5 The Rational Canonical Form

Here p is a monic irreducible factor of the minimum polynomial m_T and is not necessarily of degree one.

Let F_p denote the field constructed earlier in the course, consisting of all matrices of the form $f(B)$, $f \in F[x]$, where $B = C(p)$, the companion matrix of p. (We saw that if $\deg p = n$, then

$$F_p = \{a_0 I_n + \cdots + a_{n-1} B^{n-1} | \, a_0, \ldots, a_{n-1} \in F\}.$$

Let $\overline{f} = f(B)$, where $f \in F[x]$. Then this new symbol has the following properties:

(i) $\overline{f} + \overline{g} = \overline{f+g}$; $\overline{f}\,\overline{g} = \overline{fg}$;

(ii) $\overline{f} = \overline{0} \Leftrightarrow p \mid f$;

(iii) $\overline{f} = \overline{g} \Leftrightarrow p \mid (f-g)$;

(iv) $\overline{f}^{-1}$ exists $\Leftrightarrow p$ does not divide f.

Note: If $p = x - c$, then $F_p = F$.

THEOREM 5.1

$N_{h,p}$ becomes a vector space over F_p if we define

$$\overline{f}v = fv = f(T)(v).$$

First we must verify that the above definition is well–defined, that is, independent of the particular polynomial f used to define the field element $\overline{f}$. So suppose $\overline{f} = \overline{g}$. Then $f = g + kp$, $k \in F[x]$. Hence

$$fv = (g + kp)v = gv + k(pv) = gv + k0 = gv,$$

as $v \in \operatorname{Im} p^{h-1}(T) \cap \operatorname{Ker} p(T)$ and consequently $pv = 0$.

The four addition axioms hold as V is already a vector space over F; The remaining vector space axioms then follow from the left $F[x]$–module axioms:

(i) $(\overline{f} + \overline{g}) = \overline{(f+g)}v = (f+g)v = fv + gv = \overline{f}v + \overline{g}v$;

(ii) $\overline{f}(v+w) = f(v+w) = fv + fw = \overline{f}v + \overline{f}w$;

(iii) $\overline{f}(\overline{g}v) = \overline{f}(gv) = f(gv) = (fg)v = (\overline{fg})v$;

(iv) $\overline{1}v = 1v = v$.

Remark: An F–basis for $N_{h,\,p}$ will be an F_p–spanning family for $N_{h,\,p}$, but will not, in general, be an F_p–basis for $N_{h,\,p}$. The precise connection between F–independence and F_p–independence is given by the following theorem:

THEOREM 5.2

Vectors $v_1, \ldots, v_r$ form an F_p–basis for $N_{h,\,p}$ if and only if the vectors

$$\begin{array}{cccc} v_1, & T(v_1), & \ldots, & T^{n-1}(v_1) \\ v_2, & T(v_2), & \ldots, & T^{n-1}(v_2) \\ \vdots & \vdots & \vdots & \vdots \\ v_r, & T(v_r), & \ldots, & T^{n-1}(v_r) \end{array}$$

form an F–basis for $N_{h,\,p}$.

COROLLARY 5.1

$$\nu_{h,\,p} = \dim_{F_p} N_{h,\,p} = \frac{1}{\deg p} \dim_F N_{h,\,p} = \frac{\nu(p^h(T)) - \nu(p^{h-1}(T))}{\deg p}.$$

The exposition for $p = x - c$ now goes over to general p, with small changes. We again have the decreasing sequence of dimensions:

$$\nu_{1,\,p} \geq \cdots \geq \nu_{b,\,p} \geq 1,$$

where $\nu_{1,\,p} = \dim_{F_p} \operatorname{Ker} p(T) = \frac{\nu(p(T))}{\deg p}$.

Also

$$\nu_{1,\,p} + \cdots + \nu_{b,\,p} = \frac{\nu(p^b(T))}{\deg p}, \tag{24}$$

where $p^b \parallel \mathrm{m}_T$.

There is a corresponding dot diagram where the number of dots in the h–th row from the bottom represents the integer $\nu_{h,\,p}$. We also have a similar theorem to an earlier one, in terms of the conjugate partition

$$e_1 \geq \cdots \geq e_\gamma \geq 1$$

of the partition (24) above, where $\gamma = \nu_{1,\,p} = \dim_{F_p} \operatorname{Ker} p(T) = \frac{\nu(p(T))}{\deg p}$.

THEOREM 5.3

Vectors $v_1, \ldots, v_\gamma \in V$ can be found with the property that

$$p^{e_1-1}v_1, \ldots, p^{e_\gamma-1}v_\gamma$$

form an F_p–basis for $\operatorname{Ker} p(T)$. *Moreover*

(i) $m_{T,\,v_j} = p^{e_j}$;

(ii) $\operatorname{Ker} p^b(T) = C_{T,\,v_1} \oplus \cdots \oplus C_{T,\,v_\gamma}$.

In conclusion, if $m_T = p_1^{b_1} \cdots p_t^{b_t}$, we now have the direct sum decomposition

$$V = \bigoplus_{i=1}^{t} \bigoplus_{j=1}^{\gamma_i} C_{T,\,v_{ij}},$$

where $m_{T,\,v_{ij}} = p_i^{e_{ij}}$ and

$$e_{i1} = b_i \geq \ldots \geq e_{i\gamma_i}$$

form the conjugate partition for the dot diagram corresponding to p_i. Here

$$\gamma_i = \frac{\nu(p_i(T))}{\deg p_i}.$$

Taking T–cyclic bases β_{ij} for $C_{T,\,v_{ij}}$, then gives a basis

$$\beta = \bigcup_{i=1}^{t} \bigcup_{j=1}^{\gamma_i} \beta_{ij}$$

for V. Moreover

$$[T]_\beta^\beta = \bigoplus_{i=1}^{t} \bigoplus_{j=1}^{\gamma_i} C(p_i^{e_{ij}}).$$

The matrix on the right is said to be in **rational canonical form**.

If instead, we take the following basis β'_{ij} for $C_{T,\,v_{ij}}$

$$\beta'_{ij} : \begin{cases} v_{ij}, & T(v_{ij}), & \ldots, & T^{n-1}(v_{ij}) \\ p_i(T)(v_{ij}), & Tp_i(T)(v_{ij}), & \ldots, & T^{n-1}p_i(T)(v_{ij}) \\ \vdots & \vdots & \vdots & \vdots \\ p_i^{e_{ij}-1}(T)(v_{ij}), & Tp_i^{e_{ij}-1}(T)(v_{ij}), & \ldots, & T^{n-1}p_i^{e_{ij}-1}(T)(v_{ij}), \end{cases}$$

(with $n = \deg p_i$) which reduces to the Jordan basis when $p_i = x - c_i$, it is not difficult to verify that we get a corresponding matrix $H(p_i^{e_{ij}})$ called a **hypercompanion** matrix, which reduces to the elementary Jordan matrix $J_{e_{ij}}(c_i)$ when $p_i = x - c_i$:

$$H(p_i^{e_{ij}}) = \begin{bmatrix} C(p_i) & 0 & \cdots & 0 \\ N & C(p_i) & \cdots & 0 \\ 0 & N & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & N & C(p_i) \end{bmatrix},$$

where there are e_{ij} blocks on the diagonal and N is a square matrix of same size as $C(p_i)$ which is everywhere zero, except in the top right–hand corner, where there is a 1. The overall effect is an unbroken subdiagonal of $1'$s.

We then get the corresponding rational canonical form:

$$[T]_{\beta'}^{\beta'} = \bigoplus_{i=1}^{t} \bigoplus_{j=1}^{\gamma_i} H(p_i^{e_{ij}}).$$

Computational Remark:

We can do our computations completely over F, without going into F_p, as follows. Suppose $v_1, \ldots, v_r$ form an F–spanning family for $N_{h,\,p}$. Then we could, in principle, perform the LRA over F_p on this spanning family and find an F_p–basis $v_{c_1}, \ldots, v_{c_R}$. A little thought reveals that if we had instead applied the LRA algorithm over F to the expanded sequence:

$$v_1,\, T(v_1),\, \ldots, T^{n-1}(v_1); \ldots;\, v_r,\, T(v_r), \ldots, T^{n-1}(v_r),$$

we would have obtained the F–basis for $N_{h,\,p}$:

$$v_{c_1},\, T(v_{c_1}),\, \ldots, T^{n-1}(v_{c_1}); \ldots;\, v_{c_R},\, T(v_{c_R}), \ldots, T^{n-1}(v_{c_R})$$

from which we select the desired F_p–basis $v_{c_1}, \ldots, v_{c_R}$.

Let $A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 2 \\ 1 & 0 & 0 & 0 & 2 & 1 \\ 0 & 1 & 0 & 0 & 2 & 2 \\ 2 & 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \in M_{6\times 6}(\mathbb{Z}_3)$.

Here $m_A = p^2,\ p = x^2 + x + 2 \in F[x],\ F = \mathbb{Z}_3$.

$$p(A) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 2 & 1 & 0 \\ 0 & 1 & 2 & 2 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 2 \\ 0 & 0 & 1 & 2 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \nu(p(A)) = 4,\ \nu_{1,p} = \tfrac{\nu(p(A))}{\deg p} = 2.$$

$$p^2(A) = 0,\ \nu(p^2(A)) = 6,\ \nu_{2,p} = \frac{\nu(p^2(A)) - \nu(p(A))}{\deg p} = \frac{6-4}{2} = 1.$$

Hence we have a corresponding F_p dot diagram:



We have to find an F_p–basis $p(A)v_{11}$ for $N_{2,p}$ and extend this to an F_p–basis $p(A)v_{11},\ v_{12}$ for $N(p(A))$.

An F–basis for $N(p^2(A))$ is $E_1, \ldots, E_6$. Then

$$N_{2,p} = \langle p(A)E_1, \ldots, p(A)E_6 \rangle$$

and the LRA give $p(A)E_2$ as an F_p–basis for $N_{2,p}$ so we can take $v_{11} = E_2$.

We find the columns of the following matrix form an F–basis for $N(p(A))$:

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

We place $p(A)E_2$ in front and then pad the resulting matrix to get

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

The first four columns $p(A)E_2,\ Ap(A)E_2,\ E_1,\ AE_1$ of this matrix form a LR F–basis for $N(p(A))$ and hence $p(A)E_2,\ E_1$ form an F_p–basis for $N(p(A))$. So we can take $v_{12} = E_1$.

Then $V_6(\mathbb{Z}_3) = N(p^2(A)) = C_{T_A,\, v_{11}} \oplus C_{T_A,\, v_{12}}$.
Then joining hypercompanion bases for $C_{T_A,\, v_{11}}$ and $C_{T_A,\, v_{12}}$:

$$v_{11},\ Av_{11},\ p(A)v_{11},\ Ap(A)v_{11} \quad \text{and} \quad v_{12},\ Av_{11}$$

gives a basis v_{11}, Av_{11}, $p(A)v_{11}$, $Ap(A)v_{11}; v_{12}$, Av_{11} for $V_6(\mathbb{Z}_3)$. Finally if P is the non–singular matrix whose columns are these vectors, we transform A into direct sum of hypercompanion matrices:

$$P^{-1}AP = H(p^2) \oplus H(p) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

Explicitly, we have

$$P = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 2 & 0 & 0 & 1 \\ 0 & 1 & 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

5.1 Uniqueness of the Rational Canonical Form

Suppose that $T : V \to V$ is a linear transformation over F and that β is a basis for V such that

$$[T]_\beta^\beta = \bigoplus_{i=1}^{t} \bigoplus_{j=1}^{\gamma_i} C(p_i^{e_{ij}}). \tag{25}$$

where

$$e_{i1} \geq \ldots \geq e_{i\gamma_i} \geq 1 \tag{26}$$

and $p_1, \ldots, p_t$ are distinct monic irreducible polynomials.

We show that the polynomials p_i and the sequences (26) are determined by the transformation T.

First, it is not difficult to show that

$$\beta = \bigcup_{i=1}^{t} \bigcup_{j=1}^{\gamma_i} \beta_{ij},$$

where

$$\beta_{ij} : v_{ij},\, T(v_{ij}), \ldots, T^{n_{ij}-1}(v_{ij})$$

and $n_{ij} = \deg p_i^{e_{ij}}$ and $m_{T,\,v_{ij}} = p_i^{e_{ij}}$. Then we have the direct sum decomposition

$$V = \bigoplus_{i=1}^{t} \bigoplus_{j=1}^{\gamma_i} C_{T,\,v_{ij}}.$$

Also if we write $b_i = e_{i1}$, we have

$$\operatorname{Ker} p_i^{b_i}(T) = \bigoplus_{j=1}^{\gamma_i} C_{T,\,v_{ij}}$$

and hence

$$V = \bigoplus_{i=1}^{t} \operatorname{Ker} p_i^{b_i}(T).$$

Then from equation (25) above, it follows that

$$\mathrm{m}_T = \operatorname{lcm} p_i^{e_{ij}} = p_1^{b_1} \cdots p_t^{b_t},$$

thereby determining $p_1, \ldots, p_t$ up to order.

Then it can be shown that if $1 \leq h \leq b_i$, then $N_{h,\,p_i}$ has F_{p_i} basis

$$p_i^{e_{i1}-1} v_{i1}, \ldots, p_i^{e_{ij_h}-1} v_{ij_h},$$

where $e_{i1}, \ldots, e_{ij_h}$ are the integers not less than h.

There are consequently $\dim_{F_{p_i}} N_{h,\,p_i} = \nu_{h,\,p_i}$ such integers and hence the number of integers $e_{i1}, \ldots, e_{i\gamma_i}$ equal to h is equal to $\nu_{h,\,p_i} - \nu_{h+1,\,p_i}$, which depends only on T. In other words, for each i, the sequence $e_{i1}, \ldots, e_{i\gamma_i}$ depends only on T.

5.2 Deductions from the Rational Canonical Form

THEOREM 5.4

$$\frac{\nu(p_i^{b_i}(T))}{\deg p_i} = a_i$$

where $p_i^{a_i} || \operatorname{ch}_T$, *and* $p_i^{b_i} || m_T$.

Note that this determines b_i—we may evaluate

$$\frac{\nu(p_i^h(T))}{\deg p_i}$$

for $h = 1, 2, \ldots$ until we get a value of a_i. Then that $h = b_i$.
PROOF $\exists$ a basis for V such that

$$A = [T]_\beta^\beta = \bigoplus_{i=1}^{t} \bigoplus_{j=1}^{\gamma_i} C(p_i^{e_{ij}}).$$

So

$$\mathrm{ch}_T = \prod_{i=1}^{t} \prod_{j=1}^{\gamma_i} \mathrm{ch}_{B_{i,j}}$$

where, for brevity, we write $B_{i,j} = C(p_i^{e_{ij}})$. Hence

$$\begin{aligned} \mathrm{ch}_T &= \prod_{i=1}^{t} \prod_{j=1}^{\gamma_i} p_i^{e_{ij}} \\ &= \prod_{i=1}^{t} p_i^{\sum_{j=1}^{\gamma_i} e_{ij}} \\ &= \prod_{i=1}^{t} p_i^{\frac{\nu(p_i^{b_i}(T))}{\deg p_i}} \end{aligned}$$

as required.

THEOREM 5.5
$\mathrm{ch}_T = m_T \Leftrightarrow \exists$ *a basis* β *for* V *such that*

$$[T]_\beta^\beta = C(p_1^{b_1}) \oplus \ldots \oplus C(p_t^{b_t})$$

where $p_1, \ldots, p_t$ *are distinct monic irreducibles and* $b_1 \geq \ldots b_t \geq 1$.

Note that if $\mathrm{ch}_A = m_A$ (i.e. $T = T_A$ in the above), we say that the matrix A is **non-derogatory**.
PROOF

$\Leftarrow$

$$\begin{aligned} \mathrm{ch}_T &= \prod_{i=1}^{t} \mathrm{ch}_{C(p_i^{b_i})} = \prod_{i=1}^{t} p_i^{b_i}, \\ m_T &= \mathrm{lcm}\,(p_1^{b_1}, \ldots, p_t^{b_t}) = p_1^{b_1} \ldots p_t^{b_t} = \mathrm{ch}_T\,. \end{aligned}$$

$\Rightarrow$ Suppose that $\mathrm{ch}_T = m_T$.

We deduce that the dot diagram for each p_i consists of a single column of b_i dots, where $p_i^{b_i} || m_T$; that is,

$$\dim_{F_p} N_{h,p_i} = 1 \qquad \text{for } h = 1, 2, \ldots, b_i.$$

Observe that

$$\frac{\nu(p_i^h(T))}{\deg p_i} \in \mathbb{N},$$

for it may be written

$$\begin{aligned} & \sum_{j=1}^{h} \frac{\nu(p_i^j(T)) - \nu(p_i^{j-1}(T))}{\deg p_i} \\ = & \sum_{j=1}^{h} \dim_{F_p} N_{j,p_i} \in \mathbb{N}. \end{aligned}$$

Then, for each $i = 1, 2, \ldots, t$ we have the following sequence of positive integers:

$$1 \leq \frac{\nu(p_i(T))}{\deg p_i} < \frac{\nu(p_i^2(T))}{\deg p_i} < \ldots < \frac{\nu(p_i^{b_i}(T))}{\deg p_i} = a_i.$$

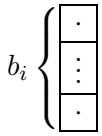
But $a_i = b_i$ here, as we are assuming that $\mathrm{ch}_T = m_T$. In particular, it follows that

$$\frac{\nu(p_i^h(T))}{\deg p_i} = h \qquad \text{for } h = 1, 2, \ldots, b_i$$

and $h = 1$ gives

$$\frac{\nu(p_i(T))}{\deg p_i} = 1 = \gamma_i.$$

So the bottom row of the i-th dot diagram has only one element; it looks like this:

b_i 

and we get the secondary decomposition

$$\operatorname{Ker} p_i^{b_i}(T) = C_{T,v_{i1}}.$$

Further, if $\beta = \beta_{11} \cup \cdots \cup \beta_{t1}$, where β_{i1} is the T–cyclic basis for $C_{T,v_{i1}}$, then

$$\begin{aligned} [T]_\beta^\beta &= \bigoplus_{i=1}^{t} \bigoplus_{j=1}^{\gamma_i} C(p_i^{e_{ij}}) \\ &= \bigoplus_{i=1}^{t} C(p_i^{b_i}) \\ &= C(p_1^{b_1}) \oplus \ldots \oplus C(p_t^{b_t}) \end{aligned}$$

as required.

THEOREM 5.6

$m_T = p_1 p_2 \ldots p_t$, a product of distinct monic irreducibles, if and only if $\exists$ a basis β for V such that

$$\begin{aligned} [T]_\beta^\beta = \quad & \underbrace{C(p_1) \oplus \ldots \oplus C(p_1)}_{\gamma_1 \text{ times}} \oplus \ldots\ldots \\ \oplus \quad & \underbrace{C(p_t) \oplus \ldots \oplus C(p_t)}_{\gamma_t \text{ times}}. \end{aligned} \tag{27}$$

Note: This is a generalization of an earlier result, namely that a transformation is diagonable if and only if its minimum polynomial splits into a product of distinct linear factors.

PROOF

$\Leftarrow$ Assume $\exists \beta$ such that (27) holds. Then

$$\begin{aligned} m_T &= \operatorname{lcm}(\underbrace{p_1, \ldots, p_1}_{\gamma_1}, \ldots, \underbrace{p_t, \ldots, p_t}_{\gamma_t}) \\ &= \operatorname{lcm}(p_1, \ldots, p_t) \\ &= p_1 p_2 \ldots p_t. \end{aligned}$$

$\Rightarrow$ Assume $m_T = p_1 \ldots p_t$. Then $b_i = 1$ for $i = 1, \ldots, t$ (i.e. the i-th dot diagram has height 1) and $\exists \beta$ such that

$$[T]_\beta^\beta = \bigoplus_{i=1}^{t} \bigoplus_{j=1}^{\gamma_i} C(p_i),$$

as $e_{ij} = 1 \ \forall i, j$.

5.3 Elementary divisors and invariant factors

5.3.1 Elementary Divisors

DEFINITION 5.1

The polynomials $p_i^{e_{ij}}$ occurring in the rational canonical form of T are called the **elementary divisors** *of T. Similarly the elementary divisors of a matrix $A \in M_{n\times n}(F)$ are the polynomials $p_i^{e_{ij}}$ occurring in the rational canonical form of A.*

THEOREM 5.7

Linear transformations $T_1, T_2 : V \to V$ have the same elementary divisors if and only if there exists an isomorphism $L : V \to V$ such that

$$T_2 = L^{-1}T_1L.$$

PROOF

"only if". Suppose that T_1 and T_2 have the same elementary divisors. Then $\exists$ bases β, γ for V such that

$$[T_1]_\beta^\beta = [T_2]_\gamma^\gamma = A.$$

Then we have the equations

$$\begin{aligned} \phi_\beta T_1 &= T_A\phi_\beta \\ \phi_\gamma T_2 &= T_A\phi_\gamma. \end{aligned}$$

Hence

$$\phi_\beta T_1\phi_\beta^{-1} = T_A = \phi_\gamma T_2\phi_\gamma^{-1},$$

so

$$\phi_\gamma^{-1}\phi_\beta T_1\phi_\beta^{-1}\phi_\gamma = T_2,$$

or

$$L^{-1}T_1L = T_2,$$

where $L = \phi_\beta^{-1}\phi_\gamma$ is an isomorphism.

"if". Suppose that $L^{-1}T_1L = T_2$. Then

$$m_{T_1} = m_{T_2} = p_1^{b_1}\cdots p_t^{b_t}, \quad \text{say;}$$

also for all i and h, because

$$p_i^h(T_2) = p_i^h(L^{-1}T_1L) = L^{-1}p_i^h(T_1)L,$$

we have

$$\nu(p_i^h(T_2)) = \nu(p_i^h(T_1)).$$

Hence for each p_i, the corresponding dot diagrams for T_1 and T_2 are identical and consequently the elementary divisors for T_1 and T_2 are identical.

COROLLARY 5.2

Let A, $B \in M_{n\times n}(F)$. Then A is similar to B if and only if A and B have the same elementary divisors.

PROOF

$$\begin{aligned} A \text{ is similar to } B &\Leftrightarrow \exists\ P \text{ non–singular, with } P^{-1}AP = B \\ &\Leftrightarrow \exists\ P \text{ non–singular, with } T_P^{-1}T_AT_P = T_B \\ &\Leftrightarrow \exists\ L \text{ an isomorphism, with } L^{-1}T_AL = T_B. \end{aligned}$$

5.3.2 Invariant Factors

THEOREM 5.8

Let $T : V \to V$ be a linear transformation over F. Then there exist non–constant monic polynomials $d_1, \ldots, d_s \in F[x]$, such that

(i) d_k divides d_{k+1} for $1 \le k \le s-1$;

(ii) vectors $v_1, \ldots, v_s \in V$ exist such that

$$V = \bigoplus_{k=1}^{s} C_{T,\,v_k},$$

where $m_{T,\,v_k} = d_k$.

REMARK: If β is the basis for V obtained by stringing together the T–cyclic bases for each $C_{T,\,v_k}$, we obtain the matrix direct sum

$$[T]_\beta^\beta = \bigoplus_{k=1}^{s} C(d_k).$$

This matrix is also said to be in rational canonical form.
PROOF

Let $s = \max(\gamma_1, \ldots, \gamma_t)$ and if $1 \leq i \leq t$ and $\gamma_i < j \leq s$, define $e_{ij} = 0$ and $v_{ij} = 0$, the zero vector of V. Now arrange the polynomials $p_i^{e_{ij}}$, $1 \leq i \leq t$; $1 \leq j \leq s$ as a $t \times s$ rectangular array:

$$\begin{array}{|c|c|c|} \hline p_1^{e_{1s}} & \cdots & p_1^{e_{11}} \\ \hline \vdots & \vdots & \vdots \\ \hline p_t^{e_{ts}} & \cdots & p_t^{e_{t1}} \\ \hline \end{array}$$

Let

$$d_1 = p_1^{e_{1s}} \cdots p_t^{e_{ts}}, \ldots, d_s = p_1^{e_{11}} \cdots p_t^{e_{t1}}$$

be the products along columns of the array, from left to right. Then $d_1, \ldots, d_s$ are monic non–constant polynomials and

$$d_1 | d_2 | \cdots | d_s.$$

Also $C_{T, v_{ij}} = \{0\}$ if $v_{ij} = 0$, so V is the direct sum of the following ts T–cyclic subspaces:

$$\begin{array}{|c|c|c|} \hline C_{T, v_{1s}} & \cdots & C_{T, v_{11}} \\ \hline \vdots & \vdots & \vdots \\ \hline C_{T, v_{ts}} & \cdots & C_{T, v_{t1}} \\ \hline \end{array}$$

Then by Problem Sheet 5, Question 15(b), if we let

$$v_1 = v_{1s} + \cdots v_{ts}, \ldots, v_s = v_{11} + \cdots v_{t1},$$

we have $m_{T, v_1} = d_1, \ldots, m_{T, v_s} = d_s$ and

$$\begin{aligned} C_{T, v_1} &= C_{T, v_{1s}} \oplus \cdots \oplus C_{T, v_{ts}} \\ \vdots \quad & \quad \vdots \\ C_{T, v_s} &= C_{T, v_{11}} \oplus \cdots \oplus C_{T, v_{t1}}. \end{aligned}$$

Consequently

$$V = C_{T, v_1} \oplus \cdots \oplus C_{T, v_s}.$$

DEFINITION 5.2

Polynomials $d_1, \ldots, d_s$ satisfying the conditions of the above theorem are called **invariant factors** *of T.*

There is a similar definition for matrices: if $A \in M_{n \times n}(F)$ is similar to a direct sum

$$\bigoplus_{k=1}^{s} C(d_k),$$

where $d_1, \ldots, d_s$ are non–constant monic polynomials in $F[x]$ such that d_k divides d_{k+1} for $1 \leq k \leq s-1$, then $d_1, \ldots, d_s$ are called **invariant factors** *of A. So the invariant factors of A are the invariant factors of T_A.*

THEOREM 5.9

The invariant factors of a linear transformation $T : V \to V$ are uniquely defined by T.

PROOF

Reverse the construction in the proof of the above theorem using Question 15(a) of Problem Sheet 5, thereby recapturing the rectangular array of elementary divisors, which in turn is uniquely determined by T.

EXAMPLE 5.1

Suppose $T : V \to V$ has elementary divisors

$$p_1^2,\ p_1^3,\ p_1^3;\ p_2,\ p_2^2,\ p_2^2,\ p_2^4;\ p_3,\ p_3,\ p_3^4,\ p_3^5,\ p_3^5.$$

Form the rectangular array

1	1	p_1^2	p_1^3	p_1^3
1	p_2	p_2^2	p_2^2	p_2^4
p_3	p_3	p_3^4	p_3^5	p_3^5

Then the invariant factors of T are obtained by respectively multiplying along columns:

$$\begin{aligned} d_1 &= p_3 \\ d_2 &= p_2 p_3 \\ d_3 &= p_1^2 p_2^2 p_3^4 \\ d_4 &= p_1^3 p_2^2 p_3^5 \\ d_5 &= p_1^3 p_2^4 p_3^5. \end{aligned}$$

THEOREM 5.10

If $d_1, \ldots, d_s$ are the invariant factors of $T : V \to V$, then

(i) $m_T = d_s$;

(ii) $\mathrm{ch}_T = d_1 \cdots d_s$.

PROOF

Suppose $B = [T]_\beta^\beta = \bigoplus_{k=1}^s C(d_k)$ is the canonical form corresponding to the invariant factors $d_1, \ldots, d_s$ of T. Then

$$\begin{aligned} m_T = m_B &= \operatorname{lcm}(m_{C(d_1)}, \ldots, m_{C(d_s)}) \\ &= \operatorname{lcm}(d_1, \ldots, d_s) = d_s. \end{aligned}$$

Also

$$\operatorname{ch}_T = \operatorname{ch}_B = \prod_{k=1}^s \operatorname{ch}_{C(d_k)} = \prod_{k=1}^s d_k.$$

We shall soon see that the invariant factors of a linear transformation or matrix are of independent interest. For example the invariant factors allow us to calculate the dimension of the vector space $Z_{L,M}$ consisting of all linear transformations $N : U \to V$ which satisfy the equation $MN = NL$, where $L : U \to U$ and $M : V \to V$ are given linear transformations over F.

It turns out that there is a more direct way of finding the invariant factors of T. To introduce this algorithm, we need to discuss an interesting equivalence relation on $M_{m\times n}(F[x])$, which in turn leads to the so-called **Smith canonical form** of a matrix over $F[x]$.