6-1-2017

A Polynomial in A of the Diagonalizable and Nilpotent Parts of A

Kathryn Wilson
Western Oregon University

A Polynomial in A of the Diagonalizable and Nilpotent Parts of A

By
Kathryn Wilson

An Honors Thesis Submitted in Partial Fulfillment of the
Requirements for Graduation from the
Western Oregon University Honors Program

Dr. Scott Beaver,
Thesis Advisor

Dr. Gavin Keulks,
Honor Program Director

June 2017

Acknowledgements

I would first like to thank my advisor, Dr. Scott Beaver. He has been an amazing resources throughout this entire process. Your support was what made this project possible. I would also like to thank Dr. Gavin Keulks for his part in running the Honors Program and allowing me the opportunity to be part of it. Further I would like to thank the WOU Mathematics department for making my education here at Western immensely fulfilling, and instilling in me the quality of thinking that will create my success outside of school. Lastly, to my parents for always supporting me.

Contents

Abstract

Any square matrix A can be decomposed into a sum of the diagonal (D_A) and nilpotent (N_A) parts as $A = D_A + N_A$. The components D_A and N_A commute with each other and with A. For many matrices A, B, if B commutes with A, then B is a polynomial in A; this holds for D_A and N_A. Following a Herbert A. Medina preprint, this paper shows how to construct the polynomials $p(A) = N_A$ and $q(A) = D_A$. Further, the Jordan canonical form J is a conjugate QAQ^{-1} of A; this paper demonstrates that the conjugation relating J and A also relates N_A and N_J and D_A and D_J, respectively.

1 Introduction

The contents of this paper are based on work one in a preprint by Herbert A. Medina [1]. We will focus on proving one theorem, and the bulk of the paper will consist of various lemmas that support the theorem. First, we introduce some basic definitions that will be used in the paper, all from [2], [4], or [5]. Then, we will put forth some lemmas and a theorem that will be referenced within multiple parts of the paper. The proofs for those may be found in the Appendix.

Definition 1.1 Basis *A basis of a vector space V is a list of vectors in V that are linearly independent and spans V.*

Definition 1.2 Linear Map *A linear map from V to W is a function $T : V \to W$ with the following properties:*

additivity $T(u + v) = Tu + Tv$ *for all* $u, v \in V$*;*

homogeneity $T(\lambda v) = \lambda T(v)$ *for all* $\lambda \in \mathbb{F}$ *and all* $v \in V$.

Definition 1.3 Linear Operator *The set of all linear maps such that $V = W$ is called a linear operator and denoted by $L(V)$.*

Definition 1.4 Jordan Basis *Suppose $T \in L(V)$. A basis of V is called a Jordan basis for T if with respect to this basis T has a block diagonal matrix*

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_p \end{pmatrix},$$

where each A_j is an upper triangular matrix of the form

$$A_j = \begin{pmatrix} \lambda_j & 1 & & 0 \\ & \ddots & \ddots & \\ & & & 1 \\ 0 & & & \lambda_j \end{pmatrix}.$$

Definition 1.5 Diagonalizable *An operator $T \in L(V)$ is diagonalizable if the operator has a diagonal matrix with respect to some basis of V.*

Definition 1.6 Nilpotent *An operator $T \in L(V)$ is called nilpotent if some positive power n of L equals the zero operator.*

Definition 1.7 Notation *For a matrix $A \in M_{n\times n}(\mathbb{C})$, a_{i*} represents the i^{th} row of A for all columns 1 through n. Also, a_{*j} represents the j^{th} column of A for all rows 1 through n.*

Definition 1.8 Polynomial *A polynomial is defined as*

$$p(x) = \sum_{k=0}^{n} a_k x^k = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0.$$

Definition 1.9 Inner Product *An inner product on V is a function that takes each ordered pair (u, v) of elements of V to a number $\langle u, v\rangle \in \mathbb{F}$ that has the following properties:*

positivity $\langle v, v\rangle \geq 0$ *for all* $v \in V$*;*

definiteness $\langle v, v\rangle = 0$ *if and only if* $v = 0$*;*

additivity in first slot $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ *for all* $u, v, w \in V$*;*

homogeneity in first slot $\langle \lambda u, v\rangle = \lambda\langle u, v\rangle$ *for all* $\lambda \in \mathbb{F}$ *and all* $u, v \in V$*;*

conjugate symmetry $\langle u, v\rangle = \overline{\langle v, u\rangle}$ *for all* $u, v \in V$*.*

Definition 1.10 Simultaneously Diagonalizable *Matrices $A, B \in M_{n\times n}(\mathbb{C})$ are considered simultaneously diagonalizable if there exists an invertible matrix P such that PAP^{-1} and PBP^{-1} are diagonal.*

Definition 1.11 Jordan Chevalley Decomposition *A Jordan Chevalley decomposition is an expression of the sum of a linear operator x as*

$$x = x_{ss} + x_n$$

where x_{ss} is the semisimple part of x and x_n is the nilpotent part of x. Furthermore, x_{ss} and x_n commute. If such a decomposition exists, it is unique.

Lemma 1.12 *If A, B, $C \in M_{n\times n}(\mathbb{C})$, then*
(i) $(A(B + C))_{ij} = (AB)_{ij} + (AC)_{ij}$
(ii) $((A + B)C)_{ij} = (AC)_{ij} + (BC)_{ij}$.

Lemma 1.13 *If A, $B \in M_{n\times n}(\mathbb{C})$ and $AB = BA$, then $Ap(B) = p(B)A$ where p is any polynomial of B.*

Lemma 1.14 *If A and B are nilpotent matrices and commute, then $A - B$ is also nilpotent.*

Theorem 1.15 *If A and B are $n \times n$ matrices, diagonalizable, and $AB = BA$, then A and B are simultaneously diagonalizable.*

These definitions, lemmas, and theorem provide context to the main theorem of this paper, which is the following:

Theorem 1.16 *Let $A \in M_{n\times n}(\mathbb{C})$ and let $A = D_A + N_A$ be a decomposition of A where D_A is diagonalizable, N_A is nilpotent and $N_A D_A = D_A N_A$. Then there exists polynomials $p(x)$, $q(x) \in P(C)$ such that $p(A) = N_A$ and $q(A) = D_A$. Moreover, N_A and D_A are unique.*

The first step in proving Theorem 1.16 is showing the existence of a polynomial, which we shall do by constructing said polynomial. We note that the polynomial we construct will only form the nilpotent part of the matrix. However, we know that by the Jordan Chevalley decomposition $A = D_A - N_A$, it follows that $D_A = A - N_A$. If we can find a polynomial such that $p(A) = N_A$, then $D_A = A - p(A)$. So, if we define $q(A) = D_A$, then simply finding $p(A)$ will show the existence of $q(A)$.

2 Initial Statements

Many linear algebra texts attest that the matrix A as defined in Theorem 1.16 can also be expressed in the form $A = QJQ^{-1}$, where $Q \in M(\mathbb{C})$ is invertible and J is a Jordan canonical matrix [3]. For the rest of this paper we shall represent a Jordan block J by:

$$J = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda_i & 1 & & \cdots & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ & & & & & 0 \\ & & & & \lambda_i & 1 \\ 0 & \cdots & & & 0 & \lambda_i \end{pmatrix}.$$

It is possible that the Jordan block can be a 1×1 with λ_i as the sole element of the matrix,

Multiple Jordan blocks can then form the Jordan canonical form. The Jordan canonical form is a matrix that consists of individual Jordan blocks along the the main diagonal of the matrix and zeros in all the other entries. Thus, we define the Jordan canonical form J as the following:

$$J = \begin{pmatrix} J^1_{\lambda_1} & 0 & \cdots & & & & & & & 0 \\ 0 & \ddots & & & & & & & & \vdots \\ \vdots & & J^{m_1}_{\lambda_1} & & & & & & & \\ & & & J^1_{\lambda_2} & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & J^{m_2}_{\lambda_2} & & & & \\ & & & & & & \ddots & & & \\ & & & & & & & J^1_{\lambda_q} & & \vdots \\ \vdots & & & & & & & & \ddots & 0 \\ 0 & \cdots & & & & & & \cdots & 0 & J^{m_q}_{\lambda_q} \end{pmatrix}.$$

where each $J^j_{\lambda_i}$ is an individual Jordan block. We define the individual Jordan blocks of a matrix J in Jordan Canonical form as $J^j_{\lambda_i}$, where λ_i is the diagonal value and $j \in \{1, 2, \ldots, m_i\}$ identifies the Jordan block with that value. We note that the j becomes vitally important when we have a Jordan canonical matrix, as a Jordan canonical matrix can have multiple Jordan blocks with the same diagonal. We also note that

$$
\begin{aligned}
J^j_{\lambda_i} &= \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda_i & 1 & & \cdots & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ & & & & & 0 \\ & & & & \lambda_i & 1 \\ 0 & \cdots & & & 0 & \lambda_i \end{pmatrix} \\
&= \underbrace{\begin{pmatrix} \lambda_i & 0 & \cdots & & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & & & & \\ & & & & 0 \\ 0 & \cdots & & 0 & \lambda_i \end{pmatrix}}_{D_J} + \underbrace{\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ & \ddots & \ddots & & \vdots \\ & & & & \\ \vdots & & & 0 & 1 \\ 0 & \cdots & & & 0 \end{pmatrix}}_{N_J}
\end{aligned}
$$

Furthermore, this shows that $J = D_J + N_J$. Next, we will show that D_J and N_J commute and are unique.

Lemma 2.1 *For a Jordan matrix J, the diagonalizable and nilpotent parts, D_J and N_J respectively, commute.*

Proof: Consider $e_j = (0 \cdots 0 \underbrace{1}_{j^{\text{th}}\text{position}} 0 \cdots 0)$. Thus, we can represent D_J and N_J as $D_J = \begin{pmatrix} \lambda_i e_1 \\ \lambda_i e_2 \\ \vdots \\ \lambda_i e_n \end{pmatrix}$ and $N_J = \begin{pmatrix} 0 & e_1^T & e_2^T & \cdots & e_{n-1}^T \end{pmatrix}$. We find when multiplying the vectors that,

$$
e_j e_i^T = \delta_{ij} = \begin{cases} 1 & j = i; \\ 0 & j \neq i. \end{cases}
$$

Thus,

$$
\begin{aligned}
D_J N_J &= \begin{pmatrix} \lambda_i e_1 \\ \lambda_i e_2 \\ \vdots \\ \lambda_i e_n \end{pmatrix} \times \begin{pmatrix} 0 & e_1^T & e_2^T & \cdots & e_{n-1}^T \end{pmatrix} \\
&= \begin{pmatrix} 0 & \lambda_i e_1 e_1^T & \lambda_i e_1 e_2^T & \lambda_i e_1 e_3^T & \cdots & \lambda_i e_1 e_{n-1}^T \\ 0 & \lambda_i e_2 e_1^T & \lambda_i e_2 e_2^T & \lambda_i e_2 e_3^T & \cdots & \lambda_i e_2 e_{n-1}^T \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & \lambda_i e_{n-1} e_1^T & \lambda_i e_{n-1} e_2^T & \lambda_i e_{n-1} e_3^T & \cdots & \lambda_i e_{n-1} e_{n-1}^T \\ 0 & \lambda_i e_n e_1^T & \lambda_i e_n e_2^T & \lambda_i e_n e_3^T & \cdots & \lambda_i e_n e_{n-1}^T \end{pmatrix}
\end{aligned}
$$

$$= \begin{pmatrix} 0 & \lambda_i & 0 & \cdots & 0 \\ 0 & 0 & \lambda_i & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

Next, the other side. We have $N_J = \begin{pmatrix} e^2 \\ e^3 \\ \vdots \\ e^n \\ 0 \end{pmatrix}$ and $D_J = \begin{pmatrix} \lambda_i e_1^T & \lambda_i e_2^T & \cdots & \lambda_i e_n^T \end{pmatrix}$. Thus,

$$N_J D_J = \begin{pmatrix} e^2 \\ e^3 \\ \vdots \\ e^n \\ 0 \end{pmatrix} \times \begin{pmatrix} \lambda_i e_1^T & \lambda_i e_2^T & \cdots & \lambda_i e_n^T \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_i e_2 e_1^T & \lambda_i e_2 e_2^T & \lambda_i e_2 e_3^T & \cdots & \lambda_i e_2 e_n^T \\ \lambda_i e_3 e_1^T & \lambda_i e_3 e_2^T & \lambda_i e_3 e_3^T & \cdots & \lambda_i e_3 e_n^T \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda_i e_{n-1} e_1^T & \lambda_i e_{n-1} e_2^T & \lambda_i e_{n-1} e_3^T & \cdots & \lambda_i e_{n-1} e_n^T \\ \lambda_i e_n e_1^T & \lambda_i e_n e_2^T & \lambda_i e_n e_3^T & \cdots & \lambda_i e_n e_n^T \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & \lambda_i & 0 & \cdots & 0 \\ 0 & 0 & \lambda_i & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

As we can see, $D_J N_J = N_J D_J$. ■

Lemma 2.2 *For a Jordan matrix J, the diagonalizable and nilpotent parts, D_J and N_J respectively, are unique.*

Proof: We proceed by contradiction. Suppose there exists $D_J^{'}$ and $N_J^{'}$ such that $J = D_J^{'} + N_J^{'}$. Since $J = D_J + N_J$, we have

$$D_J + N_J = D_J^{'} + N_J^{'} \Longrightarrow D_J - D_J^{'} = N_J^{'} - N_J. \tag{1}$$

By Theorem 1.15 and since diagonal matrices commute we know there exists an invertible matrix P such that PD_JP^{-1} and $PD_J^{'}P^{-1}$ are diagonal. So, we find that

$$PD_JP^{-1} - PD_J^{'}P^{-1} = PN_J^{'}P^{-1} - PN_JP^{-1}.$$

By Lemma 1.14, $PN_J^{'}P^{-1} - PN_JP^{-1}$ is nilpotent. Therefore, $PD_JP^{-1} - PD_J^{'}P^{-1}$ is diagonal and nilpotent, $PD_JP^{-1} - PD_J^{'}P^{-1} = 0$, and hence $D_J = D_J^{'}$. Therefore, by (1), $N_J^{'} = N_J$. It follows that

D_J and N_J are unique. ∎

Primarily the strategy for proving Theorem 1.16 will involve finding $p(x)$, where $p(x)$ is the polynomial representation of N_A, because that will make it simple to find $q(x) = x - p(x)$ as we elaborated on earlier. The first step is to actually search for the polynomial representation of N_J. Later, we will show that $N_A = QN_JQ^{-1}$. We know that $A = QJQ^{-1}$ and $J = D_J + N_J$, and will go from there. But first, we begin by showing that $p(x)$ exists.

3 Construction of the Polynomial

In order to construct the polynomial that represents the nilpotent part of the matrix J, we assess the individual $J^j_{\lambda_i}$ for $j \in \{1, 2, \ldots, m_i\}$. The final polynomial that represents N_J, which we call $p(x)$, is constructed as a sum of polynomials each corresponding to the unique λ_i, which we call $p_i(x)$. Since each λ_i can have multiple Jordan blocks, it follows that the polynomial corresponding to an individual λ_i will also involve multiple polynomials that will be computed into one larger all-encompassing polynomial. These individual pieces will be called $p_{i,h}(x)$ where i identifies the λ that forms the main diagonal of the Jordan block the polynomial is manipulating and h denotes which term of $p_i(x)$ it is.

In order to better visualize and understand the steps involved in this process, consider the following example:

$$J = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 \end{pmatrix}.$$

As we can see, there are four Jordan blocks. We shall set the first 4×4 block with the 2's along the diagonal as the centerpiece of the $p_i(x)$. If we consider the labeling process of the $J^j_{\lambda_i}$, in the example $\lambda_i = 2$, essentially. As a note, when referring to the general case for the polynomial, it will be labeled $p(x)$. The specific example will use $p(J)$.

To start the general construction of $p_i(x)$, we first want to find $p_{i,1}(x)$. To begin, we focus on eliminating the Jordan blocks pertaining to the other λ_r where $r \neq i$. In order to do this, we use the factorization of the characteristic polynomial of the $J^j_{\lambda_r}$. So, we will have $q-1$ factors of $(x-\lambda_r)^{m_r}$ where m_r is the size of the largest Jordan block of λ_r and q is the number of unique λ_r. Within the context of the example, that means we will have $(J-3)^3(J-4)$ as the first few factors of the construction. Note, we shall use a $\sim$ to denote the association between the polynomial and its unfinished construction. Only when the polynomial is complete can we use an equals sign. Thus,

$$p_{i,1}(J) \sim (J-3)^3(J-4)$$

and we can create a general case for the first part of $p_{i,1}(x)$, where

$$p_{i,1}(x) \sim (x-\lambda_2)^{m_2} \cdots (x-\lambda_r)^{m_r}.$$

Doing those calculations gives us:

$$p_{i,1}(J) \sim (J-3)^3(J-4) = \begin{pmatrix} 2 & -7 & 9 & -5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & -7 & 9 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -7 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

As was mentioned, we focus on the largest block of λ_i. The calculations for any smaller blocks of λ_i would be affected by the calculations in the exact same way as the largest block, just applied to a smaller matrix. This is proven by the following lemma.

Lemma 3.1 Same k^{th} Diagonal *Let $A = (a_{ij})$, $B = (b_{ij})$ be $n \times n$ upper triangular matrices such that $a_{ij} = a_{km}$ and $b_{ij} = b_{km}$ whenever $j - i = k - m$. The $\{a_{ij} | j - i = k - m\}$, $\{a_{km} | j - i = k - m\}$, $\{b_{ij} | j - i = k - m\}$, and $\{b_{km} | j - i = k - m\}$ form the diagonals of their respective matrices. Moreover, the elements in the diagonals are constant. We then find that $AB = (c_{ij})$ also has the property that $c_{ij} = c_{km}$ whenever $j - i = m - k$. Further, if $P = (p_{ij})$, $Q = (q_{ij})$ are upper triangular $m \times m$ matrices, such that $m \leq n$ and the k^{th} diagonal of P is the same as the k^{th} diagonal of A, just adjusted for the size P and the k^{th} diagonal of Q is the same as the k^{th} diagonal of B, adjusted for the size of Q, then the k^{th} diagonal of $(AB) = k^{th}$ diagonal of (PQ).*

First an example to hopefully make the lemma clearer: Let

$$A = \begin{pmatrix} 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \text{ and } B = \begin{pmatrix} 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \implies AB = C = \begin{pmatrix} 0 & 0 & 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Further, if we have $A' = \begin{pmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ and $B' = \begin{pmatrix} 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \implies A'B' = C' = \begin{pmatrix} 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

Now, on to the proof.

Proof: Let $A = (a_{ij})$, $B = (b_{ij})$ be $n \times n$ upper triangular matrices. Let $j \geq i$ and $m \geq k$ with $j - i = m - k$. Suppose that $a_{ij} = a_{km}$ and $b_{ij} = b_{km}$.

$$AB = c_{ij} = \sum_{r=i}^{j} a_{ir} b_{rj}$$

Note, since we set $j - i = m - k$, when $r = j$, we find that $r - i = m - k$ when $m = k + r - i$. Also, when r is in the place of i, we find that $j - r = m - k$ when $k = r - j + m$. It follows that $a_{ir} = a_{k(r-i+k)}$ and

$b_{rj} = b_{(j-r+m)m}$. Therefore, by substitution

$$c_{ij} = \sum_{r=i}^{j} a_{k(r-i+k)} b_{(r-j+m)j}. \tag{2}$$

Let $s = r - i + k$. When $r = i$, $s = i - i + k = k$; and when $r = j$, $s = j - i + k = m - k + k = m$, since we previously noted that $j - i = m - k$. Also, since $s = r - i + k$, we know $r = s + i - k$. Hence, $r - j + m = s + i - k - j + m = s - (j - i) + (m - k) = s$. Thus we can substitute s in place of r in (2). Therefore,

$$c_{ij} = \sum_{s=k}^{m} a_{ks} b_{sm} = c_{km}. \tag{3}$$

Equation (3) shows that the elements in the product matrix only depend on the values in the diagonal entries and not on the size of the matrix. This is because the *kth* diagonals are the same when $j - i = m - k$, but that doesn't not necessarily mean that $m = j$ and $i = k$. ■

Back to the example, since we want the nilpotent part of the matrix, to keep the diagonal as all 1's, we will need to divide the polynomial by the product of each of the $(\lambda_i - \lambda_r)^{m_r}$. So, in the example we multiply by 1/2. For the general case we have:

$$p_{i,1}(x) \sim \frac{(x - \lambda_2)^{m_2} \cdots (x - \lambda_m)^{m_r}}{(\lambda_i - \lambda_2)^{m_2} (\lambda_i - \lambda_r)^{m_r}}.$$

Thus, we have:

$$p_{i,1}(J) \sim \frac{(J-3)^3(J-4)}{(2-3)^3(2-4)} = \begin{pmatrix} 1 & -7/2 & 9/2 & -5/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -7/2 & 9/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -7/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Once that is done, all that is left is the 4×4 block that we are focusing on. We want to shift the diagonal of 1's into the superdiagonal. We do this by multiplying $(x - \lambda_i)$, which is $(J - 2)$ in the example, by the prior construction of $p_{i,1}(x)$. Thus, the final polynomial is

$$p_{i,1}(x) = \frac{(x - \lambda_i)(x - \lambda_1)^{m_1} \cdots (x - \lambda_r)^{m_r}}{(\lambda_i - \lambda_1)^{m_1} \cdots (\lambda_i - \lambda_r)^{m_r}}.$$

And for the example we have:

$$p_{i,1}(J) = \frac{(J-2)(J-3)^3(J-4)}{(2-3)^3(2-4)} = \begin{pmatrix} 0 & 1 & -7/2 & 9/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -7/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

However, doing this will create nonzero numbers in the entries above the diagonal. These nonzero entries in the superdiagonals will be in constant, which makes it rather simple to get rid of them. We use another polynomial to get rid of those numbers. Note that the current polynomial creates 1's in the diagonal. If we are able to shift that diagonal up, we then simply need to multiply it by the value we want to subtract. To do so, this second polynomial we create will be subtracted from the first. Hence, we begin to construct $p_{i,2}(x)$.

This second polynomial, called $p_{i,2}(x)$, will shift the Jordan block corresponding to λ_i up. So, we begin construction by taking the first polynomial and multiplying by another $(J - \lambda_i)$. This has the effect of shifting the elements of the matrix up one diagonal. For the general case, we have:

$$p_{i,2}(x) = \frac{(x-\lambda_i)^2(x-\lambda_1)^{m_1}\cdots(x-\lambda_r)^{m_r}}{(\lambda_i-\lambda_1)^{m_1}\cdots(\lambda_i-\lambda_r)^{m_r}}.$$

And for the example we have:

$$p_{i,2}(J) = \frac{(J-2)^2(J-3)^3(J-4)}{(2-3)^3(2-4)} = \begin{pmatrix} 0 & 0 & 1 & -7/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

We will prove this using the following lemma.

Lemma 3.2 Shifting Matrices *Let* $N \in M_{n\times n}(\mathbb{C})$ *such that* $N = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0\cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & & 0 & 1 \\ 0 & \cdots & & 0 & 0 \end{pmatrix}$.

(i) For $k \geq 1$,

$$(N^k)_{ij} = \begin{cases} 1 & j-i=k; \\ 0 & \textit{otherwise.} \end{cases}$$

(ii.) Let $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ *be any* $n \times n$ *matrix. Then* $N^k A = \begin{pmatrix} a_{(k+1)1} & \cdots & a_{(k+1)n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \\ 0 & \cdots & 0 \\ \vdots & \cdots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$.

Proof of (i): Let $N = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0\cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & & 0 & 1 \\ 0 & \cdots & & 0 & 0 \end{pmatrix}$ be an $n \times n$ matrix. We proceed by induction on k.

Base Case: Let $k = 1$. We can see that by construction of N, the 1's are in the positions where $j - i = 1$. Thus, the statement holds for $k = 1$.

Induction Hypothesis: Assume that for $k \geq 1$,

$$N^k_{ij} = \begin{cases} 1 & j - i = k; \\ 0 & \text{otherwise.} \end{cases}$$

We want to show that

$$N^{k+1}_{ij} = \begin{cases} 1 & j - i = k+1; \\ 0 & \text{otherwise.} \end{cases}$$

We begin by noting that $N^{k+1} = NN^k$. Consider $e_j = (0 \cdots 0 \underbrace{1}_{j^{\text{th}}\text{position}} 0 \cdots 0)$. Thus, we can represent N and N^k as

$$N = \begin{pmatrix} e_2 \\ e_3 \\ \vdots \\ e_n \\ 0 \end{pmatrix}$$

and

$$N^k = \begin{pmatrix} 0 & \cdots & 0 & e_1^T & e_2^T & \cdots & e_{n-k}^T \end{pmatrix}$$

When multiplying vectors $e_j e_i^T$, we find

$$e_j e_i^T = \delta_{ij} = \begin{cases} 1 & j = i; \\ 0 & j \neq i. \end{cases}$$

It follows that

$$N^{k+1} = NN^k = \begin{pmatrix} e_2 \\ e_3 \\ \vdots \\ e_n \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 & \cdots & 0 & e_1^T & e_2^T & \cdots & e_{n-k}^T \end{pmatrix}$$

$$= \begin{pmatrix} 0 & \cdots & 0 & e_2e_1^T & e_2e_2^T & e_2e_3^T & \cdots & e_2e_{n-k}^T \\ 0 & \cdots & 0 & e_3e_1^T & e_3e_2^T & e_3e_3^T & \cdots & e_3e_{n-k}^T \\ \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & e_{n-k}e_1^T & e_{n-k}e_2^T & e_{n-k}e_3^T & \cdots & e_{n-k}e_{n-k}^T \\ \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & e_ne_1^T & e_ne_2^T & e_ne_3^T & \cdots & e_ne_{n-k}^T \end{pmatrix}$$

$$= \begin{pmatrix} 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 1 \\ \vdots & \ddots & & \ddots & \ddots & \ddots & \ddots & \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

We find that

$$N^{k+1} = NN^k = \begin{pmatrix} 0 & \cdots & 0 & \underbrace{0}_{j^{\text{th}}\text{position}} & \underbrace{e_1^T}_{(j+1)^{\text{st}}\text{column}} & e_2^T & \cdots & e_{n-k-1}^T \end{pmatrix}$$

.

Thus, the e_j^T in N^k, are in the e_{j-1}^T column we find that NN^k. Therefore, for N^{k+1}, $k = j-1-i$ and $k+1 = j-i$. In conclusion,

$$N_{ij}^k = \begin{cases} 1 & j-i=k; \\ 0 & \text{otherwise.} \end{cases}$$

holds for all $k \geq 1$. ■

Proof of (ii): Let $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ be any $n \times n$ matrix and

$$(N^k)_{ij} = \begin{cases} 1 & j-i=k; \\ 0 & \text{otherwise.} \end{cases}$$

Representing N^k with the vectors we defined in the part (i), we have

$$N^k = \begin{pmatrix} e_{k+1} \\ e_{k+2} \\ \vdots \\ e_n \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

For readability, we also have

$$A = \begin{pmatrix} a_{*1} & a_{*2} & \cdots & a_{*n} \end{pmatrix}.$$

Note that $e_j = (0 \cdots 0 \underbrace{1}_{j^{\text{th}}\text{position}} 0 \cdots 0)$ and $a_{*m} = \begin{pmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{pmatrix}$. It follows that $e_j a_{*m} = a_{jm}$.

Then,

$$
\begin{aligned}
N^k A &= \begin{pmatrix} e_{K+1} \\ e_{k+2} \\ \vdots \\ e_n \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{pmatrix} a_{*1} & a_{*2} & \cdots & a_{*n} \end{pmatrix} \\
&= \begin{pmatrix} e_{k+1}a_{*1} & e_{k+1}a_{*2} & \cdots & e_{k+1}a_{*n} \\ e_{k+2}a_{*1} & e_{k+2}a_{*2} & \cdots & e_{k+2}a_{*n} \\ \vdots & & & \vdots \\ e_n a_{*1} & e_n a_{*2} & \cdots & e_n a_{*n} \\ 0 & \cdots & 0 & \\ \vdots & \cdots & \vdots & \\ 0 & \cdots & 0 & \end{pmatrix} \\
&= \begin{pmatrix} a_{(k+1)1} & \cdots & a_{(k+1)n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \\ 0 & \cdots & 0 \\ \vdots & \cdots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}
\end{aligned}
$$

In conclusion, $N^k A = \begin{pmatrix} a_{(k+1)1} & \cdots & a_{(k+1)n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \\ 0 & \cdots & 0 \\ \vdots & \cdots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$. ■

Back to the example, we now have 1's in the diagonal we want to get rid of. So, the next step is to multiply the matrix by a number so that the constant diagonal of 1's will be equal to the element in the first upper diagonal of the first polynomial formed matrix. Note, we call these $d_i(k)$ where the i identifies which λ_i the polynomial is focusing on, and k the diagonal the $d_i(k)$ is located in. In the example, $d_i(1) = -7/2$. So, we multiply $p_{i,2}(x)$ by $d_i(1)$ and then subtract that product from $p_{i,1}(x)$. We continue this process for as many $p_{i,h}(x)$ are present.

Then, we would repeat this process for the rest of the $J_{\lambda_i}^j$. Finally, we would add all the $p_i(x)$

together to get $p(x)$. For the example, we eventually find that:

$$p_1(J) = \frac{(J-2)(J-3)^3(J-4) + (7/2)[(J-2)^2(J-3)^3(J-4)] + (31/4)[(J-2)^3(J-3)^3(J-4)]}{(2-3)^3(2-4)}.$$

Further, using the construction we find:

$$p_2(J) = \frac{(J-3)(J-2)^4(J-4) - (3)[(J-3)^2(J-2)^4(J-4)]}{(3-2)^4(3-4)}.$$

Since the last Jordan block is a 1×1, it doesn't have any 1's in the diagonal. Thus, $p_3(J) = 0$.
In conclusion,

$$p(J) = \frac{(J-2)(J-3)^3(J-4) + (7/2)[(J-2)^2(J-3)^3(J-4)] + (31/4)[(J-2)^3(J-3)^3(J-4)]}{(2-3)^3(2-4)}$$
$$+ \frac{(J-3)(J-2)^4(J-4) - (3)[(J-3)^2(J-2)^4(J-4)]}{(3-2)^4(3-4)}.$$

We currently do not have a generalized way of finding $d_i(k)$. However, we do have a way of extrapolating that there exists a way to generalize them with the following lemma.

Lemma 3.3 Entries in Further Diagonals *For each i, k,*

$$\frac{(J^i_{\lambda_1} - \lambda_k)^{s_k}}{(\lambda_1 - \lambda_k)^{s_k}} = \begin{pmatrix} 1 & b_{s_k}(1) & b_{s_k}(2) & \cdots & \cdots & b_{s_k}(m-2) & b_{s_k}(m-1) \\ 0 & 1 & b_{s_k}(1) & \cdots & \cdots & \cdots & b_{s_k}(m-2) \\ \vdots & & \ddots & \ddots & & & \vdots \\ & & & & 1 & b_{s_k}(1) & b_{s_k}(2) \\ \vdots & & & & & 1 & b_{s_k}(1) \\ 0 & \cdots & & & & 0 & 1 \end{pmatrix}$$

where

$$b_{s_k}(j) = \begin{cases} \frac{\binom{s_k}{j}}{(\lambda_1 - \lambda_k)^j} & j \leq s_k; \\ 0 & j > s_k. \end{cases}$$

We shall also provide an example for this lemma. Referring back to the matrix used in the broad explanation of how to form the polynomial $p(x)$, we shall focus on the Jordan block J_2 with the λ_k corresponding to the 3's in the diagonal of the 3×3 block. So,

$$\frac{(J_2 - 3I)^3}{(2-3)^3} = \begin{pmatrix} 1 & -3 & 3 & -1 \\ 0 & 1 & -3 & 3 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Hence, $\frac{\binom{3}{1}}{(2-3)^1} = -3$, $\frac{\binom{3}{2}}{(2-3)^2} = 3$, and $\frac{\binom{3}{3}}{(2-3)^3} = -1$ as desired.
Proof: We argue by induction on s_k.

Base Case: Let $s_k = 1$. Therefore, we consider the first $\frac{(J^i_{\lambda_1} - \lambda_k)}{(\lambda_1 - \lambda_k)}$. Note that

$$J^i_{\lambda_1} = \begin{pmatrix} \lambda_1 & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & 0 \\ & & & \lambda_1 & 1 \\ 0 & \cdots & & 0 & \lambda_1 \end{pmatrix}.$$

Thus,

$$J^i_{\lambda_1} - \lambda_k = \begin{pmatrix} \lambda_1 - \lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & 0 \\ & & & \lambda_1 - \lambda_k & 1 \\ 0 & \cdots & & 0 & \lambda_1 - \lambda_k \end{pmatrix}$$

Furthermore,

$$\frac{(J^i_{\lambda_1} - \lambda_k)}{(\lambda_1 - \lambda_k)} = \begin{pmatrix} 1 & \frac{1}{\lambda_1 - \lambda_k} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & 0 \\ & & & 1 & \frac{1}{\lambda_1 - \lambda_k} \\ 0 & \cdots & & 0 & 1 \end{pmatrix}$$

Since $\binom{1}{1} = 1$, we conclude that when $s_k = 1$ the statement holds.
Induction Hypothesis: Assume the result for $s_k = \ell$, where $\ell \geq 1$.

$$\frac{(J^i_{\lambda_1} - \lambda_k)^{\ell+1}}{(\lambda_1 - \lambda_k)^{\ell+1}} = \begin{pmatrix} 1 & \frac{1}{\lambda_1 - \lambda_k} & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \frac{1}{\lambda_1 - \lambda_k} & 0 & \cdots & 0 \\ \vdots & \ddots & & \ddots & \ddots & \vdots \\ & & & & 1 & \frac{1}{\lambda_1 - \lambda_k} \\ 0 & \cdots & & & 0 & 1 \end{pmatrix} \times$$

$$\begin{pmatrix} 1 & b_\ell(1) & b_\ell(2) & \cdots & \cdots & b_\ell(m-2) & b_\ell(m-1) \\ 0 & 1 & b_\ell(1) & \cdots & \cdots & \cdots & b_\ell(m-2) \\ \vdots & & \ddots & \ddots & & & \vdots \\ & & & & 1 & b_\ell(1) & b_\ell(2) \\ & & & & & 1 & b_\ell(1) \\ 0 & \cdots & & & & 0 & 1 \end{pmatrix}$$

The main diagonal of the resulting product is clear, as it still is a diagonal of $1's$. We find, that products when not the diagonal, take the form of $b_l(n) + \frac{b_l(n-1)}{\lambda_i - \lambda_k}$. Note that for $j \geq i$, where j denotes the b_l being used and i is the size of the Jordan block of λ_i, the i,j component of the matrix for the case $l = s_k$ is $b_l(j-i)$. Therefore, we have to show that the i,j component of the matrix resulting from the product of the above is $b_{l+1}(j-i)$. The i,j component in the product is $1 * b_l(j-i) + \frac{1}{\lambda_i - \lambda_k} * b_l(j-i-1)$. This expression is computed in three cases:

$$\begin{cases} \frac{\binom{l}{j-i}}{(\lambda_1-\lambda_k)^{j-i}} + \frac{1}{(\lambda_1-\lambda_k)} \frac{\binom{l+1}{j-i-1}}{(\lambda_1-\lambda_k)^{j-i-1}} = \frac{\binom{l+1}{j-i}}{(\lambda_1-\lambda_k)^{j-i}} = b_{l+1}(j-i) & j-i \le l; \\ 0 + \frac{1}{(\lambda_1-\lambda_k)} \frac{\binom{l+1}{j-i-1}}{(\lambda_1-\lambda_k)^{j-i-1}} = \frac{\binom{l+1}{j-i}}{(\lambda_1-\lambda_k)^{j-i}} = b_{l+1}(j-i) & j-i \le l+1; \\ 0 = b_{l+1}(j-i) & j-i \ge l+2. \end{cases} \blacksquare$$

To summarize, we have created a polynomial where:

$$p_{i,1}(J^j_{\lambda_i}) = \begin{pmatrix} 0 & 1 & d_i(1) & d_i(2) & \cdots & \cdots & d_i(m-2) & d_i(m-1) \\ 0 & 0 & 1 & d_i(1) & \cdots & \cdots & \cdots & d_i(m-2) \\ \vdots & \vdots & & \ddots & \ddots & & & \vdots \\ & & & & & & & \\ & & & & & 1 & d_i(1) & d_i(2) \\ \vdots & \vdots & & & & & 1 & d_i(1) \\ & & & & & & 0 & 1 \\ 0 & 0 & \cdots & & & & \cdots & 0 \end{pmatrix}$$

and

$$p_{i,2}(J^j_{\lambda_i}) = \begin{pmatrix} 0 & 0 & 1 & d_i(1) & d_i(2) & \cdots & \cdots & d_i(m-2) & d_i(m-2) \\ 0 & 0 & 0 & 1 & d_i(1) & \cdots & \cdots & \cdots & d_i(m-3) \\ \vdots & \vdots & & & \ddots & \ddots & & & \vdots \\ & & & & & & & & \\ & & & & & & & 1 & d_i(1) \\ \vdots & \vdots & & & & & & & 1 \\ & & & & & & & & 0 \\ 0 & 0 & \cdots & & & & & \cdots & 0 \end{pmatrix}.$$

So, broadly speaking, the general case of the polynomial is the following:

$$p_{i,k}(x) = \frac{(x-\lambda_i)^k (x-\lambda_1)^{m_1} \cdots (x-\lambda_r)^{m_r}}{(\lambda_i-\lambda_1)^{m_1} \cdots (\lambda_i-\lambda_r)^{m_r}}.$$

Furthermore, we define $p_1(x)$ as

$$p_1(x) = p_{i,1}(x) - \alpha_1(2)p_{i,2}(x) - \alpha_1(3)p_{i,3}(x) - \cdots - \alpha_1(m_1-1)p_{i,(m_1-1)}(x)$$

where α_1 is a function defined recursively by

$$\begin{aligned} \alpha_1(2) &= d_1(1) \\ \alpha_1(3) &= -\alpha_1(2)d_1(1) + d_1(2) \\ &\vdots \end{aligned}$$

$$\alpha_1(n) = -\alpha_1(n-1)d_1(1) - \alpha_1(n-2)d_1(1) - \cdots - -\alpha_1(2)d_1(n-1) + d_1(n)$$

In conclusion, we let

$$p(x) = \sum_{k=1}^{q} p_k(x).$$

So, we have constructed a polynomial representing the $p(x)$ parts of the matrix. Note that since $J = D_J + N_J$, we know $D_J = J - N_J$. Since $p(J) = N_J$, we also have shown the existence of $q(J)$ which would represent D_J.

To conclude Theorem 1.16, we want to show that $D_A = QD_JQ^{-1}$ and $N_A = QN_JQ^{-1}$. We use the following two lemmas to do so:

Lemma 3.4 *Let $A \in M_{n\times n}(\mathbb{C})$ and let J be its Jordan canonical form such that $A = QJQ^{-1}$. Let $J = D_J + N_J$ be the Jordan Chevalley decomposition of J into the sum of its diagonal and nilpotent parts. If $p(x)$ is a polynomial such that such that $p(J) = N_J$, then $p(A) = QN_JQ^{-1}$.*

Proof: Let $A \in M_{n\times n}(\mathbb{C})$, where J is the Jordan canonical form. Assume $A = QJQ^{-1}$ and that $J = D_J + N_J$. Also, suppose there exists $p(x)$ such that $p(J) = N_J$. So,

$$p(A) = p(QJQ^{-1}).$$

We note by Lemma 1.12 that it follows that $p(A) = Qp(J)Q^{-1}$. So, $p(A) = QN_JQ^{-1}$ as desired. ■

To make the final step to $N_A = QN_JQ^{-1}$, we have the following.

Lemma 3.5 *Let $A \in M_{n\times n}(\mathbb{C})$ and let J be its Jordan canonical form such that $A = QJQ^{-1}$.Let $J = D_J + N_J$ be J's Jordan Chevalley decomposition into the sum of a diagonal matrix D_J and a nilpotent matrix N_J. Suppose that $A = D_A + N_A$ is the Jordan Chevalley decomposition of A where D_A is diagonalizable, N_A is nilpotent, and $D_AN_A = N_AD_A$. Then $D_A = QD_JQ^{-1}$ and $N_A = QN_JQ^{-1}$.*

Proof: Let $A \in M_{n\times n}(\mathbb{C})$ and let J be its Jordan canonical form such that $A = QJQ^{-1}$. Let $J = D_J + N_J$ be J's decomposition into the sum of a diagonal matrix D_J and a nilpotent matrix N_J. Suppose that $A = D_A + N_A$ is a decomposition of A where D_A is diagonalizable, N_A is nilpotent, and $D_AN_A = N_AD_A$. By assumption and Lemma 1.12, we know that

$$A = QJQ^{-1} = Q(D_J + N_J)Q^{-1} = QD_JQ^{-1} + QN_JQ^{-1}$$

By Lemma 3.4 we can say there exist polynomials $p(x)$ and $q(x)$ such that $p(x) = QD_JQ^{-1}$ and $q(x) = QN_JQ^{-1}$. By Lemma 1.13, the fact that D_A and N_A commute with A, we conclude that D_A commutes with QD_JQ^{-1} and N_A commutes with QN_JQ^{-1}. By Lemma 1.15 we know there exists an invertible matrix P such that PD_AP^{-1} and $P(QD_JQ^{-1})P^{-1}$ are diagonal. So, we find that

$$PAP^{-1} = PD_AP^{-1} + PN_AP^{-1} = P(QD_JQ^{-1})P^{-1} + P(QN_JQ^{-1})P^{-1}$$

From this we obtain the key relation

$$PD_AP^{-1} - P(QD_JQ^{-1})P^{-1} = P(QN_JQ^{-1})P^{-1} - PN_AP^{-1}$$

Because QN_JQ^{-1} and N_A commute, by Lemma 1.14, $P(QN_JQ^{-1})P^{-1} - PN_AP^{-1}$ is nilpotent. Hence, $PD_AP^{-1} - P(QD_JQ^{-1})P^{-1}$ is both diagonal and nilpotent; therefore it must be zero. Therefore, $D_A = QD_JQ^{-1}$ and $N_A = QN_JQ^{-1}$. ■

Thus concludes the proof of Theorem 1.16. We have shown the existence of a polynomial representation of N_J, then showed that $N_A = QN_JQ^{-1}$. We also know that Jordan-Chevalley decompositions are unique, so D_A and N_A are unique. Therefore, we have proven Theorem 1.16.

A Appendix

Lemma 1.12 If A, B, and $C \in M_{n\times n}(\mathbb{C})$, then
(i) $(A(B+C))_{ij} = (AB)_{ij} + (AC)_{ij}$
(ii) $((A+B)C)_{ij} = (AC)_{ij} + (BC)_{ij}$

Proof: Let A, B, and $C \in M_{n\times n}(\mathbb{C})$. Let $i, j \in Z$ such that $1 \le i, j \le n$. Note, when multiplying matrices, by the definition of the inner product,

$$(AB)_{ij} = \langle \overline{A}_{i*}, B_{*j} \rangle$$

Thus, by matrix multiplication and additivity of the inner product:

$$\begin{aligned}(A(B+C))_{ij} &= \langle \overline{A}_{i*}, (B+C)_{*j} \rangle \\ &= \langle \overline{A}_{i*}, B_{*j} \rangle + \langle \overline{A}_{i*}, C_{*j} \rangle \\ &= (AB)_{ij} + (AC)_{ij}\end{aligned}$$

Hence, we have shown (i). Next, by matrix multiplication, conjugate symmetry, and additivity of the inner product:

$$\begin{aligned}((A+B)C)_{ij} &= \langle (\overline{A}_{i*} + \overline{B}_{i*}), C_{*j} \rangle \\ &= \overline{\langle C_{*j}, (\overline{A}_{i*} + \overline{B}_{i*}) \rangle} \\ &= \overline{\langle C_{*j}, \overline{A}_{i*} \rangle + \langle C_{*j}, \overline{B}_{i*} \rangle} \\ &= \langle \overline{A}_{i*}, C_{*j} \rangle + \langle \overline{B}_{i*}, C_{*j} \rangle \\ &= (AC)_{ij} + (BC)_{ij}\end{aligned}$$

Therefore, we have shown (ii). In conclusion, if A, B, and $C \in M_{n\times n}(C)$, then $(A(B+C))_{ij} = (AB)_{ij} + (AC)_{ij}$ and $((A+B)C)_{ij} = (AC)_{ij} + (BC)_{ij}$. ■

Lemma 1.13 If A, $B \in M_{n\times n}(\mathbb{C})$ and $AB = BA$, then $Ap(B) = p(B)A$.

Proof: Let A, $B \in M_{n\times n}(\mathbb{C})$. Suppose $AB = BA$. We know by the definition of a polynomial that $p(B) = c_0 + c_1B + c_2B^2 + ... + c_mB^m$ such that $c_i \in Z$ where $i = 0, 1, 2, ..., m$. Thus, by Lemma 1.1 and $AB = BA$,

$$\begin{aligned}Ap(B) &= A(c_0 + c_1B + c_2B^2 + ... + c_mB^m) \\ &= Ac_0 + Ac_1B + Ac_2B^2 + ... + Ac_mB^m \\ &= c_0A + c_1AB + c_2AB^2 + ... + c_mAB^m\end{aligned}$$

$$
\begin{aligned}
&= c_0A + c_1BA + c_2B^2A + ... + c_mB^mA \\
&= (c_0 + c_1B + c_2B^2 + ... + c_mB^m)A \\
&= p(B)A
\end{aligned}
$$

Therefore, as desired when A, $B \in M_{n\times n}(C)$ and $AB = BA$, then $Ap(B) = p(B)A$. ■

Lemma 1.14 If A and B are nilpotent matrices and commute, then $A - B$ is also nilpotent.

Proof: Let $A, B \in M_{n\times n}(\mathbb{C})$ and be nilpotent. By definition, $\exists\ n, m \in \mathbb{Z}$ such that $A^n = 0$ and $B^m = 0$. Without loss of generality, let $n \leq m$. Let $k = n + m - 1$ and $i = \{1, 2, \ldots, k-1\}$.Thus, using the binomial theorem, we find that $(A-B)^k = A^k + c_1A^{k-1}B^1 + c_2A^{k-2}B^2 + \cdots + c_{m-1}A^{k-(m-1)}B^{m-1} + c_mA^{k-m}B^m + c_{m+1}A^{k-(m+1)}B^{m+1} + \cdots + c_{k-1}A^1B^{k-1} + B^k$
where $c_i \in \mathbb{Z}$. Note that $k - (m-1) = n$. By this fact and the assumption we can conclude that

$$c_{m-1}A^{k-(m-1)}B^{m-1} = 0$$

$$c_mA^{k-m}B^m = 0$$

$$c_{m+1}A^{k-(m+1)}B^{m+1} = 0.$$

Furthermore, it follows that every other term of $(A-B)^k$ is equal to 0. Hence, $(A-B)^k = 0$. Therefore, by definition, $A - B$ is nilpotent. ■

Theorem 1.15 If A and B are $n \times n$ matrices, diagonalizable, and $AB = BA$, then A and B are simultaneously diagonalizable.

Proof: Let A, $B \in M_{n\times n}(\mathbb{C})$, be diagonalizable, and $AB = BA$. Let E_b be an eigenbasis of $\mathbb{C}^n$ which is made of eigenvectors of B, which is invariant. Note that E_b is also invariant under A, since if we have a vector v in E_b, then $ABv = BAv = \lambda Av$. This implies that Av is also an eigenvector of B. Essentially, A and B respect each other's eigenspaces. We look at a restriction of A to E_b. This restriction is a one-dimensional space, which implies that A must take all of the elements of E_b to a scalar multiple of themselves. So, by our understanding of eigenvectors and basis, and that eigenbasis are linearly independent, this implies that E_b is an eigenbasis of A. Note, that A can't have more eigenvectors because an eigenbasis is a spanning set and if A had more than it wouldn't span the entire space. Hence, $Av_i = \lambda_i v_i$, where $v_i \in E_b$. Set

$$X = (v_1 \cdots v_n)$$

and

$$\Lambda = (v_1 \cdots v_n) \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

So, we have $XAX^{-1} = \Lambda$. Also, since if $v_i \in E_b$ it is an eigenvector of B, we can conclude that $XBX^{-1} = \Lambda$ as well. Therefore, since A and B share an eigenbasis, they are simultaneously diagonalizable. ■

References

[1] H. Medina, The Diagonalizable and Nilpotent Parts of a Matrix, pre-print, pp. 1-11.

[2] S. Axler, Linear Algebra Done Right, 3^{rd} ed., Springer, New York, 2015.

[3] K. Hoffman & R. Kunze, Linear Algebra, 2^{nd} ed., Prentice-Hall, New Jersey, 1971.

[4] Jordan Chevalley decomposition, online, https://en.wikipedia.org/wiki/Jordan-Chevalley-decomposition.

[5] Diagonalizable Matrix, online, https://en.wikipedia.org/wiki/Diagonalizable-matrix.