Jurnal Matematika Integratif. p-ISSN 1412-6184, e-ISSN 2549-903
Vol. 22, No. 1 (2026), pp. 37–58. doi 10.24198/jmi.v22.n1.68420.37-58

On the Structural Relationship Between the Characteristic and Minimal Polynomials of a Linear Operator

Istiqomah Suharto, Edi Kurniadi, and Ema Carnia

Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran
Jl. Raya Bandung Sumedang KM.21, Hegarmanah, Kec. Jatinangor, Kabupaten Sumedang, Jawa Barat 45363, Indonesia
Email: *istiqomah.work@gmail.com

Abstract

The relationship between the characteristic and minimal polynomial of a linear operator is well-known. The condition under which they coincide is known to be when each eigenvalue has geometric multiplicity one, or equivalently, when the matrix has exactly one Jordan block for each distinct eigenvalue over an algebraically closed field. While this equivalence is well-established in the literature, it is typically proved using Jordan canonical form theory.
In this paper, we review this theory and present a more detailed proof of the equivalence. Our main contribution is to demonstrate that when the characteristic and minimal polynomials are equal, the minimal polynomial can be recovered from a single Krylov sequence. Specifically, if every eigenvalue has geometric multiplicity one, then for any cyclic vector $v \in \mathbb{R}^n$, the first linear dependence in the sequence $\{v, Av, A^2v, \ldots, A^nv\}$ uniquely determines the minimal polynomial.
To illustrate these ideas, we analyze examples using the Leverrier-Faddeev algorithm and Mertens' algorithm. For companion matrices, a cyclic vector yields the minimal polynomial from a single Krylov sequence, since companion matrices are cyclic. For nilpotent matrices with index $k < n$, no cyclic vectors exist, requiring multiple iterations of Mertens' algorithm. These examples highlight when a matrix is cyclic, a single Krylov sequence from a cyclic vector determines the minimal polynomial; otherwise, multiple starting vectors are necessary.
Keywords *Characteristic Polynomial, Minimal Polynomial, Jordan Canonical Form, Leverrier-Faddeev Algorithm*

1. Introduction

The determinant and eigenvalues of a matrix appear frequently in problems of differential equations, numerical analysis, control theory, and dynamical systems. The characteristic polynomial of a matrix A, denoted by $p_A(X)$, provides key information about the eigenvalues of a linear transformation, allowing analysis of stability and long-term behavior [1]. From the characteristic polynomial, one can determine whether a matrix is invertible and, in the context

2020 Mathematics Subject Classification: 15A21, 15A15
Received 03-12-2025, accepted 28-12-2025.

of linear dynamical systems, assess stability and asymptotic behavior (i.e., the long-term rate of decay or growth and the presence of oscillations). For example, in Fatoorehchi et al.'s 2024 work[2], Markov parameters, which are associated to the system's characteristic polynomial, are used to construct a Hankel matrix, which is then employed in the stability analysis of linear time-invariant (LTI) systems.

There are, however, some properties that characteristic polynomials are unable to detect, such as whether or not a matrix is diagonalizable or the size of its Jordan blocks. This is where the minimal polynomial, denoted by $m(X)$, comes in [3]. Minimal polynomials are defined as the unique monic polynomial of least degree that annihilates a given algebraic element or linear operator [4]. Historically, the concept originates in the work of Dedekind and Weber who called them "normal bases" in their 1882 work "Theorie der algebraischen Functionen einer Veränderlichen" [5]. Now, it is widely used in many areas of applied mathematics such as multivariable linear systems theory, coding theory, control theory, and in the spectral theory of rational and polynomial matrices [6]. Dopico et al. in "Block minimal bases ℓ-ifications of matrix polynomials" [6], for example, show how to rewrite matrix polynomials as lower-degree equivalents without changing their fundamental structure. This is done using minimal bases, which serve the same purpose for matrix polynomials that minimal polynomials do for single matrices. This analogy highlights the foundational role of the minimal polynomial. In linear algebra, the minimal polynomial of a matrix or operator contains information about nilpotency, diagonalizability, and the structure of its Jordan canonical form [3], [7]. Because characteristic and minimal polynomials have been vital to mathematical research for over two centuries, they remain indispensable tools [8], needing continued studies into their construction and application.

The relationship between the characteristic and minimal polynomials of a matrix is well-established in the literature. These polynomials coincide when the matrix satisfies a structural condition, such as cyclicity or, equivalently, the existence of a single Jordan block for each eigenvalue (see e.g. Bolotnikov's work in [9]). However, this equivalence is usually presented as a minor consequence of Jordan canonical form theory. It is not often highlighted as a fundamental principle with its own important algebraic and computational implications.

This paper examines the well-established case where the characteristic and minimal polynomials coincide. While this theory is classical, we provide an alternative, algorithmic proof. Furthermore, a key observation is that for such cyclic matrices, Mertens' algorithm [10] simplifies. We then use the resulting algorithmic framework to obtain alternative proofs of fundamental results, including the fact that for a companion matrix C, the characteristic and minimal polynomials coincide with its generating polynomial, $p_C(X) = m_C(X) = f(X)$. We also analyze nilpotent matrices as a contrasting case which is that for a nilpotent matrix N of index k, we recover the classical result that the minimal polynomial is given by $m_N(X) = X^k$.

2. Theoretical Framework

In this chapter, we will be going over the definitions and theorems that are necessary to understanding the results in chapter 3.

2.1. Basic Linear Algebra Theory.

The following theorems are definitions are basic knowledge in Linear Algebra that are essential to characteristic and minimal polynomials and the Jordan structure of a matrix.

Theorem 2.1. *[4] Let A be an $n \times n$ matrix. Then λ is an eigenvalue of A if and only if* $\det(A - \lambda I) = 0$.

Definition 2.2. *A Jordan Block of size m and value λ is a matrix $J_m(\lambda)$ having the value λ repeated along the main diagonal, ones along the superdiagonal and zeroes everywhere else.*

Example 2.3. *Consider the following matrices*

$$J_2(7) = \begin{bmatrix} 7 & 1 \\ 0 & 7 \end{bmatrix}, \qquad J_4(\frac{1}{3}) = \begin{bmatrix} \frac{1}{3} & 1 & 0 & 0 \\ 0 & \frac{1}{3} & 1 & 0 \\ 0 & 0 & \frac{1}{3} & 1 \\ 0 & 0 & 0 & \frac{1}{3} \end{bmatrix}$$

Definition 2.4. *A Jordan form is a block diagonal matrix consisting of several Jordan blocks.*

Example 2.5. *Consider the following matrix*

$$J = \begin{bmatrix} J_2(7) & 0 & 0 \\ 0 & J_2(7) & 0 \\ 0 & 0 & J_4(\frac{1}{3}) \end{bmatrix} = \begin{bmatrix} 7 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 7 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 7 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{3} & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} \end{bmatrix}$$

Theorem 2.6. *Let $A \in \mathbb{R}^{n\times n}$. Then there exists an invertible matrix P and a block-diagonal matrix J, called a* Jordan form *of A, such that*

$$A = PJP^{-1}.$$

The matrix J consists of Jordan blocks

$$J_{m_1}(\lambda_1), \ldots, J_{m_r}(\lambda_r),$$

where $\lambda_1, \ldots, \lambda_r$ are the eigenvalues of A (not necessarily distinct) and

$$m_1 + \cdots + m_r = n.$$

Remark 2.7. *The matrix J is not unique, as the Jordan blocks may be ordered in different ways. Since all Jordan forms are similar, and all quantities studied in this paper are invariant under similarity, the particular choice of J is immaterial.*

Definition 2.8. *[7] For a matrix $A \in \mathbb{R}^{n\times n}$ and a vector $v \in \mathbb{R}^n$, the Krylov sequence of length k is*

$$span\{v, Av, A^2v, \ldots, A^{k-1}v\}.$$

2.2. **Characteristic and Minimal Polynomial of a Linear Transformation.** In this section, we will be going over definitions and theorems pertaining to the characteristic and minimal polynomials of a matrix and their affects on the Jordan structure.

Definition 2.9. *[4] The characteristic polynomial of a square matrix A is the polynomial $p_A(X)$ defined by*

$$p_A(X) = \det(A - XI).$$

The characteristic equation is the equation $p_A(X) = 0$.

Definition 2.10. *[4] The algebraic multiplicity of an eigenvalue λ of a linear transformation f or a matrix A is the multiplicity of λ as a root of $p_A(X)$ or $p_f(X)$. The algebraic multiplicity of an eigenvalue λ is typically denoted by $am(\lambda)$.*

Remark 2.11. *Over an algebraically closed field, the characteristic polynomial of a matrix A can be factored as*

$$p_A(X) = \prod_{i=1}^{k} (X - \lambda_i)^{am(\lambda_i)},$$

where $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of A and $am(\lambda_i)$ denotes the algebraic multiplicity of λ_i.

Definition 2.12. *[4] If λ is an eigenvalue of a linear transformation $f : V \to V$, then the subspace $ker(f - \lambda I)$ is called the λ-eigenspace of f. Its dimension is called the geometric multiplicity of the eigenvalue λ. The geometric multiplicity of an eigenvalue λ is typically denoted by $gm(\lambda)$.*

Example 2.13. *Consider the following matrix*

$$A = \begin{bmatrix} 2 & 1 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda - 2 & -1 & 1 \\ 0 & \lambda - 1 & -1 \\ 0 & 0 & \lambda - 1 \end{bmatrix} = (\lambda - 2)(\lambda - 1)^2$$

So the eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = 1$. The algebraic multiplicities are $am(2) = 1$ and $am(1) = 2$.

For $\lambda_1 = 2$, we get

$$A - 2I = \begin{bmatrix} 0 & 1 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

The eigenspace is $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$, *so $gm(2) = 1$.*

For $\lambda_2 = 1$, we get

$$A - I = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

The eigenspace is $\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$, *so $gm(1) = 1$.*

Definition 2.14. *[4] Consider the polynomial over a field F.*

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

where $a_r \in F, r = 0, 1, 2, \ldots, n$. The polynomial is monic, or is a monomial, if $a_n = 1$.

Definition 2.15. *[4] Let V be a vector space over a field F. Given a linear transformation $f : V \to V$, the minimal polynomial of f, denoted by $m_f(X)$, is the monic polynomial $p(X)$ of least degree which has $p(f) = 0$. The minimal polynomial of a square matrix $A, m_A(X)$, is the monic polynomial $p(X)$ of least degree such that $p(A) = 0$.*

Remark 2.16. *Over an algebraically closed field, the minimal polynomial of a matrix A can be factored as*

$$m_A(X) = \prod_{i=1}^{k} (X - \lambda_i)^{m_i},$$

where $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of A and $1 \le m_i \le am(\lambda_i)$

Definition 2.17. *[7] An $n \times n$ matrix A, over a field $\mathbb{F}$ is said to be cyclic if its characteristic polynomial $p_A(X)$ is equal to its minimal polynomial $m_A(X)$.*

Theorem 2.18. ***(The Cayley-Hamilton Theorem)*** *[4] If A is any square matrix over a field $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and $p_A(X)$ is its characteristic polynomial, then $p_A(A) = 0$.*

Corollary 2.19. *[4] For a square matrix $A, m_A(X)$ divides $p_A(X)$.*

Theorem 2.20. ***(Jordan canonical form)*** *[7] Let $A \in \mathbb{F}^{n\times n}$. Then A is similar to a block-diagonal matrix*

$$J = \begin{pmatrix} J(\lambda_1) & 0 & \cdots & 0 \\ 0 & J(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J(\lambda_m) \end{pmatrix},$$

where each block $J(\lambda_j)$ is of the size $am(\lambda_j) \times am(\lambda_j)$ and has the form

$$J(\lambda_j) = J_{k_1}(\lambda_j) \oplus J_{k_2}(\lambda_j) \oplus \cdots \oplus J_{k_r}(\lambda_j)$$

for some sizes $k_1 \geq k_2 \geq \cdots \geq k_r$ (with $k_1 + \cdots + k_r = n_j$), where each $k_i \times k_i$ Jordan block is

$$J_{k_i}(\lambda_j) = \begin{pmatrix} \lambda_j & 1 & 0 & \cdots & 0 \\ 0 & \lambda_j & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_j & 1 \\ 0 & 0 & \cdots & 0 & \lambda_j \end{pmatrix}.$$

The sizes and multiplicities of these Jordan blocks are uniquely determined by the dimensions

$$d_k = \dim\ker(A - \lambda_j I)^k, \quad k = 1, 2, \ldots$$

Specifically, the number of Jordan blocks of size k equals $(d_k - d_{k-1}) - (d_{k+1} - d_k)$.

Remark 2.21. *We can then get that the characteristic polynomial encodes the multiset of eigenvalues of A. In particular, the algebraic multiplicity of each eigenvalue restricts the Jordan form by determining the total size of all Jordan blocks associated to that eigenvalue.*

For an $n \times n$ matrix, A, with distinct eigenvalues, λ_i, the characteristic polynomial is

$$p_A(X) = \prod_{i=1}^{k} (X - \lambda_i)^{am(\lambda_i)}$$

In the Jordan canonical form of A, this means that all Jordan blocks associated to λ_i have sizes

$$J_{k_{i1}}(\lambda_i),\ J_{k_{i2}}(\lambda_i),\ \ldots,\ J_{k_{ir_i}}(\lambda_i),$$

where the block sizes satisfy

$$k_{i1} + k_{i2} + \cdots + k_{ir_i} = am(\lambda_i).$$

Thus, the characteristic polynomial imposes a constraint on the Jordan structure, which is for each eigenvalue the algebraic multiplicity is the sum of the sizes of its associated Jordan blocks. However, it does not determine how the multiplicity p_i is partitioned among individual Jordan blocks; that additional information will be provided (in part) by the minimal polynomial.

Theorem 2.22. *[11] For an $n \times n$ matrix, A, with distinct eigenvalues, λ_i, the minimal polynomial is*

$$m_A(X) = \prod_{i=1}^{k} (X - \lambda_i)^{m_i}$$

where m_i is the maximal size of a Jordan block of type λ_i in the Jordan canonical form of A.

Remark 2.23. *From 2.21 and 2.22 , we know that if λ_0 is an eigenvalue of an $n \times n$ matrix A, and the characteristic polynomial contains the factor $(\lambda - \lambda_0)^{am(\lambda_0)}$, then the Jordan blocks associated with λ_0 must satisfy*

$$k_1 + k_2 + \cdots = am(\lambda_0), \qquad \max\{k_1, k_2, \ldots\} = m_0,$$

where m_0 is the degree of $(X - \lambda_0)$ in $m_A(X)$. However, these conditions alone do not determine the individual block sizes $k_1, k_2, \ldots$. The characteristic polynomial only fixes the sum of the block sizes (the algebraic multiplicity), not how that total is partitioned into blocks.

In other words, even when $am(\lambda_0)$ is known, several different Jordan canonical forms are possible unless the minimal polynomial is also known. This shows that the characteristic polynomial by itself is insufficient to fully determine the Jordan structure of A.

Example 2.24. *Let A be an $n \times n$ matrix with characteristic polynomial*

$$p_A(X) = (X - \lambda_0)^4.$$

Although the algebraic multiplicity of λ_0 is 4, the Jordan form of A may take any of the following configurations

(1) $J_4(\lambda_0)$
(2) $J_3(\lambda_0), J_1(\lambda_0)$
(3) $J_2(\lambda_0), J_2(\lambda_0)$
(4) $J_2(\lambda_0), J_1(\lambda_0), J_1(\lambda_0)$
(5) $J_1(\lambda_0), J_1(\lambda_0), J_1(\lambda_0), J_1(\lambda_0)$

Each option corresponds to a different Jordan structure despite sharing the same characteristic polynomial.

2.3. **Algorithms to compute the minimal and characteristic polynomial.** By Corollary 2.19, the degree of the minimal polynomial of a matrix A is bounded above by the degree of its characteristic polynomial,

$$\deg m_A(X) \leq \deg p_A(X).$$

As a consequence, classical approaches to computing the minimal polynomial, such as those presented in [4], proceed by first determining the characteristic polynomial and then testing its monic divisors of the form

$$q(X) = \prod_{i=1}^{k} (X - \lambda_i)^{e_i}, \quad 1 \leq e_i \leq am(\lambda_i),$$

to identify the polynomial of least degree satisfying $q(A) = 0$. While theoretically sound, this approach becomes computationally expensive for large matrices due to the combinatorial growth in the number of candidate divisors.

In this work, we focus on two iterative algorithms drawn from the literature: the Leverrier–Faddeev algorithm [12] for computing the characteristic polynomial, and Mertens' algorithm [10] for computing the minimal polynomial. These algorithms are chosen for conceptual rather than computational reasons.

First, both methods are constructive and avoid explicit computation of eigenvalues or canonical forms, allowing the underlying algebraic structure of the problem to be examined directly. Second, their iterative nature makes it possible to analyze the computation step by step, which is useful for understanding how structural properties of the matrix influence the resulting polynomials. Finally, Mertens' algorithm is closely tied to Krylov subspaces generated by repeated application of A to a vector, making it a natural framework for studying cyclic matrices and, in particular, the case in which the minimal and characteristic polynomials coincide.

The Leverrier-Faddeev algorithm was originally due to Urbain Le Verrier (1840) and later refined by Faddeev and Sominskii, and and more recently presented and applied by Nikolaos Halidias in his paper "Computation of the minimal polynomial and applications" (2022) to find the characteristic polynomial of a matrix using trace identities [12].

Algorithm 1 Leverrier-Faddeev Algorithm [12]

Input $A \in \mathbb{R}^{n \times n}$
Output Coefficients $a_1, \ldots, a_n$ of the characteristic polynomial $p_A(\lambda)$

1: $N_1 = I$
2: $a_1 = -\frac{1}{1}\,\text{trace}(AN_1)$
3: for $k = 2, 3, \ldots, n$, compute $N_k = AN_{k-1} + a_{k-1}I$ and $a_k = -\frac{1}{k}\,\text{trace}(AN_k)$
4: The characteristic polynomial is

$$p_A(\lambda) = \lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \cdots + a_{n-1}\lambda + a_n.$$

Next, we present an algorithm to compute the minimal polynomial. This algorithm should be attributed to Mertens(2015)[10]. Remember that the minimal polynomial is the monic polynomial of least degree that annihilates the entire space, and therefore every vector $v \in \mathbb{R}^n$ produces a *local* annihilating polynomial. The minimal polynomial is obtained as the least common multiple of all such local polynomials.

Consider the Krylov sequence generated by a vector v

$$v,\ Av,\ A^2v,\ \ldots,\ A^kv,$$

and suppose that this sequence becomes linearly dependent at step d. This gives a relation

$$A^dv = c_{d-1}A^{d-1}v + \cdots + c_1Av + c_0v,$$

from which we obtain the monic polynomial

$$\widetilde{\mu}(X) = X^d - (c_{d-1}X^{d-1} + \cdots + c_1X + c_0),$$

which annihilates v. By repeating this process for enough vectors to span $\mathbb{R}^n$, we ensure that the least common multiple of the polynomials obtained is the minimal polynomial of A.

We now state the algorithm formally.

Algorithm 2 Mertens' Algorithm for the Minimal Polynomial [10]

Input $A \in \mathbb{R}^{n \times n}$
Output $m_A(X)$

1: Select any non-zero vector $v \in \mathbb{R}^n$
2: Starting with vector v, repeatedly multiply to generate the sequence $v, Av, A^2v, \ldots$ until $\exists k \ni A^kv = C_{k-1}A^{k-1}v + C_{k-2}A^{k-2}v + \cdots + C_2A^2v + C_1A^1v + C_0Av$ for some $C_{k-1}, C_{k-2}, \ldots, C_2, C_1, C_0$.
3: Suppose the linear dependence first occurs at power k. We then form the normalized linear dependency

$$A^kv - (C_{k-1}A^{k-1}v + C_{k-2}A^{k-2}v + \cdots + C_2A^2v + C_1Av + C_0v) = 0$$
$$A^kv - C_{k-1}A^{k-1}v - C_{k-2}A^{k-2}v - \cdots - C_2A^2v - C_1Av - C_0v = 0$$
$$(A^k - C_{k-1}A^{k-1} - C_{k-2}A^{k-2} - \cdots - C_2A^2 - C_1A - C_0)v = 0$$

This corresponds to the polynomial

$$\widetilde{\mu}(X) = X^k - C_{k-1}X^{k-1} - C_{k-2}X^{k-2} - \cdots - C_2X^2 - C_1X - C_0$$

This polynomial is monic and $\widetilde{\mu}(A)v = 0$. This is a divisor of the minimal polynomial $m_A(X)$.
4: construct the set $S = \{v, Av, A^2v, \ldots, A^{k-1}v\}$. If S does not span $\mathbb{R}^n$, choose a vector not in their span and repeat the steps 3 and 4.
5: The minimal polynomial is the least common multiple of all divisors.

3. Result and Discussion

Theorem 3.1. *[9] Let $A \in \mathbb{F}^{n\times n}$, where $\mathbb{F}$ is an algebraically closed field. Then* $m_A(X) = p_A(X) \iff$ *for each eigenvalue λ_i, A has exactly one Jordan block associated with λ_i.*

Proof. Assume the underlying field is algebraically closed.

(⇒) Suppose $m_A(X) = p_A(X)$. Write the characteristic polynomial as

$$p_A(X) = \prod_{i=1}^{k} (X - \lambda_i)^{am(\lambda_i)},$$

where $am(\lambda_i)$ is the algebraic multiplicity of λ_i. Now, write the minimal polynomial as

$$m_A(X) = \prod_{i=1}^{k} (X - \lambda_i)^{m_i}$$

where m_i is the size of the largest Jordan block corresponding to λ_i. Since $m_A(X) = p_A(X)$, we must have $m_i = am(\lambda_i)$ for every i. Now fix an eigenvalue λ_i. Let the sizes of its Jordan blocks be

$$k_{i1}, k_{i2}, \ldots, k_{is_i},$$

From the characteristic and minimal polynomial, we get that

$$k_{i1} + \cdots + k_{is_i} = am(\lambda_i), \qquad \max\{k_{i1}, \ldots, k_{is_i}\} = m_i.$$

But $am(\lambda_i) = m_i$ (since $m_A(X) = p_A(X)$). Thus, the largest block already has size $am(\lambda_i)$, so every other block must have size 0 (which is impossible) or 1. However, adding any block of size 1 gives a total strictly greater than $am(\lambda_i)$. Therefore the only possibility is that there is *exactly one* block, and it must have size $am(\lambda_i)$. Hence, for each eigenvalue λ_i, A has exactly one Jordan block.

(⇐) Conversely, suppose that for each eigenvalue λ_i, there is exactly one Jordan block of size $am(\lambda_i)$. Then the largest block corresponding to λ_i also has size $am(\lambda_i)$. Thus the minimal polynomial contains the factor $(X - \lambda_i)^{am(\lambda_i)}$, and no smaller exponent is sufficient. Therefore

$$m_A(X) = \prod_{i=1}^{k} (X - \lambda_i)^{am(\lambda_i)} = p_A(X).$$

This completes the proof. □

Theorem 3.2. *[9] If A is an $n \times n$ matrix. Then*

$$m_A(X) = p_A(X) \iff \text{for every eigenvalue } \lambda_i,\ gm(\lambda_i) = 1$$

The following example illustrates 3.2 by contrasting two matrices with identical characteristic polynomials but different Jordan structures (one of the matrices does not have a geometric multiplicity of 1).

Example 3.3. *Consider the matrices A and B*

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 2 & 1 & 5 & 0 \\ 0 & 0 & 2 & -3 & 4 \\ 0 & 0 & 0 & 4 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 2 & 1 & 5 & 2 \\ 0 & 0 & 2 & -3 & 4 \\ 0 & 0 & 0 & 4 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Note that both A and B are upper triangular matrices, therefore their characteristic polynomials can be easily derived from their diagonal entries and coincide $p_A(X) = p_B(X) = (X-1)^2(X-2)^2(X-4)$.

Matrix A has an eigenvalue with geometric multiplicity greater than one, while matrix B has geometric multiplicity one for every eigenvalue. According to Theorem 3.2, this should imply $m_A(X) \neq p_A(X)$ and $m_B(X) = p_B(X)$.

Although the characteristic polynomial follows immediately from triangularity, we include the Leverrier–Faddeev computation to illustrate Algorithm 1.

Applying the Leverrier-Faddeev algorithm to matrix A

(1) ***Initialization*** $(k=1)$

$$N_1 = I_5 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Since $AN_1 = A$

$$a_1 = -\frac{1}{1}\operatorname{trace}(AN_1) = -\operatorname{trace}(A) = -(1+2+2+4+1) = -10$$

(2) ***Iteration*** $k=2$

$$N_2 = AN_1 + a_1 I = A - 10I = \begin{bmatrix} -9 & 2 & 3 & 4 & 5 \\ 0 & -8 & 1 & 5 & 0 \\ 0 & 0 & -8 & -3 & 4 \\ 0 & 0 & 0 & -6 & -1 \\ 0 & 0 & 0 & 0 & -9 \end{bmatrix}$$

$$AN_2 = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 2 & 1 & 5 & 0 \\ 0 & 0 & 2 & -3 & 4 \\ 0 & 0 & 0 & 4 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -9 & 2 & 3 & 4 & 5 \\ 0 & -8 & 1 & 5 & 0 \\ 0 & 0 & -8 & -3 & 4 \\ 0 & 0 & 0 & -6 & -1 \\ 0 & 0 & 0 & 0 & -9 \end{bmatrix} = \begin{bmatrix} -9 & -14 & -19 & -27 & 10 \\ 0 & -16 & -6 & -34 & -20 \\ 0 & 0 & -16 & 18 & -40 \\ 0 & 0 & 0 & -24 & 4 \\ 0 & 0 & 0 & 0 & -9 \end{bmatrix}$$

$$a_2 = -\frac{1}{2}\operatorname{trace}(AN_2) = -\frac{1}{2}(-9-16-16-24-9) = -\frac{1}{2}(-74) = 37$$

(3) ***Iteration*** $k=3$

$$N_3 = AN_2 + a_2 I = \begin{bmatrix} -9 & -14 & -19 & -27 & 10 \\ 0 & -16 & -6 & -34 & -20 \\ 0 & 0 & -16 & 18 & -40 \\ 0 & 0 & 0 & -24 & 4 \\ 0 & 0 & 0 & 0 & -9 \end{bmatrix} + 37I = \begin{bmatrix} 28 & -14 & -19 & -27 & 10 \\ 0 & 21 & -6 & -34 & -20 \\ 0 & 0 & 21 & 18 & -40 \\ 0 & 0 & 0 & 13 & 4 \\ 0 & 0 & 0 & 0 & 28 \end{bmatrix}$$

$$AN_3 = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 2 & 1 & 5 & 0 \\ 0 & 0 & 2 & -3 & 4 \\ 0 & 0 & 0 & 4 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 28 & -14 & -19 & -27 & 10 \\ 0 & 21 & -6 & -34 & -20 \\ 0 & 0 & 21 & 18 & -40 \\ 0 & 0 & 0 & 13 & 4 \\ 0 & 0 & 0 & 0 & 28 \end{bmatrix} = \begin{bmatrix} 28 & 28 & 56 & 84 & 52 \\ 0 & 42 & -12 & -68 & -40 \\ 0 & 0 & 42 & -36 & -80 \\ 0 & 0 & 0 & 52 & 16 \\ 0 & 0 & 0 & 0 & 28 \end{bmatrix}$$

$$a_3 = -\frac{1}{3}\operatorname{trace}(AN_3) = -\frac{1}{3}(28+42+42+52+28) = -\frac{1}{3}(192) = -64$$

(4) ***Iteration*** $k=4$

$$N_4 = AN_3 + a_3 I = \begin{bmatrix} 28 & 28 & 56 & 84 & 52 \\ 0 & 42 & -12 & -68 & -40 \\ 0 & 0 & 42 & -36 & -80 \\ 0 & 0 & 0 & 52 & 16 \\ 0 & 0 & 0 & 0 & 28 \end{bmatrix} - 64I = \begin{bmatrix} -36 & 28 & 56 & 84 & 52 \\ 0 & -22 & -12 & -68 & -40 \\ 0 & 0 & -22 & -36 & -80 \\ 0 & 0 & 0 & -12 & 16 \\ 0 & 0 & 0 & 0 & -36 \end{bmatrix}$$

$$AN_4 = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 2 & 1 & 5 & 0 \\ 0 & 0 & 2 & -3 & 4 \\ 0 & 0 & 0 & 4 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -36 & 28 & 56 & 84 & 52 \\ 0 & -22 & -12 & -68 & -40 \\ 0 & 0 & -22 & -36 & -80 \\ 0 & 0 & 0 & -12 & 16 \\ 0 & 0 & 0 & 0 & -36 \end{bmatrix} = \begin{bmatrix} -36 & -16 & -26 & -52 & -48 \\ 0 & -44 & -24 & -136 & -80 \\ 0 & 0 & -44 & 72 & 160 \\ 0 & 0 & 0 & -48 & -64 \\ 0 & 0 & 0 & 0 & -36 \end{bmatrix}$$

$$a_4 = -\frac{1}{4}\,\text{trace}(AN_4) = -\frac{1}{4}(-36-44-44-48-36) = -\frac{1}{4}(-208) = 52$$

(5) ***Iteration*** $k = 5$

$$N_5 = AN_4 + a_4 I = \begin{bmatrix} -36 & -16 & -26 & -52 & -48 \\ 0 & -44 & -24 & -136 & -80 \\ 0 & 0 & -44 & 72 & 160 \\ 0 & 0 & 0 & -48 & -64 \\ 0 & 0 & 0 & 0 & -36 \end{bmatrix} + 52I = \begin{bmatrix} 16 & -16 & -26 & -52 & -48 \\ 0 & 8 & -24 & -136 & -80 \\ 0 & 0 & 8 & 72 & 160 \\ 0 & 0 & 0 & 4 & -64 \\ 0 & 0 & 0 & 0 & 16 \end{bmatrix}$$

$$AN_5 = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 2 & 1 & 5 & 0 \\ 0 & 0 & 2 & -3 & 4 \\ 0 & 0 & 0 & 4 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 16 & -16 & -26 & -52 & -48 \\ 0 & 8 & -24 & -136 & -80 \\ 0 & 0 & 8 & 72 & 160 \\ 0 & 0 & 0 & 4 & -64 \\ 0 & 0 & 0 & 0 & 16 \end{bmatrix} = \begin{bmatrix} 16 & 0 & 0 & 0 & 0 \\ 0 & 16 & 0 & 0 & 0 \\ 0 & 0 & 16 & 0 & 0 \\ 0 & 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & 0 & 16 \end{bmatrix} = 16I$$

$$a_5 = -\frac{1}{5}\,\text{trace}(AN_5) = -\frac{1}{5}\,\text{trace}(16I) = -\frac{1}{5}(16 \times 5) = -\frac{1}{5}(80) = -16$$

After completing all iterations, we obtain the coefficients

$$a_1 = -10, \quad a_2 = 37, \quad a_3 = -64, \quad a_4 = 52, \quad a_5 = -16$$

Thus, the characteristic polynomial of A *is*

$$p_A(\lambda) = \lambda^5 + a_1\lambda^4 + a_2\lambda^3 + a_3\lambda^2 + a_4\lambda + a_5 = \lambda^5 - 10\lambda^4 + 37\lambda^3 - 64\lambda^2 + 52\lambda - 16$$

The same procedure applied to matrix B *yields identical coefficients (the computation is similar but slightly different due to the non-zero entry in position (2,5)), confirming that* $p_A(\lambda) = p_B(\lambda)$.

Now, we calculate their respective minimal polynomials using Mertens algorithm.

Applying Mertens algorithm to matrix A

(1) ***Iteration 1*** *Choose initial vector* $v_1 = e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$.

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad Av_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

We immediately observe that $Av_1 = v_1$, *giving us the linear dependence*

$$Av_1 - v_1 = 0 \iff (A - I)v_1 = 0$$

Thus, the first divisor is

$$\widetilde{\mu}_1(X) = X - 1$$

$S_1 = \{v_1\}$ *does not span* $\mathbb{R}^5$

(2) ***Iteration 2*** *Choose* $v_2 = e_2 = (0,1,0,0,0)^T$

$$v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad Av_2 = \begin{bmatrix} 2 \\ 2 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad A^2v_2 = \begin{bmatrix} 6 \\ 4 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$A^2v_2 - 3Av_2 + 2v_2 = 0$, *giving* $(A^2 - 3A + 2I)v_2 = 0$. *The second divisor is*

$$\widetilde{\mu}_2(X) = X^2 - 3X + 2 = (X-1)(X-2)$$

The set $S_1 \cup (S_2 = \{v_2, Av_2\})$ *does not span* $\mathbb{R}^5$

(3) ***Iteration 3*** *Choose* $v_3 = e_3 = (0,0,1,0,0)^T$

$$v_3 = \begin{bmatrix}0\\0\\1\\0\\0\end{bmatrix}, \quad Av_3 = \begin{bmatrix}3\\1\\2\\0\\0\end{bmatrix}, \quad A^2v_3 = \begin{bmatrix}11\\4\\4\\0\\0\end{bmatrix}, \quad A^3v_3 = \begin{bmatrix}31\\12\\8\\0\\0\end{bmatrix}$$

$(A^3 - 5A^2 + 8A - 4I)v_3 = 0$. *The third divisor is*

$$\widetilde{\mu}_3(X) = X^3 - 5X^2 + 8X - 4 = (X-1)(X-2)^2$$

Now $S_1 \cup S_2 \cup (S_3 = \{v_3, Av_3, A^2v_3\})$ *does not span* $\mathbb{R}^5$

(4) ***Iteration 4*** *Choose* $v_4 = e_4 = (0,0,0,1,0)^T$

$$v_4 = \begin{bmatrix}0\\0\\0\\1\\0\end{bmatrix}, \quad Av_4 = \begin{bmatrix}4\\5\\-3\\4\\0\end{bmatrix}, \quad A^2v_4 = \begin{bmatrix}21\\27\\-18\\16\\0\end{bmatrix}, \quad A^3v_4 = \begin{bmatrix}85\\116\\-84\\64\\0\end{bmatrix} \quad A^4v_4 = \begin{bmatrix}321\\468\\-360\\256\\0\end{bmatrix}$$

$(A^4 - 9A^3 + 28A^2 - 36A + 16I)v_4 = 0$. *The third divisor is*

$$\widetilde{\mu}_4(X) = X^4 - 9X^3 + 28X^2 - 36X + 16 = (X-1)(X-2)^2(X-4)$$

Now $S_1 \cup S_2 \cup S_3 \cup (S_4 = \{v_4, Av_4, A^2v_4, A^3v_4\})$ *does not span* $\mathbb{R}^5$

(5) ***Iteration 5*** *Choose* $v_5 = e_5 = (0,0,0,0,1)^T$

$$v_5 = \begin{bmatrix}0\\0\\0\\0\\1\end{bmatrix}, \quad Av_5 = \begin{bmatrix}5\\0\\4\\-1\\1\end{bmatrix}, \quad A^2v_5 = \begin{bmatrix}18\\-1\\-15\\-5\\1\end{bmatrix}, \quad A^3v_5 = \begin{bmatrix}46\\-12\\49\\-21\\1\end{bmatrix} \quad A^4v_5 = \begin{bmatrix}90\\-80\\165\\-85\\1\end{bmatrix}$$

$(A^4 - 9A^3 + 28A^2 - 36A + 16I)v_5 = 0$. *The third divisor is*

$$\widetilde{\mu}_5(X) = X^4 - 9X^3 + 28X^2 - 36X + 16 = (X-1)(X-2)^2(X-4)$$

Since the vectors $v_1, \ldots, v_5$ *are the standard basis vectors of* $\mathbb{R}^5$, *their union is linearly independent and spans* $\mathbb{R}^5$. $S_1 \cup S_2 \cup S_3 \cup S_4 \cup S_5$ *spans* $\mathbb{R}^5$

(6) ***The Minimal Polynomial of A*** *The minimal polynomial is the least common multiple of all divisors*

$$m_A(X) = lcm(\widetilde{\mu}_1(X), \widetilde{\mu}_2(X), \widetilde{\mu}_3(X), \widetilde{\mu}_4(X), \widetilde{\mu}_5(X))$$

$$m_A(X) = (X-1)(X-2)^2(X-4)$$

Applying Mertens algorithm to matrix B

(1) ***Iteration 1*** *Choose initial vector* $v_1 = e_1 = (1,0,0,0,0)^T$.

$$v_1 = \begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix}, \quad Bv_1 = \begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix}$$

We observe that $Bv_1 = v_1$, *giving the linear dependence*

$$Bv_1 - v_1 = 0 \iff (B - I)v_1 = 0$$

Thus, the first divisor is

$$\widetilde{\mu}_1(X) = X - 1$$

$S_1 = \{v_1\}$ *does not span* $\mathbb{R}^5$.

(2) ***Iteration 2*** *Choose* $v_2 = e_2 = (0, 1, 0, 0, 0)^T$.

$$v_2 = \begin{bmatrix}0\\1\\0\\0\\0\end{bmatrix}, \quad Bv_2 = \begin{bmatrix}2\\2\\0\\0\\0\end{bmatrix}, \quad B^2v_2 = \begin{bmatrix}6\\4\\0\\0\\0\end{bmatrix}$$

$(B^2 - 3B + 2I)v_2 = 0$. *The second divisor is*

$$\widetilde{\mu}_2(X) = X^2 - 3X + 2 = (X-1)(X-2)$$

The set $S_1 \cup S_2$ *does not span* $\mathbb{R}^5$.

(3) ***Iteration 3*** *Choose* $v_3 = e_3 = (0, 0, 1, 0, 0)^T$.

$$v_3 = \begin{bmatrix}0\\0\\1\\0\\0\end{bmatrix}, \quad Bv_3 = \begin{bmatrix}3\\1\\2\\0\\0\end{bmatrix}, \quad B^2v_3 = \begin{bmatrix}11\\4\\4\\0\\0\end{bmatrix}, \quad B^3v_3 = \begin{bmatrix}31\\12\\8\\0\\0\end{bmatrix}$$

$(B^3 - 5B^2 + 8B - 4I)v_3 = 0$. *The third divisor is*

$$\widetilde{\mu}_3(X) = X^3 - 5X^2 + 8X - 4 = (X-1)(X-2)^2$$

The set $S_1 \cup S_2 \cup S_3$ *does not span* $\mathbb{R}^5$.

(4) ***Iteration 4*** *Choose* $v_4 = e_4 = (0, 0, 0, 1, 0)^T$.

$$v_4 = \begin{bmatrix}0\\0\\0\\1\\0\end{bmatrix}, \quad Bv_4 = \begin{bmatrix}4\\5\\-3\\4\\0\end{bmatrix}, \quad B^2v_4 = \begin{bmatrix}21\\27\\-18\\16\\0\end{bmatrix}, \quad B^3v_4 = \begin{bmatrix}85\\116\\-84\\64\\0\end{bmatrix} \quad B^4v_4 = \begin{bmatrix}321\\468\\-360\\256\\0\end{bmatrix}$$

$(B^4 - 9B^3 + 28B^2 - 36B + 16I)v_4 = 0$. *The fourth divisor is*

$$\widetilde{\mu}_4(X) = X^4 - 9X^3 + 28X^2 - 36X + 16 = (X-1)(X-2)^2(X-4)$$

The set $S_1 \cup S_2 \cup S_3 \cup S_4$ *does not span* $\mathbb{R}^5$.

(5) ***Iteration 5*** *Choose* $v_5 = e_5 = (0, 0, 0, 0, 1)^T$.

$$v_5 = \begin{bmatrix}0\\0\\0\\0\\1\end{bmatrix}, \quad Bv_5 = \begin{bmatrix}5\\2\\4\\-1\\1\end{bmatrix}, \quad B^2v_5 = \begin{bmatrix}22\\5\\15\\-5\\1\end{bmatrix}, \quad B^3v_5 = \begin{bmatrix}62\\2\\49\\-21\\1\end{bmatrix} B^4v_5 = \begin{bmatrix}134\\-50\\165\\-85\\1\end{bmatrix}, \quad B^5v_5 = \begin{bmatrix}194\\-358\\589\\-341\\1\end{bmatrix}$$

$(B^5 - 10B^4 + 37B^3 - 64B^2 + 52B - 16I)v_5 = 0$. *The fifth divisor is*

$$\widetilde{\mu}_5(X) = X^5 - 10X^4 + 37X^3 - 64X^2 + 52X - 16 = (X-1)^2(X-2)^2(X-4)$$

Since the vectors $v_1, \ldots, v_5$ *are the standard basis vectors of* $\mathbb{R}^5$, *their union is linearly independent and spans* $\mathbb{R}^5$. *Hence the union of all sets* S_i *spans* $\mathbb{R}^5$.

(6) ***The Minimal Polynomial of B*** *The minimal polynomial is the least common multiple of all divisors*

$$m_B(X) = lcm(\widetilde{\mu}_1(X), \widetilde{\mu}_2(X), \widetilde{\mu}_3(X), \widetilde{\mu}_4(X), \widetilde{\mu}_5(X))$$

$$m_B(X) = lcm\Big((X-1), (X-1)(X-2), (X-1)(X-2)^2, (X-1)(X-2)^2(X-4), (X-1)^2(X-2)^2(X-4)\Big)$$

$$m_B(X) = (X-1)^2(X-2)^2(X-4)$$

From carrying out the calculations above, we get that

$$m_A(X) \neq p_A(X), \quad m_B(X) = p_B(X)$$

A simple look at their geometric multiplicities provides the same information.

For matrix A, the eigenvalue $\lambda = 1$ has algebraic multiplicity 2 but possesses two linearly independent eigenvectors

$$\mathbf{v}_{1,1} = (1,0,0,0,0)^\top, \quad \mathbf{v}_{1,2} = (0,4/3,-3,1/3,1)^\top$$

So $gm(1) = 2$, the degree of the minimal polynomial is strictly less than that of the characteristic polynomial.

Conversely, for matrix B, the geometric multiplicity of $\lambda = 1$ is restricted to 1, yielding only

$$\mathbf{v}_1 = (1,0,0,0,0)^\top$$

Since the geometric multiplicities for $\lambda = 4$, $\lambda = 2$, and $\lambda = 1$ are all 1 in matrix B, thus $m_B(X) = p_B(X)$.

Since the geometric multiplicity equals the number of Jordan blocks associated with an eigenvalue, the conclusions above follow directly from the Jordan canonical form characterization stated in 3.1.

From 3.2, we also get that the computation of both matrices can be used by one algorithm only, and when we use Mertens algorithm in 2.3 the process is shortened to one iteration.

Proposition 3.4. *Let $A \in \mathbb{R}^{n\times n}$ be a matrix such that the geometric multiplicity of every eigenvalue is 1*

$$gm(\lambda_i) = 1 \qquad \text{for every eigenvalue } \lambda_i.$$

Let v be a cyclic vector in $\mathbb{R}^n$ (that is, $\{v, Av, A^2v, \ldots, A^{n-1}v\}$ spans $\mathbb{R}^n$), the minimal polynomial of A is uniquely determined by the linear dependence relation of the Krylov sequence

$$\{v, Av, A^2v, \ldots, A^n v\}.$$

Specifically, if we write $A^n v$ as a linear combination of the previous vectors

$$A^n v = c_{n-1}A^{n-1}v + c_{n-2}A^{n-2}v + \cdots + c_1 Av + c_0 v,$$

then the minimal polynomial of A is

$$m_A(X) = X^n - (c_{n-1}X^{n-1} + c_{n-2}X^{n-2} + \cdots + c_1 X + c_0).$$

Proof. Recall that the geometric multiplicity $gm(\lambda_i)$ equals the number of Jordan blocks associated with the eigenvalue λ_i. Since $gm(\lambda_i) = 1$ for every eigenvalue, there is exactly one Jordan block for each distinct eigenvalue. For such matrices, the minimal polynomial $m_A(X)$ coincides with the characteristic polynomial $p_A(X)$. Because $p_A(X)$ has degree n for an $n \times n$ matrix, we obtain

$$\deg(m_A) = n. \tag{1}$$

Since A has exactly one Jordan block for each eigenvalue, it is a cyclic matrix. That is, there exists at least one vector $v \in \mathbb{R}^n$ such that the Krylov sequence

$$\{v, Av, A^2v, \ldots, A^{n-1}v\}$$

is linearly independent and hence forms a basis of $\mathbb{R}^n$.

Fix such a cyclic vector v.

Because this set is a basis, the next vector in the sequence must be linearly dependent on it. Hence there exist unique scalars $c_0, c_1, \ldots, c_{n-1}$ such that

$$A^n v = c_{n-1}A^{n-1}v + c_{n-2}A^{n-2}v + \cdots + c_1 Av + c_0 v.$$

Define the monic polynomial

$$p(X) = X^n - \sum_{k=0}^{n-1} c_k X^k.$$

Applying this polynomial to A and evaluating at v gives

$$\begin{aligned} p(A)v &= A^n v - \sum_{k=0}^{n-1} c_k A^k v \\ &= 0. \end{aligned}$$

Since $p(A)$ is a polynomial in A, it commutes with A. Therefore, for each $k = 0, 1, \ldots, n-1$,

$$p(A)(A^k v) = A^k(p(A)v) = 0.$$

But $p(A)v = 0$, so

$$A^k(p(A)v = A^k(0) = 0$$

Hence,

$$p(A)(A^k v) = 0, \forall k = 0, 1, \ldots, n-1.$$

Thus $p(A)$ sends every vector in the basis

$$\{v, Av, A^2v, \ldots, A^{n-1}v\}$$

to zero.

By linearity, a linear operator that annihilates every vector of a basis must be the zero operator. Hence

$$p(A) = 0.$$

By definition of the minimal polynomial, $m_A(X)$ divides every polynomial that annihilates A. Thus there exists a polynomial $q(X)$ such that

$$p(X) = m_A(X)\, q(X).$$

Taking degrees and using (1), we obtain

$$n = \deg(p) = \deg(m_A) + \deg(q),$$

which forces $\deg(q) = 0$. Since both $p(X)$ and $m_A(X)$ are monic, this implies $q(X) = 1$, and therefore

$$p(X) = m_A(X).$$

□

Lemma 3.5. *Let $A \in \mathbb{R}^{n\times n}$ be a matrix whose minimal polynomial satisfies*

$$m_A(X) = p_A(X).$$

Equivalently, each eigenvalue of A has geometric multiplicity one. Then the set of cyclic vectors

$$\mathcal{C}(A) := \{v \in \mathbb{R}^n : \operatorname{span}\{v, Av, \ldots, A^{n-1}v\} = \mathbb{R}^n\}$$

is a dense open subset of $\mathbb{R}^n$. In particular, a generic choice of v yields a Krylov sequence whose first linear dependence occurs at step n, and this dependence uniquely determines the minimal polynomial of A.

Proof. Since $m_A(X) = p_A(X)$, each eigenvalue of A has geometric multiplicity equal to one, and hence A has exactly one Jordan block for each distinct eigenvalue. In particular, A is cyclic, so there exists at least one vector whose Krylov sequence spans $\mathbb{R}^n$.

A vector $v \in \mathbb{R}^n$ is non-cyclic if and only if the vectors

$$\{v, Av, \ldots, A^{n-1}v\}$$

are linearly dependent. Define the Krylov matrix

$$K_A(v) := \begin{pmatrix} v & Av & A^2v & \cdots & A^{n-1}v \end{pmatrix} \in \mathbb{R}^{n\times n}.$$

Then v is non-cyclic precisely when $\det(K_A(v)) = 0$.

Writing $v = (v_1, \ldots, v_n)^T$, each column $A^k v$ depends linearly on the entries of v. Hence each entry of $K_A(v)$ is a polynomial in $v_1, \ldots, v_n$, and therefore

$$\Phi_A(v) := \det(K_A(v))$$

is a polynomial function on $\mathbb{R}^n$. The set of non-cyclic vectors is

$$\mathcal{C}(A)^c = \{v \in \mathbb{R}^n : \Phi_A(v) = 0\}.$$

We still need to show that Φ_A is not identically zero. Let $A = PJP^{-1}$ be the Jordan form of A. Since each eigenvalue has geometric multiplicity equal to one, the Jordan form J consists of exactly one Jordan block per eigenvalue. Let u be the vector obtained by summing the first standard basis vector from each Jordan block of J. Then the Krylov sequence

$$\{u, Ju, J^2u, \dots, J^{n-1}u\}$$

is linearly independent and spans $\mathbb{R}^n$, so $\det(K_J(u)) \neq 0$.

Setting $v := Pu$, we obtain

$$K_A(v) = PK_J(u),$$

and since P is invertible, it follows that

$$\det(K_A(v)) = \det(P)\det(K_J(u)) \neq 0.$$

Thus Φ_A is not zero polynomial.

Consequently, $\mathcal{C}(A)^c$ is a proper algebraic subset of $\mathbb{R}^n$. Its complement $\mathcal{C}(A)$ is therefore open and dense. □

Example 3.6. *Let B be the matrix from example 3.3, for which we established that $gm(\lambda) = 1$ for every eigenvalue λ. We previously found that $m_B(X) = p_B(X)$ using 5 iterations of Mertens' algorithm 2.3 using $e_1 = (1, 0, 0, 0, 0)^T$.*

Notice that B is upper triangular.For a general upper triangular matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix},$$

when we compute the Krylov sequence starting from e_1, the Krylov sequence is $\{e_1, a_{11}e_1, a_{11}^2e_1, \dots a_{11}^{n-1}e_1\}$ which spans only a 1-dimensional subspace. For our specific matrix B, $a_{11} = 1$ so we obtain the Krylov sequence $\{e_1, e_1, e_1, \dots e_1\}$. Therefore e_1 is not cyclic in $\mathbb{R}^5$. Hence why proposition 3.4 was not satisfied and the calculation to obtain the minimal polynomial took 5 iterations.

This time, we choose the vector

$$v = (1, 1, 1, 1, 1)^\top.$$

Computing the Krylov sequence, we obtain

$$v = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad Bv = \begin{bmatrix} 15 \\ 10 \\ 3 \\ 3 \\ 1 \end{bmatrix}, \quad B^2v = \begin{bmatrix} 61 \\ 40 \\ 1 \\ 11 \\ 1 \end{bmatrix},$$

$$B^3v = \begin{bmatrix} 193 \\ 138 \\ -27 \\ 43 \\ 1 \end{bmatrix}, \quad B^4v = \begin{bmatrix} 565 \\ 466 \\ -179 \\ 171 \\ 1 \end{bmatrix}, \quad B^5v = \begin{bmatrix} 1649 \\ 1610 \\ -867 \\ 683 \\ 1 \end{bmatrix}.$$

The determinant of the Krylov matrix

$$\det \begin{bmatrix} v & Bv & B^2v & B^3v & B^4v \end{bmatrix} \neq 0,$$

is not zero, but when we include B^5v the determinant of the extended Krylov matrix is 0.

$$\det \begin{bmatrix} v & Bv & B^2v & B^3v & B^4v & B^5v \end{bmatrix} = 0,$$

This shows the first occurrence of linear dependence is at B^5v. The linear dependence is $B^5v - 10B^4v + 37B^3v - 64B^2v + 52Bv - 16v = 0$

This corresponds directly to the minimal polynomial

$$m_B(X) = X^5 - 10X^4 + 37X^3 - 64X^2 + 52X - 16 = (X-1)^2(X-2)^2(X-4)$$

which matches the result we obtained in the multi-iteration algorithm.

This example illustrates that when all eigenvalues of a matrix have geometric multiplicity one, the minimal polynomial can be recovered from a single Krylov sequence generated by a generic vector, without the need for multiple iterations of Mertens' algorithm.

Remark 3.7. *For an upper triangular matrix, the standard basis vector e_i is always an eigenvector (with eigenvalue equal to the i-th diagonal entry). Eigenvectors are never cyclic for matrices with dimension $n \geq 2$, because their Krylov sequences are*

$$\{v, \lambda v, \lambda^2 v, \ldots, \lambda^{n-1} v\},$$

which span only a 1-dimensional space.

To obtain a cyclic vector, we must choose v such that it has nonzero components in the directions corresponding to each Jordan block. For the matrix B in Example 3.3, which has Jordan blocks of sizes $2, 2, 1$ (corresponding to eigenvalues $1, 2, 4$ with algebraic multiplicities $2, 2, 1$), a vector with all nonzero entries is generically a good choice, hence why in Example 3.6, proposition 3.4 was satisfied.

Theorem 3.8. *[13] Let $N \in \mathbb{R}^{n\times n}$ be a nilpotent matrix such that*

$$N^k = 0 \quad \text{and} \quad N^{k-1} \neq 0,$$

for some $k < n$. Then

$$m_N(X) = X^k.$$

Proof. We interpret the result through the behavior of Mertens' algorithm and its associated Krylov sequences.

Since $N^{k-1} \neq 0$, there exists a vector $u_1 \in \mathbb{R}^n$ such that $N^{k-1}u_1 \neq 0$. It is convenient to begin our Krylov process with the vector

$$v = N^{k-1}u_1,$$

which is nonzero by construction.

Now, when we apply N to this vector, we immediately obtain

$$Nv = N(N^{k-1}u_1) = N^k u_1 = 0.$$

Thus the Krylov sequence generated by v,

$$v, \quad Nv, \quad N^2v, \; \ldots,$$

collapses at the first step (first linear dependence occurs at power 1), and the corresponding monic polynomial is simply

$$\widetilde{\mu}_1(X) = X.$$

This polynomial annihilates v, but of course it does not yet annihilate the entire space, so it cannot be the minimal polynomial of N.

To force higher powers of X, we must look at other starting vectors. Since $N^{k-2} \neq 0$, we may choose another vector u_2 such that $N^{k-2}u_2 \neq 0$. Starting instead with

$$v_2 = N^{k-2}u_2,$$

the Krylov sequence becomes longer

$$v_2, \quad Nv_2 = N^{k-1}u_2, \quad N^2v_2 = N^k u_2 = 0.$$

Here the first dependence occurs at power 2, giving the polynomial

$$\widetilde{\mu}_2(X) = X^2,$$

which now annihilates both v_2 and v.

Continuing this pattern, each time we start with a vector of the form $v_j = N^{k-j}u_j \neq 0$, the Krylov sequence collapses at exactly j steps (first linear dependence at power j), producing the monic polynomial

$$\widetilde{\mu}_j(X) = X^j$$

At the final stage, where $j = k$, take a vector v_k with $N^0 v_k = v_k \neq 0$ but $N^k v_k = 0$. Its Krylov sequence is

$$v_k,\ Nv_k,\ N^2v_k,\ \ldots,\ N^k v_k = 0$$

and the first linear dependence appears at power k. This yields the monic polynomial

$$\widetilde{\mu}_k(X) = X^k,$$

which is the smallest polynomial that annihilates all the vectors considered. By Mertens' algorithm, the minimal polynomial is the least common multiple of all such polynomials, and therefore

$$m_N(X) = \mathrm{lcm}(X, X^2, \ldots, X^k) = X^k.$$

This confirms that the minimal polynomial of a nilpotent matrix whose nilpotency index is k is exactly X^k. □

Theorem 3.8 illustrates a fundamental limitation of Krylov-based determination of minimal polynomials. For a nilpotent matrix N of index $k < n$, the minimal polynomial satisfies $m_N(X) = X^k$, while the characteristic polynomial is $p_N(X) = X^n$. Hence $m_N(X) \neq p_N(X)$.

As a consequence, no single Krylov sequence of the form

$$\{v, Nv, N^2v, \ldots, N^n v\}$$

can determine the minimal polynomial through a linear dependence at degree n. Instead, Krylov sequences terminate strictly earlier, reflecting the presence of multiple Jordan blocks.

This behavior contrasts with the situation described in Proposition 3.4, where the equality $m_A(X) = p_A(X)$ guarantees the existence of a vector whose Krylov sequence uniquely determines the minimal polynomial. Nilpotent matrices therefore provide a canonical example in which this condition fails.

We now present a complementary situation to the nilpotent case, in which the minimal and characteristic polynomials always coincide and can be recovered from a single Krylov sequence.

Theorem 3.9. *[14] Let C be the companion matrix of a monic polynomial $f(X) = X^n + c_{n-1}X^{n-1} + \cdots + c_1X + c_0$*

$$C = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & 0 & \cdots & 0 & -c_2 \\ \vdots & & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & -c_{n-2} \\ 0 & 0 & 0 & \cdots & 1 & -c_{n-1} \end{pmatrix}$$

Then the characteristic polynomial $p_C(X) = f(X)$ and the minimal polynomial $m_C(X) = f(X)$.

Proof. We will prove this by computing both polynomials using the Leverrier–Faddeev algorithm and Mertens' algorithm respectively.

We want to prove that applying the Leverrier-Faddeev algorithm to the companion matrix C gives the coefficients of the polynomial $f(X)$. Specifically, we want to show that for each step k, the computed coefficient is

$$a_k = c_{n-k}$$

This means $a_1 = c_{n-1}$, $a_2 = c_{n-2}$, and so on.

We begin with $N_1 = I$. The trace of the companion matrix C is the sum of its diagonal elements. Since the only non-zero diagonal element is the last one

$$\mathrm{trace}(C) = -c_{n-1}.$$

Substituting this into the algorithm's formula

$$a_1 = -\frac{1}{1}\,\mathrm{trace}(CN_1) = -\,\mathrm{trace}(C) = -(-c_{n-1}) = c_{n-1}.$$

We compute the next matrix N_2.

$$N_2 = CN_1 + a_1 I = C + c_{n-1}I.$$

The diagonal of C is $(0, 0, \ldots, 0, -c_{n-1})$. The diagonal of $c_{n-1}I$ is $(c_{n-1}, c_{n-1}, \ldots, c_{n-1}, c_{n-1})$. Adding these together

$$\mathrm{diag}(N_2) = (c_{n-1}, c_{n-1}, \ldots, c_{n-1}, 0).$$

Notice that the last element becomes zero because $-c_{n-1} + c_{n-1} = 0$.

Now we calculate a_2. The formula is

$$a_2 = -\frac{1}{2}\,\mathrm{trace}(CN_2).$$

Substituting $N_2 = C + c_{n-1}I$, we obtain

$$\begin{aligned}\mathrm{trace}(CN_2) &= \mathrm{trace}(C(C + c_{n-1}I)) \\ &= \mathrm{trace}(C^2) + c_{n-1}\,\mathrm{trace}(C) \\ &= \mathrm{trace}(C^2) - c_{n-1}^2,\end{aligned}$$

where we used $\mathrm{trace}(C) = -c_{n-1}$.

Rather than computing $\mathrm{trace}(C^2)$ directly, we use Newton's identities, which relate the power sums $s_k = \mathrm{trace}(C^k)$ (the sums of k-th powers of eigenvalues) to the coefficients of the characteristic polynomial. For $k = 2$, Newton's identity states

$$s_2 + c_{n-1}s_1 + 2c_{n-2} = 0.$$

Since $s_1 = \mathrm{trace}(C) = -c_{n-1}$ and $s_2 = \mathrm{trace}(C^2)$, we have

$$\mathrm{trace}(C^2) + c_{n-1}(-c_{n-1}) + 2c_{n-2} = 0,$$

which gives

$$\mathrm{trace}(C^2) = c_{n-1}^2 - 2c_{n-2}.$$

Substituting back:

$$\begin{aligned}\mathrm{trace}(CN_2) &= (c_{n-1}^2 - 2c_{n-2}) - c_{n-1}^2 \\ &= -2c_{n-2}.\end{aligned}$$

Finally, we compute a_2:

$$a_2 = -\frac{1}{2}\,\mathrm{trace}(CN_2) = -\frac{1}{2}(-2c_{n-2}) = c_{n-2}.$$

Assume that the properties hold for all steps $j < k$. Specifically, we assume that $a_{j-1} = c_{n-(j-1)}$, and that

$$N_{k-1} = C^{k-2} + c_{n-1}C^{k-3} + \cdots + c_{n-k+2}I.$$

We use the definition from the Leverrier–Faddeev algorithm to compute N_k

$$N_k = CN_{k-1} + a_{k-1}I.$$

Substitute the expression for N_{k-1} from the Inductive Hypothesis and substitute the known coefficient $a_{k-1} = c_{n-k+1}$

$$N_k = C\left(C^{k-2} + c_{n-1}C^{k-3} + \cdots + c_{n-k+2}I\right) + c_{n-k+1}I.$$

$$N_k = \left(C^{k-1} + c_{n-1}C^{k-2} + c_{n-2}C^{k-3} + \cdots + c_{n-k+2}C\right) + c_{n-k+1}I.$$

$$N_k = C^{k-1} + c_{n-1}C^{k-2} + c_{n-2}C^{k-3} + \cdots + c_{n-k+1}I.$$

This proves Property I for step k.

The algorithm defines a_k using the trace of CN_k

$$a_k = -\frac{1}{k}\,\mathrm{trace}(CN_k).$$

First, we find the matrix product CN_k by multiplying the expression derived above by C

$$CN_k = C^k + c_{n-1}C^{k-1} + c_{n-2}C^{k-2} + \cdots + c_{n-k+1}C.$$

Now, we take the trace of this sum of matrices. Since the trace is a linear operator, we can separate the terms

$$\mathrm{trace}(CN_k) = \mathrm{trace}(C^k) + c_{n-1}\,\mathrm{trace}(C^{k-1}) + c_{n-2}\,\mathrm{trace}(C^{k-2}) + \cdots + c_{n-k+1}\,\mathrm{trace}(C).$$

Let $s_j = \mathrm{trace}(C^j)$ represent the power sum of the eigenvalues (sum of j-th powers of the roots of $f(X)$). The expression becomes

$$\mathrm{trace}(CN_k) = s_k + c_{n-1}s_{k-1} + c_{n-2}s_{k-2} + \cdots + c_{n-k+1}s_1.$$

This specific linear combination of power sums and polynomial coefficients is directly related by Newton's Identities

$$s_k + c_{n-1}s_{k-1} + c_{n-2}s_{k-2} + \cdots + c_{n-k+1}s_1 + kc_{n-k} = 0.$$

Solving this identity for the trace term

$$\mathrm{trace}(CN_k) = s_k + c_{n-1}s_{k-1} + \cdots + c_{n-k+1}s_1 = -kc_{n-k}.$$

Finally, substituting this result back into the formula for a_k

$$a_k = -\frac{1}{k}\,(-kc_{n-k}) = c_{n-k}.$$

This proves Property II for step k.

By the principle of mathematical induction, the Leverrier–Faddeev algorithm produces coefficients $a_k = c_{n-k}$ for all $k = 1, \ldots, n$. Since the characteristic polynomial is defined as

$$p_C(X) = X^n + a_1X^{n-1} + a_2X^{n-2} + \cdots + a_n,$$

substituting the results yields

$$p_C(X) = X^n + c_{n-1}X^{n-1} + c_{n-2}X^{n-2} + \cdots + c_0 = f(X).$$

The characteristic polynomial of the companion matrix C is indeed the polynomial $f(X)$ from which C was constructed.

Now we apply Mertens' algorithm to find the minimal polynomial of C. Choose the starting vector $v = e_1 = (1, 0, 0, \ldots, 0)^T$. Compute the Krylov sequence $\{v, Cv, C^2v, \ldots\}$.

For the companion matrix, we have

$$v = e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$Cv = Ce_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = e_2$$

$$C^2 v = Ce_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} = e_3$$

$$\vdots$$

$$C^{n-1} v = e_n$$

See that $\{e_1, e_2, \ldots, e_n\}$ are the standard basis vectors, which are linearly independent. Therefore, $\{v, Cv, C^2v, \ldots, C^{n-1}v\}$ are linearly independent and span $\mathbb{R}^n$. We then compute $C^n v$

$$C^n v = C \cdot e_n = \begin{pmatrix} -c_0 \\ -c_1 \\ -c_2 \\ \vdots \\ -c_{n-1} \end{pmatrix}$$

We can express this as a linear combination of the previous vectors

$$\begin{aligned} C^n v &= -c_0 e_1 - c_1 e_2 - c_2 e_3 - \cdots - c_{n-1} e_n \\ &= -c_0 v - c_1 Cv - c_2 C^2 v - \cdots - c_{n-1} C^{n-1} v \end{aligned}$$

$$C^n v + c_{n-1} C^{n-1} v + c_{n-2} C^{n-2} v + \cdots + c_1 Cv + c_0 v = 0$$

This linear dependence corresponds to the polynomial

$$\widetilde{\mu}(X) = X^n + c_{n-1} X^{n-1} + c_{n-2} X^{n-2} + \cdots + c_1 X + c_0 = f(X)$$

Since the Krylov subspace span$\{v, Cv, \ldots, C^{n-1}v\} = \mathbb{R}^n$, we don't need to find additional vectors. The minimal polynomial is

$$m_C(X) = \widetilde{\mu}(X) = f(X)$$

Therefore

$$p_C(X) = m_C(X) = f(X)$$

This completes the proof that for a companion matrix, the characteristic polynomial equals the minimal polynomial, and both equal the original polynomial $f(X)$ for which the companion matrix was constructed. □

Remark 3.10 (Computational limitations)**.** *Both the Leverrier–Faddeev algorithm for computing characteristic polynomials and Mertens' Krylov-based algorithm for minimal polynomials are primarily of theoretical interest and are practical only for small matrices.*

The Leverrier–Faddeev algorithm relies on repeated matrix multiplications. As a result, it is computationally expensive and quickly destroys sparsity. This makes it unsuitable for large or sparse matrices, and it may also suffer from numerical instability.

Mertens' algorithm is simpler, since it uses only matrix–vector products. However, in general one must test several starting vectors, and the Krylov vectors often become dense even when the original matrix is sparse. These factors limit its usefulness for large-scale computations.

In summary, while these algorithms are useful for understanding the structure of minimal and characteristic polynomials, they are not intended as efficient methods for large or sparse matrices, where modern numerical algorithms are more appropriate.

4. Conclusion

The relationship between the characteristic polynomial and the minimal polynomial of a matrix is a well-known result in linear algebra. In particular, it is already known that these two polynomials are equal if and only if each eigenvalue of the matrix has geometric multiplicity one, or equivalently, if the matrix has exactly one Jordan block for each eigenvalue. This result is classical and is usually proved using Jordan canonical form theory.

The goal of this paper was not to claim this result as new, but to study it from a different point of view. We focused on what happens structurally and algorithmically when the characteristic and minimal polynomials are equal. In particular, we provided an alternative proof of the classical equivalence then showed that when $m_A(X) = p_A(X)$, the minimal polynomial can be recovered from a single Krylov sequence generated by a cyclic vector. In this case, the sequence $\{v, Av, A^2v, \dots, A^nv\}$ spans the whole space and its first linear dependence occurs at step n. This makes Mertens' algorithm especially simple to apply. When this condition does not hold, multiple starting vectors are required, and the computation becomes more complicated.

We showed these ideas through two contrasting examples. For companion matrices, which are cyclic, we showed that starting with the cyclic vector 4_1 yields the minimal and characteristic polynomials immediately from a single Krylov sequence. For nilpotent matrices with nilpotency index $k < n$, which are not cyclic, we showed that multiple iterations of Mertens' algorithm are necessary, as no single vector generates a Krylov sequence spanning the entire space. Together, these examples highlight the insight of Proposition 3.4: when a matrix is cyclic and we start with a cyclic vector, the minimal polynomial can be recovered from a single Krylov sequence; otherwise, multiple starting vectors are required.

While the algorithms discussed here are not intended for large or sparse matrices, they provide a clear and useful framework for understanding the connection between matrix structure and algorithmic behavior.

Words of Appreciation.

Thank you to my academic advisors and my one-eyed cat.

References

[1] J. Abbott, A. M. Bigatti, E. Palezzato, and L. Robbiano, "Computing and Using Minimal Polynomials," *arXiv preprint arXiv:1702.07262*, 2017. doi: 10.1016/j.jsc.2019.07.022.
[2] K. Conrad, "The Minimal Polynomial and Some Applications," Conference Notes, 2008. Available: `https://api.semanticscholar.org/CorpusID:449089`
[3] R. Kaye and R. Wilson, *Linear Algebra.* Oxford University Press, UK, 1998.
[4] H. Woerdeman, *Advanced Linear Algebra.* CRC Press, Textbooks in Mathematics, 2015.

[5] N. Halidias, “On the Computation of the Minimum Polynomial and Applications,” *Asian Research Journal of Mathematics*, vol. 18, no. 11, pp. 301–319, 2022. doi: 10.9734/arjom/2022/v18i11603.

[6] M. H. Mertens, “The Minimal Polynomial,” Unpublished manuscript, Oct. 2015.

[7] K. Korkeathikhun, B. Khuhirun, S. Sriwongsa, and K. Wiboonton, “More on Characteristic Polynomials of Lie Algebras,” *arXiv preprint arXiv:2308.04618*, 2023. Available: `https://arxiv.org/abs/2308.04618`

[8] B. Jacob, *Linear Algebra*. Freeman, 1990.

[9] A. Magidin, “When are Minimal and Characteristic Polynomials the Same?” *Mathematics Stack Exchange*, Nov. 2011. Available: `https://math.stackexchange.com/q/81473`

[10] M. van Leeuwen, “Nilpotent Operator Minimal Polynomial,” *Mathematics Stack Exchange*, Jun. 2019. Available: `https://math.stackexchange.com/q/3247237`

[11] Y. Filmus, “The Characteristic and Minimal Polynomials of a Companion Matrix,” *Mathematics Stack Exchange*, May 2015. Available: `https://math.stackexchange.com/q/10217`

[12] J. Liesen and V. Mehrmann, “The Characteristic Polynomial and Eigenvalues of Matrices,” in *Linear Algebra*. Cham: Springer International Publishing, 2015, pp. 101–113. doi: 10.1007/978-3-319-24346-7_8.

[13] H. Fatoorehchi and M. Ehrhardt, “A Novel Technique for Stability Analysis of Linear Time-Invariant and Nonlinear Continuous-Time Control Systems Based on Markov Parameters and Hankel Matrices,” *Transactions of the Institute of Measurement and Control*, vol. 46, 2024. doi: 10.1177/01423312231216209.

[14] F. M. Dopico, J. Pérez, and P. Van Dooren, “Block Minimal Bases ℓ-ifications of Matrix Polynomials,” *Linear Algebra and its Applications*, vol. 562, pp. 163–204, 2019. doi: 10.1016/j.laa.2018.10.010.

[15] R. Dedekind and H. Weber, “Theorie der algebraischen Functionen einer Veränderlichen,” Walter de Gruyter, Berlin/New York, 1882.

[16] N. A. Loehr, *Advanced Linear Algebra*. CRC Press, Textbooks in Mathematics, 2024, p. 218.

[17] P. M. Neumann and C. E. Praeger, “Cyclic Matrices Over Finite Fields,” *Journal of the London Mathematical Society*, vol. 52, no. 2, pp. 263–284, 1995. doi: 10.1112/jlms/52.2.263.

[18] V. Bolotnikov, “Cyclic Matrices and Polynomial Interpolation over Division Rings,” *Linear Algebra and its Applications*, vol. 646, pp. 132–174, 2022. doi: 10.1016/j.laa.2022.03.030.