

Matrix pairs and 2D systems analysis

Ettore Fornasini, Giovanni Marchesini

Maria Elena Valcher

Dipartimento di Elettronica ed Informatica

Università di Padova, Padova, Italy

Abstract

Pairs of linear transformations on a finite dimensional vector space are of great relevance in the analysis of two-dimensional (2D) systems evolutions. In this paper, special properties of matrix pairs, such as finite memory, separability, property L and property P, as well as their dynamical interpretations, are investigated. Practical criteria for testing property L and property P in a finite number of steps are also presented.

The nonnegativity requirement on a matrix pair allows for much stronger characterizations of finite memory and separability, which in fact prove to be structural properties. Finally, the irreducibility and primitivity notions of positive matrix pairs are discussed, and connected with the dynamical behavior of the associated 2D systems.

1 Introduction

The theory of pairs and, more generally, k -tuples of linear transformations on a finite dimensional vector space dates from about the end of the last century. Nevertheless, in spite of its relatively long history, important general results, such as the invariant theory of $n \times n$ matrices (see [16] and the references therein) and a wealth of more specialized topics, recently appeared in mathematical journals, showing the permanent fertility of this field.

1D system theory has often taken advantage of these results (e.g. in Fliess' representation of bilinear systems) and, conversely, problems arising in modelling and control have provided a permanent source of new mathematical questions involving two or more linear transformations. Even more interestingly, system theoretic methods opened new vistas on purely algebraic problems on matrix sets, as in the case of matrix pairs with common eigenvectors [18].

In 2D system theory, virtually all the problems of modelling, realization and control involve a pair of linear transformations, whose spectral and combinatorial properties characterize the pattern of dynamical trajectories. Furthermore, 2D

systems endowed with special structures, such as finite memory, separability, positivity etc., bring into focus special classes of matrix pairs, that deserve on their own a careful investigation. As several definitions and problems that one might hope to extend from 1D dynamical models do, in fact, carry over to 2D models, classical system theoretic concepts often suggest the guideline for tackling with pairs of linear transformations. It should be stressed, however, that the mathematical tools needed in the analysis of matrix pairs are substantially more difficult, and the mechanics of the proofs more involved, than most others occurring in Linear Algebra.

This paper deals with the circle of ideas that includes the theory of matrix pairs, two-dimensional signals and 2D state space models. In the next section the main properties of the characteristic polynomial and trace series of an unconstrained matrix pair are summarized. In section 3 some special classes of pairs, as well as their connections with the dynamical behavior of 2D state models, are investigated. Section 4 is devoted to nonnegative 2D systems and pairs. Detailed proofs of the results presented in sections 2 and 3 can be found in [6, 7], while for positive pairs we refer the interested reader to [8, 9]. The concluding section presents some open problems and further issues on matrix pairs arising from current researches.

2 Characteristic polynomials and trace series

2D systems theory connotes a large collection of problems and methods held together by a central theme: the analysis and control of processes and devices whose dynamics depends on two independent variables. In this note, however, we restrict our attention to the very special class of quarter plane causal unforced 2D state models, whose description essentially depends on a pair of square matrices associated with the shift operators along the coordinate axes [4]:

$$\mathbf{x}(h+1, k+1) = A_1 \mathbf{x}(h, k+1) + A_2 \mathbf{x}(h+1, k). \quad (1)$$

The *local states* $\mathbf{x}(h, k)$ are defined on (a suitable region of) the discrete plane $\mathbb{Z} \times \mathbb{Z}$ and take values in \mathbb{R}^n , A_1 and A_2 are real $n \times n$ matrices and the initial conditions are usually assigned by specifying all local state values $\mathbf{x}(-\ell, \ell)$ on the *separation set* $\mathcal{C}_0 := \{(-\ell, \ell) : \ell \in \mathbb{Z}\}$. Upon resorting to formal power series, initial conditions are represented by the *global state* $\mathcal{X}_0 := \sum_{\ell=-\infty}^{+\infty} \mathbf{x}(-\ell, \ell) z_1^{-\ell} z_2^\ell$ and the induced sequence of local states in the half-plane $\{(h, k) : h+k \geq 0\}$ is described by

$$X(z_1, z_2) = \sum_{h,k} \mathbf{x}(h, k) z_1^h z_2^k = (I - A_1 z_1 - A_2 z_2)^{-1} \mathcal{X}_0 = \left[\sum_{i,j=0}^{\infty} A_1^i \mathfrak{W}^j A_2 z_1^i z_2^j \right] \mathcal{X}_0.$$

where the *Hurwitz products* $A_1^i \mathfrak{W}^j A_2$, $i, j \in \mathbb{N}$, are the matrix coefficients of the power series expansion of $(I - A_1 z_1 - A_2 z_2)^{-1}$. Each Hurwitz product decomposes as

$A_1^i \mathfrak{W}^j A_2 = \sum_{\nu_1, \nu_2, \dots, \nu_{i+j}} A_{\nu_1} A_{\nu_2} \dots A_{\nu_{i+j}}$, the summation being extended to all products that include the factors A_1 and A_2 , i and j times respectively. In particular, when assuming $\mathbf{x}(-\ell, \ell) = 0$ for every $\ell \neq 0$, the state in (h, k) is given by

$$\mathbf{x}(h, k) = \sum_{\nu_1, \nu_2, \dots, \nu_{i+j}} A_{\nu_1} A_{\nu_2} \dots A_{\nu_{i+j}} \mathbf{x}(0, 0), \quad (2)$$

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.

Most analytic and combinatorial properties of the pair (A_1, A_2) obviously reflect into the behavior of (1), and it is often useful to translate an abstract question on the pair into a dynamical problem for the corresponding state model. This is well illustrated by the *characteristic polynomial* and the *trace series* of (A_1, A_2) , defined as

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

and

$$T_{A_1, A_2}(z_1, z_2) := \sum_{h=1}^{\infty} \left(\sum_{i+j=h} \text{tr}(A_1^i \mathfrak{W}^j A_2) z_1^i z_2^j \right),$$

respectively. Like the characteristic polynomial of a single matrix, which in general does not capture its Jordan structure, Δ_{A_1, A_2} does not identify the similarity orbit of the pair and hence does not provide a complete information on the underlying linear transformations. Nevertheless, important aspects of the 2D motion completely rely on it. There is, first of all, the internal stability of system (1), which depends only on the variety of the zeros 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} .

We shall not dwell here with stability issues, and concentrate instead on the remaining topics.

Proposition 2.1 *Let $\Delta_{A_1, A_2}(z_1, z_2) = 1 - \sum_{h=1}^n \delta_h(z_1, z_2)$ and $T_{A_1, A_2}(z_1, z_2) = \sum_{h=1}^{\infty} \tau_h(z_1, z_2)$, with $\delta_h(z_1, z_2)$ and $\tau_h(z_1, z_2)$ homogeneous forms of degree h , be the characteristic polynomial and trace series, respectively, of the $n \times n$ matrix pair (A_1, A_2) . Then*

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

$$\begin{aligned} \tau_1(z_1, z_2) - \delta_1(z_1, z_2) &= 0 \\ \tau_2(z_1, z_2) - \delta_1(z_1, z_2)\tau_1(z_1, z_2) - 2\delta_2(z_1, z_2) &= 0 \\ \dots & \\ \tau_n(z_1, z_2) - \delta_1(z_1, z_2)\tau_{n-1}(z_1, z_2) - \dots - n\delta_n(z_1, z_2) &= 0 \end{aligned} \quad (3)$$

and, for all $k > 0$,

$$\tau_{n+k}(z_1, z_2) - \sum_{i=1}^n \tau_{n+k-i}(z_1, z_2)\delta_i(z_1, z_2) = 0; \quad (4)$$

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

$$\mathrm{tr}(A_1^i \mathbf{u}^j A_2) = \sum_{0 < r+s < i+j} d_{rs} \mathrm{tr}(A_1^{i-r} \mathbf{u}^{j-s} A_2) + (i+j)d_{ij}, \quad (5)$$

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

Equation (5) has some simple, but useful consequences. First of all, it provides an algorithm for recursively computing the traces of $A_1^i \mathbf{u}^j A_2$ 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 $\Delta_{A_1, A_2}(z_1, z_2)$, is made possible. Actually, if an upper bound \bar{n} on the degree of $\Delta_{A_1, A_2}(z_1, z_2)$ is known, then assigning $\mathrm{tr}(A_1^i \mathbf{u}^j A_2)$ for $i+j \leq \bar{n}$ allows to recover both $\Delta_{A_1, A_2}(z_1, z_2)$ and the traces of $A_1^i \mathbf{u}^j A_2$ for $i+j > \bar{n}$. Finally, it is easy to realize that T_{A_1, A_2} is a rational power series, since its coefficients satisfy the recursive relation (5). Further connections of T_{A_1, A_2} with the structure of the characteristic polynomial are enlightened by the following proposition.

Proposition 2.2 *Let $\Delta_{A_1, A_2}(z_1, z_2) = 1 - \sum_{h=1}^n \delta_h(z_1, z_2)$ 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*

$$T_{A_1, A_2}(z_1, z_2) = \frac{\delta_1(z_1, z_2) + 2\delta_2(z_1, z_2) + \dots + n\delta_n(z_1, z_2)}{\Delta(z_1, z_2)}. \quad (6)$$

Moreover, if Δ_{A_1, A_2} factorizes as $\Delta(z_1, z_2) = \prod_{i=1}^t \Delta_i(z_1, z_2)^{\nu_i}$, with $\Delta_i(z_1, z_2) = 1 - \sum_{j=1}^{r_i} \delta_j^{(i)}(z_1, z_2)$ irreducible distinct factors, $\nu_i \in \mathbb{N}$ and $\delta_j^{(i)}(z_1, z_2)$, $i = 1, 2, \dots, t$, homogeneous polynomials of degree j , then

$$T_{A_1, A_2}(z_1, z_2) = \sum_{i=1}^t \nu_i \frac{\sum_{j=1}^{r_i} j \delta_j^{(i)}(z_1, z_2)}{\Delta_i(z_1, z_2)}. \quad (7)$$

In (7) the trace series $T_{A_1, A_2}(z_1, z_2)$ is expressed 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 irreducible factors with multiplicity one. On the other hand, when an irreducible rational function $T(z_1, z_2)$ has been given, (7) suggests a way to check whether $T(z_1, z_2)$ can be expanded into a trace series.

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

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

Basing on the previous results, useful sets of complete invariants for the equivalence relation \sim are easily derived, involving Hurwitz products and linear combinations of the matrix pairs, as illustrated in the following proposition.

Proposition 2.3 *Let (A_1, A_2) and (\hat{A}_1, \hat{A}_2) be two (complex-valued) matrix pairs, possibly of different size. 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 $\alpha, \beta \in \mathbb{C}$, $\Lambda_0(\alpha A_1 + \beta A_2) = \Lambda_0(\alpha \hat{A}_1 + \beta \hat{A}_2)$, where $\Lambda_0(A)$ denotes the set of nonzero eigenvalues of the matrix A , each of them counted according with the corresponding algebraic multiplicity;
- iii) for all $\alpha, \beta \in \mathbb{C}$ and $k \in \mathbb{N}_+$, $\text{tr}(\alpha A_1 + \beta A_2)^k = \text{tr}(\alpha \hat{A}_1 + \beta \hat{A}_2)^k$;
- iv) for all $(i, j) \neq (0, 0)$, $\text{tr}(A_1^i \mathbf{w}^j A_2) = \text{tr}(\hat{A}_1^i \mathbf{w}^j \hat{A}_2)$.
- v) $T_{A_1, A_2}(z_1, z_2) = T_{\hat{A}_1, \hat{A}_2}(z_1, z_2)$;

3 Pairs of matrices with special structure

In this section we aim to specifically focus on matrix pairs endowed with property L or property P. Pairs with property L occur quite frequently in the applications: indeed, the important classes of finite memory and separable 2D systems are described by pairs with this property. A pair of $n \times n$ matrices with entries in \mathbb{C} , (A_1, A_2) , has *property L* if the eigenvalues of A_1 and A_2 can be ordered into two n -tuples

$$\Lambda(A_1) = (\lambda_1, \lambda_2, \dots, \lambda_n) \quad \text{and} \quad \Lambda(A_2) = (\mu_1, \mu_2, \dots, \mu_n) \quad (8)$$

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

$$\Lambda(\alpha A_1 + \beta A_2) = (\alpha \lambda_1 + \beta \mu_1, \dots, \alpha \lambda_n + \beta \mu_n). \quad (9)$$

Clearly matrices $\hat{A}_1 = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ and $\hat{A}_2 = \text{diag}\{\mu_1, \mu_2, \dots, \mu_n\}$ constitute a pair with property L. Any other pair (A_1, A_2) of the same dimension, with property L w.r.t. the orderings (8), satisfies, by Proposition 2.3, $\Lambda(\alpha A_1 + \beta A_2) = \Lambda(\alpha \hat{A}_1 + \beta \hat{A}_2)$. A straightforward consequence is the chain of equivalences listed in the following proposition.

Proposition 3.1 *Let A_1 and A_2 be in $\mathbb{C}^{n \times n}$. The following statements are equivalent:*

- L) (A_1, A_2) has property L w.r.t. the orderings (8) of their spectra;
- L₁) $\Delta_{A_1, A_2}(z_1, z_2) = \prod_{i=1}^n (1 - \lambda_i z_1 - \mu_i z_2)$;
- L₂) for all $\alpha, \beta \in \mathbb{C}$ and $k \in \mathbb{N}$, $\text{tr}(\alpha A_1 + \beta A_2)^k = \sum_{i=1}^n (\alpha \lambda_i + \beta \mu_i)^k$;
- L₃) for every $(h, k) \in \mathbb{N} \times \mathbb{N}$, $\text{tr}(A_1^h \mathbf{w}^k A_2) = \binom{h+k}{h} \sum_{i=1}^n \lambda_i^h \mu_i^k$;
- L₄) $T_{A_1, A_2}(z_1, z_2) = \sum_{h+k>0} \text{tr}(A_1^h \mathbf{w}^k A_2) z_1^h z_2^k = \sum_{i=1}^n \frac{\lambda_i z_1 + \mu_i z_2}{1 - \lambda_i z_1 - \mu_i z_2}$.

Hankel matrix theory provides a direct method 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 introduce [3] the infinite Hankel matrix $\mathcal{H}(s)$ of a formal power series s in the commuting variables z_1 and z_2 . If $s := \sum_{h,k} \langle s, z_1^h z_2^k \rangle z_1^h z_2^k$, with $\langle s, z_1^h z_2^k \rangle$ denoting the coefficients of the monomial $z_1^h z_2^k$ in s , we have

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

Rows and columns of $\mathcal{H}(s)$ take indices in the multiplicative monoid of the terms $\mathcal{T} := \{z_1^h z_2^k : h, k \in \mathbb{N}\}$. For all $M', M'' \in \mathbb{N}$, we 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 homogeneous degree not greater than M' (M'').

We are now in a position to state the following criterion.

Proposition 3.2 *Let (A_1, A_2) be a pair of matrices in $\mathbb{C}^{n \times n}$ and consider the power series*

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

(A_1, A_2) has property L if and only if

$$\text{rank } \mathcal{H}_{(n-1) \times (n-1)}(R_{A_1, A_2}) = \text{rank } \mathcal{H}_{n \times n}(R_{A_1, A_2}) \leq n.$$

The results of Proposition 3.1 provide a convenient framework for understanding the internal dynamics of finite memory and separable 2D state models. A 2D system (1) that reaches the zero (global) state within a finite number of steps, independently of its initial conditions, is called *finite memory* [5, 6]. It is almost immediate to show that the finite memory (FM) property corresponds to assume that $(I - A_1 z_1 - A_2 z_2)^{-1}$ is a polynomial matrix or, equivalently, that the characteristic polynomial of the pair (A_1, A_2) is unitary. The following proposition summarizes a useful set of different characterizations of the FM property.

Proposition 3.3 *Let A_1 and A_2 be matrices in $\mathbb{C}^{n \times n}$. The following facts are equivalent*

- $FM_1)$ $\Delta_{A_1, A_2}(z_1, z_2) = 1$;
- $FM_2)$ $\Lambda(\alpha A_1 + \beta A_2) = (0, 0, \dots, 0)$, $\forall \alpha, \beta \in \mathbb{C}$;
- $FM_3)$ $\Lambda(\nu A_1 + A_2) = \Lambda(A_1 + \nu A_2) = (0, 0, \dots, 0)$, for every $\nu = 1, \dots, n+1$;
- $FM_4)$ $\text{tr}(A_1^i \mathbf{1}^j A_2) = 0$, $\forall (i, j) \neq (0, 0)$;
- $FM_5)$ $A_1^i \mathbf{1}^j A_2 = 0$, for $i + j \geq n$.

2D systems whose characteristic polynomials factorize into the product of a polynomial in z_1 and a polynomial in z_2 are called *separable* [5, 6], and are usually thought of as the simplest examples of 2D systems with infinite impulse response.

Actually, many properties one may hope to extrapolate from an understanding of 1D systems carry over to separable systems, and 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 3.4 *Let A_1 and A_2 be matrices in $\mathbb{C}^{n \times n}$. 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 $\Lambda(A_1) = (\lambda_1, \dots, \lambda_\rho, 0, \dots, 0, 0, \dots, 0)$, $\Lambda(A_2) = (0, \dots, 0, \mu_1, \dots, \mu_\sigma, 0, \dots, 0)$, so that $\Lambda(\alpha A_1 + \beta A_2) = (\alpha \lambda_1, \dots, \alpha \lambda_\rho, \beta \mu_1, \dots, \beta \mu_\sigma, 0, \dots, 0)$ for every $\alpha, \beta \in \mathbb{C}$;
- $S_3)$ $\text{tr}(A_1^i \mathbf{1}^j A_2) = 0$ whenever i and j are both nonzero;
- $S_4)$ $\text{tr}(\alpha A_1 + \beta A_2)^k = \text{tr}(\alpha A_1)^k + \text{tr}(\beta A_2)^k$, $\forall \alpha, \beta \in \mathbb{C}, k \in \mathbb{N}_+$.

Before turning to the discussion of property P, we have to introduce some background notations on polynomials and series in noncommutative variables. Given an alphabet $\Xi = \{\xi_1, \xi_2\}$, the free monoid Ξ^* with base Ξ is the set of all words $w = \xi_{i_1} \xi_{i_2} \cdots \xi_{i_m}$, $m \in \mathbb{N}$, $\xi_{i_h} \in \Xi$. The integer m is called the length of w and denoted by $|w|$, while $|w|_i$ represents the number of occurrences 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 $|1| = 0$.

$\mathbb{C}\langle \xi_1, \xi_2 \rangle$ and $\mathbb{C}\langle\langle \xi_1, \xi_2 \rangle\rangle$ 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 $\mathbb{C}^{n \times n}$, 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 $\mathbb{C}\langle \xi_1, \xi_2 \rangle$ into $\mathbb{C}^{n \times n}$. The ψ -image of a polynomial $\mathcal{P}(\xi_1, \xi_2) \in \mathbb{C}\langle \xi_1, \xi_2 \rangle$ is denoted by $\mathcal{P}(A_1, A_2)$.

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

$$\Lambda(A_1) = (\lambda_1, \lambda_2, \dots, \lambda_n), \quad \Lambda(A_2) = (\mu_1, \mu_2, \dots, \mu_n), \quad (10)$$

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

$$\Lambda(\mathcal{P}(A_1, A_2)) = (\mathcal{P}(\lambda_1, \mu_1), \mathcal{P}(\lambda_2, \mu_2), \dots, \mathcal{P}(\lambda_n, \mu_n)). \quad (11)$$

According to a celebrated result of Mc Coy [12], property P is equivalent to simultaneous triangularizability, a feature that allows for good insights into the geometry of system (1), which can be viewed as a cascade of 2D systems having dimension one. Moreover, when considering state models with inputs and outputs, triangular matrix pairs provide 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 [2]. It should be stressed that the same is not true, however, if we consider only commutative matrix pairs.

As property P trivially implies property L, while examples can be given [14, 15] disproving the converse, the set of all pairs with property P is properly included in the set of pairs with property L. On the other hand, we have seen in Proposition 3.1 that a pair is endowed with property L if and only if its characteristic polynomial factorizes into a product of linear terms, and it is obvious that any such polynomial corresponds also to some matrix pair with property P. Consequently, no possibility is left of describing property P basing only on the characteristic polynomial. The appropriate tools are, instead, certain noncommutative polynomials and power series associated with the pair, as well as the corresponding Hankel matrices.

Proposition 3.5 *Let A_1 and A_2 be $n \times n$ matrices with entries in \mathbb{C} , and consider the orderings of their spectra given in (10). The following statements are equivalent:*

- P_1) (A_1, A_2) has property P w.r.t. the orderings (10);
- P_2) for any $w \in \Xi^*$, with $|w|_1 = h$ and $|w|_2 = k$, $\text{tr}(w(A_1, A_2)) = \sum_{i=1}^n \lambda_i^h \mu_i^k$;
- P_3) the noncommutative power series $\mathcal{N} = \sum_{w \in \Xi^*} \text{tr}(w(A_1, A_2))w$ can be represented as $\mathcal{N} = \sum_{i=1}^n (1 - \lambda_i \xi_1 - \mu_i \xi_2)^{-1}$, and hence is recognizable [3];
- P_4) for any $w \in \Xi^*$, with $|w|_1 = h$ and $|w|_2 = k$, $\det(zI - w(A_1, A_2)) = \prod_{i=1}^n (z - \lambda_i^h \mu_i^k)$.

An effective method for testing property P depends on the Hankel matrix [3, 17] $\mathcal{H}(\mathcal{N})$ of the noncommutative power series \mathcal{N} , i.e. the infinite matrix whose rows and columns are indexed by the words of Ξ^* , and whose element with indexes u and v is equal to $\langle \mathcal{N}, uv \rangle$. The words in Ξ^* , and consequently the row and column indexes in $\mathcal{H}(\mathcal{N})$, are ordered according to their length, while the lexicographical order is adopted for words of the same length. For all $M', M'' \in \mathbb{N}$, 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'').

Proposition 3.6 *Let A_1 and A_2 be $n \times n$ matrices with entries in \mathbb{C} , and consider the associated noncommutative power series $\mathcal{N} = \sum_{w \in \Xi^*} \text{tr}(w(A_1, A_2))w$. The following statements are equivalent:*

- i) (A_1, A_2) has property P;
- ii) $\text{rank } \mathcal{H}_{(n^2-1) \times (n^2-1)}(\mathcal{N}) = \bar{n} \leq n$ and, for all pairs of words w, \bar{w} with length not greater than $2\bar{n}$,

$$|w|_i = |\bar{w}|_i, \quad i = 1, 2 \quad \Rightarrow \quad \text{tr}(w(A_1, A_2)) = \text{tr}(\bar{w}(A_1, A_2)); \quad (12)$$

- iii) (12) holds for all pairs of words w, \bar{w} with length not greater than $2n^2$.

4 Nonnegative matrix pairs

The interest in nonnegative matrix pairs is largely motivated, apart from theoretical reasons, by their possible applications in building discrete models of dynamical systems that involve only nonnegative variables, such as pressures, concentrations, levels of population etc.[10, 11, 19].

The results so far obtained have extended to positive pairs the available characterizations of finite memory, separability and property L. Furthermore, some basic notions of positive matrix theory, like irreducibility, primitivity and the Perron-Frobenius theorem, particularly meaningful for their relevance in (1D) positive system dynamics, have been generalized to the 2D case.

In this section the following notation will be adopted: given a matrix $A = [a_{ij}]$, we will write $A \gg 0$ (A *strictly positive*), if $a_{ij} > 0$ for all i, j ; $A > 0$ (A *positive*), if $a_{ij} \geq 0$ for all i, j , and $a_{hk} > 0$ for at least one pair (h, k) ; $A \geq 0$ (A *nonnegative*), if $a_{ij} \geq 0$ for all i, j .

4.1 Nonnegative pairs with special properties

Introducing the nonnegativity assumption allows to obtain more penetrating characterizations of finite memory and separable matrix pairs. In particular, finite memory and separability turn out to be “structural properties” of a pair, in the sense that they depend only on the zero patterns of the matrices and not on the specific values their nonzero entries take.

This is a quite interesting feature. Indeed, in many cases the information available on the physical process one aims to model allows to assume that no interaction exists among certain variables, and, consequently, that some entries of the matrices A_1 and A_2 are exactly 0, whereas the others can be assumed nonnegative, and known with some level of uncertainty. This is always the case of compartmental models [10], where nonzero entries correspond to the existence of flows between different compartments, and physical or biological reasons guarantee that some pairs of compartments have no direct interaction at all.

Proposition 4.1 *For a pair of $n \times n$ nonnegative matrices (A_1, A_2) , the following statements are equivalent*

- i) $\Delta_{A_1, A_2}(z_1, z_2) = 1$;
- ii) $A_1 + A_2$ is a nilpotent (and, a fortiori, a reducible) matrix;
- iii) $A_1^i \mathbf{w}^j A_2$ is nilpotent, for all $(i, j) \neq (0, 0)$;
- iv) $w(A_1, A_2)$ is nilpotent, for all $w \in \Xi^* \setminus \{1\}$;
- v) there exists a permutation matrix P such that $P^T(A_1 + A_2)P$ is upper triangular with zero diagonal entries.

Notice that in the general case, when the matrix entries assume both positive and negative values, condition *ii*) is necessary, but not sufficient, for guaranteeing the finite memory property, which depends on the nilpotency of all linear combinations

$\alpha A + \beta A_2$, $\alpha, \beta \in \mathbb{C}$. On the contrary, examples can be given [8] showing that conditions *iii*) and *iv*) are sufficient, but not necessary, for the finite memory property. Moreover, while for a general finite memory pair (A_1, A_2) we can only guarantee that the Hurwitz products $A^i \mathbf{1}^j A_2$ are zero when $i + j \geq n$, in the nonnegative case this property extends to all matrix products $w(A_1, A_2)$, $w \in \Xi^*$ and $|w| \geq n$.

In analyzing nonnegative separable pairs we end up with some results that strictly parallel those obtained in the finite memory case. A fairly complete spectral characterization of separability is summarized in the following proposition.

Proposition 4.2 *For a nonnegative matrix pair (A_1, A_2) , the following statements are equivalent:*

- i) $\Delta_{A_1, A_2}(z_1, z_2) = r(z_1)s(z_2)$;*
- ii) $\det[I - (A_1 + A_2)z] = \det[I - A_1z] \det[I - A_2z]$;*
- iii) $A_1^i \mathbf{1}^j A_2$ is nilpotent for all (i, j) with $i, j > 0$;*
- iv) $w(A_1, A_2)$ is nilpotent, for all $w \in \Xi^*$ such that $|w|_i > 0$, $i = 1, 2$;*
- v) there exists a permutation matrix P such that $P^T A_1 P$ and $P^T A_2 P$ are conformably partitioned into block triangular matrices*

$$P^T A_1 P = \begin{bmatrix} [A_1]_{11} & * & * & * \\ & [A_1]_{22} & * & * \\ & & \ddots & * \\ & & & [A_1]_{tt} \end{bmatrix} \quad P^T A_2 P = \begin{bmatrix} [A_2]_{11} & * & * & * \\ & [A_2]_{22} & * & * \\ & & \ddots & * \\ & & & [A_2]_{tt} \end{bmatrix}, \quad (13)$$

with $[A_1]_{ii} \neq 0$ implying $[A_2]_{ii} = 0$;

- vi) there exists a nonsingular matrix $T \in \mathbb{C}^{n \times n}$ such that $\hat{A}_1 = T^{-1} A_1 T$ and $\hat{A}_2 = T^{-1} A_2 T$ are upper triangular matrices, and the Hadamar product $A_1 * A_2$ has zero diagonal entries.*

The characterizations given in points v) of Propositions 4.1 and 4.2 have a combinatorial nature and make it clear that finite memory and separability are preserved under all perturbations of the positive entries. Property L and property P, instead, are not structural ones, as examples can be given [8] showing that they depend on the specific values assumed by the nonzero entries.

4.2 Irreducibility and primitivity

In positive matrix theory, irreducibility property of a matrix A can be introduced in combinatorial terms: indeed, irreducible matrices are those that cannot be reduced, by means of cogredience transformations, to block-triangular form [13]. Several equivalent descriptions are also available: algebraic ones, which connect irreducibility to the zero-patterns of the powers of A , and graph-theoretic and system theoretic characterizations, which relate it to the structure of the corresponding directed graph $\mathcal{D}(A)$ and to the behavior of the associated state model.

More precisely, a positive matrix $A \in \mathbb{R}^{n \times n}$ is irreducible if and only if positive integers h and T can be found, such that

$$\sum_{i=t+1}^{t+h} A^i \gg 0, \quad \text{for every } t \geq T, \quad (14)$$

or equivalently, if and only if for every pair of vertices i and j in $\mathcal{D}(A)$ there exists a path connecting i to j or, finally, if and only if for every positive initial condition $\mathbf{x}(0) > 0$ the dynamical model $\mathbf{x}(t+1) = A\mathbf{x}(t)$, $t = 0, 1, \dots$, produces state vectors satisfying

$$\sum_{i=t+1}^{t+h} \mathbf{x}(i) \gg \mathbf{0}, \quad (15)$$

for sufficiently large values of t .

Searching for a natural extension of the irreducibility definition to positive matrix pairs, we can try to generalize anyone of the above characterizations. For instance, we can look for a two-dimensional extension of (15), and refer to the state evolution of the 2D system (1), corresponding to an arbitrary set \mathcal{X}_0 of nonnegative local states. Working on the discrete grid $\mathbb{Z} \times \mathbb{Z}$, it seems reasonable to replace the interval $[t+1, t+h]$ appearing in (15) with some finite set $\mathcal{F} \subset \mathbb{Z} \times \mathbb{Z}$ and to define irreducible any nonnegative pair (A_1, A_2) for which a finite set $\mathcal{F} \subset \mathbb{Z} \times \mathbb{Z}$ can be found such that for every nonnegative \mathcal{X}_0 the following condition

$$\sum_{[i \ j] \in [h \ k] + \mathcal{F}} \mathbf{x}(i, j) \gg 0, \quad \forall [h \ k] \in \mathbb{Z} \times \mathbb{Z} \text{ s.t. } h + k \geq T, \quad (16)$$

holds true for some suitable positive integer T . If \mathcal{X}_0 consists of a finite number of nonzero local states, however, or no upper bound exists on the distance between consecutive nonzero local states on \mathcal{C}_0 , condition (16) can be satisfied only for $T \rightarrow +\infty$. So, in order to make our definition more consistent, we confine ourselves to *admissible* sets of initial conditions, namely to nonnegative sequences \mathcal{X}_0 such that $\sum_{\ell=h}^{h+N} \mathbf{x}(\ell, -\ell) > 0$ holds true for some $N > 0$ and every $h \in \mathbb{Z}$. Irreducibility can now be characterized as follows.

Definition A pair (A_1, A_2) of $n \times n$ positive matrices is *irreducible* if there is a finite set $\mathcal{F} \subset \mathbb{Z} \times \mathbb{Z}$ such that for every admissible set of initial conditions \mathcal{X}_0 a positive integer T can be found such that

$$\sum_{[i \ j] \in [h \ k] + \mathcal{F}} \mathbf{x}(i, j) \gg 0, \quad \forall [h \ k] \in \mathbb{Z} \times \mathbb{Z} \text{ s.t. } h + k \geq T. \quad (17)$$

An alternative description of irreducible matrix pairs can be obtained by replacing in (14) the power matrices with the Hurwitz products, and the half-line $[T, +\infty)$ with a suitable solid convex cone. In fact, intuitively speaking, a positive matrix pair (A_1, A_2) should be thought of as irreducible if there are a finite

“window” and a solid convex cone such that, independently of the window position within the cone, the sum of all Hurwitz products $A_1^r \mathfrak{W}^s A_2$ corresponding to integer pairs in the window, is strictly positive.

Proposition 4.5 *A pair of $n \times n$ positive matrices (A_1, A_2) is irreducible if and only if there are a solid convex cone $\mathcal{K}^* \subset \mathbb{R}_+^2$ and a finite set $\mathcal{F} \subset \mathbb{N}^2$ such that*

$$\sum_{[r \ s] \in [h \ k] + \mathcal{F}} A_1^r \mathfrak{W}^s A_2 \gg 0, \quad \forall [h \ k] \in \mathbb{N}^2 \text{ s.t. } [h \ k] + \mathcal{F} \subset \mathcal{K}^*. \quad (18)$$

The minimal value of the cardinality of a set \mathcal{F} for which anyone of the above equivalent conditions, (17) or (18), holds true is called the *imprimitivity index* of the pair (A_1, A_2) . This index can be related to the number of points in which the variety $\mathcal{V}(\Delta_{A_1, A_2})$ of the characteristic polynomial $\Delta_{A_1, A_2}(z_1, z_2)$ intersects the polydisc $\mathcal{P}_{r^{-1}} := \{(z_1, z_2) \in \mathbb{C}^2 : |z_1| \leq r^{-1}, |z_2| \leq r^{-1}\}$, with $r = \rho(A_1 + A_2)$ the spectral radius of $A_1 + A_2$. More precisely, it has been shown [9] that $\mathcal{V}(\Delta_{A_1, A_2})$ intersects $\mathcal{P}_{r^{-1}}$ only in (r^{-1}, r^{-1}) , and in a finite number of points $(r^{-1}e^{i\theta_1}, r^{-1}e^{i\theta_2})$ of its distinguished boundary $\mathcal{T}_{r^{-1}} := \{(z_1, z_2) \in \mathbb{C}^2 : |z_1| = r^{-1}, |z_2| = r^{-1}\}$, where the pairs (θ_1, θ_2) satisfy the congruence relations $h\theta_1 + k\theta_2 \equiv 0 \pmod{2\pi}$, for all $(h, k) \in \text{supp}(\Delta_{A, B})$. This result can be viewed as the two-dimensional analogue of the Perron-Frobenius theorem.

For sake of completeness, we mention that irreducibility property of a matrix pair admits also a graph-theoretic interpretation, based on the notion of 2D-directed graph, i.e. a directed graph with two classes of arcs, corresponding to the nonzero entries of A_1 and A_2 , respectively. For more details, the interested reader is referred to [9].

A *primitive* pair of positive matrices is an irreducible pair with unitary imprimitivity index. The following proposition illustrates an interesting set of equivalent primitivity conditions that motivate the above definition.

Proposition 4.7 *Let (A_1, A_2) be an irreducible pair of $n \times n$ positive matrices, with $\rho(A_1 + A_2) = r$. The following facts are equivalent:*

- i) (A_1, A_2) is primitive;
- ii) there exists a strictly positive Hurwitz product;
- iii) there is a solid convex cone \mathcal{K}^* in \mathbb{R}_+^2 such that for all $(h, k) \in \mathbb{N}^2 \cap \mathcal{K}^*$ the Hurwitz product $A_1^h \mathfrak{W}^k A_2$ is strictly positive;
- iv) for every admissible set of initial conditions there is a positive integer T such that $\mathbf{x}(h, k) \gg 0$ for all $(h, k) \in \mathbb{N}^2$, $h + k \geq T$;
- v) the variety $\mathcal{V}(\Delta_{A_1, A_2})$ intersects the polydisk $\mathcal{P}_{r^{-1}}$ only in (r^{-1}, r^{-1}) .

5 Concluding remarks

In this paper we have presented an outline of some results about matrix pairs, in particular nonnegative ones, and of their relevance in the analysis of 2D systems. Further research will, it is to be hoped, better clarify the connections between the spectral and combinatorial properties of a pair and the dynamics of the associated state model. An interesting question, for instance, concerns the existence of a dominating vector representing the asymptotic direction of all local states $\mathbf{x}(h, k)$ as $h + k$ goes to infinity. Preliminary results have been obtained in the scalar case, in the case of 2D Markov chains and when (A_1, A_2) is a k -commuting pair.

The positive realization problem for 2D rational functions provides a wide set of research issues involving positive matrix pairs. In particular, basing on the 1D analogue, it is reasonable to expect that the singularities of a positively realizable transfer function fulfil some regularity constraints, inherited from the Perron-Frobenius structure of the peripheral spectrum of a positive pair.

Finally, an interesting topic of research is the investigation of the structure of matrix pairs over finite fields, which arise in the synthesis of encoders and decoders for 2D convolutional codes.

References

- [1] A. Berman and R.J. Plemmons, *Nonnegative matrices in the mathematical sciences*, Academic Press, New York (NY), (1979).
- [2] M. Bisiacco, E. Fornasini, and G. Marchesini, 2D partial fraction expansions and minimal commutative realizations, *IEEE Trans. Circ. Sys.* **35** (1988), pp.1533–1538.
- [3] M. Fliess, Un outil algébrique: Les séries formelles non commutatives, *Springer Lect.Notes in Econ. Math. Sys.* **131**, (1975), pp.122–149.
- [4] E. Fornasini and G. Marchesini, Doubly indexed dynamical systems, *Math. Sys. Theory* **12**, (1978), pp.59–72.
- [5] E. Fornasini and G. Marchesini, Properties of pairs of matrices and state-models for 2D systems, In C.R.Rao, editor, *Multivariate Analysis: Future Directions*. North Holland Series in Probability and Statistics **5**, (1993), pages 131–80.
- [6] E. Fornasini, G. Marchesini, and M.E. Valcher, On the structure of finite memory and separable 2D systems, *Automatica* **29**, (1994), pp.347–350.
- [7] E. Fornasini and M.E. Valcher, Matrix pairs in 2D systems: an approach based on trace series and Hankel matrices, *SIAM J. Contr. Optim.* **33**, 4 (1995), pp.1127–1150.
- [8] E. Fornasini and M.E. Valcher, On the spectral and combinatorial structure of 2D positive systems, *Lin. Alg. Appl.* **245**, (1996), pp.223–258.

- [9] E. Fornasini and M.E. Valcher, Directed graphs, 2D state models and characteristic polynomial of irreducible matrix pairs, to appear in *Lin. Alg. Appl.*, (1997).
- [10] E. Fornasini and M.E. Valcher, Recent developments in 2D positive system theory , to appear in *J. of Applied Mathematics*, (1997).
- [11] W.P. Heath. Self-tuning control for two-dimensional processes, J.Wiley & Sons, (1994).
- [12] N.H. McCoy, On the characteristic roots of matrix polynomials, *Bull. Amer.Math.Soc.* **42**, (1936), pp.592–600.
- [13] H. Minc, *Nonnegative Matrices*, J.Wiley & Sons, New York, (1988).
- [14] T.S. Motzkin and O. Taussky, Pairs of matrices with property L, *Trans. Amer. Math.Soc.* **73**, (1952), pp.108–114.
- [15] T.S. Motzkin and O. Taussky, Pairs of matrices with property L (II), *Trans. Amer. Math.Soc.* **80**, (1955), pp.387–401.
- [16] C. Procesi, The invariant theory of $n \times n$ matrices. *Advances in Mathematics* **19**, (1976), pp.306–381.
- [17] A. Salomaa and M. Soittola, *Automata theoretic aspects of formal power series*, Springer-Verlag, (1978).
- [18] D. Shemesh, Common eigenvectors of two matrices, *Linear Algebra Appl.* **62**, (1984), pp. 11-18.
- [19] S. Vomiero, Un'applicazione dei sistemi 2D alla modellistica dello scambio sangue-tessuto, PhD thesis, (in Italian), Univ.di Padova, Italy, (1992).

Prof. Ettore Fornasini
 Prof. Giovanni Marchesini
 Dr. Maria Elena Valcher
 Dipartimento di Elettronica ed Informatica, Università di Padova,
 via Gradenigo 6/A, 35131 Padova, Italy
 phones: + 39-49-827-7605 (7610-7795)
 fax: + 39-49-827-7699
 e-mails: fornasini@dei.unipd.it, meme@dei.unipd.it