Multidimensional systems with finite support behaviors: signal structure, generation and detection

Ettore Fornasini Maria Elena Valcher *

March 3, 2014

Abstract

The main features of finite multidimensional behaviors are introduced as properties of the trajectories supports, and connected with the polynomial matrices adopted for their description.
Observability and local detectability are shown to be equivalent to the kernel representation of a behavior via some parity check matrix H^T.
The main properties of locally undetectable behaviors as well as their connections with the notion of constrained variables are investigated, and a general representation result for finite support behaviors is derived.
The input/output representation via generator matrices is finally discussed, and some connections between matrix primeness and the constraints every trajectory imposes on the support of the corresponding input are analysed.

multidimensional systems, behavior theory, polynomial matrices, parity checks, observability, local detectability/undetectability
93C35, 93B25, 93A25, 94B10

1 Introduction

Behavior theory is the study of the trajectories a dynamical system produces according to its evolution laws. It originated in the analysis of 1D systems, and was developed in a complete and useful form by J.C.Willems in the last two decades. In a series of papers [17, 18, 19], Willems provided a thorough description of the ways a system interacts with its environment, as well as a clear conceptual apparatus for analysing and identifying the attributes a family

*Dipartimento di Elettronica ed Informatica, Università di Padova - via Gradenigo 6a, 35131 Padova, ITALY - phone: +39-49-827-7605, fax: +39-49-827-7699 - (e-mail: fornasini@dei.unipd.it, meme@dei.unipd.it).

of trajectories possibly exhibits. Perhaps the most important of the notions he introduced is external controllability, which displays the way memory function operates and hence constitutes a powerful tool for obtaining state space models of infinite behaviors.

Recently, purely ring-theoretic extensions of Willems theory have been obtained by F.Fagnani, S.K.Mitter and S.Zampieri in [2, 21]. The new field of research is relevant for the investigation of many classes of systems, and makes it quite clear how several concepts of behavior theory depend on the nature of the underlying algebraic structures. Nevertheless, a certain continuity with Willems former results is apparent, if for no other reason that the analysis is normally developed and thought of in a standard one-dimensional time domain.

A second stage in the development of behavior theory, initiated by P.Rocha and J.C.Willems at the end of the eighties [12, 13], resulted in the absorption of two-dimensional (2D) signals into the theory. The analysis of 2D behaviors has led to new insights in the classical theory of 2D systems and to new investigations of Laurent polynomial operators, centering around the algebra of matrix pairs and various primeness conditions for polynomial matrices.

Another development in behavior theory is G.D.Forney's work on the behavioral approach to group systems [7]. Like the original work on minimal bases of rational spaces [6], Forney's papers find several applications in the theory of convolutional codes. At the same time, however, they draw on duality theory, and suggest new problems on observability and memory span. Also, they emphasize the importance of topological groups in behavior and coding theory.

During the last few years, there has been an increasing interest in convolutional coding of multidimensional (nD) data [3, 15], motivated to large extent by the possibility of investigating code performances and properties in a behavior context. Also, multidimensional convolutional codes have been a fruitful source of problems and conjectures, both in polynomial modules algebra and in signal processing of discrete data arrays [14].

The aim of this paper is to present, in as self contained a manner as possible, the behavior theory of finite support multidimensional signals. The finite support assumption is motivated by the fact that in several applications the independent variables represent spatial coordinates and the phaenomenon one aims to model regards only a finite region of the space. So, infinite behaviors, which constitute the core of Willems theory, here are only marginally touched on. A detailed analysis of the main connections between finite and infinite nD signals falls within the scope of duality theory, and as far as the 2D case is concerned has been carried on in [15].

Not intending to be inclusive of all aspects of the subject, we have concentrated on what seem to be the most interesting topics to be investigated, and have included some preliminary material, as necessary for the discussion. Particular attention has been devoted to the supports of the signals, and to certain elementary operations (restriction, extension and concatenation) which have a concrete meaning from the signal processing standpoint. Actually, several "internal" properties of a behavior have been introduced in terms of these operations, and expressed as possibilities of "cutting and pasting together" pieces of different trajectories into a new one.

As each of these features mirrors into a particular polynomial matrix representation, an explicit link between the parity checks description of an nD behavior and the concept of observability is derived; indeed, the support of the parity check matrix measures the range of action of the system laws and provides useful bounds on the region where parity checks apply when detecting if some signal is a legal sequence.

The trajectories of an observable behavior can be expressed as the solutions of a system of multidimensional difference equations, and hence can be recognized by means of local testing procedures. Locally undetectable behaviors, instead, exhibit opposite properties, because every finite signal can be completed into a legal trajectory and no local recognition procedure can be successfully implemented.
Interestingly enough, these two classes of behaviors allow every finite behavior to be described via intersection operations.

A point of view somewhat complementary to detection calls for an input/output analysis of the way behavior trajectories are generated, and the supports of the trajectories are related to the corresponding inputs. This problem appears particularly relevant when the behavior sequences are injectively generated, and hence a given trajectory is produced by a unique input. Although no general statement can be made on the way these supports are related, specific assumptions on the structure of the generating matrices allow to uniformly confine the support of each input signal into a suitable extension of the support of the associated output trajectory.

The use of nD Laurent polynomial (L-polynomial) matrices is pervasive all over the paper; no attempt has been made, however, to give a complete account of their algebraic properties. For these we may refer the reader to recent books [1] and articles [4, 16, 20] dealing with that part of abstract algebra. A certain attention, however, has been paid to the analysis of the supports of nD L-polynomial vectors, and some results obtained in this context seem to be original.

The paper is organized as follows. The first part introduces the basic definitions and properties of nD finite behaviors; in particular, operations which involve only the supports are sufficient to define the notions of (external) controllability and observability. While controllability is well-estabilished [7, 13], and in the context of finite support signals it follows from linearity and shift invariance, the observability definition we will adopt comes from duality issues, and is fully justified when a parity checks description of the behavior trajectories is adopted. Actually, as shown in Section 3, an observable behavior is characterized by a finite set of parity checks one has to apply in order to recognize its trajectories. This result allows observable behaviors to be identified with kernels of polynomial matrix operators or, in more abstract terms, as maximal submodules of given rank in the module of all finite support signals.

In Section 4 the notions of unconstrained variables and locally undetectable behaviors are introduced. A general representation result is then provided, showing that every finite behavior can be expressed as the intersection of an observable and a (generally not unique) locally undetectable behavior.

In the last part of the paper we develop the theory of input-output generation of nD behaviors, and present some relevant connections between support conditions on the input-output pairs and primeness requirements on the generator matrices.

2 Finite support behaviors: preliminary definitions and basic properties

Let $\mathbb{F}$ be an arbitrary field and denote by $\mathbf{z}$ the n-tuple $(z_1, z_2, ..., z_n)$, so that $\mathbb{F}[\mathbf{z}]$ and $\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]$ are shorthand notations for the polynomial and the Laurent polynomial (L-polynomial) rings in the indeterminates $z_1, ..., z_n$, respectively. For any sequence $\mathbf{w} = \{\mathbf{w}(\mathbf{h})\}_{\mathbf{h}\in\mathbb{Z}^n}$, taking values in $\mathbb{F}^p$, the *support* of $\mathbf{w}$ is the set of points where $\mathbf{w}$ is nonzero, i.e., $\mathrm{supp}(\mathbf{w}) := \{\mathbf{h} = (h_1, h_2, ..., h_n) \in \mathbb{Z}^n : \mathbf{w}(\mathbf{h}) \neq 0\}$. Also, $\mathbf{w}$ can be represented via a formal power series

$$\sum_{h_i\in\mathbb{Z}} \mathbf{w}(h_1, h_2, ..., h_n)\ z_1^{h_1} z_2^{h_2} \cdots z_n^{h_n} = \sum_{\mathbf{h}\in\mathbb{Z}^n} \mathbf{w}(\mathbf{h})\ \mathbf{z}^{\mathbf{h}},$$

where $\mathbf{h}$ stands for the n-tuple $(h_1, h_2, ..., h_n)$ and $\mathbf{z}^{\mathbf{h}}$ for the term $z_1^{h_1} z_2^{h_2} ... z_n^{h_n}$. On the other hand, power series can be viewed as representing vectors with entries in $\mathcal{F}_\infty := \mathbb{F}^{\mathbb{Z}^n}$, thus setting a bijective map between nD sequences taking values in $\mathbb{F}^p$ and formal power series with coefficients in $\mathbb{F}^p$. This allows us to identify nD sequences with the associated power series, in particular, finite support nD signals with L-polynomial vectors, and to denote both of them with the same symbol $\mathbf{w}$. Sometimes, mostly when a power series $\mathbf{w}$ is obtained as a Cauchy product, it will be useful to denote the coefficient of $\mathbf{z}^{\mathbf{h}}$ in $\mathbf{w}$ as $(\mathbf{w}, \mathbf{z}^{\mathbf{h}})$.

Linear operators on the sequence space are represented by appropriate matrices with elements in $\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]$, whose primeness features find a counterpart in terms of properties of the associated operators. The main primeness notions which arise in the nD context are the following:

An L-polynomial matrix $G \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^{p\times m}$, $p \geq m$, is

- unimodular if $p = m$ and $\det G$ is a unit in $\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]$, i.e., $\det G = c\mathbf{z}^{\mathbf{h}}$ for some nonzero $c \in \mathbb{F}$ and some $\mathbf{h} \in \mathbb{Z}^n$;

- right factor prime (rFP) if in every factorization $G = \bar{G}T$, with $\bar{G} \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^{p\times m}$ and $T \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^{m\times m}$, T is a unimodular matrix;
- right minor prime (rMP) if its maximal order minors have no common factors;
- right variety prime (rVP) if the ideal $\mathcal{I}_G$, generated by its maximal order minors, includes (nonzero) L-polynomials in $\mathbb{F}[z_i, z_i^{-1}]$ for every $i = 1, 2, ..., n$;
- right zero prime (rZP) if the ideal $\mathcal{I}_G$ is the ring $\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]$ itself.

The support of a matrix $G \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^{p\times m}$ is the union of the supports of its elements.

An nD (finite) *behavior* $\mathfrak{B}$ *with* p *components* is a set of finite support signals (trajectories) taking values in $\mathbb{F}^p$ and endowed with the following properties:

(L) [Linearity] *If* $\mathbf{w}_1$ *and* $\mathbf{w}_2$ *belong to* $\mathfrak{B}$*, then* $\alpha\mathbf{w}_1 + \beta\mathbf{w}_2 \in \mathfrak{B}$*, for all* α*,* β *in* $\mathbb{F}$*;*

(SI) [Shift-Invariance] $\mathbf{w} \in \mathfrak{B}$ *implies* $\mathbf{v} = \mathbf{z}^{\mathbf{h}}\mathbf{w} \in \mathfrak{B}$ *for every* $\mathbf{h} \in \mathbb{Z}^n$*, i.e.,* $\mathfrak{B}$ *is invariant w.r.t. the shifts along the coordinate axes in* $\mathbb{Z}^n$*.*

As every nD behavior $\mathfrak{B}$ can be viewed as an $\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]$-submodule of $\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p$, which is a Noetherian module [10], $\mathfrak{B}$ is finitely generated, i.e., there exists a finite set of column vectors $\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_m$ in $\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p$ such that

$$\mathfrak{B} \equiv \left\{\sum_{i=1}^{m} \mathbf{g}_i u_i : u_i \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]\right\} = \{\mathbf{w} = G\mathbf{u} : \mathbf{u} \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^m\} =: \mathrm{Im}G. \quad (2.1)$$

The L-polynomial matrix $G := \mathrm{row}\{\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_m\}$ is called *generator matrix* of $\mathfrak{B}$.

$G_1 \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^{p\times m_1}$ and $G_2 \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^{p\times m_2}$ are generator matrices of the same behavior if and only if there exist $P_1 \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^{m_1\times m_2}$ and $P_2 \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^{m_2\times m_1}$ such that $G_1P_1 = G_2$ and $G_2P_2 = G_1$. Consequently, G_1 and G_2 have the same rank r over the field of rational functions $\mathbb{F}(\mathbf{z})$. Being an invariant w.r.t. all generator matrices of $\mathfrak{B}$, r is called the *rank of* $\mathfrak{B}$. It somehow represents a complexity index of the behavior, as r independent trajectories can be found in $\mathfrak{B}$, while $r + 1$ trajectories $(\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_{r+1})$ always satisfy an autoregressive equation $\mathbf{w}_1p_1 + \mathbf{w}_2p_2 + ... + \mathbf{w}_{r+1}p_{r+1} = \mathbf{0}$, with at least one nonzero $p_i \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]$.

A behavior $\mathfrak{B}$ of rank r is *free* if it admits a full column rank generator matrix, that is a generator matrix G with r columns. This amounts to saying that each trajectory $\mathbf{w}$ in $\mathfrak{B}$ is uniquely expressed as a linear combination $\mathbf{w} = \mathbf{g}_1u_1 + \mathbf{g}_2u_2 + \cdots + \mathbf{g}_ru_r$, $u_i \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]$, of the columns of G.

The main properties of a finite behavior $\mathfrak{B}$ are connected with certain elementary operations we can perform on the system trajectories. These operations essentially reduce to "pasting" pieces of different trajectories into legal elements of $\mathfrak{B}$, or to "cutting" a set of samples out of a given trajectory, so as to obtain a new behavior sequence.

One of the pillars of Willems behavior theory is the notion of (external) controllability. For 1D controllable behaviors the past has no lasting implications about the future [18], which means that the restriction of a 1D trajectory to $(-\infty, t]$ does not provide any information about the values the trajectory takes on $[t+\delta, +\infty)$, when $\delta > 0$ is properly chosen. In a multidimensional context the notions of "past" and "future" are quite elusive and, in many cases, unsuitable for classifying and processing the available data. What seems more reasonable, instead, is to investigate to what extent the values a trajectory $\mathbf{w}$ assumes on a subset $\mathcal{S}_1 \subset \mathbb{Z}^n$ influence the values on a subset $\mathcal{S}_2$, disjoint from $\mathcal{S}_1$, and to check if there exists a lower bound on the distance

$$d(\mathcal{S}_1, \mathcal{S}_2) := \min \left\{ \sum_{i=1}^{n} |h_i - k_i| : (h_1, h_2, ..., h_n) \in \mathcal{S}_1, (k_1, k_2, ..., k_n) \in \mathcal{S}_2 \right\}, \tag{2.2}$$

which guarantees that $\mathbf{w}|\mathcal{S}_2$, the *restriction to* $\mathcal{S}_2$ *of the sequence* $\mathbf{w}$, is independent of $\mathbf{w}|\mathcal{S}_1$. This point of view led to the following definition [12].

(C$_1$) [Controllability] *A finite behavior* $\mathfrak{B}$ *is* controllable *if there exists an integer* $\delta > 0$ *such that, for any pair of nonempty subsets* $\mathcal{S}_1, \mathcal{S}_2$ *of* $\mathbb{Z}^n$*, with* $d(\mathcal{S}_1, \mathcal{S}_2) \geq \delta$*, and any pair of trajectories* $\mathbf{w}_1$ *and* $\mathbf{w}_2 \in \mathfrak{B}$ *, there exists* $\mathbf{v} \in \mathfrak{B}$ *such that*

$$\mathbf{v}|\mathcal{S}_1 = \mathbf{w}_1|\mathcal{S}_1 \qquad \text{and} \qquad \mathbf{v}|\mathcal{S}_2 = \mathbf{w}_2|\mathcal{S}_2. \tag{2.3}$$

While definition (C$_1$) requires pasting together different signals into a new one, the following statement refers to the possibility of finding a legal extension for every portion $\mathbf{w}|\mathcal{S}$ of a behavior trajectory $\mathbf{w}$, by adjusting the sample values in a small area surrounding $\mathcal{S}$. More precisely, by introducing for $\varepsilon \geq 0$ the ε-*extension* of the set $\mathcal{S}$

$$\mathcal{S}^\varepsilon := \{\mathbf{h} \in \mathbb{Z}^n : d(\mathbf{h}, \mathcal{S}) \leq \varepsilon\},$$

one can give the following definition.

(C$_2$) [Zero-controllability] *A finite behavior* $\mathfrak{B}$ *is* zero-controllable *if there exists an integer* $\varepsilon > 0$ *such that, for any nonempty set* $\mathcal{S}$ *of* $\mathbb{Z}^n$ *and any* $\mathbf{w} \in \mathfrak{B}$*, there exists* $\mathbf{v} \in \mathfrak{B}$ *satisfying*

$$\mathbf{v}|\mathcal{S} = \mathbf{w}|\mathcal{S} \qquad \text{and} \qquad \text{supp}(\mathbf{v}) \subseteq \mathcal{S}^\varepsilon. \tag{2.4}$$

Properties (C_1) and (C_2) make sense both for finite and infinite support behaviors, and the proof of (C_1) $\Leftrightarrow$ (C_2) given below holds for both of them. However, while conditions (C_1) and (C_2) are always met by a finite behavior $\mathfrak{B}$, and essentially follow from the module structure of $\mathfrak{B}$, for an infinite behavior controllability constitutes an additional constraint w.r.t. linearity and shift invariance [12, 13].

Controllability and zero controllability are equivalent.

(C_1) $\Rightarrow$ (C_2) Assume that $\mathfrak{B}$ meets condition (C_1). Given $\mathbf{w} \in \mathfrak{B}$ and $\mathcal{S} \subset \mathbb{Z}^n$, take in ($C_1$) $\mathbf{w}_1 = \mathbf{w}$, $\mathbf{w}_2 = \mathbf{0}$, $\mathcal{S}_1 = \mathcal{S}$ and $\mathcal{S}_2 = \mathcal{C}\mathcal{S}^\delta$, where $\mathcal{C}\mathcal{S}$ denotes the complementary set of $\mathcal{S}$. Then the trajectory $\mathbf{v}$ which fulfills (2.3) satisfies (2.4) with $\varepsilon = \delta$.

(C_2) $\Rightarrow$ (C_1) Assume that $\mathfrak{B}$ satisfies condition (C_2). Given $\mathbf{w}_1$ and $\mathbf{w}_2 \in \mathfrak{B}$ and $\mathcal{S}_1, \mathcal{S}_2 \subset \mathbb{Z}^n$, with $d(\mathcal{S}_1, \mathcal{S}_2) > \varepsilon$, by ($C_2$) there exist $\mathbf{v}_1$ and $\mathbf{v}_2 \in \mathfrak{B}$ such that

$$\mathbf{v}_i|\mathcal{S}_i = \mathbf{w}_i|\mathcal{S}_i, \quad \mathrm{supp}(\mathbf{v}_i) \subset \mathcal{S}_i^\varepsilon, \qquad i = 1, 2.$$

Thus $\mathbf{v} := \mathbf{v}_1 + \mathbf{v}_2 \in \mathfrak{B}$ satisfies $\mathbf{v}|\mathcal{S}_i = \mathbf{w}_i|\mathcal{S}_i$, $i = 1, 2$, and (C_1) holds for $\delta = \varepsilon + 1$.

A finite behavior $\mathfrak{B}$ is controllable.

Suppose that $G \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^{p \times m}$ is a generator matrix of $\mathfrak{B}$ and let η be a positive integer such that $B(\mathbf{0}, \eta)$, the ball of radius η and center in the origin, includes $\mathrm{supp}(G)$. Consider any set $\mathcal{S} \subset \mathbb{Z}^n$ and $\mathbf{w} = G\mathbf{u} \in \mathfrak{B}$. If $\bar{\mathbf{u}}$ is the sequence which coincides with $\mathbf{u}$ on $\mathcal{S}^\eta$ and is zero elsewhere, the trajectory $\mathbf{v} := G\bar{\mathbf{u}}$ satisfies $\mathbf{v}|\mathcal{S} = \mathbf{w}|\mathcal{S}$, and has support which does not exceed $\mathcal{S}^{2\eta}$. So (C_2) is met with $\varepsilon = 2\eta$.

Given two disjoint sets $\mathcal{S}_1$ and $\mathcal{S}_2$ which are far enough apart, controllability expresses the possibility of steering any behavior sequence known in $\mathcal{S}_1$ into another element of $\mathfrak{B}$ assigned on $\mathcal{S}_2$, meanwhile producing a legal trajectory. Like controllability, also observability will be introduced without reference to the concept of state, according to some recent works of Forney et al. [7, 11]. Observability formalizes the possibility of pasting into a legal sequence any pair of trajectories that take the same values on a sufficiently large subset of $\mathbb{Z}^n$. This is equivalent to saying that, however a sequence $\mathbf{w} \in \mathfrak{B}$ and a subset $\mathcal{S} \subset \mathbb{Z}^n$ are chosen, the possible extensions of $\mathbf{w}|\mathcal{S}$ only depend on the values of $\mathbf{w}$ on a boundary region of $\mathcal{S}$.

Under this viewpoint, observability endows a behavior with a "separation property" that allows to take into account only a small amount of data in order to extend a portion of behavior sequence. Furthermore, once we think of the samples in $\mathcal{S}$ as the information about the past dynamics of the system, observability enables us to design the "future" evolution by considering only the most "recent" data (those on the boundary), this way reminding of the notion of state.

($\mathbf{O}_1$) [Observability] *A finite behavior* $\mathfrak{B}$ *is* observable *if there exists an integer* $\delta > 0$ *such that, for any pair of nonempty subsets* $\mathcal{S}_1, \mathcal{S}_2$ *of* $\mathbb{Z}^n$*, with* $d(\mathcal{S}_1, \mathcal{S}_2) \geq \delta$*, and any pair of trajectories* $\mathbf{w}_1, \mathbf{w}_2 \in \mathfrak{B}$*, satisfying* $\mathbf{w}_1|\mathcal{C}(\mathcal{S}_1 \cup \mathcal{S}_2) = \mathbf{w}_2|\mathcal{C}(\mathcal{S}_1 \cup \mathcal{S}_2)$*, the trajectory*

$$\mathbf{v}(\mathbf{h}) = \begin{cases} \mathbf{w}_1(\mathbf{h}) & \mathbf{h} \in \mathcal{S}_1 \\ \mathbf{w}_1(\mathbf{h}) = \mathbf{w}_2(\mathbf{h}) & \mathbf{h} \in \mathcal{C}(\mathcal{S}_1 \cup \mathcal{S}_2) \\ \mathbf{w}_2(\mathbf{h}) & \mathbf{h} \in \mathcal{S}_2 \end{cases} \tag{2.5}$$

is an element of $\mathfrak{B}$*.*

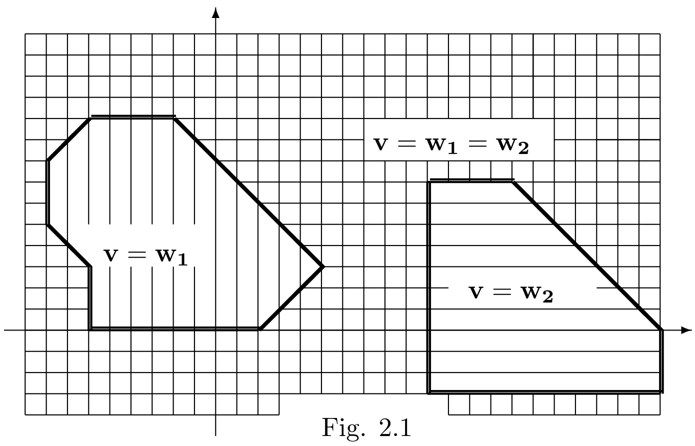

Fig. 2.1

Observability can be equivalently restated as follows: if the support of a behavior sequence $\mathbf{w}$ can be partitioned into a pair of disjoint subsets, which are far enough apart, the restrictions of $\mathbf{w}$ to each subset represent legal trajectories.

($\mathbf{O}_2$) [Zero-observability] *A finite behavior* $\mathfrak{B}$ *is* zero-observable *if there exists an integer* $\varepsilon > 0$ *such that for any* $\mathbf{w} \in \mathfrak{B}$ *satisfying* $\mathbf{w}|(\mathcal{S}^\varepsilon \setminus \mathcal{S}) = \mathbf{0}$*,* $\mathcal{S}$ *a nonempty set in* $\mathbb{Z}^n$*, the sequence*

$$\mathbf{v}(\mathbf{h}) = \begin{cases} \mathbf{w}(\mathbf{h}) & \mathbf{h} \in \mathcal{S} \\ 0 & \textit{elsewhere} \end{cases} \tag{2.6}$$

belongs to $\mathfrak{B}$*.*

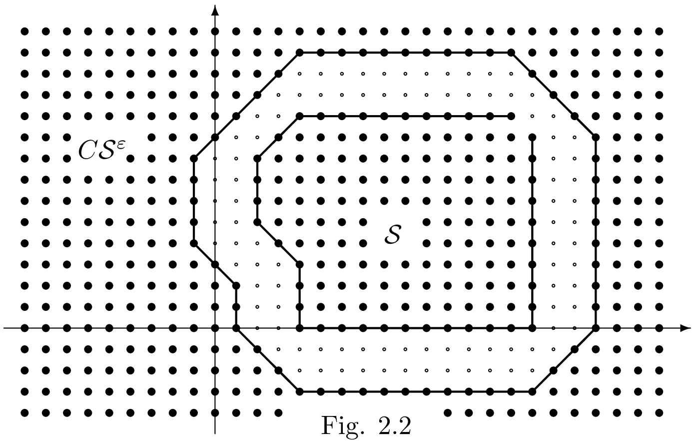

Fig. 2.2

Observability and zero observability are equivalent.

$(\mathrm{O}_1) \Rightarrow (\mathrm{O}_2)$ Assume that $\mathfrak{B}$ fulfills condition (O_1). Given $\mathcal{S} \subset \mathbb{Z}^n$ and $\mathbf{w} \in \mathfrak{B}$ such that $\mathbf{w}|(\mathcal{S}^\delta \setminus \mathcal{S}) = \mathbf{0}$, take in (O_1) $\mathbf{w}_1 = \mathbf{w}$, $\mathbf{w}_2 = \mathbf{0}$, $\mathcal{S}_1 = \mathcal{S}$ and $\mathcal{S}_2 = \mathcal{C}\mathcal{S}^\delta$. The trajectory $\mathbf{v} \in \mathfrak{B}$ satisfying (2.5), satisfies also (2.6) with $\varepsilon = \delta$.

$(\mathrm{O}_2) \Rightarrow (\mathrm{O}_1)$ Assume that $\mathfrak{B}$ fulfills condition (O_2). Given $\mathcal{S}_1, \mathcal{S}_2 \subset \mathbb{Z}^n$, with $d(\mathcal{S}_1, \mathcal{S}_2) > \varepsilon$, and $\mathbf{w}_1$, $\mathbf{w}_2 \in \mathfrak{B}$ satisfying $\mathbf{w}_1|\mathcal{C}(\mathcal{S}_1 \cup \mathcal{S}_2) = \mathbf{w}_2|\mathcal{C}(\mathcal{S}_1 \cup \mathcal{S}_2)$, the sequence $\mathbf{w}_1 - \mathbf{w}_2 \in \mathfrak{B}$ satisfies $(\mathbf{w}_1 - \mathbf{w}_2)|\mathcal{C}(\mathcal{S}_1 \cup \mathcal{S}_2) = \mathbf{0}$. As a consequence, the sequence $\mathbf{w}$ given by

$$\mathbf{w}(\mathbf{h}) = \begin{cases} \mathbf{w}_1(\mathbf{h}) - \mathbf{w}_2(\mathbf{h}) & \mathbf{h} \in \mathcal{S}_1 \\ 0 & \text{elsewhere} \end{cases}$$

is in $\mathfrak{B}$, and $\mathbf{v} := \mathbf{w} + \mathbf{w}_2 \in \mathfrak{B}$ fulfills (2.5). So, (O_1) holds for $\delta = \varepsilon + 1$.

3 Parity checks and trajectory recognition

Underlying the definition of controllability is the idea of driving a portion of trajectory into another one, provided that there is room enough for adjustements. In rough terms, the objective one has in mind is that of manipulating the control variables to cause the system to behave in $\mathcal{S}_2$ in a more desirable manner than is expected by watching the system trajectory on $\mathcal{S}_1$. So, controllability is naturally connected with the generation of $\mathfrak{B}$ as the image of some matrix G, acting on the input space.

Observability is somehow related with the "dual" issue of recognizing whether a given sequence $\mathbf{v} \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p$ is an element of $\mathfrak{B}$. This problem, that typically arises in fault detection and convolutional encoding contexts, can be managed by resorting to a linear filter (residual generator or syndrome former) that produces an identically zero output signal when the input is an admissible trajectory of $\mathfrak{B}$. From a mathematical point of view, this requires to find a set of sequences (parity checks) endowed with the property that their convolution with every element of $\mathfrak{B}$ is zero.

So, for a given behavior $\mathfrak{B} \subseteq \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p$, a (finite) *parity check* is a column vector $\mathbf{s} \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p$ that satisfies $\mathbf{s}^T\mathbf{w} = \mathbf{0}$, for all $\mathbf{w} \in \mathfrak{B}$. The set $\mathfrak{B}^\perp$ of all finite parity checks of $\mathfrak{B}$ is the *orthogonal behavior*, and as a submodule of $\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p$, it is generated by the columns of some matrix $H \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^{p\times q}$, that is

$$\mathfrak{B}^\perp = \{\mathbf{s} \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p : \mathbf{s} = H\mathbf{x}, \mathbf{x} \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^q\} = \mathrm{Im}H. \tag{3.1}$$

Condition $\mathbf{s}^T\mathbf{w} = \mathbf{0}$, $\forall\ \mathbf{s} \in \mathfrak{B}^\perp$, however, need not imply $\mathbf{w} \in \mathfrak{B}$. In general

$$\mathfrak{B}^{\perp\perp} := \{\mathbf{w} \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p : \mathbf{s}^T\mathbf{w} = \mathbf{0}, \forall\ \mathbf{s} \in \mathfrak{B}^\perp\} \tag{3.2}$$

properly includes $\mathfrak{B}$, and is the set of all L-polynomial vectors obtained by combining the columns of G over the field of rational functions $\mathbb{F}(\mathbf{z})$. It is clear that $\mathfrak{B}$ can be identified via a finite set of parity checks if and only if $\mathfrak{B} = \mathfrak{B}^{\perp\perp}$ or, equivalently,

$$\mathfrak{B} = \ker H^T := \{\mathbf{w} \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p : H^T\mathbf{w} = \mathbf{0}\}. \tag{3.3}$$

In this setting, observability finds a somewhat more substantial interpretation. Actually, if $\mathfrak{B} = \ker H^T$, the restriction of a trajectory to a set $\mathcal{S}$ still provides a legal signal every time the distance between $\mathcal{S}$ and the remaining support of the trajectory exceeds the range of action of the parity check matrix H.

Proposition 3.1 below shows that kernel representations are possible, as it can be expected, only for observable behaviors, and makes it clear that observability induces further constraints on the structure of $\mathfrak{B}$, in addition to linearity and shift invariance.

A behavior $\mathfrak{B} \subseteq \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p$ is observable if and only if there exist an integer $h > 0$ and an L-polynomial matrix $H^T \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^{h\times p}$ such that $\mathfrak{B} = \ker H^T$.

The proof of the proposition depends on a couple of technical lemmas.

Let R be an integral domain, and consider the polynomial in $R[z]$

$$m(z) = \alpha_0\ z^r - \alpha_1\ z^{r-1} + \alpha_2\ z^{r-2} \cdots + (-1)^r\alpha_r.$$

For any $\rho \geq 0$ there is $p(z) \in R[z]$ such that $p(z)m(z) \in R[z^{\rho+1}]$.

Let Q be the field of fractions of R and L the algebraic closure of Q. Then $m(z)$ can be written as $m(z) = \alpha_0 \prod_{i=1}^r (z - \xi_i)$, where $\xi_i \in L,\ i = 1, 2, \ldots, n$, and

$$\sum_i \xi_i = \alpha_1/\alpha_0, \qquad \sum_{i<j} \xi_i\xi_j = \alpha_2/\alpha_0, \quad \ldots \quad \xi_1\xi_2\cdots\xi_r = \alpha_r/\alpha_0. \tag{3.4}$$

Consider the following polynomial in $L[z]$

$$\begin{aligned}\tilde{p}(z) &= \prod_{i=1}^{r}(z^\rho + \xi_i z^{\rho-1} + \xi_i^2 z^{\rho-2} + \cdots + \xi_i^{\rho-1} z + \xi_i^\rho) \\ &= \sum_{0\le i_1,i_2,\dots,i_r\le\rho} z^{r\rho-i_1-i_2-\cdots-i_r}\xi_1^{i_1}\xi_2^{i_2}\cdots\xi_r^{i_r} = \sum_{t=0}^{r\rho} z^{r\rho-t} \sum_{\substack{i_1+i_2+\dots+i_r=t \\ 0\le i_1,i_2,\dots,i_r\le\rho}} \xi_1^{i_1}\xi_2^{i_2}\cdots\xi_r^{i_r}.\end{aligned}$$

Each coefficient of $\tilde{p}(z)$ is a symmetric polynomial in the indeterminates $\xi_1, \xi_2, ..., \xi_r$, with integer coefficients, and hence it is expressible [10] as a polynomial in the elementary symmetric polynomials defined in (3.4), again with integer coefficients. Thus $\tilde{p}(z)$ is in $Q[z]$, the denominators of its coefficients are powers of α_0, and there exists a positive integer ν such that $p(z) := \alpha_o^\nu\, \tilde{p}(z)$ belongs to $R[z]$. To conclude the proof, we note that $p(z)m(z)$ is an element of $R[z^{\rho+1}]$ since

$$p(z)m(z) = \alpha_o^\nu \prod_{i=1}^{r}[(z^\rho+\xi_i z^{\rho-1}+\xi_i^2 z^{\rho-2}+\cdots+\xi_i^\rho)(z-\xi_i)] = \alpha_o^\nu \prod_{i=1}^{r}(z^{\rho+1}-\xi_i^{\rho+1}).$$

Let $m(\mathbf{z})$ be in $\mathbb{F}[\mathbf{z}]$. For any integer $\rho > 0$ there is $p(\mathbf{z}) \in \mathbb{F}[\mathbf{z}]$ s.t.

$$m(\mathbf{z})p(\mathbf{z}) \;\in\; \mathbb{F}[\mathbf{z}^\rho] := \mathbb{F}[z_1^\rho, ..., z_n^\rho]. \tag{3.5}$$

As $m(\mathbf{z}) = m(z_1, ..., z_n)$ can be viewed as an element of $\mathbb{F}[z_1, ..., z_{n-1}]\ [z_n]$, by Lemma 3.2 there exists $p_1(z_1, \dots, z_n) \in \mathbb{F}[z_1, \dots, z_{n-1}][z_n]$ such that

$$m_1(z_1, ..., z_n^\rho) := m(z_1, ..., z_n)p_1(z_1, \dots, z_n) \in \mathbb{F}[z_1, ..., z_{n-1}][z_n^\rho].$$

Looking at $m_1(z_1, ..., z_n^\rho)$ as a polynomial in $\mathbb{F}[z_1, \dots, z_{n-2}, z_n^\rho][z_{n-1}]$, we know that there exists $p_2(z_1, \dots, z_{n-1}, z_n^\rho)$ such that

$$m_2(z_1, ..., z_{n-1}^\rho, z_n^\rho) := m_1(z_1, ..., z_{n-1}, z_n^\rho)p_2(z_1, \dots, z_{n-1}, z_n^\rho) \in \mathbb{F}[z_1, \dots, z_{n-2}, z_n^\rho][z_{n-1}^\rho].$$

In n steps we end up with a polynomial

$$\begin{aligned} m_n(z_1^\rho, z_2^\rho, \dots, z_n^\rho) : &= m(z_1, \dots, z_n)p_1(z_1, \dots, z_n) \\ &\cdot\; p_2(z_1, z_2, \dots, z_{n-1}, z_n^\rho) \dots p_n(z_1^\rho, z_2^\rho, ..., z_n^\rho) \in \mathbb{F}[z_1^\rho, z_2^\rho, \dots, , z_n^\rho], \end{aligned}$$

and (3.5) holds with $p = p_1\ p_2\ \dots\ p_n$.

(*of Proposition 3.1*) Assume that $\mathfrak{B} = \mathrm{Im}G$, $G \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^{p\times m}$, is an observable behavior, and let $\mathfrak{B}^\perp = \mathrm{Im}\, H$, $H \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^{p\times q}$, denote the orthogonal behavior introduced in (3.1). We will show that $\mathfrak{B} \equiv \ker H^T$. Since $H^T G = \mathbf{0}$, it is clear that $\ker H^T \supseteq \mathfrak{B}$. To prove the converse, express $\mathbf{w} \in \ker H^T$ as $\mathbf{w} = G\mathbf{n}/d(\mathbf{z})$, $d \in \mathbb{F}[\mathbf{z}], \mathbf{n} \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^{m\times 1}$. By Lemma 3.3, for every integer $\rho > 0$ there is a suitable polynomial $p(\mathbf{z})$ such that $p(\mathbf{z})d(\mathbf{z}) \in \mathbb{F}[z_1^\rho, \dots, z_n^\rho]$. If property ($\mathrm{O}_2$) holds for $\varepsilon > 0$, and $r > 0$ is an integer such that $\mathrm{supp}(\mathbf{w}) \subseteq B(\mathbf{0}, r)$, we choose $\rho > 2r + \varepsilon$. So, the behavior sequence $p(\mathbf{z})d(\mathbf{z})\mathbf{w} = G\mathbf{n}p(\mathbf{z})$ can be written as

$$\sum_{i_1,i_2,\dots,i_n} c_{i_1,i_2,\dots,i_n} z_1^{\rho i_1} z_2^{\rho i_2} \cdots z_n^{\rho i_n} \mathbf{w},$$

and thus is the sum of disjoint shifted copies of $\mathbf{w}$, and the distance between two arbitrary copies exceeds ε. So, by (O_2), each copy of $\mathbf{w}$, and hence $\mathbf{w}$ itself, is in $\mathfrak{B}$.

Conversely, let $\mathfrak{B} = \ker H^T$, and set $\varepsilon = 2s$, with $s > 0$ an integer s.t. $B(\mathbf{0}, s) \supseteq \mathrm{supp}(H^T)$. If $\mathcal{S}$ is a subset of $\mathbb{Z}^n$ and $\mathbf{w} \in \mathfrak{B}$ satisfies $\mathbf{w}|(\mathcal{S}^\varepsilon \setminus \mathcal{S}) = \mathbf{0}$, the sequence

$$\mathbf{v}(\mathbf{h}) = \begin{cases} \mathbf{w}(\mathbf{h}) & \mathbf{h} \in \mathcal{S} \\ 0 & \text{elsewhere} \end{cases}$$

is in $\ker H^T$ and hence in $\mathfrak{B}$. Consequently, $\mathfrak{B}$ is zero-observable.

The kernel description given in Proposition 3.1 leads to new insights into the internal structure of an observable behavior. Observability, indeed, expresses a sort of "localization" of the system laws or, equivalently, the existence of a bound on the size of all windows (in $\mathbb{Z}^n$) we have to look at when deciding whether a signal belongs to $\mathfrak{B}$. Denoting by $\mathfrak{B}|\mathcal{S} := \{\mathbf{w}|\mathcal{S} : \mathbf{w} \in \mathfrak{B}\}$ the set of all restrictions to $\mathcal{S}$ of behavior trajectories, the above localization property finds a formal statement as follows:

(O_3) [Local-detectability] *A finite behavior $\mathfrak{B}$ is* locally-detectable *if there is an integer $\nu > 0$ such that every signal $\mathbf{w}$ satisfying $\mathbf{w}|\mathcal{S} \in \mathfrak{B}|\mathcal{S}$ for every $\mathcal{S} \subset \mathbb{Z}^n$ with $\mathrm{diam}(\mathcal{S}) \le \nu$, is in $\mathfrak{B}$.*

Local detectability and observability are equivalent.

Assume that $\mathfrak{B}$ satisfies (O_3) for a certain $\nu > 0$. Given $\mathcal{S} \subset \mathbb{Z}^n$ and $\mathbf{w} \in \mathfrak{B}$ such that $\mathbf{w}|(\mathcal{S}^\nu \setminus \mathcal{S}) = \mathbf{0}$, define $\mathbf{v}$ as follows

$$\mathbf{v}(\mathbf{h}) = \begin{cases} \mathbf{w}(\mathbf{h}) & \mathbf{h} \in \mathcal{S}^\nu \\ 0 & \text{elsewhere.} \end{cases} \tag{3.6}$$

Consider any window $\mathcal{W}$, with $\mathrm{diam}(\mathcal{W}) \le \nu$. If $\mathcal{W}$ is included in $\mathcal{S}^\nu$, then $\mathbf{v}|\mathcal{W} = \mathbf{w}|\mathcal{W} \in \mathfrak{B}|\mathcal{W}$, otherwise we have $\mathcal{W} \cap \mathcal{S} = \emptyset$, and therefore

$$\mathbf{v}|\mathcal{W} = \mathbf{0}|\mathcal{W} \in \mathfrak{B}|\mathcal{W}. \tag{3.7}$$

So, by (O_3), $\mathbf{v}$ is a legal trajectory, and (O_2) holds for $\varepsilon = \nu$.

Conversely, assume that $\mathfrak{B}$ is observable. By Proposition 3.1, there exists an L-polynomial matrix $H \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^{p\times q}$ such that $\mathfrak{B} = \ker H^T$. Let $\nu > 0$

be an integer such that $\mathrm{supp}(H^T) \subseteq B(\mathbf{0}, \nu)$, and suppose that $\mathbf{v}$ is any signal satisfying $\mathbf{v}|\mathcal{S} \in \mathfrak{B}|\mathcal{S}$ for every $\mathcal{S} \subset \mathbb{Z}^n$ with $\mathrm{diam}(\mathcal{S}) \leq 2\nu$. If $\bar{\mathcal{S}} := -\mathrm{supp}(H^T)$, the computation of the coefficient of $\mathbf{z}^{\mathbf{k}}$ in $H^T\mathbf{v}$ involves only samples of $\mathbf{v}$ indexed in

$$\mathbf{k} + \bar{\mathcal{S}} := \{\mathbf{h} \in \mathbb{Z}^n : \mathbf{h} - \mathbf{k} \in \bar{\mathcal{S}}\} = -\mathrm{supp}(\mathbf{z}^{\mathbf{k}} H^T). \tag{3.8}$$

On the other hand, since $\mathrm{diam}(\mathbf{k} + \bar{\mathcal{S}}) \leq 2\nu$, there exists $\mathbf{w}_{\mathbf{k}} \in \mathfrak{B}$ which satisfies $\mathbf{v}|(\mathbf{k}+\bar{\mathcal{S}}) = \mathbf{w}_{\mathbf{k}}|(\mathbf{k}+\bar{\mathcal{S}})$, and this result holds for every $\mathbf{k} \in \mathbb{Z}^n$. So, the coefficient of $\mathbf{z}^{\mathbf{k}}$ in $H^T\mathbf{v}$ is the same as in $H^T\mathbf{w}_{\mathbf{k}} \equiv \mathbf{0}$, and hence $\mathbf{v} \in \ker H^T = \mathfrak{B}$.

The equivalent descriptions of observability given in $(\mathrm{O}_1) \div (\mathrm{O}_3)$ rely on the trajectories' supports, whereas Proposition 3.1 involves parity checks and kernel representations. A different approach to this notion consists of regarding behaviors with p components as elements in the lattice of submodules of $\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p$, and investigating whether observable elements enjoy some special ordering properties.

Keeping in with the same spirit, we may investigate how an observable behavior is affected by certain "extension operations" that merge lattice elements into larger ones. There are essentially two natural ways to perform these extensions: one consists of embedding $\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p$, and therefore each of its submodules, in the rational vector space $\mathbb{F}(\mathbf{z})^p$, the other of considering $\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p$ as a submodule of $\mathcal{F}_\infty^p$, the set of nD trajectories with p components, whose supports possibly extend to the whole space $\mathbb{Z}^n$.

Once a behavior $\mathfrak{B}$ with p components is given, in the first case we have to consider the smallest vector subspace of $\mathbb{F}(\mathbf{z})^p$ including $\mathfrak{B}$

$$\mathfrak{B}_{\mathrm{rat}} := \Big\{ \sum_{i=1}^{r} \mathbf{w}_i a_i : \mathbf{w}_i \in \mathfrak{B},\ a_i \in \mathbb{F}(\mathbf{z}),\ r \in \mathbb{N} \Big\}, \tag{3.9}$$

and restrict our attention to the submodule $\mathfrak{B}_{\mathrm{rat}} \cap \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p$ of finite support sequences. In general this properly includes $\mathfrak{B}$, and hence is a larger element of the lattice. In the other case, we merge $\mathfrak{B}$ in

$$\mathfrak{B}_\infty := \Big\{ \sum_{i=1}^{r} \mathbf{w}_i a_i : \mathbf{w}_i \in \mathfrak{B},\ a_i \in \mathcal{F}_\infty,\ r \in \mathbb{N} \Big\}, \tag{3.10}$$

the smallest $\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]$-submodule of $\mathcal{F}_\infty^p$ which includes $\mathfrak{B}$. Again we have to confine ourselves to the set of its finite elements $\mathfrak{B}_\infty \cap \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p$, which clearly includes all trajectories of $\mathfrak{B}$.

Let $\mathfrak{B} \subseteq \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p$ be a behavior of rank r. The following statements are equivalent:

(1) $\mathfrak{B}$ is observable;
(2) $\mathfrak{B} \equiv \mathfrak{B}_\infty \cap \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p$;
(3) $\mathfrak{B} \equiv \mathfrak{B}_{\mathrm{rat}} \cap \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p$;

(4) $\mathfrak{B}$ is maximal in the set of all submodules of $\mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^p$ of rank r;

(5) $s\mathbf{w} \in \mathfrak{B} \;\Rightarrow\; \mathbf{w} \in \mathfrak{B}$, for every $\mathbf{w} \in \mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^p$ and every nonzero $s \in \mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]$;

(6) $\mathfrak{B} = \mathfrak{B}^{\perp\perp}$.

(1) $\Rightarrow$ (2) As $\mathfrak{B}$ is observable, there exists $H \in \mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^{p\times q}$ such that $\mathfrak{B} = \ker H^T = \{\mathbf{w} \in \mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^p : H^T\mathbf{w} = \mathbf{0}\}$. If $\mathbf{w} \in \mathfrak{B}_\infty \cap \mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^p$, then $\mathbf{w} = \sum_i \mathbf{w}_i a_i$, $a_i \in \mathcal{F}_\infty, \mathbf{w}_i \in \mathcal{B}$, and therefore $H^T\mathbf{w} = H^T\left(\sum_i \mathbf{w}_i a_i\right) = \sum_i (H^T\mathbf{w}_i)a_i = \mathbf{0}$. Thus $\mathbf{w} \in \ker H^T = \mathfrak{B}$, which implies $\mathfrak{B} \supseteq \mathfrak{B}_\infty \cap \mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^p$. The reverse inclusion is obvious.

(2) $\Rightarrow$ (3) Follows immediately from $\mathfrak{B} \subseteq \mathfrak{B}_{\rm rat} \cap \mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^p \subseteq \mathfrak{B}_\infty \cap \mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^p$.

(3) $\Rightarrow$ (4) If $\mathfrak{B}' \supseteq \mathfrak{B}$ and rank$\mathcal{B}'$ = rank$\mathfrak{B}$, it is clear that $\mathfrak{B}$ and $\mathfrak{B}'$ generate the same $\mathbb{F}(\mathbf{z})$-subspace of $\mathbb{F}(\mathbf{z})^p$ and, consequently, $\mathfrak{B}_{\rm rat} \cap \mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^p = \mathfrak{B}'_{\rm rat} \cap \mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^p$. So, the inclusions chain $\mathfrak{B}_{\rm rat}\cap\mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^p \supseteq \mathfrak{B}' \supseteq \mathfrak{B}$ and assumption (3) together imply $\mathfrak{B}' = \mathfrak{B}$, which means that $\mathfrak{B}$ is maximal.

(4) $\Rightarrow$ (5) Suppose $s\mathbf{w} \in \mathfrak{B}$, $s \in \mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]$. The behavior $\mathfrak{B}'$ generated by $\mathfrak{B}$ and $\mathbf{w}$ has the same rank of $\mathfrak{B}$, and hence, by the maximality assumption, coincides with $\mathfrak{B}$.

(5) $\Rightarrow$ (6) As $\mathfrak{B}$ and $\mathfrak{B}^{\perp\perp}$ have the same rank r and $\mathfrak{B}^{\perp\perp} \supseteq \mathfrak{B}$, both behaviors generate the same $\mathbb{F}(\mathbf{z})$-subspace of $\mathbb{F}(\mathbf{z})^p$. In particular, $\mathbf{w} \in \mathfrak{B}^{\perp\perp}$ implies $\mathbf{w} \in (\mathfrak{B}^{\perp\perp})_{\rm rat} = \mathfrak{B}_{\rm rat}$. So, there exist $p_i, s_i \in \mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]$ and $\mathbf{w}_i \in \mathfrak{B}$, such that $\mathbf{w} = \sum_{i=1}^r \mathbf{w}_i\, p_i/s_i$, which implies $s\mathbf{w} \in \mathfrak{B}$, $s = \ell$.c.m.$\{s_i\}$. By assumption (5), also $\mathbf{w}$ is in $\mathfrak{B}$.

(6) $\Rightarrow$ (1) Since $\mathfrak{B}^\perp$ is a submodule of $\mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^p$, there exists a suitable L-polynomial matrix H such that $\mathfrak{B}^\perp = {\rm Im}H$. So

$$\mathfrak{B}^{\perp\perp} = (\mathfrak{B}^\perp)^\perp = \{\mathbf{w} \in \mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^p : \mathbf{v}^T\mathbf{w} = \mathbf{0}, \forall \mathbf{v} \in {\rm Im}H\} = \ker H^T.$$

By assumption (6), $\mathfrak{B}$ coincides with $\ker H^T$, and hence is observable.

4 Behavior decomposition

In this section we take a first step towards a structural analysis of finite support behaviors. The scope of structure theory is to describe general behaviors in terms of some simpler ones, simpler in some perceptible way, perhaps in terms of concreteness, perhaps in terms of tractability. Of essential importance, after one has decided upon these simpler objects, is to find a method of passing down to them and to discover how they weave together to yield the general behavior with which we began.

Observable behaviors constitute good candidates for these simpler objects, as each behavior can be embedded into an observable one. In order to represent a general behavior $\mathfrak{B}$, then, we have to slice out of its enveloping observable behavior $\mathcal{B}^{\perp\perp}$ a certain part. This can be done by intersecting $\mathcal{B}^{\perp\perp}$ with a suitable, not necessarily unique, element of a behavior class that exhibits properties which are as far as possible from observability and hence from local detectability. The definition of this class depends on the notion of constrained variables which we now introduce.

Let $\mathfrak{B} \subseteq \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p$ be a finite support behavior and $\{i_1, i_2, \dots, i_r\}$, $r < p$, a subset of $\{1, 2, \dots, p\}$. We call $w_{i_1}, w_{i_2}, \dots, w_{i_r}$ constrained variables of $\mathfrak{B}$ if for every pair of trajectories $\mathbf{v}, \mathbf{v}' \in \mathfrak{B}$, $v_j = v'_j$ for every $j \notin \{i_1, i_2, \dots, i_r\}$ implies $\mathbf{v} = \mathbf{v}'$.

As shown in the following lemma, the maximum number of constrained variables of a behavior $\mathfrak{B}$ in $\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p$ can be expressed in terms of the rank and the number of components of $\mathfrak{B}$.

Let $\mathfrak{B} \subseteq \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p$ be a behavior of rank r. The maximum number of constrained variables of $\mathfrak{B}$ is $p - r$.

Let $G \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^{p \times m}$ be a generator matrix of $\mathfrak{B}$ and suppose, for sake of simplicity, that the first r rows of G are linearly independent, so that in

$$G = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} \begin{array}{l} \}r \\ \}p-r \end{array} ,$$

G_1 has full row rank. The components w_i, $i = r+1, r+2, \dots, n$, are constrained variables. If not, there would be a trajectory $\mathbf{w} = \begin{bmatrix} \mathbf{0} \\ \mathbf{w}_2 \end{bmatrix}$ in $\mathfrak{B}$, with $\mathbf{w}_2 \neq 0$, and hence an L-polynomial vector $\mathbf{u} \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^m$ s.t. $\begin{bmatrix} G_1 \\ G_2 \end{bmatrix} \mathbf{u} = \begin{bmatrix} \mathbf{0} \\ \mathbf{w}_2 \end{bmatrix}$. This is a contradiction, however, because

$$\text{rank } G_1 = \text{rank} \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} \qquad \Rightarrow \qquad \ker G_1 = \ker \begin{bmatrix} G_1 \\ G_2 \end{bmatrix}.$$

It remains to prove that the number of constrained variables cannot exceed $p-r$. Suppose, instead, that $k > p - r$ variables of $\mathfrak{B}$, say the last k, are constrained, and partition the generator matrix G into

$$G = \begin{bmatrix} \hat{G}_1 \\ \hat{G}_2 \end{bmatrix} \begin{array}{l} \}p-k \\ \}k \end{array} .$$

As r =rank $G >$ rank$\hat{G}_1$, ker$\hat{G}_1$ properly includes ker$G$. Consequently, there exists $\mathbf{u}$ s.t. $\hat{G}_2\mathbf{u} \neq \mathbf{0}$ and both $\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}$ and $\begin{bmatrix} \mathbf{0} \\ \hat{G}_2\mathbf{u} \end{bmatrix}$, $\hat{G}_2\mathbf{u} \neq \mathbf{0}$, are in $\mathfrak{B}$, which contradicts the assumption that the last k components are constrained.

A behavior $\mathfrak{B}$ devoid of constrained variables exhibits the very peculiar feature that for every finite set $\mathcal{S} \subset \mathbb{Z}^n$, $\mathfrak{B}|\mathcal{S}$ coincides with $\mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^p \mid \mathcal{S}$. This property, which appears somehow opposite to local detectability, makes it is impossible to recognize the trajectories of $\mathfrak{B}$ by resorting to a local checking procedure.

(LU) [Local-undetectability] *A behavior* $\mathfrak{B} \subseteq \mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^p$ *is* locally undetectable *if there exists* $\delta > 0$ *s.t. for every sequence* $\mathbf{v} \in \mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^p$ *and every set* $\mathcal{S} \subset \mathbb{Z}^n$*, a trajectory* $\mathbf{w} \in \mathfrak{B}$ *can be found, satisfying*

$$\mathbf{w}|\mathcal{S} = \mathbf{v}|\mathcal{S} \qquad \text{and} \qquad \mathrm{supp}(\mathbf{w}) \subseteq \mathcal{S}^{\delta}. \tag{4.1}$$

Let $\mathfrak{B} \subseteq \mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^p$ be a finite support behavior. The following facts are equivalent:

i) $\mathfrak{B}$ is devoid of constrained variables;

ii) $\mathfrak{B}$ is the image of some L-polynomial matrix $G \in \mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^{p\times m}$ with rank p;

iii) $\mathfrak{B}$ is locally undetectable.

i) $\Leftrightarrow$ ii) Immediate from Lemma 4.1.

ii) $\Rightarrow$ iii) Let $\mathbf{v}$ be an arbitrary sequence in $\mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^p$ and $\mathcal{S}$ a finite set. If $\mathfrak{B}$ = ImG, for some $G \in \mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^{p\times m}$ of rank p, $\mathbf{v}$ can be obtained as the image of some rational vector $\mathbf{u} \in \mathbb{F}(\mathbf{z})^p$, i.e., $\mathbf{v} = G\mathbf{u}$. Consider a power series expansion of $\mathbf{u}$ with support in a suitable cone of $\mathbb{Z}^n$, and introduce the finite sequence

$$\bar{\mathbf{u}}(\mathbf{h}) := \begin{cases} \mathbf{u}(\mathbf{h}) & \mathbf{h} \in \mathcal{S}^{\varepsilon} \\ \mathbf{0} & \text{elsewhere} \end{cases},$$

where ε is the radius of a ball centered in the origin and including the support of G. The behavior sequence $\bar{\mathbf{v}} := G\bar{\mathbf{u}}$ coincides with $\mathbf{v}$ on $\mathcal{S}$ and has support included in $\mathcal{S}^{2\varepsilon}$. So, (4.1) holds with $\delta = 2\varepsilon$.

iii) $\Rightarrow$ ii) Suppose that $\mathfrak{B}$ is locally undetectable and assume, by contradiction, $\mathfrak{B} = \mathrm{Im}G$, for some $G \in \mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^{p\times m}$ with rank less than p. Then there exixts a nonzero L-polynomial vector $\mathbf{h} \in \mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^p$ satisfying $\mathbf{h}^T G = 0$. Consider a sequence $\mathbf{v} \in \mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^p$ s.t. $\mathbf{h}^T\mathbf{v} \neq 0$ and a set $\mathcal{T} \subset \mathbb{Z}^n$ which includes both supp($\mathbf{v}$) and supp($\mathbf{h}^T\mathbf{v}$) and define $\mathcal{S} := \mathcal{T}^{\rho}$, where ρ is the radius of a ball, centered in the origin, which includes supp($\mathbf{h}$). If property (LU) holds for some $\delta > 0$, there is a trajectory $\mathbf{w} \in \mathfrak{B}$ that can be expressed as $\mathbf{w} = \mathbf{v} + \mathbf{r}$, for some $\mathbf{r}$ with support in $\mathcal{S}^{\delta} \setminus \mathcal{S}$ (see Fig. 4.1).

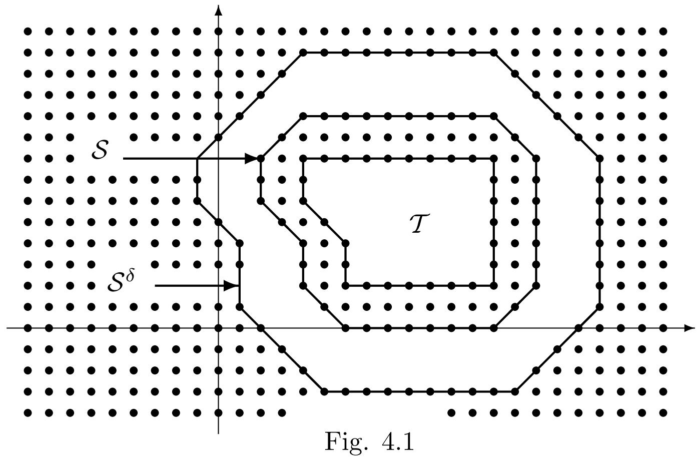

Fig. 4.1

As $\mathbf{w} = G\mathbf{a}$, for some $\mathbf{a} \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^m$, it follows that

$$0 = \mathbf{h}^T G\mathbf{a} = \mathbf{h}^T \mathbf{w} = \mathbf{h}^T \mathbf{v} + \mathbf{h}^T \mathbf{r}.$$

This is not possible, however, since $\mathcal{T}$ includes the support of $\mathbf{h}^T\mathbf{v}$ without intersecting $\mathrm{supp}(\mathbf{h}^T\mathbf{r})$.

For every behavior $\mathfrak{B} \subseteq \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p$ there exist an observable behavior $\mathfrak{B}_0$ and a locally undetectable behavior $\mathfrak{B}_{\mathrm{lu}}$ in $\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p$ s.t.

$$\mathfrak{B} = \mathfrak{B}_0 \cap \mathfrak{B}_{\mathrm{lu}}. \tag{4.2}$$

Moreover, $\mathfrak{B}_0$ is uniquely determined as $\mathfrak{B}^{\perp\perp}$, the smallest observable behavior including $\mathfrak{B}$.

Let $\mathfrak{B} = \mathrm{Im}G$ and $\mathfrak{B}_0 := \mathcal{B}^{\perp\perp} = \ker H^T$. Clearly, $\mathfrak{B}_0$ is an observable behavior including $\mathfrak{B}$. If G has rank r, we can assume, for sake of simplicity, that its first r rows are linearly independent. So, G can be partitioned as

$$G = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} \begin{matrix} \}r \\ \}p-r \end{matrix} \quad ,$$

where G_1 is a full rank matrix. Let

$$G_{\mathrm{lu}} := \begin{bmatrix} G_1 & 0 \\ 0 & I_{p-r} \end{bmatrix},$$

and $\mathfrak{B}_{\mathrm{lu}} := \mathrm{Im}G_{\mathrm{lu}}$. Clearly, $\mathfrak{B}_{\mathrm{lu}}$ is a locally undetectable behavior, and it includes $\mathfrak{B}$ as

$$G = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} = \begin{bmatrix} G_1 & 0 \\ 0 & I_{p-r} \end{bmatrix} \begin{bmatrix} I \\ G_2 \end{bmatrix}.$$

So, one obviously gets $\mathfrak{B} \subset \mathfrak{B}_0 \cap \mathfrak{B}_{\rm lu}$.
To prove the reverse inclusion, consider $\mathbf{w} \in \mathfrak{B}_0 \cap \mathcal{B}_{\rm lu}$. Clearly, $\mathbf{w}$ satisfies $H^T\mathbf{w} = 0$ and can be expressed as

$$\mathbf{w} = \begin{bmatrix} G_1\mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}.$$

Factoring G into the product of a (full column rank) right factor prime matrix $\bar{G}$ and a full row rank rational matrix Q [16], one gets

$$\begin{bmatrix} G_1 \\ G_2 \end{bmatrix} = G = \bar{G}Q = \begin{bmatrix} \bar{G}_1 \\ \bar{G}_2 \end{bmatrix} Q.$$

As the columns of $\bar{G}$ generate the $\mathbb{F}(\mathbf{z})$-vector space orthogonal to the rows of H^T, there exists $\mathbf{v} \in \mathbb{F}(\mathbf{z})^r$ s.t. $\mathbf{w} = \bar{G}\mathbf{v}$. But then $\bar{G}_1\mathbf{v} = G_1\mathbf{u}_1 = \bar{G}_1 Q\mathbf{u}_1$ implies $\mathbf{v} = Q\mathbf{u}_1$, and thus $\mathbf{u}_2 = \bar{G}_2\mathbf{v} = \bar{G}_2 Q\mathbf{u}_1$ and

$$\mathbf{w} = \begin{bmatrix} G_1\mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix} = \begin{bmatrix} \bar{G}_1 \\ \bar{G}_2 \end{bmatrix} Q\mathbf{u}_1 = G\mathbf{u}_1.$$

This implies that $\mathbf{w}$ is in $\mathfrak{B}$.

It remains to prove the uniqueness of $\mathfrak{B}_0$ in the above representation. To this end we need the following technical lemma.

Let $\mathfrak{B}_i \subset \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p$, $i = 1, 2$, be finite support behaviors with p components. If $\mathfrak{B} = \mathfrak{B}_1 \cap \mathfrak{B}_2$, then

$$\mathfrak{B}_{\rm rat} = (\mathfrak{B}_1)_{\rm rat} \cap (\mathfrak{B}_2)_{\rm rat}. \tag{4.3}$$

Let G, G_1 and G_2 be generator matrices of $\mathfrak{B}$, $\mathcal{B}_1$ and $\mathfrak{B}_2$, respectively. Clearly, $\mathfrak{B}_{\rm rat} = {\rm Im}_{\mathbb{F}(\mathbf{z})} G := \{\mathbf{v} \in \mathbb{F}(\mathbf{z})^p : \mathbf{v} = G\mathbf{u}, \mathbf{u} \text{ rational}\}$, and similarly $(\mathfrak{B}_i)_{\rm rat} = {\rm Im}_{\mathbb{F}(\mathbf{z})} G_i$, $i = 1, 2$. Therefore, $\mathbf{v} \in \mathfrak{B}_{\rm rat}$ implies $\mathbf{v} = G\mathbf{n}/d$, for some L-polynomial vector $\mathbf{n}$ and some L-polynomial d, and hence $d\mathbf{v} \in \mathfrak{B} = \mathfrak{B}_1 \cap \mathfrak{B}_2$. But then, $d\mathbf{v} = G_1\mathbf{u}_1 = G_2\mathbf{u}_2$, for suitable L-polynomial vectors $\mathbf{u}_1$ and $\mathbf{u}_2$, which implies $\mathbf{v} \in (\mathfrak{B}_1)_{\rm rat} \cap (\mathfrak{B}_2)_{\rm rat}$.
Conversely, if $\mathbf{v} \in (\mathfrak{B}_1)_{\rm rat} \cap (\mathfrak{B}_2)_{\rm rat}$, it can be expressed as $\mathbf{v} = G_1\mathbf{n_1}/d_1 = G_2\mathbf{n_2}/d_2$, for suitable L-polynomial vectors $\mathbf{n}_i$ and L-polynomials d_i, $i = 1, 2$. If $d := \ell.\text{c.m.}(d_1, d_2)$, it is clear that $d\mathbf{v}$ is an element of $\mathfrak{B}_1 \cap \mathfrak{B}_2$, and hence of $\mathfrak{B}$. Consequently, $\mathbf{v}$ is in $\mathfrak{B}_{\rm rat}$.

We now return to the proof of the uniqueness of $\mathfrak{B}_0$. Suppose, by contradiction, that $\mathfrak{B} = \hat{\mathfrak{B}}_0 \cap \hat{\mathfrak{B}}_{\rm lu}$, for some observable behavior $\hat{\mathfrak{B}}_0 \neq \mathfrak{B}^{\perp\perp}$ and some locally undetectable behavior $\hat{\mathfrak{B}}_{\rm lu}$. As $\mathfrak{B}^{\perp\perp}$ is the smallest observable behavior including $\mathfrak{B}$ and is maximal in the class of modules of rank r, $\hat{\mathfrak{B}}_0$ must have rank greater than r. Consequently, $(\hat{\mathfrak{B}}_0)_{\rm rat} \underset{\neq}{\supset} (\mathfrak{B}^{\perp\perp})_{\rm rat}$. On the other hand

$$(\mathfrak{B}_{\rm lu})_{\rm rat} = (\hat{\mathfrak{B}}_{\rm lu})_{\rm rat} = \mathbb{F}(\mathbf{z})^p,$$

and therefore $(\mathfrak{B}^{\perp\perp})_{\rm rat} \cap (\mathfrak{B}_{\rm lu})_{\rm rat} = (\mathfrak{B}^{\perp\perp})_{\rm rat} \subsetneq (\hat{\mathfrak{B}}_0)_{\rm rat} = (\hat{\mathfrak{B}}_0)_{\rm rat} \cap (\hat{\mathfrak{B}_{\rm lu}})_{\rm rat}$. But this is not possible, as $\mathfrak{B}^{\perp\perp} \cap \mathfrak{B}_{\rm lu} = \mathfrak{B} = \hat{\mathfrak{B}}_0 \cap \hat{\mathfrak{B}_{\rm lu}}$ should imply, by the above Lemma, $(\mathfrak{B}^{\perp\perp})_{\rm rat} \cap (\mathfrak{B}_{\rm lu})_{\rm rat} = (\hat{\mathfrak{B}}_0)_{\rm rat} \cap (\hat{\mathfrak{B}_{\rm lu}})_{\rm rat}$.

5 Input-output description and trajectory generation

The analysis we carried out in the previous sections focused on the properties of behavior trajectories, without concern for the way they are generated. Once a behavior $\mathfrak{B}$ is represented via a finite set of generators $\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_m$, however, it is natural to look at $G := [\mathbf{g}_1\ \mathbf{g}_2\ ...\ \mathbf{g}_m]$ as a transfer matrix, and hence to consider $\mathfrak{B}$ as the image of an input-output map acting on $\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^m$. This point of view seems particularly appropriate when $\mathfrak{B}$ is a convolutional code [3, 15], as it is customary to regard it as the result of an encoding process, and, consequently, its trajectories (codewords) as the outputs of a dynamical encoder. In a wider context, the trajectories of $\mathfrak{B}$ are obtained from certain processing operations applied to multidimensional data, or from different transformations (desired or not) performed on the original signal. In both cases the analysis of the algebraic properties of the generator matrices makes possible a detailed knowledge of the behavior structure.

When an input/output description is adopted, it is often imperative to associate trajectories of $\mathfrak{B}$ and input sequences bijectively. In data transmission the meaning of this requirement is clear, as input signals represent information messages to be retrieved from the received codewords, and an unambiguous decision at the decoding stage is possible when each codeword encodes a unique information sequence. This amounts to saying that the *encoder* G induces a 1-1 map.

Throughout this section we steadily assume that $\mathfrak{B}$ has a full column rank generator matrix G, and hence is free. Under this assumption, G admits (possibly infinitely many) rational left inverses G^{-1}. Each of them, when applied to a behavior trajectory $\mathbf{w} = G\mathbf{u}$, uniquely retrieves the (finite) input sequence $\mathbf{u}$. When $\mathfrak{B}$ represents a finite convolutional code this implies that every estimate $\hat{\mathbf{w}} \in \mathfrak{B}$ of the codeword $\mathbf{w}$ produces a finite error $\mathbf{e}_\mathbf{u} := \mathbf{u} - G^{-1}\hat{\mathbf{w}} = G^{-1}(\mathbf{w} - \hat{\mathbf{w}})$ in reconstructing the information sequence $\mathbf{u}$. Consequently, when a codeword estimator is available, no catastrophic error can arise [3, 5]. However, if we apply the "decoder" G^{-1} directly to the noisy sequence $\mathbf{v} = \mathbf{w} + \mathbf{r}$, as $\mathbf{r}$ generally is not an element of $\mathfrak{B}$, the decoding algorithm possibly gives an infinite support sequence, which differs from the correct input in infinitely many points and clearly is not even an admissible information sequence. This drawback can be avoided if and only if G^{-1} is an L-polynomial matrix.

Proposition 5.1 below provides equivalent conditions for the existence of an

L-polynomial inverse, and in particular shows that such an inverse exists if and only if G is left zero-prime.

Let G be in $\mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^{p\times m}$ and $\hat{G}=\mathbf{z}^{\mathbf{h}}G=z_1^{h_1}\cdots z_n^{h_n}G$ in $\mathbb{F}[\mathbf{z}]^{p\times m}$ for some $\mathbf{h}\in\mathbb{N}^n$. If $\mathbb{K}$ denotes the algebraic closure of $\mathbb{F}$, the L-variety $\mathcal{V}^L(G)$ of the maximal order minors of G is the algebraic set

$$\mathcal{V}^L(G):=\mathcal{V}(\hat{G})\setminus\Big\{(k_1,k_2,...,k_n):k_i\in\mathbb{K},\ \prod_i k_i=0\Big\},\tag{5.1}$$

where $\mathcal{V}(\hat{G})$ denotes the variety (in $\mathbb{K}$) of the maximal order minors of $\hat{G}$.
The above definition is well-posed, as (5.1) does not depend on the choice of $\hat{G}$.

Let G be a $p\times m$ L-polynomial matrix. The following statements are equivalent:
i) G is right zero-prime;
ii) there exists an L-polynomial matrix $P\in\mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^{m\times p}$ such that $PG=I_m$;
iii) $\mathcal{V}^L(G)$ is empty;
iv) Im $G^T=\mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]^m$.

i) $\Rightarrow$ *ii)* Let $m_i(G)$ denotes the i-th maximal order minor of G, $i=1,2,...,\binom{p}{m}$. By the zero-primeness assumption, there exist L-polynomials $h_i\in\mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]$ such that $\sum_i h_i m_i(G)=1$. If S_i is the $m\times p$ matrix which selects in G the m rows corresponding to $m_i(G)$, from $I_m=\sum_i h_i m_i(G)I_m=\sum_i h_i\big(\mathrm{adj}(S_iG)\big)(S_iG)$, we find that $P:=\sum_i h_i\big(\mathrm{adj}(S_iG)\big)S_i$ is a left inverse of G with elements in $\mathbb{F}[\mathbf{z},\mathbf{z}^{-1}]$.

ii) $\Rightarrow$ *iii)* Let $\mathbf{z}^{\mathbf{r}}=z_1^{r_1}\cdots z_n^{r_n}$ be a suitable term such that $\hat{P}=\mathbf{z}^{\mathbf{r}}P$ is in $\mathbb{F}[\mathbf{z}]^{m\times p}$. By applying the Binet-Cauchy formula [8] to equation $\hat{P}\hat{G}=z_1^{h_1+r_1}\cdots z_n^{h_n+r_n}I_m$, we get

$$\sum_i m_i(\hat{P})m_i(\hat{G})=z_1^{m(h_1+r_1)}z_2^{m(h_2+r_2)}\cdots z_n^{m(h_n+r_n)},\tag{5.2}$$

where $m_i(\hat{P})$ and $m_i(\hat{G})$ are corresponding maximal order minors of $\hat{P}$ and $\hat{G}$, respectively. Then $\mathcal{V}(\hat{G})$ is included in the variety of $z_1^{m(h_1+r_1)}z_2^{m(h_2+r_2)}\cdots z_n^{m(h_n+r_n)}$, which is a subset of $\mathcal{K}:=\{(k_1,k_2,...,k_n):k_i\in\mathbb{K},\ \prod_i k_i=0\}$.

iii) $\Rightarrow$ *i)* As $\mathcal{K}$ is the variety of $\mathbf{z}=z_1\cdots z_n$, by assumption *iii)* $\mathcal{V}(\hat{G})$ is included in the variety of $\mathbf{z}$. So, by Hilbert's Nullstellensatz [10], an integer $r>0$ exists such that $z_1^r\cdots z_n^r$ belongs to the ideal generated in $\mathbb{F}[\mathbf{z}]$ by the maximal order minors of $\hat{G}$:

$$z_1^r\cdots z_n^r=\sum_i\bar{h}_i\ m_i(\hat{G}),\qquad \bar{h}_i\in\mathbb{F}[\mathbf{z}].\tag{5.3}$$

As each maximal order minor $m_i(\hat{G})$ differs from $m_i(G)$ in a unit of $\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]$, the zero-primeness of G easily follows after dividing both members of (5.3) by $z_1^r \cdots z_n^r$.

ii) $\Leftrightarrow$ *iv)* Obvious.

When the generator matrix G has an L-polynomial inverse, a uniform bound can be found on the support of the input sequences which correspond to the behavior trajectories. Actually, if P is such an inverse, $\mathbf{w} \in \mathfrak{B}$ is generated by the input signal $\mathbf{u} = P\mathbf{w}$ whose support cannot exceed "too much" that of $\mathbf{w}$. This feature, we will refer to as *wrapping input property*, is quite appealing, as the mere recognition of the support of a trajectory allows the derivation of a uniformly tight bound on the support of the corresponding input sequence. In particular, in the context of finite convolutional codes, the above property guarantees that small errors in the codeword estimate reflect into small errors in the information sequence reconstruction.

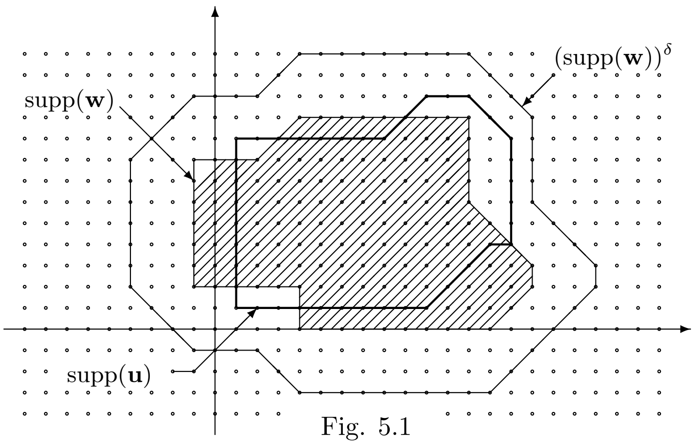

Fig. 5.1

(WI) [Wrapping input property] *A finite behavior* $\mathfrak{B}$ *has the* wrapping input property *if there exist a full column rank generator matrix G and a positive integer δ such that* $\mathbf{w} = G\mathbf{u}$ *implies*

$$\mathrm{supp}(\mathbf{u}) \subseteq \big(\mathrm{supp}(\mathbf{w})\big)^{\delta}. \tag{5.4}$$

It is worthwhile to notice that property (WI) does not depend on the particular full column rank generator matrix of $\mathfrak{B}$ we are considering. In fact, it is easily seen that if (5.4) holds for anyone of these generator matrices, then it holds for all of them (in general for a different δ). On the other hand, when non-injective generator matrices of $\mathfrak{B}$ are considered, and the uniqueness of the input sequence producing a given trajectory is lost, a particular input can be found, however, whose support satisfies (5.4), as shown by the following proposition.

Assume that $\mathfrak{B}$ has the (WI) property for some full column rank matrix G and some integer $\delta > 0$. Then for every generator matrix $\bar{G} \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^{p\times q}$ an integer $\bar{\delta} > 0$ can be found s.t. each trajectory $\mathbf{w} \in \mathfrak{B}$ can be expressed as $\mathbf{w} = \bar{G}\bar{\mathbf{u}}$ for some input $\bar{\mathbf{u}}$ with $\mathrm{supp}(\bar{\mathbf{u}}) \subseteq \big(\mathrm{supp}(\mathbf{w})\big)^{\bar{\delta}}$.

Since G and $\bar{G}$ are generator matrices of the same behavior, there exists a full column rank L-polynomial matrix Q, such that $G = \bar{G}Q$. Let τ be the radius of a ball, with center in the origin, including $\mathrm{supp}(Q)$, and consider $\mathbf{w} \in \mathfrak{B}$. By property (WI), there is $\mathbf{u}$ such that $\mathbf{w} = G\mathbf{u}$ and $\mathrm{supp}(\mathbf{u}) \subseteq \big(\mathrm{supp}(\mathbf{w})\big)^{\delta}$. So, $\bar{\mathbf{u}} := Q\mathbf{u}$ satisfies $\mathbf{w} = G\mathbf{u} = \bar{G}Q\mathbf{u} = \bar{G}\bar{\mathbf{u}}$, and $\mathrm{supp}(\bar{\mathbf{u}}) = \mathrm{supp}(Q\mathbf{u}) \subseteq \big(\mathrm{supp}(\mathbf{u})\big)^{\tau} \subseteq \big(\mathrm{supp}(\mathbf{w})\big)^{\tau+\delta}$. Consequently, the proposition holds for $\bar{\delta} = \tau + \delta$.

Interestingly enough, the zero primeness of G is not only sufficient but also necessary for property (WI). So, free behaviors satisfying property (WI) can be identified with behaviors that are generated by ℓZP matrices.

A finite behavior $\mathfrak{B}$ has the (WI) property if and only if it admits a right zero-prime generator matrix.

The "if" part has already been proved. To show the converse, we need the following characterization of rZP matrices.

Let $G \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^{p\times q}$ be a Laurent polynomial matrix and denote by $\mathbb{F}[[\mathbf{z}, \mathbf{z}^{-1}]]$ the space of bilateral scalar formal power series in the indeterminates $z_1, \dots, z_n$. Then G is right zero prime if and only if

$$G\mathbf{s} = \mathbf{0} \tag{5.5}$$

for some sequence $\mathbf{s} \in \mathbb{F}[[\mathbf{z}, \mathbf{z}^{-1}]]^q$ implies $\mathbf{s} = \mathbf{0}$.

Introduce in $\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^q \times \mathbb{F}[[\mathbf{z}, \mathbf{z}^{-1}]]^q$ the following nondegenerate bilinear form

$$\langle \cdot, \cdot \rangle_q : \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^q \times \mathbb{F}[[\mathbf{z}, \mathbf{z}^{-1}]]^q \to \mathbb{F}$$

defined by : $\langle \mathbf{u}, \mathbf{v} \rangle_q = (\mathbf{u}\mathbf{v}^T, 1) = \sum_{\mathbf{h}\in\mathbb{Z}^n} u(\mathbf{h})v^T(-\mathbf{h})$.
With this position, the space $\mathbb{F}[[\mathbf{z}, \mathbf{z}^{-1}]]^q$ is naturally viewed as $L(\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^q)$, the algebraic dual of $\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^q$ [9, 15]. In fact, we can associate with every $\mathbf{v} \in \mathbb{F}[[\mathbf{z}, \mathbf{z}^{-1}]]^q$ the linear functional on $\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^q$ defined by

$$f_{\mathbf{v}}(\cdot) = \langle \cdot, \mathbf{v} \rangle_q \tag{5.6}$$

and, conversely, every linear functional on $\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^q$ can be represented as in (5.6) for an appropriate choice of $\mathbf{v} \in \mathbb{F}[[\mathbf{z}, \mathbf{z}^{-1}]]^m$. Upon identifying $\mathbb{F}[[\mathbf{z}, \mathbf{z}^{-1}]]^q$ with $L(\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^q)$, we can resort to the well known relation

$$\begin{aligned} \ker_\infty G &:= \{\mathbf{s} \in \mathbb{F}[[\mathbf{z}, \mathbf{z}^{-1}]]^q : G\mathbf{s} = \mathbf{0}\} \\ &\equiv \{\mathbf{s} \in \mathbb{F}[[\mathbf{z}, \mathbf{z}^{-1}]]^q : \mathbf{s}^T\mathbf{v} = \mathbf{0},\ \forall \mathbf{v} \in \mathrm{Im}G^T\} =: (\mathrm{Im}G^T)^{\perp}. \end{aligned}$$

If $\mathrm{Im}G^T = \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^q$, all canonical vectors $\mathbf{e}_i$ and all monomial vectors $z_1^h z_2^k \mathbf{e}_i$ belong to $\mathrm{Im}G^T$, and therefore $\mathbf{s} \in (\mathrm{Im}G^T)^{\perp\!\!\!\perp}$ implies $\langle z_1^h z_2^k \mathbf{e}_i, \mathbf{s}\rangle_q = 0$, $h, k \in \mathbb{Z}$ and $i = 1, 2, \ldots, q$, and hence $\mathbf{s} = 0$. So, it is clear that $\ker_\infty G = \{\mathbf{0}\}$ if and only if $\mathrm{Im}G^T = \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^q$, and this happens iff G is rZP.

Suppose, now, that $\mathfrak{B}$ has the (WI) property w.r.t. some positive integer δ and some full column rank generator matrix G. We aim to prove that G is rZP. If not, there would be a sequence $\mathbf{s} \in \mathbb{F}[[\mathbf{z}, \mathbf{z}^{-1}]]^q$ satisfying (5.5). Let η be the radius of a ball, $B(\mathbf{0}, \eta)$, centered in the origin and including $\mathrm{supp}(G)$. If $\mathbf{k}$ is an element of $\mathrm{supp}(\mathbf{s})$, the finite support sequence

$$\mathbf{u}(\mathbf{h}) := \begin{cases} \mathbf{s}(\mathbf{h}) & \mathbf{h} \in B(\mathbf{k}, 2\delta + \eta) \\ \mathbf{0} & \text{elsewhere} \end{cases}$$

generates a behavior sequence $\mathbf{w} := G\mathbf{u}$ that does not fulfill (5.4).

The (WI) property introduces very severe constraints on the supports of the input sequences which produce the behavior trajectories. So, it is not unexpected that it reflects into the strongest primeness property a generator matrix can be endowed with, namely zero-primeness. Obviously, weaker requirements on the supports of the generating sequences correspond to weaker primeness properties of G. In particular, minor primeness guarantees that the signal producing a behavior sequence $\mathbf{w}$ exhibits a support which slightly exceeds a parallelepipedal box including $\mathrm{supp}(\mathbf{w})$, whereas variety primeness ensures that each projection of $\mathbf{u}$ and $\mathbf{w}$ onto a coordinate hyperplane gives a pair of signals with the (WI) property.

A standpoint which proves to be quite fruitful in analysing the above-mentioned connections is to regard an arbitrary finite support sequence $\mathbf{w} \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^p$ as a vector with entries in certain L-polynomial rings that properly include $\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]$. Actually, $\mathbf{w}$ can be thought of as an element of $\mathbb{F}(z_i^c)[z_i, z_i^{-1}] := \mathbb{F}(z_1, .., z_{i-1}, z_{i+1}, .., z_n)[z_i, z_i^{-1}]$

$$\mathbf{w} = \sum_{h_i \in \mathbb{Z}} \mathbf{w}_{h_i}(z_i^c) z_i^{h_i},$$

or as an element of $\mathbb{F}(z_i)[z_i^c, (z_i^c)^{-1}] := \mathbb{F}(z_i)[z_1, ..., z_{i-1}, z_{i+1}, ..., z_n, z_1^{-1}, .., z_{i-1}^{-1}, z_{i+1}^{-1}, .., z_n^{-1}]$

$$\mathbf{w} = \sum_{h_1,...,h_{i-1},h_{i+1},..,h_n \in \mathbb{Z}} \mathbf{w}_{h_1,..,h_{i-1},h_{i+1},..,h_n}(z_i) z_1^{h_1} \cdots z_{i-1}^{h_{i-1}} z_{i+1}^{h_{i+1}} \cdots z_n^{h_n}.$$

Correspondingly, we are lead to introduce the following support sets

$$\begin{aligned} \mathrm{supp}_i(\mathbf{w}) &:= \{(h_1, ..., h_n) \in \mathbb{Z}^n : \mathbf{w}_{h_i}(z_i^c) \neq 0\} \\ \mathrm{supp}_{i^c}(\mathbf{w}) &:= \{(h_1, ..., h_n) \in \mathbb{Z}^n : \mathbf{w}_{h_1,...,h_{i-1},h_{i+1},...,h_n}(z_i) \neq 0\}. \end{aligned}$$

[16] Let $G \in \mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]^{p \times m}$ be a full column rank matrix. Then
i) G is rMP if and only if G is right (zero) prime in $\mathbb{F}(z_i^c)[z_i, z_i^{-1}]$ for every $i = 1, 2, ..., n$;
ii) G is rVP if and only if G is rZP in $\mathbb{F}(z_i)[z_i^c, (z_i^c)^{-1}]$, for every $i = 1, 2, ..., n$.

Let $\mathfrak{B}$ be a finite behavior. Then
i) $\mathfrak{B}$ has a rMP generator matrix if and only if there exist an integer $\delta > 0$ and a full column rank generator matrix G, such that $\mathbf{w} \in \mathfrak{B}$ implies $\mathbf{w} = G\mathbf{u}$ with

$$\text{supp}(\mathbf{u}) \subseteq \bigcap_{i=1}^{n} \Big(\text{supp}_i(\mathbf{w})\Big)^{\delta}; \tag{5.7}$$

ii) $\mathfrak{B}$ has a rVP generator matrix if and only if there exist an integer $\delta > 0$ and a full column rank generator matrix G, such that $\mathbf{w} \in \mathfrak{B}$ implies $\mathbf{w} = G\mathbf{u}$ with

$$\text{supp}(\mathbf{u}) \subseteq \bigcap_{i=1}^{n} \Big(\text{supp}_{i^c}(\mathbf{w})\Big)^{\delta}. \tag{5.8}$$

i) It is easy to realize that condition (5.7) is equivalent to the set of conditions $\text{supp}_i(\mathbf{u}) \subseteq \Big(\text{supp}_i(\mathbf{w})\Big)^{\delta}$, $i = 1, 2, ..., n$. These hold true if and only if G is a rZP matrix in $\mathbb{F}(z_i^c)[z_i, z_i^{-1}]$, for every $i = 1, 2, ..., n$, namely G is rMP in $\mathbb{F}[\mathbf{z}, \mathbf{z}^{-1}]$.
ii) The result is shown along the same lines of *i)*, after replacing $\mathbb{F}(z_i^c)[z_i, z_i^{-1}]$ with $\mathbb{F}(z_i)[z_i^c, (z_i^c)^{-1}]$.

6 Conclusions

In this paper we have focused on some features of finite support nD behaviors which are relevant for multidimensional signal generation and recognition. Two opposite situations have been considered, namely the case when a local testing procedure suffices to decide whether a given signal belongs to the behavior, and the case when every finite signal can be completed into a legal trajectory, and hence behavior sequences cannot be recognized by means of local checks.

Observable and locally undetectable behaviors, which correspond to these two situations, have been characterized both in terms of their internal properties and of their polynomial matrix descriptions. Any finite support behavior, being the intersection of an observable and an unconstrained behavior, exhibits intermediate properties.

Finally, adopting an input/output point of view, the connections between the support of a behavior trajectory and that of its generating input have been enlightened.

Aknowledgments

The authors are indebted with S.Zampieri for shortening the proof of Proposition 5.3.

References

[1] N.K.Bose, *Multidimensional Systems Theory*, D.Reidel, Boston, MA, 1985

[2] F.Fagnani and S.Zampieri, *Matrix shifts operators and controllable systems over principal ideal domains*, Sc. Norm. Sup. Pisa, preprints di Matematica, 36 (1994)

[3] E.Fornasini and M.E.Valcher, *Algebraic aspects of 2D convolutional codes*, IEEE Trans.Info.Th., 40 (1994), pp. 1068–1082

[4] E.Fornasini and M.E.Valcher, *On some primeness properties in the analysis of nD finite support behaviors*, in Proc. SPRANN'94 Conf., Lille (F), 1994, pp. 89–92

[5] G.D.Forney, *Convolutional codes I: Algebraic structure*, IEEE Trans. Info. Th., 16 (1970), pp. 720–738

[6] G.D.Forney, *Minimal bases of rational vector spaces, with applications to multivariable linear systems*, SIAM J. of Control, 13 (1975), pp. 493–521

[7] G.D.Forney and M.D.Trott, *The dynamics of group codes: state spaces, trellis diagrams and canonical encoders*, IEEE Trans. Info. Th., 39 (1993), pp. 1491–1513

[8] F.R.Gantmacher, *Theory of Matrices*, Chelsea Publishing Co., New York, NY, 1959

[9] W.Greub, *Linear algebra*, Springer-Verlag, NY, 1975

[10] S.Lang, *Algebra*, Addison-Wesley, Menlo Park, CA, 1993

[11] H.A.Loeliger, G.D.Forney, T.Mittelholzer and M.D.Trott, *Minimality and observability of group systems*, Lin. Alg. and Appl., 205–206, (1994), pp. 937–963

[12] P.Rocha, *Structure and Representation of 2D Systems*, Ph.D.Thesis, Rijksuniversiteit Groningen, NL, 1990

[13] P.Rocha and J.C.Willems, *Controllability of 2D systems*, IEEE Trans. Aut. Contr., 36 (1991), pp. 413–423

[14] S.Sakata, *On determining the independent point set for doubly periodic arrays and encoding two-dimensional cyclic codes and their duals*, IEEE Trans.Info.Th., 27 (1981), pp. 556–565

[15] M.E.Valcher and E.Fornasini, *On 2D finite support convolutional codes: an algebraic approach*, Multid. Sys. Sign.Process., 5 (1994), pp. 231–243

[16] M.E.Valcher, *Modellistica ed Analisi dei Sistemi 2D con Applicazioni alla Codifica Convoluzionale*, Ph.D.Thesis, Univ.of Padova, I, 1995

[17] J.C.Willems, *From time series to linear systems, part I: finite dimensional linear time invariant systems*, Automatica, 22 (1986), pp. 561–580

[18] J.C.Willems, *Models for Dynamics*, in Dynamics Reported, vol.2, U.Kirchgraber and H.O.Walther eds., J.Wiley & Sons Ltd. and B.G.Teubner, 1989, pp. 171–269

[19] J.C.Willems, *Paradigms and puzzles in the theory of dynamical systems*, IEEE Trans. Aut.Contr., 36 (1991), pp. 259–294

[20] D.C.Youla and G.Gnavi, *Notes on n-dimensional system theory*, IEEE Trans. Circ. and Sys., 26 (1979), pp. 105–111

[21] S.Zampieri and S.K.Mitter, *Linear systems over Noetherian rings in the behavioral approach*, J. of Math. Syst., Estimation and Control, 6 (1996), pp. 235–238