# 11.D Jordan Normal Form

We consider triagonalisable operators $f$ on finite dimensional $K$-vector spaces $V$. We shall improve the description of such operators given in 11.C.6.

Suppose that $\chi_f = (X-\lambda_1)^{\alpha_1} \cdots (X-\lambda_r)^{\alpha_r}$ is the characteristic polynomial of $f: V \longrightarrow V$ with pairwise distinct zeros $\lambda_1, \ldots, \lambda_r \in K$ of multiplicities $\alpha_1, \ldots, \alpha_r$. Then the minimal polynomial of $f$ is of the form $\mu_f = (X-\lambda_1)^{\beta_1} \cdots (X-\lambda_r)^{\beta_r}$, with $0 \le \beta_i \le \alpha_i$ for $i = 1, \ldots, r$. The primary-components $V_i = \mathrm{Ker}\,(f - \lambda_i \mathrm{id}_V)^{\beta_i}$, $i = 1, \ldots, r$, in general can be further decomposed into $f$-invariant subspaces, on these subspaces $f$ operates in particular in a simple way. For this we may assume that $V$ itself is such a primary component, see 11.C.6. Then $\mu_f = (X-\lambda)^\beta$. We consider the chain of so-called <u>higher eigen-spaces</u>

$$V_f^{(i)} := V^{(i)} := \mathrm{Ker}\,(f - \lambda\,\mathrm{id})^i, \quad i = 1, \ldots, \beta.$$

Then $0 = V^{(0)} \subsetneq V^{(1)} \subsetneq \cdots \subsetneq V^{(\beta)} = V$

all these inclusions are strict. For the proof we need the following lemma:

**11.D.1 Lemma** *Let $g: V \longrightarrow V$ be an operator on the $K$-vector space $V$. For every $i \in \mathbb{N}$, $g$ induces an injective homomorphism*

$$\mathrm{Ker}\,g^{i+2} / \mathrm{Ker}\,g^{i+1} \xrightarrow{\bar{g}} \mathrm{Ker}\,g^{i+1} / \mathrm{Ker}\,g^i .$$

*In particular, $\operatorname{Ker} g^i = \operatorname{Ker} g^{i_0}$ for all $i \geq i_0$, if $\operatorname{Ker} g^{i_0+1} = \operatorname{Ker} g^{i_0}$.*

**Proof** Clearly for every $i \in \mathbb{N}$, $g$ maps the subspace $\operatorname{Ker} g^{i+1}$ in $\operatorname{Ker} g^i$. Therefore there exists the given homomorphism $\overline{g}$. For the proof of its injectivity, let $x \in \operatorname{Ker} g^{i+2}$ with $\overline{g}(\overline{x}) = \overline{g(x)} = 0$, i.e. $g(x) \in \operatorname{Ker} g^i$. Then $0 = g^i(g(x)) = g^{i+1}(x)$, i.e. $x \in \operatorname{Ker} g^{i+1}$ and so $\overline{x} = 0$ in $\operatorname{Ker} g^{i+2}/\operatorname{Ker} g^{i+1}$.

We now apply 11.D.1 to $g := f - \lambda\, \mathrm{id}$. Since $g^{\beta-1} \neq 0$, $\operatorname{Ker} g^{\beta-1} =: V^{(\beta-1)} \subsetneq V = \operatorname{Ker} g^{\beta}$. This proves that all inclusions $V^{(i)} \subsetneq V^{(i+1)}$ are strict for $i = 0, \ldots, \beta-1$. Further, we have the following chain of injective homomorphisms:

$$0 = V^{(\beta+1)}/V^{(\beta)} \xrightarrow{\overline{g}} V^{(\beta)}/V^{(\beta-1)} \xrightarrow{\overline{g}} V^{(\beta-1)}/V^{(\beta-2)} \xrightarrow{\overline{g}}$$

$$\cdots \xrightarrow{\overline{g}} V^{(2)}/V^{(1)} \xrightarrow{\overline{g}} V^{(1)}/V^{(0)} = V^{(1)}.$$

Now, for $j = 1, \ldots, \beta$, suppose that

$$v_1^{(j)}, \ldots, v_{n_j}^{(j)}, \quad n_j := \operatorname{Dim}_K V^{(j)}/V^{(j-1)} - \operatorname{Dim}_K V^{(j+1)}/V^{(j)}$$

are vectors in $V^{(j)}$ such that their residue-classes in $V^{(j)}/V^{(j-1)}$ form a basis of a complement of $\overline{g}\left(V^{(j+1)}/V^{(j)}\right)$ in $V^{(j)}/V^{(j-1)}$. Then

$$v_1^{(1)}, \dots, v_{n_1}^{(1)}, \quad v_1^{(2)}, g v_1^{(2)}, \dots, v_{n_2}^{(2)}, g v_{n_2}^{(2)},$$

$$\dots\dots$$

$$v_1^{(\beta)}, g v_1^{(\beta)}, \dots, g^{\beta-1} v_1^{(\beta)}, \dots, v_{n_\beta}^{(\beta)}, g v^{(\beta)}, \dots, g^{\beta-1} v_{n_\beta}^{(\beta)}$$

is clearly a basis of $V = V^{(\beta)}$. Further, since $g(x) = (f - \lambda\,\mathrm{id})(x) = f(x) - \lambda x$, i.e.

$$f(x) = \lambda x + g(x)$$

and $g^j v_1^{(j)} = \dots = g^j v_{n_j}^{(j)} = 0$ for $j = 1, \dots, \beta$,

the matrix of $f$ with respect to this basis is the diagonal-block matrix

$$\mathrm{Diag}\left(\mathfrak{N}_1^{(1)}, \dots, \mathfrak{N}_{n_1}^{(1)}, \dots, \mathfrak{N}_1^{(\beta)}, \dots, \mathfrak{N}_{n_\beta}^{(\beta)}\right),$$

where the matrices $\mathfrak{N}_i^{(j)}$ are of the form:

$$\mathcal{J}^{(j)}(\lambda) = \begin{pmatrix} \lambda & 0 & \cdots & 0 & 0 \\ 1 & \lambda & & 0 & 0 \\ \vdots & \vdots & & & \\ 0 & 0 & & \lambda & 0 \\ 0 & 0 & & 1 & \lambda \end{pmatrix} \in M_j(K)$$

Conversely, if the matrix of $f$ with respect to a basis of $V$ is of such a form, then the number of the blocks $\mathcal{J}^{(j)}(\lambda)$ of length $j$ in this matrix is equal to

$$n_j = \mathrm{Dim}_K V^{(j)}/V^{(j-1)} - \mathrm{Dim}_K V^{(j+1)}/V^{(j)}$$

$$= 2\,\mathrm{Dim}_K V^{(j)} - \mathrm{Dim}_K V^{(j-1)} - \mathrm{Dim}_K V^{(j+1)},$$

$j \geq 1$ and hence independent of the choice of the basis and is uniquely determined by $f$. The matrices of the type $\mathcal{J}^{(j)}(\lambda)$, $\lambda \in K$, are called Jordan-matrices over $K$ and a matrix in the above block-form with Jordan-matrices in the main-diagonal, where the diagonal elements $\lambda$ for distinct blocks are also distinct, is said to be in Jordan-Normal form. With this we have proved:

**11.D.2 Jordan-Normal-Form** For every triagonalisable operator $f$ on a finite dimensional $K$-vector space $V$, there exists a basis of $V$ such that the matrix of $f$ with respect this basis is a matrix in Jordan-normal-form. Moreover, the number $n_j(\lambda)$ of Jordan-blocks $\mathcal{J}^{(j)}(\lambda)$ of length $j$ corresponding to the eigen-value $\lambda$, appearing are uniquely determined and is equal to

$$2\,\mathrm{Dim}_K \operatorname{Ker}(f-\lambda\,\mathrm{id})^j - \mathrm{Dim}_K \operatorname{Ker}(f-\lambda\,\mathrm{id})^{j-1} - \mathrm{Dim}_K \operatorname{Ker}(f-\lambda\,\mathrm{id})^{j+1}.$$

For matrices the formulation is: Every triagonalisable matrix in $\mathrm{M}_n(K)$ is similar to a matrix in Jordan-normal form. Moreover, this is uniquely determined upto an order of the Jordan-blocks.

**11.D.3 Example** We consider the operator $f: \mathbb{R}^5 \to \mathbb{R}^5$ with the matrix with respect to the standard basis

$$\mathfrak{A} := \begin{pmatrix} 4 & -4 & 9 & 7 & 11 \\ 1 & 0 & 4 & 4 & 6 \\ 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 & 3 \end{pmatrix} \in \mathbb{M}_5(\mathbb{R}).$$

Then $\chi_f = \chi_{\mathfrak{A}} = \operatorname{Det}(X\mathfrak{E}_5 - \mathfrak{A}) = (X-2)^5$.

Further, $V^{(1)} := \operatorname{Ker}(f - 2\,\mathrm{id}) = V_f(2)$

$$= \mathbb{R}\,{}^t(2,1,0,0,0) \oplus \mathbb{R}\,{}^t(-2,0,0,-1,1)$$

$$V^{(2)} = \operatorname{Ker}(f-2\,\mathrm{id})^2 =$$

$$= V^{(1)} \oplus \mathbb{R}\,{}^t(1,0,0,0,0) \oplus \mathbb{R}\,{}^t(0,0,-1,1,0)$$

$$V^{(3)} = \operatorname{Ker}(f-2\,\mathrm{id})^3 = \mathbb{R}^5 = V^{(2)} \oplus \mathbb{R}\,{}^t(0,0,0,0,1)$$

The vector $v_1^{(3)} = (0,0,0,0,1)$ extends a basis of $V^{(2)}$ to a basis of $V^{(3)} = V$ and therefore form a basis of $V^{(3)}/V^{(2)}$. Moreover, $g(v_1^{(3)}) = {}^t(11,6,-1,0,1) \in V^{(2)}$ with $g := f - 2\,\mathrm{id}$.
The vector $v_1^{(2)} := {}^t(0,0,-1,1,0)$ together with $g(v_1^{(3)})$ and a ~~basis of~~ $V^{(1)}$ to a basis of $V^{(2)}$. Its residue class in $V^{(2)}/V^{(1)}$, therefore generate a complement of $\bar{g}(V^{(3)}/V^{(2)})$. It follows that the matrix of $f$ with respect to the basis $v_1^{(2)} = {}^t(0,0,-1,1,0)$, $gv_1^{(2)} = {}^t(-2,0,0,-1,1)$, $v_1^{(3)} = {}^t(0,0,0,0,1)$, $gv_1^{(3)} = {}^t(11,6,-1,0,1)$, $g^2v_1^{(3)} = {}^t(0,1,0,-1,1)$ is the Jordan-form:

$$\left(\begin{array}{cc|ccc} 2 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ \hline 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 & 2 \end{array}\right).$$

**11.D.4 Remark** One can deduce Theorem 11.D.2 on the Jordan Normal form very elegantly and in a more general form (any field $K$), from the Elementary divisor theorem 10.A.31 for polynomials in $K[X]$. At the same time these methods supply an important algorithmic process to transform a matrix to the Jordan-normal form.

Let $f: V \longrightarrow V$ be a $K$-linear operator on the $n$-dimensional $K$-vector space $V$. Using the primary decomposition theorem 11.C.5, we may assume that the characteristic polynomial $\chi_f$ of $f$ is the power $P^\alpha$, $\alpha \geq 1$, of a monic prime polynomial $P \in K[X]$ of degree $m$.[1] Then $n = m\alpha$ and the minimal polynomial $\mu_f$ of $f$ is $P^\beta$, $1 \leq \beta \leq \alpha$.

Let $\underline{v} = \{v_1, \ldots, v_n\}$ be a basis of $V$ and $\mathfrak{A} = \mathfrak{M}_{\underline{v}}^{\underline{v}}(f) = (a_{ij})$ be the matrix of $f$ with respect to $\underline{v}$ (For $j \in I$, $f(v_j) = \sum_{i \in I} a_{ij} v_i$, $I = \{1, \ldots, n\}$). We consider the surjective $K$-linear map $\varepsilon : K[X]^n \longrightarrow V$ with

---

[1] This reduction is often unnecessary. The following process supply a basis of $V$ such that the matrix of $f$ is a diagonal-block matrix $\mathrm{Diag}(\mathfrak{A}_{E_1}, \ldots, \mathfrak{A}_{E_n})$, where the polynomials $E_1, \ldots, E_n$ are the elementary divisors of the matrix $X\mathfrak{E}_n - \mathfrak{A} \in \mathrm{M}_n(K[X])$ and $\mathfrak{A}_{E_1}, \ldots, \mathfrak{A}_{E_n}$ are their companion matrices. This representation of $f$ (resp. $\mathfrak{A}$) is known as the first normal-form or the rational canonical form of the operator $f$ (resp. the matrix $\mathfrak{A}$).

$\varepsilon(F_1, \cdots, F_n) := \sum_{j=1}^{n} F_j(f)(v_j)$. The characteristic matrix $\mathfrak{X} := X\mathfrak{E}_n - \mathfrak{A} \in M_n(K[X])$ and its determinant $\operatorname{Det}\mathfrak{X} = \chi_f$ is the $\in K[X]$ characteristic polynomial of $f$ which is $\chi_f = P^\alpha$ by assumption. The matrix $\mathfrak{X}$ defines the $K$-linear map $K[X]^n \xrightarrow{\mathfrak{X}} K[X]^n$ with $F = {}^t(F_1, \cdots, F_n) \longmapsto \mathfrak{X}F$ defined by the matrix multiplication. Then the following sequence is exact:

$$0 \longrightarrow K[X]^n \xrightarrow{\mathfrak{X}} K[X]^n \xrightarrow{\varepsilon} V \longrightarrow 0$$

Proof First we shall prove that $\operatorname{Im}\mathfrak{X} \subseteq \operatorname{Ker}\varepsilon$: i.e. $\varepsilon \circ \mathfrak{X} = 0$. Let $F = {}^t(F_1, \cdots, F_n) \in K[X]^n$. Then

$$\varepsilon(\mathfrak{X}F) = \varepsilon(XF - \mathfrak{A}F) = \sum_{j=1}^{n} f F_j(f)(v_j) - \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} F_j(v_i) = \sum_{j=1}^{n} F_j(f)(f(v_j)) - \sum_{j=1}^{n} F_j(f)(f(v_j)) \; (= 0.)$$

(remember $f(v_j) = \sum_{i=1}^{n} a_{ij} v_i$ for every $j = 1, \cdots, n$)

From the Theorem 10.A.34, we have

$\operatorname{Dim}_K \operatorname{Coker}\mathfrak{X} = \deg(\operatorname{Det}\mathfrak{X}) = \deg \chi_f = n = \operatorname{Dim}_K V$. Therefore $\operatorname{Im}\mathfrak{X} = \operatorname{Ker}\varepsilon$. Moreover, since $\operatorname{Det}\mathfrak{X} \neq 0$, $\operatorname{Rank}\mathfrak{X} = n$ and hence multiplication by $\mathfrak{X}$ is injective. ■

By the Elementary divisor theorem 10.A.31, there exists elementary matrices

$\in M_{\alpha_i m}(K)$. This is an $\alpha_i \times \alpha_i$ block-matrix with entries in $M_m(K)$, where

$$\mathfrak{A}_P = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & -a_{m-2} \\ 0 & 0 & \cdots & 1 & -a_{m-1} \end{pmatrix} \in M_m(K)$$

is the companion matrix of the polynomial $P$ and

$$\mathcal{E}_{1m} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & & \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \in M_m(K)$$

In the special case $P = X - a$, the (standard) Jordan-matrix $\mathfrak{J}^{(\alpha_i)}(X-a)$ is simply denoted by $\mathfrak{J}^{(\alpha_i)}(a)$. Altogether, we have proved that:

**11.D.5 Theorem** Let $f: V \longrightarrow V$ be a $K$-linear operator on the finite dimensional $K$-vector space $V$. Then there exist a decomposition $V = V_1 \oplus \cdots \oplus V_r$ of $V$ into $f$-invariant subspaces $V_1, \ldots, V_r$ of $V$ such that there is a basis of $V_i$ such that the matrix of $f|V_i : V_i \longrightarrow V_i$ is of the form $\mathfrak{J}^{(\alpha_i)}(P_i)$ with a prime polynomial $P_i \in K[X]$ and $\alpha_i \in \mathbb{N}^*$.

In the situation of Theorem 11.D.5 the characteristic polynomial of $f|V_i$ (and the minimal

Note that: since $\sum_{i=1}^{n} \alpha_i = \alpha \leq n$, many of the $\alpha_i$'s are 0. When exactly all $\alpha_i \neq 0$ ?.

Let $h_i : K[X]/_{K[X]P^{\alpha_i}} \xrightarrow{\approx} V_i$ be defined by using the isomorphism $h : \mathrm{Coker}\,\vartheta \xrightarrow{\approx} V$ and let $u_i = h_i(\bar{1})$. Clearly $h_i(H) = H(f)(u_i)$ for all $H \in K[X]$. We put $x := \bar{X} \in K[X]/_{K[X]P^{\alpha_i}}$

Then the K-basis (of $K[X]/_{K[X]P^{\alpha_i}}$)

$$1, x, \dots, x^{m-1}, P(x), xP(x), \dots, x^{m-1}P(x), \dots$$
$$P^{\alpha_i-1}(x), xP^{\alpha_i-1}(x), \dots, x^{m-1}P^{\alpha_i-1}(x)$$

of $K[X]/_{K[X]P^{\alpha_i}}$ correspond to the K-basis

$$u_i, f(u_i), \dots, f^{m-1}(u_i), P(f)(u_i), fP(f)(u_i), \dots, f^{m-1}P(f)(u_i), \dots$$
$$P^{\alpha_i-1}(f)(u_i), fP^{\alpha_i-1}(f)(u_i), \dots, f^{m-1}P^{\alpha_i-1}(f)(u_i)$$

of $V_i$.

If $P = X^m + a_{m-1}X^{m-1} + \cdots + a_1X + a_0$, then the matrix of $f|V_i : V_i \longrightarrow V_i$ with respect to this basis is the <u>generalised Jordan-matrix</u> $\mathcal{J}^{(\alpha_i)}(P)$:

$$\mathcal{J}^{(\alpha_i)}(P) = \begin{pmatrix} \mathfrak{M}_P & 0 & 0 & \cdots & 0 & 0 \\ \mathcal{E}_{1m} & \mathfrak{M}_P & 0 & \cdots & 0 & 0 \\ 0 & \mathcal{E}_{1m} & \mathfrak{M}_P & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \mathfrak{M}_P & 0 \\ 0 & 0 & 0 & \cdots & \mathcal{E}_{1m} & \mathfrak{M}_P \end{pmatrix}$$

$\sigma_1, \dots, \sigma_p ; \tau_1, \dots, \tau_q \in GL_n(K[X])$

such that

$$\sigma_1 \cdots \sigma_p \, \mathfrak{X} \, \tau_1 \cdots \tau_q = \mathfrak{D} = \mathrm{Diag}(E_1, \dots, E_n)$$

is the diagonal matrix.

Since $\mathrm{Det}\,\sigma_1 = \cdots = \mathrm{Det}\,\sigma_p = \mathrm{Det}\,\tau_1 = \cdots = \mathrm{Det}\,\tau_q = 1$, we have $E_1 \cdots E_n = \mathrm{Det}\,\mathfrak{D} = \mathrm{Det}\,\mathfrak{X} = P^{\alpha}$ (by assumption).

Now, since $P$ is prime in $K[X]$, there exist elements $\varepsilon_i \in K^{\times}$ and $\alpha_i \in \mathbb{N}$ with $E_i = \varepsilon_i P^{\alpha_i}$ for $i = 1, \dots, n$. Using the invertibility of matrices $\sigma := \sigma_1 \cdots \sigma_p$ and $\tau := \tau_1 \cdots \tau_q$ in $M_n(K[X])$ the following diagram is commutativ:

$$\begin{tikzcd}
K[X]^n \arrow[r, "\mathfrak{D}"] \arrow[d, "\tau"', "\wr"] & K[X]^n \arrow[d, "\sigma^{-1}", "\wr"'] \\
K[X]^n \arrow[r, "\mathfrak{X}"] & K[X]^n
\end{tikzcd}$$

Therefore $\sigma^{-1}$ induces an isomorphism $h$ on the cokernels: $h : \mathrm{Coker}\,\mathfrak{D} \xrightarrow{\approx} \mathrm{Coker}\,\mathfrak{X} \,(= V)$.

But the cokernel $\mathrm{Coker}\,\mathfrak{D}$ of $\mathfrak{D}$ is the direct sum

$$\mathrm{Coker}\,\mathfrak{D} = \bigoplus_{i=1}^{n} K[X]/K[X]E_i = \bigoplus_{i=1}^{n} K[X]/K[X]P^{\alpha_i}$$

This direct sum decomposition of $\mathrm{Coker}\,\mathfrak{D}$ corresponds to the direct sum decomposition of $V$:

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_n \text{ with } \mathrm{Dim}_K V_i = \alpha_i m, \quad i = 1, \dots, n.$$

polynomial) of $f|V_i$ is $P_i^{\alpha_i}$. Further, we have $\chi_f = \prod_{i=1}^{r} P_i^{\alpha_i}$ and $\mu_f = \prod P^{\beta_P}$, where $\beta_P = \text{Min}\{\alpha_i \mid P_i = P\}$ and $P$ runs through all monic prime polynomials in $K[X]$.

Moreover, the matrices $\mathcal{J}^{(\alpha_i)}(P_i)$ are uniquely determined by $f$ upto an order. This corresponds to the uniqueness assertion in elementary divisor theorem and can also be proved similarly as the uniqueness assertion in 11.D.2; this is left to the reader. In particular, Theorem 11.D.5 give a complete classification of the linear operators on finite dimensional vector spaces upto similarity.

If $K = \mathbb{R}$, then all prime polynomials are either linear or quadratic. Other than the Jordan-matrices in 11.D.2, we need only to consider the following matrices (with $p, q \in \mathbb{R}$ and $p^2 < 4q$)

$$\begin{pmatrix} 0 & -q & 0 & 0 & & & \\ 1 & -p & 0 & 0 & & & \\ 0 & 1 & 0 & -q & & & \\ 0 & 0 & 1 & -p & & & \\ 0 & 0 & 0 & 1 & \ddots & & \\ & & & & \ddots & 0 & -q \\ & & & & & 1 & -p \end{pmatrix}$$