Matrix pairs in 2D systems: an approach based on trace series and Hankel matrices

Ettore FORNASINI and Maria Elena VALCHER
Dept. of Electronics and Comp. Sci., University of Padova
via Gradenigo 6/a, 35131 PADOVA, Italy
e-mail: meme@paola.dei.unipd.it

Abstract 2D system dynamics depends on matrix pairs which represent the shift operators along coordinate axes. The structure of a matrix pair is analysed referring to its characteristic polynomial and to the traces of suitable matrices in the algebra generated by the elements of the pair. Necessary and sufficient conditions for properties L and P are provided by resorting to Hankel matrix theory. Finite memory and separable systems, as well as 2D systems whose characteristic polynomials exhibit 1D structures, are finally characterized in terms of spectral properties and traces.

Keywords: 2D systems, finite memory systems, separable systems, L property, P property.

AMS Classification: 15A22, 93C35, 93B15, 93B20, 93B25

Running title: Matrix pairs in 2D systems

1 Introduction

"2D systems theory" connotes a fairly large collection of problems and methods held together by a central theme: to understand better the behaviour of processes and devices whose dynamics depends on two independent variables. Most of 2D systems theory is concerned with quarter plane causal models, whose state variable description essentially depends on a pair of square matrices associated with the shift operators along the coordinate axes. There is a long stream of research concerned with the problem of characterizing the structure of matrix pairs (see [17] for an extended bibliography). In spite of its relatively long history, however, several questions are far from a definite solution, and stimulate further research in Linear Algebra.

The purpose of this paper is to highlight how purely algebraic results on matrix pairs apply to two-dimensional system modelling. Conversely, assuming a complementary point of view, we shall show that system theoretic methodologies, connected with the realization problem, lead to a satisfactory algebraic characterization of some special matrix pairs.

Two-dimensional models we refer to are quarter plane causal 2D systems, described by the following equations [3]:

$$\begin{aligned} x(h+1,k+1) &= A_1x(h,k+1)+A_2x(h+1,k) \\ &+ B_1u(h,k+1)+B_2u(h+1,k) \\ y(h,k) &= Cx(h,k), \end{aligned} \tag{1.1}$$

where the input, state and output sequences $u(\cdot,\cdot)$, $x(\cdot,\cdot)$ and $y(\cdot,\cdot)$ are defined on the discrete plane $\mathbf{Z}\times\mathbf{Z}$ and take values in $\mathbf{R}^m$, $\mathbf{R}^n$ and $\mathbf{R}^p$ respectively. A_1, A_2, B_1, B_2 and C are real matrices of suitable dimensions. In general, the initial conditions are assigned by specifying the local state values $x(i,-i)$, $i\in\mathbf{Z}$.

The $n\times n$ matrix pair (A_1,A_2) fully encodes the free evolution of the system, providing at the same time valuable insights into the forced motion. From this point of view, (A_1,A_2) plays the same role as the state transition matrix A in the 1D discrete system:

$$\begin{aligned} x(t+1) &= Ax(t)+Bu(t) \\ y(t) &= Cx(t). \end{aligned} \tag{1.2}$$

In the 2D case, however, no decomposition of the state space into $\{A_1,A_2\}$ - invariant subspaces can be given, allowing an effective representation of the system behaviour as a superposition of elementary modes with simple structure. As the modal analysis approach to the unforced dynamics does not extend to system (1.1), we need to resort to different tools.

The characteristic polynomial of the pair (A_1,A_2),

$$\Delta_{A_1,A_2}(z_1,z_2) := \det(I-A_1z_1-A_2z_2), \tag{1.3}$$

is probably the most useful one. Like the characteristic polynomial of A, which in general does not capture the underlying Jordan structure, Δ_{A_1,A_2} does not identify the similarity orbit of the pair. Nevertheless, several important aspects of the 2D motion completely rely on it. There is, first of all, the internal stability of system (1.1), which depends only on the variety of the zeros of Δ_{A_1,A_2}. On the other hand, the 2D Cayley-Hamilton theorem implies that the free state and output evolution of (1.1) satisfies an autoregressive equation which involves only the coefficients of Δ_{A_1,A_2}. Additional insights into the structure of 2D systems come from the factorization of the characteristic polynomial. Actually, properties like finite memory and separability, and interesting features of the spectrum of $\alpha A_1 + \beta A_2$, such as property L, can be restated as conditions on the factors of Δ_{A_1,A_2}.

A different tool is constituted by the traces of suitable matrices in the algebra generated by A_1 and A_2, and the associated formal power series T_{A_1,A_2}. As the trace series T_{A_1,A_2} biuniquely corresponds to the characteristic polynomial Δ_{A_1,A_2}, in principle both of them provide an equivalent information on the pair (A_1, A_2). On the other hand, we shall see that some properties, originally defined as constraints on the structure of Δ_{A_1,A_2}, are better understood when the trace series point of view is undertaken. This happens, for instance, when finite memory and separable pairs are considered and, more generally, when each irreducible factor of the characteristic polynomial has a support included in some straight line of $\mathbf{Z} \times \mathbf{Z}$. Interestingly enough, resorting to trace series and to well established realization methodologies of system theory allows to translate property L of the pair (A_1, A_2) into a bound on the rank of the Hankel matrix associated with T_{A_1,A_2}.

As previously mentioned, some properties of a matrix pair and, consequently, of the associated 2D systems, do not reduce to conditions on the structure of the characteristic polynomial. Perhaps, the most relevant example is property P that, according to a celebrated result of Mc Coy [10], is equivalent to simultaneous triangularizability.

Actually, when considering only characteristic polynomials, property P and L prove to be indistinguishable, because both of them correspond to linear factorizations of Δ_{A_1,A_2}. Deeper insights into the structure of a matrix pair are offered by the traces of all matrix products $A_{i_1}A_{i_2}\dots A_{i_k},\ k \in \mathbf{N},\ i_j \in \{1,2\}$, and by the associated noncommutative power series. Indeed, representation methods of recognizable and exchangeable power series [2,15], borrowed from automata and languages theory, provide a general framework for analysing property P and, what is more important, a finite criterion for deciding whether a matrix pair is endowed with it.

This paper is organized as follows. In section 2 we explore the main connections existing between characteristic polynomial and traces of a matrix pair, and we present a recursive algorithm for computing the coefficients of the series T_{A_1,A_2} starting from the characteristic polynomial and viceversa. Successively a partial fraction expansion of T_{A_1,A_2}, whose terms are explicitely connected with the irreducible factors of the characteristic polynomial, is provided.

Sections 3 and 4 deal with properties L and P, and their characterizations in terms of commutative and noncommutative power series, respectively. Criteria for testing both properties are provided, based on the aforementioned Hankel matrix approach.

In the last section, the previous results are applied to investigate two important classes of state models, i.e. finite memory and separable 2D systems. As both classes have matrix pairs (A_1, A_2) with property L, it is natural to expect that a variety of different characterizations, typical af the L property, is made available. These are based on the factorization of Δ_{A_1,A_2}, on the structure of the trace series and on the spectrum of the linear combinations $\alpha A_1 + \beta A_2$. Here, however, we follow a somewhat different approach and analyze first matrix pairs (A_1, A_2) with the property that the support of Δ_{A_1,A_2} is a subset of a straight line. The corresponding 2D systems have a free state evolution which exhibits an one-dimensional pattern, and provide the building blocks for synthesizing other classes of systems, in particular finite memory and separable systems, which constitute the main concern of the section. Finally, we show how the Amitsur-Levitzki theorem, suitably revisited, allows for a neat characterization of finite memory and separable 2D systems having property P.

2 Characteristic polynomial and traces of a matrix pair

Given an autonomous 1D system

$$x(t+1) = Ax(t), \tag{2.1}$$

the motion corresponding to any initial state $x(0)$ can be represented by the power series $(I - Az)^{-1}x(0) = \sum_{t=0}^{+\infty} A^t x(0) z^t$. Thus the knowledge of the powers of A or, equivalently, of matrix $(I - Az)^{-1}$, provides a complete information on the dynamics of (2.1).

A weaker, but nevertheless significant information is given by the traces of the powers of A. Actually, as shown by the following lemma, the assigment of the traces is equivalent to that of the characteristic polynomial, which constitutes an invariant, yet not complete, relative to the similarity relation.

Lemma 2.1 *Let A be in $\mathbf{C}^{n\times n}$ and assume* $\det(I - Az) = 1 - d_1 z - d_2 z^2 - ... - d_n z^n$. *Then we have:*

$$\mathrm{tr}A - d_1 = 0 \mathrm{tr}A^2 - d_1 \mathrm{tr}A - 2d_2 = 0 \cdots \mathrm{tr}A^n - d_1 \mathrm{tr}A^{n-1} - \ldots - nd_n = 0 \tag{2.2}$$

and, for $k > 0$,

$$\mathrm{tr}A^{n+k} - d_1 \mathrm{tr}A^{n+k-1} - \ldots - d_n \mathrm{tr}A^k = 0. \tag{2.3}$$

Proof Let $\lambda_1, \lambda_2, \ldots \lambda_n$ be the eigenvalues of A, so that $\det(zI - A) = \prod_{i=1}^n (z - \lambda_i) = z^n - d_1 z^{n-1} - \ldots - d_n$. As the symmetric polynomials $s_k = \sum_{i=1}^n \lambda_i^k$ satisfy Newton's identities [7]

$$s_k - d_1 s_{k-1} - \cdots - k\ d_k = 0, \qquad k = 1, 2, \ldots,$$

and $s_k = \mathrm{tr}A^k = \sum_{i=1}^n \lambda_i^k$, then (2.2) and (2.3) follow ■

Referring to the unforced motion of system (1.1), namely

$$x(h+1, k+1) = A_1\ x(h, k+1) + A_2\ x(h+1, k), \tag{2.4}$$

the doubly indexed sequence of local states $x(h, k)$ induced by an initial global state $\mathcal{X}_0 = \sum_{\ell=-\infty}^{+\infty} x(-\ell, \ell) z_1^{-\ell} z_2^{\ell}$ is represented by the formal power series

$$X(z_1, z_2) = \sum_{h,k} x(h,k) z_1^h z_2^k = (I - A_1 z_1 - A_2 z_2)^{-1} \mathcal{X}_0 = \sum_{i,j=0}^{\infty} (A_1{}^i \sqcup\!\sqcup{}^j A_2\ z_1^i z_2^j) \mathcal{X}_0,$$

where the matrix coefficients $A_1{}^i\sqcup\!\sqcup^j A_2,\ i,j \in \mathbf{N}$, of the power series expansion of $(I - A_1z_1 - A_2z_2)^{-1}$ are inductively defined as

$$A_1{}^i\sqcup\!\sqcup^0 A_2 = A_1^i, \qquad A_1{}^0\sqcup\!\sqcup^j A_2 = A_2^j \tag{2.5}$$

and, when i and j are both greater than zero,

$$A_1{}^i\sqcup\!\sqcup^j A_2 = A_1(A_1{}^{i-1}\sqcup\!\sqcup^j A_2) + A_2(A_1{}^i\sqcup\!\sqcup^{j-1} A_2). \tag{2.6}$$

Basing on (2.5)and (2.6), one easily sees that $A_1{}^i\sqcup\!\sqcup^j A_2 = \sum_{\nu_1,\nu_2,\ldots,\nu_{i+j}} A_{\nu_1}A_{\nu_2}\ldots A_{\nu_{i+j}}$, where the summation is extended to all matrix products that include the factors A_1 and A_2, i and j times respectively. The above decomposition of matrices $A_1{}^i\sqcup\!\sqcup^j A_2$ allows to better understand how the free state evolution depends on the transition matrices A_1 and A_2. Actually, assuming $x(i,-i) = 0$ for $i \neq 0$, the state in (h,k) is given by

$$x(h,k) = \sum_{\nu_1,\nu_2,\ldots,\nu_{i+j}} A_{\nu_1}A_{\nu_2}\ldots A_{\nu_{i+j}}x(0,0), \tag{2.7}$$

and it can be interpreted as the sum of the elementary contributions along all paths connecting $(0,0)$ to (h,k) in the two-dimensional grid.

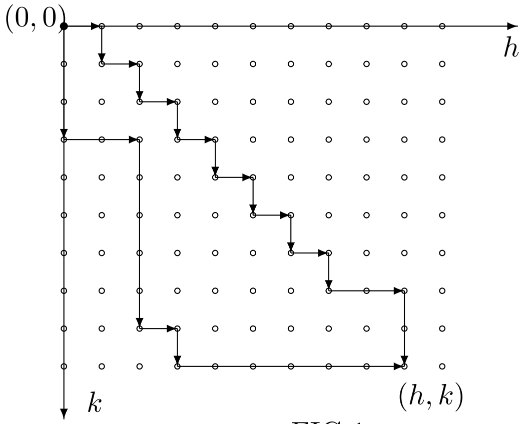

FIG.1

The analogy between the roles played by the matrix family $\{A_1{}^i\sqcup\!\sqcup^j A_2\}$ and the powers of A, can be further highlighted by extending both Cayley-Hamilton theorem and Lemma 2.1 to the 2D case.

Proposition 2.2 [2D Cayley-Hamilton theorem] [4] *Let*

$$\Delta_{A_1,A_2}(z_1,z_2) = 1 - \sum_{i+j\leq n} d_{ij}z_1^i z_2^j \tag{2.8}$$

be the characteristic polynomial of the $n \times n$ matrix pair (A_1, A_2). Then, for all pairs (h,k), with $h + k \geq n$

i) $A_1{}^h\sqcup\!\sqcup^k A_2 = \sum_{i+j\leq n} d_{ij}\ A_1{}^{h-i}\sqcup\!\sqcup^{k-j} A_2$ (2.9)

(where $A_1{}^i\sqcup\!\sqcup^j A_2$ is assumed to be zero whenever i or j is negative);

ii) $A_1{}^h \sqcup^k A_2 \in \mathrm{span}\{A_1{}^i \sqcup^j A_2 : i \leq h,\ j \leq k, i+j < n\}$. (2.10)

Moreover

iii) $\mathrm{span}\{A_1{}^h \sqcup^k A_2 : h,k \in \mathbf{N}\} = \mathrm{span}\{A_1{}^h \sqcup^k A_2 : h,k < n\}$.

PROOF *i)* Since the $(n-1)$-th order minors of $I - A_1 z_1 - A_2 z_2$ have degrees less than n, we have

$$\mathrm{adj}(I - A_1 z_1 - A_2 z_2) = \sum_{r+s<n} M_{rs} z_1^r z_2^s. \tag{2.11}$$

Replace (2.8) and (2.11) into $(I - A_1 z_1 - A_2 z_2)^{-1} \Delta_{A_1,A_2}(z_1, z_2) = \mathrm{adj}(I - A_1 z_1 - A_2 z_2)$, and use the power series expansion of $(I - A_1 z_1 - A_2 z_2)^{-1}$, obtaining

$$(\sum_{h,k} A_1{}^h \sqcup^k A_2\, z_1^h z_2^k)(1 - \sum_{r+s\leq n} d_{rs} z_1^r z_2^s) = \sum_{i+j<n} M_{ij} z_1^i z_2^j. \tag{2.12}$$

Thus, (2.9) simply states that the Cauchy product on the left hand side of (2.12) does not include nonzero monomials with degree greater than $n-1$.

ii) If $h + k = n$, the statement is trivial. If $h + k = \nu + 1 > n$, assume by induction that (2.10) holds for all $(h,k) \in \mathbf{N} \times \mathbf{N}$, with $n \leq h + k \leq \nu$. So, for $r + s > 0$, all matrices $A_1{}^{h-r} \sqcup^{k-s} A_2$ linearly depend on $\{A_1{}^i \sqcup^j A_2 : i \leq h, j \leq k, i+j < n\}$ and the same holds true for $A_1{}^h \sqcup^k A_2$, because of (2.9).

iii) It follows directly from *ii)* ■

In order to extend Lemma 2.1 to the 2D case, rewrite the characteristic polynomial of the pair (A_1, A_2) as

$$\Delta_{A_1,A_2}(z_1, z_2) = 1 - \sum_{h=1}^{n} \Big(\sum_{i+j=h} d_{ij} z_1^i z_2^j \Big) = 1 - \sum_{h=1}^{n} \delta_h(z_1, z_2), \tag{2.13}$$

and introduce the "trace series"

$$T_{A_1,A_2}(z_1, z_2) := \sum_{h=1}^{\infty} \Big(\sum_{i+j=h} \mathrm{tr}(A_1{}^i \sqcup^j A_2) z_1^i z_2^j \Big) = \sum_{h=1}^{\infty} \tau_h(z_1, z_2), \tag{2.14}$$

where $\delta_h(z_1, z_2)$ and $\tau_h(z_1, z_2)$ are homogeneous forms of degree h.

Proposition 2.3 *Let (A_1, A_2) be an $n \times n$ matrix pair with entries in $\mathbf{C}$, and $\Delta_{A_1,A_2}(z_1, z_2)$ and $T_{A_1,A_2}(z_1, z_2)$ its characteristic polynomial and trace series, respectively. Then*

i) the homogeneous components $\delta_h(z_1, z_2)$ and $\tau_h(z_1, z_2)$ satisfy

$$\begin{aligned}
&\tau_1(z_1, z_2) - \delta_1(z_1, z_2) = 0 \\
&\tau_2(z_1, z_2) - \delta_1(z_1, z_2)\tau_1(z_1, z_2) - 2\delta_2(z_1, z_2) = 0 \\
&\quad \cdots \\
&\tau_n(z_1, z_2) - \delta_1(z_1, z_2)\tau_{n-1}(z_1, z_2) - \ldots - n\delta_n(z_1, z_2) = 0
\end{aligned} \tag{2.15}$$

and, for all $k > 0$,

$$\tau_{n+k}(z_1, z_2) - \sum_{i=1}^{n} \tau_{n+k-i}(z_1, z_2)\delta_i(z_1, z_2) = 0; \tag{2.16}$$

ii) the traces of ${A_1}^i ⧢^j A_2$ *and the coefficients* d_{ij} *of* $\Delta_{A_1,A_2}(z_1, z_2)$ *satisfy*

$$\mathrm{tr}({A_1}^i ⧢^j A_2) = \sum_{0<r+s<i+j} d_{rs} \mathrm{tr}({A_1}^{i-r} ⧢^{j-s} A_2) + (i+j) d_{ij}, \tag{2.17}$$

where $d_{rs} = 0$ *for* $r + s > n$, *and* ${A_1}^r ⧢^s A_2$ *is the zero matrix whenever* r *and/or* s *is negative.*

PROOF *i)* Let $\alpha, \beta \in \mathbf{C}$, and substitute in (2.8) z_1 and z_2 for αz and βz:

$$\det[I - (\alpha A_1 + \beta A_2)z] = 1 - \sum_{h=1}^{n} \delta_h(\alpha, \beta) z^h.$$

Taking the traces on both sides of $(\alpha A_1 + \beta A_2)^h = \sum_{i=0}^{h} \alpha^i \beta^{h-i} {A_1}^i ⧢^{h-i} A_2$, one gets

$$\mathrm{tr}(\alpha A_1 + \beta A_2)^h = \sum_{i=0}^{h} \alpha^i \beta^{h-i} \mathrm{tr}({A_1}^i ⧢^{h-i} A_2). \tag{2.18}$$

As (2.18) holds for all α, β in $\mathbf{C}$, it's immediate to recognize in $\mathrm{tr}(\alpha A_1 + \beta A_2)^h$ the homogeneous forms $\tau_h(\alpha, \beta)$ of (2.14). Thus we can apply Lemma 2.1

$$\begin{aligned}
&\tau_1(\alpha, \beta) - \delta_1(\alpha, \beta) = 0 \\
&\tau_2(\alpha, \beta) - \delta_1(\alpha, \beta)\tau_1(\alpha, \beta) - 2\delta_2(\alpha, \beta) = 0 \\
&\cdots \\
&\tau_n(\alpha, \beta) - \delta_1(\alpha, \beta)\tau_{n-1}(\alpha, \beta) - \ldots - n\delta_n(\alpha, \beta) = 0
\end{aligned} \tag{2.19}$$

and, for all $k > 0$,

$$\tau_{n+k}(\alpha, \beta) - \sum_{i=1}^{n} \tau_{n+k-i}(\alpha, \beta)\delta_i(\alpha, \beta) = 0. \tag{2.20}$$

As α and β are arbitrary, (2.15) and (2.16) follow.

ii) Substitute the expressions of $\delta_h(z_1, z_2)$ and $\tau_h(z_1, z_2)$ given in (2.13) and (2.14) into (2.15) and (2.16), and equate to zero the coefficients of all monomials on the left hand side ■

Equation (2.15) has some simple, but useful consequences. First of all, it provides an algorithm for recursively computing the traces of ${A_1}^i ⧢^j A_2$ from the coefficients of the characteristic polynomial. On the other hand, once the traces are given, also the converse, i.e. the computation of the coefficients of Δ, is made possible. Actually, if an upper bound $\bar{n}$ on the degree of Δ is known, assigning $\mathrm{tr}({A_1}^i ⧢^j A_2)$ for $i + j \leq \bar{n}$ allows to recover both Δ and the traces of ${A_1}^i ⧢^j A_2$ for $i + j > \bar{n}$.

Consider the set of all matrix pairs $\mathcal{M} = \{(A_1, A_2) : A_1, A_2 \in \mathbf{C}^{n \times n}, n \in \mathbf{N}\}$, and introduce in $\mathcal{M}$ the equivalence relation

$$(A_1, A_2) \sim (\hat{A}_1, \hat{A}_2) \quad \Leftrightarrow \quad \Delta_{A_1 A_2}(z_1, z_2) = \Delta_{\hat{A}_1 \hat{A}_2}(z_1, z_2).$$

Corollary 2.4 below exhibits different sets of complete invariants for relation $\sim$. Actually, two matrix pairs have the same characteristic polynomial if and only if (the coefficients of) the corresponding trace series coincide. The equivalence relation on $\mathcal{M}$ can also be described in terms of spectra and traces of the linear combinations of the elements of each matrix pair.

Corollary 2.4 *Let A_1, A_2 be in $\mathbf{C}^{n\times n}$ and $\hat{A}_1, \hat{A}_2$ in $\mathbf{C}^{\hat{n}\times\hat{n}}$. The following statements are equivalent:*

i) $\Delta_{A_1,A_2}(z_1, z_2) = \Delta_{\hat{A}_1,\hat{A}_2}(z_1, z_2)$;

ii) for all $\alpha, \beta \in \mathbf{C}$, $\Lambda_0(\alpha A_1 + \beta A_2) = \Lambda_0(\alpha\hat{A}_1 + \beta\hat{A}_2)$, where $\Lambda_0(M)$ denotes the set of nonzero eigenvalues of the matrix M, each of them counted according with the corresponding algebraic multiplicity;

iii) for all $\alpha, \beta \in \mathbf{C}$ and $k \in \mathbf{N}_+$, $\operatorname{tr}(\alpha A_1 + \beta A_2)^k = \operatorname{tr}(\alpha\hat{A}_1 + \beta\hat{A}_2)^k$;

iv) for all $(i, j) \neq (0, 0)$, $\operatorname{tr}(A_1{}^i \sqcup\!\sqcup{}^j A_2) = \operatorname{tr}(\hat{A}_1{}^i \sqcup\!\sqcup{}^j \hat{A}_2)$.

Proof *i) $\Leftrightarrow$ ii)* As both i) and ii) are equivalent to

$$\det[I - (\alpha A_1 + \beta A_2)z] = \det[I - (\alpha\hat{A}_1 + \beta\hat{A}_2)z], \quad \forall\alpha, \beta \in \mathbf{C},$$

they are equivalent each other, too.

i) $\Leftrightarrow$ iii) $\Leftrightarrow$ iv) By Proposition 2.3, (A_1, A_2) and $(\hat{A}_1, \hat{A}_2)$ have the same characteristic polynomial if and only if the corresponding homogeneous forms $\operatorname{tr}(\alpha A_1 + \beta A_2)^k$ and $\operatorname{tr}(\alpha\hat{A}_1 + \beta\hat{A}_2)^k$, $k = 1, 2\ldots$ coincide. This, in turn, is equivalent to assume $\operatorname{tr}(A_1{}^i \sqcup\!\sqcup{}^j A_2) = \operatorname{tr}(\hat{A}_1{}^i \sqcup\!\sqcup{}^j \hat{A}_2)$, for all $(i, j) \neq (0, 0)$ $\blacksquare$

It's easy to realize that T_{A_1,A_2} has to be a rational power series, since its coefficients satisfy the recursive equations (2.17). In the remaining of the section we aim to make explicit its rational structure and what are its connections with the characteristic polynomial.

Proposition 2.5 *Let $\Delta(z_1, z_2) = 1 - \sum_{h=1}^{n} \delta_h(z_1, z_2)$ be the characteristic polynomial of the matrix pair (A_1, A_2). The corresponding trace series T_{A_1,A_2} can be expressed as*

$$T_{A_1,A_2}(z_1, z_2) = \frac{\delta_1(z_1, z_2) + 2\delta_2(z_1, z_2) + \ldots + n\delta_n(z_1, z_2)}{\Delta(z_1, z_2)}. \tag{2.21}$$

Proof Consider the linear system defined on $\mathbf{C}[\alpha, \beta]$, the ring of the polynomials in the indeterminates α and β with coefficients in $\mathbf{C}$:

$$\begin{aligned} x_{i+1} &= Fx_i + gu_i, \\ y_i &= Hx_i, \end{aligned}$$

with

$$F = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ & & & \cdots & \\ & & & \cdots & 1 \\ \delta_n(\alpha, \beta) & \delta_{n-1}(\alpha, \beta) & \delta_{n-2}(\alpha, \beta) & \cdots & \delta_1(\alpha, \beta) \end{bmatrix}, \; g = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix},$$

$$H = [\quad 0 \quad 0 \quad 0 \quad \cdots \quad 1 \quad].$$

By assuming $x_0 = 0$ and

$$u_i = \begin{cases} (i+1)\delta_{i+1}(\alpha,\beta) & \text{i = 0, 1, \ldots} \\ 0 & \text{otherwise,} \end{cases}$$

it's a matter of direct computation to check that the corresponding output sequence is given by $y_i = \tau_i(\alpha,\beta)$, $i = 1,2,...$.
As the system transfer function is $H(I - zF)^{-1}gz = z/(1 - \sum_{h=1}^{n} \delta_h(\alpha,\beta)z^h)$, the input $U(z) = u_0 + u_1 z + \ldots + u_{n-1}z^{n-1}$ produces the output

$$Y(z) = \sum_{i=0}^{\infty} \tau_i(\alpha,\beta)z^i = \frac{\sum_{h=1}^{n} h\delta_h(\alpha,\beta)z^h}{1 - \sum_{h=1}^{n} \delta_h(\alpha,\beta)z^h} \tag{2.22}$$

So, letting $z_1 = \alpha z$ and $z_2 = \beta z$, one gets (2.21) ■

Representation (2.21) of T_{A_1,A_2} is not necessarily irreducible, but its special structure makes it quite easy to obtain an irreducible one. To this purpose, consider the injective homomorphism

$$\phi : \mathbf{C}[\alpha,\beta] \to \mathbf{C}[\alpha,\beta,z] : \sum_{i=0}^{n} \delta_i(\alpha,\beta) \mapsto \sum_{i=0}^{n} \delta_i(\alpha,\beta)z^i,$$

where, as usual, $\delta_i(\alpha,\beta)$ denotes an homogeneus polynomial of degree i, and introduce the following derivation map

$$D_z : \mathbf{C}[\alpha,\beta,z] \to \mathbf{C}[\alpha,\beta,z] \ : \sum_{i=0}^{m} p_i(\alpha,\beta)z^i \mapsto \sum_{i=0}^{m} ip_i(\alpha,\beta)z^i. \tag{2.23}$$

Clearly (2.22) can be rewritten as

$$Y(z) = \frac{D_z(\phi(\Delta(\alpha,\beta)))}{\phi(\Delta(\alpha,\beta))}. \tag{2.24}$$

By assuming that Δ factorizes as $\Delta(z_1,z_2) = \prod_{i=1}^{t} \Delta_i(z_1,z_2)^{\nu_i}$, with Δ_i irreducible distinct factors, $\Delta_i(0,0) = 1$, $i = 1,2,...,t$, one easily gets

$$Y(z) = \sum_{i=1}^{t} \nu_i \frac{D_z(\phi(\Delta_i(\alpha,\beta)))}{\phi(\Delta_i(\alpha,\beta))}. \tag{2.25}$$

Thus, letting $z_1 = \alpha z$ and $z_2 = \beta z$, we have proved the following

Proposition 2.6 *Let $\Delta(z_1,z_2) = \prod_{i=1}^{t} \Delta_i(z_1,z_2)^{\nu_i}$ be a factorization of Δ, with $\Delta_i(z_1,z_2) = 1 - \sum_{j=1}^{r_i} \delta_j^{(i)}(z_1,z_2)$ irreducible distinct polynomials, $i = 1,2,...,t$. For every matrix pair (A_1,A_2) such that $\Delta_{A_1,A_2}(z_1,z_2) = \Delta(z_1,z_2)$, the corresponding trace series is given by*

$$T_{A_1,A_2}(z_1,z_2) = \sum_{i=1}^{t} \nu_i \frac{\sum_{j=1}^{r_i} j \ \delta_j^{(i)}(z_1,z_2)}{1 - \sum_{j=1}^{r_i} \delta_j^{(i)}(z_1,z_2)} \ ■ \tag{2.26}$$

Equation (2.26) expresses the trace series $T_{A_1,A_2}(z_1, z_2)$ as a partial fraction expansion, whose i-th term is the trace series of the irreducible factor $\Delta_i(z_1, z_2)$, weighted with the corresponding multiplicity ν_i. Thus the denominator of every irreducible rational function which represents a trace series factorizes into distinct irreducible factors. On the other hand, once an irreducible rational function $T(z_1, z_2)$ has been given, (2.26) suggests a quick way to check whether $T(z_1, z_2)$ can be expanded into a trace series.

3 Pairs of matrices with property L

In the next two sections we shall focus specifically on matrix pairs endowed with property L and property P.

Pairs with property L occur quite frequently in the applications: indeed, the important classes of finite memory and separable 2D systems we are going to discuss in Sec.5 are described by pairs with property L. A pair of n by n matrices, (A_1, A_2), with entries in $\mathbf{C}$, is said to have property L if the eigenvalues of A_1 and A_2 can be ordered into two n-tuples

$$\Lambda(A_1) = (\lambda_1, \lambda_2, \ldots, \lambda_n) \quad \text{and} \quad \Lambda(A_2) = (\mu_1, \mu_2, \ldots, \mu_n) \tag{3.1}$$

such that, for all α, β in $\mathbf{C}$, the spectrum of $\Lambda(\alpha A_1 + \beta A_2)$ is given by

$$\Lambda(\alpha A_1 + \beta A_2) = (\alpha\lambda_1 + \beta\mu_1, \ldots, \alpha\lambda_n + \beta\mu_n). \tag{3.2}$$

It's not difficult to show that property L corresponds to the possibility of factorizing the characteristic polynomial into linear terms [11,12]. Thus each term of the partial fraction expansion of T_{A_1,A_2} has the very special structure $(\lambda z_1 + \mu z_2)/(1 - \lambda z_1 - \mu z_2)$, which entails far reaching consequences on the possibility of characterizing property L by using the Hankel matrix theory.

Proposition 3.1 *Let A_1, A_2 be in $\mathbf{C}^{n\times n}$ and consider the orderings of their spectra given in* (3.1). *The following statements are equivalent:*

- *L)* (A_1, A_2) *has property L (w.r.t. the orderings* (3.1)*);*
- *L_1)* $\Delta_{A_1,A_2}(z_1, z_2) = \prod_{i=1}^n (1 - \lambda_i z_1 - \mu_i z_2)$*;*
- *L_2) for all* $\alpha, \beta \in \mathbf{C}$ *and* $k \in \mathbf{N}$, $\operatorname{tr}(\alpha A_1 + \beta A_2)^k = \sum_{i=1}^n (\alpha\lambda_i + \beta\mu_i)^k$*;*
- *L_3) for every* $(h, k) \in \mathbf{N} \times \mathbf{N}$, $\operatorname{tr}(A_1{}^h\sqcup{}^k A_2) = \binom{h+k}{h} \sum_{i=1}^n \lambda_i^h \mu_i^k$*;*
- *L_4)* $T_{A_1,A_2} = \sum_{h+k>0} \operatorname{tr}(A_1{}^h\sqcup{}^k A_2) z_1^h z_2^k = \sum_{i=1}^n \dfrac{\lambda_i z_1 + \mu_i z_2}{1 - \lambda_i z_1 - \mu_i z_2}$.

PROOF Clearly matrices $\bar{A}_1 = \operatorname{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ and $\bar{A}_2 = \operatorname{diag}\{\mu_1, \mu_2, \ldots, \mu_n\}$ fulfill all conditions $L) \div L_4)$ of the proposition. Any other pair (A_1, A_2), of the same dimension, with property L w.r.t. the orderings (3.1), satisfies

$$\Lambda(\alpha A_1 + \beta A_2) = \Lambda(\alpha \hat{A}_1 + \beta \hat{A}_2), \tag{3.3}$$

which corresponds to ii) of Corollary 2.4. Therefore all equivalent statements in Proposition 2.3 hold true, in particular, property L of (A_1, A_2) is equivalent to any one of the followings:

$L_1)$ $\Delta_{A_1,A_2}(z_1,z_2) = \Delta_{\hat{A}_1,\hat{A}_2}(z_1,z_2) = \prod_{i=1}^n (1-\lambda_i z_1 - \mu_i z_2)$

$L_2)$ $\mathrm{tr}(\alpha A_1 + \beta A_2)^k = \mathrm{tr}(\alpha \hat{A}_1 + \beta \hat{A}_2)^k = \sum_{i=1}^n (\alpha\lambda_i + \beta\mu_i)^k$

$L_3)$ $\mathrm{tr}(A_1{}^h\sqcup{}^k A_2) = \mathrm{tr}(\hat{A}_1{}^h\sqcup{}^k \hat{A}_2) = \binom{h+k}{k}\sum_{i+1}^n \lambda_i^h \mu_i^k$

$L_4)$ $\sum_{(h,k)\neq(0,0)} \mathrm{tr}(A_1{}^h\sqcup{}^k A_2) z_1^h z_2^k$
$= \sum_{(h,k)\neq(0,0)} \mathrm{tr}(\hat{A}_1{}^h\sqcup{}^k \hat{A}_2) z_1^h z_2^k = \sum_{i=1}^n \dfrac{\lambda_i z_1 + \mu_i z_2}{1-\lambda_i z_1 - \mu_i z_2}$ ■

Conditions $L) \div L_4)$ do not provide direct methods to check, in a finite number of steps, whether a given pair (A_1, A_2) is endowed with property L. To reach this goal, we shall analyse the rank of suitable matrices associated with the power series

$$R_{A_1,A_2}(z_1,z_2) := \sum_{i,j=0}^{\infty} \mathrm{tr}(A_1{}^i\sqcup{}^j A_2)\binom{i+j}{i}^{-1} z_1^i z_2^j. \tag{3.4}$$

Let $\mathbf{C}[[z_1,z_2]]$ be the ring of formal power series in the commuting variables z_1, z_2, and denote by

$$s := \sum_{h,k}\langle s, z_1^h z_2^k\rangle z_1^h z_2^k$$

a generic element of the ring. We associate with s the infinite Hankel matrix [2]

$$\mathcal{H}(s) := \begin{bmatrix} \langle s,1\rangle & \langle s,z_1\rangle & \langle s,z_2\rangle & \langle s,z_1^2\rangle & \langle s,z_1z_2\rangle & \langle s,z_2^2\rangle & \dots \\ \langle s,z_1\rangle & \langle s,z_1^2\rangle & \langle s,z_1z_2\rangle & \langle s,z_1^3\rangle & \langle s,z_1^2z_2\rangle & \dots & \dots \\ \langle s,z_2\rangle & \langle s,z_1z_2\rangle & \langle s,z_2^2\rangle & \dots & \dots & \dots & \dots \\ \langle s,z_1^2\rangle & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix},$$

whose rows and columns take indices in the multiplicative monoid of the (commutative) terms $\mathcal{T} := \{z_1^i z_2^j : i,j \in \mathbf{N}\}$.
For all $M', M'' \in \mathbf{N}$, we shall denote by $\mathcal{H}_{M'\times M''}(s)$ the submatrix, appearing in the upper left corner of $\mathcal{H}(s)$, whose rows (columns) are indexed by the terms of homogeneus degree not greater than M' (M'').

When rank $\mathcal{H}(s)$ is $\nu < \infty$, we can choose ν rows and ν columns, indexed by the terms $r_1, r_2, ..., r_\nu$ and $c_1, c_2, ..., c_\nu$, respectively, so that submatrix

$$N_0 := [\langle s, r_i c_j\rangle] \tag{3.5}$$

is nonsingular. Thus, for all terms $c \in \mathcal{T}$, the ν-tuple $[\langle s, r_1 c\rangle ... \langle s, r_\nu c\rangle]^T$ belongs to the range space of N_0, i.e. there exists a (unique) vector $\mathbf{x}(c) \in \mathbf{C}^\nu$ such that

$$\begin{bmatrix} \langle s, r_1 c\rangle \\ \vdots \\ \langle s, r_\nu c\rangle \end{bmatrix} = N_0\ \mathbf{x}(c).$$

Moreover, the rank assumption on $\mathcal{H}(s)$ implies

$$\langle s, rc\rangle = \sum_{j=1}^{\nu} x_j(c)\langle s, rc_j\rangle, \qquad \forall r \in \mathcal{T}. \tag{3.6}$$

We therefore have, for all $r, c \in \mathcal{T}$,

$$\langle s, rc\rangle = \Big[\langle s, rc_1\rangle, ..., \langle s, rc_\nu\rangle\Big] N_0^{-1} \begin{bmatrix} \langle s, r_1 c\rangle \\ \vdots \\ \langle s, r_\nu c\rangle \end{bmatrix}. \tag{3.7}$$

We are now in a position to state the following

Proposition 3.2 *Let A_1, A_2 be in $\mathbf{C}^{n\times n}$. (A_1, A_2) has property L if and only if*

$$\bar{n} := \text{rank } \mathcal{H}_{(n-1)\times(n-1)}(R_{A_1,A_2}) = \text{rank } \mathcal{H}_{n\times n}(R_{A_1,A_2}) \le n. \tag{3.8}$$

Proof For sake of brevity, within the proof we shall drop in R_{A_1,A_2} subscripts A_1 and A_2. Assume first that (A_1, A_2) has property L. Then, by Proposition 3.1,

$$\text{tr}(A_1{}^i \sqcup\!\sqcup{}^j A_2) = \binom{i+j}{i} \sum_{h=1}^{n} \lambda_h^i \mu_h^j$$

and, therefore,

$$R = \sum_{i,j=0}^{\infty} \Big(\sum_{h=1}^{n} \lambda_h^i \mu_h^j\Big) z_1^i z_2^j.$$

Since the Hankel matrix $\mathcal{H}(R)$ factorizes as

$$\mathcal{H}(R) = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_n \\ \mu_1 & \mu_2 & \dots & \mu_n \\ \lambda_1^2 & \lambda_2^2 & \dots & \lambda_n^2 \\ \lambda_1\mu_1 & \lambda_2\mu_2 & \dots & \lambda_n\mu_n \\ \mu_1^2 & \mu_2^2 & \dots & \mu_n^2 \\ & & \dots & \end{bmatrix} \begin{bmatrix} 1 & \lambda_1 & \mu_1 & \lambda_1^2 & \lambda_1\mu_1 & \mu_1^2 & \dots \\ 1 & \lambda_2 & \mu_2 & \lambda_2^2 & \lambda_2\mu_2 & \mu_2^2 & \dots \\ & & & \dots & & & \\ 1 & \lambda_n & \mu_n & \lambda_n^2 & \lambda_n\mu_n & \mu_n^2 & \dots \end{bmatrix}, \tag{3.9}$$

clearly rank $\mathcal{H}(R) \le n$.
Finally, apply the 2D Cayley-Hamilton theorem to the pair of matrices $Q_1 := \text{diag}\{\lambda_1, \lambda_2, ..., \lambda_n\}$ and $Q_2 := \text{diag}\{\mu_1, \mu_2, ..., \mu_n\}$. Since the matrices

$$Q_1{}^h \sqcup\!\sqcup{}^k Q_2 = \binom{h+k}{k} \text{diag}\{\lambda_1^h \mu_1^k, \lambda_2^h \mu_2^k, ..., \lambda_n^h \mu_n^k\}, \qquad h+k \ge n,$$

are linear combinations of

$$Q_1{}^i \sqcup\!\sqcup{}^j Q_2 = \binom{i+j}{i} \text{diag}\{\lambda_1^i \mu_1^j, \lambda_2^i \mu_2^j, ..., \lambda_n^i \mu_n^j\}, \quad i+j < n,$$

all rows (columns) in the first (second) factor of (3.9) that include homogeneous terms of degree greater than or equal to n, linearly depend on the previous ones. Hence rank $\mathcal{H}_{(n-1)\times(n-1)}(R) =$ rank $\mathcal{H}(R)$ and (3.8) holds.

To prove the converse, assume that (3.8) holds, and select in $\mathcal{H}_{(n-1)\times(n-1)}(R)$, $\bar{n}$ rows and $\bar{n}$ columns indexed by terms $r_1, r_2, ..., r_{\bar{n}}$ and $c_1, c_2, ..., c_{\bar{n}}$ respectively, such that the $\bar{n} \times \bar{n}$ submatrix $N_0 := [\langle R, r_i c_j\rangle]$ is nonsingular. The following matrices

$$M_1 := N_0^{-1}[\langle R, r_i z_1 c_j\rangle] \qquad M_2 := N_0^{-1}[\langle R, r_i z_2 c_j\rangle] \tag{3.10}$$

commute. Indeed, $r_i z_1 z_2 c_j$ are terms of degree not greater than $2n$, and consequently assumption (3.8) allows to resort to (3.7), which gives

$$\langle R, r_i z_1 z_2 c_j\rangle = \Big[\langle R, r_i z_1 c_1\rangle, ..., \langle R, r_i z_1 c_{\bar{n}}\rangle\Big] N_0^{-1} \begin{bmatrix} \langle R, r_1 z_2 c_j\rangle \\ \vdots \\ \langle R, r_{\bar{n}} z_2 c_j\rangle \end{bmatrix}, \ \forall\ i, j \in \{1, 2, ..., \bar{n}\}.$$

This implies

$$M_1M_2 - M_2M_1 = N_0^{-1}\Big[[\langle R, r_i z_1 c_j\rangle] N_0^{-1}[\langle R, r_i z_2 c_j\rangle] - [\langle R, r_i z_2 c_j\rangle] N_0^{-1}[\langle R, r_i z_1 c_j\rangle]\Big] = 0.$$

Introduce next the matrices

$$H := \Big[\langle R, c_1\rangle\ \langle R, c_2\rangle\ ...\ \langle R, c_{\bar{n}}\rangle\Big] \quad \text{and} \quad G := N_0^{-1} \begin{bmatrix} \langle R, r_1\rangle \\ \vdots \\ \langle R, r_{\bar{n}}\rangle \end{bmatrix}, \tag{3.11}$$

and consider the commutative power series

$$S := H(I - M_1 z_1)^{-1}(I - M_2 z_2)^{-1} G = \sum_{i,j=0}^{\infty} H M_1^i M_2^j G\ z_1^i z_2^j. \tag{3.12}$$

By resorting again to (3.7), it's easy to check that $\langle R, z_1^i z_2^j\rangle = \langle S, z_1^i z_2^j\rangle$, for $i + j \leq 2n$, and therefore $\mathcal{H}_{n\times n}(R) = \mathcal{H}_{n\times n}(S)$. As M_1 and M_2 commute, they are endowed with property L, and the power series

$$\bar{S} := \sum_{i,j=0}^{\infty} H(M_1{}^i \sqcup\!\sqcup{}^j M_2) G\ z_1^i z_2^j = H(I - M_1 z_1 - M_2 z_2)^{-1} G.$$

can be represented as a rational function of the following form

$$\bar{S} := \frac{q(z_1, z_2)}{\prod_{i=1}^{\bar{n}}(1 - \lambda_i z_1 - \mu_i z_2)}, \qquad \deg\, q < \bar{n} \leq n. \tag{3.13}$$

Note that $\bar{S}$ satisfies $\langle \bar{S}, 1\rangle = \langle S, 1\rangle = \langle R, 1\rangle = n$, and, for $0 < i + j \leq 2n$

$$\langle \bar{S}, z_1^i z_2^j\rangle = \binom{i+j}{i}\langle S, z_1^i z_2^j\rangle = \binom{i+j}{i}\langle R, z_1^i z_2^j\rangle = \langle T_{A_1,A_2}, z_1^i z_2^j\rangle. \tag{3.14}$$

On the other hand, being the trace series of an $n \times n$ matrix pair, T_{A_1,A_2} can be expressed as in (2.15) and, consequently we have

$$T_{A_1,A_2} + n = \frac{p(z_1, z_2)}{\Delta_{A_1,A_2}(z_1, z_2)}, \quad \deg p \leq n.$$

Therefore, in the rational function $\bar{S} - T_{A_1,A_2} - n$ the denominator has degree not greater than $2n$ and nonzero constant term, while the numerator has degree not greater than $2n$. As all the coefficients of the power series expansion of $\bar{S} - T_{A_1,A_2} - n$, namely $\langle \bar{S} - T_{A_1,A_2} - n, z_1^i z_2^j \rangle$, are zero for $i + j \leq 2n$, then $\bar{S} = T_{A_1,A_2} + n$.

It's clear now from (3.13) that the denominator of an irreducible representation of T_{A_1,A_2} factorizes into linear factors. Therefore, by Proposition 3.1, (A_1, A_2) has property L ■

4 Pairs of matrices with property P

Given the alphabet $\Xi = \{\xi_1, \xi_2\}$, the free monoid Ξ^* with base Ξ is the set of all words

$$w = \xi_{i_1} \xi_{i_2} \cdots \xi_{i_m}, \qquad m \in \mathbf{N}, \ \xi_{i_h} \in \Xi.$$

The integer m is called the length of the word w and denoted by $|w|$, while $|w|_i$ represents the number of occurencies of ξ_i in w, $i = 1, 2$. If $v = \xi_{j_1} \xi_{j_2} \cdots \xi_{j_p}$ is another element of Ξ^*, the product is defined by concatenation

$$wv = \xi_{i_1} \xi_{i_2} \cdots \xi_{i_m} \xi_{j_1} \xi_{j_2} \cdots \xi_{j_p}.$$

This produces a monoid with $1 = \emptyset$, the empty word, as unit element. Clearly, $|wv| = |v| + |w|$ and $|1| = 0$.
$\mathbf{C}\langle \xi_1, \xi_2 \rangle$ and $\mathbf{C}\langle\langle \xi_1, \xi_2 \rangle\rangle$ are the algebras of polynomials and formal power series respectively in the noncommuting indeterminates ξ_1 and ξ_2. For each pair of matrices A_1, A_2 in $\mathbf{C}^{n \times n}$, the map ψ defined on $\{1, \xi_1, \xi_2\}$ by the assignments $\psi(1) = I_n$ and $\psi(\xi_i) = A_i$, $i = 1, 2$, uniquely extends to an algebra morphism of $\mathbf{C}\langle \xi_1, \xi_2 \rangle$ into $\mathbf{C}^{n \times n}$. The ψ-image of a polynomial $\mathcal{P}(\xi_1, \xi_2) \in \mathbf{C}\langle \xi_1, \xi_2 \rangle$ is denoted by $\mathcal{P}(A_1, A_2)$.

A pair of $n \times n$ matrices (A_1, A_2) with elements in $\mathbf{C}$ is said to have property P if the eigenvalues of A_1 and A_2 can be ordered into two n-tuples

$$\Lambda(A_1) = (\lambda_1, \lambda_2, \ldots \lambda_n), \qquad \Lambda(A_2) = (\mu_1, \mu_2, \ldots \mu_n), \tag{4.1}$$

such that, for every polynomial $\mathcal{P}(\xi_1, \xi_2) \in \mathbf{C}\langle \xi_1, \xi_2 \rangle$,

$$\Lambda(\mathcal{P}(A_1, A_2)) = (\mathcal{P}(\lambda_1, \mu_1), \mathcal{P}(\lambda_2, \mu_2), \ldots, \mathcal{P}(\lambda_n, \mu_n)). \tag{4.2}$$

It's easy to check that property P implies property L, while examples can be given [11,16] showing that the converse is not true.

2D systems (1.1) whose transition matrices A_1, A_2 have property P are endowed with several interesting properties. Indeed, property P is equivalent to simultaneous triangularizability, a feature which allows a good insight into the geometric structure of the free state evolution. In particular, it implies that there exists a maximal chain of $\{A_1, A_2\}$-invariant subspaces of the local state space X

$$\{0\} = X_0 < X_1 < X_2 < \cdots < X_n = X$$

with $\dim(X_i) = i,\ i = 0, 1, 2, \dots n$.
When the local states $x(-\ell, \ell)$ of the initial global state $\mathcal{X}_0 = \sum_{\ell=-\infty}^{+\infty} x(-\ell, \ell) z_1^{-\ell} z_2^{\ell}$ are in X_i, all local states $x(h, k),\ h + k \geq 0$ are in X_i too. Correspondingly, systems (1.1) can be viewed as cascades of 2D systems of dimension one.

Moreover, systems with property P constitute a class of 2D systems large enough for realizing all transfer functions $p(z_1, z_2)/q(z_1, z_2)$ with denominators of the form $q(z_1, z_2) = \prod_j (1 - \lambda_j z_1 - \mu_j z_2)$, and in particular all transfer functions with separable denominators [1]. It should be stressed that the same is not true if we consider only commutative 2D systems, i.e. systems (1.1) which satisfy the (stronger) constraint $A_1 A_2 - A_2 A_1 = 0$.

As a consequence of Proposition 3.1, matrix pairs endowed with property L can be equivalently described as those, whose characteristic polynomials factorize into a product of linear terms. This class of polynomials, however, corresponds also to matrix pairs with property P; so there is no possibility of finding an equivalent description of property P which relies only on the characteristic polynomial. Appropriate tools turn out to be certain noncommutative polynomials [12] and power series associated with the pair, as well as the corresponding Hankel matrices.

Proposition 4.1 *Let A_1, A_2 be $n \times n$ matrices with entries in $\mathbf{C}$, and consider the orderings of their spectra given in* (4.1)*. The following statements are equivalent:*

- $P)$ *(A_1, A_2) has property P w.r.t. the orderings* (4.1)*;*
- $P_1)$ *for any $w \in \Xi^*$, with $|w|_1 = h$ and $|w|_2 = k$,*

$$\operatorname{tr}(w(A_1, A_2)) = \sum_{i=1}^{n} \lambda_i^h \mu_i^k; \tag{4.3}$$

- $P_2)$ *the noncommutative power series, whose coefficients are the traces of the matrices $w(A_1, A_2)$,*

$$\mathcal{N} = \sum_{w\ \in\ \Xi^*} \operatorname{tr}(w(A_1, A_2)) w, \tag{4.4}$$

can be represented as $\mathcal{N} = \sum_{i=1}^{n} (1 - \lambda_i \xi_1 - \mu_i \xi_2)^{-1}$, and hence is recognizable [2]*;*

- $P_3)$ *for any $w \in \Xi^*$, with $|w|_1 = h$ and $|w|_2 = k$,*

$$\det(zI - w(A_1, A_2)) = \prod_{i=1}^{n} (z - \lambda_i^h \mu_i^k). \tag{4.5}$$

Proof $P) \Rightarrow P_1)$ If $|w|_1 = h$ and $|w|_2 = k$, the definition of property P directly implies $\Lambda(w(A_1, A_2)) = (\lambda_1^h \mu_1^k, \dots, \lambda_n^h \mu_n^k)$, and therefore (4.3) holds.

$P_1) \Rightarrow P)$ Extend the monoid morphisms $\phi_i : \Xi^* \to \mathbf{C} :\ w \mapsto \lambda_i^{|w|_1}\mu_i^{|w|_2} = w(\lambda_i, \mu_i)$, $i = 1, 2 \ldots n$, to the algebra $\mathbf{C}\langle\xi_1, \xi_2\rangle$, letting $\phi_i(\mathcal{P}) = \mathcal{P}(\lambda_i, \mu_i),\ i = 1, 2 \ldots n$, for all $\mathcal{P}(\xi_1, \xi_2) \in \mathbf{C}\langle\xi_1, \xi_2\rangle$. Then we have

$$\phi_i(\mathcal{P}^h) = \left(\phi_i(\mathcal{P})\right)^h = \left(\mathcal{P}(\lambda_i, \mu_i)\right)^h. \tag{4.6}$$

From assumption $P_1)$ we deduce that $\operatorname{tr}(w(A_1, A_2)) = \sum_{i=1}^n \phi_i(w)$, and hence, by the linearity of the trace operator,

$$\operatorname{tr}(\mathcal{P}(A_1, A_2)) = \sum_{i=1}^n \phi_i(\mathcal{P}) = \sum_{i=1}^n \mathcal{P}(\lambda_i, \mu_i).$$

Using (4.6), for all $h \in \mathbf{N}_+$

$$\operatorname{tr}(\mathcal{P}(A_1, A_2))^h = \sum_{i=1}^n \phi_i(\mathcal{P}^h) = \sum_{i=1}^n \mathcal{P}(\lambda_i, \mu_i)^h, \tag{4.7}$$

which gives $\Lambda(\mathcal{P}(A_1, A_2)) = (\mathcal{P}(\lambda_1, \mu_1), \ldots, \mathcal{P}(\lambda_n, \mu_n))$.

$P_1) \Leftrightarrow P_2)$ Assuming P_1, we may write

$$\mathcal{N} = \sum_{w\ \in\ \Xi^*} \operatorname{tr}(w(A_1, A_2)w) = \sum_{i=1}^n \sum_{w\ \in\ \Xi^*} \lambda_i^{|w|_1}\mu_i^{|w|_2} w.$$

On the other hand, we obtain

$$\begin{array}{rcl} \sum_{i=1}^n (1 - \lambda_i\xi_1 - \mu_i\xi_2)^{-1} & = & \sum_{i=1}^n \sum_{j=0}^{+\infty} (\lambda_i\xi_1 + \mu_i\xi_2)^j \\ & = & \sum_{i=1}^n \sum_{j=0}^{+\infty} \sum_{\substack{w\ \in\ \Xi^* \\ |w|_1+|w|_2=j}} \lambda_i^{|w|_1}\mu_i^{|w|_2} w \\ & = & \sum_{i=1}^n \sum_{w\ \in\ \Xi^*} \lambda_i^{|w|_1}\mu_i^{|w|_2} w \end{array}$$

which proves (4.6). The converse can be shown in the same way.

$P_1) \Rightarrow P_3)$ Given $w \in \Xi^*$, for all $h \in \mathbf{N}$ we have

$$\operatorname{tr}(w(A_1, A_2))^h = \sum_{i=1}^n \lambda_i^{h|w|_1}\mu^{h|w|_2} = \sum_{i=1}^n (\lambda_i^{|w|_1}\mu_i^{|w|_2})^h.$$

Thus $(\lambda_1^{|w|_1}\mu_1^{|w|_2}, \ldots, \lambda_n^{|w|_1}\mu_n^{|w|_2})$ is the spectrum of $w(A_1, A_2)$, which proves (4.5).

$P_3) \Rightarrow P_1)$ Obvious $\blacksquare$

Remark As a consequence of $P_1)$, property P can be equivalently stated referring only to the words of the free monoid Ξ^*, instead of the whole algebra $\mathbf{C}\langle\xi_1, \xi_2\rangle$. Indeed, (A_1, A_2) has property P if and only if, for all $w \in \Xi^*$ we have

$$\Lambda\Big(w(A_1, A_2)\Big) = (\lambda_1^{|w|_1}\mu_1^{|w|_2}, \ldots, \lambda_n^{|w|_1}\mu_n^{|w|_2}),$$

where $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ and $(\mu_1, \mu_2, \ldots, \mu_n)$ are the spectra of A_1 and A_2, suitably ordered.

As for property L, we aim now to provide an effective method for testing property P, that depends on the study of the noncommutative power series $\mathcal{N}$ and the associated Hankel matrix [2,15]. By the Hankel matrix of $\mathcal{N}$ we mean the infinite matrix $\mathcal{H}(\mathcal{N})$, whose rows and columns are indexed by the words of Ξ^*, and whose element with indexes u and v is equal to $\langle \mathcal{N}, uv \rangle$.

It will be convenient to order the words in Ξ^*, and consequently the row and column indexes in $\mathcal{H}(\mathcal{N})$, according to their length, while the lexicographical order will be adopted for words of the same length. For all $M', M'' \in \mathbf{N}$, we shall denote by $\mathcal{H}_{M' \times M''}(\mathcal{N})$ the submatrix appearing in the upper left corner of $\mathcal{H}(\mathcal{N})$, whose rows (columns) are indexed by words of length not greater than M' (M'').

Lemma 4.2 *Let A_1, A_2 be $n \times n$ matrices with entries in $\mathbf{C}$. Then*

$$\operatorname{rank} \mathcal{H}_{(n^2-1)\times(n^2-1)}(\mathcal{N}) = \operatorname{rank} \mathcal{H}(\mathcal{N}) \leq n^2 \tag{4.8}$$

PROOF For all $w \in \Xi^*$, we have

$$\operatorname{tr}(w(A_1, A_2)) = [\,\mathbf{e}_1^T \quad \dots \quad \mathbf{e}_n^T\,] \operatorname{diag}\{w(A_1, A_2), \dots, w(A_1, A_2)\} \begin{bmatrix} \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_n \end{bmatrix}, \tag{4.9}$$

where $\mathbf{e}_1, \dots, \mathbf{e}_n$ are the vectors of the canonical basis of $\mathbf{C}^n$. It is clear that $\mathcal{H}(\mathcal{N})$ can be expressed as $\mathcal{H}(\mathcal{N}) = \mathcal{O}\mathcal{R}$, where $\mathcal{O}$ is the $\infty \times n^2$ matrix whose row of index $v \in \Xi^*$ is given by

$$[\,\mathbf{e}_1^T \quad \dots \quad \mathbf{e}_n^T\,] \ \operatorname{diag} \ \{v(A_1, A_2), \dots, v(A_1, A_2)\} \tag{4.10}$$

and, similarly, $\mathcal{R}$ is the $n^2 \times \infty$ matrix whose column of index $w \in \Xi^*$ is given by

$$\operatorname{diag} \ \{w(A_1, A_2), \dots, w(A_1, A_2)\} \begin{bmatrix} \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_n \end{bmatrix}. \tag{4.11}$$

This shows that $\operatorname{rank}\mathcal{H}(\mathcal{N}) \leq n^2$.

To complete the proof, suppose that all rows of $\mathcal{O}$ indexed by words of length ν linearly depend on the rows indexed in $\mathcal{I} := \{w \in \Xi^*, |w| < \nu\}$. We deduce that any row of index $v = u\xi_i$, $|u| = \nu$, also depends on the words indexed in $\mathcal{I}$, because

$$\begin{aligned}
&[\mathbf{e}_1^T \ \dots \ \mathbf{e}_n^T] \ \operatorname{diag}\{v(A_1, A_2), \dots, v(A_1, A_2)\} \\
&\quad = [\mathbf{e}_1^T \ \dots \ \mathbf{e}_n^T] \ \operatorname{diag}\{u(A_1, A_2), \dots, u(A_1, A_2)\} \ \operatorname{diag}\{A_i, \dots, A_i\} \\
&\quad = \textstyle\sum_{w \in \mathcal{I}} \alpha_w [\mathbf{e}_1^T \ \dots \ \mathbf{e}_n^T] \ \operatorname{diag}\{w(A_1, A_2)A_i, \dots, w(A_1, A_2)A_i\} \\
&\quad = \textstyle\sum_{w \in \mathcal{I}} \beta_w [\mathbf{e}_1^T \ \dots \ \mathbf{e}_n^T] \ \operatorname{diag}\{w(A_1, A_2), \dots, w(A_1, A_2)\}.
\end{aligned}$$

An easy inductive argument proves that all rows of $\mathcal{O}$ linearly depend on those indexed in $\mathcal{I}$. Moreover, as $\operatorname{rank}\mathcal{O} \leq n^2$, it's clear that in the definition of $\mathcal{I}$ we can assume $\nu = n^2$. The same reasoning applies to the rows of $\mathcal{R}$, showing that

$$\operatorname{rank} \mathcal{H}_{(n^2-1)\times(n^2-1)}(\mathcal{N}) = \operatorname{rank} \mathcal{H}(\mathcal{N}) \quad \blacksquare$$

Proposition 4.3 *Let A_1, A_2 be $n \times n$ matrices with entries in $\mathbf{C}$ and consider the associated noncommutative power series $\mathcal{N} = \sum_{w \in \Xi^*} \mathrm{tr}(w(A_1, A_2))w$. The following statements are equivalent:*

i) (A_1, A_2) *has property P;*

ii) rank $\mathcal{H}_{(n^2-1)\times(n^2-1)}(\mathcal{N}) = \bar{n} \leq n$ *and, for all pairs of words $w, \bar{w}$ with length not greater than $2\bar{n}$,*

$$|w|_i = |\bar{w}|_i, \quad i = 1, 2 \quad \Rightarrow \quad \mathrm{tr}(w(A_1, A_2)) = \mathrm{tr}(\bar{w}(A_1, A_2)); \tag{4.12}$$

iii) (4.12) *holds for all pairs of words $w, \bar{w}$ with length not greater than $2n^2$.*

Proof *i)* $\Rightarrow$ *ii)* By Proposition 4.1, property P implies that

$$\mathrm{tr}(w(A_1, A_2)) = \sum_{i=1}^{n} \lambda_i^{|w|_1} \mu_i^{|w|_2}, \quad \forall w \in \Xi^*, \tag{4.13}$$

for suitable orderings of the eigenvalues of A_1 and A_2. This immediately proves (4.12). Moreover, $\mathcal{H}(\mathcal{N})$ can be expressed as $\mathcal{H}(\mathcal{N}) = \mathcal{O}\mathcal{R}$, where $\mathcal{O}$ is the $\infty \times n$ matrix whose row of index $v \in \Xi^*$ is given by $[\, v(\lambda_1, \mu_1), v(\lambda_2, \mu_2), \ldots, v(\lambda_n, \mu_n)\,]$, and $\mathcal{R} = \mathcal{O}^T$.
We deduce that rank $\mathcal{H}(\mathcal{N}) \leq$ rank $\mathcal{O} \leq n$ and, therefore rank $\mathcal{H}_{(n^2-1)\times(n^2-1)}(\mathcal{N}) \leq n$.

ii) $\Rightarrow$ *iii)* By the previous lemma, rank $\mathcal{H}(\mathcal{N}) = \bar{n}$. So, there exist [2,15] $M_1, M_2 \in \mathbf{C}^{\bar{n}\times\bar{n}}, H \in \mathbf{C}^{1\times\bar{n}}$ and $G \in \mathbf{C}^{\bar{n}\times 1}$ such that $\langle \mathcal{N}, w\rangle = Hw(M_1, M_2)G, \quad \forall w \in \Xi^*$, and $\mathcal{H}(\mathcal{N})$ can be expressed as $\mathcal{H}(\mathcal{N}) = \mathcal{O}\mathcal{R}$, where $\mathcal{O}$ is the $\infty \times \bar{n}$ matrix whose row indexed by $v \in \Xi^*$ is $Hv(M_1, M_2)$, and $\mathcal{R}$ is the $\bar{n} \times \infty$ matrix whose column indexed by $w \in \Xi^*$ is $w(M_1, M_2)G$.
By the same argument used in the proof of Lemma 4.2, there exist $2\bar{n}$ words $r_1, r_2, \ldots r_{\bar{n}}$ and $c_1, c_2, \ldots c_{\bar{n}}$, of length less than $\bar{n}$, such that both the $\bar{n} \times \bar{n}$ matrices $\mathcal{O}_{\bar{n}}$, whose rows are $Hr_i(M_1, M_2)$ and $\mathcal{R}_{\bar{n}}$, whose columns are $c_j(M_1, M_2)G$, are nonsingular. Consequently, the $\bar{n} \times \bar{n}$ submatrix of $\mathcal{H}(\mathcal{N})$

$$N_0 := [\langle \mathcal{N}, r_i c_j\rangle] = \mathcal{O}_{\bar{n}}\mathcal{R}_{\bar{n}}$$

is non singular too.

Introduce next the matrices

$$\bar{M}_1 := N_0^{-1}[\langle \mathcal{N}, r_i \xi_1 c_j\rangle] \qquad \bar{M}_2 := N_0^{-1}[\langle \mathcal{N}, r_i \xi_2 c_j\rangle], \tag{4.14}$$

$$\bar{H} := [\, \langle \mathcal{N}, c_1\rangle \quad \ldots \quad \langle \mathcal{N}, c_{\bar{n}}\rangle \,] \quad \text{and} \quad \bar{G} := N_0^{-1} \begin{bmatrix} \langle \mathcal{N}, r_1\rangle \\ \vdots \\ \langle \mathcal{N}, r_{\bar{n}}\rangle \end{bmatrix}. \tag{4.15}$$

Using the assumption on the rank of $\mathcal{H}(\mathcal{N})$, we apply the same arguments as in section 3 to derive a counterpart of (3.7) for noncommutative power series. So, for all $r, c \in \Xi^*$, we have

$$\Big[\langle \mathcal{N}, rc_1\rangle \ \dots \ \langle \mathcal{N}, rc_{\bar n}\rangle\Big] N_0^{-1} \begin{bmatrix} \langle \mathcal{N}, r_1 c\rangle \\ \dots \\ \langle \mathcal{N}, r_{\bar n} c\rangle \end{bmatrix} = \langle \mathcal{N}, rc\rangle. \tag{4.16}$$

It follows that

$$\begin{aligned} \bar M_1 \bar M_2 - \bar M_2 \bar M_1 \quad &= N_0^{-1}\{[\langle \mathcal{N}, r_i \xi_1 c_j\rangle] N_0^{-1} [\langle \mathcal{N}, r_i \xi_2 c_j\rangle] - [\langle \mathcal{N}, r_i \xi_2 c_j\rangle] N_0^{-1} [\langle \mathcal{N}, r_i \xi_1 c_j\rangle]\} \\ &= N_0^{-1}\ [\langle \mathcal{N}, r_i \xi_1 \xi_2 c_j\rangle - \langle \mathcal{N}, r_i \xi_2 \xi_1 c_j\rangle] = 0, \end{aligned}$$

because of assumption (4.12). So, $\bar M_1$ and $\bar M_2$ commute.

As an immediate consequence of (4.16) and definitions (4.14) $\div$ (4.15), we get

$$M_i G = N_0^{-1} \begin{bmatrix} \langle \mathcal{N}, r_1 \xi_i\rangle \\ \vdots \\ \langle \mathcal{N}, r_{\bar n} \xi_i\rangle \end{bmatrix}, \qquad i = 1, 2 \tag{4.17}$$

and, for all $v \in \Xi^*$,

$$[\,\langle \mathcal{N}, v\xi_i c_1\rangle \quad \dots \quad \langle \mathcal{N}, v\xi_i c_{\bar n}\rangle\,] = [\,\langle \mathcal{N}, vc_1\rangle \quad \dots \quad \langle \mathcal{N}, vc_{\bar n}\rangle\,]\ N_0^{-1} M_i, \quad i = 1, 2. \tag{4.18}$$

Finally, we propose to prove that, for all $w \in \Xi^*$

$$\langle \mathcal{N}, w\rangle = \bar H w(\bar M_1, \bar M_2) \bar G, \tag{4.19}$$

which corresponds to show that $\mathrm{tr}(w(A_1, A_2)) = \mathrm{tr}(\bar w(A_1, A_2))$, for all $w, \bar w \in \Xi^*$, such that $|w|_i = |\bar w|_i,\ i = 1, 2$.
(4.19) is easily verified for $w = 1$. For any $w = \xi_{i_1} \xi_{i_2} \cdots \xi_{i_\nu},\ \nu \geq 1$, by (4.16) and (4.17), we have

$$\langle \mathcal{N}, \xi_{i_1} \xi_{i_2} \cdots \xi_{i_\nu}\rangle = [\,\langle \mathcal{N}, \xi_{i_1} \xi_{i_2} \cdots \xi_{i_{\nu-1}} c_1\rangle \dots \langle \mathcal{N}, \xi_{i_1} \xi_{i_2} \cdots \xi_{i_{\nu-1}} c_{\bar n}\rangle\,] \bar M_{i_\nu} \bar G,$$

and, by iteratively applying (4.18),

$$\langle \mathcal{N}, \xi_{i_1} \xi_{i_2} \cdots \xi_{i_\nu}\rangle = [\,\langle \mathcal{N}, c_1\rangle \quad \dots \quad \langle \mathcal{N}, c_{\bar n}\rangle\,] \bar M_{i_1} \bar M_{i_2} \dots \bar M_{i_\nu} \bar G = \bar H w(\bar M_1, \bar M_2) \bar G.$$

iii) $\Rightarrow$ *i)* By Lemma 4.2, $\bar n := \mathrm{rank}\ \mathcal{H}_{(n^2-1)\times(n^2-1)}(\mathcal{N}) = \mathrm{rank}\ \mathcal{H}(\mathcal{N})$.
Thus, as in the proof of *ii)* $\Rightarrow$ *iii)*, we can represent $\mathcal{N}$ as

$$\mathcal{N} = \bar H (I - \bar M_1 \xi_1 - \bar M_2 \xi_2)^{-1} \bar G, \tag{4.20}$$

where $\bar M_1, \bar M_2 \in \mathbf{C}^{\bar n \times \bar n}$ commute, and (4.12) holds for all words in Ξ^*, independently of their length. Therefore, for all w in Ξ^*, with $h = |w|_1$ and $k = |w|_2$, we have

$$\mathrm{tr}(w(A_1, A_2)) = \binom{h+k}{h}^{-1} \sum_{i=1}^{n} \mathrm{tr}(A_1{}^{h} \sqcup\!\sqcup^{k} A_2). \tag{4.21}$$

Taking the commutative images on both sides of (4.20), we obtain

$$T_{A_1,A_2} + n = \bar{H}(I - \bar{M}_1 z_1 - \bar{M}_2 z_2)^{-1}\bar{G},$$

and, consequently,

$$T_{A_1,A_2} = -n + \frac{\bar{H}\mathrm{adj}(I - \bar{M}_1 z_1 - \bar{M}_2 z_2)\bar{G}}{\Delta_{\bar{M}_1,\bar{M}_2}(z_1, z_2)},$$

where $\Delta_{\bar{M}_1,\bar{M}_2}(z_1, z_2)$ splits into linear factors, because of the commutativity of $\bar{M}_1$, $\bar{M}_2$. Thus the characteristic polynomial of (A_1, A_2) is given by $\Delta_{A_1,A_2}(z_1, z_2) = \prod_{i=1}^n (1 - \lambda_i z_1 - \mu_i z_2)$, and (A_1, A_2) has property L.
As Proposition 3.1 gives $\mathrm{tr}(A_1{}^h \sqcup\!\sqcup^k A_2) = \binom{h+k}{h} \sum_{i=1}^n \lambda_i^h \mu_i^k$, condition (4.21) immediately implies

$$\mathrm{tr}(w(A_1, A_2)) = \sum_{i=1}^n \lambda_i^{|w|_1} \mu_i^{|w|_2}, \quad \forall\, w \in \Xi^*,$$

which is equivalent to property P ■

5 Special factorizations of the characteristic polynomial

In this section, we consider a further property of a matrix pair (A_1, A_2) that, like property L, can be expressed as a constraint on the factors of Δ_{A_1,A_2} as well as a condition on the spectra of the linear combinations $\alpha A_1 + \beta A_2$, $\alpha, \beta \in \mathbf{C}$. Pairs we refer to are those whose characteristic polynomials split into the product of distinct polynomials $\Delta_i(z_1, z_2)$, each of them having support included in a straight line of the plane $\mathbf{Z} \times \mathbf{Z}$, passing through the origin. The interest in this property is mostly due to the fact that, as we shall see, it constitutes an immediate generalization of finite memory and separability.
To begin with, we consider a single polynomial $\Delta(z_1, z_2)$ whose support is a subset of a straight line in $\mathbf{Z} \times \mathbf{Z}$, i.e. there exists $(\ell, m) \neq (0, 0)$ in $\mathbf{N} \times \mathbf{N}$ such that

$$\mathrm{supp}(\Delta) \subset \{(k\ell, km), k \in \mathbf{N}\}. \tag{5.1}$$

2D systems having Δ as characteristic polynomial exhibit several features which strictly resemble those of 1D systems. Indeed, the local state at $(0, 0)$ determines a free evolution which is identically zero except on a "strip" that includes the straight line $\{(k\ell, km), k \in \mathbf{Z}\}$, (see Fig.2, for $\ell = 2$ and $m = 1$).

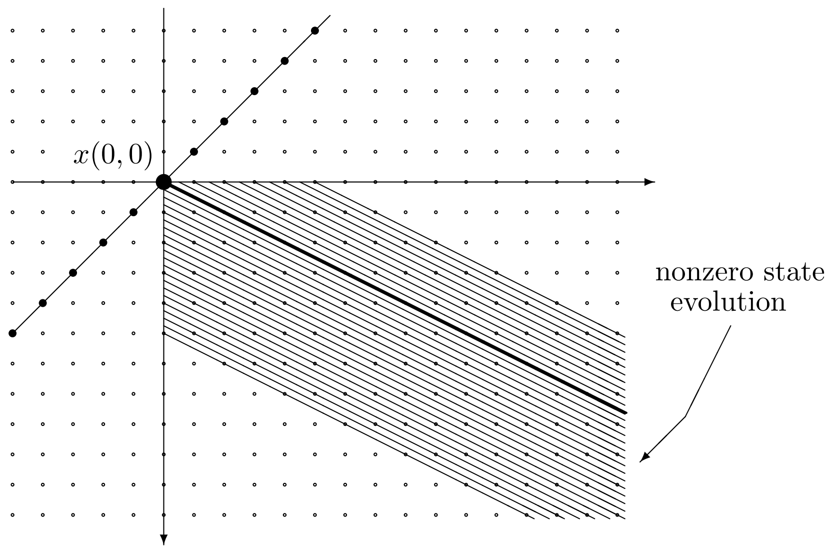

FIG.2

So, no matter how far (h,k) is from the set $\{(i,-i) : i \in \mathbf{Z}\}$ where the initial conditions are given, the local state in (h,k) is determined only by a finite subset of the initial global state, whose cardinality does not exceed a fixed integer N,

Proposition 5.1 *Let (A_1, A_2) be a pair of $n \times n$ matrices with entries in $\mathbf{C}$ and $\Delta_{A_1 A_2}(z_1, z_2)$ its characteristic polynomial. Assume moreover that (ℓ, m) is a pair of nonnegative integers and $1 = \text{g.c.d.}(\ell, m)$. The following statements are equivalent*

i) $\Delta_{A_1,A_2}(z_1, z_2) = 1 - \sum_{h=1}^{r} d_h (z_1^\ell z_2^m)^h$; (5.2)

ii) there exist $c_1, c_2, ..., c_n$ in $\mathbf{C}$ such that, for every $(\alpha, \beta) \in \mathbf{C} \times \mathbf{C}$ and every $(\ell+m)$-th root of $\alpha^\ell \beta^m$,

$$\Lambda(\alpha A_1 + \beta A_2) = \left(c_1 (\alpha^\ell \beta^m)^{\frac{1}{\ell+m}}, ..., c_n (\alpha^\ell \beta^m)^{\frac{1}{\ell+m}}\right); \tag{5.3}$$

iii) there exist $c_1, c_2, ..., c_n \in \mathbf{C}$ such that

$$\Lambda(\nu^m A_1 + \nu^{-\ell} A_2) = (c_1, c_2, ..., c_n), \quad \forall\, \nu \in \{1, 2, ...(\ell+m)n+1\}; \tag{5.4}$$

iv) $(i,j) \notin \{(k\ell, km), k \in \mathbf{N}_+\}$ *implies* $\text{tr}\ (A_1{}^i \sqcup\!\sqcup{}^j A_2) = 0$*;*

v) for all $\alpha, \beta \in \mathbf{C}$ and suitable $b_k \in \mathbf{C}$,

$$\text{tr}(\alpha A_1 + \beta A_2)^k = \begin{cases} b_k (\alpha^\ell \beta^m)^\nu & \text{if } k = (\ell+m)\nu \\ 0 & \text{otherwise;} \end{cases} \tag{5.5}$$

vi) $(i,j) \notin \mathcal{S}_n := \{(i,j) \in \mathbf{N} \times \mathbf{N} :\mid mi - \ell j \mid < n\}$ *implies* $A_1{}^i \sqcup\!\sqcup{}^j A_2 = 0$.

PROOF *i)* $\Rightarrow$ *ii)* Since $\Delta_{A_1,A_2}(z_1, z_2) \in \mathbf{C}[z_1^\ell z_2^m]$, there exist $\lambda_1, \lambda_2, ..., \lambda_r \in \mathbf{C}$ such that

$$\Delta_{A_1,A_2}(z_1, z_2) = \prod_{h=1}^{r} (1 - \lambda_h z_1^\ell z_2^m)$$

and, consequently,

$$\det(zI - \alpha A_1 - \beta A_2) = z^{n-r(\ell+m)} \prod_{h=1}^{r} (z^{\ell+m} - \lambda_h \alpha^\ell \beta^m). \tag{5.6}$$

Let $(\lambda_h)^{\frac{1}{\ell+m}}$ and $(\alpha^\ell\beta^m)^{\frac{1}{\ell+m}}$ be arbitrary $(\ell+m)$-th roots of λ_h and $\alpha^\ell\beta^m$ respectively, and ε any primitive $(\ell+m)$-th root of 1. The spectrum of $(\alpha A_1 + \beta A_2)$ is given by $\Lambda(\alpha A_1 + \beta A_2) = \left(c_1(\alpha^\ell\beta^m)^{\frac{1}{\ell+m}}, ..., c_n(\alpha^\ell\beta^m)^{\frac{1}{\ell+m}}\right)$, where

$$\begin{array}{rcll} c_{r\nu+h} & = & (\lambda_h)^{\frac{1}{\ell+m}}\varepsilon^\nu, & h = 1, ..., r \quad \text{and} \quad \nu = 1, ..., \ell+m, \\ c_\mu & = & 0, & \mu > (\ell+m)r. \end{array}$$

ii) $\Rightarrow$ *iii)* Obvious.

iii) $\Rightarrow$ *iv)* Clearly, for all $\nu \in \{1, 2, ..., (\ell+m)n+1\}$ and $h \in \mathbf{N}_+$

$$\operatorname{tr}(\nu^m A_1 + \nu^{-\ell} A_2)^h = \sum_{i=0}^{h} \nu^{(\ell+m)i-h\ell} \operatorname{tr}({A_1}^i \sqcup\!\sqcup^{h-i} A_2) = \sum_{i=1}^{n} c_i^h =: f_h,$$

whence $\sum_{i=0}^{h} \nu^{(\ell+m)i} \operatorname{tr}({A_1}^i \sqcup\!\sqcup^{h-i} A_2) - f_h \nu^{h\ell} = 0$. As in the polynomials

$$p_h(x) := \sum_{i=0}^{h} x^{(\ell+m)i} \operatorname{tr}({A_1}^i \sqcup\!\sqcup^{h-i} A_2) - f_h x^{h\ell}, \qquad h = 1, 2, ..., n,$$

the number of zeros exceeds the degree, all their coefficients have to be zero. We distinguish two cases.

Case 1: $k(\ell+m) = h\ell$, for some $k \in \mathbf{N}$. Since 1 is the unique common divisor of ℓ and m, there exists $t \in \mathbf{N}_+$ such that $k = \ell t$ and $h - k = mt$, and therefore

$$\operatorname{tr}({A_1}^i \sqcup\!\sqcup^{h-i} A_2) = \begin{cases} f_h & \text{if } (i, h-i) = t(\ell, m) \\ 0 & \text{otherwise.} \end{cases}$$

Case 2: $k(\ell+m) \neq h\ell$ for all $h \in \mathbf{N}$. Then, for $0 \le i \le h$, $\operatorname{tr}\,({A_1}^i \sqcup\!\sqcup^{h-i} A_2) = 0$.

iv) $\Rightarrow$ *v)* Obvious.

v) $\Rightarrow$ *i)* Equations (2.15) and (2.16) show that the homogeneus form δ_k of the characteristic polynomial satisfies

$$\delta_k(\alpha, \beta) = \begin{cases} d_\nu(\alpha^\ell\beta^m)^\nu & \text{if } k = (\ell+m)\nu \\ 0 & \text{otherwise.} \end{cases}$$

i) $\Rightarrow$ *vi)* Note that

$$\sum_{i,j=0}^{\infty} {A_1}^i \sqcup\!\sqcup^j A_2 \; z_1^i z_2^j = (I - A_1 z_1 - A_2 z_2)^{-1} = \frac{\operatorname{adj}(I - A_1 z_1 - A_2 z_2)}{1 - \sum_{h=1}^{r} d_h (z_1^\ell z_2^m)^h}.$$

As

$$\begin{array}{rcl} \operatorname{supp}(\operatorname{adj}(I - A_1 z_1 - A_2 z_2)) & \subseteq & \{(i,j) \in \mathbf{N} \times \mathbf{N} : i + j \le n - \ell\} \\ \operatorname{supp}((1 - \sum_{h=1}^{r} d_h (z_1^\ell z_2^m)^h)^{-1}) & \subseteq & \{(i,j) \in \mathbf{N} \times \mathbf{N} : mi = \ell j\}, \end{array}$$

it is clear that the support of $(I - A_1z_1 - A_2z_2)^{-1}$ is a subset of $\mathcal{S}_n$.

vi) $\Rightarrow$ *i)* Consider the injective ring homomorphism $\phi : \mathbf{C}[[z_1, z_2]] \to \mathbf{C}[[\eta, \xi, \xi^{-1}]]$ obtained by linearly extending the map that associates $z_1^i z_2^j$, $i, j \in \mathbf{N}$, with $\eta^k \xi^h$, where h, k are given by

$$\begin{bmatrix} h \\ k \end{bmatrix} = \begin{bmatrix} -m & \ell \\ 0 & 1 \end{bmatrix} \begin{bmatrix} i \\ j \end{bmatrix}.$$

ϕ maps any series in $\mathbf{C}[[z_1, z_2]]$ with support in $\mathcal{S}_r$ into an element of $\mathbf{C}[[\eta]][\xi, \xi^{-1}]$, the ring of Laurent polynomials [12] in the indeterminate ξ, with coefficients in $\mathbf{C}[[\eta]]$. As the support of $(I - A_1z_1 - A_2z_2)^{-1}$ is included in $\mathcal{S}_n$ and hence

$$\operatorname{supp}\Big(\det(I - A_1z_1 - A_2z_2)^{-1}\Big) \subseteq \mathcal{S}_{n^2},$$

applying the map ϕ on both sides of $\det(I - A_1z_1 - A_2z_2)^{-1}\Delta_{A_1,A_1}(z_1, z_2) = 1$, one gets

$$\phi(\det(I - A_1z_1 - A_2z_2)^{-1}\phi(\Delta_{A_1,A_2}(z_1, z_2)) = 1. \tag{5.7}$$

As both factors on the left hand side of (5.7) can be viewed as elements of $\mathbf{C}[[\eta]][\xi, \xi^{-1}]$, we have that $\phi(\Delta_{A_1,A_2}(z_1, z_2))$ is a unit of that ring, i.e. $\phi\Big(\Delta_{A_1,A_2}(z_1, z_2)\Big) = \xi^h s(\eta)$ for some $h \in \mathbf{Z}$ and $s(\eta) \in \mathbf{C}[[\eta]]$. Condition $\Delta_{A_1,A_2}(0, 0) = 1$ implies $h = 0$ and, therefore, Δ_{A_1,A_2} is a polynomial in $z_1^\ell z_2^m$ ∎

An immediate consequence of property ii) in the above proposition is the following

Corollary 5.2 *Consider A_1, A_2 in $\mathbf{C}^{n\times n}$. If* $\operatorname{supp}(\Delta_{A_1,A_2})$ *is a subset of a straight line, different from the coordinate axes, then both A_1 and A_2 are nilpotent* ∎

The results of Proposition 5.1 provide a convenient framework for understanding the internal dynamics of 2D finite memory state models, which arise quite naturally in several applications. For instance, when considering the realization of F.I.R. filters and dead beat regulators, the requirements on the state models we have to use cannot be exclusively expressed as conditions on the polynomial transfer matrix which represents the input-output map. Further aspects should be taken into account, which introduce additional constraints on the choice of the matrix pair (A_1, A_2).

a) If the state-output transfer matrix $C(I - A_1z_1 - A_2z_2)^{-1}$ is not polynomial, local states $\mathbf{x}$ exist, which give rise to free output evolutions $C(I - A_1z_1 - A_2z_2)^{-1}\mathbf{x}$ with infinite supports. Clearly such states, when induced by noise, generate infinite error sequences in the output signal.

b) If the input-state transfer matrix $(I - A_1z_1 - A_2z_2)^{-1}(B_1z_1 + B_2z_2)$ is not polynomial, finite support input sequences possibly produce infinite support sequences in the state space. Therefore the system could remain indefinitely excited by a finite signal, even though the corresponding output dies out in a finite number of steps.

Both previous drawbacks can be avoided if $(I - A_1z_1 - A_2z_2)^{-1}$ is polynomial or, equivalently, if the characteristic polynomial of the system is unitary, i.e.

$$\Delta_{A_1,A_2}(z_1, z_2) = \det(I - A_1z_1 - A_2z_2) = 1. \tag{5.8}$$

2D systems satisfying condition (5.8) are called "finite memory" [4,6], since they reach the zero state in a finite number of steps after zeroing the input signal.

Corollary 5.3 [Finite memory systems] [5,9,18] *Let A_1, A_2 be in $\mathbf{C}^{n\times n}$. The followings are equivalent*
FM_1) $\Delta_{A_1,A_2}(z_1,z_2) = 1$;
FM_2) $\Lambda(\alpha A_1 + \beta A_2) = (0, 0, ..., 0)$, $\forall \alpha, \beta \in \mathbf{C}$, *namely A_1 and A_2 are nilpotent and satisfy property L;*
FM_3) $\Lambda(\nu A_1 + A_2) = \Lambda(A_1 + \nu A_2) = (0, 0, ..., 0)$, $\nu = 1, ..., n+1$;
FM_4) $\mathrm{tr}(A_1 {}^{i}\sqcup\!\sqcup^{j} A_2) = 0$, $\forall\ (i,j) \neq (0,0)$;
FM_5) $A_1 {}^{i}\sqcup\!\sqcup^{j} A_2 = 0$, *for $i + j \geq n$.*

Proof Condition FM_1 is equivalent to assume that the support of $\Delta_{A_1A_2}(z_1,z_2)$ is a subset of both $\{(i,0) : i \in \mathbf{N}\}$ and $\{(0,j) : j \in \mathbf{N}\}$. Therefore finite memory systems are exactly those which satisfy properties $i) \div vi)$ of Proposition 5.1 both for $(\ell, m) = (1,0)$ and $(\ell, m) = (0,1)$.

FM_2) Choose first $(\ell, m) = (0,1)$ and then $(\ell, m) = (1,0)$. Then

$$\Lambda(\alpha A_1 + \beta A_2) = (c_1\beta, ..., c_n\beta) = (d_1\alpha, ..., d_n\alpha), \qquad \forall \alpha, \beta \in \mathbf{C},$$

which obviously implies $\Lambda(\alpha A_1 + \beta A_2) = (0, ..., 0)$.

FM_3) Assumptions $(\ell, m) = (0,1)$ and $(\ell, m) = (1,0)$ give $\Lambda(\alpha A_1 + \beta A_2) = (c_1, ..., c_n)$ and $\Lambda(\alpha A_1 + \beta A_2) = (d_1, ..., d_n)$, respectively. These imply

$$\Lambda(\alpha A_1 + \beta A_2) = (d_1\alpha, ..., d_n\alpha) = (c_1\beta, ..., c_n\beta), \qquad \forall \alpha, \beta \in \mathbf{C}.$$

Therefore $\Lambda(\nu A_1 + A_2) = \Lambda(A_1 + \nu A_2) = (0, ..., 0)$.

FM_4) Obvious from $iv)$ of Proposition 5.1.

FM_5) Assume first $(\ell, m) = (0,1)$ and then $(\ell, m) = (1,0)$. Point $vi)$ of Proposition 5.1 gives $A_1 {}^{i}\sqcup\!\sqcup^{j} A_2 = 0$, when $i \geq n$ or $j \geq n$. So $(I - A_1z_1 - A_2z_2)^{-1}$ is a polynomial matrix, and coincides with adj $(I - A_1z_1 - A_2z_2)$, whose support is included in $\{(i,j) : i+j < n\}$
■

The results of Proposition 5.1 partially extend to the case of a characteristic polynomial $\Delta(z_1,z_2)$ which factorizes into irreducible factors, each of them having support on a straight line through $(0,0)$. For sake of simplicity, we confine ourselves to the case when Δ factorizes as

$$\Delta(z_1,z_2) = \Delta_1(z_1,z_2)\Delta_2(z_1,z_2), \tag{5.9}$$

with

$$\Delta_i(z_1,z_2) = 1 - \sum_{j=1}^{r_i} d_j^{(i)} (z_1^{\ell_i} z_2^{m_i})^j, \quad i = 1,2 \tag{5.10}$$

and g.c.d.$(\ell_i, m_i) = 1$. The extension to the case of more than two factors is straightforward. If (A_1, A_2) is an $n \times n$ matrix pair with characteristic polynomial Δ, it can be easily shown that *i)* there exist two positive integers ρ and σ, $\rho+\sigma \leq n$, and $\rho+\sigma$ complex numbers $c_1,...,c_\rho, d_1,...,d_\sigma$, such that, for all $\alpha,\beta \in \mathbf{C}$

$$\begin{aligned}\Lambda(\alpha A_1+\beta A_2) \;\;=\;\; & \Big(c_1(\alpha^{\ell_1}\beta^{m_1})^{\frac{1}{\ell_1+m_1}},...,c_\rho(\alpha^{\ell_1}\beta^{m_1})^{\frac{1}{\ell_1+m_1}},\\ & d_1(\alpha^{\ell_2}\beta^{m_2})^{\frac{1}{\ell_2+m_2}},...,d_\sigma(\alpha^{\ell_2}\beta^{m_2})^{\frac{1}{\ell_2+m_2}},0,...,0\Big);\end{aligned}$$

ii) $\mathrm{tr}(A_1{}^i \sqcup\!\sqcup {}^j A_2) \neq 0$ implies either $(i,j)=(k\ell_1,km_1)$ or $(i,j)=(h\ell_2,hm_2)$, $h,k \in \mathbf{N}_+$.

Viceversa, each of the above properties guarantees that Δ factorizes as in (5.9)-(5.10).

We are now in a position for obtaining a fairly complete description of 2D systems whose characteristic polynomials factorize into the product of a polynomial in z_1 and a polynomial in z_2. Such systems are called "separable" [4,5] and are usually thought of as the simplest examples of I.I.R 2D systems. Actually, many properties one may hope to extrapolate from an understanding of 1D systems carry over to separable systems. Indeed, just the knowledge that the system is separable allows one to make fairly strong statements about its behaviour; in particular, internal stability can be quickly deduced from the general theory of discrete time 1D systems, as the long term performance of separable systems is determined by the eigenvalues of A_1 and A_2.

Proposition 5.4 [Separable systems] *Let A_1, A_2 be in $\mathbf{C}^{n\times n}$. The following statements are equivalent*

S_1) $\Delta_{A_1,A_2}(z_1,z_2)=r(z_1)s(z_2)$;

S_2) *A_1 and A_2 satisfy property L w.r.t. the orderings of the spectra*

$$\Lambda(A_1)=(\lambda_1,...,\lambda_\rho,0,...,0,0,...,0)$$

$$\Lambda(A_2)=(0,...,0,\mu_1,...,\mu_\sigma,0,...,0),$$

so that, for every $\alpha,\beta \in \mathbf{C}$ $\Lambda(\alpha A_1+\beta A_2)=(\alpha\lambda_1,...,\alpha\lambda_\rho,\beta\mu_1,...,\beta\mu_\sigma,0,...,0)$;

S_3) $\mathrm{tr}(A_1{}^i \sqcup\!\sqcup {}^j A_2)=0$ *if both i and j are nonzero;*

S_4) $\mathrm{tr}(\alpha A_1+\beta A_2)^k=\mathrm{tr}(\alpha A_1)^k+\mathrm{tr}(\beta A_2)^k$, $\forall \alpha,\beta \in \mathbf{C}, k \in \mathbf{N}_+$ $\blacksquare$

Property L, separability and finite memory have been introduced by progressively strengthening the constraints on the irreducible factors of $\Delta_{A_1,A_2}(z_1,z_2)$. On the other hand, the set of matrix pairs with property L properly includes the set of pairs with property P, which in turn is strictly larger than the set of commutative pairs. So, the question naturally arises whether the above constraints on the characteristic polynomial of the pair (A_1,A_2) can be related to property P and to commutativity.

We first observe that examples can be given of commutative pairs, and hence of pairs with property P, which are not finite memory and not even separable. Actually, just by taking diagonal matrices A_1 and A_2, we easily see that commutativity and property P do not imply any particular consequence on the characteristic polynomial, except that it factorizes into first order factors.

On the other hand, the following pair

$$A_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix}$$

is finite memory (and hence separable). Yet, it does not satisfy property P.

In view of this, no implication exists between commutativity and property P on one side, and finite memory and separability on the other. What is remarkable, however, is that if we restrict our analysis to matrix pairs with property P, finite memory and separable pairs can be nicely characterized in terms of semigroups of nilpotent matrices. This is made precise in the following proposition, which provides a slight extension (and an alternative proof) of the classical Amitsur-Levitzki theorem [8, pg.135].

Proposition 5.5 *Let A_1, A_2 be in $\mathbf{C}^{n\times n}$ and consider the multiplicative semigroups*

$$S := \{w(A_1, A_2),\ w \in \Xi^*,\ |w| \geq 1\}$$

and

$$\bar{S} := \{w(A_1, A_2),\ w \in \Xi^*,\ |w|_1 \geq 1,\ |w|_2 \geq 1\}.$$

The pair (A_1, A_2) has finite memory and property P (resp. separability and property P) if and only if all matrices in S (resp. in $\bar{S}$) are nilpotent.

Proof We first remark that, if (A_1, A_2) has property P, the nilpotency of all elements of S and $\bar{S}$ is equivalent to finite memory and separability, respectively. In fact, when A_1 and A_2 are in upper triangular form, the nilpotency of the elements of S and $\bar{S}$ corresponds to the assumption that the characteristic polynomial Δ_{A_1,A_2} satisfies (FM_1) of Cor. 5.3 and (S_1) of Prop 5.4, respectively.
Suppose now that the multiplicative semigroup S, generated by A_1 and A_2, is constituted by nilpotents. Since we have tr $w(A_1, A_2) = 0$ for all $w \in \Xi^*,\ |w| \geq 1$, (4.3) is clearly fulfilled. Consequently (A_1, A_2) has property P and, by the above remark, (A_1, A_2) is a finite memory pair.

On the other hand, assume that all matrices in $\bar{S}$ are nilpotent. This implies

$$\mathrm{tr}[(A_1A_2 - A_2A_1)w(A_1, A_2)] = 0,\ \forall w \in \Xi^* \tag{5.11}$$

which is a necessary and sufficient condition [14] for the pair (A_1, A_2) having property P. Again, the remark at the beginning of the proof shows that (A_1, A_2) is a separable pair ■

6 References

1. M.Bisiacco, E.Fornasini and G.Marchesini *2D partial fraction expansions and minimal commutative realizations* IEEE Trans. Circ. Sys., vol.CAS-35, pp.1533-38, 1988

2. M.Fliess *Un outil algébrique: Les séries formelles non commutatives*, Springer L.N. Econ. Math. Sys., 131, pp.122-49, 1975

3. E.Fornasini, G.Marchesini *Doubly indexed dynamical systems: state space models and structural properties* , Math. Sys. Theory, vol.12, pp.59-72, 1978

4. E.Fornasini, G.Marchesini, "Properties of pairs of matrices and state models for 2D systems. Pt.I : State dynamics and geometry of the pairs; Pt.II : Models structure and realization problems", in *Multivariate Analysis: Future Directions*, C.R.Rao Ed., North Holland Series in Probability and Statistics, vol.5, pp.131-80, 1993

5. E.Fornasini, G.Marchesini and M.E.Valcher, "On the structure of finite memory and separable 2D systems", *Automatica*, vol.30, n.2, pp. 347-350, 1994

6. E.Fornasini, S.Zampieri *A note on the state space realization of 2D FIR transfer functions*, Systems and Control Letters, vol.16, n.2, pp. 17-22, 1990

7. N.Jacobson *Basic Algebra I*, Freeman, San Francisco, 1974

8. I.Kaplansky *Fields and Rings* Chicago Univ.Press, 1972

9. B.Mathes, M.Omladic, H.Radjavi *Linear spaces of nilpotent matrices*, Lin. Alg. Appl., 149, pp.215-225, 1991

10. N.H.McCoy *On the characteristic roots of matric polynomials*, Bull. Am. Math. Soc., 42, pp. 592-600, 1936

11. T.S.Motzkin, O.Taussky *Pairs of matrices with property L*, Trans. Amer. Math. Soc., vol.73, pp.108-114, 1952

12. T.S.Motzkin, O.Taussky *Pairs of matrices with property L. (II)*, Trans. Amer. Math. Soc., vol.80, pp.387-401, 1955

13. U.Oberst *Multidimensional constant linear systems*, Acta Appl.Math., n.20, pp.1-175, 1990

14. H.Radjavi *A trace condition equivalent to simultaneous triangularizability*, Can. J.Math., vol.XXXVIII, n.2, pp.376-386, 1986

15. A.Salomaa, M.Soittola *Automata-theoretic aspects of formal power series*, Springer-Verlag,1978

16. O.Taussky *Some results concerning the transition from the L- to the P-property for pairs of finite matrices*, J. of Algebra, n.20, pp.271-83, 1972

17. O.Taussky *Commutativity in finite matrices*, Amer. Math. Monthly, vol.64, pp.229-235, 1957

18. M.E.Valcher, E.Fornasini *Polynomial inverses or 2D transfer matrices and finite memory realization via inverse systems* , Multidim. Sys. and Signal Proc., vol.4, pp.269-84, 1993