Topics in 2D Systems Theory

E.FORNASINI and G.MARCHESINI

1. Introduction

Since the early 1970s two dimensional (2D) models have been attracting the attention of the scientific community working in the area of dynamical systems theory. Several different motivations underlie the research in this field. The inspiring idea originated in the framework of two-dimensional processing (using parallel computer structures) by yhe necessity of modelling dynamical processes parametrized by two independent variables (e.g. time and one space coordinate). Beyond that, very nice system theoretical reasons exist for dealing with this kind of models, whose dynamics evolves on partially ordered sets and where the extension of "classical" systems notions does not appear a practicable road.

The conceptual and formal difficulties envisaged in the theory of 2D systems are well known and descend from the fact that most definitions and operations, which are natural in the 1D theory, have to be totally revised or abandoned when one deals with signals having support in $\mathbf{Z} \times \mathbf{Z}$.

Our aim in this contribution is to briefly discuss some research perspectives in those areas of modelling, realization and control of 2D systems where a significant set of results is already available in the literature.

2. Modelling

The class of dynamical systems called 2D systems has been introduced as the natural tool for representing the processing of discrete signals which are two-dimensional, in the sense that they are functions of two independent variables. Since the support of 2D signals is the discrete plane $\mathbf{Z} \times \mathbf{Z}$, essentially different orderings can be conceived for representing cause-effect relations and these lead to mathematical models which exhibit characteristics which are not always comparable.

In the early works [1-4] the causal structure (typically quarter plane) was a priori assumed in the model as a consequence of the choice of the product ordering in $\mathbf{Z} \times \mathbf{Z}$. In some recent works [5-7] the partial ordering definition and the induced causal structure are not a priori given, but result from the analysis of the system trajectories and the underlying causality relations. In some cases [8], inspired by the process used for discretizing partial differential equations, conditions on signals are assumed on a closed contour in $\mathbf{Z} \times \mathbf{Z}$. Thus the solution essentially relies on the evolution of a 1D system and the two-dimensional causal does not play an essential role in the computation of the response.

The theory of 2D filters [9] is the framework where models which exhibit a quarter plane causality have been initially investigated. As the input output approach is considered, 2D filters are represented by proper rational functions in two indeterminates of the following type

$$W(z_1, z_2) = \frac{\sum_{i+j\geq 1} n_{ij} z_1^i z_2^j}{1 + \sum_{i+j\geq 1} d_{ij} z_1^i z_2^j} \tag{1}$$

Some topics of 2D filter theory, such as BIBO stability and discrete circuit implementations, have deep connections with the system theoretic approach. Noteworthy, the challenging problems of determining necessary and sufficient conditions for external stability and computing the minimal number of delay elements needed in circuit implementations have not been completely solved in the general case.

The idea of associating 2D state space models with two-dimensional filters originated very naturally. However, since the beginning it appeared that the "canonical" technique based on the Nerode equivalence leads to an infinite dimensional state space [3,4] and there is not an unique way to introduce te concept of state. So, following sometimes heuristic procedures, several models have been introduced, where two different notions of state play different roles:

1. *local states* $x(h,k)$ belong to a finite dimensional vector space. They enter in the state updating equation and determine the value of the output
2. *global states* $\mathcal{X}_h = \{x(i+h, -i), i \in \mathbf{Z}\}$ provide the initial conditions on a separation set of $\mathbf{Z} \times \mathbf{Z}$. These belong to an infinite dimensional vector space, which provides an extension of the space of Nerode equivalence classes.

The most common state model with quarter plane causality is represented by the following equations [10]:

$$\begin{aligned} \mathbf{x}(h+1, k+1) =& A_1\mathbf{x}(h, k+1) + A_2\mathbf{x}(h+1, k) \\ &+ B_1\mathbf{u}(h, k+1) + B_2\mathbf{u}(h+1, k) \\ \mathbf{y}(h, k) =& C\mathbf{x}(h, k) \end{aligned} \tag{2}$$

where $\mathbf{x}(h,k) \in \mathbf{R}^n, \mathbf{u}(h,k) \in \mathbf{R}^m, \mathbf{y}(h,k) \in \mathbf{R}^p$ are the values of the local state, the input and the output at $(h,k) \in \mathbf{Z} \times \mathbf{Z}$. Since the local state at $(h+1, k+1)$ is computed by solving a first order difference equation the system (2), denoted by $\Sigma_1 = (A_1, A_2, B_1, B_2, C)$, is first order.

It has been extensively studied either in its general form (2) or under some constraints on the system matrices. The most popular particularized version of (2) is the Roesser's model [1,11], where the local state space X is the direct sum of two vector spaces X^h and X^v, and the conformably partitioned matrices of the model are constrained to have the following structure

$$A_1 = \begin{bmatrix} A_{11}^{(1)} & A_{12}^{(1)} \\ 0 & 0 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 0 \\ A_{21}^{(1)} & A_{22}^{(1)} \end{bmatrix}, B_1 = \begin{bmatrix} B_1^{(1)} \\ 0 \end{bmatrix}, B_2 = \begin{bmatrix} 0 \\ B_2^{(2)} \end{bmatrix} \tag{3}$$

Second order models are used less frequently: the typical structure of the equations is given by [3,4]:

$$\begin{aligned} \mathbf{x}(h+1,k+1) &= A_1\mathbf{x}(h,k+1) + A_2\mathbf{x}(h+1,k) + A_0\mathbf{x}(h,k) + B\mathbf{u}(h,k) \\ \mathbf{y}(h,k) &= C\mathbf{x}(h,k) \end{aligned} \tag{4}$$

In particular, Attasi's model [2] has the structure (4) with A_1 and A_2 commutative matrices and $A_1A_2 = A_0$. Although realizing only separable filters, it constitutes an interesting second order model, since the underlying theory is very close to the 1D theory.

Recently the "behaviour" approach has been extended to 2D systems. Following this theory, a 2D system is defined by a family $\mathcal{B}$ of "admissible functions" (behaviour), defined over the discrete plane. These functions are characterized by the property of belonging to the kernel of a polynomial matrix $M(z_1, z_2)$ in two variables

$$\mathcal{B} = \{\mathbf{w} = \sum_{i,j\in\mathbf{Z}} w_{ij} z_1^i z_2^j | M\mathbf{w} = 0\} \tag{5}$$

Associated with the external description provided by the behaviour $\mathcal{B}$ different "internal" representations can be given, by introducing the so called *latent* variables models. State variables constitute a particuler type of latent variables, that hold the memory of the system with respect to a notion of "past" introduced on $\mathbf{Z}\times\mathbf{Z}$. When a state description is possible, i.e. when the notions of past, present and future are allowed by the structure of $\mathcal{B}$, the behaviour is called *markovian*. Since there ins'nt any "natural" direction for the evolution in $\mathbf{Z}\times\mathbf{Z}$, the markovian property appears more general than the familiar quarter plane causality and has been exploited in the analysis of non-causal 2D dynamics [5-7].

3. Realization

In the theory of 1D systems minimal realizations of a given transfer function are reachable and observable. They are algebraically equivalent and can be canonically computed using the Nerode equivalence. This very simple picture does not fit with the 2D situation. We are not very far from the truth by saying that the minimal realization problem is the bottleneck of the entire 2D theory and that appropriate tools for investigating this problem have not yet been set up. In particular, the reachability and observability analysis, relative to local and global states,is not useful for constructing minimal realizations. Actually, local reachability and observability do not necessarily imply minimality [12]. On the other side, globally reachable and observable realizations are minimal, but the converse is not true. Is should also be noticed that in some cases (see the Example below) globally reachable and observable realizations do not exist [13].

Example 1. The transfer function $W(z_1, z_2) = z_1^2 + z_2^2$ does not admit globally reachable and observable realizations, independently of the field of the entries of the system matrices.

Several approaches to the minimal realization problem have been undertaken. One of these consists in associating an Hankel matrix with the impulse response and using a suitable version of the Ho's algorithm. It was originally conceived for scalar systems, but it can be extended to the multivariable case without any conceptual difficulty. Given a formal power series $s = \sum_{ij} s_{ij} z_1^i z_2^j$ in the commuting variables z_1 e z_2, the Hankel matrix $\mathcal{H}(s)$ is an infinite matrix with the entries indexed on the commutative semigroup of monic monomials in two indeterminates. The matrix element indexed by $(z_1^i z_2^j, z_1^h z_2^k)$ is given by the coefficient $s_{i+h,j+k}$. It has been proved [14] that,except for the case of series expansion of separable transfer functions, the rank of $\mathcal{H}(s)$ is infinite, even when s is rational. So any direct implementation of the Ho's algorithm to obtain minimal realizations of a rational transfer function becomes quite problematic.

A way to overcome this difficulty [12] descends from observing that the Hankel matrix associated with a noncommutative rational series has finite rank and allows us to use linear algorithms to construct a minimal matrix representation of the coefficients of the series. More precisely, given the alphabet $\{\xi_1, \xi_2\}$ and the noncommutative monoid $\{\xi_1, \xi_2\}^*$, consider a strictly proper rational series

$$\sigma = \sum_{w \in \{\xi_1, \xi_2\}^* \setminus \{\emptyset\}} (\sigma, w) w$$

with noncommutative indeterminates ξ_1 e ξ_2, and introduce the matrix $\mathcal{H}(\sigma)$, indexed in $\{\xi_1, \xi_2\}^*$, where the (w_1, w_2) entry is the coefficient $(\sigma, w_1 w_2)$. Ho's algorithm enables us to represent σ in the following form

$$\sigma(\xi_1, \xi_2) = C(I - A_1 \xi_1 - a_2 \xi_2)^{-1}(B_1 \xi_1 + B_2 \xi_2) \tag{6}$$

and to obtain a representation (6) where A_1 and A_2 have minimal dimension $\delta(\sigma)$. If the series expansion of the transfer function $W(z_1, z_2)$ is the commuative image of σ, the system $\Sigma_1 = (A_1, A_2, B_1, B_2, C)$ provides a realization of $W(z_1, z_2)$.

Thus a method to obtain the minimal realizations of $W(z_1, z_2)$ consists in finding the set of noncommutative rational series $\sigma(\xi_1, \xi_2)$ having $W(z_1, z_2)$ as commutative image, $\delta(\sigma)$ minimal, and then in constructing their representations (6). The algorithm for obtaining the set of such noncommutative series is intrinsically nonlinear and evidentiates that

i) the dimension of minimal realizations of $W(z_1, z_2)$ depends on the ground field of the elements of the system matrices

ii) two minimal realizations are not algebraically equivalent if they are computed applying Ho's algorithm to different noncommutative series having the same $W(z_1, z_2)$ as commutative image.

Example 2. Consider the following transfer function

$$W(z_1, z_2) = \frac{2z_1z_2}{1 + z_1^2 + z_2^2}$$

It admits a minimal realization of type (2)

$$A_1 = \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix}, A_2 = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}, B_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, B_1 = \begin{bmatrix} i \\ -i \end{bmatrix}, B_1 = \begin{bmatrix} 1 & 1 \end{bmatrix}$$

on the complex field. It can be shown that there are no realizations of type (2) on the real field having dimension two.

Example 3. The following systems

$$A_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, B_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, B_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

and

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, A_2 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, B_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, B_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

are minimal realizations of the same transfer function, but are not algebraically equivalent.

Using noncommutative power series sheds also some light on a property of 2D hidden modes, that has been recently discovered. Given an irreducible 2D transfer function $n(z_1, z_2)/d(z_1, z_2)$ and a state space realization $\Sigma = (A_1, A_2, B_1, B_2, C, D)$ of $W(z_1, z_2)$, hidden modes are the common factors of $C\text{adj}(I - A_1z_1 - A_2z_2)$ and $\det(I - A_1z_1 - A_2z_2)$.

Since in the 1D case hidden modes are associated with unreachable and/or unobservable parts, that prevent a system from being minimal, a naive extension of 1D theory would suggest that minimal 2D realizations are free of hidden modes. Actually cancellations of 2D polynomials between

$C\mathrm{adj}(I - A_1z_1 - A_2z_2)(B_1z_1 + B_2z_2)$ and the characteristic polynomial $\det(I - A_1z_1 - A_2z_2)$ are always connected with the existence of plane curves where one of the PBH controllability and reconstructibility matrices are not full rank, but this fact is not in contradiction with the minimality of the realization.

We shall sketch here an example, showing that the above intuition is false. It is based on a multistep procedure that can be summarized as follows:

i) express $W(z_1, z_2)$ as the product W_1W_2 of two irreducible transfer functions that exhibit the cancellation of some nonconstant polynomial $c(z_1, z_2)$.

ii) construct two minimal realization Σ_1 and Σ_2 of $W_1(z_1, z_2)$ and $W_2(z_1, z_2)$, respectively

iii) perform the series connection of Σ_1 and Σ_2. This provides a realization of $W(z_1, z_2)$, whose characteristic polynomial includes the factor $c(z_1, z_2)$.

The noncommutative power series σ_{12} that corresponds to the series connection of Σ_1 and Σ_2, is the product of the noncommutative power series σ_1 and σ_2 associated with Σ_1 and Σ_2 respectively. The cancellation of a commutative polynomial $c(z_1, z_2)$ in W_1W_2 needs not imply that cancellations arise in the product $\sigma_1\sigma_2 = \sigma_{12}$, when σ_1 and σ_2 are expressed via finite sums, products and inverses of noncommutative polynomials. This suggests that the series connection of Σ_1 and Σ_2 may be a minimal MR of σ_{12}, irrespective of cancellations in W_1W_2.

Example 4. Consider the transfer function $W(z_1, z_2) = (z_1 + z_2)^3 + z_2^2 + z_2$ and its factorization into

$$W(z_1, z_2) = W_1(z_1, z_2)W_2(z_1, z_2) = \left[\frac{(z_1 + z_2)^3 + z_2^2 + z_2}{1 + z_2}\right][1 + z_2]$$

Starting from a minimal realization Σ_1 of $W_1(z_1, z_2)$, given by

$$\bar{A}_1 = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}, \bar{A}_2 = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & -1 \end{bmatrix}, \bar{B}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \bar{B}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix},$$
$$\bar{C} = [1 \quad 0 \quad 0],$$

and a minimal realization Σ_2 of $W_2(z_1, z_2)$, given by $\hat{A}_1 = [0], \hat{A}_2 = [0], \hat{B}_1 = [0], \hat{B}_2 = [1], \hat{C} = [1], \hat{D} = [1]$, we compute the series connection of Σ_1 and Σ_2

$$A_1 = \begin{bmatrix} \hat{A}_1 & 0 \\ \bar{B}_1\hat{C} & \bar{A}_1 \end{bmatrix}, A_2 = \begin{bmatrix} \hat{A}_2 & 0 \\ \bar{B}_2\hat{C} & \bar{A}_2 \end{bmatrix}, B_1 = \begin{bmatrix} \hat{B}_1 \\ \bar{B}_1\hat{D} \end{bmatrix}, B_2 = \begin{bmatrix} \hat{B}_2 \\ \bar{B}_2\hat{D} \end{bmatrix},$$
$$C = [0 \quad \bar{C}]$$

The above system constitutes a minimal realization of W. The proof of this fact is rather long [15], and involves a detailed analysis of the matrix pairs A_1, A_2 in $\mathbf{C}^{3\times 3}$ that satisfy the finite memory condition $\det(I - A_1z_1 - A_2z_2) = 1$. Note that Σ_1 and Σ_2 provide minimal MR's of $\sigma_1 = (\xi_1 + \xi_2)^2(1+\xi_2)^{-1}(\xi_1+\xi_2) + \xi_2$ and $\sigma_2 = 1 + \xi_2$ respectively, and Σ provides a minimal MR of $\sigma = [(\xi_1 + \xi_2)^2(1+\xi_2)^{-1}(\xi_1+\xi_2) + \xi_2](1+\xi_2)$. No cancellation arises in the above expression because of the noncommutativity of the factors $(1+\xi_2)$ and $(\xi_1+\xi_2)$.

Every strictly proper rational functions can be realized by systems (2) and (4). In general, if we assume that the system matrices satisfy some structural constraints, only subclasses of the class of rational transfer functions can be realized.

- the transfer function of an Attasi's model is given by a separable transfer function; conversely, the whole class od separable proper rational functions can be realized using Attasi's models. It can be shown that, given a separable transfer function $W(z_1, z_2)$, the rank of its Hankel matrix $\mathcal{H}(W)$ provides the minimal dimension of the realizations having the Attasi's structure. Also, minimal realizations are algebraically equivalent and can be computed using linear algorithms based on the Ho's procedure.
- Consider now the model (4) and assume that the system matrices satisfy the constraints

$$A_0 = 0, \; A_1A_2 = A_2A_1, \tag{7}$$

In this case the class of transfer functions which can be realized coincides with the set of rational functions having series expansions of the following form

$$W(z_1, z_2) = \Big[n_0(z_1, z_2) + \sum_j \frac{n_j(z_1, z_2)}{(1 - a_{1j}z_1 - a_{2j}z_2)^{\nu_j}}\Big] z_1 z_2 \tag{8}$$

with $\deg n_j < \nu_j$ and $n_0 \in \mathbf{R}[z_1, z_2]$. The minimal realization algorithm is based on linear procedures.

- Transfer functions having series expansion (8) do not exhaust the set of all transfer functions whose denominators factorize as the product of first order polynomials. The realization of a transfer function in this set is always accomplishable by systems (2) and (4), with A_1, A_2 (and A_0) simultaneously triangularizable [17]. Algorithms for constructing minimal realizations in this class are not yet known.

4. Control

Comparing with the situation in the 1D case, we have that, as a consequence of the 2D partial ordering structure, the class of feedback schemes which can be implemented is much wider. Actually we can conceive control procedures where the values of the state and/or the output at (h,k) influence the input values at points which are not causally related to (h,k). In these cases the resulting closed loop system looses the quarter plane causality, which could be undesirable in the context of the synthesis problem we are dealing with.

A further general remark is in order about the static or dynamic nature of the feedback schemes. For the 2D systems the solution of most control problems consists in introducing causal dynamic compensators that realize state-input or output-input recursive relationships of the following form

$$\mathbf{u}(h,k) = \sum H_{ij}\mathbf{u}(h-i,k-j) + \sum K_{ij}\mathbf{x}(h-i,k-j)$$

However, as we shall see, in the case of optimal control the input at (h,h) is generated using a static noncausal control law that involves an infinite number of local states. Obviously this gives some inconveniences from the realization point of view, so that resorting to suboptimal causal dynamic compensators might be preferable.

4.1 Feedback stabilization

By definition, a 2D system $\Sigma = (A_1, A_2, B_1, B_2, C)$ is internally stable if, for any global state $\mathcal{X}_0 = \{\mathbf{x}(i,-i),\ i \in \mathbf{Z}\}$ with $\sup_i ||\mathbf{x}(i,-i)|| < \infty$, the free evolution of the local states satisfies

$$\lim_{h+k \to +\infty} \mathbf{x}(h,k) = 0. \tag{10}$$

Accordingly, a stabilizing output feedback compensator is a (possibly non strictly proper) p inputs, m outputs 2D system $\bar{\Sigma} = (\bar{A}_1, \bar{A}_2, \bar{B}_1, \bar{B}_2, \bar{C}, \bar{D})$ that makes internally stable the closed loop system resulting from the feedback interconnection of Σ and $\bar{\Sigma}$.

Both stability and stabilizability properties can be checked by analyzing the intersection of suitable polynomial varieties with the closed unit bidisk

$$\mathcal{P}_1 = \{(z_1, z_2) : |z_1| \le 1, |z_2| \le 1\}.$$

The following results have been obtained in [18,19]:

i) Σ is internally stable if an only if the characteristic polynomial

$$\det(I - A_1 z_1 - A_2 z_2) \tag{11}$$

associated with the pair (A_1, A_2) is devoid of zeros in $\mathcal{P}_1$.

ii) Σ admits a stabilizing output feedback compensator if and only if

$$[\mathcal{V}(\mathcal{R}) \cup \mathcal{V}(\mathcal{O})] \cap \mathcal{P}_1 = \emptyset \tag{12}$$

Here $\mathcal{V}(\mathcal{R})$ and $\mathcal{V}(\mathcal{O})$ denote the complex varieties of the ideals generated by the maximal order minors of the following polynomial matrices in two variables:

$$\mathcal{R} = [\, I - A_1 z_1 - A_2 z_2 \quad B_1 z_1 + B_2 z_2 \,] \tag{13}$$

and

$$\mathcal{O} = \begin{bmatrix} I - A_1 z_1 - A_2 z_2 \\ C \end{bmatrix} \tag{14}$$

The meaning of condition (12) can be fully understood by examining the analogous condition in the 1D framework. It is well known that a discrete time 1D system (A, B, C) is stabilizable via dynamic output feedback if and only if

$$[\mathcal{V}(\mathcal{R}_1) \cup \mathcal{V}(\mathcal{O}_1)] \cap \mathcal{D}_1 = \emptyset,$$

where $\mathcal{D}_1$ denotes the closed unit disk $\{z : |z| \le 1\}$, and $\mathcal{R}_1$ and $\mathcal{O}_1$ are the PBH controllability and reconstructibility matrices

$$\mathcal{R}_1 = [\, I - Az \quad Bz \,], \qquad \mathcal{O}_1 = \begin{bmatrix} I - Az \\ C \end{bmatrix}.$$

If $N_R(z) D_R^{-1}(z)$ is a right coprime MFD of the transfer matrix of (A, B, C), so that $\det D_R$ divides $\det(I - Az)$, then $[\mathcal{V}(\mathcal{R}_1) \cup \mathcal{V}(\mathcal{O}_1)]$ is the zero set of $\det(I - Az)/\det D_R$, i.e.

$$\mathcal{V}(\mathcal{R}_1) \cup \mathcal{V}(\mathcal{O}_1) = \mathcal{V}\left(\frac{\det(I - Az)}{\det D_R}\right) \tag{15}$$

This implies that the points of $[\mathcal{V}(\mathcal{R}_1) \cup \mathcal{V}(\mathcal{O}_1)]$ are naturally associated with the hidden modes of the system and 1D stabilizability reduces to have all hidden modes converging to zero.

The situation is quite different in the 2D case. Given any factor right coprime MFD $N_R(z_1, z_2) D_R^{-1}(z_1, z_2)$ of the transfer matrix $W(z_1, z_2)$ of Σ, the ideal of the maximal order minors of $[\, N_R \quad D_R \,]$ does not depend on the particular right coprime MFD and coincides with the analogous ideal of any left coprime MFD of $W(z_1, z_2)$ (see [20]). This ideal will be unambiguously denoted by $\mathcal{I}(W)$ and the points of the corresponding finite, possibly nonempty, variety $\mathcal{V}(W)$ are called "rank singularities" of Σ (or $W(z_1, z_2)$).

The variety $\mathcal{V}(W)$ turns out to be of great importance in investigating the relationship between the set of points $[\mathcal{V}(\mathcal{R}) \cup \mathcal{V}(\mathcal{O})]$ and the variety of the characteristic polynomial $\mathcal{V}(\det(I - A_1 z_1 - A_2 z_2))$. For the 2D case this relationship is more complex than (15), essentially because $\mathcal{V}(W)$ is an invariant subset of $[\mathcal{V}(\mathcal{R}) \cup \mathcal{V}(\mathcal{O})]$ with respect to the specific realization taken into account.

In [20] the following result has been obtained:

Theorem 1 *Let $W(z_1,z_2)$ be the transfer matrix of a 2D system $\Sigma = (A_1, A_2, B_1, B_2, C)$ and $N_R(z_1,z_2)D_R^{-1}(z_1,z_2)$ be a right coprime MFD of W. Then the variety of the maximal order minors of* (13) *and* (14) *is given by,*

$$\mathcal{V}(\mathcal{R}) \cup \mathcal{V}(\mathcal{O}) = \mathcal{V}(W) \cup \mathcal{V}\left(\frac{\det(I - A_1 z_1 - A_2 z_2)}{\det D_R(z_1,z_2)}\right) \tag{16}$$

Eqn. (16) shows that the set of critical points for 2D stabilizability includes both the hidden modes variety, i.e. the algebraic curve associated with the polynomial $h(z_1,z_2) := \det(I - A_1 z_1 - A_2 z_2)/\det D_R(z_1,z_2)$, and the variety of rank singularities $\mathcal{V}(W)$.

The difference between the structures of (15) and (16) has important consequences on the stabilizability of the state space realizations of input/output maps given by transfer matrices. Clearly, minimal realizations in the 1D case are stabilizable, since no hidden modes are left in the system and therefore the polynomial matrices $\mathcal{R}_1$ and $\mathcal{O}_1$ are full rank everywhere. On the other hand, as a consequence of theorem 1, the 2D stabilizability condition (12) requires that both $\mathcal{V}(h)$ and $\mathcal{V}(W)$ do not intersect $\mathcal{P}_1$. For a given transfer matrix $W(z_1,z_2)$ it is always possible to compute a realization which fulfils the requirement that $\mathcal{V}(W) \cap \mathcal{P}_1 = \emptyset$; nevertheless, if $\mathcal{V}(W) \cap \mathcal{P}_1 \neq \emptyset$, there are no stabilizable realizations of W, since the set of the rank singularities $\mathcal{V}(W)$ is independent of the realization.

Theorem 1 contains a first result about the constraints which have to be fulfilled by the closed loop polynomial variety, but it does not specify to what extent it can be modified using output feedback. The assignability of the variety has been further investigated [20] and the key result is given by the following theorem.

Theorem 2 *Let $\Sigma = (A_1, A_2, B_1, B_2, C)$ be a realization of a strictly proper transfer matrix $W(z_1,z_2)$. For any output feedback compensator $\bar{\Sigma}$, the variety of the closed loop polynomial $\bar{\Delta}(z_1,z_2)$ satisfies the inclusion*

$$\mathcal{V}(\bar{\Delta}) \supseteq \mathcal{V}(h) \cup \mathcal{V}(W)$$

Viceversa, given any algebraic curve $\mathcal{C}$ that includes $\mathcal{V}(h) \cup \mathcal{V}(W)$ and excludes the origin, there exists a compensator $\bar{\Sigma}$ such that $\mathcal{V}(\bar{\Delta}) = \mathcal{C}$.

At this point two problems naturally arise:

i) given a polynomial $c(z_1,z_2)$ in $\mathbf{R}[z_1,z_2]$, decide about the assignability of the variety $\mathcal{V}(c) := \mathcal{C}$

ii) if $\mathcal{V}(c)$ is assignable, find algorithms for realizing a compensator $\bar{\Sigma}$

The solution of the first problem consists in verifying if

$$(0,0) \notin \mathcal{C} \tag{17}$$

$$\mathcal{V}(h) \subseteq \mathcal{C} \tag{18}$$

$$\mathcal{V}(W) \subseteq \mathcal{C} \tag{19}$$

Checking (17) is trivial. Moreover, once the polynomial $h(z_1, z_2)$ has been computed, we can easily verify (18) using a linear algorithm to see if h divides $c^{\deg h}$. The condition (19) can be checked by first computing a set of generators of $\mathcal{I}(W)$ and successively exploiting them for constructing a pair of commuting matrices M_1 and M_2, with the property

$$p(z_1, z_2) \in \mathcal{I}(W) \quad \Leftrightarrow \quad p(M_1, M_2) = 0. \tag{20}$$

Thus $\mathcal{V}(W) \subseteq \mathcal{C}$ if and only if $c(M_1, M_2)$ is a nilpotent matrix. For the construction of M_1 and M_2 the reader is referred to [21].

It remains to show how to compute the polynomial h and a set of generators for $\mathcal{I}(W)$ starting from the system matrices A_1, A_2, B_1, B_2, C.

For this, let

$$[C\mathrm{adj}(I - A_1 z_1 - A_2 z_2)(B_1 z_1 + B_2 z_2)]\,[I_m \det(I - A_1 z_1 - A_2 z_2)]^{-1} = \bar{N}\bar{D}^{-1}$$

be a MFD of the transfer matrix $W(z_1, z_2)$. The generators set can be obtained by evaluating the maximal order minors $m_1, m_2, \ldots m_t$ in $[\,\bar{N}^T \quad \bar{D}^T\,]$ and then by eliminating their g.c.d. $d(z_1, z_2)$. Thus h is given by

$$h = \frac{\det(I - A_1 z_1 - A_2 z_2)}{\det D_R} = \frac{\det(I - A_1 z_1 - A_2 z_2) d(z_1, z_2)}{\det \bar{D}}$$

As far as the second problem is concerned, suppose that a variety $\mathcal{C} = \mathcal{V}(c)$ that fulfils conditions (17)-(19) has been given, and suppose we want to synthesize a compensator $\bar{\Sigma}$ that produces a closed loop polynomial $\bar{\Delta}$ whose variety is $\mathcal{C}$. The procedure can be summarized as follows:

- Evaluate a right coprime MFD $N_R D_R^{-1}$ of W. This can be performed by using the primitive factorization algorithm [11] or other algorithms that do not require primitive factorizations [22].
- Compute the maximal order minors $m_1, m_2, \ldots m_t$.
- Compute an integer r and a Gröbner basis $g_1, g_2, \ldots g_w$ such that $c^r = \sum_i m_i g_i$. A technique for performing this step has been presented in [21].
- Solve the Bézout equation $c^r I_m = X D_R + Y N_R$.
- Use the realization algorithm given in [20] for computing a coprime realization of $X^{-1} Y$.

4.2 Optimal control

The compensator synthesis procedure illustrated in section 2 enables us to obtain any preassigned variety of the closed loop polynomial, provided the constraints specified by Theorem 2 are satisfied. However the pole placement design essentially affects the asymptotic behaviour and exhibits a poor control of the short term system response. In this section we shall tackle the problem from a different point of view and outline a state feedback synthesis procedure based on the minimization of a quadratic cost functional J. The system which constitutes the end result of the optimal design is not merely internally stable, but satisfies additional requirements on the state and input evolutions that are summarized by J.

Assume that the initial global state (10) is an ℓ_2 sequence and consider the cost functional

$$J = \sum_{h+k\geq 0} \mathbf{x}^T(h,k)Q\mathbf{x}(h,k) + \mathbf{u}^T(h,k)R\mathbf{u}(h,k) \tag{21}$$

with $Q \geq 0$ and $R > 0$. Theorems 3 and 4 below provide a complete solution to the following optimal control problems:

1. given $\mathcal{X}_0$, derive conditions for the existence and the uniqueness of an input function $\mathbf{u}(\cdot,\cdot)$ that minimizes J;
2. whenever these conditions are satisfied, explicitly compute the optimal input function and the corresponding value of J.

The following theorem [23] shows that the existence and the uniqueness of a stabilizing optimal control reduce to rank conditions on polynomial matrices in two variables.

Theorem 3 *For any $\mathcal{X}_0 \in \ell_2$, there exists an ℓ_2- solution of the optimal control problem. (i.e. an input $\mathbf{u}(\cdot,\cdot)$ in ℓ_2 such that the corresponding state evolution $\mathbf{x}(\cdot,\cdot)$ is in ℓ_2 and the value of J is minimized) if and only if the polynomial matrix* (12) *has full rank on the set*

$$\mathcal{M} = \{(z_1,z_2) \in \mathbf{C}\times\mathbf{C} : |z_1| = |z_2| \leq 1\}$$

and the polynomial matrix

$$\begin{bmatrix} I - A_1z_1 - A_2z_2 \\ Q \end{bmatrix} \tag{22}$$

has full rank on the unit torus

$$\mathcal{T}_1 = \{(z_1,z_2) \in \mathbf{C}\times\mathbf{C} : |z_1| = |z_2| = 1\}$$

.

When the conditions of Theorem 3 are fulfilled, the optimal control law is obtained via an algebraic Riccati equation whose coefficients are polynomial matrices in one variable.

Theorem 4 *Assume that (12) is full rank on $\mathcal{M}$ and (22) is full rank on $\mathcal{T}_1$. Then the following algebraic Riccati equation (ARE-z)*

$$P(z) = Q + (A_1^T + A_2^T z^{-1})P(z)(A_1 + A_2 z) - (A_1^T + A_2^T z^{-1})P(z)(B_1 + B_2 z)\cdot$$
$$\cdot \left[R + (B_1^T + B_2^T z^{-1})P(z)(B_1 + B_2 z)\right]^{-1}(B_1^T + B_2^T z^{-1})P(z)(A_1 + A_2 z)$$

in the unknown matrix $P(z)$ admits a unique solution in an open annulus that includes the unit circle γ_1, with the following properties:

1. $P(e^{j\omega}) = P^*(e^{j\omega}) \geq 0, \quad \forall \omega \in [0, 2\pi]$
2. *the matrix*

$$K(z) := -\left[R + (B_1^T + B_2^T z^{-1})P(z)(B_1 + B_2 z)\right]^{-1}\cdot$$
$$\cdot (B_1^T + B_2^T z^{-1})P(z)(A_1 + A_2 z)$$

is analytic in an open annulus that includes γ_1.

The coefficients of its Laurent series expansion

$$K(z) = \sum_{i=-\infty}^{+\infty} K_i z^i \tag{23}$$

provide a stabilizing feedback law $u(h,k) = \sum_{i=-\infty}^{+\infty} K_i x(h+i, k-i)$ and the minimum value of J is given by

$$J_{\min} = \frac{1}{2\pi}\int_0^{2\pi} \hat{\mathcal{X}}_0^*(\omega)P(e^{j\omega})\hat{\mathcal{X}}_0(\omega)\, d\omega \tag{24}$$

where $\hat{\mathcal{X}}_0(\omega)$ denotes the Fourier transform of the ℓ_2-sequence $\mathcal{X}_0$.

The proofs of theorems 3 and 4 are quite long and the interested reader is referred to [23]. We shall give here two examples. The first one shows how the solvability conditions based on the rank of (12) and (22) reflect into the analytic structure of $P(e^{j\omega})$. The second one gives an idea of some difficulties involved in the computation of the feedback matrices K_i, even in dimension 1.

Example 5. Assume in (2) $m = n = 1, A_1 = B_1 = B_2 = 1, A_2 = -1$ and in (21) $R = Q = 1$. In this case the solution of (ARE-z) can be obtained in closed form as

$$P(z) = \frac{-1 \mp \sqrt{5 + 2z + 2z^{-1}}}{-2(2 + z + z^{-1})}$$

Letting $z = e^{j\omega}$, the first solution is negative and the second is given by

$$P(e^{j\omega}) = \frac{1}{-1 + \sqrt{5 + 4\cos\omega}}$$

Since we are looking for nonnegative solutions, we consider only the second one, which is positive for $\omega \in [0, 2\pi]$, except at $\omega = \pi$, where $P(e^{j\omega})$ diverges. Actually, this is not surprising because (12) is not full rank at $(1/2, -1/2) \in \mathcal{M}$. Hence for some initial global state in ℓ_2 a stabilizing optimal feedback law does not exist.

Example 6 Let's change only the sign of A_2 in the previous example. In this case the unique positive definite solution of (ARE-z) along γ_1 is given by

$$P(e^{j\omega}) = \frac{2}{\sqrt{1 + 16(1 + \cos\omega)^2} - (3 + 4\cos\omega)}$$

and the corresponding feedback matrix is

$$K(e^{j\omega}) = \frac{4(1 + \cos\omega)}{1 + \sqrt{1 + 16(1 + \cos\omega)^2}}$$

Since $P(e^{j\omega})$ attains its minimum value at $\omega = \pi$, we have

$$J_{\min}(\mathcal{X}_0) = \frac{1}{2\pi}\int_0^{2\pi} P(e^{j\omega})||\hat{\mathcal{X}}_0(\omega)||^2\, d\omega \geq ||\mathcal{X}_0||^2 P(e^{j\pi})$$

and $J_{\min}(\mathcal{X}_0)$ can be made arbitrarily close to the lower bound, if we consider initial global states whose spectral content is concentrated in a narrow neighbourhood of $e^{j\pi}$.
The computation of the values of K_h depends on the evaluation of the following integrals

$$K_h = \frac{1}{2\pi}\int_0^{2\pi} \frac{4(1 + \cos\omega)\cos(\omega h)}{1 + \sqrt{1 + (1 + \cos\omega)^2}}\, d\omega, \qquad h = 0, 1, \ldots$$

When infinitely many K_hs are different from zero, the optimal feedback law cannot be implemented by a finite dimensional device and the resulting closed loop system is an half plane causal 2D system, whose updating equation requires in principle to cope with an infinite dimensional state vector. To overcome the storage and computation problems, it seems natural to investigate whether, in case (2) satisfies the rank condition of theorem 3, the stabilizing feedback matrix could be constrained to have all elements in the bilateral polynomials ring $\mathbf{R}[z, z^{-1}]$. An obvious advantage of this control law is that $\mathbf{u}(h, k)$ would only depend on a finite number of local states, which makes the closed loop system weakly causal [24].

The question above can be positively answered. Actually, the bilateral polynomial matrix

$$K_N(z) = \sum_{i=-N}^{N} K_i z^i$$

obtained by truncation of the Laurent series (23) gives a stabilizing state feedback, provided that N is large enough. Even more, when N diverges and ℓ_2 initial states are considered, the corresponding cost functional J_N asymptotically converges to the minimum value $J_{\min}$.

REFERENCES

1. R.P.Roesser, *A discrete state-space model for linear image processing,* IEEE Trans. Aut. Contr., vol. AC-20, pp.1-10, 1975.
2. S.Attasi, *Systèmes linéaires homognes à deux indices* , Rap. Laboria, 31, 1973.
3. E.Fornasini and G.Marchesini, *Algebraic realization theory of two-dimensional filters,* in Variable Structure Systems (A.Ruberti, R.Mohler eds.), L.N. in Econ. and Math. Sys., vol. 111, Springer, 1974.
4. E.Fornasini and G.Marchesini, *State space relization theory of two-dimensional filters,* IEEE Trans. Aut. Contr., vol. AC-21, pp. 484-92, 1976.
5. P.Rocha and J.C.Willems, *State for 2D systems* Lin. Alg. Appl., Special Issue on Linear Systems, vol.122, 1989.
6. P.Rocha, *Controllability of 2D systems,* IEEE Trans. Aut. Contr, vol AC-36, pp.413-423.
7. E.Fornasini, P.Rocha and S.Zampieri, *State realization of 2D finite dimensional autonomous systems,* Proc. of ECC'91, pp.715-19, Grenoble, 1991.
8. B.Lévy, M.Adams and A.Willsky *Solution and linear estimation of 2D nearest neighbour models,* IEEE Proc., vol78, n.4, pp627-41, Apr. 1990.
9. J.L.Shanks, S.Treitel and J.H.Justice, *Stability and synthesis of two dimensional recursive filters* , IEEE Trans. Audio El., vol AU-20, pp.115-28, 1972.
10. E.Fornasini and G.Marchesini, *Doubly indexed dynamical systems: state space models and structural properties* , Math. Sys. Th., vol.12, pp. 1978.
11. S.Kung, B.Lévy and T.Kailath, *New results in 2-D systems theory. Part I : 2D polynomial matrices, factorization and coprimeness; Part II : 2D state-space models, realizations and the notions of controllability, observability and minimality* , Proc. IEEE, vol 65, n.6, 1977.

12. E. Fornasini, *On the relevance of noncommutative power series in spatial filters realization* , IEEE Trans. Circ. Sys., vol. CAS-25, pp. 290-99, 1979.
13. E.Fornasini and G.Marchesini, *Global properties and duality in 2D systems* , Sys. Contr. Letters, vol.2, pp.30-38, 1982.
14. M.Fliess, *Matrices de Hankel*, J. Math. Pures Appl., vol. 53, pp.197-224, 1974.
15. E.Fornasini and S.Zampieri, *A note on the state space realization of 2D FIR transfer functions*, Systems and Control Letters, vol.16, n.2, pp.17-22, 1991.
16. M.Bisiacco, E.Fornasini and G.Marchesini *2D systems feedback compensation: an approach based on commutative linear transformations*, Lin. Alg. Appl., vol. 121, pp. 135-50, 1989.
17. M.Bisiacco, E.Fornasini and G.Marchesini, *Some connections between algebraic properties of pairs of matrices and 2D realizations*, in Analysis and Optimization of Systems (A.Bensoussan, J.L.Lions eds.), Springer L.N. Contr. Inf. Sci., vol 63, part 2, pp. 117-29, 1984.
18. E.Fornasini and G.Marchesini, *Stability analysis of 2D systems*, IEEE Trans. Circ. Sys., vol. CAS-27, pp.1210-17, 1980.
19. M. Bisiacco, *On the state reconstruction of 2D systems*, Systems and Control Letters, vol. 5, pp. 347-353, 1985.
20. M. Bisiacco, E. Fornasini and G. Marchesini, *Dynamic regulation of 2D systems: a state space approach*, Lin. Alg. Appl., special issue on Linear Systems, vol. 122-24, pp. 195-218, 1989.
21. M. Bisiacco, E.Fornasini and G.Marchesini *2D Systems feedback compensation: an approach based on commutative linear transformations*, Lin. Alg. Appl.s, vol. 121, pp.135-50, 1989.
22. Y.S.Lai and C.T.Chen, *Coprime fraction computation of 2D rational matrices*, IEEE Trans. on Aut. Contr., vol. AC-32, pp.333-36, 1987.
23. M. Bisiacco and E. Fornasini, *Optimal control of 2D systems*, SIAM Jour. Contr. and Optimiz., vol.28, pp.582-601, June 1990.
24. J. P. Guiver and N.K. Bose, *Causal and weakly causal 2D filters with applications in stabilization*, in "Multidimensional Systems Theory", pp. 52-100, Reidel Publ. Co., 1985.

Ettore Fornasini and Giovanni Marchesini
Dipartimento di Elettronica e Informatica
Università degli Studi di Padova
via Gradenigo 6/A, 35131 Padova, Italy