THE JORDAN-FORM PROOF MADE EASY∗

LEO LIVSHITS†, GORDON MACDONALD‡, BEN MATHES†, AND HEYDAR RADJAVI§

Abstract. A derivation of the Jordan Canonical Form for linear transformations acting on finite dimensional vector spaces over \( \mathbb{C} \) is given. The proof is constructive and elementary, using only basic concepts from introductory linear algebra and relying on repeated application of similarities.

Key words. Jordan Canonical Form, block decomposition, similarity

AMS subject classifications. 15A21, 00-02

1. Introduction. The Jordan Canonical Form (JCF) is undoubtably the most useful representation for illuminating the structure of a single linear transformation acting on a finite-dimensional vector space over \( \mathbb{C} \) (or a general algebraically closed field.)

Theorem 1.1. [The Jordan Canonical Form Theorem] Any linear transformation \( T : \mathbb{C}^n \to \mathbb{C}^n \) has a block matrix (with respect to a direct-sum decomposition of \( \mathbb{C}^n \)) of the form

\[
\begin{bmatrix}
J_1 & 0 & 0 & \cdots & 0 \\
0 & J_2 & 0 & \cdots & 0 \\
0 & 0 & J_3 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & J_p
\end{bmatrix}
\]

where each \( J_i \) (called a Jordan block) has a matrix representation (with respect to some basis) of the form

\[
\begin{bmatrix}
\lambda & 1 & 0 & \cdots & 0 \\
0 & \lambda & 1 & \cdots & 0 \\
0 & 0 & \lambda & \ddots & 0 \\
\vdots & \vdots & \vdots & \ddots & 1 \\
0 & 0 & 0 & 0 & \lambda
\end{bmatrix}
\]

Of course, the Jordan Canonical Form Theorem tells you more: that the above decomposition is unique (up to permuting the blocks) while the sizes and number of blocks can be given in terms of generalized multiplicities of eigenvalues.

∗Received by the editors on ????. Accepted for publication on ????. Handling Editor: ???
†Department of Mathematics, Colby College, Waterville, Maine 04901, USA. (l.livshi@colby.edu, dbmathes@colby.edu).
‡Department of Mathematics and Computer Science, University of Prince Edward Island, Charlottetown, P.E.I., C1A 4P3, Canada (g.macdonald@upei.ca). Supported by a grant from NSERC Canada.
§Department of Mathematics and Statistics, Dalhousie University, Halifax, Nova Scotia, B3H 3J3, Canada (radjavi@mscs.dal.ca). Supported by a grant from NSERC Canada.
With the JCF theorem in hand, all the mysteries of single linear transformations are exposed and many of the other theorems of Linear Algebra (for example the Rank-Nullity Theorem and the Cayley-Hamilton Theorem) become immediately obvious.

The JCF also has many practical applications. The one to which most students of mathematics are exposed is that of linear systems of differential equations with constant coefficients. With the JCF of the coefficient matrix in hand, solving such systems is a routine exercise.

Unfortunately, the proof of the JCF is a mystery to many students. In many linear algebra courses, the JCF is given short shrift, since the traditional proofs require not only knowledge of minimal polynomials and block decompositions but also advanced knowledge regarding nilpotents, generalized eigenvalues, esoteric subspaces (usually called admissible subspaces) and polynomial ideals (usually called stuffers or conductors). The proof does not vary much from linear algebra text to linear algebra text and the exposition in Hoffman and Kunze’s text [4] is representative.

A number of papers have given alternative proofs of JCF. For example: Fletcher and Sorensen [1], Galperin and Waksman [2] and Väliaho [5]). Of these proofs, ours is most similar in flavor to that of Fletcher and Sorenson, in that our proof is given in terms of matrices and that we take an algorithmic approach. But our proof is more elementary and relies on repeated similarities to gradually bring the matrix representation to the required form.

The proof presented here requires only basic facility with row reduction, matrix representations, block decompositions, kernels and ranges of transformations, similarity, subspaces, minimal polynomials and eigenvalues. In short, the usual contents of a one semester introductory course in Linear Algebra. As far as pedagogy is concerned, this makes the JCF a natural culmination of such a course, bringing all the basic concepts together and ending the course with a bang.

We have chosen to take a comprehensive approach, and included all the details of the proof. This makes the article longer, but, we hope, also makes it more accessible to students with only the aforementioned introductory course in linear algebra.

In this paper, \( V \) and \( W \) are finite-dimensional vector spaces over \( \mathbb{C} \) (the field of complex numbers) and \( T : V \to W \), is a linear transformation mapping \( V \) to \( W \) (although all definitions and results hold for linear transformations acting on finite-dimensional vector spaces over any algebraically closed field). Some standard terminology and notation that will prove useful is:

- \( L(V, W) \) will denote the space of all linear transformations (also called operators) from the vector space \( V \) to the vector space \( W \). We also write \( L(V) \) for \( L(V, V) \).
- \( \dim(V) \) denotes the dimension of \( V \).
- The kernel (or nullspace) of a linear transformation \( T \in L(V, W) \) is \( \ker(T) = \{ v \in V : T(v) = 0 \} \).
- The minimal polynomial of \( T \) is \( m_T(x) \) (the unique monic polynomial of minimal degree such that \( m_T(T) = 0 \).)
- A direct sum decomposition \( V = V_1 \oplus V_2 \oplus \cdots \oplus V_n \) of a vector space \( V \) is an ordered list of subspaces \( \{V_j\}_{j=1}^n \) of \( V \) such that \( V_j \cap (V_1 + \cdots + V_{j-1}) = \{0\} \) for all \( 2 \leq j \leq n \) (this is one way of saying that the subspaces are independent).
and such that for all $v \in V$ there exist $v_i \in V_i$ such that $v = \sum_{i=1}^{n} v_i$.

- A block matrix of a linear transformation $T$ with respect to a direct sum decomposition such as the one above is a matrix of linear transformations $T_{ij} \in L(V_j, V_i)$ such that, with usual definition of matrix multiplication and identification of a vector $v = \sum_{i=1}^{n} v_i$ (where $v_i \in V_i$) with

$$
\left[
\begin{array}{c}
v_1 \\
v_2 \\
v_3 \\
\vdots \\
v_n
\end{array}
\right],
$$

$$T(v) =
\left[
\begin{array}{cccc}
T_{11} & T_{12} & T_{13} & \cdots & T_{1n} \\
T_{21} & T_{22} & T_{23} & \cdots & T_{2n} \\
T_{31} & T_{32} & T_{33} & \cdots & T_{3n} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
T_{n1} & T_{n2} & T_{n3} & \cdots & T_{nn}
\end{array}
\right]
\left[
\begin{array}{c}
v_1 \\
v_2 \\
v_3 \\
\vdots \\
v_n
\end{array}
\right].
$$

- A linear transformation $T : V \rightarrow V$ is said to be nilpotent if its only eigenvalue is the zero eigenvalue. By consideration of the minimal polynomial it is clear that $T^n = 0$ for some $n \in \mathbb{N}$. The smallest value of $n$ such that $T^n = 0$ is called the index (or index of nilpotency) of $T$.

2. Primary decomposition. Before getting into the technicalities of the proof of the JCF, a special case of the Primary Decomposition Theorem allows us to obtain a block diagonal matrix for a linear transformation $T$ where the eigenvalue structure of each block is as simple as possible; there is only one eigenvalue for each block.

**Theorem 2.1.** Let $T \in L(\mathbb{C}^n)$ (or $L(\mathbb{F}^n)$ for $\mathbb{F}$ an algebraically closed field) with distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$. Then there is a direct sum decomposition

$$V = K_{\lambda_1} \oplus K_{\lambda_2} \oplus \cdots \oplus K_{\lambda_n}$$

with respect to which $T$ has block matrix

$$
\left[
\begin{array}{cccc}
\lambda_1 I + N_1 & 0 & \cdots & 0 \\
0 & \lambda_2 I + N_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \lambda_n I + N_n
\end{array}
\right]
$$

where each $N_j$ is nilpotent.

We prove this using a sequence of Lemmas.

**Lemma 2.2.** Let $N$ be a nilpotent $r \times r$ matrix, $M$ an invertible $s \times s$ matrix and $F$ an arbitrary $r \times s$ matrix. Then the matrix equation $X - NXM = F$ is solvable.

**Proof.** Let $k$ be such that $N^k \neq 0$ and $N^{k+1} = 0$ (Of course, $k$ can be 0). Check that

$$X = F + NFM + N^2FM^2 + \cdots + N^kFM^k$$

3
is a solution.\footnote{}

The solution is actually unique, but we don’t need that fact here.

**Lemma 2.3.** If \( \lambda \) is not an eigenvalue of the \( s \times s \) matrix \( A \) and \( N \) is a nilpotent \( r \times r \) matrix, then the two matrices

\[
\begin{pmatrix}
\lambda I + N & E \\
0 & A
\end{pmatrix}
\quad \text{and} \quad
\begin{pmatrix}
\lambda I + N & 0 \\
0 & A
\end{pmatrix}
\]

are similar (for any choice of the \( r \times s \) matrix \( E \)).

**Proof.** It is sufficient to show there exists matrix \( X \) such that

\[
\begin{pmatrix}
\lambda I + N & E \\
0 & A
\end{pmatrix}
\begin{pmatrix}
I & X \\
0 & I
\end{pmatrix}
= 
\begin{pmatrix}
I & X \\
0 & I
\end{pmatrix}
\begin{pmatrix}
\lambda I + N & 0 \\
0 & A
\end{pmatrix}
\]

or equivalently, \( X(A - \lambda I) - NX = E \). Since \( A - \lambda I \) is invertible by hypothesis, we can rewrite this as

\[
X - NX(A - \lambda I)^{-1} = E(A - \lambda I)^{-1},
\]

which is solvable by Lemma 2.2.\footnote{\textsuperscript{4}}

**Lemma 2.4.** Every operator \( T \in L(\mathbb{C}^n) \) (or \( L(F^n) \) for \( F \) an algebraically closed field) with eigenvalue \( \lambda \) has a matrix of the form

\[
\begin{pmatrix}
\lambda I + N & 0 \\
0 & A
\end{pmatrix}
\]

where \( N \) is nilpotent and \( \lambda \) is not an eigenvalue of \( A \).

**Proof.** Use induction on \( n \), the dimension of the space. Clearly the Lemma is true if \( n = 1 \). Assuming it is true for all dimensions less than \( n \), choose a basis \( \{e_1, e_2, \ldots, e_n\} \) such that \( T(e_1) = \lambda e_1 \). Then \( T \) has block matrix

\[
\begin{pmatrix}
\lambda & E \\
0 & B
\end{pmatrix}
\]

If \( \lambda \) is not an eigenvalue of \( B \), then Lemma 2.3 gives the result. If \( \lambda \) is an eigenvalue of \( B \), since \( B \) acts on a space of dimension less than \( n \), by the induction hypothesis we have that \( T \) has a matrix of the form

\[
\begin{pmatrix}
\lambda & E_1 & E_2 \\
0 & \lambda I & 0 \\
0 & 0 & A
\end{pmatrix}
\]

where \( \lambda \) is not an eigenvalue of \( A \). Again by Lemma 2.3, this matrix is similar to

\[
\begin{pmatrix}
\lambda & E_1 & 0 \\
0 & \lambda I & 0 \\
0 & 0 & A
\end{pmatrix}
\begin{pmatrix}
\lambda I + N & 0 \\
0 & A
\end{pmatrix}
\]
where
\[ N = \begin{bmatrix} 0 & E_1 \\ 0 & 0 \end{bmatrix}. \]

Theorem 2.1 now follows by applying Lemma 2.4 inductively.

The result of this block diagonal decomposition is that we can now work only with the diagonal blocks \( \lambda_i I + N_i \), and piece things back together at the end.

3. Block simplification. It is in this section where we deviate significantly from the standard approach to the proof of the Jordan Canonical Form Theorem. The usual proof works with subspaces and shows that by choosing subspaces, their complements and bases for each judiciously, one can represent each \( \lambda_i I + N_i \) as a direct sum of Jordan blocks. Our approach de-emphasizes subspaces and bases (but they are still present) and shifts the focus to similarities. Since applying a similarity to a matrix is the same as changing the basis with respect to which the matrix is represented, this is more style than substance, but similarities are much more concrete and comprehensible to acolytes of linear algebra who are still developing the skills of visualizing subspaces.

We begin the process with the following lemma which, though elementary, is central to our exposition.

**Lemma 3.1.** Assume that \( V = V_1 \oplus V_2 \oplus V_3 \oplus V_4 \) and let \( T : V \rightarrow V \) have block matrix
\[
\begin{bmatrix}
A_{11} & A_{12} & 0 & 0 \\
A_{21} & A_{22} & B & C \\
0 & 0 & 0 & D \\
0 & 0 & 0 & E
\end{bmatrix}
\]
with respect to this decomposition. If \( B \) is left invertible (there exists \( L \) such that \( LB = I_3 \)) then there exists a linear transformation \( S \) with block matrix
\[
\begin{bmatrix}
I_1 & 0 & 0 & 0 \\
0 & I_2 & 0 & 0 \\
0 & 0 & I_3 & X \\
0 & 0 & 0 & I_4
\end{bmatrix}
\]
such that \( S^{-1}TS \) has block matrix
\[
\begin{bmatrix}
A_{11} & A_{12} & 0 & 0 \\
A_{21} & A_{22} & B & C \\
0 & 0 & 0 & D' \\
0 & 0 & 0 & E
\end{bmatrix}.
\]

(All block matrices are with respect to the given decomposition.)

**Remark 3.2.** The point is that the blocks \( A_{ij} \) and \( B \) do not change under this similarity, but \( C \) is redefined as \( 0 \).
Remark 3.3. Note that left-invertibility of $B \in L(V, W)$ is equivalent to $\ker B = \{0\}$.

Proof. Just do the computation and observe that $S^{-1}TS$

\[
\begin{bmatrix}
A_{11} & A_{12} & 0 & 0 \\
A_{21} & A_{22} & B & BX + C \\
0 & 0 & 0 & D -XE \\
0 & 0 & 0 & E
\end{bmatrix},
\]

Since $B$ is left-invertible, we can solve $BX + C = 0$ for $X$ (just pre-multiply by the left-inverse) so the lemma is proven. \(\square\)

Remark 3.4. We can see from the proof that $V_1$ can be taken to be $\{0\}$ to obtain a special case of the Lemma for the resulting $3 \times 3$ block matrix, where the first block row and column are deleted.

We shall use this lemma repeatedly in the following theorem which gives a simplified block matrix structure for nilpotents.

Theorem 3.5. Let $T : V \to V$ be a nilpotent linear transformation of index $k$. Then there exists a direct sum decomposition $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$ with respect to which $T$ has block matrix

\[
\begin{bmatrix}
0 & B_1 & 0 & 0 & \cdots & 0 \\
0 & 0 & B_2 & 0 & \cdots & 0 \\
0 & 0 & 0 & B_3 & \cdots & \vdots \\
\vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & 0 & 0 & 0 & \cdots & B_{k-1} \\
0 & 0 & 0 & 0 & \cdots & 0
\end{bmatrix}
\]

where every $B_i$ is left-invertible.

Proof. We first dispose of the case of index $k = 2$. By taking $V_1 = \ker T$ and $V_2$ any complement of $V_1$ in $V$, we obtain a block matrix for $T$ of the form

\[
\begin{bmatrix}
0 & B \\
0 & 0
\end{bmatrix}
\]

where $B$ has no kernel by the choice of $V_1$, so that it is left-invertible.

We next assume that $k \geq 3$. Consider the following standard triangulization. Since $\ker T \subseteq \ker T^2 \subseteq \cdots \subseteq \ker T^{k-1}$, we can choose $W_1$ to be the kernel of $T$, $W_2$ a complement of $W_1$ in the kernel of $T^2$, and $W_i$ a complement of $W_{i-1}$ in the kernel of $T^i$ for $i = 1, 2, \ldots, k$. Then $T$ has block representation

\[
\begin{bmatrix}
0 & B_{12} & B_{13} & B_{14} & \cdots & B_{1k} \\
0 & 0 & B_{23} & B_{24} & \cdots & B_{2k} \\
0 & 0 & 0 & B_{34} & \cdots & B_{3k} \\
\vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \\
0 & 0 & 0 & 0 & \cdots & B_{k-1,k} \\
0 & 0 & 0 & 0 & \cdots & 0
\end{bmatrix}
\]
with respect to \( V = W_1 \oplus W_2 \oplus \cdots \oplus W_k \).

Observe that for any block triangularization obtained in this manner, the \( k - 1 \) blocks \( B_{12}, B_{23}, B_{34}, \ldots, B_{k-1,k} \) are all left-invertible. This is most easily seen by noting that the \( j \)th power of \( T \) has block matrix

\[
\begin{bmatrix}
W_1 & W_2 & \cdots & W_j & W_{j+1} \\
0 & 0 & \cdots & 0 & B' \\
0 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & 0 \\
0 & 0 & \cdots & 0 & 0
\end{bmatrix}
\]

where \( B' = B_{12}B_{23}\cdots B_{j,j+1} \). Now \( W_1 \oplus W_2 \oplus \cdots \oplus W_j \) is the kernel of \( T^j \), so \( B' \) has trivial kernel and is left-invertible, hence so is \( B_{j,j+1} \).

We now apply our lemma to create zeroes above the first superdiagonal in a block matrix similar to \( T \).

We first use the special \( 3 \times 3 \) case of Lemma 3.1 with \( V_1 = \{0\} \), \( V_2 = W_1 \), \( V_3 = W_2 \) and \( V_4 = W_3 \oplus W_4 \oplus \cdots \oplus W_k \) to obtain a block matrix for \( S_{1}^{-1}TS_{1} \) of the form

\[
\begin{bmatrix}
0 & B_{12} & 0 & 0 & \cdots & 0 \\
0 & 0 & B'_{23} & B'_{24} & \cdots & B'_{2k} \\
0 & 0 & 0 & B'_{34} & \ddots & B'_{3k} \\
\vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & 0 & \cdots & B'_{k-1,k} \\
0 & 0 & 0 & 0 & \cdots & 0
\end{bmatrix}
\]

Note that the kernels of successive powers of \( S_{1}^{-1}TS_{1} \) are precisely \( W_1 \), \( W_1 \oplus W_2 \), \( W_1 \oplus W_2 \oplus W_3 \), \ldots, \( W_1 \oplus W_2 \oplus \cdots \oplus W_k \). So, by the previous comments, all the blocks \( B'_{23}, B'_{34}, \ldots, B'_{k-1,k} \) are left-invertible.

Now apply Lemma 3.1 to the operator \( S_{1}^{-1}TS_{1} \) with \( V_1 = W_1 \), \( V_2 = W_2 \), \( V_3 = W_3 \), and \( V_4 = W_4 \oplus \cdots \oplus W_k \) to obtain a block matrix for \( S_{2}^{-1}S_{1}^{-1}TS_{1}S_{2} \) of the form

\[
\begin{bmatrix}
0 & B_{12} & 0 & 0 & \cdots & 0 \\
0 & 0 & B'_{23} & 0 & \cdots & 0 \\
0 & 0 & 0 & B'_{34} & \ddots & B'_{3k} \\
\vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & 0 & \cdots & B'_{k-1,k} \\
0 & 0 & 0 & 0 & \cdots & 0
\end{bmatrix}
\]

again observing that all the blocks \( B'_{34}, \ldots, B'_{k-1,k} \) are still left-invertible.
Continue this process to obtain a block matrix for $S_{k-2}^{-1} \cdots S_2^{-1} S_1^{-1} T S_1 S_2 \cdots S_{k-2}$ of the form

$$
\begin{bmatrix}
0 & B_1 & 0 & 0 & \cdots & 0 \\
0 & 0 & B_2 & 0 & \cdots & 0 \\
0 & 0 & 0 & B_3 & \ddots & 0 \\
\vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & 0 & \cdots & B_{k-1} \\
0 & 0 & 0 & 0 & 0 & \cdots & 0
\end{bmatrix}
$$

All these block matrices are with respect to the original decomposition $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$. Now unwind the similarities. Let $S = S_1 S_2 \cdots S_{k-2}$ and consider the direct-sum decomposition $V = S(W_1) \oplus S(W_2) \oplus \cdots \oplus S(W_k)$. The block matrix of $T$ with respect to this decomposition is exactly

$$
\begin{bmatrix}
0 & B_1 & 0 & 0 & \cdots & 0 \\
0 & 0 & B_2 & 0 & \cdots & 0 \\
0 & 0 & 0 & B_3 & \ddots & 0 \\
\vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & 0 & \cdots & B_{k-1} \\
0 & 0 & 0 & 0 & 0 & \cdots & 0
\end{bmatrix}
$$

and the Theorem is proven. $\square$

Remark 3.6. The process used in the above proof of applying successive similarities can also be interpreted as a process of successively choosing appropriate complementary subspaces.

4. Block representation. Next, we arrive at an alternate block decomposition which is almost the Jordan Block decomposition. At the end of this paper we use this to give an alternative Canonical Form.

Up to now, we have been considering matrices whose entries could be considered either linear transformations, or matrices with respect to some arbitrary basis. We shall now choose specific bases and get a specific matrix representation.

We have already noted that left-invertibility of $B \in L(V, W)$ is equivalent to $\ker B = \{0\}$. But, from basic row-reduction arguments, this is tantamount to any matrix representation of $B$ being row equivalent to $\begin{bmatrix} I & 0 \end{bmatrix}$, which (via elementary matrices or other means) is equivalent to the existence of an invertible operator $C \in L(W)$ such that $CB = \begin{bmatrix} I & 0 \end{bmatrix}$. Using this we obtain the following.

Theorem 4.1. Let $T : V \rightarrow V$ be a nilpotent linear transformation of index $k$. Then there exists a direct sum decomposition $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$ with $\dim(V_{j+1}) \leq \dim(V_j)$ and bases $B_j$ for $V_j$ for $j = 1, \ldots, k$ such that the matrix
representation of $T$ with respect to the basis $\{B_\infty, B_\xi, \ldots, B_n\}$ of $V$ is

$$
\begin{bmatrix}
0 & R_1 & \cdots & 0 \\
0 & 0 & \ddots & 0 \\
\vdots & \vdots & \ddots & R_{k-1} \\
0 & 0 & 0 & 0 \\
\end{bmatrix}
$$

where the matrices $R_1, \ldots, R_{k-1}$ are all of the form $\begin{bmatrix} I & 0 \end{bmatrix}$ and decrease in size as we move down the first super-diagonal.

Proof. By Theorem 3.5, we may assume that $T$ has block matrix

$$
\begin{bmatrix}
0 & B_1 & 0 & \cdots & 0 \\
0 & 0 & B_2 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & 0 \\
0 & 0 & 0 & \cdots & B_{k-1} \\
0 & 0 & 0 & \cdots & 0 \\
\end{bmatrix}
$$

with respect to some direct-sum decomposition $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$ with $B_i$ left invertible for $i = 1, \ldots, k-1$. (From this alone we obtain that $\dim(V_{j+1}) \leq \dim(V_j)$.)

With a slight abuse of notation, choose arbitrary bases $D_j$ for $V_j$ and let $B_i$ also denote the matrix representation of $\B_i$ with respect to these bases. By comments prior to this theorem there exists an invertible matrix $C_{k-1}$ such that $C_{k-1}B_{k-1} = R_{k-1} = \begin{bmatrix} I & 0 \end{bmatrix}$. Also, $B_{k-2}C_{k-1}^{-1}$ is left-invertible so there exists an invertible matrix $C_{k-2}$ such that $C_{k-2}(B_{k-2}C_{k-1}^{-1}) = R_{k-2} = \begin{bmatrix} I & 0 \end{bmatrix}$. We can continue this way to define invertible matrices $C_i$ such that $C_i(B_iC_{i+1}^{-1}) = R_i = \begin{bmatrix} I & 0 \end{bmatrix}$. Then, letting $M$ denote the inverse of the block diagonal matrix

$$
\begin{bmatrix}
C_1 \\
C_2 \\
\vdots \\
C_{k-1} & I
\end{bmatrix}
$$

one has that

$$
M^{-1} \begin{bmatrix}
0 & B_1 & 0 & \cdots & 0 \\
0 & 0 & B_2 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & 0 \\
0 & 0 & 0 & \cdots & B_{k-1} \\
0 & 0 & 0 & \cdots & 0 \\
\end{bmatrix} M
$$
Letting $B_i$ be the basis for $V_i$ obtained by applying $M$ to each element of the basis $D_i$ (equivalently, applying $C_i^{-1}$ to each element of the basis $D_i$ where we define $C_i^{-1} = I$) the result follows. □

We now have every nilpotent represented as a matrix of zeros and ones. Our next step is to rearrange these zeros and ones to get Jordan blocks.

5. Basis rearrangement. A simple rearrangement of the basis elements (which can also be considered to be the application of a similarity derived from a permutation matrix) will convert the matrix from Theorem 4.1 into a direct sum of Jordan blocks.

**Theorem 5.1.** Let $T : V \rightarrow V$ be a nilpotent linear transformation of index $k$. Then there exists a direct sum decomposition $V = U_1 \oplus U_2 \oplus \cdots \oplus U_n$ and bases $E_i$ for $U_j$ for $j = 1, \ldots, n$ such that the matrix representation of $T$ with respect to the basis $\{E_\infty, E_1, \ldots, E_n\}$ of $V$ is

$$
\begin{bmatrix}
N_1 & 0 & 0 & \cdots & 0 \\
0 & N_2 & 0 & \cdots & 0 \\
0 & 0 & N_3 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & 0 \\
0 & 0 & 0 & 0 & N_n
\end{bmatrix}
$$

where each $N_i$ is of the form

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

**Proof.** Let $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$ with $\dim(V_{j+1}) \leq \dim(V_j)$ and bases $B_i$ for $V_j$ for $j = 1, \ldots, k$ such that the matrix representation of $T$ with respect to the basis $\{B_\infty, B_1, \ldots, B_k\}$ of $V$ is as in Theorem 4.1. If $B_i = \{v_1^i, v_2^i, \ldots, v_{\dim(V_i)}^i\}$ for $i = 1, 2, \ldots, n$ then let $U_j$ be the subspace of $V$ spanned by $\{v_j^i \in B_j : i \leq \dim(V_1)\}$
(i.e. the $j^{th}$ vectors in each of those ordered bases $B_j$ that have at least $j$ vectors). Then $E_j = \{ v_{j1}^k, \ldots, v_{j1}^1 \}$, where $k_j = \min \{ i : j \leq \dim(V_i) \}$ (i.e. the $j^{th}$ vectors in each of those ordered bases $B_j$ that have at least $j$ vectors, written in reverse order) is an ordered basis for $U_j$. Clearly $V = U_1 \oplus U_2 \oplus \cdots \oplus U_l$ and it is straightforward from the representation of $T$ in Theorem 4.1 that the matrix representation of $T$ with respect to the basis $\{ E_\infty, E_\infty, \ldots, E_\infty \}$ of $V$ is

$$
\begin{bmatrix}
N_1 & 0 & 0 & \cdots & 0 \\
0 & N_2 & 0 & \cdots & 0 \\
0 & 0 & N_3 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & 0 \\
0 & 0 & 0 & 0 & N_n
\end{bmatrix}
$$

where each $N_i$ is of the form

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

Remark 5.2. It is also clear from this proof that these blocks $N_i$ are decreasing in size as we move down the main diagonal, that the largest block is $k \times k$ where $k$ is the number of blocks in the representation in Theorem 4.1 (which was the index of nilpotency) and that the number of blocks equals $\dim(V_1)$.

6. Synthesizing the Jordan Canonical Form. So, taking it from the top, by Theorem 2.1, an arbitrary linear transformation $T : V \to V$ can be decomposed as a block matrix

$$
\begin{bmatrix}
\lambda_1 I + N_1 & 0 & 0 & \cdots & 0 \\
0 & \lambda_2 I + N_2 & 0 & \cdots & 0 \\
0 & 0 & \lambda_3 I + N_3 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & 0 \\
0 & 0 & 0 & 0 & \lambda_n I + N_n
\end{bmatrix}
$$

with respect to $V = V_1 \oplus V_2 \oplus \cdots \oplus V_n$. By Theorems 3.5, 4.1 and 5.1 we can choose bases for each $V_i$ such that the matrix of $N_i$ with respect to these bases is

$$
\begin{bmatrix}
N_{i1} & 0 & \cdots & 0 \\
0 & N_{i2} & \cdots & 0 \\
\vdots & \vdots & \ddots & 0 \\
0 & 0 & 0 & N_{in}
\end{bmatrix}
$$
where each $N_{ij}$ has the form
\[
\begin{bmatrix}
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
0 & 0 & 0 & \ddots & 0 \\
\vdots & \vdots & \ddots & \ddots & 1 \\
0 & 0 & 0 & \cdots & 0
\end{bmatrix}
\]
so $\lambda I + N_i$ has block matrix
\[
\begin{bmatrix}
N_{i1} & 0 & \cdots & 0 \\
0 & N_{i2} & \cdots & 0 \\
\vdots & \vdots & \ddots & 0 \\
0 & 0 & \cdots & N_{in}
\end{bmatrix}
= \begin{bmatrix}
\lambda I + N_{i1} & 0 & \cdots & 0 \\
0 & \lambda I + N_{i2} & \cdots & 0 \\
\vdots & \vdots & \ddots & 0 \\
0 & 0 & \cdots & \lambda I + N_{in}
\end{bmatrix}
\]
and each $J_{ij} = \lambda I + N_{ij}$ is of the form
\[
\begin{bmatrix}
\lambda_i & 1 & 0 & \cdots & 0 \\
0 & \lambda_i & 1 & \cdots & 0 \\
0 & 0 & \lambda_i & \ddots & 0 \\
\vdots & \vdots & \vdots & \ddots & 1 \\
0 & 0 & 0 & \cdots & \lambda_i
\end{bmatrix}
\]
This gives us the Jordan Canonical Form of $T$.

Contained in the above proof is a shorter proof (just omit the basis rearrangement) of the following equivalent Canonical Form.

**Theorem 6.1.** If the linear transformation $T : \mathbb{C}^n \to \mathbb{C}^n$ has eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_p$ then $T$ has a block matrix (with respect to a direct-sum decomposition of $\mathbb{C}^n$) of the form
\[
\begin{bmatrix}
K_1 & 0 & 0 & \cdots & 0 \\
0 & K_2 & 0 & \cdots & 0 \\
0 & 0 & K_3 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & 0 \\
0 & 0 & 0 & \cdots & K_p
\end{bmatrix}
\]
where each $K_i$ has a matrix representation (with respect to some basis) of the form
\[
\begin{bmatrix}
\lambda I & R_1 & 0 & \cdots & 0 \\
0 & \lambda I & R_2 & \cdots & 0 \\
0 & 0 & \lambda I & \ddots & 0 \\
\vdots & \vdots & \vdots & \ddots & R_i \\
0 & 0 & 0 & \cdots & \lambda I
\end{bmatrix}
\]
where each $R_i$ is a matrix of the form $\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$.

Just as with the Jordan Canonical Form, the above decomposition is unique up to permuting the blocks while the sizes and numbers of blocks can be given in terms of multiplicities of eigenvalues.

**Remark 6.2.** Although this proof is elementary and constructive, we are not implying that this algorithm can necessarily be modified to give a computer program which could be used to compute the Jordan Canonical Form on a finite precision machine. There are a number of technical difficulties involved in implementing such a program due to the extreme sensitivity of JCF to perturbation of eigenvalues. The interested reader can see [3] for more details.

**REFERENCES**


