State models and asymptotic behaviour of 2D positive systems

Maria Elena Valcher and Ettore Fornasini
Dipartimento di Elettronica ed Informatica, Univ. di Padova
via Gradenigo 6a, 35131 Padova, ITALY

Abstract

Homogeneous 2D positive systems are 2D state space models whose variables are always nonnegative and, consequently, are described by a pair of nonnegative square matrices (A, B). In the paper, the properties of these pairs are discussed both in the general case and under particular assumptions like finite memory, separability and property L.
Various aspects of the positive asymptotic dynamics are considered; in particular, sufficient conditions are provided guaranteeing that the local states are eventually strictly positive. Finally, some results on the convergence of the states towards a constant asymptotic distribution are presented.

Keywords: positive systems, 2D systems, reducible/irreducible matrices, finite memory, separability, property L

1 Introduction

A positive system is a system in which the state variables are always positive (or at least nonnegative) in value. The positivity constraint arises quite naturally when modelling real systems, whose state variables represent quantities that are intrinsically nonnegative, such as pressures, concentrations, population levels, etc..
Even though some interesting topics have not been clarified yet, and constitute active areas of research, the major aspects of linear 1D positive systems are already well understood. The cornerstone of the 1D theory is a family of results on positive matrices that essentially draw on the celebrated Perron-Frobenius theorem [Berman Plemmons, 1979]. The spectral characterization of irreducible and primitive matrices allows, in particular, for a complete description of the long term performance, based on the dominant eigenvalues and the associated nonnegative eigenvectors.
In this paper we consider homogeneous positive 2D systems, described by the equation:

$$x(h+1, k+1) = A\ x(h, k+1) + B\ x(h+1, k), \tag{1.1}$$

where the doubly indexed local state sequence $x(\cdot, \cdot)$ takes values in the positive cone $\mathbf{R}^n_+ := \{x \in \mathbf{R}^n : x_i \geq 0,\ i = 1, 2, ..., n\}$, A and B are nonnegative $n \times n$ matrices, and the initial conditions are assigned by specifying the nonnegative values of the local states on the *separation set* $\mathcal{C}_0 := \{(i, -i) : i \in \mathbf{Z}\}$.

There are essentially two reasons why the investigation of homogeneous positive systems is more difficult in 2D than in 1D case. First of all, the dynamics of a 2D system (1.1) is determined by the matrix pair (A, B) and, as well known, the algebraic tools we use for studying a pair of linear transformations are not as simple and effective as those available for the analysis of a single linear transformation. In particular, a natural 2D extension of the Perron-Frobenius theorem is not immediately apparent. On the other hand, the free evolution is strongly influenced by the choice of the nonnegative initial local states and, most of all, by the support of the states sequence on $\mathcal{C}_0$.

Our aim is to explore the properties of nonnegative matrix pairs and the way they influence the asymptotic behaviour of the associated homogeneous 2D systems. More precisely, in sections 2 and 3 we consider the main features of the characteristic polynomial of nonnegative matrix pairs both in the general case and for pairs endowed with particular structures. In section 4 we analyse under which conditions on the initial local states and on the pair (A, B) the states $x(h, k)$ eventually become strictly positive. Finally, in section 5, some preliminary results on the convergence of the states towards a constant asymptotic distribution are presented.

Before proceeding, it is convenient to introduce some notation for distinguishing positive (and nonnegative) vectors and matrices. If $M = [m_{ij}]$ is a matrix (in particular, a vector), we write

i) $M \gg 0$ (M *strictly positive*), if $m_{ij} > 0$ for all i, j;

ii) $M > 0$ (M *positive or strictly nonnegative*), if $m_{ij} \geq 0$ for all i, j, and $m_{hk} > 0$ for at least one pair (h, k);

iii) $M \geq 0$ (M *nonnegative*), if $m_{ij} \geq 0$ for all i, j.

Two matrices M and N, with the same dimensions, are said to have the same *zero pattern* if $m_{ij} = 0$ implies $n_{ij} = 0$ and vice versa.
In the following developments, moreover, we shall use some terminology borrowed from the semigroup theory. Given the alphabet $\Xi = \{\xi_1, \xi_2\}$, the free monoid Ξ^* with base Ξ is the set of all words

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

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, $|wv| = |w| + |v|$ and $|1| = 0$.
$\mathbf{C}\langle\xi_1, \xi_2\rangle$ is the algebra of polynomials in the noncommuting indeterminates ξ_1 and ξ_2. For each pair of matrices $A, B \in \mathbf{C}^{n\times n}$, the map ψ defined on $\{1, \xi_1, \xi_2\}$ by the assignements $\psi(1) = I_n$, $\psi(\xi_1) = A$ and $\psi(\xi_2) = B$, 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 $\wp(\xi_1, \xi_2) \in \mathbf{C}\langle\xi_1, \xi_2\rangle$ is denoted by $\wp(A, B)$.

2 Characteristic polynomial of a nonnegative matrix pair

For a given matrix pair (A, B) we define the characteristic polynomial

$$\Delta_{A,B}(z_1, z_2) := \det(I - Az_1 - Bz_2), \tag{2.1}$$

and denote by $\mathcal{V}(\Delta_{A,B})$ the corresponding variety, i.e. the set of all (complex) solutions of the equation $\Delta_{A,B}(z_1, z_2) = 0$. Several properties of the pair (A, B), like finite memory, separability and property L, that affect the dynamical behaviour of system (1.1), can be completely described in terms of its characteristic polynomial. Furthermore, the variety $\mathcal{V}(\Delta_{A,B})$ is the keystone of the stability analysis [Fornasini Marchesini, 1980] and, in some instances [Fornasini 1990], allows to establish the existence of a vector which plays the same role as the Perron-Frobenius eigenvector in the 1D case.

Our concern in this section is to show how some constraints on the zero pattern of a nonnegative pair reflect into the structure of the variety $\mathcal{V}(\Delta_{A,B})$.

As well known [Minc, 1988], a nonnegative $n \times n$ matrix M, $n \geq 2$, is called *reducible* if there exists a permutation matrix P such that

$$P^T M P = \begin{bmatrix} M_{11} & M_{12} \\ 0 & M_{22} \end{bmatrix}, \tag{2.2}$$

where M_{11} and M_{22} are square submatrices. Otherwise M is *irreducible.* Whereas there are positions (i, j) where all powers M^ν of a reducible matrix have zero entries, this is never true for an irreducible matrix.

The zero patterns of the powers M^ν of an irreducible matrix M can exhibit different behaviours for large values of ν. If there exists an integer N such that $M^\nu >> 0$ for all $\nu \geq N$, then M is called *primitive*; otherwise there exist positive integers h and t_{ij}, $i, j = 1, 2, ..., n$, such that for all $\nu \geq t_{ij}$, $[M^\nu]_{ij}$, the (i, j)-th entry of M^ν, is positive if and only if $\nu = t_{ij} + \ell h$. The two cases admit a spectral characterization. Actually, a primitive matrix M has a simple real positive eigenvalue r, whose module is strictly greater than the module of any other eigenvalue of M. On the other hand, if M is not primitive, its spectrum includes h simple eigenvalues of maximal module, i.e. $r, re^{j\frac{2\pi}{h}}, ..., re^{j\frac{2\pi}{h}(h-1)}$, $r > 0$. The integer h is called the *imprimitivity index* of M.

The following proposition shows that, under rather mild assumptions on a nonnegative matrix pair (A, B), some properties of the variety $\mathcal{V}(\Delta_{A,B})$ strictly resemble those of the spectrum of a primitive matrix.

Proposition 2.1 Let $A > 0$ and $B > 0$ be $n \times n$ matrices, whose sum $A + B$ is irreducible, with maximal eigenvalue r. Then the variety $\mathcal{V}(\Delta_{A,B})$ intersects the polydisc

$$\mathcal{P}_{r^{-1}} := \{(z_1, z_2) : |z_1| \leq r^{-1}, |z_2| \leq r^{-1}\} \tag{2.3}$$

only in (r^{-1}, r^{-1}) and in some points of its distinguished boundary $\{(z_1, z_2) : |z_1| = r^{-1}, |z_2| = r^{-1}\}$; moreover, (r^{-1}, r^{-1}) is a regular point of the variety.

Proof Let $(\rho_1 e^{j\theta_1}, \rho_2 e^{j\theta_2})$, $\rho_i \geq 0$, $i = 1, 2$, be a point of $\mathcal{V}(\Delta_{A,B})$. Then there exists a nonzero vector $v \in \mathbf{C}^n$ such that

$$(\rho_1 e^{j\theta_1} A + \rho_2 e^{j\theta_2} B)v = v, \tag{2.4}$$

and it's clear that

$$(\rho_1 A + \rho_2 B)|v| \geq |v|, \tag{2.5}$$

where $|v|$ denotes the real vector whose i-th component is $|v_i|$, $i = 1, 2, ..., n$.
Assume now that $(\rho_1 e^{j\theta_1}, \rho_2 e^{j\theta_2})$ belongs to $\mathcal{P}_{r^{-1}}$, so that

$$\rho_1 \leq r^{-1} \qquad \text{and} \qquad \rho_2 \leq r^{-1}. \tag{2.6}$$

simultaneously hold. Then (2.5) and (2.6) imply $(A+B)|v| \geq r\,|v|$. Since $|v|$ is a positive vector and $A+B$ is an irreducible matrix with maximal eigenvalue r, then $|v|$ is a strictly positive eigenvector of $A+B$ corresponding to r [Minc, 1988]. If in (2.6) ρ_1 and ρ_2 were strictly less than r^{-1}, the strict positivity of $|v|$ and the assumption $A, B > 0$ would imply

$$|v| = r^{-1}(A + B)|v| > (\rho_1 A + \rho_2 B)|v| \geq |v|, \tag{2.7}$$

a contradiction. Therefore, if $(\rho_1 e^{j\theta_1}, \rho_2 e^{j\theta_2})$ belongs to $\mathcal{P}_{r^{-1}}$, we must have

$$\rho_1 = \rho_2 = r^{-1}. \tag{2.8}$$

To prove the second part of the proposition, namely that (r^{-1}, r^{-1}) is a regular point, we show that the partial derivatives $\partial\Delta_{A,B}/\partial z_1$ and $\partial\Delta_{A,B}/\partial z_2$ are nonzero when evaluated at (r^{-1}, r^{-1}). In fact, we have

$$\begin{aligned} \frac{\partial}{\partial z_1}\det(I - Az_1 - Bz_2) &= -\sum_{i,j=1}^{n}(-1)^{i+j}\det\Big(I - Az_1 - Bz_2)(i|j)\Big)a_{ij} \\ &= -\,\mathrm{tr}\,\Big(A^T \mathrm{adj}(I - Az_1 - Bz_2)\Big), \end{aligned} \tag{2.13}$$

where $(I - Az_1 - Bz_2)(i|j)$ denotes the $(n-1)\times(n-1)$ matrix obtained from $(I - Az_1 - Bz_2)$ by deleting the i-th row and the j-th column.
Since r is a nonzero eigenvalue of $A + B$, we have also

$$\Big(I - (A + B)r^{-1}\Big)\mathrm{adj}\Big(I - (A + B)r^{-1}\Big) = r^n I_n \det(rI - (A + B)) = 0,$$

that is, each nonzero column of $\mathrm{adj}(I - (A + B)r^{-1})$ is an eigenvector corresponding to r and thus it is either strictly positive or strictly negative.
Applying the same reasonings to A^T and B^T, the above conclusions hold also for the columns of $\mathrm{adj}(I - (A + B)^T r^{-1})$, that is for the rows of $\mathrm{adj}(I - (A + B)r^{-1})$.

Thus each row and column of adj $(I - (A + B)r^{-1})$ is either strictly positive or negative or zero, and at least one of the rows and of the columns is nonzero. It follows that adj $(I - (A + B)r^{-1})$ is either strictly positive or strictly negative and, as $A > 0$, at least one row of $A^T \mathrm{adj}(I - (A + B)r^{-1})$ is strictly positive or strictly negative. So (2.13), evaluated at $(z_1, z_2) = (r^{-1}, r^{-1})$, is nonzero. The same reasoning proves also the inequality $\partial\Delta_{A,B}/\partial z_2\,|_{(z_1,z_2)=(r^{-1},r^{-1})} \neq 0$ ■

3 Finite memory, separability and property L

When dealing with a generic (i.e. nonnecessarily positive) 2D system, some natural assumptions on the structure of the characteristic polynomial allow to single out important classes of systems, whose dynamical behaviour exhibits very peculiar, distinguishing features. As an example, finite memory systems [Bisiacco, 1985], i.e. 2D systems whose state evolution goes to zero in a finite number of steps, are characterized by the condition

$$\Delta_{A,B}(z_1, z_2) = 1. \tag{3.1}$$

On the other hand, separable systems [Fornasini Marchesini Valcher, 1993], that are usually viewed as the simplest examples of IIR 2D systems, are those which satisfy the factorization property

$$\Delta_{A,B}(z_1, z_2) = r(z_1)s(z_2), \tag{3.2}$$

for suitable polynomials $r(z_1) \in \mathbf{R}[z_1]$ and $s(z_2) \in \mathbf{R}[z_2]$.
A more general class, which encompasses both finite memory and separable systems, refers to the so-called property L [Motzkin Taussky, 1952 and 1955], of the matrix pair (A, B), and corresponds to the possibility of factorizing $\Delta_{A,B}(z_1, z_2)$ into the product of linear factors, as follows

$$\Delta_{A,B}(z_1, z_2) = \prod_{i=1}^{n} (1 - \lambda_i z_1 - \mu_i z_2). \tag{3.3}$$

The nonnegativity hypothesis introduces further constraints on the structure of the above systems, we will explore in some detail. To this purpose, we introduce the Hurwitz products of a matrix pair, that naturally arise in the analysis of the local state evolution. The Hurwitz products of two square matrices A and B are inductively defined as

$$A^i \sqcup\!\sqcup^0 B = A^i, \qquad A^0 \sqcup\!\sqcup^j B = B^j \tag{3.4}$$

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

$$A^i \sqcup\!\sqcup^j B = A(A^{i-1} \sqcup\!\sqcup^j B) + B(A^i \sqcup\!\sqcup^{j-1} B). \tag{3.5}$$

Basing on (3.4) and (3.5), we easily see that

$$A^i \sqcup\!\sqcup^j B = \sum_{\substack{w \,\in\, \Xi^* \\ |w|_1 = i, |w|_2 = j}} w(A, B).$$

Note that the sequence of local states $x(\cdot, \cdot)$ one obtains by assuming zero initial conditions on $\mathcal{C}_0$, except at $(0, 0)$, is represented by the power series

$$\begin{aligned} X(z_1, z_2) &= \sum_{h,k \geq 0} x(h, k) z_1^h z_2^k = (I - Az_1 - Bz_2)^{-1} x(0, 0) \\ &= \sum_{h,k \geq 0} A^h \sqcup\!\sqcup^k B \; x(0, 0) z_1^h z_2^k. \end{aligned} \tag{3.6}$$

As a consequence, the local state at (h,k)

$$x(h,k) = A^{h} ⧢^{k} B \; x(0,0) = \sum_{\substack{w \,\in\, \Xi^* \\ |w|_1 = h, \; |w|_2 = k}} w(A,B) \; x(0,0)$$

has to be interpreted as the sum of the elementary contributions along all paths connecting $(0,0)$ to (h,k) in the two-dimensional grid [Fornasini Marchesini, 1993].

For 2D finite memory systems there is only a finite number of nonzero Hurwitz products. This property and the positivity constraint allow for a simple characterization of the finite memory property, provided by the following proposition.

Proposition 3.1 [Finite memory] For a pair of $n \times n$ nonnegative matrices (A,B), the followings are equivalent

i) $\Delta_{A,B}(z_1,z_2) = 1$;

ii) $A + B$ is nilpotent;

iii) there exists a permutation matrix P such that $P^T AP$ and $P^T BP$ are both upper triangular matrices with zero diagonal.

Proof *i)* $\Rightarrow$ *ii)* Letting $z_1 = z_2 = z$ in $\Delta_{A,B}(z_1,z_2) = 1$, we get $\det(I - (A+B)z) = 1$, which implies the nilpotency of $A + B$.

ii) $\Rightarrow$ *iii)* Since $(A+B)^\nu = 0$, for all $\nu \geq n$, $A+B$ is reducible and, consequently, there exists a permutation matrix P_1 such that

$$P_1^T (A+B) P_1 = \begin{bmatrix} C_{11} & C_{12} \\ 0 & C_{22} \end{bmatrix}.$$

As C_{11} and C_{22}, in turn, are nilpotent, we can apply the above procedure to both diagonal blocks. By iterating this reasoning, we end up with one dimensional nilpotent diagonal blocks and, therefore, with an upper triangular matrix, as follows

$$P^T (A+B) P = P^T AP + P^T BP = \begin{bmatrix} 0 & * & * \\ & \ddots & * \\ & & 0 \end{bmatrix}. \tag{3.7}$$

As $P^T AP$ and $P^T BP$ are nonnegative, both of them are upper triangular with zero diagonal.

iii) $\Rightarrow$ *i)* Obvious ■

As a consequence of *iii)*, if we perform a permutation on the basis of the local state space, so as to reduce matrices A and B into upper triangular form, it is easy to realize that, for any initial global state on $\mathcal{C}_0$, the last t components of the local states on the separation set $\mathcal{C}_t = \{(h,k) : h + k = t\}$ are identically zero.

Remark In the general case, the finite memory condition depends on the nilpotency of all linear combinations $\alpha A + \beta B$, $\alpha, \beta \in \mathbf{C}$ [Fornasini Marchesini Valcher, 1993], whereas the nonnegativity assumption allows to consider only one linear combination $A + B$. Moreover, while for an arbitrary finite memory pair we can only guarantee that $A^i \sqcup\!\sqcup^j B = 0$, for $i + j \geq n$, in the nonnegative case all matrix products $w(A, B)$, $w \in \Xi^*$, are zero when $|w| \geq n$.

Proposition 3.2 [Separability] For a pair of $n \times n$ matrices $A > 0$ and $B > 0$, the followings are equivalent:

i) $\Delta_{A,B}(z_1, z_2) = r(z_1)s(z_2)$;

ii) tr $w(A, B) = 0$, for all $w \in \Xi^*$ such that $|w|_i > 0,\ i = 1, 2$;

iii) $w(A, B)$ is nilpotent, for all $w \in \Xi^*$ such that $|w|_i > 0,\ i = 1, 2$.

Proof *i)* ⇔ *ii)* To prove this equivalence we refer to a characterization of separability, presented in [Fornasini Marchesini Valcher, 1993], which states that (A, B) is separable if and only if

$$\mathrm{tr}(A^i \sqcup\!\sqcup^j B) = 0, \qquad \forall\ (i, j), \quad i > 0,\ j > 0. \tag{3.8}$$

As tr $(A^i \sqcup\!\sqcup^j B) = \sum_{|w|_1 = i, |w|_2 = j}$ tr $w(A, B)$, and all the words $w(A, B)$ are nonnegative, (3.8) implies *ii)*. The converse is always true.

ii) ⇔ *iii)* By assumption *ii)*, for each $w \in \Xi^*$, with $|w|_1 > 0$ and $|w|_2 > 0$, we have

$$\mathrm{tr}(w(A, B))^k = 0, \qquad k = 1, 2, ...,$$

which implies the nilpotency of $w(A, B)$.
Conversely, the nilpotency of $w(A, B)$ trivially implies that tr $w(A, B) = 0$ ■

Separable nonnegative pairs can be reduced to two different canonical forms. One is obtained by resorting to permutation matrices, i.e. to a reordering of the basis of the local state space, while the other is based on a (complex) similarity transformation, namely a more general change of basis. To construct the canonical forms we need the following Lemma:

Lemma 3.3 Let $A > 0$ and $B > 0$ constitute a separable pair; then $A+B$ is a reducible matrix.

Proof Consider any $w = \xi_{i_1} \xi_{i_2} \cdots \xi_{i_m} \in \Xi^*$, with $|w|_1 > 0$ and $|w|_2 > 0$. Because of the characterization *ii)* of separability given in Proposition 3.2, each diagonal element of $w(A, B)$ is zero, and therefore for any sequence of integers $\ell_1, \ell_2, ..., \ell_m \in \{1, 2, ..., n\}$

$$[\psi(\xi_{i_1})]_{\ell_1 \ell_2} [\psi(\xi_{i_2})]_{\ell_2 \ell_3} ... [\psi(\xi_{i_m})]_{\ell_m \ell_1} = 0. \tag{3.9}$$

As A and B are nonzero, there exist entries $[A]_{ij} > 0$ and $[B]_{hk} > 0$. If $A + B$ were irreducible, there would be integers p and q such that $[(A+B)^p]_{jh} > 0$ and $[(A+B)^q]_{ki} > 0$. Consequently

$$[\psi(\xi_{t_1})]_{j\ell_1} [\psi(\xi_{t_2})]_{\ell_1 \ell_2} ... [\psi(\xi_{t_p})]_{\ell_{p-1} h} > 0$$

and

$$[\psi(\xi_{s_1})]_{k r_1} [\psi(\xi_{s_2})]_{r_1 r_2} ... [\psi(\xi_{s_q})]_{r_{q-1} i} > 0$$

for appropriate choices of ξ_{t_ν} and ξ_{s_μ} and of the indexes ℓ_ν and r_μ. Therefore

$$[A]_{ij}\ [\psi(\xi_{t_1})]_{j\ell_1}[\psi(\xi_{t_2})]_{\ell_1\ell_2}...[\psi(\xi_{t_p})]_{\ell_{p-1}h}\ [B]_{hk}\ [\psi(\xi_{s_1})]_{kr_1}[\psi(\xi_{s_2})]_{r_1r_2}...[\psi(\xi_{s_q})]_{r_{q-1}i} > 0,$$

which contradicts (3.9) ■

Proposition 3.4 Let (A, B) be an $n \times n$ nonnegative matrix pair, then the followings are equivalent:

i) $\Delta_{A,B}(z_1, z_2) = r(z_1)s(z_2)$;

ii) there exists a permutation matrix P such that P^TAP and P^TBP are conformably partitioned into block triangular matrices

$$P^TAP = \begin{bmatrix} A_{11} & * & * & * \\ & A_{22} & * & * \\ & & \ddots & * \\ & & & A_{tt} \end{bmatrix} \quad P^TBP = \begin{bmatrix} B_{11} & * & * & * \\ & B_{22} & * & * \\ & & \ddots & * \\ & & & B_{tt} \end{bmatrix}, \tag{3.10}$$

where $A_{ii} \neq 0$ implies $B_{ii} = 0$. It entails no loss of generality assuming that the nonzero diagonal blocks in P^TAP and P^TBP are irreducible;

iii) there exists a nonsingular matrix T such that $\hat{A} = T^{-1}AT$ and $\hat{B} = T^{-1}BT$ are upper triangular matrices and $\hat{a}_{ii} \neq 0$ implies $\hat{b}_{ii} = 0$.

Proof If one of the matrices is zero, the proposition is trivially true, so we will confine ourselves to the case of A and B both nonzero.

i) ⇒ *ii)* By the previous Lemma, there exists a permutation matrix P_1 s.t.

$$P_1^T(A+B)P_1 = \begin{bmatrix} C_{11} & C_{12} \\ 0 & C_{22} \end{bmatrix}$$

and, consequently,

$$P_1^TAP_1 + P_1^TBP_1 = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{bmatrix},$$

where A_{ii}, B_{ii} and C_{ii}, $i = 1, 2$, are square submatrices. As the nonnegative matrix pairs (A_{ii}, B_{ii}) are separable, we can apply the same procedure as before to both of them. By iterating this method we end up with a pair of matrices with the structure (3.10).

ii) ⇒ *iii)* Let P be a permutation matrix that reduces A and B as in (3.10). Recalling that every square matrix is similar to an upper triangular matrix, consider the matrix $Q = \text{diag}\{Q_{11}, Q_{22}, ..., Q_{tt}\}$, where Q_{ii} are nonsingular square matrices such that $Q_{ii}^{-1}(A_{ii} + B_{ii})Q_{ii}$ is upper triangular. Then $T = PQ$ is the nonsingular matrix we are looking for.

iii) ⇒ *i)* Obvious ■

A pair of $n \times n$ matrices (A, B) is said to have property L if the eigenvalues of A and B can be ordered into two n-tuples

$$\Lambda(A) = (\lambda_1, \lambda_2, ..., \lambda_n) \quad \text{and} \quad \Lambda(B) = (\mu_1, \mu_2, ..., \mu_n) \tag{3.11}$$

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

$$\Lambda(\alpha A + \beta B) = (\alpha\lambda_1 + \beta\mu_1, \alpha\lambda_2 + \beta\mu_2, ..., \alpha\lambda_n + \beta\mu_n). \tag{3.12}$$

In other words, property L means that the spectrum of any linear combination of A and B is the linear combination of the spectra $\Lambda(A)$ and $\Lambda(B)$. Under appropriate assumptions, nonnegativity of A and B allows for some precise statements concerning the coupling of their maximal eigenvalues.

Proposition 3.5 [Property L] Let (A, B) be a nonnegative $n \times n$ matrix pair, endowed with property L w.r.t. the orderings (3.11), and assume $A + B$ irreducible.
Then there exists a unique index i such that

$$\lambda_i, \mu_i \in \mathbf{R}_+, \qquad \lambda_i \geq |\lambda_j|, \qquad \mu_i \geq |\mu_j|, \qquad j = 1, 2, ..., n,$$

and, for each $\alpha, \beta > 0$, $\alpha\lambda_i + \beta\mu_i$ is the maximal positive eigenvalue of the irreducible matrix $\alpha A + \beta B$.

Proof Denoting by $\nu_1(\alpha), \nu_2(\alpha), ..., \nu_n(\alpha)$ the eigenvalues of $\alpha A + (1 - \alpha)B$, property L implies that

$$\nu_j(\alpha) = \alpha\lambda_j + (1 - \alpha)\mu_j, \qquad j = 1, 2, ..., n. \tag{3.13}$$

Moreover, for all $\alpha \in (0, 1)$, the matrix $\alpha A + (1 - \alpha)B$, having the same zero-pattern as $A + B$, is irreducible and hence has a simple maximal eigenvalue $\nu_{\max}(\alpha)$. We aim to prove that there exists an integer i such that for all α, $\nu_{\max}(\alpha) = \alpha\lambda_i + (1 - \alpha)\mu_i$, where λ_i and μ_i are real positive eigenvalues of A and B, respectively.
Note first that the characteristic polynomial

$$\Delta_{A,B}(z_1, z_2) = \prod_{i=1}^{n}(1 - \lambda_i z_1 - \mu_i z_2). \tag{3.14}$$

belongs to $\mathbf{R}[z_1, z_2]$. So, if one factor $1 - \lambda_i z_1 - \mu_i z_2$ has not real coefficients, also $1 - \bar{\lambda}_i z_1 - \bar{\mu}_i z_2$ appears in (3.14). That amounts to say that, when a nonreal pair (λ_j, μ_j) appears in (3.11), also the conjugate pair $(\bar{\lambda}_j, \bar{\mu}_j)$ does, and hence both $\nu_j(\alpha) = \alpha\lambda_j + (1 - \alpha)\mu_j$ and $\nu_k(\alpha) = \alpha\bar{\lambda}_j + (1 - \alpha)\bar{\mu}_j$ belong to $\Lambda(\alpha A + (1 - \alpha)B)$. Moreover, $\nu_j(\alpha)$ is real if and only if $\nu_k(\alpha)$ is, and they take the same value. As $\nu_{\max}(\alpha)$, $0 < \alpha < 1$, has to be simple, it cannot coincide with any eigenvalue $\nu_j(\alpha)$ associated with a nonreal pair (λ_j, μ_j).
Therefore, an integer $j(\alpha)$ exists, possibly depending on α, such that (λ_j, μ_j) is a real pair and

$$\nu_{\max}(\alpha) = \nu_{j(\alpha)}(\alpha). \tag{3.15}$$

Because of the linear structure of (3.13), we can determine finitely many points, $\alpha_1, \alpha_2, ...,$ α_r, $0 < \alpha_1 < \alpha_2 < ... < \alpha_r < 1$, with the property that the index $j(\alpha)$ in (3.15) remains constant on each interval $(\alpha_\mu, \alpha_{\mu+1})$, $\mu = 1, 2, ..., r - 1$, and takes different values on different intervals. If r were greater than zero, $\nu_{\max}(\alpha_\mu)$, $\mu = 1, 2, ..., r$, would be a multiple eigenvalue of the irreducible matrix $\alpha A + (1 - \alpha)B$, a contradiction. So r has to be zero and $j(\alpha)$ takes in $(0, 1)$ a unique value i.

Next, we show that λ_i and μ_i are maximal eigenvalues of A and B. Suppose, for instance, that A possesses a positive eigenvalue $\lambda_h > \lambda_i$. As the eigenvalues of $\alpha A + (1-\alpha)B$ are continuous functions of α, $|\nu_h(\alpha)|$ would be greater than $|\nu_i(\alpha)|$ for all values of α in a suitable neighbourhood of 1, a contradiction.
Finally, letting $\bar{\alpha} = \alpha/(\alpha+\beta)$ and $1-\bar{\alpha} = \beta/(\alpha+\beta)$, we have that $\bar{\alpha}\lambda_i + (1-\bar{\alpha})\mu_i$ is the maximal positive eigenvalue of $\bar{\alpha}A + (1-\bar{\alpha})B = \frac{1}{\alpha+\beta}(\alpha A + \beta B)$ and, consequently, $\alpha\lambda_i + (1-\alpha)\mu_i$ is the maximal positive eigenvalue of $\alpha A + \beta B$ ■

Example 1 The pair

$$A = \begin{bmatrix} 0 & 1/2 \\ 1 & 1/2 \end{bmatrix} \qquad B = \begin{bmatrix} 2/5 & 3/10 \\ 2/5 & 4/5 \end{bmatrix}$$

is endowed with property L w.r.t. the orderings

$$\Lambda(A) = (1,\ -1/2), \qquad \Lambda(B) = (1,\ 1/5).$$

For each $\alpha > 0$ and $\beta > 0$, the maximal eigenvalue $\alpha+\beta$ of $\alpha A + \beta B$ is obtained as a linear combination of the maximal eigenvalues of A and B. Note that $A+B$ is strictly positive, and hence irreducible. When we drop the irreducibility assumption, as, for instance, with the pair

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \qquad B = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix},$$

the maximal eigenvalues of A and B are not necessarily coupled w.r.t. the orderings of the spectra, and hence do not appear in the same linear factor of the characteristic polynomial $\Delta_{A,B}(z_1, z_2)$.

4 Strictly positive asymptotic dynamics

An issue that arises quite naturally when considering the asymptotic behaviour of positive systems is that of guaranteeing that the states eventually become strictly positive vectors. For 1D positive systems

$$x(h+1) = A\ x(h), \qquad\qquad x(0) > 0,$$

the primitivity of the system matrix A [Minc, 1988] is necessary and sufficient for $x(h)$ beeing strictly positive when h is large enough.
For 2D systems described as in (1.1), we say that the state evolution eventually becomes strictly positive if there exists a positive integer T such that $x(h,k) \gg 0$ for all (h,k), $h+k \geq T$. Clearly it's impossible that every nonzero initial global state $\mathcal{X}_0 = \{x(i,-i) : i \in \mathbf{Z}\}$ produces a strictly positive asymptotic dynamics. Actually, when $\mathcal{X}_0$ includes only a finite number of nonzero states, the support of the free evolution is included in a quarter plane causal cone of $\mathbf{Z} \times \mathbf{Z}$.
As a consequence, we have to take into account not only the properties of the matrix pair (A,B), but also the zero-pattern of the nonnegative initial global state $\mathcal{X}_0$, and we

will confine our attention to global states which satisfy the following condition: there exists an integer M such that

$$\sum_{h=1}^{M} x(i+h, -i-h) > 0, \qquad \forall i \in \mathbf{Z}. \tag{4.1}$$

In other words, the maximal distance between two consecutive positive states on the separation set $\mathcal{C}_0$ is upper bounded by M.

In the sequel we will provide a set of sufficient conditions on the pair (A, B) guaranteeing a strictly positive asymptotic dynamics for all initial global states satisfying (4.1).

Proposition 4.1 Suppose that $A > 0$ and $B > 0$ are $n \times n$ positive matrices and there exists (i, j) such that $A^i \sqcup\!\sqcup^j B$ is primitive. Then, for each initial global state satisfying (4.1), there exists a positive integer T such that $x(h, k) \gg 0$ whenever $h + k \geq T$.

Proof If $A^i \sqcup\!\sqcup^j B$ is primitive, then $(A^i \sqcup\!\sqcup^j B)^p \gg 0$ for some $p > 0$, and therefore $A^{pi} \sqcup\!\sqcup^{pj} B \geq (A^i \sqcup\!\sqcup^j B)^p \gg 0$. So in the sequel we will assume that there exists (h, k) s.t. $A^h \sqcup\!\sqcup^k B$ is strictly positive. If we suppose $x(0, 0) > 0$, we have

$$x(h, k) \geq (A^h \sqcup\!\sqcup^k B)\, x(0, 0) \gg 0,$$

and, as A and B are nonzero matrices, both $x(h + 1, k)$ and $x(h, k + 1)$ are nonzero vectors. Consequently, if M denotes the maximal distance between two consecutive nonzero local states on $\mathcal{C}_0$, the maximal distance on $\mathcal{C}_{h+k+1}$ is not greater than $M - 1$. An inductive argument shows that all local states on $\mathcal{C}_{(M-1)(h+k+1)}$ are nonzero, and, recalling $A^h \sqcup\!\sqcup^k B \gg 0$, we see that all local states on $\mathcal{C}_{M(h+k+1)-1}$ are strictly positive. Consequently, we can choose $T = M(h + k + 1) - 1$ ■

Remark The existence of a primitive Hurwitz product $A^i \sqcup\!\sqcup^j B$ implies that, for a suitable $p > 0$, $(A^i \sqcup\!\sqcup^j B)^p$, and hence $(A + B)^{(i+j)p}$, are strictly positive. Therefore $A + B$ is primitive. The converse in general is not true, as shown by the following example. The pair of irreducible matrices

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{4.2}$$

has a primitive sum. However, as $B = A^2$, each Hurwitz product can be expressed as $A^h \sqcup\!\sqcup^k B = \binom{h+k}{h} A^{h+2k}$, and hence is not primitive.

Corollary 4.2 If $A > 0$ and $B > 0$ are $n \times n$ matrices and there exists a word $w \in \Xi^*$ s.t. $w(A, B)$ is primitive, then for all initial global states satisfying (4.1) the asymptotic behaviour of system (1.1) is strictly positive.

Proof If $|w|_1 = i$ and $|w|_2 = j$, then $A^i \sqcup\!\sqcup^j B \geq w(A, B)$ is primitive too and we can resort to Proposition 4.1 ■

In order to apply Corollary 4.2 above, in some cases it is enough to have at disposal a quite poor information on the zero pattern of A and B. For instance it's sufficient to know that A is primitive and $B > 0$.

The following proposition shows that, when dealing with a pair of irreducible matrices, it is enough to check whether their imprimitivity indexes are coprime.

Proposition 4.3 Assume that A and B are irreducible matrices with imprimitivity indexes h_A and h_B. If g.c.d.$(h_A, h_B) = 1$, then there exists $w \in \Xi^*$ such that $w(A, B)$ is primitive.

PROOF The proof depends on the following properties of a pair of $n \times n$ nonnegative matrices R and S.

1) If R is irreducible, with imprimitivity index h_R, and $S = \text{diag}\{S_{11}, S_{22}, ..., S_{\nu\nu}\}$, S_{ii} strictly positive, $i = 1, 2, ..., \nu$, then SRS is irreducible, and its imprimitivity index divides h_R.

2) Let R be an $n \times n$ nonnegative block matrix,

$$R = \begin{bmatrix} R_{11} & R_{12} & ... & R_{1\nu} \\ R_{21} & R_{22} & ... & R_{2\nu} \\ & & ... & \\ R_{\nu 1} & R_{\nu 2} & ... & R_{\nu\nu} \end{bmatrix}$$

whose diagonal blocks R_{ii} are square matrices. If all blocks R_{ij} are either zero or strictly positive, (i.e. $R_{ij} > 0 \Rightarrow R_{ij} \gg 0$) and $\tilde{R}$ is a $\nu \times \nu$ matrix whose (i, j)-th element $\tilde{R}_{ij}$ is non zero if and only if $R_{ij} \gg 0$, then $\tilde{R}$ is irreducible if and only if R is. When this is the case, R and $\tilde{R}$ have the same imprimitivity indexes.

The proofs of the above properties are rather lengthy but elementary, and will be omitted for sake of brevity.

If h_A and/or h_B are unitary, the result of Proposition 4.3 is obvious. So, assume that both h_A and h_B are greater than one. Let $(A_0, B_0) := (A, B)$ and suppose, without loss of generality, that A_0 is in superdiagonal canonical form [Minc, 1988]

$$A_0 = \begin{bmatrix} 0 & A_{12} & 0 & \cdots & 0 & 0 \\ 0 & 0 & A_{23} & \cdots & 0 & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & \cdots & & 0 & A_{h_{A_0}-1, h_{A_0}} \\ A_{h_{A_0} 1} & 0 & \cdots & & 0 & 0 \end{bmatrix}, \qquad h_{A_0} = h_A.$$

STEP 1 : If N is large enough, $A^{h_{A_0} N}$ is block diagonal, with strictly positive diagonal blocks. Moreover the matrix

$$\tilde{B}_1 := A^{h_{A_0} N} B_0 A^{h_{A_0} N},$$

when partitioned conformably with A_0, is constituted by blocks which are either zero or strictly positive and, by property 1), its imprimitivity index $h_{\tilde{B}_1}$ divides h_{B_0}. As property 2) implies $h_{\tilde{B}_1} \leq h_{A_0}$, and h_{A_0} and $h_{\tilde{B}_1}$ are coprime, one gets $h_{\tilde{B}_1} < h_{A_0}$.

If $h_{\tilde{B}_1} = 1$, $\tilde{B}_1$ is primitive and there exists a strictly positive power of $\tilde{B}_1$, which proves the theorem.

STEP 2 : If $h_{\tilde{B}_1} > 1$, we introduce the matrix

$$\tilde{A}_1 := A^{h_{A_0}N} A_0 \ A^{h_{A_0}N}.$$

It has the same zero block pattern of A_0 and $\tilde{B}_1$ and all its nonzero blocks are strictly positive. Therefore its imprimitivity index is $h_{\tilde{A}_1} = h_{A_0}$ and $(h_{\tilde{B}_1}, h_{\tilde{A}_1}) = 1$. The matrices A_1 and B_1, of dimension $h_{A_0} \times h_{A_0}$, obtained by substituting the zero/nonzero blocks of $\tilde{A}_1$ and $\tilde{B}_1$ by 0 and 1 respectively, have imprimitivity indices h_{A_0} and $h_{\tilde{B}_1}$. Possibly using a cogredience transformation [Minc, 1988] on both A_1 and B_1, we can assume that B_1 is in superdiagonal block canonical form

$$B_1 = \begin{bmatrix} 0 & B_{12} & 0 & \cdots & 0 & 0 \\ 0 & 0 & B_{23} & \cdots & 0 & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & \cdots & & 0 & B_{h_{B_1}-1, h_{B_1}} \\ A_{h_{B_1}1} & 0 & \cdots & & 0 & 0 \end{bmatrix}, \qquad h_{B_1} = h_{\tilde{B}_1},$$

and proceed as in step 1. Namely we consider a large N such that $B_1^{h_{B_1}N}$ is block diagonal, with strictly positive diagonal blocks, and introduce

$$\tilde{A}_2 := B_1^{h_{B_1}N} A_1 B_1^{h_{B_1}N}.$$

Now $h_{\tilde{A}_2}$ divides h_{A_1}, and $h_{\tilde{A}_2} \leq h_{B_1}$. Moreover, $h_{\tilde{A}_2} = h_{B_1}$ would imply that h_{A_1} and h_{B_1}, and hence $h_{\tilde{A}_1}$ and $h_{\tilde{B}_1}$, have a common factor. Consequently, $h_{\tilde{A}_2} < h_{B_1}$. If $h_{\tilde{A}_2} = 1$, $\tilde{A}_2$ and hence the matrix product

$$(A^{h_{A_0}N} B_0 A^{h_{A_0}N})^{h_{B_1}N} (A^{h_{A_0}N} B_0 A^{h_{A_0}N}) (A^{h_{A_0}N} B_0 A^{h_{A_0}N})^{h_{B_1}N}$$

are primitive.
If $h_{\tilde{A}_2} > 1$, we go through a new step of the procedure, and so on. Clearly, as

$$h_A = h_{A_0} > h_{\tilde{B}_1} = h_{B_1} > h_{\tilde{A}_2} = h_{A_2} > ...,$$

in a finite number of steps the procedure terminates with a matrix whose imprimitivity index is unitary, and therefore with a primitive word $w(A, B)$ ■

Example 2 The coprimeness assumption of the above proposition is by no means necessary for guaranteeing the existence of a strictly positive $w(A, B)$. Consider, for instance, the pair

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Both matrices are irreducible, with imprimitivity index 3, and it's easy to check that $w(A, B) = ABABABAB$ is strictly positive.

A somewhat complementary point of view underlies the following proposition, where sufficient conditions are given ensuring that, for all $w \in \Xi^*$, $w(A,B)$ is not primitive.

Proposition 4.4 Suppose that $A \geq 0$ and $B \geq 0$ are simultaneously triangularizable by a similarity transformation

$$T^{-1}AT = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ & & \ddots & * \\ 0 & & & \lambda_n \end{bmatrix} \qquad T^{-1}BT = \begin{bmatrix} \mu_1 & * & * & * \\ 0 & \mu_2 & * & * \\ & & \ddots & * \\ 0 & & & \mu_n \end{bmatrix}.$$

Suppose moreover that λ_i, μ_i are the maximal eigenvalues of A and B respectively, and λ_j, μ_j, $j \neq i$, satisfy $|\lambda_j| = \lambda_i$, $|\mu_j| = \mu_i$. Then, for all $w \in \Xi^*$, $w(A,B)$ is not primitive.

Proof For each word $w \in \Xi^*$, the spectrum of $w(A,B)$ is given by

$$\Lambda(w(A,B)) = (w(\lambda_1,\mu_1), w(\lambda_2,\mu_2), ..., w(\lambda_n,\mu_n)),$$

and $|w(\lambda_j,\mu_j)| = w(\lambda_i,\mu_i)$. As the spectrum of $w(A,B)$ includes at least two eigenvalues of maximal module, $w(A,B)$ cannot be primitive ■

Example 3 Matrices A and B in (4.2) commute and, therefore are simultaneously triangularizable. As the eigenvalues of both matrices have unitary modules, the assumptions of Proposition 4.4 are fulfilled. So, none of the words $w(A,B)$, $w \in \Xi^*$, is primitive, according to the remark of Proposition 4.1. When (4.1) holds, the same matrices provide also an example of a 2D positive system with irreducible matrices A and B, and a nonstrictly positive asymptotic dynamics. Actually, assuming on $\mathcal{C}_0$ a periodic pattern $x(i,-i) = x(i+3,-i-3)$, $i \in \mathbf{Z}$, with

$$x(0,0) = [\ 1\ \ 0\ \ 0\ ]^T, \quad x(1,-1) = [\ 0\ \ 1\ \ 0\ ]^T, \quad x(2,-2) = [\ 0\ \ 0\ \ 1\ ]^T,$$

one easily checks that no local state $x(h,k)$, $h+k \geq 0$, is strictly positive.

5 Further aspects of the asymptotic dynamics

The problems that will be addressed in this section concern some aspects of the two-dimensional dynamics which entail a finer analysis of the asymptotic behaviour. Indeed our interest here does not merely concentrate on nonzero patterns; it involves also the values of the local state and a qualitative description of the vectors distribution along the separation sets $\mathcal{C}_t = \{(i,j) : i+j = t\}$ as t goes to infinity.
The first problem is that concerning the zeroing of state oscillations on the separation sets $\mathcal{C}_t$, as $t \to +\infty$, when scalar positive systems are considered.

Definition 1 : A scalar (nonnecessarily nonnegative) global state $\mathcal{X}_0 = \{x(i,-i) : i \in \mathbf{Z}\}$ has (finite) mean value μ if, given any $\varepsilon > 0$, there exists a positive integer $N(\varepsilon)$ such that, for all $\nu \geq N(\varepsilon)$ and $h \in \mathbf{Z}$

$$\left| \frac{1}{\nu} \sum_{i=h}^{h+\nu-1} x(i,-i) - \mu \right| < \varepsilon. \qquad (5.1)$$

The mean value will be denoted as $\mu = \lim_{\nu\to\infty} \nu^{-1} \sum x(i,-i)$, where the summation is extended to all intervals of length ν, and the convergence is uniform w.r.t. the position of the interval along the separation set $\mathcal{C}_0$.

The following properties are straightforward consequences of Definition 1:

i) if $\mathcal{X}_0$ has mean value μ, $\mathcal{X}_0 - \mu = \{x(i,-i) - \mu : i \in \mathbf{Z}\}$ has mean value zero;

ii) if $\mathcal{X}_0$ has mean value μ, $\mathcal{X}_t = \{x(i,j), i+j = t\}$ has mean value $(A+B)^t\mu$;

iii) if $\mathcal{X}_0$ has mean value μ, then $\mathcal{X}_0$ is bounded, i.e. there exists a positive integer M such that $|x(i,-i)| < M, \quad i \in \mathbf{Z}$;

iv) the set of scalar global states constitutes a complete subspace of $\ell_\infty(\mathbf{Z})$, the space of bilateral bounded sequences.

Given a bounded scalar global state $\mathcal{X}_0$ with mean value μ, the *oscillation* and (when $\mu \neq 0$) the *oscillation rate* of $\mathcal{X}_0$ are defined as

$$\mathrm{Osc}(\mathcal{X}_0) := \sup_{i,j\in\mathbf{Z}} |x(i,-i) - x(j,-j)| \tag{5.2}$$

and

$$\mathrm{osc}(\mathcal{X}_0) := \frac{\mathrm{Osc}(\mathcal{X}_0)}{|\mu|}, \tag{5.3}$$

respectively. The following technical lemma shows that a convexity assumption on the pair (A,B) guarantees that the oscillations of the local states on the separation sets $\mathcal{C}_t$ are damped down to zero by the 2D system structure, as $t \to +\infty$.

Lemma 5.1 Assume that in the scalar 2D system (1.1) A and B are both positive, and $A + B = 1$. Then, for all global states $\mathcal{X}_0$ satisfying the mean value condition (5.1), $\mathrm{Osc}(\mathcal{X}_t) \to 0$ as $t \to \infty$.

Proof As the amplitude of the oscillations along the separation set is unaffected when a constant value is added to all initial local states, there is no loss of generality in assuming that $\mathcal{X}_0$, and hence $\mathcal{X}_t,\ t = 1,2,\ldots$, have zero mean.

Let ε be an arbitrary real number in $(0,1)$. By the zero mean assumption, there exists an integer $N_1 \geq 0$ such that, for all $\nu \geq N_1$,

$$\bar{x}^{(\nu)}(i,-i) := \frac{1}{2\nu+1} \sum_{j=-\nu}^{\nu} x(i+j,-i-j) \tag{5.4}$$

satisfy, for all $i \in \mathbf{Z}$, $|\bar{x}^{(\nu)}(i,-i)| < \varepsilon/4$.

In the sequel, we shall compare the asymptotic behaviour of (1.1) induced by the original global state $\mathcal{X}_0 = \{x(i,-i), i \in \mathbf{Z}\}$ with that induced by the global state $\bar{\mathcal{X}}_0^{(\nu)} = \{\bar{x}^{(\nu)}(i,-i), i \in \mathbf{Z}\}$.

When the initial conditions are provided by $\bar{\mathcal{X}}_0^{(\nu)}$, we get

$$|\bar{x}^{(\nu)}(t+h,-h)| \leq \frac{\varepsilon}{4} \sum_{i=0}^{t} \binom{t}{i} A^{t-i} B^i = \frac{\varepsilon}{4} \tag{5.5}$$

for all $t \geq 0$ and $h \in \mathbf{Z}$. So, all local states $\bar{x}^{(\nu)}(i,j)$ in the half plane $\{(i,j), i+j \geq 0\}$ have an absolute value less than $\varepsilon/4$ and, consequently, $\mathrm{Osc}(\bar{\mathcal{X}}_t^{(\nu)}) \leq \varepsilon/2$ for all $t \geq 0$. Moreover,

$$\bar{x}^{(\nu)}(t+h,-h) = \sum_{i=-\nu}^{t+\nu} x(i+h,-i-h)\frac{1}{2\nu+1}\sum_{\lambda=-\nu}^{\nu}\binom{t}{i-\lambda}A^{t-i+\lambda}B^{i-\lambda}, \tag{5.6}$$

where $\binom{t}{i-\lambda}$ is zero if $i-\lambda > t$ or $i-\lambda < 0$.
On the other hand, when the initial conditions are provided by $\mathcal{X}_0$, we get

$$x(t+h,-h) = \sum_{i=0}^{t} x(i+h,-i-h)\binom{t}{i}A^{t-i}B^{i}. \tag{5.7}$$

So, comparing (5.6) and (5.7), we see that the dynamics induced by $\bar{\mathcal{X}}_0^{(\nu)}$ approximates $x(\cdot,\cdot)$ on $\mathcal{C}_t$ within an error given by

$$\begin{aligned} & e(t+h,-h) := \bar{x}^{(\nu)}(t+h,-h) - x(t+h,-h) \\ = & \sum_{i\in[-\nu,-1]\cup[t+1,t+\nu]} x(i+h,-i-h)\frac{1}{2\nu+1}\sum_{\lambda=-\nu}^{\nu}\binom{t}{i-\lambda}A^{t-i+\lambda}B^{i-\lambda} \\ + & \sum_{i=0}^{t} x(i+h,-i-h)\Big[\sum_{\lambda=-\nu}^{\nu}\binom{t}{i-\lambda}\frac{1}{2\nu+1}A^{t-i+\lambda}B^{i-\lambda} - \binom{t}{i}A^{t-i}B^{i}\Big]. \end{aligned} \tag{5.8}$$

We consider separately the behaviour of the two addenda in (5.8), as ν and t go to infinity.

(i) Since $\mathcal{X}_0 \in \ell_\infty(\mathbf{Z})$, a positive M exists, such that, for all $i \in \mathbf{Z}$, $|x(i,-i)| < M$. Once ν has been fixed, there exists a positive integer N_2 such that, for all $t \geq N_2$, both $\binom{t}{\nu}A^{t-\nu}B^{\nu}$ and $\binom{t}{t-\nu}A^{\nu}B^{t-\nu}$ are less than $\varepsilon/4M(2\nu+1)$, and therefore the modulus of the first addendum in (5.8) is less than $\varepsilon/4$.

(ii) We resort to the following statement of the classical Bernoulli theorem [Cramer, 1971]: "Let $\sigma \in (0,1)$, and consider

$$\omega := \sum_{tB(1-\sigma)<i<tB(1+\sigma)} \binom{t}{i}A^{t-i}B^{i}. \tag{5.9}$$

Then the ratio $\omega/(1-\omega)$ may be made to exceed any given quantity by choosing t sufficiently large".
So, given σ, a positive N_3 exists, such that $1-\omega < \varepsilon/16M$ for all $t \geq N_3$. Moreover, as the values of $\binom{t}{i}A^{t-i}B^{i}$ at the boundaries of the interval $(tB(1-\sigma), tB(1+\sigma))$ can be made as small as convenient if t is large enough, we can assume also

$$\binom{t}{i}A^{t-i}B^{i} < \frac{\varepsilon}{2\nu+1}\frac{1}{8M}$$

when $i - \lfloor tB(1-\sigma)\rfloor = 1, 2, \ldots, \nu$ or $\lfloor tB(1+\sigma)\rfloor - i = 0, 1, \ldots, \nu - 1$. Consequently, for all $t \geq N_3$, the summation in the second addendum of (5.8), when restricted to the values of i satisfying $|i - tB| > t\sigma$, gives

$$\sum_{|i-tB|>t\sigma} x(i+b, -i-h)\Big[\sum_{\lambda=-\nu}^{\nu} \binom{t}{i-\lambda} \frac{1}{2\nu+1} A^{t-i-\lambda} B^{i-\lambda} - \binom{t}{i} A^{t-i} B^{i}\Big] < \frac{\varepsilon}{4}. \quad (5.10)$$

Finally, we look for a suitable bound for the complementary part, namely

$$\sum_{tB(1-\sigma)<i<tb(1+\sigma)} x(i+h, -i-h)\Big[\sum_{\lambda=-\nu}^{\nu} \binom{t}{i-\lambda} \frac{1}{2\nu+1} A^{t-i-\lambda} B^{i-\lambda} - \binom{t}{i} A^{t-i} B^{i}\Big]. \quad (5.11)$$

Letting $i = t(B + \delta)$, the term in square brackets can be rewritten as

$$\begin{aligned} \mathcal{T}_i \;&=\; \binom{t}{i} A^{t-i} B^{i}\Big[-1 + \frac{1}{2\nu+1}\Big(1 + \frac{(1+\frac{\delta}{B})}{(1+\frac{\delta}{A}+\frac{1}{At})} + \frac{(1-\frac{\delta}{A})}{(1+\frac{\delta}{B}+\frac{1}{Bt})} \\ &+ \frac{(1+\frac{\delta}{B})(1+\frac{\delta}{B}+\frac{1}{tB})}{(1+\frac{\delta}{A}+\frac{1}{At})+(1+\frac{\delta}{A}+\frac{2}{At})} + \ldots + \frac{(1-\frac{\delta}{A})(1-\frac{\delta}{A}-\frac{1}{At})\ldots(1-\frac{\delta}{A}-\frac{\nu+1}{At})}{(1+\frac{\delta}{B}+\frac{1}{Bt})\ldots(1+\frac{\delta}{B}+\frac{\nu}{Bt})}\Big)\Big] \end{aligned}$$

As t goes to infinity, all terms k/At and k/Bt can be neglected. Moreover, for small values of σ, $|\delta|$ is a fortiori small, and all powers $\delta^3, \delta^4 \ldots$ can be neglected w.r.t. δ^2. This gives

$$\mathcal{T}_i \cong \binom{t}{i} A^{t-i} B^{i}(-1 + \frac{1}{2\nu+1}(1 + 2\nu + \gamma\delta^2)) = \binom{t}{i} A^{t-i} B^{i} \frac{\gamma\delta^2}{2\nu+1} \quad (5.12)$$

where γ is a suitable constant. As the absolute value of (5.11) is not greater than

$$\sum_{tB(1-\sigma)<i<tB(1+\sigma)} M\binom{t}{i} A^{t-i} B^{i} \frac{|\gamma|\sigma^2}{2\nu+1} \leq \frac{M|\gamma|}{2\nu+1}\sigma^2,$$

it can be made smaller than $\varepsilon/4$ when σ is small enough. Therefore, for large values of t, we have $|e(t+h, -h)| < \varepsilon/2$ and consequently

$$\mathrm{Osc}(\mathcal{X}_t) \leq \mathrm{Osc}(\bar{\mathcal{X}}_t^{(\nu)}) + 2\ \mathrm{sup}_h |e(t+h, -h)| \leq \frac{3}{2}\varepsilon \ \blacksquare \quad (5.13)$$

The following proposition is now an immediate consequence of Lemma 5.1.

Proposition 5.2 Consider an homogeneous 2D system (1.1) with $n = 1$ (scalar local states) and $A, B > 0$. Assume moreover that the initial global state $\mathcal{X}_0$ has a mean value $\mu > 0$. Then the oscillation rate $\mathrm{osc}(\mathcal{X}_t)$ goes to zero as t goes to infinity.

PROOF Lemma 5.1 implies that the oscillation of the global state $\hat{\mathcal{X}}_t$ in the system

$$\hat{x}(h+1, k+1) = \hat{A}\ \hat{x}(h, k+1) + \hat{B}\ \hat{x}(h+1, k),$$

$\hat{A} = A/(A+B)$, $\hat{B} = B/(A+B)$, goes to zero when its initial conditions are given by $\hat{\mathcal{X}}_0 = \mathcal{X}_0$. Since we have $x(i+t,-i) = (A+B)^t \hat{x}(i+t,-i)$ and the mean value of $\mathcal{X}_t$ is $(A+B)^t \mu$,

$$\operatorname{osc}(\mathcal{X}_t) = \frac{\operatorname{Osc}(\mathcal{X}_t)}{\mu(A+B)^t} = \frac{\operatorname{Osc}(\hat{\mathcal{X}}_t)}{\mu} \tag{5.14}$$

goes to zero as $t \to \infty$ ■

If we drop the hypothesis that (1.1) is a scalar system, the qualitative description of the asymptotic dynamics is by far more interesting, and more difficult. Actually, there is a diversity of questions one may ask, concerning the shape $\mathcal{X}_t$ eventually reaches as t goes to infinity, and answers depend like enough both on the pair (A, B) and on the structure of $\mathcal{X}_0$.

There is, first of all, the question of guaranteeing that the normalized state vector $x(h,k)/||x(h,k)||$ converges towards a unique vector v as $h + k \to \infty$. That is, how can a particular direction in the local state space be recognized as the 2D analogue of a 1D dominant eigenvector? If such a direction exists, a natural issue is to analyse the properties of the scalar sequences $(||x(i+t,-i)||)_{i\in\mathbf{Z}}$ and the possibility of obtaining global states $\mathcal{X}_t$ eventually free from oscillations. Finally, the questions above can be viewed as particular instances of the more general problem of classifying the asymptotic behaviours of the global states and detecting recurrencies that underlie their limiting structure.

The results so far available deal with two rather restrictive classes of positive 2D systems, that is 2D Markov chains [Fornasini 1990] and 2D systems with commutative A and B. Further research will lead, it is hoped, to more comprehensive theorems. For sake of brevity, we discuss only some aspects of commutative 2D systems, that partly supplement the treatment of this subject presented in [Fornasini Marchesini 1993].

Lemma 5.3 Let $A > 0$ and $B > 0$ be $n \times n$ commutative matrices, whose sum $A + B$ is irreducible. Then A and B have a strictly positive common eigenvector v

$$Av = r_A v, \qquad Bv = r_B v \tag{5.15}$$

and r_A, r_B are the spectral radii of A and B, respectively.

Proof Assume first that A is irreducible, and let $v \gg 0$ be the eigenvector of A corresponding to the eigenvalue r_A, that is $Av = r_A v$. The commutativity of A and B and the assumption $B > 0$ imply $A(Bv) = r_A(Bv)$ and $Bv > 0$ respectively. Since an irreducible matrix has exactly one eigenvector [Minc 1988] in $E^n := \{x \in \mathbf{R}^n_+ : \sum_{i=1}^n x_i = 1\}$, and both v and Bv are positive eigenvectors of A, we have

$$Bv = \lambda v, \quad \lambda > 0 \tag{5.16}$$

Consequently, v is a strictly positive eigenvector of B, corresponding to its maximal eigenvalue r_B, and in (5.16) $\lambda = r_B$.
Assume now that $A + B$ is irreducible, and let $A_\varepsilon := A + \varepsilon B, B_\varepsilon := B + \varepsilon A$, where ε is an arbitrary positive real number. As A_ε and B_ε commute and are both irreducible, the first part of the proof gives, for all $\varepsilon > 0$ $A_\varepsilon v^{(\varepsilon)} = r_{A_\varepsilon} v^{(\varepsilon)}, \quad B_\varepsilon v^{(\varepsilon)} = r_{B_\varepsilon} v^{(\varepsilon)}$ where

$v^{(\varepsilon)} \gg 0$ is a common eigenvector of A_ε and B_ε, uniquely determined by the condition $v^{(\varepsilon)} \in E^n$, and $r_{A_\varepsilon}, r_{B_\varepsilon}$ are the spectral radii of A_ε and B_ε respectively.
Now eigenvalues and eigenvectors are continuous functions of the entries of the matrices. Hence $A_\varepsilon \to A, \quad B_\varepsilon \to B, \quad r_{A_\varepsilon} \to r_A, \quad r_{B_\varepsilon} \to r_B$ as $\varepsilon \to 0^+$. Moreover, there exists $v \in E^n$ such that $v^{(\varepsilon)} \to v$, and v is a common eigenvector of A and B which fulfills equations (5.15). To conclude the proof, it remains to show that the limiting vector v is strictly positive. Indeed, (5.21) gives $(A+B)v = (r_A + r_B)v$ So, v is a positive eigenvector of the irreducible matrix $A+B$, which implies $v \gg 0$ ■

We are now in a position to provide a stronger version of some results published in [Fornasini Marchesini, 1993], and summarized in the following lemma

Lemma 5.4 Suppose that in system (1.1) A, B and the initial global state $\mathcal{X}_0$ satisfy the following assumptions:
(i) A and B are positive commuting matrices
(ii) A and B have a strictly positive common dominant eigenvector v
(iii) There exists ℓ and L, both positive, such that

$$0 < \ell[1 \quad 1 \quad \dots \quad 1]^T \le x(i,-i) \le L[1 \quad 1 \quad \dots \quad 1]^T, \quad \forall i \in \mathbf{Z}. \tag{5.17}$$

Then

$$\lim_{h+k \to +\infty} \frac{x(h,k)}{||x(h,k)||} = \frac{v}{||v||} \ ■$$

Proposition 5.5 Suppose that in system (1.1)
a) A and B are primitive commuting matrices
b) there exist an integer $M > 0$ and two positive real numbers r and R such that

$$r \le [1 \quad 1 \quad \dots \quad 1] \sum_{h=1}^{M} x(i+h, -i-h) \le R, \quad \forall i \in \mathbf{Z}$$

Then

$$\lim_{h+k \to +\infty} \frac{x(h,k)}{||x(h,k)||} = \frac{v}{||v||}$$

where $v \gg 0$ is a common eigenvector of A and B.

Proof As $A+B$ is irreducible, by Lemma 5.3 there exists $v \gg 0$ that satisfies equations (5.15). The primitivity assumption guarantees that v is a dominant eigenvector of both A and B. Thus conditions (i) and (ii) of Lemma 5.4 are fulfilled. On the other hand, when N is large enough, all matrices $A^{\nu} \sqcup\!\sqcup^{N-\nu} B$, $0 \le \nu \le N$, are strictly positive. So, denoting by s_N and S_N their minimum and maximum entries

$$s_N := \min_\nu \min_{h,k} [A^{\nu} \sqcup\!\sqcup^{N-\nu} B]_{hk} > 0, \qquad S_N := \max_\nu \max_{h,k} [A^{\nu} \sqcup\!\sqcup^{N-\nu} B]_{hk} > 0$$

respectively, and assuming $N \ge M$, we have

$$\begin{aligned} x_j(i+N, -i) &= \sum_{\nu=0}^{N} \mathrm{row}_j(A^{\nu} \sqcup\!\sqcup^{N-\nu} B) x(i+N-\nu, -i+\nu-N) \\ &\ge s_N \sum_{\nu=0}^{N} [1 \quad 1 \quad \dots \quad 1] x(i+N-\nu, -i-N+\nu) \ge s_N r \qquad j = 1, 2, \dots, n \end{aligned}$$

and

$$x_j(i+N,-i) \leq S_N \sum_{\nu=0}^{N} [1 \quad 1 \quad \dots \quad 1] x(i+N-\nu, -i-N+\nu) \leq S_N N R \qquad j = 1, 2, \dots, n.$$

Therefore, for large values of N, $\mathcal{X}_N$ fulfills condition (iii) of Lemma 5.4 , with $\ell = s_N r$ and $L = S_N N R$, and the proof is complete ■

6 References

A.Berman, R.J.Plemmons, (1979), *Nonnegative matrices in the mathematical sciences*, Academic Press publ.

M.Bisiacco, (1985), *State and output feedback stabilizability od 2D systems*, IEEE Trans. Circ. Sys., vol CAS-32, pp. 1246-54

H.Cramer (1971), *Mathematical methods of statistics*, Princeton Univ. Press

E.Fornasini, (1990) *2D Markov chains*, Lin. Alg. Appl. vol 140, pp. 101-27

E.Fornasini, G.Marchesini, (1980), *Stability analysis of 2D systems*, IEEE Trans. Circ. Sys., vol CAS-27, pp. 1210-17

E.Fornasini, G.Marchesini (1993), *Properties of pairs of matrices and state models for 2D systems. Pt.1: state dynamics and geometry of the pairs*, in "Multivariate Analysis: Future Directions", C.R.Rao ed., Elsevier Sci.Publ.

E.Fornasini, G.Marchesini, M.E.Valcher (1994), *On the structure of finite memory and separable 2D systems*, Automatica, vol 30

H.Minc, (1988), *Nonnegative matrices*, J.Wiley Publ.

T.S.Motzkin, O.Taussky, (1952) Pairs of matrices with propert L, Trans. Amer. Soc., vol. 73, pp. 108-114

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