JORDAN DECOMPOSITION VIA THE $\mathcal{Z}$-TRANSFORM

ASIA AGE, THOMAS CREDEUR, TRAVIS DIRLE, AND LAFANIQUE REED

Abstract. Well-known proofs of the decomposition of a complex-valued matrix into a linear combination of projections and nilpotents rely on some heavy machinery. We give a new proof of the Jordan Decomposition using the $\mathcal{Z}$-transform. In this way, we arrive at the result using only calculus.

1. Introduction

In 1986, Schmidt wrote a paper [S86] on canonical forms of complex-valued matrices. In his paper, he gave a proof of the Jordan decomposition of a matrix by using the Laplace transform. This method of proof only involves elementary calculus and an in-depth study of the matrix equation $e^{A(s+t)} = e^{As}e^{At}$, as opposed to more advanced topics like other known proofs. In this paper, we prove a similar result, written below and proven in Theorem 6.2, using only the $\mathcal{Z}$-transform and a careful study of the matrix equation $A^{k+\ell} = A^{\ell}A^{k}$.

Theorem 1.1. *Let A be an $n \times n$ matrix with characteristic polynomial $c_A(z) = (z-\lambda_1)^{m_1} \cdots (z-\lambda_r)^{m_r}$. Then, there exists a unique decomposition*

$$A^k = \sum_{i=1}^{r} \lambda_i^k P_i + \sum_{q=1}^{m_i-1} N_i^q \varphi_{q,\lambda_i}(k)$$

with the following properties

(1) $P_iN_i = N_iP_i = N_i$
(2) $P_iP_j = 0$ *if* $i \neq j$
$P_iP_i = P_i$
(3) $N_i^{m_i} = 0$
(4) $P_iN_j = N_jP_i = 0$ *if* $i \neq j$
(5) $A = \sum_{i=1}^{r} \lambda_i P_i + N_i$
(6) $I = \sum_{i=1}^{r} P_i$.

Properties (2), (3), and (5) together state that any matrix A can be written in a simple way as a combination of projections and nilpotents. This leads us to the following definitions and example of the idea.

Definition 1.2. A **projection** is any square matrix P such that $P^2 = P$.

Definition 1.3. A **nilpotent matrix** is a nonzero square matrix N such that $N^k = 0$ for some positive integer k.

Example 1.4. Consider the matrix

$$A = \begin{bmatrix} 1 & 4 \\ -1 & -3 \end{bmatrix}.$$

Notice that we have the following decomposition

$$A = (-1)\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \begin{bmatrix}2 & 4\\ -1 & -2\end{bmatrix}$$

Here, the first matrix is a projection and the second matrix is nilpotent, as can easily be checked.

Let us build some of the needed prerequisites.

2. Prerequisites

The $\mathcal{Z}$-transform is a useful operator that often comes up as a method to solve difference equations, but we will only use some basic properties of it. Here is its definition.

Definition 2.1. *Let $y(k)$ be a sequence of complex numbers. We define the $\mathcal{Z}$-**transform** of y to be the function $\mathcal{Z}\{y\}(z)$, where z is a complex variable, by the following formula:*

$$\mathcal{Z}\{y\}(z) = \sum_{k=0}^{\infty} \frac{y(k)}{z^k}.$$

There are some interesting functions which we are most interested in applying the $\mathcal{Z}$-transform to. These φ functions will come up in all that follow. Before we define them, we need to understand the falling factorial.

Definition 2.2. Let n be a nonnegative integer. The **falling factorial** is the sequence $k^{\underline{n}}$, with k= 0,1,2,..., given by the following formula

$$k^{\underline{n}} = k(k-1)(k-2)...(k-n+1).$$

If k were allowed to be a real variable, then $k^{\underline{n}}$ could be characterized as the unique monic polynomial of degree n that vanishes at $0, 1, ..., n-1$.

Remark 2.3. We can recover the familiar factorial from the falling factorial. Notice that $k^{\underline{n}}\,|_{k=n} = n!$.

Let us take a look at an example to illustrate how falling factorials work.

Example 2.4. We will compute $6^{\underline{4}}$. Using the definition above, we see that

$$6^{\underline{4}} = 6(6-1)(6-2)(6-3) = 6(5)(4)(3) = 360.$$

Now that we have defined the falling factorial, we can give the definition of the φ functions.

Definition 2.5. *Let a be a complex number and n be a nonnegative integer. The following sequences φ functions are sequences defined as such,*

$$\varphi_{n,a}(k) = \begin{cases} \frac{a^{k-n}k^{\underline{n}}}{n!} & a \neq 0 \\ \delta_n(k) & a = 0, \end{cases}$$

where $\delta_n(k)$ is the sequence which is 0 for all $k \neq n$ and $\delta_n(n) = 1$.

Since these functions are so important to everything that follows, we will compute a few examples of them below.

Example 2.6.

$$\varphi_{0,2}(k) = 2^k = (1, 2, 4, 8, 16, ...)$$
$$\varphi_{1,2}(k) = 2^{k-1}k = (0, 1, 4, 12, 32, ...)$$
$$\varphi_{2,0}(k) = \delta_2(k) = (0, 0, 1, 0, 0, 0, 0...)$$

It turns out that the $\mathcal{Z}$-transform of the φ functions is particularly useful for our purposes. This result is written below (see Proposition 5 of [T12] for the proof).

Proposition 2.7. *Let $a \in \mathbb{C}$ and $n \in \mathbb{N}$. Then,*

$$\mathcal{Z}\{\varphi_{n,a}(k)\}(z) = \frac{z}{(z-a)^{n+1}}.$$

One of the results that will be particularly useful for us is a canonical decomposition of the power of a matrix. This decomposition was given in [T12] and is written below.

Proposition 2.8. *Let A be an $n \times n$ matrix with complex entries. Let $c_A(z) = \det(zI - A)$ be the characteristic polynomial. Assume $a_1, \dots, a_R$ are distinct roots with corresponding multiplicities $M_1, \dots, M_R$. Then for each r, $1 \le r \le R$, and m, $0 \le m \le M_r - 1$, there are $n \times n$ matrices $B_{r,m}$ such that*

$$A^k = \sum_{r=1}^{R} \sum_{m=0}^{M_r - 1} B_{r,m}\varphi_{m,a_r}(k)$$

Proof. We borrow the following proof from [T12].
Our assumptions imply that we can factor c_A in the following way:

$$c_A(z) = \prod_{r=1}^{R} (z - a_r)^{M_r}.$$

The (i, j) entry of $(zI - A)^{-1}$ is of the form $\frac{p_{i,j}(z)}{c_A(z)}$ where $p_{i,j}(z)$ is some polynomial with degree less than n. Using partial fractions, we can write

$$\frac{p_{i,j}(z)}{c_A(z)} = \sum_{r=1}^{R} \sum_{m=0}^{M_r - 1} \frac{b_{r,m}(i, j)}{(z - a_r)^{m+1}}$$

where $b_{r,m} \in \mathbb{C}$. It follows then that

$$z(zI - A)^{-1} = \sum_{r=1}^{R} \sum_{m=0}^{M_r - 1} \frac{zB_{r,m}}{(z - a_r)^{m+1}}$$

where $B_{r,m}$ is the $n \times n$ matrix whose (i, j) entry is $b_{r,m}(i, j)$ for each pair (r, m). By Corollary 8 and Proposition 5 of [T12] we get

$$A^k = \sum_{r=1}^{R} \sum_{m=0}^{M_r - 1} B_{r,m}\varphi_{m,a_r}(k).$$

$\square$

3. An Example

Given a matrix, it is possible to observe that the matrices in its Jordan Decomposition are projections and nilpotents having the properties in Theorem 1.1 without actually knowing what the projections and nilpotents themselves are. We will show an elementary example of this below, and it will be a guiding light towards the proof of Theorem 1.1.

Example 3.1. Let

$$A = \begin{bmatrix} 1 & 4 \\ -1 & -3 \end{bmatrix}.$$

We must first find the characteristic polynomial $c_A(z) = \det(zI - A)$:

$$\begin{aligned} c_A(z) &= \det(zI - A) \\ &= \begin{bmatrix} z-1 & -4 \\ 1 & z+3 \end{bmatrix} \\ &= (z-1)(z+3) - (-4) \\ &= z^2 + 2z + 1 \\ &= (z+1)^2 \end{aligned}$$

From the characteristic polynomial, we can find the eigenvalues a_r and the multiplicities M_r and apply Proposition 2.8. In this example there is only one eigenvalue $a_1 = -1$ and its multiplicity is $M_1 = 2$, so the A^k equation becomes

$$\begin{aligned} A^k &= \sum_{r=1}^{1} \sum_{m=0}^{1} B_{r,m} \varphi_{m,a_r}(k) \\ &= B_{1,0}\varphi_{0,-1}(k) + B_{1,1}\varphi_{1,-1}(k). \end{aligned}$$

Let us define $M = B_{1,0}$ and $N = B_{1,1}$ for simplicity, so the equation then becomes

$$A^k = M\varphi_{0,-1}(k) + N\varphi_{1,-1}(k).$$

Similarly,

$$\begin{aligned} A^\ell &= \sum_{r=1}^{1} \sum_{m=0}^{1} B_{r,m} \varphi_{m,a_r}(\ell) \\ &= B_{1,0}\varphi_{0,-1}(\ell) + B_{1,1}\varphi_{1,-1}(\ell) \\ &= M\varphi_{0,-1}(\ell) + N\varphi_{1,-1}(\ell) \end{aligned}$$

and

$$\begin{aligned}
A^{k+\ell} &= \sum_{r=1}^{1}\sum_{m=0}^{1} B_{r,m}\varphi_{m,a_r}(k+\ell) \\
&= B_{1,0}\varphi_{0,-1}(k+\ell) + B_{1,1}\varphi_{1,-1}(k+\ell) \\
&= M\varphi_{0,-1}(k+\ell) + N\varphi_{1,-1}(k+\ell).
\end{aligned}$$

Now we can substitute these into the equation

$$A^k A^\ell = A^{k+\ell}$$

and use the definition of $\varphi_{n,a}(k)$ to get:

$$(M(-1)^k + Nk(-1)^{k-1})(M(-1)^\ell + N\ell(-1)^{\ell-1}) = M(-1)^{k+\ell} + N(k+\ell)(-1)^{k+\ell-1}.$$

After expanding the left side of the equation we get:

$$\begin{aligned}
& (M(-1)^k + Nk(-1)^{k-1})(M(-1)^\ell + N\ell(-1)^{\ell-1}) \\
= \; & M^2(-1)^k(-1)^\ell + MN(-1)^k\ell(-1)^{\ell-1} \\
+ \; & NMk(-1)^{k-1}(-1)^\ell + N^2k(-1)^{k-1}\ell(-1)^{\ell-1}.
\end{aligned}$$

Therefore

$$A^k A^\ell = A^{k+\ell}$$

implies

$$\begin{aligned}
M(-1)^{k+\ell} + NN(k+\ell)(-1)^{k+\ell-1} &= M^2(-1)^k(-1)^\ell + MN(-1)^k\ell(-1)^{\ell-1} \\
&\quad + NMk(-1)^{k-1}(-1)^\ell + N^2k(-1)^{k-1}\ell(-1)^{\ell-1}.
\end{aligned}$$

Divide both sides of the equation by $(-1)^k(-1)^\ell$ to get

$$M - N(k+\ell) = M^2 - MN\ell - NMk + N^2kl.$$

Since this equation must hold for all k and ℓ, we will choose specific ones to show that M and N have the right properties.
Let $k = 0,\ \ell = 0$. Then we get that $M = M^2$, so M is a projection.
We can then subtract M from both sides of the equation to get

$$N(k+\ell) = MN\ell + NMk - N^2kl.$$

Let $k = 0,\ \ell = 1$. Then we get that $N = MN$. We can make a similar choice to see that $N = NM$. Therefore our equation becomes

$$N(k+\ell) = Nk + N\ell - N^2k\ell.$$

This shows us that $N^2 = 0$, so we know that:
$M^2 = M$
$MN = NM = N$
$N^2 = 0$.
Hence, M is a projection and N is nilpotent.

Remark 3.2. Notice that we were able to conclude that M and N had the properties in Theorem 1.1 without knowing what the matrices are. This will be one of the ideas in the proof.

4. Linear Independence of the φ-functions

Another key idea in the proof of our main theorem (Theorem 1.1) is the linear independence of the φ functions. This is proven below.

Theorem 4.1. *Let $A \in M_n(\mathbb{C})$. Let the characteristic polynomial of A,*

$$c_A(s) = (s - \lambda_1)^{m_1}(s - \lambda_2)^{m_2} \cdots (s - \lambda_n)^{m_n}$$

Then the set

$$\mathbb{B} = \{\varphi_{0,\lambda_1}, \varphi_{1,\lambda_1}, \cdots, \varphi_{m_1,\lambda_1}, \cdots, \varphi_{m_n,\lambda_n}\}$$

is linearly independent.

Proof. Because the $\mathcal{Z}$ transform is an injective linear transformation, it preserves linear independence.
Therefore, if $\mathcal{Z}(\mathbb{B})$ is linearly independent, so is $\mathbb{B}$
By the definition of the $\mathcal{Z}$ transform,

$$\mathcal{Z}(\mathbb{B}) = \{\frac{z}{z - \lambda_1}, \frac{z}{(z - \lambda_1)^2}, \cdots \frac{z}{(z - \lambda_1)^{m_1}}, \cdots, \frac{z}{(z - \lambda_n)^{m_n}}\}$$

Let $c_1, c_2, \cdots c_l$ be constants, such that

$$\frac{c_1 z}{z - \lambda_1} + \frac{c_2 z}{(z - \lambda_1)^2} + \cdots + \frac{c_{m_1} z}{(z - \lambda_1)^{m_1}} + \cdots + \frac{c_l z}{(z - \lambda_n)^{m_n}} = 0$$

Rewrite this as

$$\frac{z \cdot p_1(z)}{(z - \lambda_1)^{m_1}} + \cdots + \frac{z \cdot p_n(z)}{(z - \lambda_n)^{m_n}}$$

by making common denominators. Then, take the limit as z approaches λ_1
Since the sum above is identically zero, it's limit must be zero everywhere. This means, in particular, that

$$\lim_{z \to \lambda_1} \frac{z \cdot p_1(z)}{(z - \lambda_1)^{m_1}}$$

is finite, because the limit as z approaches λ_1 is finite for all other terms in the sum, and must add up to zero.
This implies that $p_1(z) = 0$. Since

$$p_1(z) = c_1(z - \lambda_1)^{m_n - 1} + c_2(z - \lambda_1)^{m_n - 2} + \cdots c_{m_n}$$

c_1 is the only term that has z raised to the $m_n - 1$ power, so $c_1 = 0$.
Reapply this logic to conclude that $c_i = 0$, $\forall 1 \leq i \leq l$ Therefore, $\mathcal{Z}(\mathbb{B})$ is linearly independent, which implies that $\mathbb{B}$ is linearly independent. $\square$

5. Two More Examples

Now that we have shown that the φ-functions are linearly independent, we will provide two more examples that illustrate the method of proof of Theorem 1.1. First we will give an example using a 2×2 matrix.

Example 5.1. Let

$$A = \begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix}.$$

We must first find the characteristic polynomial $C_A(z) = \det(zI - A)$:

$$\begin{aligned} C_A(z) &= \det \begin{bmatrix} z-3 & -2 \\ 1 & z \end{bmatrix} \\ &= z^2 - 3z + 2 \\ &= (z-2)(z-1) = 0. \end{aligned}$$

From the characteristic polynomial, we can find the eigenvalues a_r and the multiplicities M_r and apply Proposition 2.8. In this example there are two eigenvalues $a_1 = 2$ and $a_2 = 1$ with multiplicities $M_1 = 1$ and $M_2 = 1$, so the A^k equation becomes

$$\begin{aligned} A^k &= \sum_{r=0}^{R} \sum_{m=0}^{M_r - 1} B_{r,m} \varphi_{m,a_r}(k) \\ &= \sum_{r=1}^{2} \sum_{m=0}^{0} B_{1,0} \varphi_{0,2}(k) + B_{2,0} \varphi_{0,1}. \end{aligned}$$

Let us define $M = B_{1,0}$ and $N = B_{2,0}$ for simplicity, so the equation now becomes

$$A^k = M\varphi_{0,2}(k) + N\varphi_{0,1}(k).$$

Similarly,

$$\begin{aligned} A^\ell &= \sum_{r=0}^{R} \sum_{m=0}^{M_r - 1} B_{r,m} \varphi_{m,a_r}(\ell) \\ &= \sum_{r=1}^{2} \sum_{m=0}^{0} B_{1,0} \varphi_{0,2}(k) + B_{2,0} \varphi_{0,1} \\ &= M\varphi_{0,2}(l) + N\varphi_{0,1}(\ell) \end{aligned}$$

and

$$\begin{aligned} A^{k+\ell} &= \sum_{r=0}^{R}\sum_{m=0}^{M_r-1} B_{r,m}\varphi_{m,a_r}(k+\ell) \\ &= \sum_{r=1}^{2}\sum_{m=0}^{0} B_{1,0}\varphi_{0,2}(k) + B_{2,0}\varphi_{0,1} \\ &= M\varphi_{0,2}(k+\ell) + N\varphi_{0,1}(k+\ell). \end{aligned}$$

Now we can substitute these into the equation

$$A^k + A^\ell = A^{k+\ell}$$

to get:

$$M\varphi_{0,2}(k+\ell) + N\varphi_{0,1}(k+\ell) = \left[M\varphi_{0,2}(k) + N\varphi_{0,1}(k)\right]\left[M\varphi_{0,2}(\ell) + N\varphi_{0,1}(\ell)\right].$$

After expanding the right side, the equation becomes:

$$\begin{aligned} M\varphi_{0,2}(k+\ell) + N\varphi_{0,1}(k+\ell) &= M^2\varphi_{0,2}(k+\ell) + M\varphi_{0,2}(k)N\varphi_{0,1}(\ell) \\ &+ N\varphi_{0,1}(k)M\varphi_{0,2}(\ell) + N^2\varphi_{0,1}(k+\ell). \end{aligned}$$

which can be rewritten as

$$\begin{aligned} (M\varphi_{0,2}(k))\varphi_{0,2}(l) + (N\varphi_{0,1}(k))\varphi_{0,1}(\ell) &= (M^2\varphi_{0,2}(k) + NM\varphi_{0,1}(k))\varphi_{0,2}(\ell) \\ &+ (MN\varphi_{0,2}(k) + N^2\varphi_{0,1}(k))\varphi_{0,1}(\ell). \end{aligned}$$

Note that $\mathcal{Z}$-transform is 1-1 (one to one) and linear, and $\varphi's$ on both sides of the equation are the same. Thus we can apply Theorem 4.1. The linear independence of the φ-functions implies

$$\begin{aligned} M\varphi_{0,2}(k) &= M^2\varphi_{0,2}(k) + NM\varphi_{0,1}(k) \\ N\varphi_{0,1}(k) &= MN\varphi_{0,2}(k) + N^2\varphi_{0,1}(k). \end{aligned}$$

Linear independence is applied a second time; thus

$$M^2 = M,\ N^2 = N,\ NM = 0, \text{ and } MN = 0.$$

We conclude that M and N are projections.

In the next example, we illustrate our method of proof with an example involving a 3×3 matrix.

Example 5.2. Now let A be the 3×3 matrix

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

Like the previous examples, we must find the characteristic polynomial for A first:

$$\begin{aligned} c_A(z) &= \det \begin{bmatrix} z-5 & -4 & -3 \\ 2 & z+1 & 2 \\ 0 & 0 & z-2 \end{bmatrix} \\ &= (z-1)(z-2)(z-3). \end{aligned}$$

We know

$$A^k = \sum_{r=1}^{R} \sum_{m=0}^{Mr-1} B_{r,m} \varphi_{m,a_r}(k)$$

where a_r represents the eigenvalues $a_1 = 1$ and $a_2 = 2$ and $a_3 = 3$, all with multiplicity 1.
So A^k becomes:

$$A^k = B_{1,0}\varphi_{0,1}(k) + B_{2,0}\varphi_{0,2}(k) + B_{3,0}\varphi_{0,3}(k).$$

We will say $M = B_{1,0}$, $N = B_{2,0}$ and $P = B_{3,0}$ for simplicity, so the equation becomes

$$A^k = M\varphi_{0,1}(k) + N\varphi_{0,2} + P\varphi_{0,3}(k).$$

Similarly

$$\begin{aligned} A^\ell &= B_{1,0}\varphi_{0,1}(\ell) + B_{2,0}\varphi_{0,2}(\ell) + B_{3,0}\varphi_{0,3}(\ell) \\ &= M\varphi_{0,1}(\ell) + N\varphi_{0,2}(\ell) + P\varphi_{0,3}(\ell) \end{aligned}$$

and

$$\begin{aligned} A^{k+\ell} &= B_{1,0}\varphi_{0,1}(k+\ell) + B_{2,0}\varphi_{0,2}(k+\ell) + B_{3,0}\varphi_{0,3}(k+\ell) \\ &= M\varphi_{0,1}(k+\ell) + N\varphi_{0,2}(k+\ell) + P\varphi_{0,3}(k+\ell). \end{aligned}$$

We can now substitute the three equations into the original equation

$$A^k A^\ell = A^{k+\ell}$$

to get:

$$[M\varphi_{0,1}(k) + N\varphi_{0,2}(k) + P\varphi_{0,3}(k)][M\varphi_{0,1}(\ell) + N\varphi_{0,2}(\ell) + P\varphi_{0,3}(\ell)]$$
$$= M\varphi_{0,1}(k+l) + N\varphi_{0,2}(k+l) + P\varphi_{0,3}(k+l).$$

After expanding the left side of the equation we get:

$$\begin{aligned} M\varphi_{0,1}(k+\ell)+N\varphi_{0,2}(k+\ell)+P\varphi_{0,3}(k+\ell) &= M^2\varphi_{0,1}(k)\varphi_{0,1}(\ell)+MN\varphi_{0,1}(k)\varphi_{0,2}(\ell)\\ &\quad + MP\varphi_{0,1}(k)\varphi_{0,3}(\ell)+NM\varphi_{0,2}(k)\varphi_{0,1}(\ell)\\ &\quad + N^2\varphi_{0,2}(k)\varphi_{0,2}(\ell)+NP\varphi_{0,2}(k)\varphi_{(0,3)}(\ell)\\ &\quad + PM\varphi_{0,3}(k)\varphi_{0,1}(\ell)+PN\varphi_{0,3}(k)\varphi_{0,2}(\ell).\\ &\quad + P^2\varphi_{0,3}(k)\varphi_{0,3}(\ell). \end{aligned}$$

This equation becomes

$$\begin{aligned} M\varphi_{0,1}(k)\varphi_{0,1}(\ell)+N\varphi_{0,2}(k)\varphi_{0,1}(\ell)+P\varphi_{0,3}(k)\varphi_{0,3}(\ell) &= M^2\varphi_{0,1}(k)\varphi_{0,1}(\ell)+MN\varphi_{0,1}(k)\varphi_{0,2}(\ell)\\ &+ MP\varphi_{0,1}(k)\varphi_{0,3}(\ell)+NM\varphi_{0,2}(k)\varphi_{0,1}(\ell)\\ &+ N^2\varphi_{0,2}(k)\varphi_{0,2}(\ell)+NP\varphi_{0,2}(k)\varphi_{0,3}(\ell)\\ &+ PM\varphi_{0,3}(k)\varphi_{0,1}(\ell)+PN\varphi_{0,3}(k)\varphi_{0,2}(\ell)\\ &+ P^2\varphi_{0,3}(k)\varphi_{0,3}(\ell). \end{aligned}$$

You can group the terms on the right side together by which $\varphi(\ell)$ they are being multiplied by, and then apply the linear independence of the φ functions that we proved in Theorem 4.1 to get that

$$\begin{aligned} M\varphi_{0,1}(k) &= M^2\varphi_{0,1}(k)+NM\varphi_{0,2}(k)+PM\varphi_{0,3}(k)\\ N\varphi_{0,2}(k) &= MN\varphi_{0,1}(k)+N^2\varphi_{0,2}(k)+PN\varphi_{0,3}(k)\\ P\varphi_{0,3}(k) &= MP\varphi_{0,1}(k)+NP\varphi_{0,2}(k)+P^2\varphi_{0,1}(k). \end{aligned}$$

Apply linear independence one more time to get that:
$M^2 = M$
$P^2 = P$
$N^2 = N$
$MP = PM = 0$
$MN = NM = 0$
$PN = NP = 0.$
Therefore, M, N, and P are all projections, and the properties we are trying to show hold.

6. Main Results - The Jordan Decomposition

Before we begin the proof of the Jordan Decomposition, we prove a lemma involving falling factorials, which mimics the standard binomial theorem.

Lemma 6.1. *Let x, $y \in \mathbb{Z}$, $n \in \mathbb{N}$. Then*

$$(x+y)^{\underline{n}} = \sum_{k=0}^{n} \binom{n}{k} x^{\underline{k}}\, y^{\underline{n-k}} \tag{6.1}$$

Proof. This proof relies on the fact that $x^{\underline{k+1}} = x \cdot (x-1)^{\underline{k}}$
We will do an inductive argument on n.
When $n = 0$, this equation reduces to $1 = 1$, which is true.
Therefore, assume there exists an $n \in \mathbb{N}$ such that (6.1) holds, and show that this implies (6.1) holds for $n+1$.

$$\sum_{k=0}^{n+1} \binom{n+1}{k} x^{\underline{k}}\, y^{\underline{n+1-k}} = \sum_{k=0}^{n} \binom{n+1}{k} x^{\underline{k}}\, y^{\underline{n+1-k}} + x^{\underline{n+1}} \tag{6.2}$$

Because we can just take out the last term of the sum.
Next, we will use Pascal's Rule (which can be found in [KP01] or most any treatise on combinatorics), which states that

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}$$

Apply this to the right hand side of (6.2), to get

$$\sum_{k=0}^{n+1} \binom{n+1}{k} x^{\underline{k}}\, y^{\underline{n+1-k}} = \sum_{k=0}^{n} \left(\binom{n}{k-1} + \binom{n}{k} \right) x^{\underline{k}}\, y^{\underline{n+1-k}} + x^{\underline{n+1}} \tag{6.3}$$

This is the same as writing

$$\sum_{k=0}^{n+1} \binom{n+1}{k} x^{\underline{k}}\, y^{\underline{n+1-k}} = \sum_{k=0}^{n} \binom{n}{k-1} x^{\underline{k}}\, y^{\underline{n+1-k}} + \sum_{k=0}^{n} \binom{n}{k} x^{\underline{k}}\, y^{\underline{n+1-k}} + x^{\underline{n+1}} \tag{6.4}$$

But the first term on the right hand side of (6.4) is zero, so we can shift our index and keep the same answer. This gives us

$$\sum_{k=0}^{n+1} \binom{n+1}{k} x^{\underline{k}}\, y^{\underline{n-k}} = \sum_{k=0}^{n-1} \binom{n}{k} x^{\underline{k+1}}\, y^{\underline{n+1-k}} + \sum_{k=0}^{n} \binom{n}{k} x^{\underline{k}}\, y^{\underline{n+1-k}} + x^{\underline{n+1}} \tag{6.5}$$

Add the term that was split off back in to get

$$\sum_{k=0}^{n+1} \binom{n+1}{k} x^{\underline{k}}\, y^{\underline{n-k}} = \sum_{k=0}^{n} \binom{n}{k} x^{\underline{k+1}}\, y^{\underline{n+1-k}} + \sum_{k=0}^{n} \binom{n}{k} x^{\underline{k}}\, y^{\underline{n+1-k}} \tag{6.6}$$

Now we can use the fact that was mentioned at the start of the proof, namely that $x^{\underline{k+1}} = x \cdot (x-1)^{\underline{k}}$.
Apply this to the right hand side of (6.6) to get

$$\sum_{k=0}^{n+1} \binom{n+1}{k} x^{\underline{k}}\, y^{\underline{n-k}} = x \cdot \sum_{k=0}^{n} \binom{n}{k} (x-1)^{\underline{k}}\, y^{\underline{n+1-k}} + y \cdot \sum_{k=0}^{n} \binom{n}{k} x^{\underline{k}}\, y^{\underline{n-k}} \tag{6.7}$$

Now we can invoke our inductive hypothesis, and get that

$$\sum_{k=0}^{n+1} \binom{n+1}{k} x^{\underline{k}}\, y^{\underline{n-k}} = x \cdot (x+y-1)^{\underline{n}} + y \cdot (x+y-1)^{\underline{n}} \tag{6.8}$$

The right hand side of (6.8) is the same as $(x+y)(x+y-1)^{\underline{n}}$, so using the fact mentioned at the start of the proof again, we get that

$$\sum_{k=0}^{n+1} \binom{n+1}{k} x^{\underline{k}}\, y^{\underline{n-k}} = (x+y)^{\underline{n+1}} \tag{6.9}$$

This is what we were trying to show, so by the principle of mathematical induction, (6.1) is true. □

Now we present our main result.

Theorem 6.2. *Let A be an $n \times n$ matrix with characteristic polynomial $c_A(z) = (z-\lambda_1)^{m_1} \cdots (z-\lambda_r)^{m_r}$. Then, there exists a decomposition*

$$A^k = \sum_{i=1}^{r} \lambda_i^k P_i + \sum_{q=1}^{m_i-1} N_i^q \varphi_{q,\lambda_i}(k)$$

with the following properties

(1) $P_i N_i = N_i P_i = N_i$
(2) $P_i P_j = 0$ *if* $i \neq j$
$P_i P_i = P_i$
(3) $N_i^{m_i} = 0$
(4) $P_i N_j = N_j P_i = 0$ *if* $i \neq j$
(5) $A = \sum_{i=1}^{r} \lambda_i P_i + N_i$
(6) $I = \sum_{i=1}^{r} P_i$.

Proof. This proof is essentially a careful study of the equation $A^{k+l} = A^l A^k$. From Proposition 9 of [T12], we have a canonical decomposition of the matrix power

$$A^{k+l} = \sum_{i=1}^{r} \sum_{j=0}^{M_i-1} M_{i,j} \varphi_{j,\lambda_i}(k+l).$$

Substituting the definition $\varphi_{j,\lambda_i}(k+l) = \sum_{a=0}^{j} \binom{j}{a} \frac{\lambda_i^{k-j+l} k^{\underline{j-a}} l^{\underline{a}}}{j!}$, the right-hand side becomes

$$\sum_{i=1}^{r} \sum_{j=0}^{\infty} \sum_{a=0}^{j} M_{i,j} \binom{j}{a} \frac{\lambda_i^{k-j+l} k^{\underline{j-a}} l^{\underline{a}}}{j!}.$$

Swapping the inner two sums and making the appropriate changes to their limits, we have

$$\sum_{i=1}^{r} \sum_{a=0}^{\infty} \sum_{j=a}^{\infty} M_{i,j} \binom{j}{a} \frac{\lambda_i^{k-j+l} k^{\underline{j-a}} l^{\underline{a}}}{j!}.$$

Letting $j = j + a$, rearranging, and simplifying, we get

$$\sum_{i=1}^{r} \sum_{a=0}^{\infty} \sum_{j=0}^{\infty} M_{i,j+a} \frac{\lambda_i^{k-j} k^{\underline{j}} \lambda_i^{l-a} l^{\underline{a}}}{j! a!}.$$

Now, let $\lambda_i^{k-j} = \sum_{b=1}^{r} \lambda_b^{k-j} \delta_i(b)$. Substituting into the above, our expression for A^{k+l} becomes

$$\sum_{i=1}^{r} \sum_{a=0}^{\infty} \sum_{b=1}^{r} \sum_{j=0}^{\infty} M_{i,j+a} \delta_i(b) \frac{\lambda_b^{k-j} k^{\underline{j}} \lambda_i^{l-a} l^{\underline{a}}}{j! a!}.$$

Using the definition of the φ functions, we can rewrite this as

$$\sum_{i=1}^{r} \sum_{a=0}^{\infty} \sum_{b=1}^{r} \sum_{j=0}^{\infty} M_{i,j+a} \delta_i(b) \varphi_{j,\lambda_b}(k) \varphi_{a,\lambda_i}(l).$$

Now we turn our attention to $A^l A^k$. Using Proposition 9 of [T12], we have

$$\begin{aligned} A^l A^k &= \sum_{i=1}^{r}\sum_{a=0}^{\infty} M_{i,a}\varphi_{a,\lambda_i}(l) \cdot \sum_{b=1}^{r}\sum_{j=0}^{\infty} M_{b,j}\varphi_{j,\lambda_b}(k) \\ &= \sum_{i=1}^{r}\sum_{a=0}^{\infty}\sum_{b=1}^{r}\sum_{j=0}^{\infty} M_{i,a}M_{b,j}\varphi_{j,\lambda_b}(k)\varphi_{a,\lambda_i}(l). \end{aligned}$$

Now, we set equal our expressions for A^{k+l} and $A^l A^k$:

$$\sum_{i=1}^{r}\sum_{a=0}^{\infty}\sum_{b=1}^{r}\sum_{j=0}^{\infty} M_{i,j+a}\delta_i(b)\varphi_{j,\lambda_b}(k)\varphi_{a,\lambda_i}(l) = \sum_{i=1}^{r}\sum_{a=0}^{\infty}\sum_{b=1}^{r}\sum_{j=0}^{\infty} M_{i,a}M_{b,j}\varphi_{j,\lambda_b}(k)\varphi_{a,\lambda_i}(l)$$

Invoking the linear independence of the φ functions given in Theorem 4.1, we have a collection of equations

$$\sum_{b=1}^{r}\sum_{j=0}^{\infty} M_{i,j+a}\delta_i(b)\varphi_{j,\lambda_b}(k) = \sum_{b=1}^{r}\sum_{j=0}^{\infty} M_{i,a}M_{b,j}\varphi_{j,\lambda_b}(k),$$

one for each i and a. Using Theorem 4.1 once again on each of these equations, we conclude that

$$M_{i,j+a}\delta_i(b) = M_{i,a}M_{b,j}.$$

The six properties in the statement of the theorem come from a careful analysis of this equation. Set $P_i = M_{i,0}$ and $N_i = M_{i,1}$. Notice that

$$P_i N_i = M_{i,0}M_{i,1} = M_{i,1}\delta_i(i) = N_i \text{ and } N_i P_i = M_{i,1}M_{i,0} = M_{i,1}\delta_i(i) = N_i,$$

proving (1). To prove (2), note that if $i \neq j$

$$P_i P_j = M_{i,0}M_{j,0} = M_{i,0}\delta_i(j) = 0$$

and if $i = j$

$$P_i P_j = M_{i,0}M_{i,0} = M_{i,0}\delta_i(i) = P_i$$

To prove (3), first use induction to show that $N_i^m = M_{i,m}$.
From there, we can use the fact that if $m > m_i - 1$, then $M_{i,m} = 0$.
Therefore, $N_i^{m_i} = M_{i,m_i} = 0$
For (4), if $i \neq j$,

$$P_i N_j = M_{i,0}M_{j,1} = M_{i,1}\delta_i(j) = M_{j,1}M_{i,0} = N_j P_i = 0$$

(5) and (6) follow by setting $k = 1$ and $k = 0$, respectively.

□

7. Conclusion

Every matrix has a Jordan decomposition, which happens to consist of projections and nilpotents. These matrix components have many properties. Using the $\mathcal{Z}$-transform, we notice that the φ functions are linearly independent. It is the linear independence that allows us to make the conclusion that we do. Observing the fact that $A^{k+\ell} = A^k A^\ell$, we can rewrite the equation as linear combinations of the φ functions and arrive at the properties of the matrix decomposition.

8. Acknowledgements

We would like to thank the SMILE@LSU program, Prof. Davidson, our graduate mentor Jacob Matherne, and the National Science Foundation.

References

[KP01] Walter G. Kelley and Allan C. Peterson, *Difference Equations: An Introduction with Applications*, Academic Press, San Diego, 2001.

[S86] E. J. P. Georg Schmidt, *An Alternative Approach to Canonical Forms of Matrices*, American Mathematical Monthly **93** (1986), no. 3, 176-184.

[T12] Casey Tsai, *The Cayley-Hamilton Theorem via the $\mathcal{Z}$-Transform*, Rose-Hulman Undergraduate Mathematics Journal **13** (2012), no. 2, 44-52.

SMILE@LSU, Louisiana State University, Baton Rouge, Louisiana
E-mail address: aage@xula.edu

SMILE@LSU, Louisiana State University, Baton Rouge, Louisiana
E-mail address: thomas.credeur@gmail.com

SMILE@LSU, Louisiana State University, Baton Rouge, Louisiana
E-mail address: tdirle@go.olemiss.edu

SMILE@LSU, Louisiana State University, Baton Rouge, Louisiana
E-mail address: lar257@msstate.edu