# 7 Various Applications of Rational Canonical Forms

## 7.1 An Application to commuting transformations

**THEOREM 7.1 (Cecioni 1908, Frobenius 1910)**

*Let $L : U \mapsto U$ and $M : V \mapsto V$ be given LTs. Then the vector space $Z_{L,M}$ of all LTs $N : U \mapsto V$ satisfying*

$$MN = NL$$

*has dimension*

$$\sum_{k=1}^{s} \sum_{l=1}^{t} \deg \gcd(d_k, D_l),$$

*where $d_1, \ldots, d_s$ and $D_1, \ldots, D_t$ are the invariant factors of $L$ and $M$ respectively.*

**COROLLARY 7.1**

*Now take $U = V$ and $L = M$. Then $Z_{L,L}$ the vector space of LTs satisfying*

$$NL = LN,$$

*has dimension*

$$\sum_{k=1}^{s} (2s - 2k + 1) \deg d_k.$$

proof Omitted, but here's a hint:

$$\gcd(d_k, d_l) = \begin{cases} d_k \text{ if } k \leq l \text{ ; i.e. if } d_k \mid d_l \\ d_l \text{ if } k > l \text{ ; i.e. if } d_l \mid d_k. \end{cases}$$

**N.B.** Let $P_L$ be the vector space of all LTs of the form

$$f(L) : U \mapsto U \qquad f \in F[x].$$

Then $P_L \subseteq Z_{L,L}$ and we have the following...

**THEOREM 7.2**

$$P_L = Z_{L,L} \Leftrightarrow m_L = \mathrm{ch}_L\,.$$

proof First note that $\dim P_L = \deg m_L$ as

$$I_V, L, \ldots, L^{\deg m_L - 1}$$

form a basis for $P_L$. So, since $P_L \subseteq Z_{L,L}$ we have

$$\begin{aligned} P_L = Z_{L,L} &\Leftrightarrow \dim P_L = \dim Z_{L,L} \\ &\Leftrightarrow \deg m_L = \sum_{k=1}^{s} (2s - 2k + 1) \deg d_k \\ &\Leftrightarrow s = 1 \\ &\Leftrightarrow \mathrm{ch}_L = m_L. \end{aligned}$$

proof (a sketch) of Cecioni-Frobenius theorem.
We start with the invariant factor decompositions

$$U = \bigoplus_{k=1}^{s} C_{L,u_k} \qquad \text{and} \qquad V = \bigoplus_{l=1}^{t} C_{M,v_l}$$

where $m_{L,u_k} = d_k$ for $k = 1, \ldots, s$, and $m_{M,v_l} = D_l$ for $l = 1, \ldots, t$.
Let $MN = NL \ldots$

$$\begin{aligned} &\Rightarrow \quad M^n N = N L^n \qquad \forall n \geq 1 \\ &\Rightarrow \quad f(M)N = N f(L) \qquad \forall f \in F[x]. \end{aligned}$$

Define vectors $w_1, \ldots, w_s \in V$ by $w_k = N(u_k)$, and observe

$$\begin{aligned} d_k(M)(w_k) &= d_k(M)(N(u_k)) \\ &= N(d_k(L)(u_k)) \\ &= N(0) = 0. \end{aligned}$$

Then we have the
**Definition:** Let $W$ be the set of all $(w_1, \ldots, w_s)$ such that $w_1, \ldots, w_s \in V$ and

$$d_k(M)(w_k) = 0 \qquad \forall k = 1, \ldots, s.$$

We assert that $W$ is a vector space and the mapping

$$N \mapsto (w_1, \ldots, w_s)$$

is an isomorphism between $Z_{L,M}$ and $W$; proof is left as an exercise.

Now let

$$w_k = \sum_{l=1}^{t} c_{kl}(M)(v_l) \qquad k = 1, \ldots, s \text{ and } c_{kl} \in F[x].$$

**N.B.**

$$\begin{aligned} & f(M)(v_l) = g(M)(v_l) \text{ say} \\ \Leftrightarrow \quad & D_l \mid f - g. \end{aligned}$$

So, if we restrict $c_{kl}$ by the condition

$$\deg c_{kl} < \deg D_l \quad \text{if } c_{kl} \neq 0 \tag{29}$$

then the $c_{kl}$ are uniquely defined for each $k$.
Exercise: Now let

$$g_{kl} = \gcd(d_k, D_l).$$

Then from the condition $d_k(M)(w_k) = 0$, show that

$$\frac{D_l}{g_{kl}} \mid c_{kl} \tag{30}$$

i.e. that

$$c_{kl} = b_{kl}\frac{D_l}{g_{kl}} \qquad b_{kl} \in F[x]. \tag{31}$$

Then the matrices $[c_{kl}]$, where $c_{kl}$ satisfy (30), form a vector space (call it $X$) which is isomorphic to $W$.

Then in (31),

$$(29) \iff \deg b_{kl} < \deg g_{kl} \qquad \text{if } b_{kl} \neq 0.$$

Clearly then,

$$\dim X = \dim Z_{L,M} = \sum_{k=1}^{s}\sum_{l=1}^{t} \deg g_{kl}$$

as required.

**EXAMPLE 7.1**

*(of the vector space $X$, when $s = t = 2$)*

*Say*

$$[\deg g_{kl}] = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}.$$

*Then $X$ consists of all matrices of the form*

$$\begin{aligned}[c_{kl}] &= \begin{bmatrix} (a_0+a_1x)\cdot\frac{D_1}{g_{11}} & 0\cdot\frac{D_2}{g_{12}} \\ b_0\cdot\frac{D_1}{g_{21}} & (c_0+c_1x+c_2x^2)\cdot\frac{D_2}{g_{22}} \end{bmatrix} \\ &= a_0\begin{bmatrix} \frac{D_1}{g_{11}} & 0 \\ 0 & 0 \end{bmatrix} + a_1\begin{bmatrix} \frac{xD_1}{g_{11}} & 0 \\ 0 & 0 \end{bmatrix} + b_0\begin{bmatrix} 0 & 0 \\ \frac{D_1}{g_{21}} & 0 \end{bmatrix} + \cdots \end{aligned}$$

*... and so on.*

**EXAMPLE 7.2**

**The most general $3\times 3$ matrix which commutes with others.**

*Let $A \in M_{3\times 3}(\mathbb{Q})$ such that there exists non-singular $P \in M_{3\times 3}(\mathbb{Q})$ with*

$$\begin{aligned} P^{-1}AP &= C(x-1)\oplus C((x-1)^2) \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix} = J, \text{ say,} \end{aligned}$$

*where $C(p)$ denotes the companion matrix of $p$, as usual.*

*Then $P = [u_1 \mid u_2 \mid T(u_2)]$ where $T = T_A$ and*

$$m_{T,u_1} = x-1, \qquad m_{T,u_2} = (x-1)^2.$$

*Also $V_3(\mathbb{Q}) = C_{T,u_1}\oplus C_{T,u_2}$.*

*Note that the invariant factors of $T$ are $(x-a)$ and $(x-1)^2$.*

*We find all $3\times 3$ matrices $B$ such that*

$$\begin{aligned} & BA = AB, \\ \text{i.e.} \quad & T_BT_A = T_AT_B. \end{aligned}$$

*Let $N = T_B$. Then $N$ must satisfy*

$$\begin{aligned} N(u_1) &= Bu_1 = c_{11}u_1 + c_{12}u_2 \qquad \text{and} \\ N(u_2) &= Bu_2 = c_{21}u_1 + c_{22}u_2 \qquad \text{where } c_{kl} \in \mathbb{Q}[x]. \end{aligned} \tag{32}$$

*Now*

$$[\deg \gcd(d_k, d_l)] = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

*so*

$$[c_{kl}] = \begin{bmatrix} a_0 & b_0(x-1) \\ c_0 & d_0+d_1x \end{bmatrix}$$

*where $a_0$ etc. $\in \mathbb{Q}$, so (32) gives*

$$\begin{aligned} Bu_1 &= a_0u_1 + b_0(x-1)u_2 \\ &= a_0u_1 - b_0u_2 + b_0T(u_2) \qquad (33) \\ Bu_2 &= c_0u_1 + (d_0 + d_1x)u_2 \\ &= c_0u_1 + d_0u_2 + d_1T(u_2). \qquad (34) \end{aligned}$$

*Noting that*

$$\begin{aligned} & m_{T,u_1} = x - 1 \Rightarrow T(u_1) = u_1 \\ \text{and} \quad & m_{T,u_2} = (x-1)^2 = x^2 - 2x + 1 \Rightarrow T^2(u_2) = 2T(u_2) - u_2, \end{aligned}$$

*we have from (34) that*

$$\begin{aligned} T(Bu_2) &= c_0T(u_1) + d_0T(u_2) + d_1T^2(u_2) \\ &= c_0u_1 - d_1u_2 + (d_0 + 2d_1)T(u_2). \end{aligned}$$

*In terms of matrices,*

$$\begin{aligned} B[u_1|u_2|T(u_2)] &= [u_1|u_2|T(u_2)] \begin{bmatrix} a_0 & c_0 & c_0 \\ -b_0 & d_0 & -d_1 \\ b_0 & d_1 & d_0 + 2d_1 \end{bmatrix} \\ \text{i.e.} \quad & BP = PK, \quad \text{say} \\ \text{or} \quad & B = PKP^{-1}. \end{aligned}$$

*This gives the most general matrix $B$ such that*

$$BA = AB.$$

**Note***: $BA = AB$ becomes*

$$\begin{aligned} PKP^{-1}\,PJP^{-1} &= PJP^{-1}\,PKP^{-1} \\ \Leftrightarrow \quad KJ &= JK. \end{aligned}$$

## 7.2 Tensor products and the Byrnes-Gauger theorem

We next apply the Cecioni-Frobenius theorem to derive a third criterion for deciding whether or not two matrices are similar.

**DEFINITION 7.1**

*(Tensor or Kronecker product)*

*If $A \in M_{m_1 \times n_1}(F)$ and $B \in M_{m_2 \times n_2}(F)$ we define*

$$A \otimes B = \left[\begin{array}{c|c|c} a_{11}B & a_{12}B & \cdots \\ \hline a_{21}B & a_{22}B & \cdots \\ \hline \vdots & \vdots & \ddots \end{array}\right] \in M_{m_1 m_2 \times n_1 n_2}(F).$$

*In terms of elements,*

$$(A \otimes B)_{(i,j),(k,l)} = a_{ij} b_{kl}$$

*—the element at the intersection of the $i$-th row block, $k$-th row sub-block, and the $j$-th column block, $l$-th column sub-block.*[4]

**EXAMPLE 7.3**

$$A \otimes I_p = \left[\begin{array}{ccc|c} a_{11} & & & \\ & \ddots & & \cdots \\ & & a_{11} & \\ \hline a_{21} & & & \\ & \ddots & & \cdots \\ & & a_{21} & \\ \hline & & & \\ & \vdots & & \ddots \end{array}\right],$$

$$I_p \otimes A = \left[\begin{array}{c|c|c} A & 0 & \cdots \\ \hline 0 & A & \cdots \\ \hline \vdots & \vdots & \ddots \end{array}\right].$$

*(Tensor-product-taking is obviously far from commutative!)*

### 7.2.1 Properties of the tensor product of matrices

(i) $(tA) \otimes B = A \otimes (tB) = t(A \otimes B),\ t \in F$;

(ii) $A \otimes B = 0 \Leftrightarrow A = 0$ or $B = 0$;

(iii) $A \otimes (B \otimes C) = (A \otimes B) \otimes C$;

(iv) $A \otimes (B + C) = (A \otimes B) + (A \otimes C)$;

(v) $(B + C) \otimes D = (B \otimes D) + (C \otimes D)$;

[4] That is, the $((i-1)m_2 + k, (j-1)n_2 + l)$-th element in the tensor product is $a_{ij}b_{kl}$.

(vi) $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$;

(vii) $(B \oplus C) \otimes D = (B \otimes D) \oplus (C \otimes D)$;

(viii) $P(A \otimes (B \oplus C))P^{-1} = (A \otimes B) \oplus (A \otimes C)$ for a suitable row permutation matrix $P$;

(ix) $\det(A \otimes B) = (\det A)^n (\det B)^m$ if $A$ is $m \times m$ and $B$ is $n \times n$;

(x) Let $f(x, y) = \sum_{i=0}^{m} \sum_{j=0}^{n} c_{ij} x^i y^j \in F[x, y]$ be a polynomial in $x$ and $y$ over $F$ and define

$$f(A; B) = \sum_{i=0}^{m} \sum_{j=0}^{n} c_{ij} (A^i \otimes B^j).$$

Then if $\mathrm{ch}_A = \prod_{k=1}^{s} (x - \lambda_k)$ and $\mathrm{ch}_B = \prod_{l=1}^{t} (x - \mu_l)$, we have

$$\mathrm{ch}_{f(A;B)} = \prod_{k=1}^{s} \prod_{l=1}^{t} (x - f(\lambda_k, \mu_l));$$

(xi) Taking $f(x, y) = xy$ gives

$$\mathrm{ch}_{A \otimes B} = \prod_{k=1}^{s} \prod_{l=1}^{t} (x - \lambda_k \mu_l);$$

(xii) Taking $f(x, y) = x - y$ gives

$$\mathrm{ch}_{(A \otimes I_n - I_m \otimes B)} = \prod_{k=1}^{s} \prod_{l=1}^{t} (x - (\lambda_k - \mu_l));$$

**Remark:** (ix) can be proved using the uniqueness theorem for alternating $m$–linear functions met in MP174; (x) follows from the the equations

$$P^{-1} A P = J_1 \quad \text{and} \quad Q^{-1} B Q = J_2,$$

where $J_1$ and $J_2$ are the Jordan forms of $A$ and $B$, respectively. Then $J_1$ and $J_2$ are lower triangular matrices with the eigenvalues $\lambda_k$, $1 \leq k \leq m$ and $\mu_l$, $1 \leq l \leq n$ of $A$ and $B$ as diagonal elements.

Then

$$P^{-1} A^i P = J_1^i \quad \text{and} \quad Q^{-1} B^j Q = J_2^j$$

and more generally

$$(P \otimes Q)^{-1} \sum_{i=0}^{s} \sum_{j=0}^{t} c_{ij}(A^i \otimes B^j)(P \otimes Q) = \sum_{i=0}^{s} \sum_{j=0}^{t} c_{ij}(J_1^i \otimes J_2^j).$$

The matrix on the right–hand side is lower triangular and has diagonal elements

$$f(\lambda_k, \mu_l),\ 1 \leq k \leq m,\ 1 \leq l \leq n.$$

**THEOREM 7.3**

*Let $\beta$ be the standard basis for $M_{m\times n}(F)$—i.e. the basis consisting of the matrices*

$$E_{11}, \ldots\ldots, E_{mn}$$

*and $\gamma$ be the standard basis for $M_{p\times n}(F)$.*

*Let $A$ be $p \times m$, and*

$$T_1 : M_{m\times n}(F) \mapsto M_{p\times n}(F)$$

*be defined by $T_1(X) = AX$. Then*

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

*Similarly if $B$ is $n \times p$, and*

$$T_2 : M_{m\times n}(F) \mapsto M_{m\times p}(F)$$

*is defined by $T_2(Y) = YB$, then*

$$[T_2]_\beta^\delta = A \otimes I_n$$

*(where $\delta$ is the standard basis for $M_{m\times p}(F)$).*

proof Left for the intrepid reader. A hint:

$$E_{ij}E_{kl} = \begin{cases} 0 & \text{if } j \neq k, \\ E_{il} & \text{if } j = k \end{cases}$$

**COROLLARY 7.2**

*Let $A$ be $m \times m$,*
*$B$ be $n \times n$,*
*$X$ be $m \times n$, and*

$$T : M_{m\times n}(F) \mapsto M_{m\times n}(F)$$

*be defined by* $T(X) = AX - XB$.

*Then*

$$[T]_\beta^\beta = A \otimes I_n - I_m \otimes B^t,$$

*where* $\beta$ *is the standard basis for* $M_{m\times n}(F)$.

**DEFINITION 7.2**

*For brevity in the coming theorems, we define*

$$\nu_{A,B} = \nu(A \otimes I_n - I_m \otimes B^t)$$

*where* $A$ *is* $m \times m$ *and* $B$ *is* $n \times n$.

**THEOREM 7.4**

$$\begin{aligned} \nu_{A,B} &= \nu(A \otimes I_n - I_m \otimes B^t) \\ &= \sum_{k=1}^{s}\sum_{l=1}^{t} \deg \gcd(d_k, D_l) \end{aligned}$$

*where*

$$\begin{aligned} &d_1 \mid d_2 \mid \cdots \mid d_s \quad \text{and} \\ &D_1 \mid D_2 \mid \cdots \mid D_t \end{aligned}$$

*are the invariant factors of* $A$ *and* $B$ *respectively.*

proof With the transformation $T$ from corollary 7.2 above, we note that

$$\begin{aligned} \nu_{A,B} &= \text{nullity}\, T \\ &= \dim\{\, X \in M_{m\times n}(F) \mid AX = XB \,\} \\ &= \dim\{\, N \in \text{Hom}\,(V_n(F), V_m(F)) \mid T_A N = N T_B \,\} \end{aligned}$$

and the Cecioni-Frobenius theorem gives the result.

**LEMMA 7.1 (Byrnes-Gauger)**

*(This is needed in the proof of the Byrnes-Gauger theorem following.)*

*Suppose we have two monotonic increasing integer sequences:*

$$\begin{cases} m_1 \le m_2 \le \cdots \le m_s \\ n_1 \le n_2 \le \cdots \le n_s \end{cases} \quad \text{and}$$

*Then*

$$\sum_{k=1}^{s}\sum_{l=1}^{s}\{\min(m_k,m_l)+\min(n_k,n_l)-2\min(m_k,n_l)\}\geq 0.$$

*Further, equality occurs iff the sequences are identical.*

proof

**Case 1**: $k=l$.

The terms to consider here are of the form

$$m_k+n_k-2\min(m_k,n_k)$$

which is obviously $\geq 0$. Also, the term is equal to zero iff $m_k=n+k$.

**Case 2**: $k\neq l$; without loss of generality take $k<l$.

Here we pair the off-diagonal terms $(k,l)$ and $l,k$.

$$\begin{aligned}
&\{\min(m_k,m_l)+\min(n_k,n_l)-2\min(m_k,n_l)\}\\
&\qquad+\{\min(m_l,m_k)+\min(n_l,n_k)-2\min(m_l,n_k)\}\\
=\;&\{m_k+n_l-2\min(m_k,n_l)\}+\{m_l+n_k-2\min(m_l,n_k)\}\\
\geq\;&0,\qquad\text{obviously.}
\end{aligned}$$

Since the sum of the diagonal terms and the sum of the pairs of sums of off-diagonal terms are non-negative, the sum is non-negative. Also, if the sum is zero, so must be the sum along the diagonal terms, making

$$m_k=n_k\qquad\forall k.$$

**THEOREM 7.5 (Byrnes-Gauger)**

*If $A$ is $m\times m$ and $B$ is $n\times n$ then*

$$\nu_{A,A}+\nu_{B,B}\geq 2\nu_{A,B}$$

*with equality if and only if $m=n$ and $A$ and $B$ are similar.*

proof

$$\begin{aligned}
&\nu_{A,A}+\nu_{B,B}-2\nu_{A,B}\\
=\;&\sum_{k_1=1}^{s}\sum_{k_2=1}^{s}\deg\gcd(d_{k_1},d_{k_2})+\sum_{l_1=1}^{t}\sum_{l_2=1}^{t}\deg\gcd(D_{l_1},D_{l_2})\\
&\qquad-2\sum_{k=1}^{s}\sum_{l=1}^{t}\deg\gcd(d_k,D_l).
\end{aligned}$$

We now extend the definitions of $d_1, \ldots, d_s$ and $D_1, \ldots, D_t$ by renaming them as follows, with $N = \max(s,t)$:

$$\begin{aligned} \underbrace{1, \ldots, 1}_{N-s}, d_1, \ldots, d_s &\mapsto f_1, \ldots, f_N \\ \text{and} \quad \underbrace{1, \ldots, 1}_{N-t}, D_1, \ldots, D_t &\mapsto F_1, \ldots, F_N. \end{aligned}$$

This is so we may rewrite the above sum of three sums as a single sum, viz:

$$\begin{aligned} \nu_{A,A} + \nu_{B,B} - 2\nu_{A,B} &= \sum_{k=1}^{N} \sum_{l=1}^{N} \{\deg \gcd(f_k, f_l) + \deg \gcd(F_k, F_l) \\ &\qquad -2 \deg \gcd(f_k, F_l)\}. \end{aligned} \tag{35}$$

We now let $p_1, \ldots, p_r$ be the distinct monic irreducibles in $m_A m_B$ and write

$$\left.\begin{array}{rcccc} f_k = & p_1^{a_{k1}} & p_2^{a_{k2}} & \ldots & p_r^{a_{kr}} \\ F_k = & p_1^{b_{k1}} & p_2^{b_{k2}} & \ldots & p_r^{b_{kr}} \end{array}\right\} \qquad 1 \le k \le N$$

where the sequences $\{a_{ki}\}_{i=1}^r$, $\{b_{ki}\}_{i=1}^r$ are monotonic increasing non-negative integers. Then

$$\begin{aligned} & \gcd(f_k, F_l) = \prod_{i=1}^{r} p_i^{\min(a_{ki}, b_{li})} \\ \Rightarrow \qquad & \deg \gcd(f_k, F_l) = \sum_{i=1}^{r} \deg p_i \ \min(a_{ki}, b_{li}) \\ \text{and} \quad & \deg \gcd(f_k, f_l) = \sum_{i=1}^{r} \deg p_i \ \min(a_{ki}, a_{li}) \\ \text{and} \quad & \deg \gcd(F_k, F_l) = \sum_{i=1}^{r} \deg p_i \ \min(b_{ki}, b_{li}). \end{aligned}$$

Then equation (35) may be rewritten as

$$\begin{aligned} & \nu_{A,A} + \nu_{B,B} - 2\nu_{A,B} \\ &= \sum_{k=1}^{N} \sum_{l=1}^{N} \sum_{i=1}^{r} \deg p_i \{\min(a_{ki}, a_{li}) + \min(b_{ki}, b_{li}) \\ &\qquad -2 \min(a_{ki}, b_{li})\} \\ &= \sum_{i=1}^{r} \deg p_i \sum_{k=1}^{N} \sum_{l=1}^{N} \{\min(a_{ki}, a_{li}) + \min(b_{ki}, b_{li}) \\ &\qquad -2 \min(a_{ki}, b_{li})\}. \end{aligned}$$

The latter double sum is of the form in lemma 7.1 and so, since $\deg p_i > 0$, we have

$$\nu_{A,A} + \nu_{B,B} - 2\nu_{A,B} \geq 0,$$

proving the first part of the theorem.

Next we show that equality to zero in the above is equivalent to similarity of the matrices:

$$\begin{aligned}
& \nu_{A,A} + \nu_{B,B} - 2\nu_{A,B} = 0 \\
\Leftrightarrow \quad & \sum_{i=1}^{r} \deg p_i \sum_{k=1}^{N} \sum_{l=1}^{N} \{\min(a_{ki}, a_{li}) + \min(b_{ki}, b_{li}) \\
& \qquad -2\min(a_{ki}, b_{li})\} = 0 \\
\Leftrightarrow \quad & \text{sequences } \{a_{ki}\}, \{b_{ki}\} \text{ identical (by lemma 7.1)} \\
\Leftrightarrow \quad & A \text{ and } B \text{ have same invariant factors} \\
\Leftrightarrow \quad & A \text{ and } B \text{ are similar } (\Rightarrow m = n).
\end{aligned}$$

**EXERCISE 7.1**

*Show if if*

$$P^{-1}A_1P = A_2 \qquad \text{and} \qquad Q^{-1}B_1Q = B_2$$

*then*

$$\begin{aligned}
& (P^{-1} \otimes Q^{-1})(A_1 \otimes I_m - I_n \otimes B_1^t)(P \otimes Q) \\
= \quad & A_2 \otimes I_n - I_m \otimes B_2^t.
\end{aligned}$$

*(This is another way of showing that if $A$ and $B$ are similar then*

$$\nu_{A,A} + \nu_{B,B} - 2\nu_{A,B} = 0.)$$