Jordan Canonical Form

S. Kumaresan
School of Math. and Stat.
University of Hyderabad
Hyderabad 500046
kumaresa@gmail.com

Let V be a vector space over $\mathbb{C}$ (more generally, over an algebraically closed field). Let $A\colon V \to V$ be a linear map. The purpose of this article is to give a variety of short, simple and direct proofs of the existence of a Jordan canonical form for linear maps over an algebraically closed field. We also include a proof of uniqueness for completeness sake.

Let us start with a very general observation. If $A\colon V \to V$ is linear, then a vector subspace $W \subset V$ is said be A-stable or invariant under A if $Aw \in W$ for all $w \in W$. If $V = U \oplus W$ is a direct sum of A-invariant subspaces and if $\{u_1, \dots, u_r\}$ is an ordered basis of U and $\{w_1, \dots, w_s\}$ is an ordered basis of W, then $\{u_1, \dots, u_r, w_1, \dots, w_s\}$ is an ordered basis of V. With respect to this basis, the matrix representation of A is of the form $\begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix}$ where B (respectively C) is an $r \times r$ (respectively $s \times s$) matrix. Thus to find basis of V so that the matrix representation of A takes a simple form is equivalent to expressing V as a direct sum of subspaces invariant under A.

Definition 1. A vector subspace W of V is said to be *cyclic* if there exists a scalar λ, an integer $m \geq 1$ and a nonzero vector v such that $(A - \lambda I)^m v = 0$ but $(A - \lambda I)^{m-1} v \neq 0$ and we have

$$W = \operatorname{Span}\{v, (A - \lambda I)v, \dots, (A - \lambda I)^{m-1} v\}.$$

We observe that such a subspace is invariant under A. It has dimension m. For, let

$$\sum_{j=0}^{m-1} a_j (A - \lambda I)^j v = 0.$$

Assume that r is the first integer j such that $a_j \neq 0$. Applying $(A - \lambda I)^{m-1-r}$ to both sides, we find that $a_r (A - \lambda I)^{m-1} v = 0$. Since $(A - \lambda I)^{m-1} v \neq 0$, we deduce that $a_r = 0$.

Example 2. If v is a (nonzero) eigen vector with eigenvalue λ, then $\operatorname{Span}\{v\}$ is a cyclic subspace.

Example 3. If $A\colon \mathbb{C}^2 \to \mathbb{C}^2$ is given by $Ae_1 = e_2$ and $Ae_2 = 0$, then $\mathbb{C}^2 = \{e_1, Ae_1\}$ is cyclic.

Remark 4. Let W be the cyclic subspace as in Def. 1. We find the matrix representation of $A|_W$ with respect to the ordered basis $\{w_1 := v, w_2 := (A-\lambda I)v, \ldots, w_m := (A-\lambda I)^{m-1}v\}$.

$$\begin{aligned}
Aw_1 = Av &= (A-\lambda I)v + \lambda v = \lambda w_1 + w_2 \\
Aw_2 = A\left((A-\lambda I)v\right) &= (A-\lambda I)w_2 + \lambda w_2 = \lambda w_2 + w_3 \\
&\vdots \\
Aw_{m-1} = A\left((A-\lambda I)^{m-2}v\right) &= (A-\lambda I)w_{m-1} + \lambda w_{m-1} = \lambda w_{m-1} + w_m \\
Aw_m = A\left((A-\lambda I)^{m-1}v\right) &= (A-\lambda I)w_m + \lambda w_m = \lambda w_m.
\end{aligned}$$

Hence the matrix of A is

$$\begin{pmatrix}
\lambda & 0 & 0 & 0 & \ldots & 0 \\
1 & \lambda & 0 & 0 & \ldots & 0 \\
0 & 1 & \lambda & 0 & \ldots & 0 \\
0 & 0 & \ddots & \ddots & & \\
0 & 0 & 0 & 1 & \lambda & 0 \\
0 & 0 & 0 & 0 & 1 & \lambda
\end{pmatrix}.$$

If we work with the ordered basis $\{v_1 := w_m, v_2 := w_{m-1}, \ldots, v_m := w_1\}$, the matrix of A is the standard Jordan block $J_m(\lambda)$ where

$$J_m(\lambda) = \begin{pmatrix}
\lambda & 1 & 0 & 0 & \ldots & 0 \\
0 & \lambda & 1 & 0 & \ldots & 0 \\
 & & \ddots & \ddots & \ddots & \\
0 & 0 & 0 & \lambda & 1 & 0 \\
0 & 0 & 0 & & \lambda & 1 \\
0 & 0 & 0 & & & \lambda
\end{pmatrix}.$$

A Jordan canonical form of $A\colon V \to V$ is got if we can choose an ordered basis

$$\{v_{11}, \ldots, v_{1n_1}, \ldots, v_{k1}, \ldots, v_{kn_k}\}$$

with the following property

$$Av_{i1} = \lambda_i v_{i1} \text{ and } A(v_{ij}) = \lambda_i v_{ij} + v_{ij-1} \text{ if } j > 1, 1 \le i \le k.$$

We call the ordered set $\{v_{i1}, \ldots, v_{in_i}\}$ as a *string* headed by v_{i1} corresponding to the eigenvalue λ_i. The corresponding matrix is then of the form

$$\operatorname{diag}\left(J_{n_1}(\lambda_1), \ldots, J_{n_k}(\lambda_k)\right).$$

Thus to prove the existence of the Jordan Canonical form, it is enough to prove that V is a direct sum of cyclic subspaces. Our first proof uses this observation and is due to I. Gohberg and S. Goldberg.

First Proof of the existence of a Jordan Form

Idea of the proof: Assume that A is singular. The proof is by induction on n. Let F be an $n-1$ dimensional subspace containing AV. Then F is A-invariant. The proof is split into two cases. In the first case, there exists a vector $\varphi \notin F$ such that $A\varphi = 0$. In this case, $V = F \oplus \mathrm{Span}\{\varphi\}$. The second case is when no such φ exists. The trick here is to replace one of the cyclic subspaces of F by a cyclic subspace of A in V whose dimension is larger by one, while keeping the other cyclic spaces unaffected.

Remark 5. We need the following fact in the proof: Let $W = H \oplus \mathrm{Span}\{\varphi, A\varphi, \dots, A^{m-1}\varphi\}$ with $A^{m-1}\varphi \neq 0$ and $A^m\varphi = 0$. Assume that H is an A-invariant subspace of V and that $A^m H = \{0\}$. Given any $h \in H$, we let $\psi := \varphi + h$. Then we claim that

$$W = H \oplus \mathrm{Span}\{\psi, A\psi, \dots, A^{m-1}\psi\}.$$

We need only check that the sum is direct. Let $\sum_{j=0}^{m-1} a_j A^j \psi \in H$, say h_1. Then $\sum_{j=0}^{m-1} a_j A^j \varphi = h_1 - \sum_{j=0}^{m-1} a_j A^j h \in H$. This contradicts our assumption that the given sum is direct.

Theorem 6 (Jordan Decomposition Theorem)**.** *Let V be a vector space over an algebraically closed field and $A\colon V \to V$ be linear. Then V can be written as a direct sum of cyclic subspaces of A.*

Proof. We prove the result by induction on the dimension of V. The result is trivial if $\dim V = 1$. Assume that the result is true for all $n-1$ dimensional vector spaces. Let A be a linear map on an n dimensional vector space. We first prove the result assuming that A is singular. Thus, the range AV is at most $n-1$ dimensional. Let W be any $n-1$ dimensional vector subspace which contains AV. Then W is A-invariant, since $AW \subset AV \subset W$. By induction hypothesis, W is the sum of cyclic subspaces W_j, $1 \leq j \leq k$. Let

$$W_j := \mathrm{Span}\{w_j, (A-\lambda_j I)w_j, \dots, (A-\lambda_j I)^{m_j-1}w_j\},$$

where we assume that $(A-\lambda_j I)^{m_j} w_j = 0$ but $(A-\lambda_j I)^{m_j-1}w_j \neq 0$.

We shall assume that W_j are indexed in such a way that $m_{j-1} \leq m_j$.

Let $\varphi \in V \setminus W$. We claim that $A\varphi$ is of the form

$$A\varphi = \sum_{j\in S} a_j w_j + Aw, \text{ where } w \in W \text{ and } S = \{j : \lambda_j = 0\}. \tag{1}$$

(If $S = \emptyset$, then $A\varphi = Aw$.)

To prove the claim, we note that $A\varphi \in AV \subset W$. Hence $A\varphi$ is a linear combination of vectors of the form $(A-\lambda_j I)^r w_j$, $0 \leq r \leq m_j - 1$, $1 \leq j \leq k$. Now for $\lambda_j = 0$, the vectors $Aw_j, \dots, A^{m_j-1}w_j$ are in AW.

Let $\lambda_j \neq 0$. Then we can use the binomial expansion of $(A-\lambda_j I)^{m_j}$ in the equation $(A-\lambda_j I)^{m_j} w_j = 0$ and conclude that w_j is a linear combination of $A^q w_j$ for $1 \leq q \leq m_j$.

(See Remark 7 below for more details.) Thus all the vectors $(A-\lambda_j I)^r w_j$ lie in AW. This completes the proof of our claim.

Let $\psi := \varphi - w$ where w is as (1). Observe that $\psi \notin W$. From (1), we have

$$A\psi = \sum_{j\in S} a_j w_j. \tag{2}$$

If $A\psi = 0$, we are through. For, then, Span $\{\psi\}$ is cyclic and we have $V = W \oplus \text{Span}\,\{\psi\}$. Suppose that $A\psi \neq 0$. Let p be the largest of the integers j in (2) for which $a_j \neq 0$. Let $v := (1/a_p)\psi$. Then

$$Av = w_p + \sum_{j\in S, j<p} \frac{a_j}{a_p} w_j. \tag{3}$$

Define $U := \oplus_{j\in S, j<p} W_j$. The subspace U is A-invariant. Since $m_j \leq m_p$ for $j < p$, it follows that $A^{m_p}(U) = \{0\}$. Thus by the observation made in Remark 5 and (3), it follows that

$$U \oplus W_p = U \oplus \text{Span}\,\{Av, \dots, A^{m_p}v\}.$$

Hence

$$W = (\oplus_{j\neq p} W_j) \oplus \text{Span}\,\{Av, \dots, A^{m_p}v\}.$$

Since $v \notin W$, we have

$$V = W \oplus \text{Span}\,\{v\} = (\oplus_{j\neq p} W_j) \oplus \text{Span}\,\{v, Av, \dots, A^{m_p}v\}.$$

This completes the proof of the theorem under the assumption that A is singular.

Since the filed is algebraically closed, there exists an eigen value α of A on V. The map $A - \alpha I$ is singular and we can apply the previous result to this map. □

Remark 7. Let $\lambda \neq 0$. Assume that $w \in V$ is such that $(A-\lambda I)^m w = 0$ but $(A-\lambda I)^{m-1} w \neq 0$ for $m \geq 1$. Then, from the binomial expansion, we get

$$0 = (A-\lambda I)^m w = (-\lambda)^m w + \sum_{r=1}^{m} \binom{m}{r} (-\lambda)^{m-r} A^r w,$$

so that

$$w = (-1)^{m+1} \lambda^{-m} \sum_{r=1}^{m} \binom{m}{r} (-\lambda)^{m-r} A^r w.$$

In particular, w is a linear combination of $A^r w$, $1 \leq r \leq m$.

We now show how the above proof gives us an algorithm to find the Jordan decomposition.

Example 8. Consider the matrix

$$\begin{pmatrix} 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then $Ae_1 = 0$, $Ae_2 = e_1$, $Ae_3 = e_1 + e_2 = Ae_5$, $Ae_4 = e_2$. The range of A is Span $\{e_1, e_2\}$. We take $W = \text{Span}\,\{e_1, e_2, e_3, e_4\}$. Clearly, $\{e_4, e_2 = Ae_4, e_1 = A^2e_4\}$ is a cyclic subspace of W. To write it as a direct sum of cyclic subspaces, we select e_3 which is not in the cyclic subspace. We have

$$Ae_3 = e_2 + e_1 = Ae_4 + A^2e_4 = A(e_4 + e_2).$$

Hence we consider $e_3 - e_4 - e_2$. Then $A(e_3 - e_4 - e_2) = 0$. It follows that

$$W = \text{Span}\,\{e_4, Ae_4, A^2e_4\} \oplus \text{Span}\,\{e_3 - e_4 - e_2\}.$$

We now take $e_5 \notin W$. Look at $Ae_5 = e_2 + e_1 = A(e_4 + e_2)$. So, we take $e_5 - e_2 - e_4$ and find that $A(e_5 - e_2 - e_4) = 0$. Thus

$$\mathbb{C}^5 = \text{Span}\,\{e_4, Ae_4, A^2e_4\} \oplus \text{Span}\,\{e_3 - e_2 - e_4\} \oplus \text{Span}\,\{e_5 - e_2 - e_4\}$$

is a direct sum of cyclic subspaces.

Example 9. Let

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & a \\ 0 & 0 & 0 & 0 & b \\ 0 & 0 & 0 & 1 & c \\ 0 & 0 & 0 & 0 & d \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then $Ae_2 = e_1$, $Ae_1 = 0$, $Ae_4 = e_3$, $Ae_3 = 0$. We take

$$W = \text{Span}\,\{e_1, e_2, e_3, e_4\} = \text{Span}\,\{e_2, Ae_2\} \oplus \text{Span}\,\{e_4, Ae_4\}.$$

Now $e_5 \notin W$. We have

$$Ae_5 = ae_1 + be_2 + ce_3 + de_4 = (be_2 + de_4) + A(ae_2 + ce_4).$$

Case (1): If $d \neq 0$, we take $v = (e_5 - ae_2 - ce_4)/d$. Then $Av = e_4 + (b/d)e_2$, $A^2v = e_3 + (b/d)e_1$ so that

$$\mathbb{C}^5 = \text{Span}\,\{e_2, Ae_2\} \oplus \text{Span}\,\{v, Av, A^2v\}.$$

Case (2): If $d = 0$ but $b \neq 0$, we take $v = (e_5 - ae_2 - ce_4)/b$. Then $Av = e_2$, $A^2v = Ae_2 = e_1$ and hence

$$\mathbb{C}^5 = \text{Span}\,\{e_4, Ae_4\} \oplus \text{Span}\,\{v, Av, A^2v\}.$$

Case (3): Finally, if $d = b = 0$, then take $v = e_5 - ae_2 - ce_4$. Then $Av = 0$ so that

$$\mathbb{C}^5 = \text{Span}\,\{e_2, Ae_2\} \oplus \text{Span}\,\{e_4, Ae_4\} \oplus \text{Span}\,\{v\}.$$

Example 10. Consider the matrix $B := \begin{pmatrix} 0 & 4 & 2 \\ -3 & 8 & 3 \\ 4 & -8 & -2 \end{pmatrix}$. The only eigen value of B is $\lambda = 2$. Let $A := B - 2I$. The range $R(A) := \text{Span}\,\{(2, 3, -4)\}$. We let $W := \text{Span}\,\{(2, 3, -4), e_2\}$. We let $\varphi := e_1 \notin W$. Since $A\varphi = -(2, 3, -4)$, we let $v_1 := -e_1, v_2 := (2, 3, -4), v_3 = e_1 + e_3$. Then

$$Av_1 = v_2, Av_2 = 0, Av_3 = 0,$$

so that the Jordan canonical form of A is $\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Hence the Jordan canonical form of the given matrix B is $\begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

Second Proof of the existence of a Jordan Form

The following proof is taken from my article "Structure Theorems for Linear Maps." I include it here for the sake of comparison.

Theorem 11 (Jordan Basis for Nilpotent Maps)**.** *Let $A \colon V \to V$ be a nilpotent linear map. Assume that k is the smallest positive integer such that $A^k = 0$. Let $S \subset V$ be any linearly independent set such that* $\text{Span } S \cap \text{Ker } A^{k-1} = \{0\}$*. Then S can be extended to a Jordan basis for A.*

Proof. We prove the result by induction on k. For $k = 1$, the result is obvious.

Assume that $k \geq 2$. We extend S to a linearly independent subset S' such that $V = \text{Span } S' \oplus \text{Ker } A^{k-1}$. We observe that A is one-one on Span S'. For, if $Av = 0$ for some $v \in \text{Span } S'$, then $v \in \text{Ker } A \subset \text{Ker } A^{k-1}$. Hence $v = 0$ thanks to our hypothesis on Span S'. It follows that AS' is a linearly independent subset of Ker A^{k-1}.

We claim that $A(\text{Span } S') \cap \text{Ker } A^{k-2} = \{0\}$. For, if $w = Av$ for some $v \in \text{Span } S'$, then

$$w \in \text{Ker } A^{k-2} \iff A^{k-1}v = 0 \iff v \in \text{Span } S' \cap \text{Ker } A^{k-1} = \{0\}.$$

We invoke the induction hypothesis to the restriction of A on the invariant subspace Ker A^{k-1}. It follows that AS' can be extended to a Jordan basis J' for A on Ker A^{k-1}. We combine S' with J' in such a way that each $v \in S'$ comes after Av. The result is easily seen to be a Jordan basis of V for A. □

Example 12. Consider the matrix of Example 8. Then

$$\begin{aligned} \text{Ker } A &= \text{Span}\,\{e_1, e_2 + e_4 - e_3, e_3 - e_5\} \\ \text{Ker } A^2 &= \text{Span}\,\{e_1, e_2, e_2 + e_4 - e_3, e_3 - e_5\} \end{aligned}$$

We take $S' = \{e_4\}$. Then a Jordan basis for A is given by

$$\{v_1 = e_1, v_2 = e_2, v_3 = e_4, v_4 = e_2 + e_4 - e_3, v_5 = e_3 - e_5\}$$

and the corresponding Jordan canonical form is $\big(J_3(0), J_1(0), J_1(0)\big)$, using a standard notation.

Exercise. Carry out a similar exercise for the matrix in Example 9.

Remark 13. Using the factorization of the characteristic (or minimal) polynomial of A, we can decompose V into a direct sum of generalized eigenspaces. The restriction of the map $A - \lambda I$ will be nilpotent on the generalized eigenspace corresponding to λ. Using this fact, one can derive the Jordan basis theorem for any linear map. For details, consult my article referred to above or any standard book on linear algebra. One can also derive the Jordan canonical form for general linear maps as follows.

Theorem 14. *Let $T: V \to V$ be a linear map on a finite dimensional vector space over an algebraically closed field. Then there exists a Jordan basis for T on V so that the matrix representation A of T with respect to this basis is of the form*

$$A = \operatorname{diag}\left(J_{k_1}(\lambda_1), \dots, J_{k_r}(\lambda_r)\right).$$

Proof. We prove the result by induction on the dimension of V. If $\dim V = 1$, the result is obvious. We assume that $n \geq 2$. Let λ be an eigen value of T. Let $S := T - \lambda I$.

Consider the chain of vector subspaces $\text{Ker } S \subseteq \text{Ker } S^2 \subseteq \cdots$. This is a nondecreasing chain and since $\dim V$ is finite, $(\dim \text{Ker } S^k)$ cannot be a strictly increasing sequence of integers. Assume that m is the least positive integer such that $\text{Ker } S^m = \text{Ker } S^{m+1}$.

We claim that $\text{Ker } S^m = \text{Ker } S^{m+i}$ for $i \in \mathbb{N}$. We prove this by induction on i. The claim is true for $i = 1$ by our assumption on m. Assume the result for i. Let $v \in \text{Ker } S^{m+i+1}$. Then $Sv \in \text{Ker } S^{m+i} = \text{Ker } S^m$. It follows that $v \in \text{Ker } S^{m+1} = \text{Ker } S^m$.

We now show that $V = \text{Ker } S^m \oplus \operatorname{Im} S^m$. First of all we observe that $\text{Ker } S^m \cap \operatorname{Im} S^m = (0)$. For, if $v \in \operatorname{Im} S^m$, say, $v = S^m w$, then $v \in \text{Ker } S^m$ iff $S^m v = S^{2m} w = 0$. But then $w \in \text{Ker } S^{2m} = \text{Ker } S^m$. We conclude that $v = S^m w = 0$. By the rank-nullity theorem $\dim V = \dim \text{Ker } S^m + \dim \operatorname{Im} S^m$. As a consequence, $\text{Ker } S^m + \operatorname{Im} S^m = V$ and hence $V = \text{Ker } S^m \oplus \operatorname{Im} S^m$.

Note that each of the summands $\text{Ker } S^m$ and $\operatorname{Im} S^m$ are invariant under S and that S is nilpotent on $\text{Ker } S^m$. If $V = \text{Ker } S^m$, then S is nilpotent on V and the result follows form Theorem 11. If $\text{Ker } S^m$ is a proper subspace of V, then $\dim \text{Ker } S^m \geq 1$, since λ is an eigen value of T. Thus $\text{Ker } S^m$ and $\operatorname{Im} S^m$ are both invariant under S and of dimension strictly less than that of V. By induction hypothesis, there exist Jordan basis $\mathcal{B}_1$ for $\text{Ker } S^m$ and $\mathcal{B}_2$ for $\operatorname{Im} S^m$. Concatenating them gives a Jordan basis for S (and hence for T) on V. (That is, let $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ in such a way that the elements are listed in such a way that the order is preserved among elements belonging to the same set and any element of $\mathcal{B}_2$ follows after all elements of $\mathcal{B}_1$.) □

Remark 15. Let $A: V \to V$ be linear. Let m be the least positive integer such that $\text{Ker } A^m = \text{Ker } A^{m+1}$. Then we always have the direct sum decomposition $V = \text{Ker } A^m \oplus \operatorname{Im} A^m$. Note that A is nilpotent when restricted to $\text{Ker } A^m$ and is invertible when restricted to $\operatorname{Im} A^m$, since $\text{Ker } A \subset \text{Ker } A^m$.

Such a decomposition is unique, that is, if $V = K \oplus L$ such that K and L are invariant under A, A is nilpotent on K and invertible on L, then $K = \text{Ker } A^m$ and $L = \operatorname{Im} A^m$. For, $K \subset \cup_r \text{Ker } A^r = \text{Ker } A^m$. Hence $\operatorname{Im} A^m = A^m V = A^m L = L$, since A is invertible on L. It follows that $\dim K = \dim \text{Ker } A^m$ so that $K = \text{Ker } A^m$.

Third Proof of the existence of a Jordan Form

We now give a proof of the existence of Jordan canonical form following a method of Filippov. To bring out the simplicity of the proof, we establish the result in the case of a nilpotent map. We later deal with the general case.

Theorem 16 (Jordan canonical form for nilpotent operators)**.** *Let $A\colon V \to V$ be a nilpotent operator on a finite dimensional vector space over a field F. Assume that the all the roots of the characteristic polynomial lie in F. Then there exists an A-Jordan basis of V.*

Proof. We prove this result by induction on the dimension of V. If $\dim V = 1$, then $A = 0$ and hence any nonzero element is a Jordan basis of V. The result is also true if $A = 0$ and whatever be the dimension of V. So, we now assume that the result is true for all nonzero nilpotent operators on any finite dimensional vector space with dimension less than n where $n > 1$.

Let V be of dimension n and $A\colon V \to V$ be nonzero and nilpotent. Since Ker $A \neq \{0\}$, $\dim \operatorname{Im} A < n$. It is also invariant under A. Thus the restriction of A to $W = \operatorname{Im} A$, which we denote by A again, is a nilpotent operator on W. We can therefore apply the induction hypothesis. We then get a Jordan basis of W, say, $J = J_1 \cup \cdots \cup J_k$ where each J_i is a Jordan string:

$$J_i = \{v_{i1}, \dots, v_{in_i}\} \text{ with } Av_{i1} = 0 \text{ and } Av_{ij} = v_{ij-1} \text{ for } 2 \leq j \leq n_i.$$

We have, of course, $n_1 + \cdots + n_k = \dim \operatorname{Im} A$.

Suggestion: The reader may assume that there is only one Jordan string during the first reading of the proof below. He may also would like to understand the proof in a special case, say, $A\colon \mathbb{R}^5 \to \mathbb{R}^5$ given by

$$Ae_1 = 0 = Ae_2, Ae_3 = e_4, Ae_4 = e_5, Ae_5 = 0.$$

By very assumption that J is a basis of $\operatorname{Im} A$, the set $\{v_{i1} : 1 \leq i \leq k\}$ (of the first elements of the Jordan strings J_i) is a linearly independent subset of V and it is a subset of Ker A. We extend this set to a basis of Ker A, say, $\{v_{11}, \dots, v_{k1}, z_1, \dots, z_r\}$. Each last element $v_{in_i} \in J_i$ lies in $\operatorname{Im} A$ and hence we can find $v_{in_i+1} \in V$ such that $Av_{in_i+1} = v_{in_i}$. We now let $B_i := J_i \cup \{v_{in_i+1}\}$ and $B := \cup_{i=1}^k B_i \cup \{z_1, \dots, z_r\}$. Using rank-nullity theorem, we see that $|B| = n$. We claim that B is linearly independent subset of V. Let

$$(a_{11}v_{11} + \cdots + a_{1n_1+1}v_{1n_1+1}) + \cdots + (a_{k1}v_{k1} + \cdots + a_{1n_k+1}v_{1n_k+1}) + b_1z_1 + \cdots + b_rz_r = 0. \quad (4)$$

We apply A to both sides. Since $z_j, v_{i1} \in$ Ker A for $1 \leq j \leq r$ and $1 \leq i \leq k$, we arrive at the following equation:

$$A\left([a_{12}v_{12} + \cdots + a_{1n_1+1}v_{1n_1+1}] + \cdots + [a_{k2}v_{k2} + \cdots + a_{kn_k+1}v_{1n_k+1}]\right) = 0.$$

Since $Av_{ij} = v_{ij-1}$ for $1 \leq i \leq k$ and $2 \leq j \leq n_i + 1$, we get

$$(a_{12}v_{11} + \cdots + a_{1n_1+1}v_{1n_1}) + \cdots + (a_{k2}v_{k1} + \cdots + a_{kn_k+1}v_{1n_k}) = 0.$$

Since v_{ij}'s that appear in the above equation are linearly independent, we deduce that $a_{ij} = 0$ for $1 \leq i \leq k$ and $2 \leq j \leq n_i + 1$. Thus (4) becomes

$$a_{11}v_{11} + \cdots + a_{k1}v_{k1} + b_1 z_1 + \cdots + b_r z_r = 0.$$

Th vectors that appear in the equation above form a basis of Ker A and hence we deduce all the coefficients in the equation are zero. Thus, we have show that all the coefficients in (4) are zero and hence B is linearly independent. □

We now prove the general case adapting Flippov's argument above.

Proof. The proof is by induction on $\dim V$. The result is true for when $\dim V = 1$. Assume the result for any linear map on a vector space of dimension less than n. Assume that $A\colon V \to V$ be linear and that $\dim V = n$. We shall first prove the result assuming that A is singular and extend it to all linear maps.

Let $W := \operatorname{Im} A$. Since A is singular, $\dim W < n$. Since W is invariant under A, we may apply the induction hypothesis to the restriction of A to W. We then get a Jordan basis for $A|_W$. Let $\{\lambda_1, \ldots, \lambda_m\}$ be the eigen values of $A|_W$. Let

$$\{v_{11}, v_{12}, \ldots, v_{1n_1}, \ldots, v_{m1}, v_{m2}, \ldots, v_{mn_m}\}$$

be a Jordan basis of $A|_W$. Recall that this means that

$$Av_{i1} = \lambda_i v_{i1} \quad \text{and } Av_{ij} = \lambda_i v_{ij} + v_{ij-1} \text{ for } j > 1.$$

Without loss of generality, we shall assume that the eigen values are indexed in such a way that $\lambda_i = 0$ for $1 \leq i \leq p$ and $\lambda_i \neq 0$ for $i > p$.

The proof is quite simple and the idea on which the proof is based is given below. We hope that the reader is not put off by the notation.

Idea of the proof: The set $\{v_{i1} : 1 \leq i \leq p\}$ is a basis of Ker $(A|_W)$. Extend this to a basis of Ker A by adjoining $\{u_1, \ldots, u_s\}$. Let $w_i \in V$ be such that $Aw_i = v_{in_i}$. Then

$$\{u_1, \ldots, u_s, v_{11}, v_{12}, \ldots, v_{1n_1}, w_1, \ldots, v_{m1}, v_{m2}, \ldots, v_{mn_m}, w_m\}$$

is a Jordan basis for A on V.

Since $Av_{i1} = 0$ for $1 \leq i \leq p$, we see that $v_{i1} \in \text{Ker}\,(A|_W) \subset \text{Ker } A$. We claim that

$$\text{Ker}\,(A|_W) = \text{Span}\,\{v_{11}, \ldots, v_{p1}\}. \tag{5}$$

To prove the claim, let

$$w := \sum_{i=1}^{m} \sum_{j=1}^{n_i} \alpha_{ij} v_{ij} = \sum_{i=1}^{p} \sum_{j=1}^{n_i} a_{ij} v_{ij} + \sum_{i=p+1}^{m} \sum_{j=1}^{n_i} b_{ij} v_{ij} \in \text{Ker}\,(A|_W).$$

We need to show that the only possible nonzero coefficients in the above sum are a_{i1}'s. To establish this, we compute Aw. We get

$$0 = \sum_{i=1}^{p} \sum_{j=2}^{n_i} a_{ij} v_{ij-1} + \sum_{i=p+1}^{m} \lambda_i b_{i1} v_{i1} + \sum_{i=p+1}^{m} \sum_{j=2}^{n_i} \left(b_{ij} v_{ij-1} + \lambda_i b_{ij} v_{ij} \right).$$

Since v_{ij}'s are linearly independent, if any of them occurs only once in the above sum, then its coefficient must be zero. Hence, the coefficient $a_{ij} = 0$ for $j \geq 2$, as v_{ij-1} for $1 \leq i \leq p$ appears in the sum of Aw exactly once. For similar reasons, $\lambda_i b_{in_i} = 0$, for $p+1 \leq i \leq m$. Since $\lambda_i \neq 0$ for $i > p$, we deduce that $b_{in_i} = 0$ for each $p+1 \leq i \leq m$. Let us now fix i and examine the summand corresponding to this i.

$$0 = \lambda_i b_{i1} v_{i1} + (b_{i2} v_{i1} + \lambda_i b_{i2} v_{i2}) + \cdots + \left(b_{in_i} v_{in_i - 1} + \lambda_i b_{in_i} v_{in_i} \right).$$

For brevity sake we ignore the index i and let $k = n_i$ in the expression above and rewrite it as

$$0 = \lambda b_1 v_1 + (b_2 v_1 + \lambda b_2 v_2) + \cdots + (b_k v_{k-1} + \lambda b_k v_k). \tag{6}$$

Since v_k appears only once in the above linear combination on the right hand side, we deduce that $\lambda b_k = 0$. Since $\lambda \neq 0$, it follows that $b_k = 0$. Using this in (6), we get a similar equation with k replaced by $k-1$:

$$0 = \lambda b_1 v_1 + (b_2 v_1 + \lambda b_2 v_2) + \cdots + (b_{k-1} v_{k-2} + \lambda b_{k-1} v_{k-1}).$$

Since v_{k-1} appears only once, as earlier, we conclude that $b_{k-1} = 0$ and so on. We thus find $b_j = 0$ for $1 \leq j \leq k$. Going back to the general case, what we have shown is that $b_{ij} = 0$ for all $p+1 \leq i \leq m$ and $1 \leq j \leq n_i$. This completes the proof of the claim (5).

We extend the linearly independent set $\{v_{i1} : 1 \leq i \leq p\}$ to a basis $\{v_{11}, \ldots, v_{p1}, u_1, \ldots, u_s\}$ of Ker A. Since $v_{in_i} \in \operatorname{Im} A$ for $1 \leq i \leq p$, we can find $v_i \in V$ such that $Av_i = v_{in_i}$. We claim that

$$\mathcal{B} := \{u_1, \ldots, u_s, v_{11}, \ldots, v_{1n_1}, v_1, \ldots, v_{p1}, \ldots, v_{pn_p}, v_{(p+1)1}, \ldots, v_{(p+1)n_p}, \ldots, v_{m1}, \ldots, v_{mn_m}, v_m\}$$

is a basis of V. We need only show that $\mathcal{B}$ is linearly independent. (Why?)

To prove this, let us assume that

$$0 = \sum_{i=1}^{p} \sum_{j=1}^{n_i} a_{ij} v_{ij} + \sum_{i=p+1}^{m} \sum_{j=1}^{n_i} a_{ij} v_{ij} + \sum_{r=1}^{p} b_r v_r + \sum_{k=1}^{s} c_k u_k. \tag{7}$$

Since u_k's are in Ker A and are therefore linearly independent of the vectors in $\operatorname{Im} A$. It follows that $c_k = 0$ for all $1 \leq k \leq s$. Thus (7) becomes

$$0 = \sum_{i=1}^{p} \sum_{j=1}^{n_i} a_{ij} v_{ij} + \sum_{i=p+1}^{m} \sum_{j=1}^{n_i} a_{ij} v_{ij} + \sum_{r=1}^{p} b_r v_r. \tag{8}$$

Applying A to both sides of (8), we get

$$0 = \sum_{i=1}^{p} \sum_{j=2}^{n_i} a_{ij} v_{ij-1} + \sum_{i=p+1}^{m} a_{i1} \lambda_i v_{i1} + \sum_{i=p+1}^{m} \sum_{j=2}^{n_i} a_{ij} \left(\lambda_i v_{ij} + v_{ij-1} \right) + \sum_{r=1}^{p} b_r v_{rn_r}. \tag{9}$$

The only terms that involve v_{rn_r} for $1 \le r \le p$ have coefficients b_r and hence we conclude that $b_r = 0$, $1 \le r \le p$. Using this information in (8), we get

$$0 = \sum_{i=1}^{p}\sum_{j=1}^{n_i} a_{ij} v_{ij} + \sum_{i=p+1}^{m}\sum_{j=1}^{n_i} a_{ij} v_{ij}. \tag{10}$$

The right hand side involves v_{ij}'s only once and hence their coefficients $b_{ij} = 0$ for all i, j. This establishes the linear independence of $\mathcal{B}$.

Each of u_j contributes to $J_1(0)$. Each of the string $\{v_{i1}, \ldots, v_{in_i}, v_i\}$, $1 \le i \le p$, contributes to $J_{n_i}(0)$. Each of the strings $\{v_{i1}, \ldots, v_{in_i}\}$, $p+1 \le i \le m$, contributes to $J_{n_i}(\lambda_i)$. □

Remark 17. The following are some of the important features of the Jordan canonical form of a linear map and they are very useful in determining the Jordan canonical form.

(i) The sum of the sizes of the blocks involving a fixed eigenvalue equals the algebraic multiplicity of the eigenvalue, that is, the multiplicity of the the eigenvalue as a root of the characteristic polynomial.
(ii) The number of blocks involving an eigenvalue equals its geometric multiplicity, that is, the dimension of the corresponding eigenspace .
(iii) The largest block involving an eigenvalue equals the multiplicity of the eigenvalue as a root of the minimal polynomial.

Let J be a Jordan canonical form of A. Then A and J are similar. Hence their characteristic polynomials are the same. Statement (i) follows if we observe that the eigenvalues of a Jordan block $J_k(\lambda)$ is λ with algebraic multiplicity k.

Statement (ii) follows from the observation that the eigenvalue λ of similar matrices (or linear maps) have the same geometric multiplicity and the fact that any Jordan block $J_k(\lambda)$ has one dimensional eigenspace.

Statement (iii) follows from the observations: (a) the map $T := J_k(\lambda) - \lambda I_{k\times k}$ is nilpotent with index k, that is, $T^k = 0$ but $T^{k-1} \ne 0$ and (b) if $J = \operatorname{diag}(J_{n_1}(\lambda_1), \ldots, J_{n_k}(\lambda_k))$, then its minimal polynomial is the product of the minimal polynomials of $J_{n_i}(\lambda_i)$.

Theorem 18 (Uniqueness of the Jordan Form)**.** *The Jordan form is unique apart from a permutation of the Jordan blocks.*

Proof. Let us assume that A is similar to two Jordan forms J_1 and J_2. Then there is some eigenvalue λ of A such that the corresponding blocks in J_1 and J_2 differ. As observed in the above remark (Property (ii), more precisely), the number of blocks corresponding to λ in J_1 and J_2 will be the geometric multiplicity, say, k of λ. Let $m_1 \ge m_2 \ge \cdots \ge m_k$ be the sizes of the blocks of J_1 corresponding to the eigenvalue λ. Let $n_1 \ge n_2 \ge \cdots \ge n_k$ be the sizes of the blocks in J_2. It follows that there exists some $1 \le j \le k$ such that $m - i = n_i$ for all $1 \le i \le j-1$ but $m_j \ne n_j$. Assume without loss of generality that $n_j > m_j$. But then $(J_1 - \lambda I)^{m_j} = 0$ but $(J_2 - \lambda I)^{m_j} \ne 0$. This is absurd since J_1 and J_2 are similar. □

Example 19. Consider the matrix in Example 8. It is easy to see that it is a nilpotent matrix of index 3 so that its minimal polynomial is X^3 while the characteristic polynomial is X^5. It is equally easy to see that e_1, $e_2 + e_4 - e_3$ and $e_2 + e_4 - e_5$ span the 0-eigen space. Hence the Jordan form must be $\operatorname{diag}(J_3(0), J_1(0), J_1(0)) = \operatorname{diag}(J_3(0), 0, 0)$.

Example 20. We wish to find the Jordan form of the matrix

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

The characteristic polynomial of A is $(X-1)^4$. The rank of the matrix

$$A - I_{4\times 4} = \begin{pmatrix} -3 & 5 & 1 & 0 \\ -2 & 3 & 1 & 0 \\ -1 & 2 & 0 & 0 \\ -1 & 2 & 0 & 0 \end{pmatrix}$$

is 2. Hence the geometric multiplicity of the eigenvalue $\lambda = 1$ is 2. Hence there must be two Jordan blocks in the Jordan form of A. The Jordan form of A is therefore either $\operatorname{diag}(J_3(1), J_1(1))$ or $\operatorname{diag}(J_2(1), J_2(1))$. The matrix $(A-I)^2$ can be seen to be nonzero. Hence the minimal polynomial of A cannot be $(X-1)^2$. Thus the Jordan form of A is $\operatorname{diag}(J_3(1), J_1(1))$.

Example 21. Let us consider

$$\begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

It is easy to show that the minimal polynomial of A is X^3 and the geometric multiplicity of the eigenvalue $\lambda = 0$ is 1. Hence the Jordan form of A is $J_3(0)$.

Example 22. We show that the Jordan form of $A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ -1 & 1 & 1 & 1 \\ -1 & 1 & 0 & 2 \end{pmatrix}$ is $\operatorname{diag}(J_2(1), 1, 1)$.
We expand the $\det(A - \lambda I)$ by the third column and then by the second row to see that the characteristic polynomial is $(X-1)^4$. Its is easy to see that the rank of $A - 1 \cdot I$ is 1. Hence the geometric multiplicity of $\lambda = 1$ is three. The claim follows form these observations.

Example 23. Consider $A := \begin{pmatrix} -2 & 0 & -1 & 1 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}$. Then $\lambda = -2$ is the eigenvalue of A with algebraic multiplicity 4. It is easily seen that the rank of $B := A + 2I$ is 2. Thus the Jordan form of B is either $\operatorname{diag}(J_2(0), J_2(0))$ or $\operatorname{diag}(J_3(0), 0)$. Since $B^2 = 0$, the first case occurs. Hence the Jordan form of A is $\operatorname{diag}(J_2(-2), J_2(-2))$.

Ex. 24. The characteristic polynomial of A is $(X-1)^3(X-2)^2$ and its minimal polynomial is $(X-1)^2(X-2)$. What is its Jordan form?

Ex. 25. The characteristic polynomial of A is $(X-1)^3(X-2)^2$. Write down all possible Jordan forms of A.

Ex. 26. Find all possible Jordan forms of an 8×8 matrix whose characteristic polynomial is $(X-1)^4(X-2)^4$ and the minimal polynomial $(X-1)^2(X-2)^2$ if the geometric multiplicity of the eigenvalue $\lambda = 1$ is three.

Ex. 27. Show that any square matrix A is similar to its transpose. *Hint:* If A is similar to J what is A^T similar to? Also, see Remark 4.

Ex. 28. Two triangulable matrices are similar iff they have the 'same' Jordan forms.

Ex. 29 ($S+N$ decomposition)**.** Given $A\colon V \to V$, there exist linear maps S and N with the following properties: (i) $A = S+N$, (ii) S is diagonalizable and N is nilpotent, (iii) $NS = SN$.

Ex. 30. Consider the two matrices

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Show that their characteristic polynomial is $(X-1)^4$ and the minimal polynomial is $(X-1)^2$, but they do not have the same Jordan form. (Question: What are the Jordan forms of the given matrices?) Thus for two matrices to be similar it is necessary but not sufficient that they have the same characteristic and the same minimal polynomial.

Ex. 31. Show that if $A \in M(n, \mathbb{C})$ is such that $A^n = I$, then A is a diagonalizable.

Ex. 32. Prove that if $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A and if $p(X)$ is a polynomial, then $p(\lambda_i)$, $1 \le i \le n$, are the eigenvalues of $p(A)$.

Ex. 33. If $A := \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$ show that $A^{50} = 2^{50} \begin{pmatrix} -24 & 25 \\ -25 & 26 \end{pmatrix}$.

Acknowledgements: I thank Dr. Jayanthan of Goa, Dr. Sushama Agrawal and Thiru Sukumar of Ramanujan Institute, Chennai for a careful reading ans pointing out typographical mistakes and making suggestions for the clarity of the exposition.