On the computation of the Jordan canonical form of regular matrix polynomials

G. Kalogeropoulos[1], P. Psarrakos[2] and N. Karcanias[3]

Dedicated to Professor Peter Lancaster on the occasion of his 75th birthday

Abstract

In this paper, an algorithm for the computation of the Jordan canonical form of regular matrix polynomials is proposed. The new method contains rank conditions of suitably defined block Toeplitz matrices and it does not require the computation of the Jordan chains or the Smith form. The Segré and Weyr characteristics are also considered.

Keywords: matrix polynomial; companion linearization; Jordan canonical pair; Jordan chain; Segré characteristic; Weyr characteristic

AMS Subject Classifications: 65F30; 15A21; 15A22; 34A30; 93C05

1 Introduction and preliminaries

Consider a linear system of ordinary differential equations of the form

$$A_m\, q^{(m)}(t) + A_{m-1}\, q^{(m-1)}(t) + \cdots + A_1\, q^{(1)}(t) + A_0\, q(t) \;=\; f(t), \tag{1}$$

where $A_j \in \mathbb{C}^{n\times n}$ $(j = 0, 1, \ldots, m)$ and $q(t), f(t)$ are continuous vector functions with values in $\mathbb{C}^n$ (the indices on $q(t)$ denote derivatives with respect to the independent variable t). Applying the Laplace transformation to (1) yields the associated *matrix polynomial*

$$P(\lambda) \;=\; A_m\lambda^m + A_{m-1}\lambda^{m-1} + \cdots + A_1\lambda + A_0, \tag{2}$$

whose spectral analysis leads to the general solution of (1). The suggested references on matrix polynomials and their applications to differential equations

[1]Department of Mathematics, University of Athens, Panepistimioupolis 15784, Athens, GREECE (E-mail: `gkaloger@math.uoa.gr`).

[2]Department of Mathematics, National Technical University, Zografou Campus 15780, Athens, GREECE (E-mail: `ppsarr@math.ntua.gr`).

[3]Control Engineering Research Centre, School of Engineering and Mathematical Sciences, City University, Northampton Square, London EC1V 0HB, U.K. (E-mail: `n.karcanias@city.ac.uk`).

are [4, 11, 15]. To facilitate the presentation, a brief description of the Jordan structure of the matrix polynomial $P(\lambda)$ in (2) follows.

A system of vectors $\{x_0, x_1, \ldots, x_k\}$, where $x_j \in \mathbb{C}^n$ $(j = 0, 1, \ldots, k)$ and $x_0 \neq 0$, is called a *Jordan chain* of length $k+1$ of $P(\lambda)$ corresponding to the *eigenvalue* λ_0 and the *eigenvector* x_0 if it satisfies the equations

$$\sum_{j=0}^{\xi} \frac{1}{j!} P^{(j)}(\lambda_0)\, x_{\xi-j} \;=\; 0 \;\;;\quad \xi = 0, 1, \ldots, k.$$

The vectors in a Jordan chain are not uniquely defined, and for $m > 1$, they need not be linearly independent [4]. The set of all finite eigenvalues of $P(\lambda)$, that is, $\sigma_F(P) = \{\lambda \in \mathbb{C} : \det P(\lambda) = 0\}$, is called the *finite spectrum* of $P(\lambda)$. In this note, we assume that $P(\lambda)$ in (2) is *regular*, i.e., $\det P(\lambda)$ is not identically zero. As a consequence, $\sigma_F(P)$ contains no more than nm elements.

Let $\lambda_0 \in \sigma_F(P)$ and let $\{x_{1,0}, x_{2,0}, \ldots, x_{\rho,0}\}$ be a basis of the null space $\mathrm{Null}\, P(\lambda_0)$. Suppose that for every $l = 1, 2, \ldots, \rho$, a maximal Jordan chain of $P(\lambda)$ corresponding to λ_0 and $x_{l,0}$ is of the form $\{x_{l,0}, x_{l,1}, \ldots, x_{l,s_l-1}\}$ and has length s_l. Then define $J_P(\lambda_0) = J_1 \oplus J_2 \oplus \cdots \oplus J_\rho$, where J_l $(l = 1, 2, \ldots, \rho)$ is the $s_l \times s_l$ *(elementary) Jordan block* having all its diagonal elements equal to λ_0, ones on the super diagonal and zeros elsewhere. Choosing the order of the Jordan chains such that $s_1 \geq s_2 \geq \cdots \geq s_\rho$, the set $\mathcal{F}_P(\lambda_0) = \{s_1, s_2, \ldots, s_\rho\}$ is uniquely defined [4, 10] and it is said to be the *Segré characteristic* of $P(\lambda)$ at the eigenvalue λ_0 (for linear pencils, see [8, 9, 14]). In this way, the matrix $J_P(\lambda_0)$ is also uniquely defined. Furthermore, if

$$X_P(\lambda_0) \;=\; \left[\begin{array}{ccccccccc} x_{1,0} & x_{1,1} & \cdots & x_{1,s_1-1} & x_{2,0} & \cdots & x_{2,s_2-1} & \cdots & x_{\rho,s_\rho-1} \end{array}\right]$$

is the $n \times (s_1 + s_2 + \cdots + s_\rho)$ matrix whose columns are the vectors of the Jordan chains above, then the pair $(X_P(\lambda_0), J_P(\lambda_0))$ is known as a *Jordan pair* of $P(\lambda)$ corresponding to λ_0 [4, p. 184].

The value $\lambda = \infty$ is said to be an eigenvalue of $P(\lambda)$ if $\hat{\lambda} = 0$ is an eigenvalue of the *algebraic dual matrix polynomial* of $P(\lambda)$, namely,

$$\hat{P}(\lambda) \;=\; \lambda^m P(1/\lambda) \;=\; A_0\lambda^m + A_1\lambda^{m-1} + \cdots + A_{m-1}\lambda + A_m, \tag{3}$$

or equivalently, if the matrix A_m is singular. The *spectrum* of $P(\lambda)$ is defined by $\sigma(P) = \sigma_F(P)$ when A_m is nonsingular, and by $\sigma(P) = \sigma_F(P) \cup \{\infty\}$ when A_m is singular. Moreover, a *Jordan pair* of $P(\lambda)$ corresponding to the infinity is defined by $(X_{P,\infty}, J_{P,\infty}) = (X_{\hat{P}}(0), J_{\hat{P}}(0))$ [4, p. 185]. If the finite spectrum of $P(\lambda)$ is $\sigma_F(P) = \{\lambda_1, \lambda_2, \ldots, \lambda_k\}$, then denote $X_{P,F} = [\, X_P(\lambda_1)\; X_P(\lambda_2)\; \cdots\; X_P(\lambda_k) \,]$ and $J_{P,F} = \oplus_{j=1}^{k} J_P(\lambda_j)$. The $nm \times nm$ matrix $J_P = J_{P,F} \oplus J_{P,\infty}$ and the pair $([\, X_{P,F}\; X_{P,\infty} \,], J_{P,F} \oplus J_{P,\infty})$ are known as the *Jordan canonical form* (which is unique up to permutations of the diagonal Jordan blocks of $J_{P,F}$) and a *Jordan canonical pair* of $P(\lambda)$, respectively, and provide a deep insight into the structure of the system (1) [3, 4, 6, 10, 13, 15]. In particular, suppose that the size of $J_{P,F}$ is $\zeta \times \zeta$, ν is the smallest positive integer satisfying $J_{P,\infty}^{\nu} = 0$ and the vector function $f(t)$ is $(\nu + m - 1)$-times continuously differentiable. Then the general solution of (1) is given by the formula [4, Theorem 8.1]

$$q(t) \;=\; X_{P,F}\, e^{t\,J_{P,F}} c + \int_{t_0}^{t} X_{P,F}\, e^{(t-s)\,J_{P,F}} Y_{P,F} f(s)\,ds - \sum_{j=0}^{\nu-1} X_{P,\infty} J_{P,\infty} Y_{P,\infty} f^{(j)}(t),$$

where the vector $c \in \mathbb{C}^{\zeta}$ is arbitrary, and the matrices $Y_{P,F} \in \mathbb{C}^{\zeta \times n}$ and $Y_{P,\infty} \in \mathbb{C}^{(nm-\zeta)\times n}$ are directly computable by $([\, X_{P,F}\; X_{P,\infty}\,]\,, J_{P,F} \oplus J_{P,\infty})$ [4].

In [10, 15], one can find a second (algebraic) method for the construction of the Jordan matrix J_P, which is based on the notions of the elementary divisors and the Smith form of $P(\lambda)$. In general, this alternative procedure brings also some inconvenience since it requires the computation of the (monic) greatest common divisor of all $j \times j$ minors of $P(\lambda)$ for every $j = 1, 2, \ldots, n$.

In this paper, generalizing results of Karcanias and Kalogeropoulos [8] (see Theorem 1 below), we obtain a new methodology for the construction of J_P, which contains rank conditions of suitably defined block Toeplitz matrices (Theorem 2). Note that the calculation of the ranks can be done in an efficient and reliable way by using existing algorithms based on the well known singular value decomposition or the rank-revealing QR factorizations (see [1, 12] and the references therein). The new method does not require the computation of the Jordan chains or the elementary divisors of $P(\lambda)$, and it can also be formulated in terms of the notion of the Weyr characteristic [8, 9, 14]. Moreover, in Section 3, we give an overall algorithm and two illustrative examples.

2 A piecewise arithmetic progression property

Let $P(\lambda) = A_m\lambda^m + \cdots + A_1\lambda + A_0$ be an $n \times n$ regular matrix polynomial as in (2). An $nm \times nm$ linear pencil $S_1\lambda + S_0$ is called a *linearization* of $P(\lambda)$ if

$$E(\lambda)(S_1\lambda + S_0)G(\lambda) \;=\; P(\lambda) \oplus I_{n(m-1)}$$

for some $nm \times nm$ *unimodular* matrix polynomials $E(\lambda)$ and $G(\lambda)$, i.e., with constant nonzero determinants. The *companion linearization* of $P(\lambda)$ is defined by the $nm \times nm$ linear pencil

$$L_P(\lambda) \;=\; (I_{n(m-1)} \oplus A_m)\lambda - C_P$$

$$= \begin{bmatrix} I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \\ 0 & 0 & \cdots & 0 & A_m \end{bmatrix} \lambda - \begin{bmatrix} 0 & I & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & I \\ -A_0 & -A_1 & \cdots & -A_{m-2} & -A_{m-1} \end{bmatrix}. \quad (4)$$

By [4, p. 186] (see also [10, 11, 13]),

$$E(\lambda)L_P(\lambda)G(\lambda) \;=\; P(\lambda) \oplus I_{n(m-1)}, \quad (5)$$

where the $nm \times nm$ matrix polynomials

$$E(\lambda) = \begin{bmatrix} E_1(\lambda) & E_2(\lambda) & \cdots & E_{m-1}(\lambda) & I \\ -I & 0 & \cdots & 0 & 0 \\ 0 & -I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -I & 0 \end{bmatrix}$$

with $E_m(\lambda) = A_m$ and $E_j(\lambda) = A_j + \lambda E_{j+1}(\lambda)$ $(j = m-1, m-2, \ldots, 1)$, and

$$G(\lambda) = \begin{bmatrix} I & 0 & \cdots & 0 & 0 \\ I\lambda & I & \cdots & 0 & 0 \\ I\lambda^2 & I\lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ I\lambda^{m-1} & I\lambda^{m-2} & \cdots & I\lambda & I \end{bmatrix} \tag{6}$$

are unimodular. In particular, $\det E(\lambda) = \pm 1$ and $\det G(\lambda) = 1$, and the matrix functions $E(\lambda)^{-1}$ and $G(\lambda)^{-1}$ are also (unimodular) matrix polynomials. By (5), it is obvious that $\sigma(L_P) \equiv \sigma(P)$.

At this point, we remark that the notion of the Jordan canonical form of $P(\lambda)$ defined in the previous section is equivalent to the *Weierstrass canonical form*

$$T(\lambda) = (I\lambda - J_{P,F}) \oplus (J_{P,\infty}\lambda - I)$$

of the companion linearization $L_P(\lambda)$. The linear pencil $T(\lambda)$ is also a linearization of $P(\lambda)$ and satisfies $T(\lambda) = ML_P(\lambda)N$ for some nonsingular matrices $M, N \in \mathbb{C}^{nm \times nm}$ [4, Theorems 7.3 and 7.6]. Furthermore, the connection between $J_{P,\infty}$ and the Smith form of $P(\lambda)$ at infinity (see [15] for definitions and details) is given by [13, Theorem 1] (see also [15, Proposition 4.40]).

By the uniqueness of the Weierstrass canonical form of a linear pencil [3], one can obtain the following lemma, and for clarity, we give a simple proof.

Lemma 1 *The matrix polynomial $P(\lambda)$ in (2) and its companion linearization $L_P(\lambda)$ in (4) have exactly the same Jordan canonical form.*

Proof By (5) and [10, Theorem 7.7.1] (see also [13, Proposition 2]), it is clear that $J_{P,F} = J_{L_P,F}$. For the infinite spectrum, consider the algebraic dual matrix polynomial $\hat{P}(\lambda)$ in (3) and its companion linearization $L_{\hat{P}}(\lambda)$. For the $nm \times nm$ unimodular matrix polynomial $G(\lambda)$ in (6) and the algebraic dual pencil of $L_P(\lambda)$, that is,

$$\hat{L}_P(\lambda) = \begin{bmatrix} I & -I\lambda & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & -I\lambda \\ A_0\lambda & A_1\lambda & \cdots & A_{m-2}\lambda & A_{m-1}\lambda + A_m \end{bmatrix},$$

we have

$$
\hat{L}_P(\lambda)G(\lambda)^T = \begin{bmatrix} I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \\ A_0\lambda & A_0\lambda^2 + A_1\lambda & \cdots & A_0\lambda^{m-1} + \cdots + A_{m-2}\lambda & \hat{P}(\lambda) \end{bmatrix}.
$$

Thus, $\hat{L}_P(\lambda)$ is equivalent to $I_{n(m-1)} \oplus \hat{P}(\lambda)$. Moreover, by the discussion on the companion linearizations, it follows that there exist two $nm \times nm$ unimodular matrix polynomials $F_1(\lambda)$ and $F_2(\lambda)$ such that $I_{n(m-1)} \oplus \hat{P}(\lambda) = F_1(\lambda)L_{\hat{P}}(\lambda)F_2(\lambda)$. Hence, the linear pencils $L_{\hat{P}}(\lambda)$ and $\hat{L}_P(\lambda)$ are equivalent, and by [10, Theorem 7.7.1], $J_{L_{\hat{P}},0} = J_{\hat{L}_P,0}$. Consequently,

$$
J_{P,\infty} = J_{\hat{P},0} = J_{L_{\hat{P}},0} = J_{\hat{L}_P,0} = J_{L_P,\infty}. \quad \square
$$

Keeping in mind the above lemma, we can compute $J_P = J_{P,F} \oplus J_{P,\infty}$ by applying to $L_P(\lambda)$ the following result [8].

Theorem 1 (For linear pencils)
Consider an $n \times n$ regular linear pencil $A\lambda + B$ and an eigenvalue $\lambda_0 \in \sigma(A\lambda + B)$. For $k = 1, 2, \ldots,$ define the $nk \times nk$ matrix

$$
\mathcal{Q}_k(\lambda_0) = \begin{bmatrix} A\lambda_0 + B & 0 & \cdots & 0 & 0 \\ A & A\lambda_0 + B & \cdots & 0 & 0 \\ 0 & A & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & A & A\lambda_0 + B \end{bmatrix}
$$

when $\lambda_0 \neq \infty$, and

$$
\mathcal{Q}_k(\infty) = \begin{bmatrix} A & 0 & \cdots & 0 & 0 \\ B & A & \cdots & 0 & 0 \\ 0 & B & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & B & A \end{bmatrix}
$$

when $\lambda_0 = \infty$. Then define the sequence

$$
\nu_k(\lambda_0) = \dim \mathrm{Null}\, \mathcal{Q}_k(\lambda_0) \ ; \quad k = 1, 2, \ldots
$$

and set $\nu_0(\lambda_0) = 0$.

(i) *For every $k = 1, 2, \ldots,$ $2\,\nu_k(\lambda_0) \geq \nu_{k-1}(\lambda_0) + \nu_{k+1}(\lambda_0)$.*

(ii) *If $2\,\nu_k(\lambda_0) = \nu_{k-1}(\lambda_0) + \nu_{k+1}(\lambda_0)$ for some $k = 1, 2, \ldots,$ then there is no $k \times k$ Jordan block of $A\lambda + B$ corresponding to λ_0.*

(iii) *If* $2\,\nu_k(\lambda_0) > \nu_{k-1}(\lambda_0) + \nu_{k+1}(\lambda_0)$ *for some* $k = 1, 2, \ldots,$ *then the Jordan canonical form of* $A\lambda + B$ *has exactly* $2\,\nu_k(\lambda_0) - (\nu_{k-1}(\lambda_0) + \nu_{k+1}(\lambda_0))$ *Jordan blocks of order* k *corresponding to* λ_0.

Moreover, there exists a positive integer τ_0 *such that*

$$\nu_0(\lambda_0) \;<\; \nu_1(\lambda_0) \;<\; \cdots \;<\; \nu_{\tau_0-1}(\lambda_0) \;<\; \nu_{\tau_0}(\lambda_0) \;=\; \nu_{\tau_0+1}(\lambda_0) \;=\; \cdots \;.$$

It is worth noting that for the companion linearization $L_P(\lambda)$ in (4), $\mathcal{Q}_k(\lambda_0)$ is an $nmk \times nmk$ matrix. Thus, the problem of obtaining a direct rank condition for $P(\lambda)$ arises in a natural way. Notice also that the derivative of the linear pencil $A\lambda+B$ in Theorem 1 is $(A\lambda+B)^{(1)} = A$, and that for every $j = 2, 3, \ldots,$ $(A\lambda + B)^{(j)} = 0$.

Theorem 2 (For matrix polynomials)
Let $P(\lambda) = A_m\lambda^m + \cdots + A_1\lambda + A_0$ *be an* $n \times n$ *regular matrix polynomial and let* $\lambda_0 \in \sigma(P)$. *For* $k = 1, 2, \ldots,$ *consider the* $nk \times nk$ *matrix*

$$\mathcal{R}_k(\lambda_0) \;=\; \begin{bmatrix} P(\lambda_0) & 0 & \cdots & 0 & 0 \\ \frac{1}{1!}P^{(1)}(\lambda_0) & P(\lambda_0) & \cdots & 0 & 0 \\ \frac{1}{2!}P^{(2)}(\lambda_0) & \frac{1}{1!}P^{(1)}(\lambda_0) & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{1}{(k-1)!}P^{(k-1)}(\lambda_0) & \frac{1}{(k-2)!}P^{(k-2)}(\lambda_0) & \cdots & \frac{1}{1!}P^{(1)}(\lambda_0) & P(\lambda_0) \end{bmatrix}$$

when $\lambda_0 \neq \infty$, *and*

$$\mathcal{R}_k(\infty) \;=\; \begin{bmatrix} A_m & 0 & \cdots & 0 & 0 \\ A_{m-1} & A_m & \cdots & 0 & 0 \\ A_{m-2} & A_{m-1} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ A_{m-k+1} & A_{m-k+2} & \cdots & A_{m-1} & A_m \end{bmatrix}$$

when $\lambda_0 = \infty$, *where it is assumed that for* $j = -1, -2, \ldots,$ $A_j = 0$. *Define the sequence*

$$\nu_k(\lambda_0) \;=\; \dim \operatorname{Null} \mathcal{R}_k(\lambda_0) \;\;;\quad k = 1, 2, \ldots$$

and set $\nu_0(\lambda_0) = 0$.

(i) *For every* $k = 1, 2, \ldots,$ $2\,\nu_k(\lambda_0) \geq \nu_{k-1}(\lambda_0) + \nu_{k+1}(\lambda_0)$.

(ii) *If* $2\,\nu_k(\lambda_0) = \nu_{k-1}(\lambda_0) + \nu_{k+1}(\lambda_0)$ *for some* $k = 1, 2, \ldots,$ *then there is no* $k \times k$ *Jordan block of* $P(\lambda)$ *corresponding to* λ_0.

(iii) *If* $2\,\nu_k(\lambda_0) > \nu_{k-1}(\lambda_0) + \nu_{k+1}(\lambda_0)$ *for some* $k = 1, 2, \ldots,$ *then the Jordan canonical form of* $P(\lambda)$ *has exactly* $2\,\nu_k(\lambda_0) - (\nu_{k-1}(\lambda_0) + \nu_{k+1}(\lambda_0))$ *Jordan blocks of order* k *corresponding to* λ_0.

Moreover, there is a positive integer τ_0 *such that*

$$\nu_0(\lambda_0) \;<\; \nu_1(\lambda_0) \;<\; \cdots \;<\; \nu_{\tau_0-1}(\lambda_0) \;<\; \nu_{\tau_0}(\lambda_0) \;=\; \nu_{\tau_0+1}(\lambda_0) \;=\; \cdots \;. \tag{7}$$

Proof Assume that $\lambda_0 \in \sigma_F(P)$. In the case $\lambda_0 = \infty$, the proof is similar (see also Remark 1 below). For any $k = 1, 2, \ldots,$ consider the $nmk \times nmk$ matrix

$$\mathcal{Q}_{L_P,k}(\lambda_0) = \begin{bmatrix} L_P(\lambda_0) & 0 & \cdots & 0 & 0 \\ I_{n(m-1)} \oplus A_m & L_P(\lambda_0) & \cdots & 0 & 0 \\ 0 & I_{n(m-1)} \oplus A_m & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_{n(m-1)} \oplus A_m & L_P(\lambda_0) \end{bmatrix},$$

where $L_P(\lambda)$ is the companion linearization of $P(\lambda)$ in (4). Recall (5) and observe that $\mathcal{Q}_{L_P,k}(\lambda_0)$ has the same rank with the $nmk \times nmk$ matrix

$$\mathcal{M}_{L_P,k}(\lambda_0) = \begin{bmatrix} E(\lambda_0) & 0 & \cdots & 0 \\ 0 & E(\lambda_0) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & E(\lambda_0) \end{bmatrix} \mathcal{Q}_{L_P,k}(\lambda_0) \begin{bmatrix} G(\lambda_0) & 0 & \cdots & 0 \\ 0 & G(\lambda_0) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & G(\lambda_0) \end{bmatrix}$$

$$= \begin{bmatrix} P(\lambda_0) \oplus I_{n(m-1)} & 0 & \cdots & 0 & 0 \\ S(\lambda_0) & P(\lambda_0) \oplus I_{n(m-1)} & \cdots & 0 & 0 \\ 0 & S(\lambda_0) & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & S(\lambda_0) & P(\lambda_0) \oplus I_{n(m-1)} \end{bmatrix},$$

where the $nm \times nm$ matrix polynomial $S(\lambda) = E(\lambda)(I_{n(m-1)} \oplus A_m)G(\lambda)$ is of the form

$$S(\lambda) = \begin{bmatrix} B_1(\lambda) & B_2(\lambda) & \cdots & B_{m-1}(\lambda) & B_m(\lambda) \\ -I & 0 & \cdots & 0 & 0 \\ -I\lambda & -I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -I\lambda^{m-2} & -I\lambda^{m-3} & \cdots & -I & 0 \end{bmatrix}$$

with

$$B_l(\lambda) = (m+1-l)A_m\lambda^{m-l} + \cdots + 2A_{l+1}\lambda + A_l \;\; ; \quad l = 1, 2, \ldots, m.$$

For $j = k, k-1, \ldots, 2$, by straightforward computations, we eliminate all the $n \times n$ submatrices $-I, -I\lambda_0, \ldots, -I\lambda_0^{m-2}$ in the $(j, j-1)$-th block $S(\lambda_0)$ of the matrix $\mathcal{M}_{L_P,k}(\lambda_0)$ by using the $n \times n$ identity matrices on the main diagonal of the (j, j)-th block of $\mathcal{M}_{L_P,k}(\lambda_0)$, that is,

$$P(\lambda_0) \oplus I_{n(m-1)} = \begin{bmatrix} P(\lambda_0) & 0 & 0 & \cdots & 0 \\ 0 & I & 0 & \cdots & 0 \\ 0 & 0 & I & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I \end{bmatrix}.$$

This elimination process leads to an $nmk \times nmk$ matrix of the form

$$
\begin{array}{c}
\begin{bmatrix}
P(\lambda_0) & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & I & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\
0 & 0 & \cdots & I & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
\frac{1}{1!}P^{(1)}(\lambda_0) & B_2(\lambda_0) & \cdots & B_m(\lambda_0) & P(\lambda_0) & 0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 & 0 & I & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & I & 0 & \cdots & 0 \\
\frac{1}{2!}P^{(2)}(\lambda_0) & * & \cdots & * & \frac{1}{1!}P^{(1)}(\lambda_0) & * & \cdots & * & P(\lambda_0) & \cdots & 0 \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & I
\end{bmatrix} . \\
\begin{array}{cccc}
\uparrow & \uparrow & \uparrow & \\
\text{1-st column} & \text{(m+1)-th column} & \text{(2m+1)-th column} & \cdots
\end{array}
\end{array}
$$

Consider this matrix as an $mk \times mk$ block matrix with blocks of order n. Then every diagonal block coincides with either the matrix $P(\lambda_0)$ or the identity matrix. Moreover, for every $\xi \in \{1, 2, \ldots, mk\} \setminus \{1, m+1, \ldots, (k-1)m+1\}$, its ξ-th row is

$$
\begin{array}{ccccccccc}
0 & 0 & \cdots & 0 & I & 0 & \cdots & 0 & 0 \\
 & & & & \uparrow & & & & \\
 & & & & \xi\text{-th column} & & & &
\end{array}
$$

and for every $j = 0, 1, \ldots, k-1$, its $(jm+1)$-th column is

$$
\begin{array}{rc}
 & 0 \\
 & \vdots \\
 & 0 \\
(jm+1)\text{-th row} \rightarrow & P(\lambda_0) \\
 & 0 \\
 & \vdots \\
 & 0 \\
(jm+m+1)\text{-th row} \rightarrow & \frac{1}{1!}P^{(1)}(\lambda_0) \\
 & 0 \\
 & \vdots \\
 & 0 \\
(jm+2m+1)\text{-th row} \rightarrow & \frac{1}{2!}P^{(2)}(\lambda_0) \\
 & 0 \\
\vdots & \vdots
\end{array} .
$$

By changing the order of rows and columns, it follows that $\mathcal{M}_{L_P,k}(\lambda_0)$ is equivalent to an $nmk \times nmk$ matrix of the form

$$\begin{bmatrix} \mathcal{R}_k(\lambda_0) & * \\ 0 & I_{n(m-1)k} \end{bmatrix}.$$

As a consequence,

$$\operatorname{rank} \mathcal{Q}_{L_P,k}(\lambda_0) \;=\; \operatorname{rank} \mathcal{M}_{L_P,k}(\lambda_0) \;=\; n(m-1)k + \operatorname{rank} \mathcal{R}_k(\lambda_0),$$

or equivalently,

$$\dim \operatorname{Null} \mathcal{Q}_{L_P,k}(\lambda_0) \;=\; nk - \operatorname{rank} \mathcal{R}_k(\lambda_0) \;=\; \dim \operatorname{Null} \mathcal{R}_k(\lambda_0).$$

The proof is completed by Lemma 1 and Theorem 1. $\square$

Remark 1 For the algebraic dual matrix polynomial $\hat{P}(\lambda)$ in (3) and for $\lambda = 0$, we have $\frac{1}{j!}\,\hat{P}^{(j)}(0) = A_{m-j}$ $(j = 0, 1, \ldots, m)$. Hence, the motivation for the definition of the matrix $\mathcal{R}_k(\infty)$ $(k = 1, 2, \ldots)$ in the above theorem is clear. Note also that for $\lambda_0 = \infty$, a simple proof of the second part of Theorem 2 (for the infinite spectrum) follows by applying the first part of the theorem (for the finite spectrum) to $\hat{P}(\lambda)$ and its eigenvalue $\hat{\lambda}_0 = 0$.

Remark 2 Theorem 2 admits an alternative direct proof, which is independent from the results in [8], based on the following considerations. Let $\lambda_0 \in \sigma(P)$ and consider the Segré characteristic $\mathcal{F}_P(\lambda_0) = \{s_1, s_2, \ldots, s_\rho\}$ $(s_1 \geq s_2 \geq \cdots \geq s_\rho \geq 1)$. Let also

$$x_{l,0},\, x_{l,1},\, \ldots,\, x_{l,s_l-1} \;\; ; \quad l = 1, 2, \ldots, \rho$$

be a set of corresponding (not uniquely defined) maximal Jordan chains. If for any $k = 1, 2, \ldots,$ we denote $\zeta_l = \min\{k, s_l\}$ $(l = 1, 2, \ldots, \rho)$, then by [5, Lemma 2.5], the vectors

$$\begin{bmatrix} 0 \\ \vdots \\ 0 \\ x_{l,0} \\ x_{l,1} \\ \vdots \\ x_{l,j} \end{bmatrix} \in \mathbb{C}^{nk} \;\; ; \quad l = 1, 2, \ldots, \rho,\; j = \zeta_l - 1, \zeta_l - 2, \ldots, 1, 0$$

form a basis of the null space of $\mathcal{R}_k(\lambda_0)$. Thus, $\dim \operatorname{Null} \mathcal{R}_k(\lambda_0) = \sum_{l=1}^{\rho} \zeta_l$, and the statements of Theorem 2 follow readily.

The positive integer τ_0 that satisfies (7) is said to be the *index of annihilation* of $J_P(\lambda_0)$. This index coincides with the largest length of Jordan chains of $P(\lambda)$

corresponding to λ_0, i.e., with the size of the largest Jordan blocks of $J_P(\lambda_0)$ (see [16] for meromorphic matrix functions). The set

$$\mathcal{W}_P(\lambda_0) \,=\, \{w_j(\lambda_0) = \nu_j(\lambda_0) - \nu_{j-1}(\lambda_0),\, j = 1, 2, \ldots, \tau_0\}$$

is made up from positive integers and it is called the *Weyr characteristic* of $J_P(\lambda_0)$. By [8, Remark 5.2] and Theorem 2, it follows that for every $j = 1, 2, \ldots, \tau_0$, the difference $w_j(\lambda_0) = \nu_j(\lambda_0) - \nu_{j-1}(\lambda_0)$ is the number of the Jordan blocks of λ_0 of order at least j.

An eigenvalue λ_0 of the matrix polynomial $P(\lambda)$ in (2) is called *semisimple* if all the corresponding Jordan blocks are of order 1, i.e., the matrix $J_P(\lambda_0)$ is diagonal. Semisimple eigenvalues play an important role in the study of the stability of matrix polynomials under perturbations (see [7] and its references). Theorem 2 leads directly to a characterization of these eigenvalues.

Corollary 1 *An eigenvalue* $\lambda_0 \in \sigma_F(P)$ *is semisimple if and only if*

$$\dim \mathrm{Null}\, P(\lambda_0) \,=\, \dim \mathrm{Null} \begin{bmatrix} P(\lambda_0) & 0 \\ P^{(1)}(\lambda_0) & P(\lambda_0) \end{bmatrix}.$$

Proof Let $\lambda_0 \in \sigma_F(P)$ and consider the sequence $\nu_0(\lambda_0), \nu_1(\lambda_0), \nu_2(\lambda_0), \ldots$ defined in Theorem 2. Then λ_0 is a semisimple eigenvalue of $P(\lambda)$ if and only if the index of annihilation of $J_P(\lambda_0)$ is $\tau_0 = 1$, or equivalently, if and only if $\nu_1(\lambda_0) = \nu_2(\lambda_0)$. $\square$

3 Numerical examples

Let $P(\lambda) = A_m\lambda^m + \cdots + A_1\lambda + A_0$ be an $n \times n$ regular matrix polynomial, and let $\lambda_0 \in \sigma(P)$. Then we can compute the Segré characteristic of $P(\lambda)$ at the eigenvalue λ_0,

$$\mathcal{F}_P(\lambda_0) \,=\, \{s_1, s_2, \ldots, s_\rho\} \quad (s_1 \geq s_2 \geq \cdots \geq s_\rho \geq 1),$$

or equivalently, the Jordan matrix $J_P(\lambda_0)$, by applying the following algorithm.

Step I. Let $\nu_0 = 0$, $\nu_1 = n - \mathrm{rank}\, P(\lambda_0)$ and $k = 1$.

Step II. While $\nu_k \neq \nu_{k-1}$, repeat:

- **(a)** set $k = k + 1$,
- **(b)** if $\lambda_0 \neq \infty$, then construct $P^{(k-1)}(\lambda_0)$,
- **(c)** construct the matrix $\mathcal{R}_k(\lambda_0)$ as in Theorem 2,
- **(d)** compute $\nu_k = nk - \mathrm{rank}\, \mathcal{R}_k(\lambda_0)$.

Step III. Let $\tau_0 = k - 1$ be the index of annihilation of $J_P(\lambda_0)$.

Step IV. For $j = 1, 2, \ldots, \tau_0$, compute the differences

$$d_j \;=\; 2\,\nu_j - (\nu_{j-1} + \nu_{j+1}).$$

Step V. Let $\eta = 0$.

Step VI. For $j = \tau_0, \tau_0 - 1, \ldots, 2, 1$, repeat:

if $d_j > 0$, then set $s_{\eta+1} = \cdots = s_{\eta+d_j} = j$ and $\eta = \eta + d_j$.

Step VII. Print the numbers $s_1 (= \tau_0) \,\geq\, s_2 \,\geq\, \cdots \,\geq\, s_\rho \,\geq\, 1$.

Our results are illustrated in the following two examples. In the first example we construct the Jordan canonical form of a matrix polynomial with finite spectrum. In the second one, we consider the infinite case.

Example 1 [2, Example 6.1]
Consider the 3×3 quadratic matrix polynomial

$$P(\lambda) \;=\; I_3 \lambda^2 + \frac{1}{3} \begin{bmatrix} 8 & 0 & \sqrt{2} \\ 0 & 12 & 0 \\ \sqrt{2} & 0 & 16 \end{bmatrix} \lambda + \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 8 \end{bmatrix}.$$

Then it is easy to see that $\det P(\lambda) = (\lambda + 2)^6$. Hence, $P(\lambda)$ has exactly one eigenvalue, -2, of algebraic multiplicity 6. By [2], the Jordan canonical form of $P(\lambda)$ is

$$J_P \;\equiv\; J_{P,F} \;\equiv\; J_P(-2) \;=\; \begin{bmatrix} -2 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & -2 \end{bmatrix} \oplus \begin{bmatrix} -2 & 1 \\ 0 & -2 \end{bmatrix}. \qquad (8)$$

We can verify that

$$\begin{aligned}
\nu_0(-2) \;&=\; 0 \\
\nu_1(-2) \;&=\; \dim \operatorname{Null} \mathcal{R}_1(-2) \;=\; \dim \operatorname{Null} P(-2) \;=\; 2 \\
\nu_2(-2) \;&=\; \dim \operatorname{Null} \mathcal{R}_2(-2) \;=\; \dim \operatorname{Null} \begin{bmatrix} P(-2) & 0 \\ P^{(1)}(-2) & P(-2) \end{bmatrix} \;=\; 4 \\
\nu_3(-2) \;&=\; \dim \operatorname{Null} \mathcal{R}_3(-2) \\
&=\; \dim \operatorname{Null} \begin{bmatrix} P(-2) & 0 & 0 \\ P^{(1)}(-2) & P(-2) & 0 \\ \frac{1}{2} P^{(2)}(-2) & P^{(1)}(-2) & P(-2) \end{bmatrix} \;=\; 5 \\
\nu_4(-2) \;&=\; \dim \operatorname{Null} \mathcal{R}_4(-2) \\
&=\; \dim \operatorname{Null} \begin{bmatrix} P(-2) & 0 & 0 & 0 \\ P^{(1)}(-2) & P(-2) & 0 & 0 \\ \frac{1}{2} P^{(2)}(-2) & P^{(1)}(-2) & P(-2) & 0 \\ 0 & \frac{1}{2} P^{(2)}(-2) & P^{(1)}(-2) & P(-2) \end{bmatrix} \;=\; 6 \\
\nu_5(-2) \;&=\; \dim \operatorname{Null} \mathcal{R}_5(-2) \;=\; \cdots \;=\; 6.
\end{aligned}$$

Thus, our methodology implies that $P(\lambda)$ has exactly

$$2\,\nu_4(-2) - (\nu_3(-2) + \nu_5(-2)) \;=\; 12 - 11 \;=\; 1$$

Jordan block of order 4 and

$$2\,\nu_2(-2) - (\nu_1(-2) + \nu_3(-2)) \;=\; 8 - 7 \;=\; 1$$

Jordan block of order 2 corresponding to the eigenvalue $\lambda_0 = -2$, and that the index of annihilation of $J_P(-2)$ is $\tau_0 = 4$, confirming Theorem 2. Furthermore, by writing the Weyr characteristic $\mathcal{W}_P(-2) = \{w_j(-2) = \nu_j(-2) - \nu_{j-1}(-2),\ j = 1,2,3,4\}$ in the *Ferrer diagram*

$$\begin{array}{rclclcl}
w_1(-2) &=& \nu_1(-2) - \nu_0(-2) &=& 2-0 &=& 2 & \rightarrow & * & & * \\
w_2(-2) &=& \nu_2(-2) - \nu_1(-2) &=& 4-2 &=& 2 & \rightarrow & * & & * \\
w_3(-2) &=& \nu_3(-2) - \nu_2(-2) &=& 5-4 &=& 1 & \rightarrow & * & & \\
w_4(-2) &=& \nu_4(-2) - \nu_3(-2) &=& 6-5 &=& 1 & \rightarrow & * & & \\
 & & & & & & & & \uparrow & & \uparrow \\
 & & & & & & & & \text{Jordan block} & & \text{Jordan block} \\
 & & & & & & & & \text{of order 4} & & \text{of order 2}
\end{array}$$

and observing that $\nu_5(-2) = \nu_4(-2) = 6$, we conclude once again that the Jordan canonical form of the matrix polynomial $P(\lambda)$ is given by (8).

Example 2 [4, Example 7.1]
For the 2×2 matrix polynomial

$$Q(\lambda) \;=\; \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} \lambda^3 + \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix} \lambda^2 + \begin{bmatrix} -3 & 1 \\ 0 & 1 \end{bmatrix} \lambda + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

it is known [4] that a Jordan pair corresponding to the infinity is

$$(X_{Q,\infty}, J_{Q,\infty}) \;=\; \left(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right).$$

One can see that

$$\begin{aligned}
\nu_0(\infty) &= 0 \\
\nu_1(\infty) &= \dim \mathrm{Null}\, \mathcal{R}_1(\infty) \;=\; \dim \mathrm{Null} \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} \;=\; 1 \\
\nu_2(\infty) &= \dim \mathrm{Null}\, \mathcal{R}_2(\infty) \;=\; \dim \mathrm{Null} \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 3 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \;=\; 2 \\
\nu_3(\infty) &= \dim \mathrm{Null}\, \mathcal{R}_3(\infty)
\end{aligned}$$

$$= \quad \dim \mathrm{Null} \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -3 & 1 & 3 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \; = \; 2.$$

Hence, according to Theorem 2, the matrix polynomial $Q(\lambda)$ has exactly

$$2\,\nu_2(\infty) - (\nu_1(\infty) + \nu_3(\infty)) \; = \; 4 - 3 \; = \; 1$$

Jordan block of order 2 corresponding to the infinity.

Acknowledgments

The authors acknowledge with thanks an anonymous referee for suggesting the direct proof described in Remark 2.

References

[1] C.H. Bischof, G. Quintana-Orti, Computing rank-revealing QR factorizations of dense matrices, ACM Transactions on Mathematical Software 24 (1998) 226-253.

[2] P. Freitas, P. Lancaster, On the optimal value of the spectral abscissa for a system of linear oscillators, SIAM J. Matrix Anal. Appl. 21 (1999) 195-208.

[3] F.R. Gantmacher, The Theory of Matrices, 2nd edition, Chelsea Publishing Company, New York, 1959.

[4] I. Gohberg, P. Lancaster, L. Rodman, Matrix Polynomials, Academic Press, New York, 1982.

[5] I. Gohberg, L. Rodman, Analytic matrix functions with prescribed local data, J. Anal. Math. 40 (1981) 90-128.

[6] G.E. Hayton, A.C. Pugh, P. Fretwell, Infinite elementary divisors of a matrix polynomial and its implications, Internat. J. Control 47 (1988) 53-64.

[7] R. Hryniv, P. Lancaster, On the perturbations of analytic matrix functions, Integral Equations and Operator Theory 34 (1999) 325-338.

[8] N. Karcanias, G. Kalogeropoulos, On the Segré characteristics of right (left) regular matrix pencils, Internat. J. Control 44 (1986) 991-1015.

[9] N. Karcanias, G. Kalogeropoulos, The prime and generalized nullspaces of right regular pencils, Circuits Systems and Signal Process 14 (1995) 485-524.

[10] P. Lancaster, M. Tismenetsky, The Theory of Matrices, 2nd edition, Academic Press, Orlando, 1985.

[11] A.S. Markus, Introduction to the Spectral Theory of Polynomial Operator Pencils, Amer. Math. Society, Providence, Translations of Math. Monographs, Vol. 71, 1988.

[12] G. Quintana-Orti, E.S Quintana-Orti, Paraller codes for computing the numerical rank, Linear Algebra Appl. 275/276 (1998) 451-470.

[13] L. Tan, A.C. Pugh, Spectral structures of the generalized companion form and applications, Systems Control Lett. 46 (2002) 75-84.

[14] H.W. Turnbull, A.C. Aitken, An Introduction to the Theory of Canonical Matrices, Dover Publications, New York, 1961.

[15] A.I.G. Vardulakis, Linear Multivariable Control, John Willey & Sons Ltd, Chichester U.K., 1991.

[16] F. Zhou, A rank criterion for the order of a pole of a matrix function, Linear Algebra Appl. 362 (2003) 287-292.