

# Revealing the Hidden Mathematical Beauties of the Cayley-Hamilton Method

Haiduke Sarafian

The Pennsylvania State University, University College, York, PA, USA

Email: has2@psu.edu

**How to cite this paper:** Sarafian, H. (2024) Revealing the Hidden Mathematical Beauties of the Cayley-Hamilton Method. *American Journal of Computational Mathematics*, 14, 257-263.  
<https://doi.org/10.4236/ajcm.2024.142010>

**Received:** May 15, 2024

**Accepted:** June 23, 2024

**Published:** June 26, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

The inversion of a non-singular square matrix applying a Computer Algebra System (CAS) is straightforward. The CASs make the *numeric* computation efficient but mock the mathematical characteristics. The algorithms conducive to the output are sealed and inaccessible. In practice, other than the CPU timing, the applied inversion method is irrelevant. This research-oriented article discusses one such process, the Cayley-Hamilton (C.H.) [1]. Pursuing the process *symbolically* reveals its unpublished hidden mathematical characteristics even in the original article [1]. This article expands the general vision of the original named method without altering its practical applications. We have used the famous CAS *Mathematica* [2]. We have briefed the theory behind the method and applied it to different-sized *symbolic* and numeric matrices. The results are compared to the named CAS's sealed, packaged library commands. The codes are given, and the algorithms are unsealed.

## Keywords

Cayley-Hamilton Method, Matrix Inversion, Linear Algebra, Computer Algebra System, *Mathematica*

## 1. Introduction

In science and engineering, one routinely encounters the eigenvector and eigenvalue equations. For instance, in physics, e.g., in quantum mechanics, the wave function,  $\psi$  being an eigenvector of a stationary state, is subject to the Hamiltonian-based approach conducive to the Schrodinger equation [3] [4] [5],

$$\hat{H}\psi = E\psi, \quad (1)$$

where  $\hat{H}$  is the Hamiltonian operator corresponding to the eigenvalue energy,  $E$ . Alternatively, if one envisions it to be represented by a squared matrix, the

Heisenberg approach [3] [4] [5], according to mathematical notations [6] in general (1), is written as,

$$\hat{A}\bar{y} = \lambda\bar{y} \quad (2)$$

Here,  $\bar{y}$  is the eigenvector and the,  $\lambda$  its corresponding eigenvalue. Arranging (2) in the linear algebra format [6], (2) yields,

$$(\hat{A} - \lambda)\bar{y} = 0 \quad (3)$$

The solution to (3) yields the nonzero eigenvectors—values that require criteria.

$$\det(\hat{A} - \lambda\mathbf{I}) = 0.$$

where  $\det$  is the determinant, and  $\mathbf{I}$  is an identity matrix. The solution of this determinant equation leads to the needed,  $\lambda$  and the,  $\bar{y}$ . This is a soft, short idea behind the eigenvector-value concept. In practice, following these steps, one gets the needed information.

Cayley and Hamilton [1] (C-H) have taken another step, directed to a different objective: inversion of the matrix, within the issue's framework.

Since the determinant equation yields a  $\lambda$  dependent polynomial equation, their approach replaced the  $\lambda$ 's in the latter with the original matrix! *i.e.*  $\lambda \xrightarrow{\text{by}} \text{matrix}$ . Consequently, the  $\lambda$  dependent polynomial equation becomes a matrix polynomial equation. Obviously, in doing so, both polynomials sustain their corresponding polynomial coefficients. A bit of manipulation leads to the *inverted* matrix. In practice, most exercises are *numeric* [7]-[13] and work successfully. However, this numeric method misses my point, which is the objective of this article.

In many words, because numeric matrices are used, there is no reason to pay attention to the “source-origin” of these coefficients. Rightfully, this is because of the objective and the goal of the search. Hence, one may think, as is the current trend of thought that these coefficients are “structureless-patternless,” but we have shown otherwise.

Our conclusion hinges on a *symbolic*, not numeric, approach, which we owe to CAS's symbolic capabilities [2].

The layout of this work has four sections. In addition to Sect. 1, Introduction, Sect. 2, is Formulation. It is composed of three subsections that *symbolically* detail the main idea. In Sect. 3, Examples, via one numeric example, the formulation of Section 2 is in action. The last Sect., Conclusion are the highlights of the achievements.

## 2. Formulation

### 2.1. Example: A 3 × 3 Matrix

We are transitioning from the  $\lambda$  dependent polynomial to its associated matrix polynomial calls for replacing  $\lambda \rightarrow \mathbf{matrix}$ . This operation doesn't alter the coefficients of the former or the latter polynomials. In practice, numeric matrices aim at the product, *i.e.*, calculating the eigenvectors' values, so one ignores the

details! source of these values. However, we reveal a few hidden facts by symbolically pursuing the same idea by manipulating these coefficients. Specifically, we consider manageably sized matrices, e.g., a  $3 \times 3$  and a  $4 \times 4$ , adopting a CAS, e.g., *Mathematica* [2], where we show the details.

Consider a  $3 \times 3$  matrix,  $B$ . The needed elements of the determinant equation, the equation on the line below (3) is,

```
B=Table[bi,j,{i,1,3},{j,1,3}];
λBI=DiagonalMatrix[λ{1,1,1}];
(B-λBI)//MatrixForm;
Det[(B-λBI)];
polyλB=Collect[Det[(B-λBI)],λ]==0
```

$$-\lambda^3 - b_{1,3}b_{2,2}b_{3,1} + b_{1,2}b_{2,3}b_{3,1} + b_{1,3}b_{2,1}b_{3,2} - b_{1,1}b_{2,3}b_{3,2} - b_{1,2}b_{2,1}b_{3,3} + b_{1,1}b_{2,2}b_{3,3} + \lambda^2 (b_{1,1} + b_{2,2} + b_{3,3}) + \lambda (b_{1,2}b_{2,1} - b_{1,1}b_{2,2} + b_{1,3}b_{3,1} + b_{2,3}b_{3,2} - b_{1,1}b_{3,3} - b_{2,2}b_{3,3}) = 0 \quad (4)$$

Equation (4) is a *complete* third-order descending monotonic polynomial with coefficients that appear to be structureless-patternless. A careful analysis has proven otherwise. First, following the C-H approach, the  $\lambda$ 's in (4) is replaced with matrices, *i.e.*,  $\lambda \xrightarrow{\text{by}} B$ . Then, by manipulating the coefficients of (4), it yields (5),

$$-B^3 + \text{Tr}(B)B^2 - \left\{ \det \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} + \det \begin{pmatrix} b_{22} & b_{23} \\ b_{32} & b_{33} \end{pmatrix} + \det \begin{pmatrix} b_{11} & b_{13} \\ b_{31} & b_{33} \end{pmatrix} \right\} B + \det(B) = 0, \quad (5)$$

Noticing *five* features: 1) the coefficient of the highest ordered  $B$ , in this case, order 3 has a  $(-)$  sign. As we have shown with other examples, this comes from  $(-)^n$ ;  $n$  being the size of the matrix on hand;  $n = 3$ . 2) The coefficient of the next lowered order of  $B$ , *i.e.*,  $B^2$ , is the Trace of the matrix accompanied by a positive sign,  $\text{Tr}$ , which stands for Trace. 3) the coefficient of the next lower order of  $B$  with an over sign of  $(-)$  is written as a sum of three  $2 \times 2$  determinants, denoted by **det**. The number of terms in the sum increases for larger matrices. Nonetheless, each term of the sum is a  $2 \times 2$  matrix. 4) the  $B$  independent term, *i.e.*,  $B^0$ , is always a determinant of the matrix on hand. 5) The terms of the monotonic polynomial alternate signs, as shown in (5), the highest ordered term has a minus sign that is  $(-)^n$  and, for  $n = 3$  is a minus.

As mentioned, had this been analyzed numerically, the five bullet points mentioned would have been missed, as is the status in the literature [7]-[13].

Taking advantage of the  $B$  independent term in the polynomial, *i.e.*,  $B^0$  term in (5) by multiplying (5) from left by  $B^{-1}$  “unintentionally,” the C-H approach yields the inverted matrix, in our case symbolically. With a bit of manipulation, this gives,

$$B^{-1} = -\frac{1}{\det(B)} \left[ -B^2 + \text{Tr}(B)B - \left\{ \det \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} + \det \begin{pmatrix} b_{22} & b_{23} \\ b_{32} & b_{33} \end{pmatrix} + \det \begin{pmatrix} b_{11} & b_{13} \\ b_{31} & b_{33} \end{pmatrix} \right\} \right] \quad (6)$$

The format of (6) is not reported in the literature [7]-[13]. The only needed criteria applying (6) is a command calculating **(matrix)<sup>power</sup>**, a library command

in any chosen CAS.

### 2.2. Example: A 4 × 4 Matrix

Next, we consider a larger matrix, a 4 × 4. Following the same previous routine explained in Sect. 2a we have,

```
C4=Table[Ci,j,{i,1,4},{j,1,4}];λCI=DiagonalMatrix[λ{1,1,1,1}];(C4-λCI)//MatrixForm;Det[(C4-λI)];λ4ordered={λ4,λ3,λ2,λ,0}//MatrixForm;arrows5=Table["→",5]//MatrixForm;
```

It yields after replacing the λ's in the output polynomial λ->matrix C and re-writing the coefficients in a format compatible with this article's objective (8). Because the expression is long, the output is shortened.

```
{CoefficientList[Collect[Det[C4-λCI],λ],λ]//Reverse//MatrixForm,arrows5,λ4ordered}//Short
```

$$\left\{ \begin{pmatrix} 1 \\ -C_{1,1} - C_{2,2} - C_{3,3} - C_{4,4} \\ -C_{1,2}C_{2,1} + \langle\langle 14 \rangle\rangle + C_{2,2}C_{4,4} + C_{3,3}C_{4,4} \\ \langle\langle 1 \rangle\rangle \\ \langle\langle 1 \rangle\rangle \end{pmatrix}, \begin{pmatrix} " \rightarrow " \\ " \rightarrow " \\ " \rightarrow " \\ " \rightarrow " \\ " \rightarrow " \end{pmatrix}, \begin{pmatrix} \lambda^4 \\ \lambda^3 \\ \lambda^2 \\ \lambda \\ 0 \end{pmatrix} \right\} \quad (7)$$

$$+ C^4 - \text{Tr}(C)C^3 + \left\{ \det \begin{pmatrix} c_{11} & c_{13} \\ c_{31} & c_{33} \end{pmatrix} + \det \begin{pmatrix} c_{22} & c_{23} \\ c_{32} & c_{33} \end{pmatrix} + \det \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \right.$$

$$+ \left. \det \begin{pmatrix} c_{11} & c_{14} \\ c_{41} & c_{44} \end{pmatrix} + \det \begin{pmatrix} c_{22} & c_{24} \\ c_{42} & c_{44} \end{pmatrix} + \det \begin{pmatrix} c_{33} & c_{34} \\ c_{43} & c_{44} \end{pmatrix} \right\} C^2 \quad (8)$$

$$- \left\{ C_{11}C_{22} \det \begin{pmatrix} c_{33} & c_{34} \\ c_{43} & c_{44} \end{pmatrix} - c_{11}c_{23} \det \begin{pmatrix} c_{32} & c_{34} \\ c_{42} & c_{44} \end{pmatrix} \dots \right\} C + \det(c)C^0 = 0$$

The rows of the first column matrix in (7) are the coefficients of the ordered λ polynomial displayed in the third column matrix. The (8) follows all five bullets mentioned in section 2a. By little manipulation, this yields (8) the inverted matrix, *i.e.*, C<sup>-1</sup>.

$$C^{-1} = -\frac{1}{\det[C]} \left\{ +C^4 - \text{Tr}(C)C^3 + \left\{ \det \begin{pmatrix} c_{11} & c_{13} \\ c_{31} & c_{33} \end{pmatrix} + \det \begin{pmatrix} c_{22} & c_{23} \\ c_{32} & c_{33} \end{pmatrix} \right. \right.$$

$$+ \left. \det \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} + \det \begin{pmatrix} c_{11} & c_{14} \\ c_{41} & c_{44} \end{pmatrix} + \det \begin{pmatrix} c_{22} & c_{24} \\ c_{42} & c_{44} \end{pmatrix} + \det \begin{pmatrix} c_{33} & c_{34} \\ c_{43} & c_{44} \end{pmatrix} \right\} C^2 \quad (9)$$

$$- \left\{ C_{11}C_{22} \det \begin{pmatrix} c_{33} & c_{34} \\ c_{43} & c_{44} \end{pmatrix} - C_{11}C_{23} \det \begin{pmatrix} c_{32} & c_{34} \\ c_{42} & c_{44} \end{pmatrix} \dots \right\} C$$

### 2.3. Example: A 2 × 2 Matrix

As a third example consider the simplest case, a 2 × 2 matrix. Although it is the simplest, it provides an opportunity to check the objective of our approach quickly.

```
G=Table[gij,{i,1,2},{j,1,2}];λGI=λIdentityMatrix[2];(G-λGI)//MatrixForm;Det[(G-λGI)]; polyλG=Collect[Det[(G-λGI)],λ]==0
```

$$\lambda^2 - g_{1,2}g_{2,1} + \lambda(-g_{1,1} - g_{2,2}) + g_{1,1}g_{2,2} = 0 \tag{10}$$

As in the previous case (10) is rewritten and manipulated, yielding,

$$G^2 - \text{Tr}(G)G + \det(G) = 0, \tag{11}$$

$$G^{-1} = -\frac{1}{\det(G)} [G - \text{Tr}(G)I] \tag{12}$$

Next, we put the mentioned formulation into action, *i.e.*, a numeric example.

### 3. Numeric Examples

#### A 3 × 3 Matrix

Consider a numeric 3 × 3 matrix. The code below generates a matrix. Its matrix elements are produced using the **Random** command. The numeric range of the elements is {-2, 10}. This range can be set arbitrarily.

```
n=3;Partition[Table[Placeholder[]/.{Placeholder[]->RandomInteger[{-2,10}]],n2],n]//MatrixForm(*//MatrixForm*)(*Det[%]*)
```

$$\begin{pmatrix} 8 & 3 & 5 \\ 7 & 8 & 10 \\ 7 & 0 & 7 \end{pmatrix}$$

```
M=%;
```

```
Det[M] (*check the determinate. If zero, repeat the previous lines*);
```

The determinant equation in the text is coded. The last output of the code is the inverted matrix. This example is in line with public literature [7]-[13].

```
expEqλ=Det[M-DiagonalMatrix[Table[λ,n]]](*formthecharacteristicpolynomial*);reverseCoffexpEqλ=Reverse[CoefficientList[expEqλ,λ]];(*theorderofthepolynomialisreversed,nextlineaccordingtoeq(?)yieldstheinvertedmatrix*)inverseM=-1/(Det[M])(∑i=1n-1(reverseCoffexpEqλ[[i]] MatrixPower[M,n-i])+reverseCoffexpEqλ[[n]] DiagonalMatrix[Table[1,n]]//MatrixForm(*tocheckthecorresctnessoftheprocedute,theproductofthematrrixbyitsinvered shouldresultidentity*formtheHayley-Hamiltoninversion*)MtimesInverted M=inverseM.M;(*checktheaccuracyoftheoutputjustifyingtheproductofthematrrixbyitsinvertedmatrixisunity*)(*checktheoutputvs.library*)libraryInverse dM=Inverse[M];(*libraryInvertedMartix*)difference=inverseM-libraryInversedM//MatrixForm;
```

$$\begin{pmatrix} \frac{8}{33} & -\frac{1}{11} & -\frac{10}{231} \\ \frac{1}{11} & \frac{1}{11} & -\frac{15}{77} \\ -\frac{8}{33} & \frac{1}{11} & \frac{43}{231} \end{pmatrix}$$

## 4. Conclusions

In this research-oriented article, we focused on the *missed* information in literature, even in the well-known respected theory of the Cayley-Hamilton method, which is conducive to the inversion of an  $n \times n$  non-singular matrix. For practical purposes, most Computer Algebra Systems (CAS) include commands that calculate matrix inversion numerically or symbolically. Philosophically speaking, this generally falls in the same category as commands of, e.g., differentiating functions or solving differential equations.

This article differentiates itself from the rest by adding a piece to the named theory. An exhausting literature search shows there is no reference to our report.

In short, with a brief overview, this article reviews the C-H theory and embeds formulation and examples, filling in the *missed* feature in the original article.

The interested reader may find [14] [15] resourceful for *Mathematica* programming.

## Acknowledgments

The author acknowledges the John T. and Page S. Smith Professorship funds for completing and publishing this work.

## Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

## References

- [1] Cayley-Hamilton Theorem. [https://en.wikipedia.org/wiki/Cayley%E2%80%93Hamilton\\_theorem](https://en.wikipedia.org/wiki/Cayley%E2%80%93Hamilton_theorem)
- [2] Mathematica Version 14.0. <http://Wolfram.com>
- [3] Davydov, A.S. (1966) Quantum Mechanics. NEO Press Technical Translated Series.
- [4] Merzbacher, E. (1997) Quantum Mechanics. John Wiley and Sons.
- [5] Powell, J. and Crasemann, B. (2015) Quantum Mechanics. Dover.
- [6] Strang, G. (2023) Introduction to Linear Algebra. 6th Edition, Wellesley Cambridge Press. <https://atozmath.com>CONM>GaussEli>
- [7] Inverse of Matrix Using Cayley Hamilton Method Calculator—Inverse of Matrix Using Cayley Hamilton Method. <https://atozmath.com/>
- [8] Cayley Hamilton Theorem Calculator Calculator Academy. [https://en.wikipedia.org/wiki/Cayley%E2%80%93Hamilton\\_theorem](https://en.wikipedia.org/wiki/Cayley%E2%80%93Hamilton_theorem)
- [9] Inverse of Matrix Using Cayley Hamilton Method. <https://atozmath.com>example>CONM>GaussEli>
- [10] <https://calculator.academy>Cayley-Hamilton-theorem-cal>
- [11] Cayley-Hamilton Matrix Inversion Procedure. <https://testbook.com>
- [12] Step-by-Step Longhand *Mathematica* Package vs. 1) *Mathematica* Builtin Library Function and 2) Microsoft Excel Spreadsheet. <https://mathematica.stackexchange.com/>

- [13] Theorem—From Wolfram MathWorld.  
<https://mathworld.wolfram.com/Hamilton-CayleyTheorem.html>
- [14] Sarafian, H. (2019) “Mathematica Graphics” Examples. 2nd Edition, Scientific Research Publishing.
- [15] Wolfram, S. (2003) The Mathematica Book. 5th Edition, Cambridge University Publications.