Multidimensional behaviors: the state-space paradigm (special issue JCW)☆

Debasattam Pal[a], Harish K. Pillai[a]

[a] *Department of Electrical Engineering, Indian Institute of Technology Bombay, Mumbai 400076, India*

Abstract

For 1-D systems, the state-space approach has perhaps become the most popular method of analyzing these systems. There have been several attempts to imitate the state-space framework for n-D systems. Introduction of behavioral theory by Jan C. Willems, has given fresh impetus to this attempt to imitate state-space framework for n-D systems. In this paper, dedicated to Jan Willems, we provide our recent attempt at obtaining a state-space framework for n-D systems.

Keywords: Behaviors, n-D systems, State space, Noether normalization

1. Introduction

Jan C. Willems, in his celebrated paper *Paradigms and puzzles in the theory of dynamical systems* [1], wrote: "In engineering, particularly in control and signal processing, there has always been a tendency to view systems as *processors*, producing output signals from input signals. In many applications in control engineering and signal processing, it will, indeed, be eminently clear what the inputs and the outputs are. However, there are also many applications where this input-output structure is not at all evident (an example at point is in the terminal behavior of an electrical circuit)." Willems pointed out a number of situations where it is indeed impractical to assume an input/output structure on the system variables. Examples include Kepler's laws of planetary motion, econometrics [2], economics (relation between production, capital cost and labor cost) and discrete event systems [1]. Willems also argued that there are dynamical systems (e.g., *Leontief economy* [2]), for which it is outright *impossible* to obtain an input/output model. In a series of works [1, 2, 3, 4, 5, 6] Willems brought about a radical change in the way a mathematical model for dynamical systems should be viewed. He showed that systems viewed as maps from inputs (plus initial conditions) to outputs is perhaps not the most suited approach as a modelling paradigm for dynamical systems. There is in fact a more fundamental object, the *behavior* of the system – that is, the collection of trajectories allowed by the laws of physics applied to the system – that provides a better mathematical model for dynamical systems. With this rudimentary object – the behavior – Willems succeeded in deducing – and refining where required – several existing notions of dynamical systems: linearity, shift/time-invariance, input/output representation, autonomy, controllability, observability, stability, stabilizability, detectability, state-variables etc. This was a remarkable feat, a *paradigm*

☆This work has been supported in part by IRCC, IIT Bombay; and INSPIRE Faculty Grant by Indian National Science Academy and Department of Science and Technology, Govt. of India.

Email addresses: `debasattam@ee.iitb.ac.in` (Debasattam Pal), `hp@ee.iitb.ac.in` (Harish K. Pillai)

shift, indeed. This approach became known as the *behavioral approach* to systems theory.

The limitations of the input/output approach become much more exposed for *multidimensional systems* (also called *nD systems*) – that is, dynamical systems with more than one independent variables. For example, consider the system described by the following partial differential equation (PDE):

$$\frac{\partial w_1}{\partial x_1} - \frac{\partial w_2}{\partial x_2} = 0.$$

Although one may view w_2 as an input and w_1 as an output, modeling the system as a mapping from w_2 to w_1 is fraught with technical problems. (For instance, the 'transfer function' here would be $\frac{s_2}{s_1}$, which has numerator and denominator sharing a common root! See [7, Remark 76] for a more elaborate discussion on this.) However, the behavior (that is, the set of solutions of the above equation) still exists, and hence, many system-theoretic questions posed in behavioral approach of 1D systems would make perfect sense for this system too. Spurred by Willems' treatment of 1D systems, issues like autonomy, controllability, observability, stability for nD systems were tackled and resolved using the behavioral approach (see [8] for $n = 2$, and [7, 9, 10] among others for general n). With this development, the behavioral approach has become one of the strongest contenders for a *grand unified theory* of systems and control.

Despite the success of the behavioral approach to nD systems, the current state-of-the-art lags far behind its 1D counterpart. On several issues, for which there has been a well-accepted solution in 1D systems, for nD systems a successful resolution has either completely evaded the community, or there have been many resolutions none of which were universally acceptable. Defining state-variables and obtaining state-space representations from a given representation of an nD system is one such issue that is still largely open. In this paper, we hope to provide a partial answer to this issue. State-space representation of nD systems has been an active field of research, especially for $n = 2$; see [11, 12, 13, 14, 8] among many notable works. However, these works suffer from a crucial drawback: their restricted applicability. Indeed, in each of the earlier works, several restrictive assumptions were made. For example, in [12, 13, 14] that deal with state-space models for discrete 2D systems, the systems concerned are assumed to satisfy a certain notion of *causality* in 2D integer grid. Needless to say, many 2D systems do not satisfy this assumption. In this paper, we provide a methodology to construct state-space and a first order evolution law for general nD systems that are described by linear partial differential/difference equations with constant real coefficients; we make only one assumption: the system is autonomous.

2. Background

2.1. nD systems

Following Willems, we define a dynamical system by a triplet $(\mathbb{T}, \mathbb{W}, \mathfrak{B})$, where $\mathbb{T}$ is the *indexing set* (the set of independent variables over which the system's variables, w, evolve), $\mathbb{W}$ is the *signal space* (the set from where the manifest variables take values), and $\mathfrak{B}$ is the *behavior* of the system (the subset of the set of all possible trajectories, $\mathbb{W}^{\mathbb{T}}$, that are allowed by the system). In this paper, we shall assume $\mathbb{W} = \mathbb{R}^{\mathtt{w}}$; $\mathtt{w}$ denotes the cardinality of the vector w. These variables w are called *manifest variables. Multidimensional (nD) systems* are characterized by the fact that they have n independent variables; that is, the indexing set $\mathbb{T}$ is either $\mathbb{Z}^n$ or $\mathbb{R}^n$. We shall use the term *continuous or discrete nD systems* for the case when $\mathbb{T} = \mathbb{R}^n$ or $\mathbb{T} = \mathbb{Z}^n$, respectively. The letter $\boldsymbol{t}$ will be used to denote the independent variable; that is, $\boldsymbol{t} \in \mathbb{R}^n$ for continuous systems, and $\boldsymbol{t} \in \mathbb{Z}^n$ for discrete systems. In this paper, we are going to look at a special kind of nD systems, namely, systems that are described by linear partial differential/difference equations with constant real coefficients.

Behaviors of continuous nD systems are sets of solutions to partial differential equations (PDEs). Such PDEs are written using polynomials in partial differential operators. Let ∂_i denote the partial differential operator with respect to the variable t_i, that is, $\partial_i = \frac{\partial}{\partial t_i}$. The polynomial ring in the variables $\{\partial_1, \dots, \partial_n\}$ is denoted by $\mathbb{R}[\partial_1, \dots, \partial_n]$. We often use the short-hand ∂ to denote the n-tuple $\{\partial_1, \dots, \partial_n\}$. The idea of solutions of PDEs intrinsically depends on the function space where solutions are sought. In this paper, we shall consider the space of *smooth functions*, denoted by $\mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}})$. Thus, a behavior $\mathfrak{B}$ of a continuous nD system, described by a set of linear partial differential equations with constant real coefficients, can be defined as

$$\mathfrak{B} := \{w \in \mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}}) \mid R(\partial)w = 0\}, \tag{1}$$

where $R(\partial) \in \mathbb{R}^{\bullet \times \mathtt{w}}[\partial]$. For obvious reasons, equation (1) is called a *kernel representation* of $\mathfrak{B}$ and $R(\partial)$ is called a *kernel representation matrix* of $\mathfrak{B}$. We write $\mathfrak{B} = \ker R(\partial)$ for brevity.

For discrete nD systems, the role of ∂_is is played by the shift operators, σ_is. In this case, the function space that we consider is the space of vector valued ($\mathtt{w}$ tuple) sequences indexed by $\mathbb{Z}^n$, i.e., $(\mathbb{R}^{\mathtt{w}})^{\mathbb{Z}^n} = \{w : \mathbb{Z}^n \to \mathbb{R}^{\mathtt{w}}\}$. In this paper, we use the symbol $\mathcal{W}(\mathbb{Z}^n, \mathbb{R}^{\mathtt{w}})$ to denote this space. The i^{th} shift operator σ_i acts on a discrete trajectory $w \in \mathcal{W}(\mathbb{Z}^n, \mathbb{R}^{\mathtt{w}})$ as

$$(\sigma_i w)(t_1, \dots, t_n) = w(t_1, \dots, t_i + 1, \dots, t_n), \tag{2}$$

for all $(t_1, \dots, t_n) \in \mathbb{Z}^n$. Note that σ_i^{-1} is a legitimate operator on $\mathcal{W}(\mathbb{Z}^n, \mathbb{R}^{\mathtt{w}})$. Thus, unlike the continuous case, the operator algebra for the discrete case contains polynomials having terms with (finite) positive as well as (finite) negative powers. Therefore, the operator algebra, in this case, is given by the n-variable *Laurent polynomial ring* in the variables $\{\sigma_1, \dots, \sigma_n\}$. We denote this ring by $\mathbb{R}[\sigma_1, \sigma_1^{-1}, \dots, \sigma_n, \sigma_n^{-1}]$. Like in the case of partial differential operators, we shall use the singleton σ to denote the n-tuple $\{\sigma_1, \dots, \sigma_n\}$, and, likewise, we shall write $\mathbb{R}[\sigma, \sigma^{-1}]$ to denote the n-variable Laurent polynomial ring in the shifts $\{\sigma_1, \dots, \sigma_n\}$. Consequently, a behavior $\mathfrak{B}$ of a discrete nD system, which is the solution set of a system of partial difference equations gets defined as

$$\mathfrak{B} := \{w \in \mathcal{W}(\mathbb{Z}^n, \mathbb{R}^{\mathtt{w}}) \mid R(\sigma)w = 0\}, \tag{3}$$

where $R(\sigma) \in \mathbb{R}^{\bullet \times \mathtt{w}}[\sigma, \sigma^{-1}]$. As in the continuous case, equation (3) is called a *kernel representation* of $\mathfrak{B}$ and $R(\sigma)$ is called a *kernel representation matrix*, while $\mathfrak{B}$ is written in short as $\mathfrak{B} = \ker R(\sigma)$.

It is apparent from the last two paragraphs that continuous and discrete systems share a common model of description – the kernel representation – with only the operator algebras and the function spaces being different. This commonality is utilized throughout this paper – so much so that we use common symbols to denote various objects that relate to both discrete and continuous systems. For example, $\mathcal{A}$ is the *operator algebra* ($\mathcal{A} = \mathbb{R}[\partial]$ (continuous), or $\mathbb{R}[\sigma, \sigma^{-1}]$ (discrete)), $\mathcal{F}_n^{\mathtt{w}}$ is the *function space* ($\mathcal{F}_n^{\mathtt{w}} = \mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}})$ (continuous), or $\mathcal{W}(\mathbb{Z}^n, \mathbb{R}^{\mathtt{w}})$ (discrete)), ξ is the n-tuple of operators ($\xi = \partial$ (continuous), or σ (discrete)). We shall use ξ^ν to denote the monomial $\xi_1^{\nu_1} \cdots \xi_n^{\nu_n}$, where $\nu = (\nu_1, \dots, \nu_n) \in \mathbb{Z}^n$. The collection of all nD systems that have $\mathtt{w}$ manifest variables and are described by linear partial differential/difference equations is denoted by $\mathfrak{L}_n^{\mathtt{w}}$. We often abuse this notation and write $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$. Thus, a continuous/discrete behavior $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$ is described as $\mathfrak{B} = \ker R(\xi) \subseteq \mathcal{F}_n^{\mathtt{w}}$, where $R(\xi) \in \mathcal{A}^{\bullet \times \mathtt{w}}$. Another representation of

$\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$, called a *latent variable representation*, is required in the sequel. In this representation, $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$ is described as

$$\mathfrak{B} := \{w \in \mathcal{F}_n^{\mathtt{w}} \mid \exists \ell \in \mathcal{F}_n^r \text{ such that } R(\xi)w = M(\xi)\ell\}\,, \tag{4}$$

where $R(\xi) \in \mathcal{A}^{g\times\mathtt{w}}$ and $M(\xi) \in \mathcal{A}^{g\times r}$. The variables ℓ are called *latent variables.*

2.2. The equation module and the quotient module

Following the trend set by Willems for 1D systems, in this paper, we treat nD systems algebraically via the operator algebra $\mathcal{A}$. In this connection, two algebraic objects are of great importance to the analysis. The first one of these algebraic objects is called the *equation module*, denoted by $\mathcal{R}$, and is defined as follows: suppose $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$ is given by a kernel representation as $\mathfrak{B} = \ker\, R(\xi)$, with $R(\xi) \in \mathcal{A}^{\bullet\times\mathtt{w}}$, then $\mathcal{R}$ is defined to be the set of all (row-)vectors in $\mathcal{A}^{1\times\mathtt{w}}$ that can be written as linear combinations of the rows of $R(\xi)$ with coefficients from $\mathcal{A}$. This $\mathcal{R}$ is a submodule of the $\mathcal{A}$-module $\mathcal{A}^{1\times\mathtt{w}}$. With $\mathcal{R}$ in place, it is easy to see that $\mathfrak{B} = \ker\, R(\xi)$ admits an alternative description given by

$$\mathfrak{B} = \{w \in \mathcal{F}_n^{\mathtt{w}} \mid r(\xi)w = 0 \text{ for all } r(\xi) \in \mathcal{R}\}\,. \tag{5}$$

Thus, given a submodule $\mathcal{R} \subseteq \mathcal{A}^{1\times\mathtt{w}}$, we can define the corresponding behavior $\mathfrak{B}(\mathcal{R})$ by equation (5). The following proposition gives a characterization of the equation module. The complete proof of the proposition can be gathered from several sources like [1, 6, 7, 15].

Proposition 2.1. *Suppose $\mathcal{R}_1, \mathcal{R}_2$ are submodules of $\mathcal{A}^{1\times\mathtt{w}}$, and let $\mathfrak{B}_1$ and $\mathfrak{B}_2$ be the corresponding behaviors. Further, suppose that the function space $\mathcal{F}_n^{\mathtt{w}} = \mathfrak{C}^\infty(\mathbb{R}^n, \mathbb{R}^{\mathtt{w}})$ for continuous systems and $\mathcal{F}_n^{\mathtt{w}} = \mathcal{W}(\mathbb{Z}^n, \mathbb{R}^{\mathtt{w}})$ for discrete systems. Then the following are true:*

1. $\mathfrak{B}_1 \subseteq \mathfrak{B}_2$ *if and only if* $\mathcal{R}_1 \supseteq \mathcal{R}_2$.
2. $\mathfrak{B}_1 = \mathfrak{B}_2$ *if and only if* $\mathcal{R}_1 = \mathcal{R}_2$.
3. $\mathfrak{B}(\mathcal{R}_1 \cap \mathcal{R}_2) = \mathfrak{B}_1 + \mathfrak{B}_2$.
4. $\mathfrak{B}(\mathcal{R}_1 + \mathcal{R}_2) = \mathfrak{B}_1 \cap \mathfrak{B}_2$.

The crux of Proposition 2.1 can be expressed as follows: $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$ and $\mathcal{R} \subseteq \mathcal{A}^{1\times\mathtt{w}}$ are in an inclusion reversing one-to-one correspondence with each other. This is also called a Galois correspondence of lattices.

The second algebraic object of import is called the *quotient module* of $\mathfrak{B}$ and is denoted by $\mathcal{M}$. This module $\mathcal{M}$ is constructed from the equation module $\mathcal{R}$ corresponding to the behavior $\mathfrak{B}$ by defining an equivalence relation $\sim$ in $\mathcal{A}^{1\times\mathtt{w}}$ as follows: $r_1(\xi) \sim r_2(\xi)$ if $r_1(\xi) - r_2(\xi) \in \mathcal{R}$. The quotient module $\mathcal{M}$ then is defined to be the module of equivalence classes under the relation $\sim$; this is denoted by $\mathcal{A}^{1\times\mathtt{w}}/\mathcal{R}$. In this paper, we use the 'bar' notation to denote equivalence class, that is, for $f(\xi) \in \mathcal{A}^{1\times\mathtt{w}}$, the symbol $\overline{f(\xi)}$ denotes the equivalence class of f. The natural map $\mathcal{A}^{1\times\mathtt{w}} \ni f(\xi) \mapsto \overline{f(\xi)} \in \mathcal{M}$ is called the *canonical surjection.* Elements from $\mathcal{M}$ acts on $w \in \mathfrak{B}$. This action is defined as follows: let $g \in \mathcal{M}$, and suppose $f(\xi) \in \mathcal{A}^{1\times\mathtt{w}}$ is such that $\overline{f(\xi)} = g$, then

$$g(w) := f(\xi)w. \tag{6}$$

It can be easily checked that this action is well-defined.

2.3. Free variables and autonomy

The basic premise of the behavioral approach to dynamical systems is that the system's variables w need not be partitioned into inputs and outputs at the time of modelling. The splitting, however, can be *deduced* once the mathematical equations of the system is known. Willems showed this for 1D systems in [5, 1]. Oberst's paper [7] generalizes this for nD systems for general n. This dichotomy of variables into inputs and outputs is based on the notion of *free* variables in a behavior. Freeness of a variable in turn is defined using the *projection* of a behavior in the following manner: suppose $w \in \mathfrak{B}$ is partitioned as $w = (w_1, w_2)$ after possibly permuting the given order of w's components. Suppose, further, that w_1 is a $\mathtt{w}_1$-tuple and w_2 is a $\mathtt{w}_2$-tuple. Then the projection of $\mathfrak{B}$ to the w_1 variables is defined as

$$\Pi_{w_1}(\mathfrak{B}) := \{w_1 \in \mathcal{F}_n^{\mathtt{w}_1} \mid \exists\, w_2 \in \mathcal{F}_n^{\mathtt{w}_2} \text{ such that } (w_1, w_2) \in \mathfrak{B}\}.$$

Definition 2.2. *Suppose $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}_1+\mathtt{w}_2}$ has its variable w partitioned as $w = (w_1, w_2)$ with w_1, w_2 having $\mathtt{w}_1$ and $\mathtt{w}_2$ components, respectively. Then w_1 is said to be* free *in $\mathfrak{B}$ if*

$$\Pi_{w_1}(\mathfrak{B}) = \mathcal{F}_n^{\mathtt{w}_1}.$$

The problem of determining whether a variable is free or not was solved by Willems for the 1D case [1]. A key observation in Willems' proof was that, for 1D systems, a behavior may be assumed without loss of generality to have a kernel representation matrix $R(\xi)$ which is *full row-rank* – that is, no non-trivial linear combination of rows of $R(\xi)$ with coefficients coming from the ring $\mathcal{A}$ can go to zero. With this, Willems showed, if $w \in \mathfrak{B}$ is partitioned as (w_1, w_2), and a full row-rank kernel representation matrix $R(\xi)$ is also partitioned conformally with (w_1, w_2) to give the describing system of equations as

$$R(\xi)w = R_1(\xi)w_1 + R_2(\xi)w_2 = 0 \tag{7}$$

then w_1 is free in $\mathfrak{B}$ if and only if $R_2(\xi)$ is full row-rank over $\mathcal{A}$.

The situation changes drastically for general nD systems with $n \geqslant 2$. In this case, it is not always possible to bring a kernel representation matrix to full row-rank. For example, with $\mathcal{A} = \mathbb{R}[\partial_1, \partial_2, \partial_3]$, the matrix representing the curl operator in Cartesian coordinates

$$R(\partial) = \begin{bmatrix} 0 & -\partial_3 & \partial_2 \\ \partial_3 & 0 & -\partial_1 \\ -\partial_2 & \partial_1 & 0 \end{bmatrix}, \tag{8}$$

has rank 2, but $\mathfrak{B} = \ker R(\partial)$ here cannot be represented by a matrix having two rows. This happens essentially because for $n \geqslant 2$, the ring of operators $\mathcal{A}$ is no longer a principal ideal domain. Because of this issue, the methods used for the 1D case do not extend to nD in a straightforward manner. Necessary and sufficient condition for freeness, in this case, can be given by utilizing the well-known fundamental principle of Ehrenpreis and Palamodov.

Lemma 2.3 (Fundamental Principle). *Consider the equation*

$$A(\xi)w_1 = w_2, \tag{9}$$

where $A(\xi) \in \mathcal{A}^{\mathtt{w}_2 \times \mathtt{w}_1}$ is given, $w_1 \in \mathcal{F}_n^{\mathtt{w}_1}$ is unknown and $w_2 \in \mathcal{F}_n^{\mathtt{w}_2}$ is given. Then there is a solution w_1 to equation (9) *if and only if for every $v(\xi) \in \mathcal{A}^{1 \times \mathtt{w}_2}$ satisfying $v(\xi)A(\xi) = 0 \in \mathcal{A}^{1 \times \mathtt{w}_1}$ it holds true that $v(\xi)w_2 = 0$.*

Using fundamental principle one can now provide a necessary and sufficient condition for freeness. This is the content of Theorem 2.4 below; the proof easily follows from Lemma 2.3.

Theorem 2.4. *Suppose* $\mathfrak{B} = \ker\, R(\xi)$*, where* $w \in \mathfrak{B}$ *is partitioned as* $w = (w_1, w_2)$ *and* $R(\xi) \in \mathcal{A}^{g\times(\mathtt{w}_1+\mathtt{w}_2)}$ *is conformally partitioned as* $R(\xi) = \begin{bmatrix} R_1(\xi) & R_2(\xi) \end{bmatrix}$ *so that the kernel representation is written as equation* (7)*. Then* w_1 *free in* $\mathfrak{B}$ *if and only if for every* $v(\xi) \in \mathcal{A}^{1\times g}$*, we have*

$$v(\xi)R_2(\xi) = 0 \ \Rightarrow\ v(\xi)R_1(\xi) = 0.$$

As an example, consider $R(\partial)$ given in equation (8). For the behavior

$$\mathfrak{B} = \left\{ w = [w_1, w_2, w_3]^T \in \mathfrak{C}^\infty(\mathbb{R}^3, \mathbb{R}^3) \mid R(\partial)w = 0 \right\},$$

each of w_1, w_2, w_3 is individually free, whereas, the variables taken as pairs are not free. In [7], Oberst extends the idea of input/output partitioning to nD systems, and generalizes the notion of transfer functions for such input/output partitioning. However, such transfer functions cannot in general be viewed as maps from inputs to outputs (see [7, Remark 76]).

A 1D behavior $\mathfrak{B}$ having no free variables was shown to be *autonomous* by Willems in [5, 1]; autonomy was defined to be the property that for every $w \in \mathfrak{B}$, the knowledge of w in the past uniquely determines w in the future. The notion of autonomy can be carried forward to nD systems by taking the property of having no free variables as the definition of autonomy. The following result then immediately follows from Theorem 2.4.

Proposition 2.5. *Let* $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$*. Then the following are equivalent:*

1. $\mathfrak{B}$ *is autonomous, that is,* $\mathfrak{B}$ *has no free variables.*
2. $\mathfrak{B} = \ker\, R(\xi)$*, where* $R(\xi)$ *is full column-rank over* $\mathcal{A}$*.*
3. $\mathcal{M}$ *is a torsion module.*
4. *The ideal* $\operatorname{ann}(\mathcal{M}) := \{f(\xi) \in \mathcal{A} \mid f(\xi)m = 0\ \forall\ m \in \mathcal{M}\}$ *is nonzero.*

Remark 2.6. Stated alternatively, $\mathfrak{B}$ is *not* autonomous if and only if $\mathcal{M}$ is *not* a torsion module, or equivalently, the ideal $\operatorname{ann}(\mathcal{M}) = \{0\}$. In this case, $\mathcal{M}$ is said to be a *faithful* $\mathcal{A}$-module [16].

A behavior $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$ is said to be *strongly autonomous* if the ideal $\operatorname{ann}(\mathcal{M})$, in addition to being non-zero, is also such that $\mathcal{A}/\operatorname{ann}(\mathcal{M})$ is a finite dimensional vector space over $\mathbb{R}$. This is also equivalent to the quotient module $\mathcal{M}$ being a finite dimensional vector space over $\mathbb{R}$.

3. Construction of states for autonomous systems

One of the most significant achievements of behavioral theory of 1D systems is a formalization of the construction of states. A key idea behind this formalization is an abstraction of state-variables through the concepts of *Markovianity* and *the axiom of state*, see [5, 8, 17]. Consequently, it was shown that having a state-variable representation is equivalent to having a first order kernel representation, which is first order on the state-variables and zeroth order on the other (manifest) variables [17]. Further, if a behavior is given by a kernel representation that is not first order, it was shown in [17] how this system of equations can be converted into a system of first order equations by constructing state-variables from the manifest variables by a *state map*.

In [18, 19], extension of these ideas to nD systems was investigated. It turns out, rather surprisingly, that a stronger notion of Markovianity is required in order to reconcile it with first order representation. However, this equivalence has been shown, so far, to hold for only special cases of nD systems. Moreover, it is not clear how to construct the state-variables and how to obtain a corresponding first order representation from a given higher order system of equations. In this paper, we propose a method to carry out the task of constructing state-variables for general nD autonomous systems. We show how these states that we construct replicate to a large extent the standard notion of state-variables. Unlike the 1D case, here we do not attempt to formulate an axiom of state in order to justify that the created latent variables are indeed state-variables. However, the justification does come, as a matter of fact, from an alternative formulation of state-variables by Willems [5]. In [5], a state-space system is defined by a quadruple $(\mathbb{T}, \mathbb{W}, \mathcal{X}, \partial)$, where $\mathbb{T}$ and $\mathbb{W}$ are as before – indexing set and signal space respectively – $\mathcal{X}$ is defined as an abstract state-space, and ∂ is a *first order evolution* law. In this paper, we show that given any nD autonomous system, we can create a state-space $\mathcal{X}$ and a first order evolution law ∂ such that the state-space system $(\mathbb{T}, \mathbb{W}, \mathcal{X}, \partial)$ is equivalent to the original system. Note that as the original system is an nD system, the evolution law (though first order), may well turn out to evolve in several directions (or dimensions). We shall see that the state-space is usually an infinite dimensional vector space that is itself a dD behavior, while the evolution is $(n-d)$ dimensional.

Our construction methodology of state-variables is inspired by a representation formula given in [20]. This methodology truly generalizes the existing state construction method for 1D autonomous systems [17], and the same for strongly autonomous 2D systems [8]; see Remarks 4.1 and 4.2 in Section 4, respectively. The key factor in this generalization is a change of coordinates in the indexing set $\mathbb{T}$. This is a feature that manifests itself only when $n \geqslant 2$. Interestingly, this coordinate change is not required for strongly autonomous nD systems. This could very well be the reason that some earlier attempts to generalize state construction for general nD systems were not completely successful.

3.1. Coordinate change and its effects

By a coordinate change on $\mathbb{T}$ we mean an invertible linear transformation on $\mathbb{T}$. Clearly, such a coordinate change can then be represented by a square invertible matrix T, which acts on $\boldsymbol{t} \in \mathbb{T}$ to produce $\widetilde{\boldsymbol{t}} = T\boldsymbol{t}$. For the continuous case $T \in \mathbb{R}^{n \times n}$ having $\det T \neq 0$, and for the discrete case $T \in \mathbb{Z}^{n \times n}$ with $\det T = \pm 1$. Such a coordinate change induces two important maps – the *pull-back,* T^*, and the *push-forward,* T_* – which will be of crucial importance to us in the sequel.

The pull-back is a map $T^* : \mathcal{F}_n^{\mathtt{w}} \to \mathcal{F}_n^{\mathtt{w}}$ defined as follows: for $w \in \mathcal{F}_n^{\mathtt{w}}$ define

$$T^*(w) = w \circ T. \tag{10}$$

That is, for $\boldsymbol{t} \in \mathbb{T}$,

$$(T^*(w))(\boldsymbol{t}) = w(T\boldsymbol{t}).$$

On the other hand, the push-forward is a map $T_* : \mathcal{A} \to \mathcal{A}$ whose definition depends on whether the system is continuous or discrete. For the continuous case, given the coordinate change, $T \in \mathbb{R}^{n \times n}$, define $\widetilde{\xi} := T^T \xi$. Then T_* is defined as: for $f(\xi) \in \mathbb{R}[\xi_1, \ldots, \xi_n]$

$$(T_*(f))(\xi) = f\left(\widetilde{\xi}\right). \tag{11}$$

For the discrete case, given the coordinate change, $T \in \mathbb{Z}^{n\times n}$, define $\widetilde{\xi}_i = \xi^{T_i}$, for all $i = 1, \ldots, n$, where T_i is the i^{th} column of T. Then T_* is defined as: for $f(\xi) \in \mathbb{R}[\xi_1, \xi_1^{-1}, \ldots, \xi_n, \xi_n^{-1}]$

$$(T_*(f))(\xi) = f\left(\widetilde{\xi}\right). \tag{12}$$

Note that, owing to the invertibility of T, both T^* and T_* turn out to be invertible. While T^* is a linear map of infinite dimensional vector spaces, T_* is an $\mathbb{R}$-algebra homomorphism that keeps the base field $\mathbb{R}$ fixed. The push-forward T_* is easily extended to $\mathcal{A}^{1\times\mathtt{w}}$ by applying it element-wise. Under this map of $\mathcal{A}^{1\times\mathtt{w}}$ to itself, a submodule $\mathcal{R} \subseteq \mathcal{A}^{1\times\mathtt{w}}$ gets mapped to another submodule because T_* is invertible and is an $\mathbb{R}$-algebra homomorphism. Since there is no risk of ambiguity, we use the same T_* to denote its extension to $\mathcal{A}^{1\times\mathtt{w}}$.

With the maps T^* and T_* we are now in a position to state the following crucial observation that relates them through a behavior $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$.

Theorem 3.1. *Suppose $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$ with equation module $\mathcal{R} \subseteq \mathcal{A}^{1\times\mathtt{w}}$. Let $T : \mathbb{T} \to \mathbb{T}$ be a coordinate change. Define $T^* : \mathcal{F}_n^{\mathtt{w}} \to \mathcal{F}_n^{\mathtt{w}}$ and $T_* : \mathcal{A}^{1\times\mathtt{w}} \to \mathcal{A}^{1\times\mathtt{w}}$ as above. Then*

$$\mathfrak{B}(\mathcal{R}) = T^*\left(\mathfrak{B}\left(T_*\left(\mathcal{R}\right)\right)\right). \tag{13}$$

Proof : See the appendix. □

3.2. Noether normalization

The main reason behind carrying out the coordinate change is a normalization process that allows us to readily construct the state-space equations. This normalization is called *Noether normalization*. While the conventional form of Noether normalization works for continuous systems, for discrete systems, however, a crucial modification is required. This modified version was worked out in [21]; the special case of it with $n = 2$ can be found in [20]. See [16, 22] for the conventional form of Noether normalization. We state the normalization process and its end result in Lemma 3.2 below. For the remaining part of this paper, we use the symbol $\mathcal{A}_d$ to denote the dD subring of $\mathcal{A}$ generated by the variables $\{\xi_1, \ldots, \xi_d\}$, i.e., for the continuous case $\mathcal{A}_d := \mathbb{R}[\xi_1, \ldots, \xi_d]$ and for the discrete case $\mathcal{A}_d := \mathbb{R}[\xi_1, \xi_1^{-1}, \ldots, \xi_d, \xi_d^{-1}]$.

Lemma 3.2. *Let $\mathcal{R} \subseteq \mathcal{A}^{1\times\mathtt{w}}$ be a submodule. Then there exists a coordinate change $T : \mathbb{T} \to \mathbb{T}$ and a positive integer $d < n$ such that the quotient module $\mathcal{M}_d := \mathcal{A}^{1\times\mathtt{w}}/T_*(\mathcal{R})$ is a finitely generated faithful*[1] *module over $\mathcal{A}_d$.*

3.3. Behaviors with a special structure

Suppose $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$ is such that its quotient module $\mathcal{M}$ is a finitely generated module over $\mathcal{A}_d$ for some non-negative integer $d < n$. Note that for the special case of $d = 0$, this assumption means $\mathcal{M}$ is a finite dimensional vector space over $\mathbb{R}$, which, in turn, implies that $\mathfrak{B}$ is strongly autonomous. Further, for discrete 2D systems, this assumption with $d = 1$ implies that there exists a finite union of horizontal lines in $\mathbb{Z}^2$ such that restriction of any trajectory $w \in \mathfrak{B}$ to that union of horizontal lines uniquely determines w (see [23]). Providing a complete system theoretic interpretation of this assumption is a subject of future research. However, see [20, Proposition 3.4] for many equivalent conditions to this assumption in terms of the equation module $\mathcal{R}$ for the special case of discrete 2D systems.

[1]Faithful means ann $\mathcal{M}_d = \{0\}$. See Remark 2.6

The module $\mathcal{M}$ being finitely generated over $\mathcal{A}_d$ means that one can fix a generating set $\mathcal{G} := \{g_1, \ldots, g_r\} \subseteq \mathcal{M}$ such that every element in $\mathcal{M}$ can be written as a linear combination of the elements of $\mathcal{G}$ with coefficients from $\mathcal{A}_d$. For the purpose of notational convenience, let us partition ξ as $\xi = (\zeta, \eta)$, such that

$$\begin{array}{rclcl} \zeta & = & (\zeta_1, \ldots, \zeta_d) & := & (\xi_1, \ldots, \xi_d) \text{ and} \\ \eta & = & (\eta_1, \ldots, \eta_{n-d}) & := & (\xi_{d+1}, \ldots, \xi_n). \end{array}$$

We now define the following $\mathcal{A}_d$-module homomorphisms $\mu_i : \mathcal{M} \to \mathcal{M}$ for $1 \leqslant i \leqslant (n-d)$ as

$$\mu_i(f) = \eta_i f \tag{14}$$

for all $f \in \mathcal{M}$. That is, μ_i is the map given by multiplication by η_i. It can be easily checked that μ_i's are indeed $\mathcal{A}_d$-module homomorphisms. Since $\mathcal{M}$ has been assumed to be finitely generated as an $\mathcal{A}_d$-module, for each $i = 1, \ldots, (n-d)$, the map μ_i is represented by a matrix $A_i(\zeta) \in \mathcal{A}_d^{r \times r}$. Indeed, suppose $\mu_i(g_j) = \sum_{k=1}^{r} a_i^{j,k}(\zeta) g_k$, where $a_i^{j,k}(\zeta) \in \mathcal{A}_d$ then

$$A_i(\zeta) := \left[a_i^{j,k}(\zeta)\right]_{1 \leqslant j \leqslant r, 1 \leqslant k \leqslant r}. \tag{15}$$

Next, suppose that $\overline{e_i} \in \mathcal{M}$ is the image of $e_i \in \mathcal{A}^{1 \times \mathtt{w}}$ – the i^{th} basis vector – under the canonical surjection $\mathcal{A}^{1 \times \mathtt{w}} \twoheadrightarrow \mathcal{M}$. Once again, since $\mathcal{M}$ is generated by the elements of $\mathcal{G}$ as an $\mathcal{A}_d$-module, there exist $c^{i,j}(\zeta) \in \mathcal{A}_d$ such that $\overline{e_i} = \sum_{j=1}^{r} c^{i,j} g_j$. Then define

$$C(\zeta) := \left[c^{i,j}(\zeta)\right]_{1 \leqslant i \leqslant \mathtt{w}, 1 \leqslant j \leqslant r}. \tag{16}$$

Finally, suppose $\mathcal{G}^{\perp}$ denotes the following set:

$$\mathcal{G}^{\perp} := \left\{ f(\zeta) \in \mathcal{A}_d^{1 \times r} \mid f(\zeta) \left[g_1, \ldots, g_r\right]^T = 0 \right\}.$$

It is easy to check that $\mathcal{G}^{\perp}$ has the structure of a sub-module of $\mathcal{A}_d^{1 \times r}$ over $\mathcal{A}_d$; it is called the *module of relations* of the generators $\{g_1, \ldots, g_r\}$ (see [9]). Since $\mathcal{A}_d^{1 \times r}$ is a Noetherian module, $\mathcal{G}^{\perp}$ is finitely generated as an $\mathcal{A}_d$-module (see [16] for a definition of Noetherian modules and the implication that $\mathcal{G}^{\perp}$ is finitely generated). Let $X(\zeta) \in \mathcal{A}_d^{\bullet \times r}$ be a matrix whose rows generate $\mathcal{G}^{\perp}$ as an $\mathcal{A}_d$-module. That is,

$$\text{rowspan } X(\zeta) = \mathcal{G}^{\perp}. \tag{17}$$

With these matrices $A_1(\zeta), \ldots, A_{n-d}(\zeta) \in \mathcal{A}_d^{r \times r}$, $C(\zeta) \in \mathcal{A}_d^{\mathtt{w} \times r}$ and $X(\zeta) \in \mathcal{A}_d^{\bullet \times r}$ we are now in a position to state the following crucial result.

Theorem 3.3. *Let $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$ be autonomous with its quotient module $\mathcal{M}$ being finitely generated as an $\mathcal{A}_d$-module for some positive integer $d < n$. Suppose $\mathcal{G} \subseteq \mathcal{M}$ is a finite generating set of $\mathcal{M}$ as an $\mathcal{A}_d$-module. Let $A_1(\zeta), \ldots, A_{n-d}(\zeta) \in \mathcal{A}_d^{r \times r}$, $C(\zeta) \in \mathcal{A}_d^{\mathtt{w} \times r}$ and $X(\zeta) \in \mathcal{A}_d^{\bullet \times r}$ be as defined in equations* (15), (16) *and* (17), *respectively. Then $\mathfrak{B}$ admits the following latent variable representation:*

$$\mathfrak{B} = \left\{ w \in \mathcal{F}_n^{\mathtt{w}} \;\middle|\; \exists x \in \mathcal{F}_n^r \text{ satisfying} \begin{bmatrix} X(\zeta) & 0 \\ \eta_1 I - A_1(\zeta) & 0 \\ \vdots & \vdots \\ \eta_{n-d} I - A_{n-d}(\zeta) & 0 \\ C(\zeta) & -I \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix} = 0 \right\}. \tag{18}$$

Proof : See the appendix. □

3.4. Constructing the state-space and the evolution equations

A remarkable feature of Theorem 3.3 lies in a special property that $X(\zeta)$ has in connection with each of the $A_1(\zeta), \ldots, A_{n-d}(\zeta)$. It follows from the construction that for each $i = 1, \ldots, (n-d)$ there exists $F_i(\zeta) \in \mathcal{A}_d^{\bullet\times\bullet}$ such that

$$X(\zeta)A_i(\zeta) = F_i(\zeta)X(\zeta) \tag{19}$$

(see [24, Lemma 4]). This observation prompts us to define the following object:

$$\mathcal{X} := \ker X(\zeta) \subseteq \mathcal{F}_d^r. \tag{20}$$

Note that $\mathcal{X} \in \mathfrak{L}_d^r$, that is, $\mathcal{X}$ is a dD behavior with r number of manifest variables. It is this dD behavior that plays the role of the state-space in our construction. Equation (19) implies that, for all $i = 1, \ldots, (n-d)$, $\mathcal{X}$ is $A_i(\zeta)$-invariant, that is, $A_i(\zeta)(\mathcal{X}) \subseteq \mathcal{X}$. Therefore, $A_i(\zeta)$'s can be viewed as maps from $\mathcal{X}$ to itself. It then makes sense to define the following state-space system:

$$\Sigma_{\text{state}} := \left(\widetilde{\mathbb{T}}, \mathcal{X}, \mathfrak{B}_{\text{state}}\right), \tag{21}$$

where the indexing set $\widetilde{\mathbb{T}} = \mathbb{R}^{n-d}$ (for continuous) or $\widetilde{\mathbb{T}} = \mathbb{Z}^{n-d}$ (for discrete), the signal space $\mathcal{X}$ is as defined by equation (20), and the behavior is defined by

$$\mathfrak{B}_{\text{state}} := \left\{\widetilde{x} : \widetilde{\mathbb{T}} \to \mathcal{X} \mid \eta_i\widetilde{x} = A_i(\zeta)\widetilde{x} \text{ for all } i = 1, \ldots, (n-d)\right\} \tag{22}$$

with η_i being the i^{th} differential (for continuous[2]) or the i^{th} shift (for discrete) operator defined on the indexing set $\widetilde{\mathbb{T}}$.

Now note that if $\widetilde{x} : \widetilde{\mathbb{T}} \to \mathcal{X}$ then $\widetilde{x}$ can be identified uniquely with an $x \in \mathcal{F}_n^r$ in the following manner: suppose $\boldsymbol{t} = (\boldsymbol{t}_1, \boldsymbol{t}_2) \in \mathbb{T}$, where $\boldsymbol{t}_1$ is a d-tuple and $\boldsymbol{t}_2$ is an $(n-d)$-tuple, then

$$x(\boldsymbol{t}) = x(\boldsymbol{t}_1, \boldsymbol{t}_2) := (\widetilde{x}(\boldsymbol{t}_2))(\boldsymbol{t}_1). \tag{23}$$

This is true because $\mathcal{X}$ is a subset of $\mathcal{F}_d^r$. With this identification, and the fact that $\mathcal{X}$ is $A_i(\zeta)$-invariant, it follows that the representation of $\mathfrak{B}$ by equation (18) of Theorem 3.3 and $\mathfrak{B}_{\text{state}}$ as given by equation (22) are equivalent. This is the content of Lemma 3.4 below.

Lemma 3.4. *Let $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$ be autonomous with its quotient module $\mathcal{M}$ being finitely generated as an $\mathcal{A}_d$-module for some positive integer $d < n$. Suppose $\mathcal{G} \subseteq \mathcal{M}$ is a finite generating set of $\mathcal{M}$ as an $\mathcal{A}_d$-module. Let $A_1(\zeta), \ldots, A_{n-d}(\zeta) \in \mathcal{A}_d^{r\times r}$, $C(\zeta) \in \mathcal{A}_d^{\mathtt{w}\times r}$ and $X(\zeta) \in \mathcal{A}_d^{\bullet\times r}$ be as defined in equations* (15), (16) *and* (17)*, respectively. Define $\mathcal{X}$ and Σ_{state} as per equations* (20) *and* (21)*, where $\mathfrak{B}_{\text{state}}$ is defined as in equation* (22)*. Then $w \in \mathfrak{B}$ if and only if there exists $\widetilde{x} \in \mathfrak{B}_{\text{state}}$ such that $w = C(\zeta)x$, where x corresponds to $\widetilde{x}$ as per equation* (23)*.*

Theorem 3.3 and Lemma 3.4 hold whenever the behavior $\mathfrak{B}$ has the special property that the quotient module $\mathcal{M}$ is a finitely generated $\mathcal{A}_d$-module. However, there are many systems which fail to satisfy this rather restrictive condition. This is where Noether normalization (Lemma 3.2) and Theorem 3.1 become crucial. Suppose an autonomous $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$, with equation module

[2]It is important to note here that for the differential operator to define an evolution on $\mathcal{X}$, the operator $A_i(\zeta)$ must define a map from $\mathcal{X}$ to its tangent bundle. However, $\mathcal{X}$ being a *linear* subspace of $\mathcal{F}_d^r$, the tangent space of $\mathcal{X}$ at a point $x \in \mathcal{X}$ is $\mathcal{X}$ itself. Therefore, $A_i(\zeta)$ does indeed map $x \in \mathcal{X}$ to a point in the tangent space of $\mathcal{X}$ at x.

$\mathcal{R}$, is such that the corresponding $\mathcal{M}$ is not finitely generated over any $\mathcal{A}_d$. Then find out a coordinate change $T : \mathbb{T} \to \mathbb{T}$ such that under the corresponding push-forward T_* we have $\mathcal{A}^{1\times \mathtt{w}}/T_*(\mathcal{R})$ to be finitely generated over an $\mathcal{A}_d$ for some positive integer $d < n$. Therefore, Lemma 3.4 now applies to the behavior $\mathfrak{B}(T_*(\mathcal{R}))$. Lemma 3.2 guarantees that such a d exists. By Theorem 3.1, this transformed behavior $\mathfrak{B}(T_*(\mathcal{R}))$ is in a one-to-one correspondence with the original $\mathfrak{B}$ by the pull-back of the coordinate change. Thus, the state representation of $\mathfrak{B}(T_*(\mathcal{R}))$ provided by Lemma 3.4 is effectively a state representation for the original $\mathfrak{B}$, too.

Theorem 3.5. *Suppose $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$ is autonomous. Then there exist a positive integer $d < n$ and matrices $A_1(\zeta), \ldots, A_{n-d}(\zeta) \in \mathcal{A}_d^{r\times r}$, $C(\zeta) \in \mathcal{A}_d^{\mathtt{w}\times r}$ and $X(\zeta) \in \mathcal{A}_d^{\bullet\times r}$ such that Σ_{state} can be defined as in equation (21) with $\mathcal{X}$ and $\mathfrak{B}_{\text{state}}$ as defined in equations (20) and (22). Further, define*

$$\mathfrak{B}_{\text{s/o}} := \{w \in \mathcal{F}_n^{\mathtt{w}} \mid w = C(\zeta)x,\ \widetilde{x} \in \mathfrak{B}_{\text{state}}\},$$

where x and $\widetilde{x}$ are related as shown in equation (23). Then there exists a coordinate change $T : \mathbb{T} \to \mathbb{T}$ such that

$$\mathfrak{B} = T^*(\mathfrak{B}_{\text{s/o}}).$$

Proof : See the appendix. $\square$

4. Remarks on the state-space and illustrative examples

In this section we collect a few salient features of the state-space $\mathcal{X}$ and the first order system $\mathfrak{B}_{\text{state}}$ defined on $\mathcal{X}$.

Remark 4.1. (special case: $n = 1$) Theorem 3.5 applied to an autonomous $\mathfrak{B} \in \mathfrak{L}_1^{\mathtt{w}}$ results in the conventional state-space representation of $\mathfrak{B}$. Indeed, for a 1D autonomous system, the quotient module $\mathcal{M}$ turns out to be a finite dimensional vector space over $\mathbb{R}$, which is a special case of being a finitely generated module over an $\mathbb{R}$-algebra. So, here $d = 0$, $\mathcal{A}_d = \mathbb{R}$, and thus, the matrices $A(\zeta), C(\zeta)$ are constant matrices. Further, since $\mathcal{M}$ is an $\mathbb{R}$-vector space, one can always choose a generating set $\mathcal{G}$ that is linearly independent over $\mathbb{R}$. Therefore, $X(\zeta) = 0$. As a result, the state-space is a finite dimensional real vector space.

Remark 4.2. (special case: strongly autonomous) When $\mathfrak{B} \in \mathfrak{L}_n^{\mathtt{w}}$ is strongly autonomous, its quotient module $\mathcal{M}$ turns out to be a finite dimensional vector space over $\mathbb{R}$. Thus, in this case too, $d = 0$ and $\mathcal{A}_d = \mathbb{R}$. Therefore in this case, there is an n-tuple of square matrices $A_1, \ldots, A_n$ that represent multiplication by $\xi_1, \ldots, \xi_n$, respectively. Thus the evolution laws are first order in n directions, specified by the A_i matrices. Once again, $X(\zeta) = 0$ as $\mathcal{M}$ is an $\mathbb{R}$-vector space and therefore $\mathcal{G}$ can be chosen to be linearly independent over $\mathbb{R}$. Note that this case does not involve any coordinate change in the independent variables. So, T is identity.

Remark 4.3. (the smallest value of d) The non-negative integer d plays a pivotal role in the state construction methodology presented in this paper. A natural question that arises now is: what is the smallest value of this d? Further, we may also ask whether this smallest value of d is an intrinsic property of $\mathfrak{B}$, or it is dependent on representations of $\mathfrak{B}$. Quite interestingly, it turns out that the smallest value of d is indeed an intrinsic property of $\mathfrak{B}$, for it is equal to a well-known algebraic invariant of the quotient module $\mathcal{M}$ called *Krull dimension* of $\mathcal{M}$. (See [16] for the definition of Krull dimension.) We give below a brief sketch of this result, and illustrate it further in Example 4.4. It has been shown in [24, Proposition 36], for the discrete case, that if d_{Krull} is the Krull dimension of $\mathcal{M}$ then there exists a coordinate change $T : \mathbb{Z}^n \to \mathbb{Z}^n$ such that $\mathcal{A}^{1\times \mathtt{w}}/T_*(\mathcal{R})$ is a finitely generated module over $\mathcal{A}_d$ for some non-negative integer $d < n$ if and only if $d \geqslant d_{\text{Krull}}$. The same can be proved for the continuous case following exactly the same chain of arguments as in [24, Proposition 36]. Now, recall from Theorems 3.3 and 3.5 that for the state construction to work it is necessary that $\mathcal{A}^{1\times \mathtt{w}}/T_*(\mathcal{R})$ be finitely generated over $\mathcal{A}_d$. Therefore, from the last observation it can be argued that the smallest possible d for which the state construction can be carried out is $d = d_{\text{Krull}}$.

Among the various possible d's for which there exists a coordinate change $T : \mathbb{T} \to \mathbb{T}$ such that $A^{1\times \mathtt{w}}/T_*(\mathcal{R})$ is a finitely generated module over $\mathcal{A}_d$, the smallest d is characterized by the fact that $A^{1\times \mathtt{w}}/T_*(\mathcal{R})$ is a *faithful* finitely generated module over $\mathcal{A}_d$ (see Remark 2.6 for the definition of a faithful module). This has been proved in [24, Lemma 37] for the discrete case; the proof of the continuous case follows along the same line. Note that Noether's normalization lemma (and its discrete version) (Lemma 3.2) provides a coordinate change T that makes $A^{1\times \mathtt{w}}/T_*(\mathcal{R})$ a faithful finitely generated module over $\mathcal{A}_d$. Therefore, the d that Lemma 3.2 provides happens to be equal to d_{Krull}, the Krull dimension of $\mathcal{M}$, and, therefore, the smallest possible. Incidentally, $A^{1\times \mathtt{w}}/T_*(\mathcal{R})$ being faithful over $\mathcal{A}_d$ also means that the state-space $\mathcal{X} = \ker\, X(\zeta) \in \mathfrak{L}^r_d$ is non-autonomous (see Remark 2.6).

Example 4.4. Let us consider a continuous 3D scalar system given by the following kernel representation:

$$\mathfrak{B} = \ker \begin{bmatrix} \partial_3^2 \\ \partial_2^2 \\ \partial_3\partial_1 - \partial_2 \end{bmatrix}.$$

In this case, clearly, $\mathcal{A} = \mathbb{R}[\partial_1, \partial_2, \partial_3]$. The equation module $\mathcal{R}$ here is the ideal $\mathfrak{a} = \langle \partial_3^2, \partial_2^2, \partial_3\partial_1 - \partial_2\rangle$. Note that the quotient module $\mathcal{M} = \mathcal{A}/\mathfrak{a}$ is a finitely generated module over $\mathcal{A}_2 := \mathbb{R}[\partial_1, \partial_2]$. The set $\{\overline{1}, \overline{\partial_3}\}$ is a generating set for $\mathcal{M}$ as a module over $\mathcal{A}_2$. One can therefore compute the various matrices $A_1(\partial_1, \partial_2), C(\partial_1, \partial_2)$ and $X(\partial_1, \partial_2)$. The matrices $A_1(\partial_1, \partial_2) \in \mathcal{A}_2^{2\times 2}, C(\partial_1, \partial_2) \in \mathcal{A}_2^{1\times 2}$ and $X(\partial_1, \partial_2) \in \mathcal{A}_2^{\bullet\times 2}$ are given by

$$A_1(\partial_1, \partial_2) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},\ C(\partial_1, \partial_2) = \begin{bmatrix} 1 & 0 \end{bmatrix},\ X(\partial_1, \partial_2) = \begin{bmatrix} \partial_2^2 & 0 \\ -\partial_2 & \partial_1 \\ 0 & \partial_2 \end{bmatrix}.$$

In this case, the state-space, $\mathcal{X} := \ker\, X(\partial_1, \partial_2) \subseteq \mathfrak{C}^\infty(\mathbb{R}^2, \mathbb{R}^2)$, is a 2D behavior. It can be easily checked that $\mathcal{X}$ is invariant under the action of $A_1(\partial_1, \partial_2)$. This is because

$$X(\partial_1, \partial_2)A_1(\partial_1, \partial_2) = \begin{bmatrix} 0 & 0 & \partial_2 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} X(\partial_1, \partial_2)$$

The evolution law is given by $\frac{\partial x}{\partial t_3} = A_1(\partial_1, \partial_2)x$ which is an evolution in only one direction, namely, along t_3. Thus the evolution law is 1-dimensional. Note however that $\mathcal{X}$ here turns out to be an *autonomous* 2D behavior because $X(\partial_1, \partial_2)$ is full column-rank. Therefore, $\mathcal{X}$ can be further reduced to a 'smaller' state-space which must be a 1D behavior.

It can be checked that the quotient ring $\mathcal{M} = \mathcal{A}/\mathfrak{a}$ is also finitely generated as a module over $\mathcal{A}_1 = \mathbb{R}[\partial_1]$. We now compute the matrices $A_1(\partial_1), A_2(\partial_1), C(\partial_1), X(\partial_1)$. First, note that

$\{\overline{1}, \overline{\partial_2}, \overline{\partial_3}\}$ is a finite generating set for $\mathcal{M}$ as a $\mathcal{A}_1$-module. The matrices $A_1(\partial_1)$ and $A_2(\partial_1)$ in this case, are given by

$$A_1(\partial_1) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A_2(\partial_1) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

A matrix of relations $X(\partial_1)$, for the chosen generating set is given by

$$X(\partial_1) = \begin{bmatrix} 0 & -1 & \partial_1 \end{bmatrix}.$$

The state-space $\mathcal{X}$ is thus given by

$$\mathcal{X} = \ker\, X(\partial_1) \subseteq \mathfrak{C}^\infty(\mathbb{R}, \mathbb{R}^3).$$

Note that $\mathcal{X}$ is indeed invariant under $A_1(\partial_1)$ and $A_2(\partial_1)$ because

$$X(\partial_1)A_1(\partial_1) = 0, \quad X(\partial_1)A_2(\partial_1) = 0.$$

The output matrix $C(\partial_1)$ turns out to be

$$C(\partial_1) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}.$$

The first order evolution on $\mathcal{X}$ then gets defined as

$$\frac{\partial x}{\partial t_2} = A_1(\partial_1)x, \qquad \frac{\partial x}{\partial t_3} = A_2(\partial_1)x, \tag{24}$$

where $x : \mathbb{R}^2 \ni (t_2, t_3) \mapsto x(t_2, t_3) \in \mathcal{X}$. It can be verified that $w \in \mathfrak{B}$ if and only if there exists $x \in \mathcal{X}^{\mathbb{R}^2}$ that satisfies equation (24) and $w(t_1, t_2, t_3) = (C(\partial_1)x(t_2, t_3))\,(t_1)$. The state-space here is a 1D behavior and the first order evolution is in two directions – namely t_2 and t_3.

Note that any further reduction in d is not possible because the original behavior $\mathfrak{B}$ is not strongly autonomous. Thus the minimum value of d in this case is $d = 1$. Note that the Krull dimension of the quotient ring $\mathcal{M}$ equals 1.

Remark 4.5. The first order evolutionary system $\mathfrak{B}_{\text{state}}$ over the state-space $\mathcal{X}$ is much akin to the model of infinite dimensional systems followed in [25]. In fact, the two are exactly the same when $d = n - 1$. However, as pointed out in Remark 4.3, it is beneficial to have d that is the smallest possible. The smallest d equals the Krull dimension of $\mathcal{M}$, and thus, is an invariant of $\mathfrak{B}$. So, the smallest d need not always be $n-1$. Moreover, another shortcoming of the model in [25] is that for many nD systems, the first principle model does not have the form of a first order evolution over a (possibly infinite dimensional) state-space. In our approach, however, for *every* autonomous nD system $\mathfrak{B}$, a state-space $\mathcal{X}$ and a first order system $\mathfrak{B}_{\text{state}}$ on $\mathcal{X}$ can be constructed that is equivalent to $\mathfrak{B}$. The key step that enables this construction for every autonomous nD system is Noether normalization (Lemma 3.2). Hence, Theorem 3.5 can also be viewed as a generalization of the first order representation given in [25] for the special case of autonomous systems. We illustrate this in Example 4.6 below.

Example 4.6. Consider a discrete 3D scalar behavior

$$\mathfrak{B} = \ker\,((\sigma_1 - 1)(\sigma_2 - 1)(\sigma_3 - 1)).$$

The equation module $\mathcal{R}$ is the ideal $\mathfrak{a} = \langle \sigma_1\sigma_2\sigma_3 - \sigma_1\sigma_2 - \sigma_2\sigma_3 - \sigma_3\sigma_1 + \sigma_1 + \sigma_2 + \sigma_3 - 1\rangle$. Clearly, the quotient module $\mathcal{M} = \mathcal{A}/\mathfrak{a}$ is not a finitely generated module over either $\mathcal{A}_1$ or $\mathcal{A}_2$. So we invoke Lemma 3.2; in this case, we take the coordinate change $T = \left[\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{smallmatrix}\right]$. Under the corresponding push-forward T_* the equation ideal $\mathfrak{a}$ gets mapped to

$$\begin{aligned} T_*(\mathfrak{a}) &= \langle \sigma_3^3\sigma_1\sigma_2 - \sigma_3^2(\sigma_1\sigma_2 + \sigma_1 + \sigma_2) + \sigma_3(\sigma_1 + \sigma_2 + 1) - 1\rangle \\ &= \langle \sigma_3^3 - \sigma_3^2(1 + \sigma_2^{-1} + \sigma_1^{-1}) + \sigma_3(\sigma_2^{-1} + \sigma_1^{-1} + \sigma_1^{-1}\sigma_2^{-1}) - \sigma_1^{-1}\sigma_2^{-1}\rangle. \end{aligned}$$

Now, $\mathcal{A}/T_*(\mathfrak{a})$ is finitely generated as an $\mathcal{A}_2$-module. A generating set can be chosen as $\{\overline{1}, \overline{\sigma_3}, \overline{\sigma_3^2}\}$. Consequently, we get

$$A_1(\sigma_1, \sigma_2) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \sigma_1^{-1}\sigma_2^{-1} & -(\sigma_2^{-1}+\sigma_1^{-1}+\sigma_1^{-1}\sigma_2^{-1}) & 1+\sigma_2^{-1}+\sigma_1^{-1} \end{bmatrix}, \quad C(\sigma_1, \sigma_2) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix},$$

and $X(\sigma_1, \sigma_2) = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$. So the first order representation of $\mathfrak{B}$ here is given by

$$x(t_3 + 1) = A_1(\sigma_1, \sigma_2)x(t_3), \qquad w = C(\sigma_1, \sigma_2)x,$$

where $x : \mathbb{Z} \ni t_3 \mapsto x(t_3) \in \mathcal{W}(\mathbb{Z}^2, \mathbb{R}^3)$. Note that the system has a state-space which is a 2D behavior ($\mathcal{W}(\mathbb{Z}^2, \mathbb{R}^3)$) and the evolution law displays evolution in only one direction (t_3).

5. Concluding remarks

In this paper, we have provided a framework for constructing equivalent state-space representations for autonomous nD behaviors. We have shown that the state-space may itself turn out to be a dD behavior (and hence infinite dimensional). We have also shown that this equivalent state-space representation has a first order evolution law attached to it. This evolution law gives equations of evolution of the behavior in $(n - d)$ directions. The state-space thus constructed is in complete agreement with one of Willems' notions about state-space systems and agrees completely with the known (and established) cases of state-space representations like 1D systems and strongly autonomous nD systems.

Appendix A. Proofs

Proof of Theorem 3.1: The continuous case has been proved in [26, Lemma 17] for scalar behaviors, which easily extends to the vector ones. The discrete case has been proved in [21, Theorem 6.5] for scalar behaviors, and in [24, Lemma 23] for vector behaviors. □

Proof of Theorem 3.3: For the proof, we need the following observation first. Given $\mathcal{M}$, a finitely generated $\mathcal{A}_d$-module, with generating set $\mathcal{G} = \{g_1, \ldots, g_r\}$, a map $\psi : \mathcal{A}_d^{1\times r} \to \mathcal{M}$ can be set up as

$$\begin{array}{rccc} \psi : & \mathcal{A}_d^{1\times r} & \twoheadrightarrow & \mathcal{M} \\ & e_i & \mapsto & g_i \end{array} \tag{A.1}$$

for $1 \leqslant i \leqslant r$, where e_i is the i^{th} standard basis row-vector in $\mathcal{A}_d^{1\times r}$. Further, with $\mathcal{A}_d$-module homomorphisms $\mu_1, \ldots, \mu_{n-d}$ as defined by equation (14), and the corresponding matrix representations $A_1(\zeta), \ldots, A_{n-d}(\zeta)$ as defined in equation (15), we get that the following diagram commutes for all $1 \leqslant j \leqslant (n - d)$.

$$\begin{array}{rcl} \mathcal{A}_d^{1\times r} & \overset{\psi}{\twoheadrightarrow} & \mathcal{M} \\ A_j(\zeta) \downarrow & & \downarrow \mu_j \\ \mathcal{A}_d^{1\times r} & \overset{\psi}{\twoheadrightarrow} & \mathcal{M} \end{array} \tag{A.2}$$

where the action $A_j(\zeta)$ on $\mathcal{A}_d^{1\times r}$ is right-multiplication. Moreover, it is straightforward to check that a repeated action of the maps μ_j is represented by the corresponding product of the matrices $A_j(\zeta)$. With this observation, we now define the following construction: given a polynomial $f(\zeta,\eta) \in \mathcal{A}$, construct first the map $f(\zeta,\mu) : \mathcal{M} \to \mathcal{M}$ by replacing every η_j by the corresponding μ_j for all $1 \leqslant j \leqslant (n-d)$; notice that multiplication of various η_j's must be replaced by compositions of μ_j's. (The order of composition does not matter, because it follows from their definition that the μ_j's commute with each other.) It then follows from the commutative diagram (equation (A.2)) that the action of $f(\zeta,\mu)$ is represented by the matrix $f(\zeta, A(\zeta)) \in \mathcal{A}_d^{r\times r}$, which is obtained by replacing every η_j in $f(\zeta,\eta)$ by the corresponding $A_j(\zeta)$ for all $1 \leqslant j \leqslant (n-d)$. The following lemma is crucial for the proof of Theorem 3.3. Its proof follows immediately from the fact that $f(\zeta,\mu)$ is represented by $f(\zeta, A(\zeta))$. However, a formal proof for the discrete case can be found in [24, Lemma 7].

Lemma Appendix A.1. *Let $f(\zeta,\eta) \in \mathcal{A}^{1\times \mathtt{w}}$ be given by $f(\zeta,\eta) = \begin{bmatrix} f_1(\zeta,\eta) & \cdots & f_{\mathtt{w}}(\zeta,\eta)\end{bmatrix}$. Let $A_1(\zeta,\eta), \ldots, A_{n-d}(\zeta)$ and $C(\zeta,\eta)$ be as defined by equations* (15) *and* (16)*, respectively. Then the following are equivalent:*

1. $f(\zeta,\eta) \in \mathcal{R}$.
2. $\psi\left(\sum_{i=1}^{\mathtt{w}} C_i(\zeta) f_i\left(\zeta, A(\zeta)\right)\right) = 0$.
3. *There exists $F(\zeta) \in \mathcal{A}^{\mathtt{w}\times \bullet}$ such that $\sum_{i=1}^{\mathtt{w}} C_i(\zeta) f_i\left(\zeta, A(\zeta)\right) = F(\zeta) X(\zeta)$.*

Here, $C_i(\zeta) := \begin{bmatrix} c^{i,1}(\zeta) & \cdots & c^{i,r}(\zeta)\end{bmatrix}$.

We now prove Theorem 3.3. First, let us define the auxiliary behavior $\mathfrak{B}_{\text{aux}} := \{w \in \mathcal{F}_n^{\mathtt{w}} \mid \exists x \in \mathcal{F}_n^r$ such that equation (A.3) is satisfied$\}$

$$\begin{bmatrix} X(\zeta) & 0 \\ \eta_1 I - A_1(\zeta) & 0 \\ \vdots & \vdots \\ \eta_{n-d} I - A_{n-d}(\zeta) & 0 \\ C(\zeta) & -I \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix} = 0. \tag{A.3}$$

Theorem 3.3 would be proven if we show that $\mathfrak{B} = \mathfrak{B}_{\text{aux}}$.
($\mathfrak{B} \subseteq \mathfrak{B}_{\text{aux}}$) Suppose $\mathcal{G} := \{g_1, \ldots, g_r\} \subseteq \mathcal{M}$ is a generating set for $\mathcal{M}$ as an $\mathcal{A}_d$-module. Recall from equation (6) the action of elements from $\mathcal{M}$ on trajectories in $\mathfrak{B}$. For $w \in \mathfrak{B}$, define

$$x := \begin{bmatrix} g_1 \\ \vdots \\ g_r \end{bmatrix} w.$$

It then follows from the construction of $X(\zeta), A_i(\zeta), C(\zeta)$ that $X(\zeta)x = 0$, $\eta_i x = A_i(\zeta)x$, $w = C(\zeta)x$. This means x, w satisfy equation (A.3). Hence, $w \in \mathfrak{B}_{\text{aux}}$.
($\mathfrak{B} \supseteq \mathfrak{B}_{\text{aux}}$) Suppose $w \in \mathfrak{B}_{\text{aux}}$. Then, there exists $x \in \mathcal{F}_n^r$ such that x, w satisfy equation (A.3). Now, let $f(\zeta,\eta) = \begin{bmatrix} f_1(\zeta,\eta) & \cdots & f_{\mathtt{w}}(\zeta,\eta)\end{bmatrix} \in \mathcal{R}$ be arbitrary. In order to show $w \in \mathfrak{B}$, it is enough that we show $f(\zeta,\eta)w = 0$. Since $w = C(\zeta)x$ we get that

$$f(\zeta)w = \sum_{i=1}^{\mathtt{w}} f_i(\zeta,\eta) w_i = \sum_{i=1}^{\mathtt{w}} f_i(\zeta,\eta) C_i(\zeta) x,$$

where $C_i(\zeta)$ is the i^{th} row of $C(\zeta)$. Clearly, we can write

$$\sum_{i=1}^{\mathtt{w}} f_i(\zeta,\eta) C_i(\zeta) x = \sum_{i=1}^{\mathtt{w}} C_i(\zeta) f_i(\zeta,\eta) x$$

because $f_i(\zeta,\eta)$ is scalar, and so, commutes with $C_i(\zeta)$. Now, recall that x satisfies $\eta_i x = A_i(\zeta)x$. Therefore, we get

$$\sum_{i=1}^{\mathtt{w}} C_i(\zeta)f_i(\zeta,\eta)x = \sum_{i=1}^{\mathtt{w}} C_i(\zeta)f_i(\zeta,A(\zeta))x = \left(\sum_{i=1}^{\mathtt{w}} C_i(\zeta)f_i(\zeta,A(\zeta))\right)x.$$

But, $f(\zeta,\eta) \in \mathfrak{R}$, which, by Lemma Appendix A.1, implies that $(\sum_{i=1}^{\mathtt{w}} C_i(\zeta)f_i(\zeta,A(\zeta))) = F(\zeta)X(\zeta)$ for some $F(\zeta) \in \mathcal{A}_d^{\mathtt{w}\times\bullet}$. Therefore,

$$\left(\sum_{i=1}^{\mathtt{w}} C_i(\zeta)f_i(\zeta,A(\zeta))\right)x = F(\zeta)X(\zeta)x = 0,$$

because x satisfies $X(\zeta)x = 0$. Thus, $f(\zeta,\eta)w = F(\zeta)X(\zeta)x = 0$, hence, $w \in \mathfrak{B}$. $\square$

Proof of Theorem 3.5: Follows from Theorem 3.1, Lemma 3.2 and Lemma 3.4. $\square$

References

[1] J. C. Willems, Paradigms and puzzles in the theory of dynamical systems, IEEE Tansactions on Automatic Control 36 (3) (1991) 259–294.

[2] J. C. Willems, From time series to linear system: part I. finite dimensional linear time invariant systems, Automatica 22 (5) (1986) 561–580.

[3] J. C. Willems, From time series to linear system: part II. exact modelling, Automatica 22 (6) (1986) 675–694.

[4] J. C. Willems, From time series to linear system: part III. approximate modelling, Automatica 23 (1) (1986) 87–115.

[5] J. C. Willems, Models for dynamics, Dynamics Reported 2 (1989) 171–269.

[6] J. W. Polderman, J. C. Willems, Introduction to Mathematical Systems Theory: A Behavioral Approach, Springer-Verlag, 1998.

[7] U. Oberst, Multidimensional constant linear systems, Acta Appl. Math. 20 (1990) 1–175.

[8] P. Rocha, J. C. Willems, State for 2-D systems, Linear Algebra and its Applications 122/123/124 (1989) 1003–1038.

[9] H. K. Pillai, S. Shankar, A behavioral approach to control of distributed systems, SIAM Journal on Control and Optimization 37 (2) (1998) 388–408.

[10] J. Pommaret, A. Quadrat, Algebraic analysis of linear multidimensional control systems, IMA Journal of Mathematical Control and Information 16 (1999). 275–297.

[11] S. Attasi, Systèmes lineaires homogènes à deux indices, Rapport LABORIA 31 (1973).

[12] R. P. Roesser, A discrete state-space model for linear image processing, IEEE Transactions on Automatic Control 20 (1975) 1–10.

[13] E. Fornasini, G. Marchesini, State-space realization theory of two-dimensional filters, IEEE Transactions on Automatic Control 21 (4) (1976) 484–492.

[14] E. Fornasini, G. Marchesini, Doubly-indexed dynamical systems: state-space models and structural properties, Mathematical systems theory 12 (1978) 59–72.

[15] S. Shankar, The lattice structure of behaviors, SIAM Journal on Control and Optimization 39 (6) (2001) 1817–1832.

[16] D. Eisenbud, Commutative Algebra with a View Toward Algebraic Geometry, Springer-Verlag, 1995.

[17] P. Rapisarda, J. C. Willems, State maps for linear systems, SIAM Journal on Control and Optimization 35 (3) (1997) 1053–1091.

[18] P. Rocha, J. C. Willems, Markov properties for systems described by PDEs and first-order representations, Systems and Control Letters 55 (2006) 538–542.

[19] P. Rocha, J. C. Willems, Markovian properties for 2-D behavioral systems described by PDE's: the scalar case, Multidimensional Systems and Signal Processing 22 (1-3) (2011) 45–53.

[20] D. Pal, H. K. Pillai, Representation formulae for discrete 2-D autonomous systems, SIAM Journal on Control and Optimization 51 (3) (2013) 2406–2441.

[21] D. Pal, Algebro-geometric analysis of multidimensional (n-D) systems, Ph.D. thesis, Indian Institute of Technology Bombay, Mumbai, India (2012).

[22] M. Kreuzer, L. Robbiano, Computational Commutative Algebra 2, Springer-Verlag, 2000.

[23] D. Pal, Every discrete 2-D autonomous system admits a finite union of parallel lines as a characteristic set, Multidimensional Systems and Signal Processing doi:10.1007/s11045-015-0330-y, (2015), 1-25.

[24] D. Pal, H. K. Pillai, Novel representation of discrete n-D autonomous systems,arXiv:1505.05684.

[25] R. Curtain, H. Zwart, An Introduction to Infinite Dimensional Linear Systems Theory, Springer-Verlag, 1995.

[26] D. Pal, H. K. Pillai, Algorithms for the theory of restrictions of scalar n-D systems to proper subspaces of $\mathbb{R}^n$, Multidimensional Systems and Signal Processing 26 (2) (2015) 439–457.