

LINEARIZATION OF REGULAR MATRIX POLYNOMIALS*

PETER LANCASTER†

Abstract. This note contains a short review of the notion of linearization of regular matrix polynomials. The objective is clarification of this notion when the polynomial has an “eigenvalue at infinity”. The theory is extended to admit reduction by locally unimodular analytic matrix functions.

Key words. Linearization, Regular matrix polynomials.

AMS subject classifications. 15A29, 74A15.

1. Introduction. The objective of this work is clarification of the notion of “linearization” of matrix polynomials, say $P(\lambda) = \sum_{j=0}^{\ell} \lambda^j A_j$, where the $A_j \in \mathbb{C}^{n \times n}$ and $P(\lambda)$ is *regular*, i.e. $\det P(\lambda)$ does not vanish identically. The main ideas to be developed first appeared in the paper [6] but seem to have been largely overlooked by the linear algebra community. In large measure, therefore, this exposition repeats what has been developed more succinctly in [6] - and also used more recently in [10]. However, one new contribution to the theory is developed in Section 3 (and was used in [1]). Many of the underlying ideas concerning *strict equivalence* and *equivalence* transformations can be found (among others) in classic works such as that of Gantmacher [5].

First, if the leading coefficient A_{ℓ} is nonsingular, the notion of linearization is quite familiar. In this case we may form the matrices $\hat{A}_j := A_{\ell}^{-1} A_j$ for $j = 0, 1, \dots, \ell - 1$ and then the “companion matrix” of $P(\lambda)$,

$$C_P = \begin{bmatrix} 0 & I & 0 & \cdots & 0 \\ 0 & 0 & I & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & I \\ -\hat{A}_0 & -\hat{A}_1 & \cdots & -\hat{A}_{\ell-2} & -\hat{A}_{\ell-1} \end{bmatrix}.$$

One linearization of $P(\lambda)$ is then the linear matrix function $\lambda I - C_P$. It is well-known that the spectral structure of this particular linearization reproduces that of $P(\lambda)$ itself. By this we mean that the eigenvalues of $P(\lambda)$ and $\lambda I - C_P$ are the same and,

*Received by the editors on August 8, 2007. Accepted for publication on January 9, 2008.
Handling Editor: Daniel Hershkowitz.

†Dept of Mathematics and Statistics, University of Calgary, Calgary, Alberta, Canada, T2N 1N4
(lancaste@ucalgary.ca).

more importantly, all of their “partial multiplicities”¹ are preserved. Thus, the beauty of the linearization lies in the preservation of these properties but in a **linear** function of λ ; hence the term linearization.

However, having obtained one such function $\lambda A - B$ we can find many more by applying transformations to $\lambda A - B$ which preserve these essential spectral properties. This is true of all “strict equivalence” transformations of $\lambda A - B$. Thus, for any nonsingular E and F , $E(\lambda I - C_P)F$ is also a linearization of $P(\lambda)$. In particular, as long as A_ℓ is nonsingular, we can take $E = \text{diag}\{I, I, \dots, I, A_\ell\}$ and $F = I_{\ell n}$ to see that

$$(1.1) \quad \lambda \begin{bmatrix} I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \\ 0 & 0 & \cdots & 0 & A_\ell \end{bmatrix} - \begin{bmatrix} 0 & I & 0 & \cdots & 0 \\ 0 & 0 & I & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & I \\ -A_0 & -A_1 & \cdots & -A_{\ell-2} & -A_{\ell-1} \end{bmatrix}$$

is a linearization of $P(\lambda)$.

But one can generalise the class of linearizations of $P(\lambda)$ further in the following way: Replace the constant nonsingular matrices E and F of the strict equivalences above by *unimodular*² matrix polynomials $E(\lambda)$ and $F(\lambda)$ for which

$$(1.2) \quad \begin{bmatrix} P(\lambda) & 0 \\ 0 & I_{n(\ell-1)} \end{bmatrix} = E(\lambda)(\lambda A - B)F(\lambda),$$

and then $\lambda A - B$ is also a linearization of $P(\lambda)$ in the sense that all of the partial multiplicities of all eigenvalues are preserved (and this is because the matrix on the left and $\lambda A - B$ have the same Smith normal form (Theorem A.1.1 of [8], for example)). Indeed, this fact is frequently used in the definition of a linearization; as in Section 7.2 of [7]. Candidates for A and B are, of course, I and C_P , or the two matrices of (1.1). But see also some more recent alternatives to the classical companion forms developed in [2] and [3].

On checking the history of the basic ideas, we find that the companion matrix itself plays an important role in the “rational canonical form” of a matrix under similarity (over any field), and this seems to date back more than a hundred years to work of Frobenius (see Chapter V of [15]). The use of a companion matrix in the context of matrix polynomials also has a long history. For example, as the book of Frazer, Duncan, and Collar shows (Section 5.5 of [4]), it was well-established in 1938.

¹Partial multiplicities can also be described as the degrees of “elementary divisors”. See the notes on Theorem 2 below.

²i.e. with determinant independent of λ and non-zero.

More recently, this idea has become a standard tool in the theory of matrices and operators (as in [1] and [13]), and systems theory (see [9] and [14], for example).

2. Admitting a singular leading coefficient. The situation becomes more interesting if A_ℓ is allowed to be singular (as in the so-called “descriptor systems”). Indeed, there are some cases in which one would want to admit $A_\ell = 0$ and this is the point where at least two different practices have emerged.

The theory developed in Chapter 7 of [7] was mainly utilised in the analysis of problems in which, if A_ℓ is singular, then A_0 is nonsingular (and $P(\lambda)$ is certainly regular). Then equation (1.2) still applies (with A and B as in (1.1)) and A is singular if and only if A_ℓ is singular. In this case transformation to the parameter $\mu = \lambda^{-1}$ admits application of earlier theory to the “reverse” polynomial

$$(2.1) \quad P^\#(\lambda) := \lambda^\ell L(\lambda^{-1}) = A_0\lambda^\ell + A_1\lambda^{\ell-1} + \cdots + \lambda A_{\ell-1} + A_\ell.$$

This is where the notion of an “infinite eigenvalue” makes an appearance, and it is natural to say that $P(\lambda)$ has an infinite eigenvalue if and only if $P^\#(\lambda)$ has a zero eigenvalue. Furthermore, the multiplicity and partial multiplicities of the infinite eigenvalue of P are defined to be those of the zero eigenvalue of $P^\#$.

This is all very natural, but a problem appears: Although the (total) algebraic multiplicities of the infinite eigenvalue of P and the zero eigenvalue of $P^\#$ agree, **the partial multiplicities of the eigenvalue at infinity depend on the choice of the linear pencil $\lambda A - B$ used in (1.2)**. This was noticed in [14] for example, and was further investigated in [10] where the following result appears.

THEOREM 2.1. *Let $P(\lambda)$ be an $n \times n$ regular matrix polynomial of degree ℓ with an eigenvalue at infinity of algebraic multiplicity $\kappa > 0$ and let $\kappa = \sum_{j=1}^p k_j$ be any partition of κ into positive integers. Then there is an $\ell n \times \ell n$ linear pencil $\lambda A - B$ which preserves the partial multiplicities of all finite eigenvalues of $P(\lambda)$ and has an eigenvalue at infinity with the p partial multiplicities k_1, \dots, k_p .*

To understand this phenomenon it can be argued that, although the treatment of a finite eigenvalue and its reciprocal - via the properties of P and $P^\#$ - are consistent (the partial multiplicities are preserved) the same is not true of the zero and infinite eigenvalues. In fact, there is an implicit assumption in [7] (see p. 185) that $A_\ell \neq 0$.³ Thus, although zero trailing coefficients A_0, A_1, \dots in P are admissible in [7] and many other works, the same is not true of trailing coefficients $A_\ell, A_{\ell-1}, \dots$ in $P^\#$.

This asymmetry is removed in the following definition of [6], which admits van-

³Although “strong linearization” is recognized in the recent work of [12], the hypothesis that the leading coefficient is nonzero is maintained.

ishing leading coefficients. In this case there must be some *a priori* understanding of the *number* of vanishing leading coefficients, and this will determine uniquely what we will call the *extended degree* of the polynomial. Thus, $P(\lambda) = \sum_{j=0}^{\ell} A_j \lambda^j$ has *extended degree* ℓ if the degree is $\ell_0 < \ell$ and $A_{\ell} = A_{\ell-1} = \dots = A_{\ell_0+1} = 0$, $A_{\ell_0} \neq 0$.

DEFINITION 2.2. An $\ell n \times \ell n$ linear matrix pencil $\lambda A - B$ is a *strong linearization* of the $n \times n$ regular matrix polynomial $P(\lambda)$ of extended degree ℓ if there are unimodular matrix polynomials $E(\lambda)$, $F(\lambda)$ such that

$$(2.2) \quad \begin{bmatrix} P(\lambda) & 0 \\ 0 & I_{n(\ell-1)} \end{bmatrix} = E(\lambda)(\lambda A - B)F(\lambda)$$

and there are unimodular matrix polynomials $H(\lambda)$, $K(\lambda)$ such that

$$(2.3) \quad \begin{bmatrix} P^{\#}(\lambda) & 0 \\ 0 & I_{n(\ell-1)} \end{bmatrix} = H(\lambda)(A - \lambda B)K(\lambda).$$

We emphasize that, with this definition, the treatment of the reciprocal eigenvalues, $\infty \in \sigma(P)$ and $0 \in \sigma(P^{\#})$ (or *vice versa*) enjoy the same symmetry as the treatment of reciprocal pairs of finite eigenvalues and, of course, this definition admits the presence of zero matrices as leading coefficients of $P(\lambda)$. Fortunately, it turns out that the pencil (1.1) is a strong linearization, and so is its “dual”

$$(2.4) \quad \lambda \begin{bmatrix} I & 0 & \dots & 0 & 0 \\ 0 & I & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & I & 0 \\ 0 & 0 & \dots & 0 & A_{\ell} \end{bmatrix} - \begin{bmatrix} 0 & 0 & \dots & 0 & -A_0 \\ I & 0 & \dots & 0 & -A_1 \\ 0 & I & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & 0 & -A_{\ell-2} \\ 0 & 0 & \dots & I & -A_{\ell-1} \end{bmatrix},$$

(see Proposition 1.1 of [6] and also Corollary 2 of [9]).

EXAMPLE 2.3. As a simple illustration consider the polynomial $P(\lambda) = \lambda - 2$, but with extended degree 3. Thus, the reverse polynomial is $P^{\#}(\lambda) = -2\lambda^3 + \lambda^2 = -2\lambda^2(\lambda - \frac{1}{2})$, and there are partial multiplicities 2 and 1 for the eigenvalues of P at infinity and at 2, respectively. The linearization of (1.1) is

$$\lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -1 & 0 \end{bmatrix},$$

with determinant equal to $P(\lambda)$.

For explicit construction of the equivalence transformations in Definition 2.1 see the proof of Proposition 1.1 of [6], for example.

It should also be noted that (given the extended degree) a regular polynomial P has associated “canonical” linearizations. For example, if the analysis is over the complex or real fields the corresponding canonical reductions under strict equivalence transformations are classical. They appear as special cases of Theorems 3.1 and 3.2 of the recent survey paper, [11]. Thus, (writing a Jordan matrix of size s with eigenvalue a as $J_s(a)$) we have the Kronecker canonical form.

THEOREM 2.4. *Every regular pencil $\lambda A - B \in \mathbb{C}^{m \times m}$ is strictly equivalent to a matrix pencil with the block-diagonal form*

$$(2.5) \quad (I_{k_1} - \lambda J_{k_1}(0)) \oplus \cdots \oplus (I_{k_r} - \lambda J_{k_r}(0)) \oplus (\lambda I_{\ell_1} - J_{\ell_1}(\lambda_1)) \oplus \cdots \oplus (\lambda I_{\ell_s} - J_{\ell_s}(\lambda_s)),$$

where $k_1 \leq \cdots \leq k_r$ and $\ell_1 \leq \cdots \leq \ell_s$ are positive integers.

Moreover, the integers k_u are uniquely determined by the pair A, B , and the part

$$(\lambda I_{\ell_1} - J_{\ell_1}(\lambda_1)) \oplus \cdots \oplus (\lambda I_{\ell_s} - J_{\ell_s}(\lambda_s))$$

is uniquely determined by A and B up to a permutation of the diagonal blocks.

Here, k_1, \dots, k_r are the partial multiplicities of the eigenvalue at infinity, the λ_j 's are the finite eigenvalues (not necessarily distinct), and the ℓ 's are their partial multiplicities.

This result tells us that linearizations are determined entirely by their Jordan or Kronecker structures and suggests a second definition in the spirit of the introduction (and consistent with Definition 2.2).

DEFINITION 2.5. An $\ell n \times \ell n$ linear matrix function $\lambda A - B$ is a *linearization* of an $n \times n$ regular matrix polynomial $P(\lambda)$ of extended degree ℓ if these two functions have the same spectrum and, for each finite eigenvalue, the two sets of partial multiplicities are the same.

If, in addition, the point at infinity is a common eigenvalue and its partial multiplicities in $\lambda A - B$ and $P(\lambda)$ agree, then the linearization is said to be *strong*.

3. Another criterion for strong linearization. A generalized criterion for identifying a linearization will be pointed out in this section. It was found to be useful in [1], and is based on the following lemma - which is proved (as Theorem A.6.6) in [8] (and is also known as the “local Smith form”).

LEMMA 3.1. *Let $A(\lambda)$ be an $n \times n$ matrix function which is analytic in a neighbourhood of an eigenvalue λ_0 . Then the partial multiplicities of λ_0 are invariant under*

multiplication of $A(\lambda)$ on the left and on the right by analytic matrix functions which are invertible at λ_0 .

THEOREM 3.2. *Let $P(\lambda)$ be an $n \times n$ regular matrix polynomial of extended degree ℓ and let $\lambda A - B$ be an $\ell n \times \ell n$ linear matrix function. Assume that, for each distinct finite eigenvalue λ_j , there exist functions $E_j(\lambda)$ and $F_j(\lambda)$ which are unimodular and analytic on a neighbourhood of λ_j and for which*

$$(3.1) \quad \begin{bmatrix} P(\lambda) & 0 \\ 0 & I_{n(\ell-1)} \end{bmatrix} = E_j(\lambda)(\lambda A - B)F_j(\lambda),$$

then $\lambda A - B$ is a linearization of $P(\lambda)$.

If $P(\lambda)$ has an eigenvalue at infinity assume also that there are functions $E_0(\lambda)$ and $F_0(\lambda)$ which are unimodular and analytic on a neighbourhood of $\lambda = 0$ and for which

$$(3.2) \quad \begin{bmatrix} P^\#(\lambda) & 0 \\ 0 & I_{n(\ell-1)} \end{bmatrix} = E_0(\lambda)(A - \lambda B)F_0(\lambda).$$

Then $\lambda A - B$ is a strong linearization of $P(\lambda)$.

Proof. The proof is clear. Together with Lemma 3, the hypotheses imply that all partial multiplicities of all eigenvalues of $\lambda A - B$ agree with those of $P(\lambda)$. Then the conclusion follows from Definition 2.5. \square

Acknowledgements. The author is grateful to an anonymous reviewer for helpful and constructive comments. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

- [1] A. Amiraslani, R. Corless, and P. Lancaster. Linearization of matrix polynomials expressed in polynomial bases. *IMA Journal of Numerical Analysis*, to appear.
- [2] E.N. Antoniou and S. Vologiannidis. A new family of companion forms of polynomial matrices. *Electronic Journal of Linear Algebra*, 11:78–87, 2004.
- [3] M. Fiedler. A note on companion matrices. *Linear Algebra and its Applications*, 372:325–331, 2003.
- [4] R.A. Frazer, W.J. Duncan, and A.R. Collar. *Elementary Matrices*, Cambridge, 1938.
- [5] F.R. Gantmacher. *The Theory of Matrices*. Chelsea, New York, 1959.
- [6] I. Gohberg, M.A. Kaashoek, and P. Lancaster. General theory of regular matrix polynomials and band Toeplitz operators. *Integral Equations and Operator Theory*, 11:776–882, 1988.
- [7] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials*, Academic Press, New York, 1982.
- [8] I. Gohberg, P. Lancaster, and L. Rodman. *Invariant Subspaces of Matrices with Applications*, SIAM, Providence, 2006. (Also John Wiley, New York, 1986.)

- [9] N.P. Karampetakis and S. Vologiannidis. Infinite elementary divisor structure-preserving transformations for polynomial matrices. *International Journal of Applied Mathematics and Computer Science*, 13:493–503, 2003.
- [10] P. Lancaster and P. Psarrakos. A note on weak and strong linearizations of regular matrix polynomials. *Manchester Centre for Computational Mathematics: nareport 470*, June 2005.
- [11] P. Lancaster and L. Rodman. Canonical forms for Hermitian matrix pairs under strict equivalence and congruence. *SIAM Review*, 47:407–443, 2005.
- [12] D.S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Vector spaces of linearizations for matrix polynomials. *SIAM Journal of Matrix Analysis and Applications*, 28:971–1004, 2006.
- [13] L. Rodman, *An Introduction to Operator Polynomials*, Birkhäuser, Basel, 1989.
- [14] L. Tan and A.C. Pugh. Spectral structures of the generalized companion form and applications. *Systems and Control Letters*, 46:75–84, 2002.
- [15] H.W. Turnbull and A.C. Aitken. *An Introduction to the Theory of Canonical Matrices*, Blackie & Son, 1932.