Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 607-624, July 2012

ELA

THE BEHAVIOUR OF THE COMPLETE EIGENSTRUCTURE OF A POLYNOMIAL MATRIX UNDER A GENERIC RATIONAL TRANSFORMATION*

VANNI NOFERINI†

Abstract. Given a polynomial matrix $P(x)$ of grade g and a rational function $x(y) = n(y)/d(y)$, where $n(y)$ and $d(y)$ are coprime nonzero scalar polynomials, the polynomial matrix $Q(y) := [d(y)]^g P(x(y))$ is defined. The complete eigenstructures of $P(x)$ and $Q(y)$ are related, including characteristic values, elementary divisors and minimal indices. A theorem on the matter, valid in the most general hypotheses, is stated and proved.

Key words. Polynomial matrix, Rational transformation, Complete eigenstructure, Elementary divisors, Matrix polynomial, Minimal indices.

AMS subject classifications. 15A18, 15A21, 15A54.

1. Introduction.

A polynomial matrix is a matrix whose entries belong to some polynomial ring R [4]. In this paper we will always assume that R is a principal ideal domain. This condition is equivalent to $R = \mathbb{F}[x]$, the ring of univariate polynomials in x with coefficients lying in some field $\mathbb{F}$.

An important property of a polynomial matrix with entries in $\mathbb{F}[x]$ is its complete eigenstructure, whose definition is given in Subsection 2.2. The name *eigenstructure* comes from the special case where $\mathbb{F} = \mathbb{C}$; in this context, polynomial matrices are usually seen instead as matrix polynomials, that is polynomials whose coefficients are matrices [7]. Any matrix polynomial is associated with a polynomial eigenvalue problem (PEP); the complete eigenstructure is strictly related with the properties of the associated PEP. More precisely, it gives the complete information about the eigenvalues, eigenvectors and Jordan chains of the matrix polynomial, and also about the Kronecker form of any strong linearization of the matrix polynomial [1]. Polynomial eigenvalue problems arise in many applications, from mathematics, science and engineering; both their algebraic properties and the numerical methods for their approximate solutions are widely studied; see, e.g., [8, 14, 12].

The aim of this paper is to investigate the link between the complete eigenstruc-

*Received by the editors on November 13, 2011. Accepted for publication on June 10, 2012. Handling Editor: João Filipe Queiró.

†Department of Mathematics, University of Pisa, Largo Bruno Pontecorvo 5, 56127 Pisa, Italy (noferini@mail.dm.unipi.it).

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 607-624, July 2012

tures of two polynomial matrices $P(x)$ and $Q(y)$ related one to another by a rational transformation $x(y)$ of the variable. In order to better explain the question we are interested in, let us consider the following example, where $R = \mathbb{C}[x]$. Suppose that we have to deal with the polynomial matrix

$$P(x) = \begin{bmatrix} x^2 - 20x & 0 & 0 \\ x - 20 & x^2 - 20x & 0 \\ 0 & 0 & x \\ 0 & 0 & x^2 \\ 0 & 0 & 0 \end{bmatrix};$$

if we choose grade$(P(x))= 2$ (the grade of $P(x)$ is an arbitrary integer g such that $g \geq \deg P(x)$; more details are given in Section 2), then the complete eigenstructure of $P(x)$ is the following:

- the elementary divisors of $P(x)$ are $(x - 20)$, x, $(x - 20)$, x^2;
- there are no right minimal indices;
- the left minimal indices of $P(x)$ are $0,\ 1$.

The rational change of variable $x(y) = \frac{16y^2-25}{y^2-y}$ induces a mapping Φ_2, as defined in (3.1), such that $\Phi_2(P(x)) = (y^2 - y)^2 P(\frac{16y^2-25}{y^2-y}) =: Q(y)$, with grade$(Q(y)) = 4$ (see Section 3) and

$$Q(y) = \begin{bmatrix} (25 - 16y^2)(2y - 5)^2 & 0 & 0 \\ (y - y^2)(2y - 5)^2 & (25 - 16y^2)(2y - 5)^2 & 0 \\ 0 & 0 & (y^2 - y)(16y^2 - 25) \\ 0 & 0 & (16y^2 - 25)^2 \\ 0 & 0 & 0 \end{bmatrix}.$$

By studying the complete eigenstructure of $Q(y)$, we find out that

- the elementary divisors of $Q(y)$ are $(y-\frac{5}{2})^2$, $(y-\frac{5}{4})$, $(y+\frac{5}{4})$, $(y-\frac{5}{2})^2$, $(y-\frac{5}{4})^2$, $(y + \frac{5}{4})^2$;
- there are no right minimal indices;
- the left minimal indices of $Q(y)$ are $0,\ 2$.

Notice that $x(\frac{5}{2}) = 20$, $x(\pm\frac{5}{4}) = 0$, and that $y = \frac{5}{2}$ is a root of multiplicity 2 of the equation $x(y) = 20$ while $y = \pm\frac{5}{4}$ are roots of multiplicity 1 of the equation $x(y) = 0$. We can therefore conjecture that if $(x - x_0)^\ell$ is an elementary divisor of $P(x)$ and y_0 is a root of multiplicity m of the equation $x(y) = x_0$ then $(y - y_0)^{m\cdot\ell}$ is an elementary divisor of $Q(y)$. Moreover, we see that apparently the minimal indices have been multiplied by a factor 2; notice that 2 is the degree of the considered rational transformation (that is the maximum of the degrees of the numerator and the denominator).

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 607-624, July 2012

ELA

The main result of the present paper is the proof that the conjectures above, which will be stated more precisely in Section 4, are true for every rational transformation of the variable $x(y)$ and every polynomial matrix $P(x)$. Moreover, analogous properties hold for infinite elementary divisors and right minimal indices.

The motivation for this work comes from the will to generalise the partial results derived in [5], where we considered the particular case of a square and regular polynomial matrix with entries in $\mathbb{C}[x]$ and without infinite elementary divisors, and the Dickson change of variable $x(y) = \frac{y^2+1}{y}$. Moreover, we wish to extend the results by D.S. Mackey and N. Mackey [11], who described the special case of rational transformations of degree 1, also known as Möbius transformations. The present contribution is offered as both a synthesis and an extension of the previous works cited above.

The results provided in this paper can be used to design numerical methods for the approximate solution of PEPs. An example in this regard, restricted to the case of the Dickson transformation, is given in [5] for the solution of the palindromic PEP.

The structure of this paper is the following: in Section 2, we expose the theoretical background we are going to work within, and we give some basic definitions that we will use later on. In Section 3, we formally define the mapping between polynomial matrices induced by a rational change of variable and we present some intermediate results. Our main result is Theorem 4.1, which is stated and commented in Section 4; Sections 5 and 6 are devoted to the proof of our result. For the sake of simplicity, in Sections 5 and 6, we assume that the underlying field is algebraically closed. In Section 7, we show how the result still holds for an arbitrary field. Finally, in Section 8, root polynomials are introduced in order to prove a technical lemma.

The first part of Theorem 4.1 was stated and proved, but only for a very special case, in [5]. Besides the generalisation to a generic rational transformation and a generic polynomial matrix, this paper also contains the analysis of what happens to minimal indices and infinite elementary divisors.

2. Preliminary definitions. In this section, we describe our notation and recall some basic definitions.

2.1. Basic facts on polynomials. Let Z be a ring and let $Z[x]$ be the ring of the univariate polynomials in the variable x with coefficients in Z. We denote the degree of $z \in Z[x]$ by the letter k, and sometimes write $k = \deg z$.

On the other hand, the *grade* [10] of a polynomial $z \in Z[x]$ is any integer $g = \text{grade}(z)$ satisfying $g \geq k$. The choice of the grade of a polynomial is arbitrary: nevertheless, some algebraic properties of polynomial matrices depend on the grade.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 607-624, July 2012

ELA

REMARK 2.1. In some sense, the degree of a polynomial is an intrinsic property while the grade depends on its representation. In fact, informally speaking, the grade depends on how many zero coefficients one wishes to add in front of the polynomial.

Let now g be the grade of $z = \sum_{i=0}^{g} a_i x^i \in Z[x]$. The *reversal* of z with respect to its grade [6, 10] is

$$\mathrm{Rev}_g z := \sum_{i=0}^{g} a_{g-i} x^i. \tag{2.1}$$

The subscript g will sometimes be omitted when the reversal is taken with respect to the degree of the polynomial, that is $\mathrm{Rev}_k z = \mathrm{Rev} z$.

Let now $\mathbb{F}$ be an arbitrary algebraically closed field.

REMARK 2.2. Although the hypothesis that $\mathbb{F}$ is algebraically closed is useful to state in a simpler way our results, it is not strictly necessary; see Section 7.

A well-known result that is crucial to us is that $\mathbb{F}[x]$ is guaranteed to be an Euclidean domain. Given $z_1, z_2 \in \mathbb{F}[x]$, not both zero, we denote by $\mathrm{GCD}(z_1, z_2)$ their greatest common divisor; we additionally require that $\mathrm{GCD}(z_1, z_2)$ is always monic so that it is uniquely defined. We say that z_1 and z_2 are *coprime* if $\mathrm{GCD}(z_1, z_2) = 1_{\mathbb{F}[x]}$.

Notice that a polynomial $z \in \mathbb{F}[x]$ can be thought of as a function $z(x) : \mathbb{F} \to \mathbb{F}$. Thus, applying (2.1), in this case the formula $\mathrm{Rev}_g z(x) = x^g z(x^{-1})$ holds.

Let now $Z^{m\times p}$ be the set of $m \times p$ matrices with entries in Z; the case $p = 1$ corresponds to the set of vectors with m elements in Z, denoted by Z^m. We are mainly interested in analysing $(\mathbb{F}[x])^{m\times p}$, the set of $m \times p$ *polynomial matrices* with entries in $\mathbb{F}[x]$. $M_m(\mathbb{F}[x]) := (\mathbb{F}[x])^{m\times m}$ is the *ring of square polynomial matrices of dimension* m. A square polynomial matrix $A \in M_m(\mathbb{F}[x])$ is said to be *regular* if $\det A \neq 0_{\mathbb{F}[x]}$ and *singular* otherwise. If A is regular and $\det A \in \mathbb{F}$ then A is called *unimodular*.

REMARK 2.3. Notice that $(\mathbb{F}[x])^{m\times p} = (\mathbb{F}^{m\times p})[x]$; or in other words, a polynomial matrix, defined as a matrix whose entries are polynomials, is also a matrix polynomial, defined as a polynomial whose coefficients are matrices.

The notions of grade and degree can be extended in a straightforward way to polynomial matrices, as follows: the grade (resp., the degree) of $A \in (Z[x])^{m\times p}$ is defined as $\max_{i,j} \mathrm{grade}(A_{ij})$ (resp., as $\max_{i,j} \deg A_{ij}$). Analogously, the reversal of a polynomial matrix is defined just as in (2.1), after replacing $a_i \in Z$ with $B_i \in Z^{m\times p}$.

2.2. Characteristic values, elementary divisors, and minimal indices. Let $A \in (F[x])^{m\times p}$, and let $\nu := \min(m, p)$. Suppose that there exist $D_1, \ldots, D_\nu \in$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 607-624, July 2012

$\mathbb{F}[x]$ such that $A_{ij} = D_i\delta_{ij}$, where δ_{ij} is the Kronecker's delta. Then we write $A = \mathrm{diag}(D_1, \ldots, D_\nu)$, and we say that A is *diagonal*. Notice that we use the notation indifferently for both square and rectangular polynomial matrices.

The following theorem, which in its most general form is due to Frobenius [3], is in point of fact valid for any matrix with entries in a principal ideal domain [4, 7].

THEOREM 2.4. *Let $P(x) \in (\mathbb{F}[x])^{m\times p}$. Then there exist two unimodular $A(x) \in M_m(\mathbb{F}[x])$ and $B(x) \in M_p(\mathbb{F}[x])$ such that*

$$S(x) = A(x)P(x)B(x) = \mathrm{diag}(d_1(x), \ldots, d_\nu(x)),$$

where $d_i(x) \in \mathbb{F}[x]$ is monic $\forall i \leq \nu := \min(m, p)$ and $d_i(x)|d_{i+1}(x) \ \ \forall i \leq \nu - 1$.

Such an $S(x) \in (\mathbb{F}[x])^{m\times p}$ is called the *Smith form* [7, 13] of $P(x)$, and the $d_i(x)$ are called its *invariant polynomials* [4, 7]. The Smith form, and thus the invariant polynomials, are uniquely determined by $P(x)$. Notice that a square polynomial matrix $P(x)$ is singular if and only if at least one of its invariant polynomials is zero.

Using the fact that $\mathbb{F}$ is algebraically closed, let us consider a factorization of the invariant polynomials over $\mathbb{F}[x]$: $d_i(x) = \prod_j (x - x_j)^{k_{j,(i)}}$. The factors $(x - x_j)^{k_{j,(i)}}$ are called the *elementary divisors* of $P(x)$ [4, 7] corresponding to the *characteristic value* x_j [4]. Notice that, from Theorem 2.4, $i_1 \leq i_2 \Rightarrow k_{j,(i_1)} \leq k_{j,(i_2)}$.

REMARK 2.5. When $\mathbb{F} = \mathbb{C}$ the characteristic values of the polynomial matrix $P(x)$ are often called the eigenvalues of the matrix polynomial $P(x)$. Given an eigenvalue x_0, there is a Jordan chain of length ℓ at x_0 if and only if $(x - x_0)^\ell$ is an elementary divisor. The number of Jordan chains at x_0 is equal to the number of invariant polynomials that have x_0 as a root [7].

Let us now denote by $\mathbb{F}(x)$ the *field of fractions* of the ring $\mathbb{F}[x]$. Let $\mathcal{V}$ be a vector subspace of $(\mathbb{F}(x))^p$, with $\dim \mathcal{V} = s$. Let $\{v_i\}$ be a polynomial basis for $\mathcal{V}$ with the property $\deg v_1 \leq \cdots \leq \deg v_s$. Often we will arrange a polynomial basis in the matrix form $V(x) = [v_1(x), \ldots, v_s(x)] \in (\mathbb{F}[x])^{p\times s}$. Clearly, polynomial bases always exist, because one may start from any basis with elements in the (vectorial) field of fractions, and then build a polynomial basis just by multiplying by the least common denominator. Let $\alpha_i := \deg v_i$ be the degrees of the vectors of such a polynomial basis; the *order* of $V(x)$ is defined [2] as $\sum_{i=1}^s \alpha_i$. A polynomial basis is called *minimal* [2] if its order is minimal amongst all the polynomial bases for $\mathcal{V}$, and the α_i are called its *minimal indices* [2]. It is possible to prove [2, 4] that, although there is not a unique minimal basis, the minimal indices are uniquely determined by $\mathcal{V}$.

The *right minimal indices* [1] of a polynomial matrix $P(x) \in (\mathbb{F}[x])^{m\times p}$ are defined as the minimal indices of $\ker P(x)$. Analogously, the *left minimal indices* [1] of $P(x)$ are the minimal indices of $\ker P(x)^T$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 607-624, July 2012

ELA

Given the grade g of $P(x)$, we say that ∞ is a characteristic value of $P(x)$ if $0_{\mathbb{F}}$ is a characteristic value of $Rev_g P(x)$. The elementary divisors corresponding to ∞ are defined [9] as the elementary divisors of $Rev_g P(x)$ corresponding to $0_{\mathbb{F}}$; if x^ℓ is an elementary divisor of $Rev_g P(x)$ we formally write that $(x-\infty)^\ell$ is an infinite elementary divisor of $P(x)$. Notice that the infinite elementary divisors of a polynomial matrix clearly depend on the arbitrary choice of its grade.

We complete this section with the following definition [1]: the *complete eigenstructure* of $P(x)$ is the set of both finite and infinite elementary divisors of $P(x)$ and of its left and right minimal indices.

3. Rational transformations of polynomial matrices. Let $n(y), d(y) \in \mathbb{F}[y]$ be two nonzero, coprime polynomials. Let us define $N := \deg n(y)$, $D := \deg d(y)$, and $G := \max(N, D)$. We will always suppose $G \geq 1$, that is $n(y)$ and $d(y)$ are not both elements of $\mathbb{F}$. We denote the coefficients of $n(y)$ and $d(y)$ as $n_i \in \mathbb{F}$, $i = 0, \ldots, N$ and $d_j \in \mathbb{F}$, $j = 0, \ldots, D$, that is $n(y) = \sum_{i=0}^{N} n_i y^i$, $d(y) = \sum_{i=0}^{D} d_i y^i$.

Let us introduce the notation $\mathbb{F}^* := \mathbb{F} \cup \{\infty\}$, having formally defined $\infty := 0_{\mathbb{F}}^{-1}$. We consider the generic rational function from $\mathbb{F}^*$ to $\mathbb{F}^*$:

$$x(y) = \frac{n(y)}{d(y)}. \tag{3.1}$$

The function (3.1) induces a mapping $\Phi_{g,n(y),d(y)} : (\mathbb{F}[x])^{m\times p} \to (\mathbb{F}[y])^{m\times p}$ defined as

$$\Phi_{g,n(y),d(y)}(P(x)) = Q(y) := [d(y)]^g P(x(y)).$$

Here g is the grade of $P(x) \in (\mathbb{F}[x])^{m\times p}$, so for any choice of g a different mapping is defined. We will usually omit the functional dependence of Φ on $n(y)$ and $d(y)$ unless the context allows any possible ambiguity; also, if the grade is chosen to be $g = k$ we will sometimes omit the subscript g, that is $\Phi(P(x)) := \Phi_{k,n(y),d(y)}(P(x))$.

Since a polynomial matrix is also a matrix polynomial, we can write $P(x) = \sum_{i=0}^{g} P_i x^i$ for some $P_i \in \mathbb{F}^{m\times p}$, $i = 0, \ldots, g$. Notice that following the same point of view we can also write $Q(y) = \sum_{i=0}^{g} P_i [n(y)]^i [d(y)]^{g-i}$.

LEMMA 3.1. *$\deg Q(y) = \deg \Phi_g(P(x))$ is less than or equal to $q := gD + \max_{i:P_i\neq 0}(iN - iD)$. If $N \neq D$ the strict equality $\deg Q(y) = q$ always holds. Moreover, $q \leq gG$.*

Proof. Writing $Q(y)$ as above, we can see it as a sum of the $k+1$ polynomial matrices $Q_i(y) = P_i[n(y)]^i[d(y)]^{g-i}$, $0 \leq i \leq k$, with either $Q_i(y) = P_i = 0$ or $\deg Q_i(y) = gD + i(N-D)$. Since the degree of the sum of two polynomials cannot exceed the greatest of the degrees of the considered polynomials, $\deg Q(y)$ cannot be greater than q. Notice that if $N = G$ then $gG \geq q = kG + (g-k)D$ and the maximum

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 607-624, July 2012

is realised by $i = k$, while otherwise the maximum is realised by the smallest index j such that $P_j \neq 0$, and $q = (g-j)G + jN$. This means that if $N < G$ and $P_0 = 0$ then $q < gG$, while $q = gG$ if $N < G$ but $P_0 \neq 0$.

Notice finally that, if $i_1 \neq i_2$, then $Q_{i_1}(y)$ and $Q_{i_2}(y)$ have the same degree if and only if $D = N$. Since $\deg Q_{i_1}(y) \neq \deg Q_{i_2}(y) \Rightarrow \deg(Q_{i_1}(y) + Q_{i_2}(y)) = \max(\deg Q_{i_1}(y), \deg Q_{i_2}(y))$, $D \neq N$ is a sufficient condition for $\deg Q(y) = q$. □

Lemma 3.1 shows that $\deg Q(y) \leq q \leq gG$. The next proposition describes the conditions under which the equality $\deg Q(y) = gG$ holds.

PROPOSITION 3.2. *Let $Q(y) = \Phi(P(x))$. It always holds $\deg Q(y) \leq gG$, and $\deg Q(y) < gG$ if and only if one of the following is true:*

1. *$N > D$ and $g > k$;*
2. *$N \leq D$, and there exist a natural number $a \geq 1$ and a polynomial matrix $\hat{P}(x) \in (\mathbb{F}[x])^{m \times p}$ such that $P(x) = (x - \hat{x})^a \hat{P}(x)$, where $\hat{x} := n_G d_G^{-1}$ if $N = D = G$ and $\hat{x} := 0_{\mathbb{F}}$ if $N < D = G$.*

Proof. Lemma 3.1 guarantees $\deg Q(y) \leq gG$. To complete the proof, there are three possible cases to be analysed.

- If $G = N > D$, we know from Lemma 3.1 that $\deg Q(y) = q$, and in this case $q = gD + kN - kD$. Therefore, $\deg Q(y) = gG \Leftrightarrow g = k$.
- If $N = D = G$, and, we get $q = gG$. Let $Q(y) = \sum_{i=0}^{gG} \Theta_i y^i$: then, $\deg Q(y) < gG \Leftrightarrow \Theta_{gG} = 0_{(\mathbb{F}[x])^{m \times p}}$. On the other hand Θ_{gG} is the coefficient of y^{gG} in $Q(y) = \sum_{i=0}^{g} P_i [n(y)]^i [d(y)]^{g-i}$, so $\Theta_{gG} = d_G^g \sum_{i=0}^{g} P_i n_G^i d_G^{-i} = d_G^g P(n_G d_G^{-1})$. Therefore, Θ_{gG} is zero if and only if every entry of $P(n_G d_G^{-1})$ is equal to $0_{\mathbb{F}[x]}$, or in other words if and only if $P(x) = (x - n_G d_G^{-1})^a \hat{P}(x)$ for some $a \geq 1$ and some suitable polynomial matrix $\hat{P}(x)$.
- If $N < D = G$, recalling the proof of Lemma 3.1 we conclude that $\deg Q(y) < gG$ if and only if $P_0 = 0$, which is equivalent to $P(x) = x^a \hat{P}(x)$ for a suitable value of $a \geq 1$ and some polynomial matrix $\hat{P}(x)$. □

The grade of $Q(y)$ is of course arbitrary, even though it must be greater than or equal to its degree. Since $\deg Q(y) \leq q \leq gG$, we shall define that the grade of $Q(y)$ is gG. This choice has an influence on the infinite elementary divisors of $Q(y)$, as they are equal to the elementary divisors corresponding to zero of the reversal of $Q(y)$ *taken with respect to its grade*, that is $\mathrm{Rev}_{(gG)} Q(y)$.

If one is interested in picking a different choice for the grade of $Q(y)$, the following proposition explains how the infinite elementary divisors change.

PROPOSITION 3.3. *Let $P(x) \in (\mathbb{F}[x])^{m \times p}$, with $k = \deg P(x)$. Then the finite elementary divisors and the minimal indices of $P(x)$ do not depend on its grade, while the*

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 607-624, July 2012

infinite elementary divisors do. Namely, let $\nu = \min(m,p)$*;* $x^{g-k}d_1(x), \ldots, x^{g-k}d_\nu(x)$ *are the invariant polynomials of* $\mathrm{Rev}_g P(x)$ *if and only if* $d_1(x), \ldots, d_\nu(x)$ *are the invariant polynomials of* $\mathrm{Rev}_k P(x)$*, for any choice of* $g \geq k$*.*

Proof. Neither Theorem 2.4 nor the properties of $\ker P(x)$ and $\ker P^T(x)$ depend on the grade, so minimal indices and finite elementary divisors cannot be affected by different choices. Let $S(x) = A(x)\mathrm{Rev}_k P(x)B(x)$ be the Smith form of $\mathrm{Rev}_k P(x)$. We have $\mathrm{Rev}_g P(x) = x^{g-k}\mathrm{Rev}_k P(x)$, which implies that $x^{g-k}S(x) = A(x)\mathrm{Rev}_g P(x)B(x)$. Clearly $d_i(x)|d_j(x) \Leftrightarrow x^{g-k}d_i(x)|x^{g-k}d_j(x)$, and therefore we conclude that $x^{g-k}S(x)$ is the Smith form of $\mathrm{Rev}_g P(x)$. □

Let $\alpha, \beta, \gamma, \delta \in \mathbb{F}$. If $G = 1$, $\Phi_{g,\alpha y+\beta,\gamma y+\delta}$ is clearly invertible and its inverse, with a little abuse of notation, is $Q(y) \to \Phi_{g,\beta-\delta x,\gamma x-\alpha}(Q(y)) = [\gamma x - \alpha]^g Q(\frac{\beta-\delta x}{\gamma x-\alpha})$. The most general case is analysed below.

PROPOSITION 3.4. *Let us denote by* $\mathbb{F}[x]_g$ *the set of the univariate polynomials in* x *whose degree is less than or equal to* g*. Given* g*,* $n(y)$*,* $d(y)$*, the mapping* $\Phi_{g,n(y),d(y)} : (\mathbb{F}[x]_g)^{m\times p} \to (\mathbb{F}[y]_{(gG)})^{m\times p}$ *is always an injective function, but it is not surjective unless* $G = 1$*.*

Proof. Notice that Φ_g can be thought as acting componentwise, sending $P(x)_{ij}$ to $Q(y)_{ij} = \Phi_g(P(x)_{ij})$. Thus, it will be sufficient to show that, in the scalar case $\Phi_g : \mathbb{F}[x]_g \to \mathbb{F}[y]_{(gG)}$, Φ_g is surjective if and only if $G = 1$. This is true because any polynomial that does *not* belong to the set $R_y := \{a(y) \in \mathbb{F}[y] : a(y) = \sum_{i=0}^{g} a_i[d(y)]^{g-i}[n(y)]^i\}$ cannot belong to the image of Φ_g, and $R_y = \mathbb{F}[y]_{(gG)}$ if and only if $G = 1$. In fact, if we require that a generic $r \in \mathbb{F}[y]_{(gG)}$ belongs to R_y, we find out that the $g+1$ coefficients a_i must satisfy $gG+1$ linear constraints.

To prove injectivity: $\Phi_g(P_1(x)) = \Phi_g(P_2(x)) \Rightarrow P_1(x(y)) = P_2(x(y)) \Rightarrow P_1(x) = P_2(x)$. □

Proposition 3.4 tells us that, unless $G = 1$ (the Möbius case), not every $Q(y)$ is such that $Q(y) = \Phi(P(x))$ for some $P(x)$.

A couple of additional definitions will turn out to be useful in the following. Let $x_0 \in \mathbb{F}^*$: we define T_{x_0} as the counterimage of x_0 under the rational function $x(y)$.

Moreover, let $\alpha, \beta \in \mathbb{F}$ be such that $\frac{\alpha}{\beta} = x_0$ and α and β are not both zero. For instance, we can pick $(\alpha, \beta) = (x_0, 1_\mathbb{F})$ if $x_0 \neq \infty$ and $(\alpha, \beta) = (1_\mathbb{F}, 0_\mathbb{F})$ otherwise. Consider the polynomial equation

$$\alpha d(y) = \beta n(y). \tag{3.2}$$

Let S be the degree of the polynomial $\alpha d(y) - \beta n(y)$. Equation (3.2) cannot have

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 607-624, July 2012

more than S finite roots. If $S < G$ then we formally say $\infty \in T_{x_0}$.

REMARK 3.5. Notice that there are three cases that lead to $S < G$:

1. $N = D = G$ and $x_0 = n_G d_G^{-1}$, so that (3.2) becomes $d_G n(y) = n_G d(y)$: in this case, S is the maximum value of i such that $n_i \neq x_0 d_i$;
2. $N < D = G$ and $x_0 = 0_{\mathbb{F}}$, so that (3.2) becomes $n(y) = 0_{\mathbb{F}}$ and $S = N$;
3. $D < N = G$ and $x_0 = \infty$, so that (3.2) is $d(y) = 0_{\mathbb{F}}$ and $S = D$.

We now define the multiplicity m_0 of any finite $y_0 \in T_{x_0}$ as the multiplicity of y_0 as a solution of the polynomial equation (3.2). If $\infty \in T_{x_0}$, its multiplicity is defined to be equal to $G - S$. Therefore, the sum of the multiplicities of all the (both finite and infinite) elements of T_{x_0} is always equal to G, while the sum of the multiplicities of all the finite elements of T_{x_0} is S.

The finite elements of T_{x_0} are characterised by the following proposition.

PROPOSITION 3.6. *Let $y_0 \in \mathbb{F}$ and $x_0 \in \mathbb{F}^*$. Then $y_0 \in T_{x_0}$ if and only if y_0 is a solution of (3.2) for $\alpha, \beta : x_0 = \frac{\alpha}{\beta}$. Moreover, $\alpha_1 d(y_0) = \beta_1 n(y_0)$ and $\alpha_2 d(y_0) = \beta_2 n(y_0)$ if and only if $\frac{\alpha_1}{\beta_1} = \frac{\alpha_2}{\beta_2}$.*

Proof. The definition of T_{x_0} implies the first part of the proposition. The second part comes from the fact that $x(y)$ is a function. □

Proposition 3.6, albeit rather obvious, has the following important implication:

COROLLARY 3.7. *$x_0 \neq x_1 \Leftrightarrow T_{x_0} \cap T_{x_1} = \emptyset$. Equivalently, $\alpha_1\beta_2 \neq \alpha_2\beta_1$ if and only if $[\beta_1 n(y) - \alpha_1 d(y)]$ and $[\beta_2 n(y) - \alpha_2 d(y)] \in \mathbb{F}[y]$ are coprime.*

In particular, for any finite $x_0 \in \mathbb{F}$, $\Phi(x - x_0)$ and $d(y)$ are coprime.

In order to clarify the latter definitions, let us consider an example. Let $\mathbb{F} = \mathbb{C}$ and take $n(y) = y^4 + y^3 - y^2 - y + 1$, $d(y) = y^4$. T_1 is the set of the solutions of the equation $n(y) = d(y)$, so in this case $T_1 = \{-1, 1, \infty\}$. Moreover, the multiplicity of -1 and 1 are, respectively, 1 and 2; since $S = 3$ and $G = 4$, the multiplicity of ∞ is by definition $G - S = 1$. Within the same example, $T_\infty = \{0\}$; 0 has multiplicity 4 because it is a root of order 4 of the equation $d(y) = 0$.

4. Main result.

We are now able to state our main theorem.

THEOREM 4.1. *Given $m, p \in \mathbb{N}_0$ and $n(y), d(y) \in \mathbb{F}[y]$, let $x_0 \in \mathbb{F}^*$ be a characteristic value of $P(x) \in (\mathbb{F}[x])^{m\times p}$, and let $(x-x_0)^{\ell_1}, \ldots, (x-x_0)^{\ell_j}$ be the corresponding elementary divisors. Let g be the grade of $P(x)$, define $G = \max(\deg n(y), \deg d(y))$ and let gG be the grade of $Q(y) = \Phi_g(P(x)) := [d(y)]^g P(\frac{n(y)}{d(y)}) \in (\mathbb{F}[y])^{m\times p}$. Then for any $y_0 \in T_{x_0}$:*

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 607-624, July 2012

- y_0 *is a characteristic value of* $Q(y)$*;*
- $(y-y_0)^{m_0\ell_1},\ \ldots,\ (y-y_0)^{m_0\ell_j}$ *are the elementary divisors corresponding to* y_0 *for* $Q(y)$*, where* m_0 *is the multiplicity of* y_0*.*

Conversely, if $Q(y)=\Phi_g(P(x))$ *for some* $P(x)$*, and if* $y_0\in\mathbb{F}^*$ *is a characteristic value of* $Q(y)$ *with corresponding elementary divisors* $(y-y_0)^{\kappa_1},\ \ldots,\ (y-y_0)^{\kappa_j}$*:*

- $x_0=\frac{n(y_0)}{d(y_0)}$ *is a characteristic value of* $P(x)$*;*
- $m_0|\kappa_i\ \forall i\leq j$*, where* m_0 *is the multiplicity of* y_0 *as an element of* T_{x_0}*, and* $(x-x_0)^{m_0^{-1}\kappa_1},\ \ldots,\ (x-x_0)^{m_0^{-1}\kappa_j}$ *are the elementary divisors corresponding to* x_0 *for* $P(x)$*.*

In addition, the following properties hold:

- *the right minimal indices of* $P(x)$ *are* $\beta_1,\ \ldots,\ \beta_s$ *if and only if the right minimal indices of* $Q(y)$ *are* $G\beta_1,\ \ldots,\ G\beta_s$*;*
- *the left minimal indices of* $P(x)$ *are* $\gamma_1,\ \ldots,\ \gamma_r$ *if and only if the left minimal indices of* $Q(y)$ *are* $G\gamma_1,\ \ldots,\ G\gamma_r$*.*

For any choice of the mapping Φ_g, Theorem 4.1 gives a thorough description of the complete eigenstructure of $\Phi_g(P(x))$ with respect to the complete eigenstructure of $P(x)$. Notice that if $x(y)$ is a Möbius transformation then $m_0\equiv 1$ and $G=1$, so the complete eigenstructure is unchanged but for the shift from one set of characteristic values to another. This is not the case for more general rational transformations, where other changes do happen.

The structure of the proof of Theorem 4.1 is the following. First we prove the first part of the theorem (the statement on elementary divisors). This is done dividing the statement in three cases:

1. $x_0\in\mathbb{F}$ and $y_0\in\mathbb{F}$;
2. $x_0\in\mathbb{F}$ and $y_0=\infty$;
3. $x_0=\infty$.

We first prove that the statement is true for case 1, then show that this implies that it is true for case 2. The validity of cases 1 and 2 implies case 3.

Finally, we prove the second part of the theorem (the statement on minimal indices) with a constructive proof: we build a minimal basis of $Q(y)$ given a minimal basis of $P(x)$, and vice versa.

5. Proof of Theorem 4.1: elementary divisors. The proof relies on the following lemma, whose statement generalises [7, Proposition 11.1]. The proof of the

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 607-624, July 2012

ELA

lemma and more details are given in Section 8.

LEMMA 5.1. *Let $P(x) \in (\mathbb{F}[x])^{m\times p}$ and let $Q(x) = A(x)P(x)B(x)$ where $A(x) \in M_m(\mathbb{F}[x])$ and $B(x) \in M_p(\mathbb{F}[x])$ are both regular, and suppose that $x_0 \in \mathbb{F}$ is neither a root of $\det A(x) \in \mathbb{F}[x]$ nor a root of $\det B(x) \in \mathbb{F}[x]$. Then $P(x)$ and $Q(x)$ have the same elementary divisors associated with x_0.*

5.1. Case 1. Define $\nu := \min(m, p)$, and let $P(x) = A(x)T(x)B(x)$ where $A(x)$ and $B(x)$ are unimodular polynomial matrices, $T(x) =: \operatorname{diag}(\delta_1(x), \dots, \delta_\nu(x))$ is the Smith form of $P(x)$, and $\delta_i(x)$ are its invariant polynomials. Let now $\hat{Q}(y) := \Phi(A(x))\Phi(T(x))\Phi(B(x))$. Clearly, $\hat{Q}(y)$ and $Q(y)$ differ only for a multiplicative factor $[d(y)]^\lambda$, $\lambda \in \mathbb{N}$; moreover, both $\det \Phi(A(x))$ and $\det \Phi(B(x))$ are nonzero whenever $d(y) \neq 0_\mathbb{F}$. Notice that, if x_0 is finite, then for any $y_0 \in T_{x_0}$ there must hold $d(y_0) \neq 0_\mathbb{F}$ (Corollary 3.7). Therefore, Lemma 5.1 implies that $Q(y)$, $\hat{Q}(y)$ and $\Phi(T(x))$ have the same elementary divisors corresponding to y_0.

Unfortunately, $\Phi(T(x))$ may not be the Smith form of $\hat{Q}(y)$, because neither $\Phi(A(x))$ nor $\Phi(B(x))$ are necessarily unimodular and also because $\Phi(\delta_i(x))$ may not be monic. Nevertheless, it has the form $\operatorname{diag}([d(y)]^{k_1}\hat{\delta}_1(y), \dots, [d(y)]^{k_\nu}\hat{\delta}_\nu(y))$, where $k_1 \geq k_2 \geq \cdots \geq k_\nu$ and $\hat{\delta}_i(y) := \Phi(\delta_i(x))$. ¿From Corollary 3.7, $\hat{\delta}_i(y)$ and $d(y)$ cannot share common roots. To reduce $\Phi(T(x))$ into a Smith form, we proceed by steps working on 2×2 principal submatrices.

In each step, we consider the submatrix $\begin{bmatrix} [d(y)]^\gamma\hat{\delta}_i(y) & 0 \\ 0 & [d(y)]^\phi\hat{\delta}_j(y) \end{bmatrix}$, where $\gamma := k_i$ and $\phi := k_j$, with $i < j$. If $\gamma = \phi$, then do nothing; if $\gamma > \phi$, premultiply the submatrix by $\begin{bmatrix} 1_\mathbb{F} & 1_\mathbb{F} \\ -b(y)q(y) & 1_\mathbb{F} - b(y)q(y) \end{bmatrix}$ and postmultiply it by $\begin{bmatrix} a(y) & -q(y) \\ b(y) & [d(y)]^{\gamma-\phi} \end{bmatrix}$, where $q(y) = \hat{\delta}_j(y)/\hat{\delta}_i(y)$ while $a(y)$ and $b(y)$ are such that $a(y)[d(y)]^\gamma\hat{\delta}_i(y) + b(y)[d(y)]^\phi\hat{\delta}_j(y) = [d(y)]^\phi\hat{\delta}_i(y)$; the existence of two such polynomials is guaranteed by Bezout's Lemma, since $[d(y)]^\phi\hat{\delta}_i(y)$ is the greatest common divisor of $[d(y)]^\gamma\hat{\delta}_i(y)$ and $[d(y)]^\phi\hat{\delta}_j(y)$. It is easy to check that both matrices are unimodular, and that the result of the matrix multiplications is $\begin{bmatrix} [d(y)]^\phi\hat{\delta}_i(y) & 0 \\ 0 & [d(y)]^\gamma\hat{\delta}_j(y) \end{bmatrix}$. Hence, by subsequent applications of this algorithm and after having defined a unimodular diagonal matrix $\Delta \in \mathbb{F}^{\nu\times\nu}$ chosen in such a way that the resulting invariant polynomials are monic, it is possible to conclude that the Smith form of $\Phi(T(x))$ is either $S(y) := \Delta \cdot \hat{S}(y)$ or $S(y) := \hat{S}(y) \cdot \Delta$ (whichever of the two products makes sense, depending on whether $m \leq p$ or not), where

$$\hat{S}(y) = \operatorname{diag}([d(y)]^{k_\nu}\hat{\delta}_1(y), \dots, [d(y)]^{k_1}\hat{\delta}_\nu(y)).$$

Thus, the ith invariant polynomial of $P(x)$ has a root of multiplicity ℓ_i at x_0 if and only if the ith invariant polynomial of $\hat{Q}(y)$ has a root of multiplicity $m_0\ell_i$ at $y_0 \in T_{x_0}$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 607-624, July 2012

5.2. Case 2. By definition, the infinite elementary divisors for a given polynomial matrix are the elementary divisors corresponding to zero of the reversal of such polynomial matrix. Therefore, in order to prove Theorem 4.1 for the case of $y_0 = \infty$, we have to analyse the polynomial matrix $Z(y) := \mathrm{Rev}_{(gG)}Q(y) = y^{gG}[d(y^{-1})]^g P(x(y^{-1}))$, and find out what its relation to $P(x)$ is, with particular emphasis to its elementary divisors corresponding to $y_0 = 0_{\mathbb{F}}$. Recalling Remark 3.5, notice that there are two distinct subcases for which $\infty \in T_{x_0}$ for a finite $x_0 \in \mathbb{F}$. We will consider them separately.

5.2.1. Subcase 2.1: $N = D = G$, $x_0 = n_G d_G^{-1}$**.** We get $x(y^{-1}) = \frac{\mathrm{Rev}n(y)}{\mathrm{Rev}d(y)}$ and $y^G d(y^{-1}) = \mathrm{Rev}d(y)$; therefore $Z(y) = [\mathrm{Rev}d(y)]^g P(\frac{\mathrm{Rev}n(y)}{\mathrm{Rev}d(y)})$. This means that we can prove analogous results for $Z(y)$ just by considering this time the new rational transformation $y \to x = \frac{\mathrm{Rev}n(y)}{\mathrm{Rev}d(y)}$. From Remark 3.5, $0_{\mathbb{F}}$ is a root of multiplicity $G - S$ for the equation $\mathrm{Rev}n(y) = x_0 \mathrm{Rev}d(y)$; moreover, since we took the reversal with respect to the degree (or also because of Corollary 3.7), $0_{\mathbb{F}}$ cannot be a root of $\mathrm{Rev}d(y)$. Therefore, following the proof given above, one can state that $P(x)$ has $(x - x_0)^{\ell_1}, \ldots, (x - x_0)^{\ell_j}$ as elementary divisors corresponding to x_0 if and only if $Z(y)$ has the j elementary divisors $y^{(G-S)\ell_1}, \ldots, y^{(G-S)\ell_j}$ corresponding to $0_{\mathbb{F}}$. The thesis follows immediately.

5.2.2. Subcase 2.2: $N < D = G$, $x_0 = 0_{\mathbb{F}}$**.** This time, we can write $x(y^{-1}) = \frac{y^{G-N}\mathrm{Rev}n(y)}{\mathrm{Rev}d(y)}$ and $Z(y) = [\mathrm{Rev}d(y)]^g \ P(\frac{y^{G-N}\mathrm{Rev}n(y)}{\mathrm{Rev}d(y)})$. It is therefore sufficient to consider the transformation $y \to x = y^{G-N}\frac{\mathrm{Rev}n(y)}{\mathrm{Rev}d(y)}$.

In fact, notice that $0_{\mathbb{F}}$ is a solution of multiplicity $G - N$ for the equation $y^{G-N}\mathrm{Rev}n(y) = 0$ ($0_{\mathbb{F}}$ is neither a root of $\mathrm{Rev}n(y)$ nor a root of $\mathrm{Rev}d(y)$, because $\mathrm{Rev}n(0_{\mathbb{F}}) = n_N \neq 0_{\mathbb{F}}$ and $\mathrm{Rev}d(0_{\mathbb{F}}) = d_D \neq 0_{\mathbb{F}}$). Thus, $P(x)$ has the j elementary divisors $x_1^{\ell}, \ldots, x_j^{\ell}$ corresponding to $0_{\mathbb{F}}$ if and only if $Z(y)$ has the j elementary divisors $y^{(G-N)\ell_1}, \ldots, y^{(G-N)\ell_j}$ corresponding to $0_{\mathbb{F}}$, and the thesis follows.

5.3. Case 3. By definition, the infinite elementary divisors of $P(x)$ are the elementary divisors corresponding to the characteristic value $0_{\mathbb{F}}$ for $R(x) := \mathrm{Rev}_g P(x) = x^g P(x^{-1})$. But let $\Psi_{g,n(y),d(y)} = \Phi_{g,d(y),n(y)}$ and $U(y) = \Psi_g(R(x))$, that is to say $U(y) = [n(y)]^g R(\frac{d(y)}{n(y)})$. A simple calculation gives

$$U(y) = [n(y)]^g [\frac{d(y)}{n(y)}]^g P([\frac{d(y)}{n(y)}]^{-1}) = [d(y)]^g P(\frac{n(y)}{d(y)}) = \Phi_g(P(y)) = Q(y).$$

One can therefore follow the proof as in the previous subsections, but starting from $R(x)$ and using a different transformation (notice that the equation $d(y) = 0_{\mathbb{F}}$ defines both T_∞ for the old transformation and $T_{0_{\mathbb{F}}}$ for the new transformation).

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 607-624, July 2012

ELA

6. Proof of Theorem 4.1: minimal indices. We shall only prove the theorem for right minimal indices. The proof for left minimal indices follows from the proof for right minimal indices and from the fact that Φ and the operation of transposition commute, that is $\Phi_g(P^T(x)) = (\Phi_g(P(x)))^T \ \forall\ P(x) \in (\mathbb{F}[x])^{m\times p}$.

6.1. Sufficiency. Let $\dim\ker P(x) = s$, and $V(x) = [v_1(x), \ldots, v_s(x)]$ be a minimal basis for $\ker P(x)$, with minimal indices $\beta_i := \deg v_i \ \forall i = 1, \ldots, s$ and order $B := \sum_{i=1}^{s} \beta_i$. For each value of i let us define $w_i(y) := \Phi_{\beta_i}(v_i(x))$; we also define $W(y) := [w_1(y), \ldots, w_s(y)]$. Clearly $\deg w_i(y) = G\beta_i$. Suppose in fact $\deg w_i(y) \neq G\beta_i$; applying Proposition 3.2 (in the case $g = k = \beta_i$), this would imply that there exists some $x_0 \in \mathbb{F}$ and some polynomial vector $u(x) \in (\mathbb{F}[x])^p$ such that $v_i(x) = (x - x_0)u(x)$. Hence, $[v_1(x), \ldots, (x - x_0)^{-1}v_i(x), \ldots, v_s(x)]$ would be a polynomial basis of order $B - 1$ for $\ker P(x)$, leading to a contradiction. In order to prove that $W(y)$ is a minimal basis for $\ker Q(y)$ we must show that it is a basis and that it is minimal.

Clearly, $w_i(y)$ lies in $\ker Q(y)$ for all i. In fact, $P(x)v_i(x) = 0$ implies that $Q(y)w_i(y) = 0$. So it is sufficient to show that $W(y)$, considered as an element of $(\mathbb{F}(y))^{p\times s}$, has rank s. Notice that $W(y) = V(x(y)) \cdot \mathrm{diag}([d(y)]^{\beta_1}, \ldots, [d(y)]^{\beta_s})$. A well-known property of the rank is that, if $A_1 = A_2A_3$ and A_3 is square and regular, then $\mathrm{rk}(A_1) = \mathrm{rk}(A_2)$. Therefore $\mathrm{rk}(W(y)) = \mathrm{rk}(V(x(y))$, because the diagonal matrix above is regular. Let $\hat{V}(x)$ be some regular $s \times s$ submatrix of $V(x)$, which exists because $\mathrm{rk}(V(x)) = s$. By hypothesis, $\det(\hat{V}(x)) \neq 0_{\mathbb{F}[x]}$, which implies $\det(\hat{V}(x(y))) \neq 0_{\mathbb{F}(y)}$. Hence, $s = \mathrm{rk}(V(x(y))) = \mathrm{rk}(W(y))$. Then $W(y)$ is a basis.

In order to prove that it is minimal, let us introduce the following lemma whose proof can be found in [2].

LEMMA 6.1. *Let $\mathcal{V}$ be a vector subspace of $\mathbb{F}(x)^p$, with $\dim\mathcal{V} = s$. Let $H = [h_1, \ldots, h_s]$ be a polynomial basis of order A for $\mathcal{V}$ and define ξ_i, $i = 1, \ldots, \binom{p}{s}$ to be the $s \times s$ minors (i.e., determinants of $s \times s$ submatrices) of H. Then the following statements are equivalent:*

- *H is a minimal basis for $\mathcal{V}$*
- *The following conditions are both true: (a) $\mathrm{GCD}(\xi_1, \ldots, \xi_r) = 1_{\mathbb{F}[x]}$ and (b) $\max_i \deg\xi_i = A$.*

So, let $\xi_i(y)$ be the $s \times s$ minors of $W(y)$. We shall prove that (a) their GCD is $1_{\mathbb{F}[y]}$ and (b) their maximal degree is $GB = G\sum_{i=1}^{s} \beta_i$. By Lemma 6.1, these two conditions imply that $W(y)$ is minimal. Recall that $w_i(y) = \Phi_{\beta_i}(v_i(x))$, that is to say $w_i(y) = [d(y)]^{\beta_i} v_i(x(y))$. Any $s \times s$ submatrix of $W(y)$ is therefore obtained from the corresponding $s \times s$ submatrix of $V(x)$ by applying the substitution $x = x(y)$ and then multiplying the ith column by $[d(y)]^{\beta_i}$ for $i = 1, \ldots, s$. Let us call $\zeta_i(x)$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 607-624, July 2012

the $s \times s$ minors of $V(x)$. From the properties of determinants we obtain the relation $\xi_i(y) = \left(\prod_{i=1}^{s}[d(y)]^{\beta_i}\right)\zeta_i(x(y)) = [d(y)^B]\zeta_i(x(y)) = \Phi_B(\zeta_i(x))$.

Now for each i let $\gamma_i := \deg \zeta_i(x)$ and $\delta_i := \deg \xi_i(y) \leq \max_{j\leq\gamma_i}(Nj - Dj) + DB$ where the maximum is taken over those values of j such that the jth coefficient of $\xi_i(y)$ is nonzero (Lemma 3.1). There are two cases. If $N \leq D = G$, $\delta_i \leq GB$, and applying Proposition 3.2 (with $g = B$), the inequality holds if and only if $(x-\hat{x})|\zeta_i(x)$, where $\hat{x} = 0_{\mathbb{F}}$ if $N < D$ and $\hat{x} = n_G d_G^{-1}$ if $N = D$; notice that there must be at least one value of i for which $\delta_i = GB$, otherwise $(x - \hat{x})$ would be a common factor of all the $\zeta_i(x)$, which is not possible because of Lemma 6.1. Finally, if $D < N = G$, $\delta_i = \gamma_i G + (B - \gamma_i)D$. Since $V(x)$ is minimal we have $\max_i(\gamma_i) = B$, which implies that also in this case $\max_i(\delta_i) = GB$. This proves condition (b).

Notice moreover that $\xi_i(y) = \Phi_B(\zeta_i(x)) = [d(y)]^{B-\gamma_i}\Phi_{\gamma_i}(\zeta_i(x))$, where the first and the second factor are coprime (because of Corollary 3.7). Let us prove the following lemma.

LEMMA 6.2. *Let $p, q, r \in \mathbb{F}[x]$ with r monic. Then, $\mathrm{GCD}_{\mathbb{F}[x]}(p, q) = r$ if and only if $\mathrm{GCD}_{\mathbb{F}[y]}(\Phi_{\deg p}(p), \Phi_{\deg q}(q)) = \kappa \cdot \Phi_{\deg r}(r)$, where $\kappa \in \mathbb{F}$ is such that $\kappa \cdot \Phi_{\deg r}(r)$ is monic.*

Proof. Let α, β be two suitable elements of $\mathbb{F}$ and let us write the prime factor decompositions $p = \alpha\cdot\prod_i(x-p_i)^{\pi_i}$, $q = \beta\cdot\prod_i(x-q_i)^{\theta_i}$, $r = \prod_i(x-r_i)^{\rho_i}$. Of course we have that $(x-r_i)^{\rho_i}|r$ if and only if there exist i_1, i_2 such that $(x-p_{i_1})^{\pi_{i_1}}|p$, $(x-q_{i_2})^{\theta_{i_2}}|q$, with $p_{i_1} = q_{i_2} = r_i$ and $\rho_i = \min(\pi_{i_1}, \theta_{i_2})$. We get $\Phi_{\deg p}(p) = \alpha \cdot \prod_i(n(y) - p_i d(y))^{\pi_i}$, $\Phi_{\deg q}(q) = \beta \cdot \prod_i(n(y) - q_i d(y))^{\theta_i}$ and $\Phi_{\deg r}(r) = \prod_i(n(y) - r_i d(y))^{\rho_i}$. The thesis follows by invoking Corollary 3.7. □

Lemma 6.2 implies condition (a). This follows from the equation $\mathrm{GCD}_i(\xi_i(y)) = \mathrm{GCD}_i([d(y)]^{B-\gamma_i}) \cdot \mathrm{GCD}_i(\Phi_{\gamma_i}(\zeta_i(x))) = 1_{\mathbb{F}[y]} \cdot 1_{\mathbb{F}[y]}$, where the first $1_{\mathbb{F}[y]}$ comes from the fact that $\max_i(\gamma_i) = B$, while the second $1_{\mathbb{F}[y]}$ comes by applying Lemma 6.2 to $\mathrm{GCD}(\Phi_{\gamma_1}(\zeta_1(x)), \ldots, \Phi_{\gamma_s}(\zeta_s(x))) = \mathrm{GCD}(\ldots \mathrm{GCD}(\Phi_{\gamma_2}(\zeta_2(x)), \Phi_{\gamma_1}(\zeta_1(x)))\ldots)$ and from the identity $\Phi_0(1_{\mathbb{F}[x]}) = 1_{\mathbb{F}[y]}$.

6.2. Necessity. To complete the proof, suppose now that $Q(y) = \Phi_g(P(x))$ for some $P(x) \in \mathbb{F}[x]$ and that $\hat{W}(y)$ is a minimal basis for $\ker Q(y)$, with minimal indices $\epsilon_1 \leq \cdots \leq \epsilon_s$. The other implication that we proved in the previous subsection implies that $G|\epsilon_i \ \forall\ i$, so define $\beta_i = \frac{\epsilon_i}{G}$. Suppose that there exists a minimal basis $\hat{V}(x) = (\hat{v}_1(x), \ldots, \hat{v}_s(x))$ for $\ker P(x)$; suppose moreover that an index $i_0 \in \{1, \ldots, s\}$ exists such that $\deg \hat{v}_{i_0} \neq \beta_{i_0}$. Applying the reverse implication, this would imply that there is a minimal basis $\tilde{W}(y) = (\tilde{w}_1(y), \ldots, \tilde{w}_s(y))$ for $\ker Q(y)$ whose i_0th right minimal index is not equal to ϵ_{i_0}. This is absurd because every minimal basis has the same minimal indices.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 607-624, July 2012

ELA

7. Extension to more relaxed hypotheses. For the sake of convenience, we have so far assumed that the field $\mathbb{F}$ is algebraically closed. This unnecessary hypothesis can be dropped. To see it, assume that $\mathbb{F}$ is not algebraically closed and let $\mathbb{K}$ be the algebraic closure of $\mathbb{F}$. Then $(\mathbb{F}[x])^{m\times p} \subseteq (\mathbb{K}[x])^{m\times p}$, so we can use Theorem 4.1 to identify the Smith forms of $P(x)$ and $Q(y) = \Phi_g(P(x))$ over the polynomial rings $\mathbb{K}[x]$ and $\mathbb{K}[y]$. We can then join back elementary divisors in $\mathbb{K}[x]$ and $\mathbb{K}[y]$ to form elementary divisors in $\mathbb{F}[x]$ and $\mathbb{F}[y]$. Of course, in this case an elementary divisor is no more necessarily associated with a characteristic value in $\mathbb{F}$. For instance, if $\mathbb{F} = \mathbb{Q}$, then the elementary divisor $x^2 + 2$ is not associated with any rational characteristic value, but if we consider the field of complex algebraic numbers $\mathbb{K} = \overline{\mathbb{Q}}$ then we can split it as $(x-\sqrt{2}i)(x+\sqrt{2}i)$ and associate it to the characteristic values $\pm\sqrt{2}i$. Similarly, the other results (e.g., Lemma 6.2) that use the algebraic closure of $\mathbb{F}$ can be straightforwardly extended to a generic field $\mathbb{F}$ via an immersion into its algebraic closure $\mathbb{K}$.

8. Proof of Lemma 5.1. Let $P(x) \in \mathbb{F}^{m\times p}[x]$. If $U(x) := [u_1(x),\ \ldots,\ u_s(x)]$ is a minimal basis for $\ker P(x)$, we define $\ker_{x_0} P(x) := \operatorname{span}(\{u_1(x_0),\ \ldots,\ u_s(x_0)\}) \subseteq \mathbb{F}^p$. In general $\ker_{x_0} P(x)$ is a subset of $\ker P(x_0)$. It is a proper subset when x_0 is a characteristic value of $P(x)$, as is illustrated by the following example: let $\mathbb{F} = \mathbb{C}$ and

$$P(x) = \begin{bmatrix} x & 1 & 0 & 0 \\ 0 & x & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x \end{bmatrix}.$$

Evaluating the polynomial at 0, we get $\ker P(0) = \operatorname{span}(\{[1,0,0,0]^T, [0,0,0,1]^T\})$. On the other hand, a minimal basis for $\ker P(x)$ is $[1, -x, x^2, 0]^T$, so $\ker_0 P(x) = \operatorname{span}(\{[1,0,0,0]^T\})$.

We need now to slightly modify a definition given in [7] in order to extend it to the case of singular and/or rectangular polynomial matrices. A polynomial vector $v(x) \in (\mathbb{F}[x])^p$ is called a *root polynomial* of order ℓ corresponding to x_0 for $P(x)$ if the following conditions are met:

1. x_0 is a zero of order ℓ for $P(x)v(x)$;
2. $v(x_0) \notin \ker_{x_0} P(x)$.

Observe that $v(x_0) \in \ker_{x_0} P(x) \Leftrightarrow \exists\ w(x) \in \ker P(x) \subseteq (\mathbb{F}(x))^p : w(x_0) = v(x_0)$. In fact, let $w(x) = U(x)c(x)$ for some $c(x) \in (\mathbb{F}(x))^s$ and $w(x_0) = v(x_0)$, then $v(x_0) = U(x_0)c(x_0) \in \ker_{x_0} P(x)$[1]. Conversely, write $v(x_0) = U(x_0)c$ for some $c \in \mathbb{F}^s$ and notice that $U(x)c \in \ker P(x)$. Hence, condition 2 implies $v(x) \notin \ker P(x)$.

[1] $c(x_0) \in \mathbb{F}^s$: otherwise, there exists an integer $\alpha \geq 1$ s.t. $d(x) := (x - x_0)^\alpha c(x)$ and $0 \neq d(x_0) \in \mathbb{F}^s$, so $0 = (x_0 - x_0)^\alpha v(x_0) = U(x_0)d(x_0)$, absurd because $U(x_0)$ has rank s.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 607-624, July 2012

In [7, Proposition 1.11], it is shown that given three *regular* polynomial matrices $P(x), A(x), B(x) \in M_n(x)$, and if x_0 is neither a root of $\det A(x)$ nor a root of $\det B(x)$, then $v(x)$ is a root polynomial of order ℓ for $A(x)P(x)B(x)$ corresponding to x_0 if and only if $B(x)v(x)$ is a root polynomial of order ℓ corresponding to x_0 for $P(x)$. The next proposition generalises this result.

PROPOSITION 8.1. *Let $P(x) \in \mathbb{F}^{m\times p}[x]$, $A(x) \in M_m(\mathbb{F}[x])$ and $B(x) \in M_p(\mathbb{F}[x])$. Suppose that both $A(x_0)$ and $B(x_0)$, with $x_0 \in \mathbb{F}$, are full rank matrices. Then $v(x)$ is a root polynomial of order ℓ corresponding to x_0 for $A(x)P(x)B(x)$ if and only if $B(x)v(x)$ is a root polynomial of order ℓ corresponding to x_0 for $P(x)$.*

Proof. Notice that if $A(x_0)$ and $B(x_0)$ are full rank then $A(x)$ and $B(x)$ are regular. In [7], a root polynomial is defined for regular square polynomial matrices, so that condition 2 reduces to $v(x_0) \neq 0$. Nevertheless, the proof given in [7, Proposition 1.11] for condition 1 does not actually use the regularity of $P(x)$, and it is therefore still valid when $P(x)$ is not a regular square polynomial matrix. To complete the proof: $v(x_0) \in \ker_{x_0} A(x)P(x)B(x) \Leftrightarrow \exists w_1(x) \in \ker A(x)P(x)B(x) : w_1(x_0) = v(x_0) \Leftrightarrow \exists w_2(x) \in \ker P(x) : w_2(x_0) = B(x_0)v(x_0) \Leftrightarrow B(x_0)v(x_0) \in \ker_{x_0} P(x)$. To build $w_2(x)$ from $w_1(x)$, simply put $w_2(x) = B(x)w_1(x)$ and use the fact that $A(x)$ is regular. To build $w_1(x)$ from $w_2(x)$, let $(B(x))^{-1}$ be the inverse matrix (over $\mathbb{F}(x)$) of $B(x)$, which exists because $B(x)$ is regular; then, put $w_1(x) = (B(x))^{-1}w_2(x)$. □

Let $v_1(x), \dots, v_s(x)$ be root polynomials corresponding to x_0 of orders $\ell_1 \leq \cdots \leq \ell_s$. We call them a *maximal set of x_0-independent root polynomials* if:

1. they are x_0-independent, i.e., $v_1(x_0), \dots, v_s(x_0)$ are linearly independent;
2. no $(s+1)$-tuple of x_0-independent root polynomials corresponding to x_0 exists;
3. there are no root polynomials corresponding to x_0 of order $\ell > \ell_s$;
4. for all $j = 1, \dots, s-1$, there does not exist a root polynomial $\hat{v}_j(x)$ of order $\hat{\ell}_j > \ell_j$ such that $\hat{v}_j(x), v_{j+1}(x), \dots, v_s(x)$ are x_0-independent.

As long as $\det B(x_0)$ and $\det A(x_0)$ are nonzero, it is easy to check that $v_1(x), \dots, v_s(x)$ are a maximal set of x_0-independent root polynomials for $A(x)P(x)B(x)$ if and only if $B(x)v_1(x), \dots, B(x)v_s(x)$ are a maximal set of x_0-independent root polynomials for $P(x)$. The next proposition completes the proof of Lemma 5.1.

PROPOSITION 8.2. *$P(x) \in (\mathbb{F}[x])^{m\times p}$ has a maximal set of x_0-independent root polynomials, of order $\ell_1, \dots, \ell_s$, if and only if $(x-x_0)^{\ell_1}, \dots, (x-x_0)^{\ell_s}$ are the elementary divisors of $P(x)$ associated with x_0.*

Proof. Let $S(x)$ be the Smith form of $P(x)$, and recall that the inverse of a unimodular polynomial matrix is still a unimodular polynomial matrix [7]. Thus, in view of Proposition 8.1 and Theorem 2.4, it suffices to prove the thesis for $S(x)$. If $S(x)$ is the zero matrix, it has neither a root polynomial nor an elementary divisor, so there is

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 607-624, July 2012

ELA

nothing to prove. Otherwise, let ν be the maximal value of i such that $(S(x))_{ii} \neq 0_{\mathbb{F}[x]}$ and for $j = 1, \ldots, p$ let $e_j \in (\mathbb{F}[x])^p$ be the polynomial vector such that $(e_j)_i = \delta_{ij}$. If $\nu < p$, $[e_{\nu+1}, \ldots, e_p]$ is a minimal basis for $\ker S(x)$ and, being of order 0, also for $\ker_{x_0} S(x)$. Suppose that $v_1(x), \ldots, v_s(x)$ is a maximal set of x_0-independent root polynomials for $S(x)$. Let $k \leq \nu$ be the largest index such that $(v_s(x_0))_k \neq 0_{\mathbb{F}}$: there must exist such an index because $v_s(x_0) \notin \ker_{x_0} S(x)$. Let $(S(x))_{kk} = (x - x_0)^\mu \theta(x)$, with $\theta(x_0) \neq 0_{\mathbb{F}}$. We get $(S(x)v_s(x))_k = (x - x_0)^\mu \theta(x)(v_s(x))_k$, so $\mu \geq \ell_s$. Actually, $\mu = \ell_s$, or e_k would be a root polynomial of order greater than ℓ_s, which is absurd. Then let $k' \leq \nu$ be the largest index not equal to k and such that $(v_{s-1}(x_0))_{k'} \neq 0_{\mathbb{F}}$ (if such and index does not exist, then $v_{s-1}(x_0)$ is, up to a vector lying in $\ker_{x_0} S(x)$, a multiple of e_k and thus $\ell_{s-1} = \ell_s$: in this case, replace without any loss of generality $v_{s-1}(x)$ by a suitable linear combination of $v_{s-1}(x)$ and $v_s(x)$). Following an argument similar as above, we can show that $(S(x))_{k'k'} = (x - x_0)^{\ell_{s-1}} \hat{\theta}(x)$, $\hat{\theta}(x_0) \neq 0_{\mathbb{F}}$. We repeat the process until we find all the s sought elementary divisors. There cannot be more, otherwise $\dim \ker S(x_0) - \dim \ker_{x_0} S(x) > s$ (notice in fact that each elementary divisor associated with x_0 corresponds to a nonzero diagonal element of $S(x)$, say $(S(x))_{ii}$, such that $(S(x_0))_{ii} = 0_{\mathbb{F}}$). Then it would be possible to find an $(s+1)$-tuple of x_0-independent root polynomials, which is absurd.

Conversely, it is easy to check that $e_{\nu-s+1}, \ldots, e_\nu$ are a maximal set of x_0-independent root polynomials. □

REMARK 8.3. Root polynomials carry all the information on Jordan chains [7]. Let $v(x) = \sum_{i=0}^{\ell-1} (x - x_0)^i v_i$, $v_i \in \mathbb{F}^m$, be a root polynomial of order ℓ corresponding to x_0 for $P(x)$. It is possible to prove that then $w(y) = \sum_{i=0}^{\ell-1} [d(y)]^{\ell-1-i} [n(y) - x_0 d(y)]^i v_i$ is a root polynomial of order $m_0 \ell$ corresponding to y_0 for $Q(y)$. The latter formula relates the Jordan chains of $Q(y)$ at y_0 to the Jordan chains of $P(x)$ at x_0.

9. Conclusions. We have shown that if $P(x)$ and $Q(y)$ are polynomial matrices whose entries belong to the ring of univariate polynomials in x (resp. y) with coefficients in any field, and if $P(x)$ and $Q(y)$ are related by a rational transformation $x(y)$, then the complete eigenstructures of $Q(y)$ and $P(x)$ are simply related.

Acknowledgment. I would like to thank N. Mackey for her talk [11] on Möbius transformations that was one of the sources of inspiration for this generalisation on generic rational functions (the other one is [5]). I am also grateful to N. Mackey and D.S. Mackey for the subsequent discussion, and to D.A. Bini and L. Gemignani for useful comments. Finally, I am indebted with an anonymous referee for valuable suggestions and remarks that helped improving the paper.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 23, pp. 607-624, July 2012

REFERENCES

[1] F. De Terán, F. Dopico, and D.S. Mackey. Linearizations of singular matrix polynomials and the recovery of minimal indices. *Electron. J. Linear Algebra*, 18, 371–402, 2009.
[2] G.D. Forney Jr. Minimal bases of rational vector spaces, with applications to multivariable linear systems. *SIAM J. Control*, 13:493–520, 1975.
[3] G. Frobenius. Theorie der linearen Formen mit ganzen Coefficienten. *J. Reine Angew. Math. (Crelle's Journal)*, 86:146–208, 1879.
[4] F.R. Gantmacher. *The Theory of Matrices*. AMS Chelsea, Providence, RI, 1998.
[5] L. Gemignani and V. Noferini. The Ehrlich-Aberth method for palindromic matrix polynomials represented in the Dickson basis. *Linear Algebra Appl.*, doi:10.1016/j.laa.2011.10.035.
[6] I. Gohberg, M.A. Kaashoek, and P. Lancaster. General theory of regular matrix polynomials and band Toeplitz operators. *Integral Equations Operator Theory*, 11:776–882, 1988.
[7] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials*. Academic Press, Inc., New York, 1982.
[8] G. Golub and H.A. van der Vorst. Eigenvalue computation in the 20th century. *J. Comput. Appl. Math.*, 123:35–65, 2000.
[9] G.E. Hayton, A.C. Pugh, and P. Fretwell. Infinite elementary divisors of a matrix polynomial and implications. *Internat. J. Control*, 47:53–64, 1988.
[10] D.S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Smith forms of palindromic matrix polynomials. *Electron. J. Linear Algebra*, 22, 53–91, 2011.
[11] N. Mackey. "Moebius Transformations of Matrix Polynomials." *Talk given on Friday, July 22, 2011*, 7th International Congress on Industrial and Applied Mathematics, Vancouver.
[12] V. Mehrmann and H. Voss. Nonlinear eigenvalue problems: A challenge for modern eigenvalue methods. *GAMM Mitt. Ges. Angew. Math. Mech.*, 27:121–152, 2005.
[13] H.J.S. Smith. On Systems of Linear Indeterminate Equations and Congruences. *Philos. Trans. R. Soc. Lond.*, 151:293–326, 1861.
[14] F. Tisseur and K. Meerbergen. The quadratic eigenvalue problem. *SIAM Rev.*, 43:235–286, 2001.