We now present some interesting applications of the Jordan canonical form.

4.4 Non–derogatory matrices and transformations

If $\mathrm{ch}_A = m_A$, we say that the matrix A is **non-derogatory**.

THEOREM 4.5

Suppose that ch_T splits completely in $F[x]$. Then $\mathrm{ch}_T = m_T \Leftrightarrow \exists$ *a basis β for V such that*

$$[T]_\beta^\beta = J_{b_1}(c_1) \oplus \ldots \oplus J_{b_t}(c_t),$$

where $c_1, \ldots, c_t$ are distinct elements of F.

PROOF.

$\Leftarrow$

$$\begin{aligned}\mathrm{ch}_T &= \prod_{i=1}^{t} \mathrm{ch}_{J_{b_i}(c_i)} = \prod_{i=1}^{t} (x - c_i)^{b_i}, \\ m_T &= \mathrm{lcm}\left((x-c_1)^{b_1}, \ldots, (x-c_t)^{b_t}\right) = (x-c_1)^{b_1} \ldots (x-c_t)^{b_t} = \mathrm{ch}_T\,.\end{aligned}$$

$\Rightarrow$ Suppose that $\mathrm{ch}_T = m_T = (x-c_1)^{a_1} \cdots (x-c_t)^{a_t}$.

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

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

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

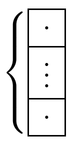
$$1 \le \nu(p_i(T)) < \nu(p_i^2(T)) < \cdots < \nu(p_i^{b_i}(T)) = a_i.$$

But $a_i = b_i$ here, as we are assuming that $\mathrm{ch}_T = m_T$. In particular, it follows that $\nu(p_i^h(T)) = h$ for $h = 1, 2, \ldots, b_i$ and $h = 1$ gives

$$\nu(p_i(T)) = 1 = \gamma_i.$$

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 elementary Jordan basis for $C_{T,v_{i1}}$, then

$$\begin{aligned} [T]_\beta^\beta &= \bigoplus_{i=1}^{t} \bigoplus_{j=1}^{\gamma_i} J_{e_{ij}}(c_i)) \\ &= \bigoplus_{i=1}^{t} J_{b_i}(c_i), \end{aligned}$$

as required.

4.5 Calculating A^m, where $A \in M_{n\times n}(\mathbb{C})$.

THEOREM 4.6

Let $c \in F$.

(a)

$$J_n^m(c) = \begin{bmatrix} c^m & 0 & \cdots & \cdots & & 0 \\ \binom{m}{1}c^{m-1} & c^m & \cdots & \cdots & & 0 \\ \binom{m}{2}c^{m-2} & \binom{m}{1}c^{m-1} & \cdots & \cdots & & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \binom{m}{m} & & \cdots & \cdots & & 0 \\ 0 & \binom{m}{m} & \cdots & \cdots & & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \binom{m}{m} & \cdots & \binom{m}{1}c^{m-1} & c^m \end{bmatrix}$$

if $1 \le m \le n-1$;

(b)

$$J_n^m(c) = \begin{bmatrix} c^m & 0 & \cdots & 0 & 0 \\ \binom{m}{1}c^{m-1} & c^m & \cdots & 0 & 0 \\ \binom{m}{2}c^{m-2} & \binom{m}{1}c^{m-1} & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ \binom{m}{n-1}c^{m-n+1} & \binom{m}{n-2}c^{m-n+2} & \cdots & \binom{m}{1}c^{m-1} & c^m \end{bmatrix}$$

if $n-1 \le m$*, where* $\binom{m}{k}$ *is the binomial coefficient*

$$\binom{m}{k} = \frac{m!}{k!(m-k)!} = \frac{m(m-1)\cdots(m-n+1)}{k!}.$$

PROOF. $J_n(c) = cI_n + N$, where N has the special property that N^k has 1 on the k–th sub–diagonal and 0 elsewhere, for $0 \le k \le n-1$.

Then because cI_n and N commute, we can use the binomial theorem:

$$\begin{aligned} J_n^m(c) &= (cI_n + N)^m \\ &= \sum_{k=0}^{m} \binom{m}{k} (cI_n)^{m-k} N^k \\ &= \sum_{k=0}^{m} \binom{m}{k} c^{m-k} N^k. \end{aligned}$$

(a). Let $1 \le m \le n-1$. Then in the above summation, the variable k must satisfy $0 \le k \le n-1$. Hence $J_n^m(c)$ is an $n \times n$ matrix having $\binom{m}{k} c^{m-k}$ on the k–th sub–diagonal, $0 \le k \le m$ and 0 elsewhere.

(b). Let $n-1 \le m$. Then

$$J_n^m(c) = \sum_{k=0}^{m} \binom{m}{k} c^{m-k} N^k = \sum_{k=0}^{n-1} \binom{m}{k} c^{m-k} N^k,$$

as $N^k = 0$ if $n \le k$. Hence $J_n^m(c)$ is an $n \times n$ matrix having $\binom{m}{k} c^{m-k}$ on the k–th sub–diagonal, $0 \le k \le n-1$ and 0 elsewhere.

COROLLARY 4.1

Let $F = \mathbb{C}$*. Then*

$$\lim_{m \to \infty} J_n^m(c) = 0 \quad \textit{if} \quad |c| < 1.$$

PROOF. Suppose that $|c| < 1$. Let $n-1 \le m$. Then

$$J_n^m(c) = \sum_{k=0}^{n-1} \binom{m}{k} c^{m-k} N^k.$$

But for fixed k, $0 \le k \le n-1$, $c^{m-k} \to 0$ as $m \to \infty$. For

$$\binom{m}{k} = \frac{m(m-1)\cdots(m-k+1)}{k!}$$

is a polynomial in m of degree k and

$$|m^j c^{m-k}| = |m^j e^{(m-k)\log c}| = m^j e^{(m-k)\log|c|} \to 0 \quad \text{as} \quad m \to \infty,$$

as $\log c = \log|c| + i \arg c$ and $\log|c| < 0$.

The last corollary gives a more general result:

COROLLARY 4.2

Let $A \in M_{n\times n}(\mathbb{C})$ and suppose that all the eigenvalues of A are less than 1 in absolute value. Then

$$\lim_{m\to\infty} A^m = 0.$$

PROOF. Suppose $ch_A = (x - c_1)^{a_1} \cdots (x - c_t)^{a_t}$, where $c_1, \ldots, c_t$ are the distinct eigenvalues of A and $|c_1| < 1, \ldots, |c_t| < 1$.

Then if J is the Jordan canonical form of A, there exists a non–singular matrix $P \in M_{n\times n}(\mathbb{C})$, such that

$$P^{-1}AP = J = \bigoplus_{i=1}^{t} \bigoplus_{j=1}^{\gamma_i} J_{e_{ij}}(c_i).$$

Hence

$$P^{-1}A^mP = (P^{-1}AP)^m = J^m = \bigoplus_{i=1}^{t} \bigoplus_{j=1}^{\gamma_i} J^m_{e_{ij}}(c_i).$$

Hence $P^{-1}A^mP \to 0$ as $m \to \infty$, because $J^m_{e_{ij}}(c_i) \to 0$.

4.6 Calculating e^A, where $A \in M_{n\times n}(\mathbb{C})$.

We first show that the matrix limit

$$\lim_{M\to\infty} \left(I_n + A + \frac{1}{2!}A^2 + \cdots + \frac{1}{M!}A^M \right)$$

exists. We denote this limit by e^A and write

$$e^A = I_n + A + \frac{1}{2!}A^2 + \cdots + \frac{1}{m!}A^m + \cdots = \sum_{m=0}^{\infty} \frac{1}{m!}A^m.$$

To justify this definition, we let $A^m = [a^{(m)}_{ij}]$. We have to show that

$$\left(I_n + A + \frac{1}{2!}A^2 + \cdots + \frac{1}{M!}A^M \right)_{ij} = a^{(0)}_{ij} + \frac{1}{1!}a^{(1)}_{ij} + \cdots + \frac{1}{M!}a^{(M)}_{ij}$$

tends to a limit as $M \to \infty$; in other words, we have to show that the series

$$\sum_{m=0}^{\infty} \frac{1}{m!} a_{ij}^{(m)}$$

converges. To do this, suppose that

$$|a_{ij}| \le \rho, \quad \forall\, i,\, j.$$

Then it is an easy induction to prove that

$$|a_{ij}^{(m)}| \le n^{m-1}\rho^m \quad \text{if} \quad m \ge 1.$$

Then the above series converges by comparison with the series

$$\sum_{m=0}^{\infty} \frac{1}{m!} n^{m-1}\rho^m.$$

4.7 Properties of the exponential of a complex matrix

THEOREM 4.7

(i) $e^0 = I_n$;

(ii) $e^{\mathrm{diag}\,(\lambda_1,\ldots,\lambda_n)} = \mathrm{diag}\,(e^{\lambda_1},\ldots,e^{\lambda_n})$;

(iii) $e^{P^{-1}AP} = P^{-1}e^A P$;

(iv) $e^{\bigoplus_{i=1}^t A_i} = \bigoplus_{i=1}^t e^{A_i}$;

(v) *if A is diagonable and has principal idempotent (spectral) decomposition:*

$$A = c_1E_1 + \cdots + c_tE_t,$$

then

$$e^A = e^{c_1}E_1 + \cdots + e^{c_t}E_t;$$

(vi)

$$\frac{d}{dt}\, e^{tA} = Ae^{tA},$$

if A is a constant matrix;

(vii) $e^A = p(A)$, *where* $p \in \mathbb{C}[x]$;

(viii) e^A *is non–singular and*

$$(e^A)^{-1} = e^{-A};$$

(ix) $e^A e^B = e^{A+B}$ *if* $AB = BA$*;*

(x)

$$e^{J_n(c)} = \begin{bmatrix} e^c & 0 & 0 & \cdots & & 0 \\ e^c/1! & e^c & 0 & \cdots & & 0 \\ e^c/2! & e^c/1! & e^c & \cdots & & 0 \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ e^c/(n-2)! & & \ddots & e^c/1! & e^c & 0 \\ e^c/(n-1)! & e^c/(n-2)! & \cdots & e^c/2! & e^c/1! & e^c \end{bmatrix}.$$

(xi)

$$e^{tJ_n(c)} = \begin{bmatrix} e^{tc} & 0 & 0 & \cdots & & 0 \\ te^{tc}/1! & e^{tc} & 0 & \cdots & & 0 \\ t^2e^{tc}/2! & e^{tc}/1! & e^{tc} & \cdots & & 0 \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ t^{n-2}e^{tc}/(n-2)! & & \ddots & te^{tc}/1! & e^{tc} & 0 \\ t^{n-1}e^{tc}/(n-1)! & t^{n-2}e^{tc}/(n-2)! & \cdots & t^2e^{tc}/2! & te^{tc}/1! & e^c \end{bmatrix}.$$

(xii) *If*

$$P^{-1}AP = J = \bigoplus_{i=1}^{t} \bigoplus_{j=1}^{\gamma_i} J_{e_{ij}}(c_i).$$

then

$$P^{-1}e^A P = J = \bigoplus_{i=1}^{t} \bigoplus_{j=1}^{\gamma_i} e^{J_{e_{ij}}(c_i)}.$$

PROOF.

(i)

$$e^0 = \sum_{m=0}^{\infty} \frac{1}{m!} 0^k = I_n;$$

(ii) Let $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Then

$$\begin{aligned}
A^m &= \operatorname{diag}(\lambda_1^m, \ldots, \lambda_n^m) \\
\sum_{m=0}^{\infty} \frac{1}{m!} A^m &= \operatorname{diag}\left(\sum_{m=0}^{\infty} \frac{\lambda_1^m}{m!}, \ldots, \sum_{m=0}^{\infty} \frac{\lambda_n^m}{m!}\right) \\
&= \operatorname{diag}(e^{\lambda_1}, \ldots, e^{\lambda_n}).
\end{aligned}$$

(iii)

$$\begin{aligned}
e^{P^{-1}AP} &= \sum_{m=0}^{\infty} \frac{1}{m!}(P^{-1}AP)^m \\
&= \sum_{m=0}^{\infty} \frac{1}{m!}(P^{-1}A^mP) \\
&= P^{-1}\left(\sum_{m=0}^{\infty} \frac{1}{m!} A^m\right) P \\
&= P^{-1}e^AP.
\end{aligned}$$

(iv) and (v) are left as exercises.

(vi) Using the earlier notation, $A^m = [a_{ij}^{(m)}]$, we have

$$\begin{aligned}
e^{tA} &= \sum_{m=0}^{\infty} \frac{1}{m!}(tA)^m \\
&= \sum_{m=0}^{\infty} \frac{t^m}{m!} A^m \\
&= \left[\sum_{m=0}^{\infty} \frac{t^m a_{ij}^{(m)}}{m!}\right] \\
\frac{d}{dt} e^{tA} &= \left[\frac{d}{dt} \sum_{m=0}^{\infty} \frac{t^m a_{ij}^{(m)}}{m!}\right] \\
&= \left[\sum_{m=1}^{\infty} \frac{t^{m-1} a_{ij}^{(m)}}{(m-1)!}\right] \\
&= \left[\sum_{m=0}^{\infty} \frac{t^m a_{ij}^{(m+1)}}{(m)!}\right]
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{m=0}^{\infty} \frac{t^m}{m!} A^{m+1} \\
&= Ae^{tA}.
\end{aligned}
$$

(vii) Let $\deg m_A = r$. Then the matrices $I_n,\ A, \ldots, A^{r-1}$ are linearly independent over $\mathbb{C}$, as if

$$m_A = x^r - a_{r-1}x^{r-1} - \cdots - a_0,$$

then

$$m_A(A) = 0 \Rightarrow A^r = a_0 I_n + a_1 A + \cdots + a_{r-1}A^{r-1}.$$

Consequently for each $m \geq 1$, we can express A^m as a linear combination over $\mathbb{C}$ of $I_n,\ A,\ \ldots,\ A^{r-1}$:

$$A^m = a_0^{(m)} I_n + a_1^{(m)} A + \cdots + a_{r-1}^{(m)} A^{r-1}$$

and hence

$$\sum_{m=0}^{M} \frac{1}{m!} A^m = \sum_{m=0}^{M} \frac{a_0^{(m)}}{m!} I_n + \sum_{m=0}^{M} \frac{a_1^{(m)}}{m!} A + \cdots + \sum_{m=0}^{M} \frac{a_{r-1}^{(m)}}{m!} A^{r-1},$$

or

$$[t_{ij}^{(M)}] = s_{0M} I_n + s_{1M} A + \cdots + s_{r-1\,M} A^{r-1},$$

say.

Now $[t_{ij}^{(M)}] \to e^A$ as $M \to \infty$.

Also the above matrix equation can be regarded as n^2 equations in

$$s_{0M},\ s_{1M}, \ldots, s_{r-1,\,M}.$$

Also the linear independence of $I_n,\ A, \ldots, A^{r-1}$ implies that this sytem has a unique solution. Consequently we can express $s_{0M},\ s_{1M}, \ldots, s_{r-1,\,M}$ as linear combinations *with coefficients independent of* M of the sequences $t_{ij}^{(M)}$. Hence, because each of the latter sequences converges, it follows that each of the sequences $s_{0M},\ s_{1M}, \ldots, s_{r-1,\,M}$ converges to $s_0,\ s_1, \ldots, s_{r-1}$, respectively. Consequently

$$\sum_{k=0}^{r-1} s_{kM} A^k \to \sum_{k=0}^{r-1} s_k A^k$$

and

$$e^A = s_0 I_n + s_1 A + \cdots s_{r-1} A^{r-1},$$

a polynomial in A.

(viii) – (ix) Suppose that $AB = BA$. Then e^{tB} is a polynomial in B and hence A commutes with e^{tB}. Similarly, A and B commute with e^{A+B}. Now let

$$C(t) = e^{t(A+B)}e^{-tB}e^{-tA},\ t \in \mathbb{R}.$$

Then $C(0) = I_n$. Also

$$\begin{aligned} C'(t) &= (A+B)e^{t(A+B)}e^{-tB}e^{-tA} \\ &\quad + e^{t(A+B)}(-B)e^{-tB}e^{-tA} \\ &\quad + e^{t(A+B)}e^{-tB}(-A)e^{-tA} \\ &= 0. \end{aligned}$$

Hence $C(t)$ is a constant matrix and $C(0) = C(1)$. That is

$$I_n = e^{A+B}e^{-B}e^{-A}, \qquad (12)$$

for any matrices A and B which commute.

The special case $B = -A$ then gives

$$I_n = e^0 e^A e^{-A} = e^A e^{-A},$$

thereby proving that e^A is non–singular and $(e^A)^{-1} = e^{-A}$.

Then multiplying both sides of equation (12) on the left by $e^A e^B$ gives the equation $e^A e^B = e^{A+B}$.

In §4.8 we give an application to the solution of a system of differential equations.

(x) Let $J_n(c) = cI_n + N$, where $N = J_n(0)$. Then

$$\begin{aligned} e^{J_n(c)} &= e^{cI_n+N} = e^{cI_n}e^N \\ &= (e^c I_n)\sum_{m=0}^{\infty}\frac{1}{m!}N^m \\ &= \sum_{m=0}^{n-1}\frac{e^c}{m!}N^m. \end{aligned}$$

(xi) Similar to above.

4.8 Systems of differential equations

THEOREM 4.8

If $X = X(t)$ satisfies the system of differential equations

$$\dot{X} = AX,$$

for $t \geq t_0$, where A is a constant matrix, then

$$X = e^{(t-t_0)A}X(t_0).$$

PROOF. Suppose $\dot{X} = AX$ for $t \geq t_0$. Then

$$\begin{aligned} \frac{d}{dt}(e^{-tA}X) &= (-Ae^{-tA})X + e^{-tA}\dot{X} \\ &= (-Ae^{-tA})X + e^{-tA}(AX) \\ &= (-Ae^{-tA})X + (Ae^{-tA})X \\ &= (-Ae^{-tA} + Ae^{-tA})X \\ &= 0X = 0. \end{aligned}$$

Hence the vector $e^{-tA}X$ is constant for $t \geq t_0$. Thus

$$e^{-tA}X = e^{-t_0A}X(t_0)$$

and

$$X = e^{tA}e^{-t_0A}X(t_0) = e^{(t-t_0)A}X(t_0).$$

EXAMPLE 4.3

Solve $\dot{X} = AX$, where

$$A = \begin{bmatrix} 0 & 4 & -2 \\ -1 & -5 & 3 \\ -1 & -4 & 2 \end{bmatrix}.$$

Solution: *$\exists P$ with*

$$\begin{aligned} P^{-1}AP &= J_2(-1) \oplus J_1(-1) \\ &= \begin{bmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \end{aligned}$$

and

$$P^{-1}(tA)P = \begin{bmatrix} -t & 0 & 0 \\ t & -t & 0 \\ 0 & 0 & -t \end{bmatrix}.$$

Thus

$$\begin{aligned} P^{-1}e^{tA}P &= e^{tJ_2(-1) \oplus J_1(-1)} \\ &= e^{tJ_2(-1)} \oplus e^{tJ_1(-1)} \\ &= \left[\begin{array}{cc|c} e^{-t} & 0 & 0 \\ te^{-t} & e^{-t} & 0 \\ \hline 0 & 0 & e^{-t} \end{array}\right] = K(t), \text{ say.} \end{aligned}$$

So $e^{tA} = PK(t)P^{-1}$. *Now*

$$\begin{aligned} X = e^{tA}X_0 &= e^{-t}P \begin{bmatrix} 1 & 0 & 0 \\ t & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} \\ &= e^{-t}P \begin{bmatrix} a \\ at+b \\ c \end{bmatrix}, \end{aligned}$$

where for brevity we have set $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = P^{-1}X_0$.