{M} = \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\mathsf{M}^3 = \mathsf{M}^2\mathsf{M} = \begin{pmatrix} 0 & 0 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

[21] See Problem 3.29 for the more general "Baker–Campbell–Hausdorff" formula.

Because $\mathsf{M}^3 = \mathsf{0}$, all higher powers are equal to the zero matrix as well.

$$\mathsf{M}^4 = \mathsf{M}^3\mathsf{M} = \mathsf{0M} = \mathsf{0}$$

$$\mathsf{M}^5 = \mathsf{M}^4\mathsf{M} = \mathsf{0M} = \mathsf{0}$$

$$\vdots$$

$$\mathsf{M}^k = \mathsf{M}^{k-1}\mathsf{M} = \mathsf{0M} = \mathsf{0}, \quad k \geq 4$$

Therefore,

$$\begin{aligned} e^{\mathsf{M}} &= \mathsf{I} + \mathsf{M} + \frac{1}{2}\mathsf{M}^2 \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0 & 0 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 & 5 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

Suppose now that

$$\mathsf{M} = \begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix}.$$

By definition, the exponential of this matrix is

$$e^{\mathsf{M}} = \mathsf{I} + \mathsf{M} + \frac{1}{2}\mathsf{M}^2 + \frac{1}{3!}\mathsf{M}^3 + \cdots.$$

As before, the higher powers of M can be found by multiplying matrices [equation (1) below]. A more systematic approach will be taken here, though, in order to show how to exponentiate more complicated matrices. Consider the eigenvalue problem for M.

$$\mathsf{Ma} = \lambda\mathsf{a}$$

Bring $\lambda\mathsf{a}$ to the left side and factor a.

$$(\mathsf{M} - \lambda\mathsf{I})\mathsf{a} = \mathsf{0}$$

Since $\mathsf{a} \neq \mathsf{0}$, the matrix in parentheses must be singular, that is,

$$\det(\mathsf{M} - \lambda\mathsf{I}) = 0$$

$$\begin{vmatrix} -\lambda & \theta \\ -\theta & -\lambda \end{vmatrix} = 0$$

$$\lambda^2 + \theta^2 = 0.$$

This is the characteristic equation of M. According to the Cayley-Hamilton theorem, every square matrix satisfies its own characteristic equation.

$$\mathsf{M}^2 + \theta^2\mathsf{I} = \mathsf{0}$$

Solve for M^2.

$$\mathsf{M}^2 = -\theta^2\mathsf{I} \tag{1}$$

Use this result to construct the higher powers of M.

$$\mathsf{M}^4 = \mathsf{M}^2\mathsf{M}^2 = (-\theta^2\mathsf{I})(-\theta^2\mathsf{I}) = \theta^4\mathsf{I} \qquad \mathsf{M}^3 = \mathsf{M}^2\mathsf{M} = (-\theta^2\mathsf{I})(\mathsf{M}) = -\theta^2\mathsf{M}$$

$$\mathsf{M}^6 = \mathsf{M}^4\mathsf{M}^2 = (\theta^4\mathsf{I})(-\theta^2\mathsf{I}) = -\theta^6\mathsf{I} \qquad \mathsf{M}^5 = \mathsf{M}^2\mathsf{M}^3 = (-\theta^2\mathsf{I})(-\theta^2\mathsf{M}) = \theta^4\mathsf{M}$$

$$\mathsf{M}^8 = \mathsf{M}^6\mathsf{M}^2 = (-\theta^6\mathsf{I})(-\theta^2\mathsf{I}) = \theta^8\mathsf{I} \qquad \mathsf{M}^7 = \mathsf{M}^2\mathsf{M}^5 = (-\theta^2\mathsf{I})(\theta^4\mathsf{M}) = -\theta^6\mathsf{M}$$

$$\vdots \qquad\qquad \vdots$$

$$\mathsf{M}^{2k} = (-1)^k\theta^{2k}\mathsf{I} \qquad \mathsf{M}^{2k+1} = (-1)^k\theta^{2k}\mathsf{M}$$

Therefore,

$$\begin{aligned}
e^{\mathsf{M}} &= \mathsf{I} + \mathsf{M} + \frac{1}{2}\mathsf{M}^2 + \frac{1}{3!}\mathsf{M}^3 + \cdots \\
&= \sum_{k=0}^{\infty} \frac{\mathsf{M}^k}{k!} \\
&= \sum_{k=0}^{\infty} \frac{\mathsf{M}^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{\mathsf{M}^{2k+1}}{(2k+1)!} \\
&= \sum_{k=0}^{\infty} \frac{(-1)^k\theta^{2k}\mathsf{I}}{(2k)!} + \sum_{k=0}^{\infty} \frac{(-1)^k\theta^{2k}\mathsf{M}}{(2k+1)!} \\
&= \mathsf{I}\sum_{k=0}^{\infty} \frac{(-1)^k\theta^{2k}}{(2k)!} + \mathsf{M}\sum_{k=0}^{\infty} \frac{(-1)^k\theta^{2k}}{(2k+1)!} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \sum_{k=0}^{\infty} \frac{(-1)^k\theta^{2k}}{(2k)!} + \begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix} \sum_{k=0}^{\infty} \frac{(-1)^k\theta^{2k}}{(2k+1)!} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \sum_{k=0}^{\infty} \frac{(-1)^k\theta^{2k}}{(2k)!} + \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \sum_{k=0}^{\infty} \frac{(-1)^k\theta^{2k+1}}{(2k+1)!} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cos\theta + \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \sin\theta \\
&= \begin{pmatrix} \cos\theta & 0 \\ 0 & \cos\theta \end{pmatrix} + \begin{pmatrix} 0 & \sin\theta \\ -\sin\theta & 0 \end{pmatrix} \\
&= \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.
\end{aligned}$$

Part (b)

Assume that M is a diagonalizable $n \times n$ matrix. Then it has n linearly independent eigenvectors, which means a similarity matrix S can be constructed. The product SMS^{-1} is a matrix with the eigenvalues corresponding to the n eigenvectors along the main diagonal and zeros elsewhere.

$$\mathsf{SMS}^{-1} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \tag{2}$$

Premultiply both sides by S^{-1}.

$$\mathsf{S}^{-1}\mathsf{SMS}^{-1} = \mathsf{S}^{-1} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

Postmultiply both sides by S.

$$\mathsf{S}^{-1}\mathsf{SMS}^{-1}\mathsf{S} = \mathsf{S}^{-1} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \mathsf{S}$$

These last two steps isolate M on the left side, since matrix multiplication is associative and $\mathsf{S}^{-1}\mathsf{S} = \mathsf{I}$.

$$\mathsf{M} = \mathsf{S}^{-1} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \mathsf{S}$$

Exponentiate both sides.

$$e^{\mathsf{M}} = \exp \left[\mathsf{S}^{-1} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \mathsf{S} \right]$$

$$= \mathsf{I} + \left[\mathsf{S}^{-1} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \mathsf{S} \right] + \frac{1}{2} \left[\mathsf{S}^{-1} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \mathsf{S} \right] \left[\mathsf{S}^{-1} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \mathsf{S} \right] + \cdots$$

$$= \mathsf{I} + \mathsf{S}^{-1} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \mathsf{S} + \frac{1}{2}\mathsf{S}^{-1} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \mathsf{S} + \cdots$$

Multiply the diagonal matrices on the right side.

$$e^{\mathsf{M}} = \mathsf{I} + \mathsf{S}^{-1}\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}\mathsf{S} + \frac{1}{2}\mathsf{S}^{-1}\begin{pmatrix} \lambda_1^2 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^2 \end{pmatrix}\mathsf{S} + \frac{1}{3!}\mathsf{S}^{-1}\begin{pmatrix} \lambda_1^3 & 0 & \cdots & 0 \\ 0 & \lambda_2^3 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^3 \end{pmatrix}\mathsf{S} + \cdots$$

Now proceed to eliminate S from the equation. Premultiply both sides by S.

$$\begin{aligned} \mathsf{S}e^{\mathsf{M}} &= \mathsf{S}\mathsf{I} + \mathsf{S}\mathsf{S}^{-1}\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}\mathsf{S} + \frac{1}{2}\mathsf{S}\mathsf{S}^{-1}\begin{pmatrix} \lambda_1^2 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^2 \end{pmatrix}\mathsf{S} + \frac{1}{3!}\mathsf{S}\mathsf{S}^{-1}\begin{pmatrix} \lambda_1^3 & 0 & \cdots & 0 \\ 0 & \lambda_2^3 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^3 \end{pmatrix}\mathsf{S} + \cdots \\ &= \mathsf{S} + \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}\mathsf{S} + \frac{1}{2}\begin{pmatrix} \lambda_1^2 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^2 \end{pmatrix}\mathsf{S} + \frac{1}{3!}\begin{pmatrix} \lambda_1^3 & 0 & \cdots & 0 \\ 0 & \lambda_2^3 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^3 \end{pmatrix}\mathsf{S} + \cdots \end{aligned}$$

Postmultiply both sides by S^{-1}.

$$\begin{aligned} \mathsf{S}e^{\mathsf{M}}\mathsf{S}^{-1} &= \mathsf{S}\mathsf{S}^{-1} + \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}\mathsf{S}\mathsf{S}^{-1} + \frac{1}{2}\begin{pmatrix} \lambda_1^2 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^2 \end{pmatrix}\mathsf{S}\mathsf{S}^{-1} + \frac{1}{3!}\begin{pmatrix} \lambda_1^3 & 0 & \cdots & 0 \\ 0 & \lambda_2^3 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^3 \end{pmatrix}\mathsf{S}\mathsf{S}^{-1} + \cdots \\ &= \mathsf{I} + \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \lambda_1^2 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^2 \end{pmatrix} + \frac{1}{3!}\begin{pmatrix} \lambda_1^3 & 0 & \cdots & 0 \\ 0 & \lambda_2^3 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^3 \end{pmatrix} + \cdots \\ &= \begin{pmatrix} 1 + \lambda_1 + \frac{1}{2}\lambda_1^2 + \frac{1}{3!}\lambda_1^3 + \cdots & 0 & \cdots & 0 \\ 0 & 1 + \lambda_2 + \frac{1}{2}\lambda_2^2 + \frac{1}{3!}\lambda_2^3 + \cdots & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 + \lambda_n + \frac{1}{2}\lambda_n^2 + \frac{1}{3!}\lambda_n^3 + \cdots \end{pmatrix} \\ &= \begin{pmatrix} \displaystyle\sum_{k=0}^{\infty}\frac{\lambda_1^k}{k!} & 0 & \cdots & 0 \\ 0 & \displaystyle\sum_{k=0}^{\infty}\frac{\lambda_2^k}{k!} & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \displaystyle\sum_{k=0}^{\infty}\frac{\lambda_n^k}{k!} \end{pmatrix} \end{aligned}$$

As a result,

$$\mathsf{S}e^{\mathsf{M}}\mathsf{S}^{-1} = \begin{pmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n} \end{pmatrix}.$$

Take the determinant of both sides.

$$\det\left(\mathsf{S}e^{\mathsf{M}}\mathsf{S}^{-1}\right) = \det \begin{pmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n} \end{pmatrix}$$

Use the fact that the determinant of a product is the product of the determinants:
$\det\left(\mathsf{S}e^{\mathsf{M}}\mathsf{S}^{-1}\right) = \det(\mathsf{S})\det(e^{\mathsf{M}})\det\left(\mathsf{S}^{-1}\right) = \det(e^{\mathsf{M}})\det(\mathsf{S})\det\left(\mathsf{S}^{-1}\right) = \det(e^{\mathsf{M}})\det\left(\mathsf{S}\mathsf{S}^{-1}\right) = \det(e^{\mathsf{M}})\det(\mathsf{I}) = \det(e^{\mathsf{M}})(1) = \det(e^{\mathsf{M}})$.

$$\det(e^{\mathsf{M}}) = \begin{vmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n} \end{vmatrix}$$

Evaluate the determinant on the right side.

$$\begin{aligned} \det(e^{\mathsf{M}}) &= e^{\lambda_1}e^{\lambda_2}\cdots e^{\lambda_n} \\ &= e^{\lambda_1+\lambda_2+\cdots+\lambda_n} \end{aligned}$$

In the exponent is the trace of the matrix in equation (2).

$$\det(e^{\mathsf{M}}) = e^{\operatorname{Tr}\left(\mathsf{S}\mathsf{M}\mathsf{S}^{-1}\right)}$$

Use the result from Problem A.17 that says three matrices in a product can be cyclically permuted within a trace argument.

$$\begin{aligned} \det(e^{\mathsf{M}}) &= e^{\operatorname{Tr}\left(\mathsf{M}\mathsf{S}^{-1}\mathsf{S}\right)} \\ &= e^{\operatorname{Tr}(\mathsf{M}\mathsf{I})} \end{aligned}$$

Therefore,

$$\det(e^{\mathsf{M}}) = e^{\operatorname{Tr}(\mathsf{M})}.$$

Part (c)

Assume that M and N are $n \times n$ matrices and that they commute: $\mathsf{MN} = \mathsf{NM}$. By definition,

$$\begin{aligned} e^{\mathsf{M}+\mathsf{N}} &= \mathsf{I} + (\mathsf{M}+\mathsf{N}) + \frac{1}{2}(\mathsf{M}+\mathsf{N})^2 + \frac{1}{3!}(\mathsf{M}+\mathsf{N})^3 + \cdots \\ &= \sum_{k=0}^{\infty} \frac{(\mathsf{M}+\mathsf{N})^k}{k!}. \end{aligned} \tag{3}$$

Consider element ij of $(\mathsf{M}+\mathsf{N})^2$.

$$\begin{aligned} [(\mathsf{M}+\mathsf{N})^2]_{ij} &= \sum_{k=1}^{n} (\mathsf{M}+\mathsf{N})_{ik}(\mathsf{M}+\mathsf{N})_{kj} \\ &= \sum_{k=1}^{n} (M_{ik} + N_{ik})(M_{kj} + N_{kj}) \\ &= \sum_{k=1}^{n} (M_{ik}M_{kj} + M_{ik}N_{kj} + N_{ik}M_{kj} + N_{ik}N_{kj}) \\ &= \sum_{k=1}^{n} M_{ik}M_{kj} + \sum_{k=1}^{n} M_{ik}N_{kj} + \sum_{k=1}^{n} N_{ik}M_{kj} + \sum_{k=1}^{n} N_{ik}N_{kj} \\ &= (\mathsf{MM})_{ij} + (\mathsf{MN})_{ij} + (\mathsf{NM})_{ij} + (\mathsf{NN})_{ij} \\ &= (\mathsf{MM} + \mathsf{MN} + \mathsf{NM} + \mathsf{NN})_{ij} \\ &= (\mathsf{MM} + \mathsf{MN} + \mathsf{MN} + \mathsf{NN})_{ij} \\ &= (\mathsf{M}^2 + 2\mathsf{MN} + \mathsf{N}^2)_{ij} \end{aligned}$$

Consequently, $(\mathsf{M}+\mathsf{N})^2 = \mathsf{M}^2 + 2\mathsf{MN} + \mathsf{N}^2$.

Now consider element ij of $(\mathsf{M}+\mathsf{N})^3$.

$$\begin{aligned}
[(\mathsf{M}+\mathsf{N})^3]_{ij} &= [(\mathsf{M}+\mathsf{N})(\mathsf{M}+\mathsf{N})^2]_{ij} \\
&= [(\mathsf{M}+\mathsf{N})(\mathsf{M}^2+2\mathsf{MN}+\mathsf{N}^2)]_{ij} \\
&= \sum_{k=1}^{n}(\mathsf{M}+\mathsf{N})_{ik}(\mathsf{M}^2+2\mathsf{MN}+\mathsf{N}^2)_{kj} \\
&= \sum_{k=1}^{n}\sum_{l=1}^{n}(M_{ik}+N_{ik})(M_{kl}M_{lj}+2M_{kl}N_{lj}+N_{kl}N_{lj}) \\
&= \sum_{k=1}^{n}\sum_{l=1}^{n}(M_{ik}M_{kl}M_{lj}+2M_{ik}M_{kl}N_{lj}+M_{ik}N_{kl}N_{lj}+N_{ik}M_{kl}M_{lj}+2N_{ik}M_{kl}N_{lj}+N_{ik}N_{kl}N_{lj}) \\
&= \sum_{k=1}^{n}\sum_{l=1}^{n}M_{ik}M_{kl}M_{lj}+2\sum_{k=1}^{n}\sum_{l=1}^{n}M_{ik}M_{kl}N_{lj}+\sum_{k=1}^{n}\sum_{l=1}^{n}M_{ik}N_{kl}N_{lj} \\
&\qquad +\sum_{k=1}^{n}\sum_{l=1}^{n}N_{ik}M_{kl}M_{lj}+2\sum_{k=1}^{n}\sum_{l=1}^{n}N_{ik}M_{kl}N_{lj}+\sum_{k=1}^{n}\sum_{l=1}^{n}N_{ik}N_{kl}N_{lj} \\
&= (\mathsf{MMM})_{ij}+2(\mathsf{MMN})_{ij}+(\mathsf{MNN})_{ij}+(\mathsf{NMM})_{ij}+2(\mathsf{NMN})_{ij}+(\mathsf{NNN})_{ij} \\
&= (\mathsf{MMM}+2\mathsf{MMN}+\mathsf{MNN}+\mathsf{NMM}+2\mathsf{NMN}+\mathsf{NNN})_{ij} \\
&= [\mathsf{MMM}+2\mathsf{MMN}+\mathsf{MNN}+(\mathsf{NM})(\mathsf{M})+2(\mathsf{NM})(\mathsf{N})+\mathsf{NNN}]_{ij} \\
&= [\mathsf{MMM}+2\mathsf{MMN}+\mathsf{MNN}+(\mathsf{MN})(\mathsf{M})+2(\mathsf{MN})(\mathsf{N})+\mathsf{NNN}]_{ij} \\
&= [\mathsf{MMM}+2\mathsf{MMN}+\mathsf{MNN}+(\mathsf{M})(\mathsf{NM})+2\mathsf{MNN}+\mathsf{NNN}]_{ij} \\
&= [\mathsf{MMM}+2\mathsf{MMN}+\mathsf{MNN}+(\mathsf{M})(\mathsf{MN})+2\mathsf{MNN}+\mathsf{NNN}]_{ij} \\
&= (\mathsf{MMM}+2\mathsf{MMN}+\mathsf{MNN}+\mathsf{MMN}+2\mathsf{MNN}+\mathsf{NNN})_{ij} \\
&= (\mathsf{MMM}+3\mathsf{MMN}+3\mathsf{MNN}+\mathsf{NNN})_{ij} \\
&= (\mathsf{M}^3+3\mathsf{M}^2\mathsf{N}+3\mathsf{MN}^2+\mathsf{N}^3)_{ij}
\end{aligned}$$

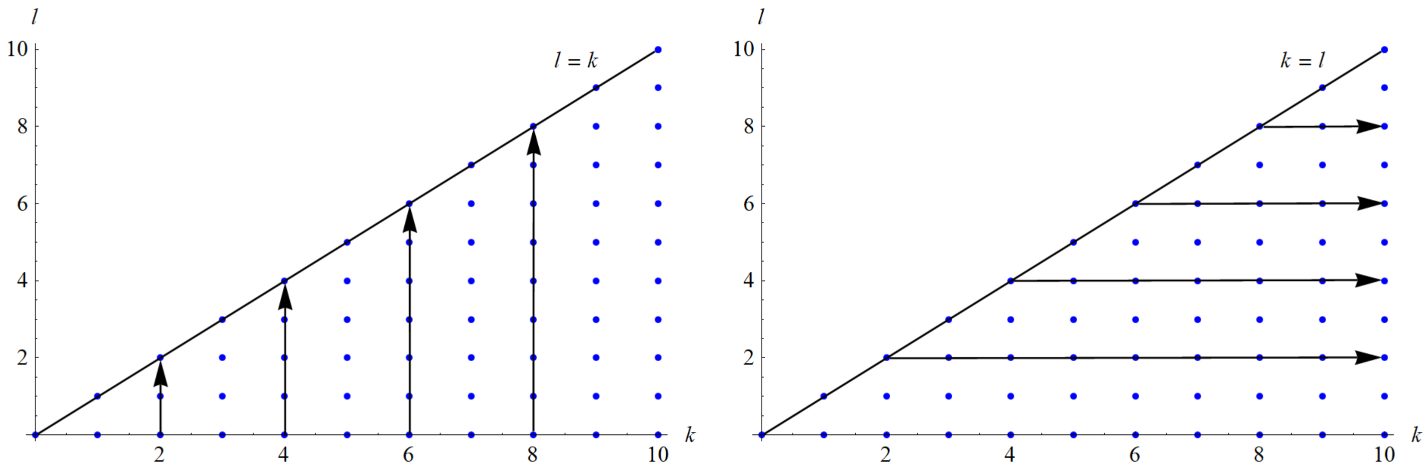
Consequently, $(\mathsf{M}+\mathsf{N})^3=\mathsf{M}^3+3\mathsf{M}^2\mathsf{N}+3\mathsf{MN}^2+\mathsf{N}^3$. The binomial theorem holds for M and N apparently because these matrices commute.

$$(\mathsf{M}+\mathsf{N})^k=\sum_{l=0}^{k}\frac{k!}{(k-l)!l!}\mathsf{M}^{k-l}\mathsf{N}^l$$

Equation (3) then becomes

$$\begin{aligned} e^{\mathsf{M}+\mathsf{N}} &= \sum_{k=0}^{\infty} \frac{(\mathsf{M}+\mathsf{N})^k}{k!} \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{l=0}^{k} \frac{k!}{(k-l)!l!} \mathsf{M}^{k-l}\mathsf{N}^l \\ &= \sum_{k=0}^{\infty} \sum_{l=0}^{k} \frac{\mathsf{M}^{k-l}\mathsf{N}^l}{(k-l)!l!}. \end{aligned}$$

The current mode of summation is shown below on the left. Each dot represents a term in the double sum.



To switch the order of summation, sum over the dots as shown on the right.

$$e^{\mathsf{M}+\mathsf{N}} = \sum_{l=0}^{\infty} \sum_{k=l}^{\infty} \frac{\mathsf{M}^{k-l}\mathsf{N}^l}{(k-l)!l!}$$

Bring l to the left side.

$$e^{\mathsf{M}+\mathsf{N}} = \sum_{l=0}^{\infty} \sum_{k-l=0}^{\infty} \frac{\mathsf{M}^{k-l}\mathsf{N}^l}{(k-l)!l!}$$

Make the substitution $p = k - l$.

$$e^{\mathsf{M}+\mathsf{N}} = \sum_{l=0}^{\infty} \sum_{p=0}^{\infty} \frac{\mathsf{M}^{p}\mathsf{N}^l}{p!l!}$$

Since there are no longer indices in the sums, the sums can be separated and ordered in any way.

$$\begin{aligned} e^{\mathsf{M}+\mathsf{N}} &= \left(\sum_{p=0}^{\infty} \frac{\mathsf{M}^p}{p!}\right)\left(\sum_{l=0}^{\infty} \frac{\mathsf{N}^l}{l!}\right) \\ &= (e^{\mathsf{M}})(e^{\mathsf{N}}) \end{aligned}$$

Therefore, provided that M and N are commuting matrices,

$$e^{\mathsf{M}+\mathsf{N}} = e^{\mathsf{M}}e^{\mathsf{N}}.$$

In order to show that this result is not true for non-commuting matrices, let

$$\mathsf{M} = \begin{pmatrix} 0 & 0 \\ -\theta & 0 \end{pmatrix} \quad \text{and} \quad \mathsf{N} = \begin{pmatrix} 0 & \theta \\ 0 & 0 \end{pmatrix} \quad \text{so that} \quad \mathsf{M} + \mathsf{N} = \begin{pmatrix} 0 & \theta \\ -\theta & 0 \end{pmatrix}.$$

The two matrix products are

$$\mathsf{MN} = \begin{pmatrix} 0 & 0 \\ -\theta & 0 \end{pmatrix}\begin{pmatrix} 0 & \theta \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & -\theta^2 \end{pmatrix} \quad \text{and} \quad \mathsf{NM} = \begin{pmatrix} 0 & \theta \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ -\theta & 0 \end{pmatrix} = \begin{pmatrix} -\theta^2 & 0 \\ 0 & 0 \end{pmatrix}.$$

$\mathsf{MN} \neq \mathsf{NM}$, so M and N do not commute. Since

$$\mathsf{M}^2 = \begin{pmatrix} 0 & 0 \\ -\theta & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ -\theta & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathsf{N}^2 = \begin{pmatrix} 0 & \theta \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & \theta \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

the higher powers of M and N vanish as well.

$$\begin{aligned} \mathsf{M}^3 &= \mathsf{M}^2\mathsf{M} = \mathsf{0M} = \mathsf{0} & \mathsf{N}^3 &= \mathsf{N}^2\mathsf{N} = \mathsf{0N} = \mathsf{0} \\ \mathsf{M}^4 &= \mathsf{M}^3\mathsf{M} = \mathsf{0M} = \mathsf{0} & \mathsf{N}^4 &= \mathsf{N}^3\mathsf{N} = \mathsf{0N} = \mathsf{0} \\ &\vdots & &\vdots \\ \mathsf{M}^k &= \mathsf{0}, \quad k \geq 2 & \mathsf{N}^k &= \mathsf{0}, \quad k \geq 2 \end{aligned}$$

And that means the respective exponential functions are

$$\begin{aligned} e^{\mathsf{M}} &= \mathsf{I} + \mathsf{M} + \frac{1}{2}\mathsf{M}^2 + \cdots & e^{\mathsf{N}} &= \mathsf{I} + \mathsf{N} + \frac{1}{2}\mathsf{N}^2 + \cdots \\ e^{\mathsf{M}} &= \mathsf{I} + \mathsf{M} & e^{\mathsf{N}} &= \mathsf{I} + \mathsf{N} \\ e^{\mathsf{M}} &= \begin{pmatrix} 1 & 0 \\ -\theta & 1 \end{pmatrix} & e^{\mathsf{N}} &= \begin{pmatrix} 1 & \theta \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Therefore, for these non-commuting matrices, M and N,

$$e^{\mathsf{M}+\mathsf{N}} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \neq \begin{pmatrix} 1 & \theta \\ -\theta & 1-\theta^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\theta & 1 \end{pmatrix}\begin{pmatrix} 1 & \theta \\ 0 & 1 \end{pmatrix} = e^{\mathsf{M}}e^{\mathsf{N}}.$$

Part (d)

Assume that H is a hermitian matrix: $\mathsf{H}^\dagger = \mathsf{H}$. The aim here is to show that $e^{i\mathsf{H}}$ is unitary, that is,

$$\left(e^{i\mathsf{H}}\right)^\dagger = \left(e^{i\mathsf{H}}\right)^{-1}.$$

Consider the exponential function of $i\mathsf{H}$.

$$\begin{aligned} e^{i\mathsf{H}} &= \mathsf{I} + (i\mathsf{H}) + \frac{1}{2}(i\mathsf{H})^2 + \frac{1}{3!}(i\mathsf{H})^3 + \cdots \\ &= \mathsf{I} + (i\mathsf{H}) + \frac{1}{2}(i\mathsf{H})(i\mathsf{H}) + \frac{1}{3!}(i\mathsf{H})(i\mathsf{H})(i\mathsf{H}) + \cdots \end{aligned}$$

Take the complex conjugate of both sides.

$$\begin{aligned} \left(e^{i\mathsf{H}}\right)^* &= \left[\mathsf{I} + (i\mathsf{H}) + \frac{1}{2}(i\mathsf{H})(i\mathsf{H}) + \frac{1}{3!}(i\mathsf{H})(i\mathsf{H})(i\mathsf{H}) + \cdots\right]^* \\ &= \mathsf{I}^* + (i\mathsf{H})^* + \left[\frac{1}{2}(i\mathsf{H})(i\mathsf{H})\right]^* + \left[\frac{1}{3!}(i\mathsf{H})(i\mathsf{H})(i\mathsf{H})\right]^* + \cdots \\ &= \mathsf{I} + (i\mathsf{H})^* + \frac{1}{2}(i\mathsf{H})^*(i\mathsf{H})^* + \frac{1}{3!}(i\mathsf{H})^*(i\mathsf{H})^*(i\mathsf{H})^* + \cdots \end{aligned}$$

Take the transpose of both sides.

$$\begin{aligned} \left[\left(e^{i\mathsf{H}}\right)^*\right]^T &= \left[\mathsf{I} + (i\mathsf{H})^* + \frac{1}{2}(i\mathsf{H})^*(i\mathsf{H})^* + \frac{1}{3!}(i\mathsf{H})^*(i\mathsf{H})^*(i\mathsf{H})^* + \cdots\right]^T \\ &= \mathsf{I}^T + [(i\mathsf{H})^*]^T + \left[\frac{1}{2}(i\mathsf{H})^*(i\mathsf{H})^*\right]^T + \left[\frac{1}{3!}(i\mathsf{H})^*(i\mathsf{H})^*(i\mathsf{H})^*\right]^T + \cdots \\ &= \mathsf{I} + [(i\mathsf{H})^*]^T + \frac{1}{2}\left[(i\mathsf{H})^*\right]^T\left[(i\mathsf{H})^*\right]^T + \frac{1}{3!}\left[(i\mathsf{H})^*\right]^T\left[(i\mathsf{H})^*\right]^T\left[(i\mathsf{H})^*\right]^T + \cdots \end{aligned}$$

Therefore,

$$\begin{aligned} \left(e^{i\mathsf{H}}\right)^\dagger &= \mathsf{I} + (i\mathsf{H})^\dagger + \frac{1}{2}(i\mathsf{H})^\dagger(i\mathsf{H})^\dagger + \frac{1}{3!}(i\mathsf{H})^\dagger(i\mathsf{H})^\dagger(i\mathsf{H})^\dagger + \cdots \\ &= \mathsf{I} + (i\mathsf{H})^\dagger + \frac{1}{2}\left[(i\mathsf{H})^\dagger\right]^2 + \frac{1}{3!}\left[(i\mathsf{H})^\dagger\right]^3 + \cdots \\ &= e^{(i\mathsf{H})^\dagger} \\ &= e^{-i\mathsf{H}^\dagger} \\ &= e^{-i\mathsf{H}} \\ &= \left(e^{i\mathsf{H}}\right)^{-1}, \end{aligned}$$

which means $e^{i\mathsf{H}}$ is unitary.