



Analytical eigenvalue methods for matrix function evaluation

Charbel Tannous

► To cite this version:

Charbel Tannous. Analytical eigenvalue methods for matrix function evaluation. Licence. Mécanique Quantique 2, UBO Brest, France. 2018, pp.11. <hal-03996134v2>

HAL Id: hal-03996134

<https://hal.science/hal-03996134v2>

Submitted on 19 Feb 2023


HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Analytical eigenvalue methods for matrix function evaluation

C. Tannous 

Université de Brest, Lab-STICC, CNRS-UMR 6285, F-29200 Brest, FRANCE

(Dated: [February 14, 2023](#))

Evaluating analytically diagonalizable matrix functions in the case of distinct or repeated eigenvalues is described. While matrix functions include inversion, powers, exponentiation, trigonometric, hyperbolic... we illustrate various methods (Infinite series recursion, Similarity transformation, Cayley-Hamilton method, Spectral Decomposition...) by evaluating, in detail, several types of matrices along with several functions and compare their respective performances. The diagonalizable and non-diagonalizable matrix cases are also discussed.

PACS numbers: 02.10.Yn, 02.10.Ud, 03.65.Fd

Keywords: Matrix theory, Linear Algebra, Algebraic methods in quantum mechanics

Contents

I. Introduction	1
A. Cayley-Hamilton theorem	2
1. Le Verrier Algorithm	3
2. Souriau-Faddeev Algorithm	4
B. Spectral decomposition and Sylvester formula	4
II. Comparative matrix function evaluation	4
A. First case: Distinct eigenvalues	5
1. Analytic continuation	5
2. Infinite Series	5
3. Cayley-Hamilton	5
4. Spectral decomposition	6
5. Landau-Lifshitz method	6
6. Similarity transformation	6
B. Second case: Repeated eigenvalues	7
1. Cayley-Hamilton	7
2. Spectral decomposition	8
C. Third case: Non-diagonalizable matrix	9
1. Cayley-Hamilton	9
2. Spectral decomposition	9
A. Polynomial interpolation	10
1. Lagrange interpolation	10
2. Hermite interpolation	10
3. Extension to matrices	11
References	11

I. INTRODUCTION

Matrix functions are used in many areas of Science and Technology. For example matrix exponentiation serves in the integration of linear systems of ODE (Ordinary Differential Equations), feedback and control systems, circuit theory...

In many areas of Classical and Quantum physics such as Mechanics, Dynamical Systems, Electromagnetism, Electronics, Neural Networks, Field Theory, Quantum mechanics, Scattering theory, Quantum Circuits, Condensed Matter physics... where time evolution or special symmetry operations such as spatial translation and rotation are needed, matrix function evaluations are required.

Mathematically, a matrix function $f(M)$ where M is a square matrix, is defined as:

$$f(M) = \sum_{i=0}^{\infty} a_i M^i \quad (1)$$

where the coefficients [1, 2] depend on the function f .

Surprisingly, we evaluate the inverse of a function with a power expansion instead of performing the standard process involving matrix of cofactors [3], adjoint and determinant evaluation.

Other matrix functions include exponentiation, trigonometric, hyperbolic... Generally, we illustrate matrix function evaluation with:

1. The exponential given by: $f(M) = \exp(M) = \sum_{i=0}^{\infty} a_i M^i$ with $a_i = \frac{1}{i!}$ by extending the exponential series coefficients to those pertaining to the matrix function.
2. Trigonometric functions such as: $\cos M = \sum_{k=0}^{\infty} (\pm) \frac{M^{2k}}{(2k)!}$ and $\sin M = \sum_{k=0}^{\infty} (\pm) \frac{M^{2k+1}}{(2k+1)!}$
3. Hyperbolic functions such as: $\cosh M = \sum_{k=0}^{\infty} \frac{M^{2k}}{(2k)!}$ and $\sinh M = \sum_{k=0}^{\infty} \frac{M^{2k+1}}{(2k+1)!}$

using Cayley-Hamilton theorem, Spectral Decomposition as well as direct methods such as Infinite Series recursion, Similarity transformation and compare their performance.

A. Cayley-Hamilton theorem

Cayley-Hamilton theorem states that for a square ($n \times n$) matrix M , the characteristic polynomial equation is satisfied by the matrix itself. It is valid whether the matrix is diagonalizable or not.

The characteristic polynomial is expressed with Descartes expansion theorem [3] in the distinct case as:

$$P(\lambda) = \det(\lambda \mathbb{1} - M) = (\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_n) = \lambda^n + q_1 \lambda^{n-1} + q_2 \lambda^{n-2} \dots + q_n \quad (2)$$

where:

$$q_1 = - \sum_{i=1}^n \lambda_i = -\text{Tr}(M), \quad q_2 = \sum_{\substack{i,j \\ i \neq j}}^n \lambda_i \lambda_j, \quad q_3 = - \sum_{\substack{i,j,k \\ i \neq j \neq k}}^n \lambda_i \lambda_j \lambda_k, \dots \quad q_n = (-)^n \lambda_1 \lambda_2 \dots \lambda_n = (-)^n \det(M) \quad (3)$$

The characteristic polynomial equation is satisfied by M according to Cayley-Hamilton theorem:

$$P(M) = M^n + q_1 M^{n-1} + \dots + q_n \mathbb{1} = 0 \quad (4)$$

thus $M^n = -q_1 M^{n-1} - \dots - q_n \mathbb{1}$.

Consequently, the inverse of M is calculated as: $M^{-1} = -\frac{1}{q_n} (M^{n-1} + q_1 M^{n-2} - \dots + q_{n-1} \mathbb{1})$.

Moreover, any arbitrary power M^m of M with $m \geq n$ can be expressed with powers of M not exceeding $n-1$ according to Cayley-Hamilton theorem allowing the evaluation of the coefficients b_i to obtain: $M^m = \sum_{i=0}^{n-1} b_i M^i$.

Consequently, any function [1, 2] can be written as: $f(M) = \sum_{i=0}^{n-1} b_i M^i$ where the b_i are evaluated from the function coefficients a_i and the characteristic polynomial.

In order to evaluate $b_0, b_1 \dots b_n$, we apply $f(M)$ to the eigenvectors, using Dirac notation [4, 5], $|\lambda_1\rangle, |\lambda_2\rangle, \dots, |\lambda_n\rangle$ corresponding to eigenvalues $\lambda_1, \lambda_2 \dots \lambda_n$. Starting from $M |\lambda_i\rangle = \lambda_i |\lambda_i\rangle$, we get for an arbitrary power m , the result $M^m |\lambda_i\rangle = \lambda_i^m |\lambda_i\rangle$ and consequently $f(M) |\lambda_j\rangle = \sum_{i=0}^{n-1} b_i M^i |\lambda_j\rangle = \sum_{i=0}^{n-1} b_i \lambda_j^i |\lambda_j\rangle$.

Thus we obtain:

$$\begin{aligned}
 f(M) |\lambda_1\rangle &= f(\lambda_1) |\lambda_1\rangle = \sum_{i=0}^{n-1} b_i \lambda_1^i |\lambda_1\rangle \\
 f(M) |\lambda_2\rangle &= f(\lambda_2) |\lambda_2\rangle = \sum_{i=0}^{n-1} b_i \lambda_2^i |\lambda_2\rangle \\
 &\vdots \\
 f(M) |\lambda_n\rangle &= f(\lambda_n) |\lambda_n\rangle = \sum_{i=0}^{n-1} b_i \lambda_n^i |\lambda_n\rangle
 \end{aligned} \tag{5}$$

This results in the following $(n \times n)$ linear (with respect to b_i coefficients) system:

$$\begin{aligned}
 f(\lambda_1) &= \sum_{i=0}^{n-1} b_i \lambda_1^i \\
 f(\lambda_2) &= \sum_{i=0}^{n-1} b_i \lambda_2^i \\
 &\vdots \\
 f(\lambda_n) &= \sum_{i=0}^{n-1} b_i \lambda_n^i
 \end{aligned} \tag{6}$$

whose solution yields the n sought coefficients $b_i, 0 = 1 \dots n - 1$. In the repeated eigenvalue case, one has $k < n$ distinct eigenvalues while the Cayley-Hamilton theorem, in order to evaluate a function requires n terms, thus the system is under-determined.

In order to retrieve the required missing equations, let us suppose one eigenvalue, say λ_i has multiplicity m_i , then we are able to restore the missing number of relations by expressing the equivalence of all derivatives of the function $f(\lambda)$ and its polynomial equivalent $P_f(\lambda) = \sum_{k=0}^{n-1} b_k \lambda^k$.

This originates from dividing the function $f(\lambda)$ by its characteristic polynomial $P(\lambda)$ to obtain $Q(\lambda)P(\lambda) + R(\lambda)$ where the rest of the division $R(\lambda) = P_f(\lambda)$.

When λ is taken as any eigenvalue λ_i we get the relation: $f(\lambda_i) = Q(\lambda_i)P(\lambda_i) + P_f(\lambda_i)$ yielding $f(\lambda_i) = P_f(\lambda_i)$ since the characteristic polynomial $P(\lambda_i) = 0$. Accounting for the fact all derivatives of the characteristic polynomial whose order is $< m_i$ are zero [3], we obtain the relations:

$$\left[\frac{d^l f(\lambda)}{d\lambda^l} \right]_{\lambda_i} = \left[\frac{d^l P_f(\lambda)}{d\lambda^l} \right]_{\lambda_i}, \quad l = 0, 1 \dots m_i - 1 \tag{7}$$

Thus we get m_i relations and the system relating the b_i coefficients is no longer under-determined.

A better possibility is to turn to Spectral decomposition, as explained further below, since it is faster than Cayley-Hamilton and calculates functions with exactly $k < n$ required terms.

1. Le Verrier Algorithm

One straightforward method for the determination of the characteristic polynomial coefficients is Le Verrier algorithm:

1. Calculate all powers M^r for $r = 2 \dots n$
2. Calculate all traces $s_r = \text{Tr} M^r$ for $r = 1 \dots n$
3. Starting from $q_0 = 1$, do $q_i = -\frac{1}{i} \sum_{j=1}^i q_{i-j} s_j$

2. Souriau-Faddeev Algorithm

In Le Verrier algorithm, the power of the matrix might lead to large values leading to conditioning and divergence whereas the Souriau-Faddeev method introduces recursively matrices M_i, B_i containing subtractions in order to avoid these problems along these lines:

For $i = 1, \dots, n$, do $M_1 = M, q_1 = \text{Tr}(M), B_1 = M_1 - q_1 \mathbb{1}, M_2 = MB_1, q_2 = \frac{1}{2} \text{Tr}(M_2), B_2 = M_2 - q_2 \mathbb{1}, \dots$

$M_i = MB_{i-1}, q_i = \frac{1}{i} \text{Tr}(M_i), B_i = M_i - q_i \mathbb{1} \dots$

For $i = n$ $B_n = M_n - q_n \mathbb{1} = 0$ and the $q_i, i = 1, \dots, n$ are the characteristic polynomial coefficients [3].

While the algorithm cures large values occurrence, it still might eventually lead to large errors.

B. Spectral decomposition and Sylvester formula

Spectral decomposition is a procedure allowing writing any non-singular $n \times n$ square M , as:

$M = \sum_{j=1}^k \lambda_j P_j$ with projection operators P_j operating toward eigen sub-spaces pertaining to eigenvalues $\lambda_j, j = 1 \dots k$ where $k \leq n$ is the number of eigenvalues.

A projection operator P_j attached to eigenvalue λ_j transforms any vector to one belonging to the eigenspace spanned by λ_j (see Appendix). Their properties are:

1. $P_j^2 = P_j$ Idempotency
2. $P_i P_j = \delta_{ij}$ Orthogonality when $i \neq j$.
3. $\sum_{i=1}^k P_i = \mathbb{1}$ Closure relation

Sylvester's formula expresses an analytic function [1, 2] $f(M)$ of a matrix M as a polynomial in M on the basis of Cayley-Hamilton theorem, in terms of the eigenvalues and eigenvectors [3] of M and the projection operators with M having a number of distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_k$ where $k \leq n$ allowing presence of repeated values.

Using λ_i the eigenvalues of M , the projection operators are (see Appendix):

$$P_j \equiv \prod_{\substack{i=1 \\ i \neq j}}^k \frac{(M - \lambda_i I)}{\lambda_j - \lambda_i}$$

built with Lagrange interpolation [3] formula extended to matrices.

Matrix M is rewritten in terms of the projectors as:

1. $M = \sum_{j=1}^k \lambda_j P_j$,
2. Its inverse is given by: $M^{-1} = \sum_{j=1}^k \lambda_j^{-1} P_j$,
3. Any arbitrary power m of M is given by $M^m = \sum_{j=1}^k \lambda_j^m P_j$
4. Any function $f(M)$ of M writes as: $f(M) = \sum_{j=1}^k f(\lambda_j) P_j$.

II. COMPARATIVE MATRIX FUNCTION EVALUATION

For a given function defined with its initial a_i coefficients, the b_i coefficients are evaluated.

For instance, we illustrate the function evaluation with the exponential of a 2×2 matrix possessing distinct eigenvalues comparing the results with several methods.

Later we work out with spectral decomposition the case of a 4×4 matrix possessing repeated eigenvalues.

A. First case: Distinct eigenvalues

Let us first consider the complex exponential $\exp(iM)$ of the matrix $M = \boldsymbol{\sigma} \cdot \mathbf{a}$ where $\boldsymbol{\sigma}$ is a vector composed of Pauli matrices $\sigma_x, \sigma_y, \sigma_z$ whereas \mathbf{a} is a simple 3D vector.

Matrices $\sigma_x, \sigma_y, \sigma_z$ pertaining to spin (1/2) are given by :

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

1. Analytic continuation

In order to evaluate $\exp(iM)$ we perform an analytic continuation of Euler formula: $\exp(ix) = \cos x + i \sin x$ from the scalar to the matrix M case. Thus $\exp(iM) = \cos M + i \sin M$.

The cosine series is given by: $\cos x = \sum_{k=0}^{\infty} (\pm) \frac{x^{2k}}{(2k)!}$ and the sine series by: $\sin x = \sum_{k=0}^{\infty} (\pm) \frac{x^{2k+1}}{(2k+1)!}$, the analytic continuation is given by: $\cos M = \sum_{k=0}^{\infty} (\pm) \frac{M^{2k}}{(2k)!}$ whereas $\sin M = \sum_{k=0}^{\infty} (\pm) \frac{M^{2k+1}}{(2k+1)!}$.

Evaluation of M powers yields: $M^2 = a^2 \mathbb{1}$, $M^3 = a^2 M$, $M^4 = a^4 \mathbb{1}$, $M^5 = a^4 M$...

Consequently, any integer power M^m of M is given by $a^m \mathbb{1}$ for m even and $a^{m-1} M$ for m odd. Thus $\cos M$ is proportional to $\mathbb{1}$ whereas $\sin M$ is proportional to M .

Substituting M powers in the analytical continuation we obtain: $\cos M = (\cos a) \mathbb{1}$ and $\sin M = (\frac{\sin a}{a}) M$.

2. Infinite Series

We start by performing direct recursive calculation of $M = \boldsymbol{\sigma} \cdot \mathbf{a}$, $M^2 = a^2 \mathbb{1}$ with $a = \sqrt{a_x^2 + a_y^2 + a_z^2}$. This entails evaluating successive powers of M such as: $M^3 = a^2 M$, $M^4 = a^4 \mathbb{1}$...

Any integer power M^m of M is given by $a^m \mathbb{1}$ for m even and $a^{m-1} M$ for m odd.

After decomposing the series of $\exp(iM) = \sum_{m=0}^{\infty} \frac{(iM)^m}{m!}$ into even powers (cosine) multiplying $\mathbb{1}$ and odd (sinus) multiplying M , we obtain: $\exp(i\boldsymbol{\sigma} \cdot \mathbf{a}) = (\cos a) \mathbb{1} + i(\frac{\sin a}{a}) \boldsymbol{\sigma} \cdot \mathbf{a}$.

3. Cayley-Hamilton

Cayley-Hamilton theorem states that for a square ($n \times n$) matrix M , the characteristic polynomial is satisfied by the matrix itself and any function $f(M)$ is expressed as $f(M) = \sum_{i=0}^{n-1} b_i M^i$ meaning that the function $f(M)$ expansion can go as far as $n - 1$ and the coefficients b_i depend on both f and M .

Let us apply these notions to the evaluation of $f(M) = \exp(iM)$ where $M = \boldsymbol{\sigma} \cdot \mathbf{a}$.

The characteristic polynomial of matrix $M = \boldsymbol{\sigma} \cdot \mathbf{a}$ is: $\lambda^2 - a^2 = 0$ with $a = \sqrt{a_x^2 + a_y^2 + a_z^2}$. Cayley-Hamilton theorem gives: $M^2 = a^2 \mathbb{1}$ since the characteristic polynomial is $\lambda^2 - a^2 = 0$ whose eigenvalues are $\lambda_1 = -a, \lambda_2 = a$. Since matrix M is (2×2) ($n=2$), any function $f(M)$ is expressed as $b_0 + b_1 M$, meaning that the function $f(M)$ expansion can go up to $n - 1 = 1$.

In order to evaluate b_0, b_1 , we apply $f(M)$ to the eigenvectors, using Dirac notation [4, 5], $|\lambda_1\rangle, |\lambda_2\rangle$ corresponding to eigenvalues $\lambda_1 = -a, \lambda_2 = a$.

Thus we get:

$$\begin{aligned} \exp(i\boldsymbol{\sigma} \cdot \mathbf{a}) |\lambda_1\rangle &= \exp(i\lambda_1) |\lambda_1\rangle = (b_0 + b_1 \lambda_1) |\lambda_1\rangle \\ \exp(i\boldsymbol{\sigma} \cdot \mathbf{a}) |\lambda_2\rangle &= \exp(i\lambda_2) |\lambda_2\rangle = (b_0 + b_1 \lambda_2) |\lambda_2\rangle \end{aligned} \quad (8)$$

This results in the following system:

$$\begin{aligned} \exp(-ia) &= b_0 - b_1 a, \\ \exp(ia) &= b_0 + b_1 a \end{aligned} \quad (9)$$

Taking the sum and difference of both equations, we get: $b_0 = \cos a, b_1 = i(\frac{\sin a}{a})$.

Finally the result writes: $\exp(i\boldsymbol{\sigma} \cdot \mathbf{a}) = (\cos a)\mathbb{1} + i(\frac{\sin a}{a})\boldsymbol{\sigma} \cdot \mathbf{a}$.

4. Spectral decomposition

Given any non-singular $n \times n$ square M , spectral decomposition is such that:

$M = \sum_{j=1}^k \lambda_j P_j$ with projection operators P_j operating toward eigen sub-spaces pertaining to eigenvalues $\lambda_j, j = 1 \dots k$ where $k \leq n$ is the number of eigenvalues.

Sylvester formula yields: $f(M) = \sum_{j=1}^k f(\lambda_j) P_j$ where projection operators are defined with Lagrange interpolation formula (see Appendix).

Matrix $M = \boldsymbol{\sigma} \cdot \mathbf{a}$ has distinct eigenvalues $\lambda_1 = -a, \lambda_2 = a$ with $a = \sqrt{a_x^2 + a_y^2 + a_z^2}$.

Thus $k = 2$, $M = \lambda_1 P_1 + \lambda_2 P_2$ and $P_1 = \frac{(M - \lambda_2 \mathbb{1})}{(\lambda_1 - \lambda_2)} = -\frac{(M - a\mathbb{1})}{2a}$. In addition, closure formula yields: $P_2 = \mathbb{1} - P_1$.

In order to evaluate $\exp(iM)$, the corresponding scalar function is: $f(x) = \exp(ix)$, hence:

$$\exp(iM) = e^{-ia} P_1 + e^{ia} (1 - P_1) = -e^{-ia} \frac{(M - a\mathbb{1})}{2a} + e^{ia} (1 + \frac{(M - a\mathbb{1})}{2a}) = \frac{1}{2}(e^{ia} + e^{-ia})\mathbb{1} + \frac{(e^{ia} - e^{-ia})}{2a} M$$

Using Euler identities: $\cos a = \frac{1}{2}(e^{ia} + e^{-ia})$, $\sin a = \frac{1}{2i}(e^{ia} - e^{-ia})$, we finally get:
 $\exp(i\boldsymbol{\sigma} \cdot \mathbf{a}) = (\cos a)\mathbb{1} + i(\frac{\sin a}{a})\boldsymbol{\sigma} \cdot \mathbf{a}$.

5. Landau-Lifshitz method

For any function $\boldsymbol{\sigma} \cdot \mathbf{a} + \mathbf{b}$, we may write: $f(\boldsymbol{\sigma} \cdot \mathbf{a} + \mathbf{b}) = \boldsymbol{\sigma} \cdot \mathbf{A} + B\mathbb{1}$. Taking z axis along vector \mathbf{a} direction, the eigenvalues of $\boldsymbol{\sigma} \cdot \mathbf{a} + \mathbf{b}$ are $b \pm a$ and consequently the corresponding values of the operator $f(\boldsymbol{\sigma} \cdot \mathbf{a} + \mathbf{b})$ will be $f(b \pm a)$. Thus we get: $\mathbf{A} = \frac{a}{2a}(f(b+a) - f(b-a))$ and $B = \frac{1}{2}(f(b+a) + f(b-a))$. In this case, we have $b = 0$ and consequently: $\exp(i\boldsymbol{\sigma} \cdot \mathbf{a}) = (\cos a)\mathbb{1} + i(\frac{\sin a}{a})\boldsymbol{\sigma} \cdot \mathbf{a}$.

We conclude that Landau-Lifshitz method uses implicitly the Cayley-Hamilton theorem while exploiting as well the analytical continuation.

6. Similarity transformation

This is the most lengthy method.

$M = \boldsymbol{\sigma} \cdot \mathbf{a}$ is diagonalized such that $D = S^{-1}MS$ (thus $M = SDS^{-1}$) with S the transformation matrix from the initial basis to the M eigenbasis M built with eigenvectors $|\lambda_1\rangle, |\lambda_2\rangle$ corresponding to eigenvalues λ_1, λ_2 . M has eigenvalues $\lambda_1 = -a, \lambda_2 = a$ (distinct case) with $a = \sqrt{a_x^2 + a_y^2 + a_z^2}$ the modulus of vector \mathbf{a} whose components $a_x, a_y, a_z \in \mathcal{R}$.

Any function of M writes as: $f(M) = f(SDS^{-1}) = S[f(D)]S^{-1}$.

Therefore we get $\exp(iM) = S \exp(iD) S^{-1} = S \begin{bmatrix} e^{-ia} & 0 \\ 0 & e^{ia} \end{bmatrix} S^{-1}$.

The eigenvectors $|\lambda_1\rangle = \begin{bmatrix} 1 \\ -\frac{(a_x + ia_y)(a + a_z)}{a_x^2 + a_y^2} \end{bmatrix} = \begin{bmatrix} 1 \\ A \end{bmatrix}$, $|\lambda_2\rangle = \begin{bmatrix} 1 \\ \frac{(a_x + ia_y)(a - a_z)}{a_x^2 + a_y^2} \end{bmatrix} = \begin{bmatrix} 1 \\ B \end{bmatrix}$ yield:

$$S = \begin{bmatrix} 1 & 1 \\ A & B \end{bmatrix} \quad S^{-1} = \frac{1}{B - A} \begin{bmatrix} B & -1 \\ -A & 1 \end{bmatrix} \quad (10)$$

with parameters $A = -\frac{(a_x + ia_y)(a + a_z)}{a_x^2 + a_y^2}$, $B = \frac{(a_x + ia_y)(a - a_z)}{a_x^2 + a_y^2}$.

Thus:

$$\begin{aligned} \exp(iM) &= \begin{bmatrix} 1 & 1 \\ A & B \end{bmatrix} \begin{bmatrix} e^{-ia} & 0 \\ 0 & e^{ia} \end{bmatrix} \frac{1}{B-A} \begin{bmatrix} B & -1 \\ -A & 1 \end{bmatrix} = \frac{1}{B-A} \begin{bmatrix} 1 & 1 \\ A & B \end{bmatrix} \begin{bmatrix} e^{-ia} & 0 \\ 0 & e^{ia} \end{bmatrix} \begin{bmatrix} B & -1 \\ -A & 1 \end{bmatrix} \\ &= \frac{1}{B-A} \begin{bmatrix} -Ae^{ia} + Be^{-ia} & 2i \sin(a) \\ -2iAB \sin(a) & -Ae^{-ia} + Be^{ia} \end{bmatrix} = \begin{bmatrix} -\frac{A}{B-A}e^{ia} + \frac{B}{B-A}e^{-ia} & 2i \sin(a) \frac{1}{B-A} \\ -2i \frac{AB}{B-A} \sin(a) & -\frac{A}{B-A}e^{-ia} + \frac{B}{B-A}e^{ia} \end{bmatrix}. \end{aligned} \quad (11)$$

Using identities:

$$\begin{aligned} B-A &= \frac{2a(a_x + ia_y)}{a_x^2 + a_y^2}, \frac{B+A}{B-A} = -\frac{a_z}{a}, \frac{2a}{B-A} = a_x - ia_y, \frac{AB}{B-A} = -\frac{a_x + ia_y}{2a}, \\ \frac{1}{B-A} &= \frac{a_x - ia_y}{2a} \frac{A}{B-A} = -\frac{a + a_z}{2a}, \frac{B}{B-A} = \frac{a - a_z}{2a} \end{aligned} \quad (12)$$

we infer:

$$\exp(iM) = \begin{bmatrix} \frac{a+a_z}{2a}e^{ia} + \frac{a-a_z}{2a}e^{-ia} & 2i \sin(a) \frac{a_x - ia_y}{2a} \\ 2i \sin(a) \frac{a_x + ia_y}{2a} & \frac{a+a_z}{2a}e^{-ia} - \frac{a-a_z}{2a}e^{ia} \end{bmatrix} = \begin{bmatrix} \cos a + ia_z \left(\frac{\sin a}{a}\right) & i(a_x - ia_y) \left(\frac{\sin a}{a}\right) \\ i(a_x + ia_y) \left(\frac{\sin a}{a}\right) & \cos a - ia_z \left(\frac{\sin a}{a}\right) \end{bmatrix}.$$

$$\text{Since } \boldsymbol{\sigma} \cdot \mathbf{a} = \begin{bmatrix} a_z & a_x - ia_y \\ a_x + ia_y & -a_z \end{bmatrix}$$

$$\text{and } i \left(\frac{\sin a}{a}\right) \boldsymbol{\sigma} \cdot \mathbf{a} = \begin{bmatrix} ia_z \left(\frac{\sin a}{a}\right) & i(a_x - ia_y) \left(\frac{\sin a}{a}\right) \\ i(a_x + ia_y) \left(\frac{\sin a}{a}\right) & -ia_z \left(\frac{\sin a}{a}\right) \end{bmatrix},$$

we finally get: $\exp(i\boldsymbol{\sigma} \cdot \mathbf{a}) = (\cos a)\mathbb{1} + i \left(\frac{\sin a}{a}\right) \boldsymbol{\sigma} \cdot \mathbf{a}.$

B. Second case: Repeated eigenvalues

We consider the binary 4×4 matrix with repeated eigenvalues: $M = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$

whose characteristic polynomial is: $\det(\lambda \mathbb{1} - M) = \lambda^4 - 4\lambda^3 + 16\lambda - 16$ that can be factored as: $(\lambda + 2)(\lambda - 2)^3$.

Thus we have two eigenvalues: $\lambda_1 = -2$ (simple) and $\lambda_2 = 2$ (triple) and the number of distinct eigenvalues is $k = 2$.

1. Cayley-Hamilton

The characteristic polynomial coefficients that can be evaluated directly or from Le Verrier algorithm are: $q_1 = -4, q_2 = 0, q_3 = 16, q_4 = -16$.

Consequently, the inverse of M is calculated as: $M^{-1} = -\frac{1}{q_n}(M^{n-1} + q_1 M^{n-2} + q_2 M^{n-3} + q_3 M^{n-4} + \dots q_{n-1} \mathbb{1})$,

thus we have: $M^{-1} = \frac{1}{16}(M^3 - 4M^2 + 16\mathbb{1})$ which yields: $M^{-1} = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$ agreeing with the direct

inversion and the Spectral decomposition below.

In order to calculate a matrix function such as the exponential, we rely on the system of relations 7 using $\lambda_1 = -2$ (simple) and $\lambda_2 = 2$ (triple). Using $P_f(\lambda) = b_0 + b_1\lambda + b_2\lambda^2 + b_3\lambda^3$ and its first and second derivatives yields the system:

$$\begin{aligned} e^{\lambda_1} &= b_0 + b_1\lambda_1 + b_2\lambda_1^2 + b_3\lambda_1^3 \\ e^{\lambda_2} &= b_0 + b_1\lambda_2 + b_2\lambda_2^2 + b_3\lambda_2^3 \\ e^{\lambda_2} &= b_1 + 2b_2\lambda_2 + 3b_3\lambda_2^2 \\ e^{\lambda_2} &= 2b_2 + 6b_3\lambda_2 \end{aligned} \quad (13)$$

After solving for the $b_i, i = 0..3$ coefficients we get:

$$\text{we get } \exp(M) = b_0 + b_1 M + b_2 M^2 + b_3 M^3 = \frac{1}{4} \begin{bmatrix} 3e^2 + e^{-2} & e^2 - e^{-2} & e^2 - e^{-2} & e^2 - e^{-2} \\ e^2 - e^{-2} & 3e^2 + e^{-2} & -e^2 + e^{-2} & -e^2 + e^{-2} \\ e^2 - e^{-2} & -e^2 + e^{-2} & 3e^2 + e^{-2} & -e^2 + e^{-2} \\ e^2 - e^{-2} & -e^2 + e^{-2} & -e^2 + e^{-2} & 3e^2 + e^{-2} \end{bmatrix}$$

Another way is to rely on spectral decomposition that requires only $k < n$ while Cayley-Hamilton requires all n terms.

2. Spectral decomposition

Given any non-singular $n \times n$ square M , spectral decomposition is such that:

$M = \sum_{j=1}^k \lambda_j P_j$ with projection operators P_j operating toward eigen sub-spaces pertaining to eigenvalues $\lambda_j, j = 1...k$ where $k \leq n$ is the number of eigenvalues.

Sylvester formula yields: $f(M) = \sum_{j=1}^k f(\lambda_j) P_j$ where projection operators are defined with Lagrange interpolation formula.

Thus $k = 2$ and the Lagrange projectors are:

$$P_1 = \frac{(M - \lambda_2 \mathbb{1})}{(\lambda_1 - \lambda_2)} = \frac{1}{4} \begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \end{bmatrix} \text{ and } P_2 = \mathbb{1} - P_1 = \frac{(M - \lambda_1 \mathbb{1})}{(\lambda_2 - \lambda_1)} = \frac{1}{4} \begin{bmatrix} 3 & 1 & 1 & 1 \\ 1 & 3 & -1 & -1 \\ 1 & -1 & 3 & -1 \\ 1 & -1 & -1 & 3 \end{bmatrix}.$$

These results can be checked by evaluating the projector from the definition $|\lambda_1\rangle \langle \lambda_1|$ where λ_1 is the normalized eigenvector corresponding to eigenvalue λ_1 given by: $\lambda_1 = \begin{bmatrix} 1/2 \\ -1/2 \\ -1/2 \\ -1/2 \end{bmatrix}$.

$$\text{Performing the operation } |\lambda_1\rangle \langle \lambda_1| \text{ as } \begin{bmatrix} 1/2 \\ -1/2 \\ -1/2 \\ -1/2 \end{bmatrix} [1/2 \quad -1/2 \quad -1/2 \quad -1/2] \text{ we recover } P_1 = \frac{1}{4} \begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \end{bmatrix}$$

Same with projector P_2 by applying Meyer [3] rule, in the degenerate case, by building the matrix $[X|0]$ where X contains the three eigenvectors of M corresponding to eigenvalue λ_2 and building matrix $[X|Y]$ where Y contains the eigenvector of M corresponding to eigenvalue λ_1 .

$$\text{Thus } [X|0] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & -1 & -1 & 0 \end{bmatrix} \text{ whereas } [X|Y] = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 1 & -1 & -1 & -1 \end{bmatrix}$$

$$\text{We recover } P_2 \text{ from the operation : } P_2 = [X|0][X|Y]^{-1} = \frac{(M - \lambda_1 \mathbb{1})}{(\lambda_2 - \lambda_1)} = \frac{1}{4} \begin{bmatrix} 3 & 1 & 1 & 1 \\ 1 & 3 & -1 & -1 \\ 1 & -1 & 3 & -1 \\ 1 & -1 & -1 & 3 \end{bmatrix}.$$

Note that in the Meyer rule, one could use non-normalized eigenvectors as above, but in the distinct case, eigenvectors should be normalized since the projector should be idempotent as shown below. Supposing a distinct eigenvalue λ_j whose projector $P_j = |\lambda_j\rangle \langle \lambda_j|$, then taking its square $P_j^2 = |\lambda_j\rangle \langle \lambda_j| |\lambda_j\rangle \langle \lambda_j| = |\lambda_j\rangle \langle \lambda_j|$ since $\langle \lambda_j | \lambda_j \rangle = 1$.

Consequently the matrix $M = \lambda_1 P_1 + \lambda_2 P_2$ and its inverse may be checked by direct inversion of M that is:

$$M^{-1} = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \text{ agreeing with } M^{-1} = \lambda_1^{-1} P_1 + \lambda_2^{-1} P_2.$$

In addition, we may write the following functions: $M^m = \lambda_1^m P_1 + \lambda_2^m P_2$ and $f(M) = f(\lambda_1) P_1 + f(\lambda_2) P_2$.

$$\text{For example, we get } \exp(M) = e^{-2} P_1 + e^2 P_2 = \frac{1}{4} \begin{bmatrix} 3e^2 + e^{-2} & e^2 - e^{-2} & e^2 - e^{-2} & e^2 - e^{-2} \\ e^2 - e^{-2} & 3e^2 + e^{-2} & -e^2 + e^{-2} & -e^2 + e^{-2} \\ e^2 - e^{-2} & -e^2 + e^{-2} & 3e^2 + e^{-2} & -e^2 + e^{-2} \\ e^2 - e^{-2} & -e^2 + e^{-2} & -e^2 + e^{-2} & 3e^2 + e^{-2} \end{bmatrix}$$

agreeing with the Cayley-Hamilton result.

C. Third case: Non-diagonalizable matrix

1. Cayley-Hamilton

We consider an example consisting of a 3×3 matrix:

$$M = \begin{pmatrix} 6 & 2 & 8 \\ -2 & 2 & -2 \\ 0 & 0 & 2 \end{pmatrix} \quad (14)$$

The characteristic polynomial is $\lambda^3 - 10\lambda^2 + 32\lambda - 32 = 0$ that is factored as: $(\lambda - 2)(\lambda - 4)^2 = 0$. Hence, the eigenvalues are: $\lambda_1 = 2$ (single), $\lambda_2 = 4$ (double).

M is non-diagonalizable since the normalized eigenvectors are: $|\lambda_1\rangle = (1/\sqrt{6}, 2/\sqrt{6}, -1/\sqrt{6})$, $|\lambda_2\rangle = (1/\sqrt{2}, -1/\sqrt{2}, 0)$, and $\lambda_2 = 4$ being double, requires two eigenvectors while only one is found. Therefore the number of linearly independent eigenvectors is smaller than the dimension of M . Consequently, the eigenvectors alone cannot form a basis in the matrix vector space in order to diagonalize it with a similarity transformation.

Using Cayley-Hamilton, the inverse of M is calculated as:

$$M^{-1} = -\frac{1}{q_n}(M^{n-1} + q_1 M^{n-2} + q_2 M^{n-3} + q_3 M^{n-4} + \dots q_{n-1} \mathbb{1}), \quad (15)$$

thus we have: $n = 3, q_1 = -10, q_2 = 32, q_3 = -32$, thus $M^{-1} = \frac{1}{32}(M^2 - 10M + 32\mathbb{1})$ yielding:

$$M^{-1} = \frac{1}{8} \begin{bmatrix} 1 & -1 & -5 \\ 1 & 3 & -1 \\ 0 & 0 & 4 \end{bmatrix} \quad (16)$$

agreeing with the cofactor method.

In order to evaluate other functions, we turn to spectral decomposition.

2. Spectral decomposition

In this case, spectral decomposition requires Hermite interpolation (see Appendix). Hermite interpolation should be used to compute any function of M with:

$$f(M) = f(\lambda_1)G_1 + f(\lambda_2)G_2 + f'(\lambda_2)(M - \lambda_2\mathbb{1})G_2 \quad (17)$$

where the first derivative of the function f is required since eigenvalue λ_2 is doubly degenerate.

In order to determine the generalized projection operators $G_i, i = 1, 2$ we apply the above to two different functions $f(x) = 1, f(x) = (x - 4)^2$ to obtain successively:

$$\begin{aligned} f(x) = 1, f'(x) = 0, f(2) = f(4) = 1 &\rightarrow f(M) = \mathbb{1} = G_1 + G_2 \\ f(x) = (x - 4)^2, f'(x) = 2(x - 4), f(2) = 2, f(4) = 0 &\rightarrow f(M) = (M - 4\mathbb{1})^2 = 4G_1 \end{aligned} \quad (18)$$

Solving the system: $G_1 + G_2 = \mathbb{1}, 4G_1 = (M - 4\mathbb{1})^2$ we extract: $G_1 = (M - 4\mathbb{1})^2/4, G_2 = \mathbb{1} - G_1$ confirming the closure relation.

Applying formula 17 to the inverse of M using the function $f(x) = 1/x, f'(x) = -1/x^2$, we get :

$$M^{-1} = \lambda_1^{-1}G_1 + \lambda_2^{-1}G_2 - \lambda_2^{-2}(M - \lambda_2\mathbb{1})G_2 \quad (19)$$

to obtain the result:

$$M^{-1} = (1/2)G_1 + (1/4)G_2 - (1/16)(M - 4\mathbb{1})G_2 = \frac{1}{8} \begin{bmatrix} 1 & -1 & -5 \\ 1 & 3 & -1 \\ 0 & 0 & 4 \end{bmatrix} \quad (20)$$

agreeing with the cofactor method and Cayley-Hamilton.

In addition, we can write the exponential as:

$$\exp(M) = \exp(\lambda_1)G_1 + \exp(\lambda_2)G_2 + \exp(\lambda_2)(M - \lambda_2\mathbb{1})G_2 \quad (21)$$

to obtain:

$$\exp(M) = \begin{pmatrix} 3e^4 & 2e^4 & 7e^4 - e^2 \\ -2e^4 & -e^4 & -4e^4 - 2e^2 \\ 0 & 0 & e^2 \end{pmatrix} \quad (22)$$

agreeing with direct series expansion.

APPENDIX A: POLYNOMIAL INTERPOLATION

1. Lagrange interpolation

Given a number of nodes $x_i, i = 1 \dots n$ it is possible to build a polynomial function going through all these distinct points, with Lagrange interpolation [3] formula:

$$p_j(x) \equiv \prod_{\substack{i=1 \\ i \neq j}}^k \frac{(x - x_i)}{(x_j - x_i)} \quad (A1)$$

such that $p_j(x_i) = \delta_{ij}$.

2. Hermite interpolation

Hermite method requires the function and its derivatives $f_i^{(l)}$ at the given points $x_i, i = 0, \dots, m$ with $x_0 < x_1 < \dots < x_m$ and $k = 0 \dots n_i - 1$ such that the interpolating polynomial writes: $p(x) = \sum_{i=0}^m \sum_{k=0}^{n_i-1} f_i^{(k)} L_{ik}(x)$ where $L_{ik}(x)$ are the generalized Lagrange polynomials and defined from the auxiliary polynomials:

$$\ell_{ik}(x) \equiv \frac{(x - x_i)^k}{k!} \prod_{\substack{j=0 \\ j \neq i}}^m \frac{(x - x_j)}{x_i - x_j}, \quad 0 \leq i \leq m, \quad 0 \leq k \leq n_i \quad (A2)$$

to get recursively, starting from $L_{i,n_i-1}(x) = \ell_{i,n_i-1}, i = 0 \dots m$, and: $L_{ik}(x) = \ell_{ik}(x) - \sum_{\nu=k+1}^{n_i-1} \ell_{ik}^{(\nu)}(x) L_{i\nu}(x)$, $k = n_i - 2, n_i - 3, \dots, 0$ obeying the constraints:

$$p(x_i) = f_i, \quad p^{(1)}(x_i) = f_i^{(1)}, \quad p^{(1)}(x_i) = f_i^{(1)}, \dots, p^{(m)}(x_i) = f_i^{(m)}, \quad i = 0, \dots, k \quad (A3)$$

where the upper index $(i), i = 0 \dots k$ indicates derivation order.

3. Extension to matrices

Extending Lagrange interpolation to the matrix case entails introducing projection operators defined from the eigenvectors using Dirac notation [4, 5] as $P_j = |\lambda_j\rangle \langle \lambda_j|$ where $|\lambda_j\rangle$ is the eigenvector corresponding to the eigenvalue λ_j .

Consequently, for any vector \mathbf{V} expressed in the eigenvalue basis as $\mathbf{V} = \sum_{i=1}^k v_i |\lambda_i\rangle$ we get the action of the projector P_j on it as: $P_j \mathbf{V} = \sum_{i=1}^k v_i |\lambda_j\rangle \langle \lambda_j | \lambda_i \rangle = \sum_{i=1}^k v_i |\lambda_i\rangle \delta_{ij} = v_j |\lambda_j\rangle$.

Thus the projector P_j transforms any vector into one along $|\lambda_j\rangle$ direction.

When λ_i the eigenvalues of M , are all distinct, the corresponding projection operators are:

$$P_j \equiv \prod_{\substack{i=1 \\ i \neq j}}^k \frac{(M - \lambda_i I)}{\lambda_j - \lambda_i} \quad (\text{A4})$$

In the degenerate case, assuming that m_j is the multiplicity of eigenvalue λ_j , then the projector is [5]:

$$P_j = \sum_{i=1}^{m_j} \left| \lambda_j^{(i)} \right\rangle \left\langle \lambda_j^{(i)} \right| \quad (\text{A5})$$

where $\left| \lambda_j^{(i)} \right\rangle, i = 1 \dots m_j$ are the different eigenvectors spanning the eigenspace corresponding to eigenvalue λ_j .

Then $M = \sum_{i=1}^k P_i \lambda_i$ and consequently: $f(M) = \sum_{i=1}^k P_i f(\lambda_i)$

Another method exists when $k = 2$ according to Meyer [3]. It consists of building the matrix $[X|0]$ where X contains the three eigenvectors of M corresponding to repeated eigenvalue λ_2 and building matrix $[X|Y]$ where Y contains the eigenvector of M corresponding to single eigenvalue λ_1 . The projector is obtained from the operation : $[X|0][X|Y]^{-1}$.

Sylvester's formula applies for any diagonalizable matrix M with k distinct eigenvalues, $\lambda_1, \lambda_2, \dots, \lambda_k$ and any function f defined on some subset of the complex numbers such that $f(M)$ is well defined. This means that every single eigenvalue λ_i and that every repeated eigenvalue λ_i with multiplicity $m_i > 1$ belong to the function f domain and f itself is $(m_i - 1)$ times differentiable [3] at λ_i .

In the non-diagonalizable case, interpolation may be done with Hermite method [6] that requires (in sharp contrast to Lagrange interpolation requiring only function values) the function and its derivatives. Thus Sylvester's method can be extended to the matrix case on the basis of Hermite interpolation polynomials [3, 7] such that:

$$f(M) = \sum_{i=1}^k \sum_{j=0}^{m_i-1} \frac{f^{(j)}(\lambda_i)}{j!} (M - \lambda_i I)^j G_j, \quad (\text{A6})$$

where G_j are generalized projection operators.

-
- [1] I. S. Gradshteyn and I. M. Ryzhik *Table of Integrals, Series and Products*, Fifth Edition, edited by Allan Jeffrey (Academic Press) (1994).
 - [2] M. Abramowitz and I.S Stegun, *Handbook of Mathematical Tables*, Dover, New-York (1960).
 - [3] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, Philadelphia (2000).
 - [4] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, Vol. 3, Pergamon, Oxford (1965).
 - [5] E. Merzbacher, *Quantum Mechanics*, Third edition, Wiley, New-York (1998).
 - [6] J. Stoer R. and Bulirsch, *Introduction to Numerical Analysis*, Second Edition, Springer-Verlag, (New York, 1992).
 - [7] N. J. Higham, *Functions of matrices: theory and computation*, SIAM, Philadelphia (2008).