Acta Appl Math (2006) 94: 83–135
DOI 10.1007/s10440-006-9068-8

Canonical State Representations and Hilbert Functions of Multidimensional Systems

Ulrich Oberst

Received: 22 June 2005 / Accepted: 24 July 2006 /
Published online: 27 September 2006

Abstract A basic and substantial theorem of one-dimensional systems theory, due to R. Kalman, says that an arbitrary input/output behavior with proper transfer matrix admits an observable state representation which, in particular, is a realization of the transfer matrix. The state equations have the characteristic property that any local, better temporal, state at time zero and any input give rise to a unique global state or trajectory of the system or, in other terms, that the global state is the unique solution of a suitable Cauchy problem. With an adaption of this state property to the multidimensional situation or rather its algebraic counter-part we prove that any behavior governed by a linear system of partial differential or difference equations with constant coefficients is isomorphic to a canonical state behavior which is constructed by means of Gröbner bases. In contrast to the one-dimensional situation, to J.C. Willems' multidimensional state space models and and to J.F. Pommaret's modified Spencer form the canonical state behavior is not necessarily a first order system. Further first order models are due E. Zerz. As a by-product of the state space construction we derive a new variant of the algorithms for the computation of the Hilbert function of finitely generated polynomial modules or behaviors. J.F. Pommaret, J. Wood and P. Rocha discussed the Hilbert polynomial in the systems theoretic context. The theorems of this paper are constructive and have been implemented in MAPLE in the two-dimensional case and demonstrated in a simple, but instructive example. A two-page example also gives the complete proof of Kalman's one-dimensional theorem mentioned above. We believe that for this standard case the algorithms of the present paper compare well with their various competitors from the literature.

Key words state • Hilbert function • behavior • multidimensional system • partial differential equation • partial difference equation • polynomial module.

Mathematics Subject Classifications (2000) 93B • 93C • 13P.

U. Oberst (✉)
Institut für Mathematik, Universität Innsbruck, Technikerstraße 25, A-6020 Innsbruck, Austria
e-mail: Ulrich.Oberst@uibk.ac.at

1 Introduction

One of the basic and substantial theorems of systems theory, originally due to Kalman [15], says that any input/output-behavior with proper transfer matrix admits an observable state realization. In more detail, let

$$\mathcal{B} := \left\{ \begin{pmatrix} y \\ u \end{pmatrix} \in \mathcal{D}'(\mathbb{R})^{p+m};\ P(\mathrm{d}/\mathrm{d}t)y = Q(\mathrm{d}/\mathrm{d}t)u \right\} \tag{1}$$

be an IO-behavior where t is the independent variable, usually *time*, $P \in \mathbb{C}[s]^{p\times p}$ and $Q \in \mathbb{C}[s]^{p\times m}$ are polynomial matrices with non-zero determinant $\det(P)$ and proper transfer matrix $P^{-1}Q$ and where the input u and output y are vector distributions. The case of real instead of complex coefficients is mathematically a special case, but usually preferred in the engineering literature. Then there is an observable Kalman system

$$\overset{\bullet}{x} = Ax + Bu,\ \ y = Cx + Du,\ \ A \in \mathbb{C}^{n\times n},\ \ B \in \mathbb{C}^{n\times m},\ \ C \in \mathbb{C}^{p\times n},\ \ D \in \mathbb{C}^{p\times m} \tag{2}$$

such that the map

$$\begin{pmatrix} C & D \\ 0 & \mathrm{id} \end{pmatrix}:\ \mathcal{B}^s := \left\{ \begin{pmatrix} x \\ u \end{pmatrix} \in \mathcal{D}'^{n+m};\ \overset{\bullet}{x} = Ax + Bu \right\} \to \mathcal{B},\ \begin{pmatrix} x \\ u \end{pmatrix} \mapsto \begin{pmatrix} Cx + Du \\ u \end{pmatrix}, \tag{3}$$

maps $\mathcal{B}^s$ into $\mathcal{B}$ and is an isomorphism. In particular the Kalman system is a realization of the transfer matrix $P^{-1}Q = D + C(s\,\mathrm{id} - A)^{-1}B$. The observability of (2) signifies that (3) is injective. The engineering significance of a Kalman or state system (2) and of the representation (3) was, of course, obvious for its principal investigator [15] and is also discussed in [14, ch.2, ch.6], [25, ch.6] and [41, ch.3], for instance. One mathematical advantage of $\overset{\bullet}{x} = Ax + Bu$ compared to $P(\mathrm{d}/\mathrm{d}t)y = Q(\mathrm{d}/\mathrm{d}t)u$ is that for given continuous input u and *local state* $x(0)$ at time 0 it has the unique solution

$$\begin{aligned} x(t) &= e^{tA}x(0) + \int_0^t e^{(t-\tau)A}Bu(\tau)\mathrm{d}\tau, \\ y(t) &= Ce^{tA}x(0) + Du(t) + \int_0^t Ce^{(t-\tau)A}Bu(\tau)\mathrm{d}\tau. \end{aligned} \tag{4}$$

In particular

$$x \mid [0,\infty) = L^s(x(0), u \mid [0,\infty)),\ \ y \mid [0,\infty) = L(x(0), u \mid [0,\infty)) \tag{5}$$

where L^s and L are linear transfer operators. In this context, $x(0)$ is considered as the memory of the system or local state (local=temporal at $t = 0$) and the existence of L^s is usually considered as the state property of x.

In [25, (4.9),Th. 4.3.5], [31, Prop. 3.1], [39] and [40] Willems, Rapisarda and Polderman call any first order system

$$E\overset{\bullet}{x} + Fx + Gw = 0 \tag{6}$$

with complex matrices E, F and G a state space model with state or latent variable x and manifest variable w, the state property being defined by a suitable concatenability of trajectories of the behavior [25, Def. 4.3.3]. Any one-dimensional behavior is trivially isomorphic to such a state space model. Indeed, let

$$\mathcal{B} := \{w \in \mathcal{D}'^q;\ \ R_n w^{(n)} + \cdots + R_0 w = 0\},\ \ R_i \in \mathbb{C}^{g\times q}, \tag{7}$$

be any behavior and define the vector

$$x = (x_0, \cdots, x_{n-1})^\top := (w^{(0)}, \cdots, w^{(n-1)})^\top \in \mathcal{D}'^{nq}$$

of derivatives of w. It satisfies the obvious equations

$$\begin{aligned} \overset{\bullet}{x}_i = x_{i+1},\ \ i = 0, \cdots, n-2,\ \ x_0 - w = 0, \\ R_n\, \overset{\bullet}{x}_{n-1} + R_{n-1} x_{n-1} + \cdots + R_0 x_0 = 0 \end{aligned} \tag{8}$$

of the form (6) and the projection

$$\mathcal{B}^s := \left\{\begin{pmatrix} x \\ w \end{pmatrix} \in \mathcal{D}'^{(n+1)q};\ \begin{pmatrix} x \\ w \end{pmatrix} \text{ satisfies (8)}\right\} \to \mathcal{B},\ \begin{pmatrix} x \\ w \end{pmatrix} \mapsto w, \tag{9}$$

is an isomorphism. Hence $\mathcal{B}^s$ is a state space model of $\mathcal{B}$ in Willems' sense. This isomorphic representation of $\mathcal{B}$ is, however, weaker than Kalman's isomorphism (3) as is also shown in [25, Section 4.5].

The main Theorems 5.12 and 5.14 of the present paper generalize the isomorphism (3) to arbitrary linear systems of partial differential or difference equations with constant coefficients and show that any multidimensional behavior is *canonically* isomorphic to a *state behavior*, the isomorphism depending only on a preselected *term order* in the sense of the Gröbner basis theory. The notions of multidimensional state and state behavior are introduced in Definition 5.1 and justified by Result 5.6 as follows: Let $w = (w_\alpha)_{\alpha\in\Delta}$ denote a trajectory in an r-dimensional state behavior $\mathcal{B}$ with r independent variables $z_1, \cdots, z_r$ and components $w_\alpha(z_1, \cdots, z_r)$, $\alpha \in \Delta$. By definition the state behavior comes equipped with a family $(S(\alpha))_{\alpha\in\Delta}$ of subsets of the index set $\{1, \cdots, r\}$, and hence every trajectory $w \in \mathcal{B}$ gives rise to a family of partial functions $w_\alpha((z_\rho)_{\rho\in S(\alpha)}, 0)$. If $S(\alpha)$ is a proper subset of $\{1, \cdots, r\}$ the corresponding component $w_\alpha((z_\rho)_{\rho\in S(\alpha)}, 0)$ is called an *initial condition* or *local state* and otherwise an *input* or *free*. For instance, for $w := \binom{x}{u} \in \mathcal{B}^s$ from (3) u is an input and $x(0)$ is a local state. In contrast to the one-dimensional situation the local states are not constants, but functions. *The defining property of a state system for various spaces of formal or convergent power series is that arbitrary local states and inputs give rise to a unique trajectory or global state*. This property generalizes (4) and signifies the unique solvability of a suitable Cauchy or initial value problem. In particular, the state of the present paper has the property required by Pommaret in [27, p. 311, line 1^+-]: "Indeed *state* is what must be given in order to integrate the control system whenever the input is given." *Integration of the system* is the classical expression for solving it in reasonable function spaces.

For other function spaces like those of C^∞-functions or distributions global states may not uniquely exist for given local state and input. These problems are well-known from the non-analytic Cauchy problem in the area of partial differential equations and apparently inevitable and are explained, but by no means eliminated, in Remark 5.7 by reference to the huge literature and its many outstanding contributors. Therefore, the state property is defined by an algebraic counter-part of the unique solvability of the Cauchy problem for arbitrary polynomial (function) modules, for instance for delay-partial differential equations.

The justification of the state property by (5), an idea which also lies behind the so-called Nerode equivalence (compare [14, pp. 470–475]), is suitable only if time is the unique independent variable and the future of a system is much more important than the past. Therefore we generalize (4) and not (5) in the multidimensional situation.

Like many other objects in multidimensional systems theory the state representations are constructed via Gröbner bases for which we need the following data. Let $\mathcal{D} := F[s_1, \cdots, s_r]$ denote the polynomial algebra in r indeterminates over a field F and ${}_{\mathcal{D}}\mathcal{F}$ a $\mathcal{D}$-module with the scalar multiplication $\circ$. In the interesting applications $\mathcal{F}$ is an F-space of functions on which the indeterminates act as partial difference or differential operators or as translations as, for instance, in the case of delay-partial differential equations. A kernel representation of an $\mathcal{F}$-behavior is given by a matrix $R \in \mathcal{D}^{g\times q}$ and the derived

$$\begin{aligned}&\text{row module } U := \mathcal{D}^{1\times g}R \subset \mathcal{D}^{1\times q},\ \text{factor module } M := \mathcal{D}^{1\times q}/U \text{ and behavior}\\ &\mathcal{B} := U^{\perp} := \{w = (w_1, \cdots, w_q)^{\top} \in \mathcal{F}^q;\ \forall f = (f_1, \cdots, f_q) \in U:\ f \circ w\\ &\qquad = f_1 \circ w_1 + \cdots + f_q \circ w_q = 0\} = \{w \in \mathcal{F}^q;\ R \circ w = 0\}.\end{aligned} \tag{10}$$

The index set $I \times \mathbb{N}^r$, $I := \{1, \cdots, q\}$, is in one-one correspondence with the *terms*

$$s^{\mu}\delta_i = s^{\mu}(0, 0, \cdots, \overset{i}{1}, \cdots, 0) = (0, 0, \cdots, \overset{i}{s^{\mu}}, \cdots, 0),$$

ie., the standard F-basis of the free $\mathcal{D}$-module $\mathcal{D}^{1\times q}$. An arbitrarily chosen term order T on $I \times \mathbb{N}^r$ in the sense of the Gröbner basis theory gives rise to the degree set $\deg(U)$ of U, its complement $\Gamma = (I \times \mathbb{N}^r) \setminus \deg(U)$ (the lattice points 'below the stairs') and a canonical disjoint decomposition

$$\Gamma = \uplus\{\{i(\alpha)\} \times (\mu(\alpha) + \mathbb{N}^{S(\alpha)});\ \alpha = (i(\alpha), \mu(\alpha)) \in \Delta \subset I \times \mathbb{N}^r\} \tag{11}$$

where Δ is a finite subset of $I \times \mathbb{N}^r$, $S(\alpha) \subseteq \{1, \cdots, r\}$ for $\alpha \in \Delta$ and $\mathbb{N}^{S(\alpha)}$ is the submonoid of $\mathbb{N}^r$ of vectors with zero-components outside $S(\alpha)$. This decomposition is originally due to Riquier [32, pp. 143–168], was taken up and proven in modern language and without involutive division in [23, pp. 269–274] and was also proven by Janet in [13, pp. 74–91] and by Gerdt in [9, Decomposition Lemma 24] by means of involutive division. The main objective of Janet's paper [13] was, by the way, the simplification of Riquier's results [13, p. 66, lines 11^-, 12^-]. The Riquier decomposition induces a so-called *Stanley decomposition* of the module M:

$$M = \oplus_{\alpha\in\Delta} F[(s_{\rho})_{\rho\in S(\alpha)}]s^{\mu(\alpha)}\overline{\delta_{i(\alpha)}}. \tag{12}$$

This, in turn, implies a *simple algorithm for the computation of the Hilbert function and polynomial* of the projectivization $\widehat{M}$ of M with the Gröbner basis algorithm (Theorem 3.6) if T is a graded term order.

A related, but different derivation of the Hilbert polynomial, but not of the Hilbert function, is also contained in [26, p. 108, line 6^-] and the pages before the quoted one with the mathematical details and also in all other books of Pommaret, for instance in [28, Cor. 2.35 on p. 330]. Gerdt also computes the Hilbert function and polynomial

[9, p. 17]. Whereas our proof uses the Riquier decomposition of the set $\Gamma = (I \times \mathbb{N}^r) \setminus \deg(U)$ of *parametric* indices or derivatives Pommaret and Gerdt use Janet's translation and extension [13] of Riquier's results and the disjoint decomposition of the set $\deg(U)$ of *principal* indices for the computation of the Hilbert data. The terms *principal* and *parametric* are due to Riquier [32, p. 169]. Other variants of the computation of the Hilbert polynomial are [18, Th. 58, p. 65], [38, Prop. 1.4.2], [1].

Already Macaulay [18] established the relation of the Hilbert polynomial with his inverse systems. Pommaret discusses the significance of the Hilbert polynomial and of its characters, i.e., its coefficients in a special representation [28, Cor. 2.35 on p. 330], for systems of partial differential equations [26, Ch.3], [28, Ch.3, Section 2]. Finally Wood et al. [42] investigated the Hilbert function in context with structure indices of a multidimensional system.

By definition the system $\mathcal{B}$ is a state behavior with respect to the term order T if and only if the leading terms of the unique reduced Gröbner basis of U with respect to T are of the form $s_\rho \delta_i$. If this is not the case we use this reduced Gröbner basis and the decomposition (11) to construct a canonical state behavior $\mathcal{B}^s$ and mutually inverse behavior isomorphisms

$$C\circ : \mathcal{B}^s \to \mathcal{B} \text{ and } E\circ : \mathcal{B} \to \mathcal{B}^s$$

where the sets $S(\alpha)$, $\alpha \in \Delta$, from (11) are exactly those which define the local states $w_\alpha((z_\rho)_{\rho\in S(\alpha)}, 0)$ of the state behavior $\mathcal{B}^s$ (Theorems 5.12 and 5.14). The matrix C has coefficients in F and corresponds to $\begin{pmatrix} C & D \\ 0 & \mathrm{id} \end{pmatrix}$ from (3) whereas E is a polynomial isomorphism and called a state map in [31, (4.2)]. In contrast to Willems' state space models and to Pommaret's *generalized Kalman form* as a specialization of the *Spencer form* [29], [30, p. 26 ff, after Ex. 3.3.14] the canonical state representation of a behavior of the present paper is in general not of the first order, i.e., of total degree one, even if $\mathcal{B}$ itself has this property. In colloquial terms: first order systems may hide non-apparent complications. Zerz [43] has constructed interesting first order representations and thus state space models in Willems' sense of arbitrary behaviors by means of linear fractional transformations, see Remark 5.8.

The algorithms of this paper have been implemented in MAPLE for two-dimensional systems. Thousands of randomly generated examples have been computed, a very simple, but instructive one is given in Example 5.16. In Example 5.15, Kalman's isomorphism (3) is derived in detail on two pages. I believe that for this standard one-dimensional case the algorithms of this paper compare well with those from the literature, for instance from [14, Section 6.7], [15], [25, Th. 6.4.2], [31, Section 6], [37, Section 2.2.1], [41, Th. 5.2.14].

In Section 2 of this paper Theorem 2.1 describes three combinatorial properties of the Riquier decomposition of the complement of $\deg(U)$ which are used in the proof of the main Theorem 5.12 to verify that the constructed behavior $\mathcal{B}^s$ is indeed a state behavior. Section 3 contains some generalities on multidimensional behaviors, the algorithm for the computation of the Hilbert function and a method to eliminate non-essential components from $\mathcal{B}$ before the construction of $\mathcal{B}^s$. In Section 4 we construct the canonical input/output structure of a behavior when a term order is chosen. The principal section Section 5 introduces and justifies the state terminology and contains the main Theorem 5.12 and its refinement 5.14 and the one-dimensional resp. two-dimensional Examples 5.15 resp. 5.16.

The present paper belongs to *Algebraic Analysis*, but restricted to linear differential or difference operators with *constant coefficients*. Its main technique is the use of algebra for the derivation of systems theoretic and analytic results. An outstanding early contributor to this field was Riquier whose pioneering work [32] has not received the recognition it deserves, neither by analysts nor by algebraists. Sturmfels in [35, ch. 10] like already Gröbner uses these techniques in *Algebraic Geometry*. In his various books (see the References) Pommaret extends the *algebraic or formal*, but not the analytic theory of systems of partial differential equations to linear ones with *variable* coefficients and even to *non-linear* ones on the basis of the work of Janet [13] and of Spencer.

2 The Combinatorics of Order Ideals

The state construction of this paper essentially relies on Buchberger's Gröbner basis algorithm for a polynomial module U and on a canonical disjoint decomposition of the complement of the degree set $\deg(U)$ *of U which is originally due to Riquier* [32, pp. 143–168]*and which we reproved in* [23, pp. 269–274] *in modern language. See the Introduction for more bibliographical details. Below we shortly repeat the algorithm for this decomposition. In Theorem 2.1 we prove several combinatorial properties of it which are important ingredients of the proof of the main Theorem 5.12 of this paper.*

Let $r > 0$ be a natural number and $[r] := \{1, \cdots, r\}$. Consider the additive monoid $\mathbb{N}^r := \{\mu = (\mu_1, \cdots, \mu_r);\ \mu_i \in \mathbb{N}\}$ of vectors of natural numbers with its basis $\epsilon_\rho = (0, \cdots, 0, \overset{\rho}{1}, 0, \cdots, 0),\ \rho = 1, \cdots, r$. On $\mathbb{N}^r$ we use the *componentwise order* $\le_{\text{cw}}$, defined by $\mu \le_{\text{cw}} \nu \Leftrightarrow \nu \in \mu + \mathbb{N}^r$, and the *lexicographic well-order* $<_{\text{lex}}$ with $\epsilon_r <_{\text{lex}} \cdots <_{\text{lex}} \epsilon_2 <_{\text{lex}} \epsilon_1$. If $\nu = \mu + \epsilon_\rho$ for some ρ we call ν resp μ a *direct successor resp. predecessor* of μ resp. ν. Let N be an *order ideal* of $\mathbb{N}^r$, i.e., a subset satisfying $N + \mathbb{N}^r = N$, $\Gamma := \mathbb{N}^r \setminus N$ its complement and $D := \min_{\text{cw}} N$ the set of minimal elements of N with respect to the componentwise order. It is of course discrete with respect to the cw-order on $\mathbb{N}^r$ and *finite* (Dickson's lemma) and $N = D + \mathbb{N}^r$.

If M is an additive monoid, for instance $\mathbb{N}$, $\mathbb{R}$ or $\mathbb{C}$, and if S is a subset of $\{1, \cdots, r\}$ with its complement $S' := \{1, \cdots, r\} \setminus S$ we *identify* M^S as subset of M^r via

$$\begin{aligned} M^S &= \{x = (x_1, \cdots, x_r) \in M;\ \forall \rho \notin S:\ x_\rho = 0\} \text{ and} \\ M^r &= M^S \times M^{S'} \ni x = (x^{(S)}, x^{(S')}) \text{ where } x^{(S)} := (x_\rho)_{\rho \in S}. \end{aligned} \tag{13}$$

The disjoint decomposition of Γ is described by a *finite* subset $\Delta \subseteq \Gamma$ and index sets $S(\alpha) \subseteq \{1, \cdots, r\},\ \alpha \in \Delta$, such that

$$\mathbb{N}^r = \Gamma \uplus N,\ \Gamma = \uplus_{\alpha \in \Delta}(\alpha + \mathbb{N}^{S(\alpha)}). \tag{14}$$

For $r = 1$ the decomposition is given as

$$\begin{cases} \Delta := \{0\},\ S(0) = \{1\} & \text{if } N = \emptyset \\ \Gamma = \Delta = \emptyset & \text{if } N = \mathbb{N} \\ \Delta = \Gamma = \{0, \cdots, d-1\},\ S(\alpha) := \emptyset \text{ for } \alpha < d & \text{if } N = d + \mathbb{N},\ 0 < d. \end{cases} \tag{15}$$

Assume now $r > 1$ and that the decomposition has already be constructed for order ideals in $\mathbb{N}^{r-1}$. We identify

$$\mathbb{N}^r := \mathbb{N} \times \mathbb{N}^{r-1} \ni \mu = (\mu_1, \mu_{II}). \tag{16}$$

We define the $(r-1)$-dimensional order ideals

$$\begin{aligned}
&N_{II}(i) = \{\mu_{II} \in \mathbb{N}^{r-1};\ (i, \mu_{II}) \in N\},\ N_{II} := \cup_{i\in\mathbb{N}} N_{II}(i),\\
&\Gamma_{II}(i) := \mathbb{N}^{r-1} \setminus N_{II}(i),\ \Gamma_{II} := \mathbb{N}^{r-1} \setminus N_{II} = \cap_{i\in\mathbb{N}} \Gamma_{II}(i) \text{ and by induction}\\
&\qquad \Delta_{II}(i) \subset \Gamma_{II}(i),\ S_{II,i}(\alpha_{II}) \subset \{2, \cdots, r\} \text{ with}\\
&\qquad \Gamma_{II}(i) = \uplus_{\alpha_{II}\in\Delta_{II}(i)} (\alpha_{II} + \mathbb{N}^{S_{II,i}(\alpha_{II})})
\end{aligned} \tag{17}$$

Let $k := \max_{d\in D} d_1$. Then

$$\begin{aligned}
&N_{II}(0) \subseteq N_{II}(1) \subseteq \cdots \subseteq N_{II}(k-1) \subset N_{II} := N_{II}(k) = N_{II}(k+1) = \cdots\\
&\Gamma_{II}(0) \supseteq \Gamma_{II}(1) \supseteq \cdots \supseteq \Gamma_{II}(k-1) \supset \Gamma_{II} = \Gamma_{II}(k) = \Gamma_{II}(k+1) = \cdots
\end{aligned}$$

and the sets Δ and $S(\alpha)$, $\alpha \in \Delta$, for the order ideal N are constructed according to [23, p. 271] as

$$\begin{aligned}
&\Delta(i) := \{i\} \times \Delta_{II}(i),\ i = 0, \cdots, k,\ \Delta := \Delta(0) \uplus \Delta(1) \uplus \cdots \uplus \Delta(k),\\
&S(\alpha) := \begin{cases} S_{II,i}(\alpha_{II}) \subseteq \{2, \cdots, r\} & \text{if } \alpha = (i, \alpha_{II}) \in \Delta(i),\ i < k,\\ \{1\} \cup S_{II,k}(\alpha_{II}) & \text{if } \alpha = (k, \alpha_{II}) \in \Delta(k). \end{cases}
\end{aligned} \tag{18}$$

The following theorem contains those properties of the data just introduced which are essential for the construction of the canonical state representation of a multidimensional behavior.

Theorem 2.1 *Data from above, i.e.,* $N = N + \mathbb{N}^r \subseteq \mathbb{N}^r$, $D := \min_{\mathrm{cw}}(N)$, $\Delta \subset \Gamma := \mathbb{N}^r \setminus N$ *and* $S(\alpha)$, $\alpha \in \Delta$, *according to* (18).

1. *If* $0 \neq \alpha \in \Delta$ *there is a direct predecessor* $\beta := \alpha - \epsilon_\rho$ *of* α *in* Δ *and of course* $\rho \notin S(\beta)$.
2. *If* $0 \neq d \in D$ *there is a direct predecessor* $\beta := d - \epsilon_\rho$ *of* d *in* Δ *and of course* $\rho \notin S(\beta)$.
3. *If* $\alpha \in \Delta$, $\rho \notin S(\alpha)$ *and* $\alpha + \epsilon_\rho = \beta + \mu \in \Gamma$ *with* $\beta \in \Delta$ *and* $\mu \in \mathbb{N}^{S(\beta)}$ *then* $\alpha <_{\mathrm{lex}} \beta$.
4. *If*

$$\begin{cases} N = \emptyset \\ N = \mathbb{N}^r \\ N \neq \emptyset,\ \Gamma \neq \emptyset \end{cases} \quad \textit{then} \begin{cases} \Gamma = \mathbb{N}^r, \Delta = \{0\},\ S(0) = \{1, \cdots, r\} \\ 0 \in D \\ 0 \in \Delta,\ S(0) \neq \{1, \cdots, r\}. \end{cases}$$

Proof The last assertion is obvious. The proof of the other assertions proceeds by induction on r. Assume first that

$r = 1$:

1. Case: $N = D = \emptyset, \Gamma = \mathbb{N}, \Delta = \{0\}, S(0) = \{1\}$: The assumptions of the three assertions are not satisfied.

2. Case: $N = \mathbb{N}$, $D = \{0\}$, $\Gamma = \Delta = \emptyset$: Again the assumptions are not satisfied.
3. Case: $N = d + \mathbb{N}$, $0 < d$, $D := \{d\}$, $\Gamma = \Delta = \{0, \cdots, d-1\}$, $S(\alpha) = \emptyset$ for $\alpha \in \Delta$:

a. If $0 < \alpha < d$ then $\alpha - \epsilon_1 = \alpha - 1 \in \Delta$.
b. $d - \epsilon_1 = d - 1 \in \Delta$.
c. If $\alpha \in \Delta$ and $\beta := \alpha + \epsilon_1 = \alpha + 1 \in \Gamma$, ie. $\alpha < \beta = \alpha + 1 < d$, then $\beta \in \Delta$ and $\alpha <_{\text{lex}} \beta$.

$r > 1$: We assume that the theorem is true for order ideals in $\mathbb{N}^{r-1}$.

1. Let $0 \neq \alpha = (\alpha_1, \alpha_{II}) \in \Delta$, whence $\alpha_{II} \in \Delta_{II}(\alpha_1)$ by construction.

1. Case: $\alpha_{II} \neq 0$: By induction applied to the order ideal $N_{II}(\alpha_1) \subseteq \mathbb{N}^{r-1}$ there is a direct predecessor $\beta_{II} = \alpha_{II} - \epsilon_\rho$ of α_{II} in $\Delta_{II}(\alpha_1)$ with $\rho \in \{2, \cdots, r\}$. Then $\beta := (\alpha_1, \beta_{II}) = \alpha - \epsilon_\rho$ is a direct predecessor of α in $\Delta(\alpha_1) \subseteq \Delta$.

2. Case: $\alpha_{II} = 0, \alpha_1 > 0$: Then

$$0 \in \Delta_{II}(\alpha_1) \subseteq \Gamma_{II}(\alpha_1) \subseteq \Gamma_{II}(\alpha_1 - 1) \Rightarrow$$
$$0 \in \Delta_{II}(\alpha_1 - 1) \text{ and } \alpha - \epsilon_1 = (\alpha_1 - 1, 0) \in \Delta(\alpha_1 - 1) \subseteq \Delta.$$

2. Let $0 \neq d = (d_1, d_{II}) \in D$. Then $d_{II} \in \min_{\text{cw}}(N_{II}(d_1))$: Indeed,

$$N = D + \mathbb{N}^r = \cup_{d' \in D}(d' + \mathbb{N}^r) \text{ implies}$$
$$N_{II}(d_1) = \cup_{d'=(d'_1, d'_{II}) \in D, d'_1 \le d_1}(d'_{II} + \mathbb{N}^{r-1}).$$

Therefore, if $\mu_{II} \le_{\text{cw}} d_{II}$ in $N_{II}(d_1)$ there is a $d' \in D$ with $d'_1 \le d_1$ and $d'_{II} \le_{\text{cw}} \mu_{II} \le_{\text{cw}} d_{II}$, hence $d' \le_{\text{cw}} d$. The discreteness of D implies $d = d'$, hence $d'_{II} = \mu_{II} = d_{II}$ and the minimality of d_{II}.

1. Case: $d_{II} \neq 0$: By induction applied to $N_{II}(d_1)$ there is a direct predecessor $\alpha_{II} = d_{II} - \epsilon_\rho$ of d_{II} in $\Delta_{II}(d_1)$ with $\rho \in \{2, \cdots, r\}$ and $d - \epsilon_\rho = (d_1, d_{II} - \epsilon_\rho)$ is a direct predecessor of d in $\Delta(d_1) \subseteq \Delta$.

2. Case: $d_{II} = 0, d_1 > 0$: Every $\mu_{II} \in N_{II}(d_1 - 1)$ is non-zero: Indeed, if $\mu_{II} = 0$ there is a $d' \in D$ with $d'_1 \le d_1 - 1$ and $d'_{II} \le_{\text{cw}} \mu_{II} = 0$, hence $d' = (d'_1, 0) <_{\text{cw}} d$, a contradiction to the discreteness of D.

1. Subcase: $N_{II}(d_1 - 1) \neq \emptyset$: The preceding consideration shows $0 \notin N_{II}(d_1 - 1)$, hence $0 \in \Delta_{II}(d_1 - 1) \subseteq \Gamma_{II}(d_1 - 1)$ and $d - \epsilon_1 = (d_1\text{-}1, 0) \in \Delta(d_1 - 1) \subseteq \Delta$.

2. Subcase: $N_{II}(d_1 - 1) = \emptyset$: Then $\Gamma_{II}(d_1 - 1) = \mathbb{N}^{r-1}$, $\Delta_{II}(d_1 - 1) = \{0\}$ and again $d - \epsilon_1 = (d_1 - 1, 0) \in \Delta(d_1 - 1) \subseteq \Delta$.

3. Let

$$\alpha = (i, \alpha_{II}) \in \Delta(i),\ \rho \notin S(\alpha) = \begin{cases} S_{II,i}(\alpha_{II}) & \text{if } i < k \\ \{1\} \cup S_{II,i}(\alpha_{II}) & \text{if } i = k \end{cases},\ \text{and}$$

$$\alpha + \epsilon_\rho = \beta + \mu \in \Gamma,\ \beta \in \Delta,\ \mu \in \mathbb{N}^{S(\beta)}.$$

1. Case: $i < k, \rho \neq 1$, hence $\rho \notin S_{II,i}(\alpha_{II})$ and $\alpha + \epsilon_\rho = (i, \alpha_{II} + \epsilon_\rho)$: Then $\alpha_{II} + \epsilon_\rho \in \Gamma_{II}(i)$ and $\alpha_{II} + \epsilon_\rho = \beta_{II} + \mu_{II}$ with $\beta_{II} \in \Delta_{II}(i)$ and $\mu_{II} \in \mathbb{N}^{S_{II,i}(\beta_{II})}$. By induction applied to the order ideal $N_{II}(i)$ we obtain that $\alpha_{II} <_{\text{lex}} \beta_{II}$. Moreover

$$\beta := (i, \beta_{II}) \in \Delta(i) \subseteq \Delta,\ S(\beta) = S_{II,i}(\beta_{II}),$$
$$\alpha + \epsilon_\rho = \beta + (0, \mu_{II}) \text{ with } (0, \mu_{II}) \in \mathbb{N}^{S(\beta)} \text{ and}$$
$$\alpha = (i, \alpha_{II}) <_{\text{lex}} \beta = (i, \beta_{II}).$$

2. Case: $\alpha_1 = i < k - 1$ and $\rho = 1$: Then

$$\alpha + \epsilon_1 = (i+1, \alpha_{II}) = \beta + \mu = (\beta_1 + \mu_1, \beta_{II} + \mu_{II}),\ \text{hence}$$
$$\beta_1 \leq i + 1 < k,\ S(\beta) = S_{II,\beta_1}(\beta_{II}) \subseteq \{2, \cdots, r\},$$
$$\mu_1 = 0 \text{ since } \mu \in \mathbb{N}^{S(\beta)} \text{ and finally } \alpha_1 = i < i + 1 = \beta_1 \text{ and } \alpha <_{\text{lex}} \beta$$

as asserted.

3. Case: $\alpha = (k-1, \alpha_{II}),\ \rho = 1$: Again

$$\alpha + \epsilon_1 = (k, \alpha_{II}) = \beta + \mu = (\beta_1 + \mu_1, \beta_{II} + \mu_{II}) \in \Gamma.$$

1. Subcase: $\mu_1 = 0$: Then $\alpha_1 = k - 1 < k = \beta_1$ and $\alpha <_{\text{lex}} \beta$ as asserted.
2. Subcase: $\mu_1 > 0$: Then $\beta_1 < k$ and hence $S(\beta) = S_{II,\beta_1}(\beta_{II}) \subset \{2, \cdots, r\}$, $\mu \in \mathbb{N}^{\{2,\cdots,r\}}$ and $\mu_1 = 0$, a contradiction. Thus the second subcase cannot occur and the assertion is proved in the third case.

4. Case: $\alpha = (k, \alpha_{II})$, i.e.,

$$\alpha_{II} \in \Delta_{II}(k),\ \rho \notin S(\alpha) = \{1\} \cup S_{II,k}(\alpha_{II}),\ \text{hence}$$
$$\rho \neq 1,\ \rho \notin S_{II,k}(\alpha_{II}),\ \alpha + \epsilon_\rho = (k, \alpha_{II} + \epsilon_\rho) \text{ and } \alpha_{II} + \epsilon_\rho \in \Gamma_{II}(k):$$

As in the first case we infer, by induction applied to $N_{II}(k)$,

$$\alpha_{II} + \epsilon_\rho = \beta_{II} + \mu_{II} \text{ with}$$
$$\beta_{II} \in \Delta_{II}(k),\ \mu_{II} \in \mathbb{N}^{S_{II,k}(\beta_{II})} \text{ and } \alpha_{II} <_{\text{lex}} \beta_{II}.$$

But then

$$\beta := (k, \beta_{II}) \in \Delta(k) \subseteq \Delta,\ \alpha = (k, \alpha_{II}) <_{\text{lex}} (k, \beta_{II}),$$
$$S(\beta) = \{1\} \cup S_{II,k}(\beta_{II}) \text{ and } \alpha + \epsilon_\rho = \beta + (0, \mu_{II}),\ (0, \mu_{II}) \in \mathbb{N}^{S(\beta)}.$$

Hence the assertion is proved in the fourth case too and the proof is completed. □

The following picture illustrates the preceding theorem for the order ideal (Figure 1) $N = D + \mathbb{N}^2$ with $D = \min_{\text{cw}}(N) := \{(2,4), (4,2), (5,1)\}$.

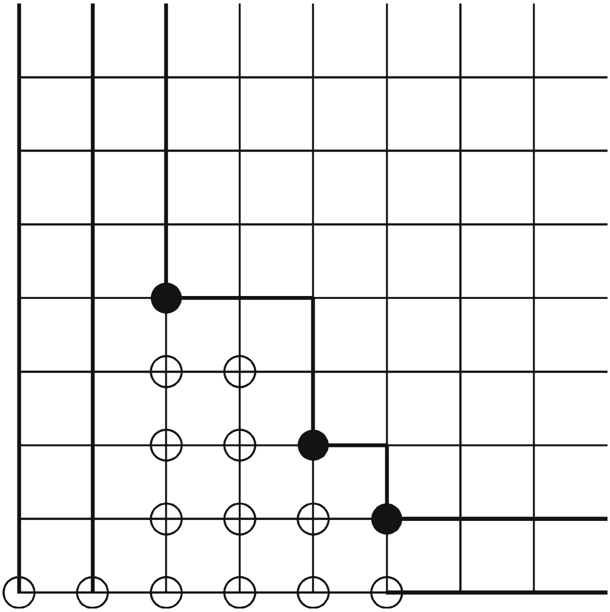

Figure 1 The order ideal N above the stair-case, the points $\bullet$ of $\min_{\mathrm{cw}}(N)$ and $\circ$ of Δ.

We generalize the preceding considerations to sets $I \times \mathbb{N}^r$ where I is an arbitrary finite index set. For $\alpha = (i, \mu) \in I \times \mathbb{N}^r$ and $\nu \in \mathbb{N}^r$ we define

$$i(\alpha) := i,\ \ \mu(\alpha) = \mu,\ \text{ hence } \alpha = (i(\alpha), \mu(\alpha)),\ \text{ and } \alpha + \nu := (i, \mu + \nu). \tag{19}$$

This addition is an action of the monoid $\mathbb{N}^r$ on the set $I \times \mathbb{N}^r$ and again induces the componentwise order

$$(i, \mu) \leq_{\mathrm{cw}} (j, \nu) \ \Leftrightarrow\ \exists \lambda \in \mathbb{N}^r \text{ with } (j, \nu) = (i, \mu) + \lambda,\ \text{ hence } i = j. \tag{20}$$

Let N be an (order) *submodule* of $I \times \mathbb{N}^r$, i.e., a subset satisfying $N + \mathbb{N}^r = N$, and $\Gamma := (I \times \mathbb{N}^r) \setminus N$ its complement. For every $i \in I$ the set

$$N_i := \{\mu \in \mathbb{N}^r;\ (i, \mu) \in N\} \text{ with } D_i := \min_{\mathrm{cw}}(N_i) \text{ and } \Gamma_i := \mathbb{N}^r \setminus N_i \tag{21}$$

is an order ideal of $\mathbb{N}^r$,

$$N := \uplus_{i \in I}\{i\} \times N_i,\ \ \Gamma = \uplus_{i \in I}\{i\} \times \Gamma_i \text{ and } D := \uplus_{i \in I}\{i\} \times D_i \tag{22}$$

where D is the set of minimal elements of N with respect to the cw-order on $I \times \mathbb{N}^r$. To each N_i we apply the decomposition from (18) and obtain

$$\mathbb{N}^r := \Gamma_i \uplus N_i,\ \ \Gamma_i = \uplus_{\mu \in \Delta_i}(\mu + \mathbb{N}^{S_i(\mu)}). \tag{23}$$

We define

$$\Delta(i) := \{i\} \times \Delta_i,\ \ \Delta := \uplus_{i \in I}\Delta(i)$$

$$S(\alpha) := S_i(\mu(\alpha)) \text{ for } \alpha = (i, \mu) \in \Delta,\ \text{ ie. } \mu \in \Delta_i, \tag{24}$$

and obtain the decomposition

$$I \times \mathbb{N}^r = \Gamma \uplus N,\ \ \Gamma = \uplus_{\alpha\in\Delta}(\alpha + \mathbb{N}^{S(\alpha)}) \tag{25}$$

which generalizes that of (18) from the monoid $\mathbb{N}^r$ to the set $I \times \mathbb{N}^r$.

Remark 2.2 The preceding construction has been implemented in MAPLE for the two-dimensional situation $\{1, \cdots, l\} \times \mathbb{N}^2$.

The following corollary of Theorem 2.1 generalizes that theorem; its proof is obvious.

Corollary 2.3 Data from above,

1. If $\alpha \in \Delta$ and $\mu(\alpha) > 0$ there is a direct predecessor $\beta := \alpha - \epsilon_\rho$ of α in Δ and of course $\rho \notin S(\beta)$.
2. If $d \in D$ and $\mu(d) > 0$ there is a direct predecessor $\beta := d - \epsilon_\rho \in \Delta$ of d in Δ and of course $\rho \notin S(\beta)$.
3. If $\alpha = (i(\alpha), \mu(\alpha)) \in \Delta,\ \rho \notin S(\alpha)$ and $\alpha + \epsilon_\rho = \beta + \mu \in \Gamma$ with $\beta \in \Delta$ and $\mu \in \mathbb{N}^{S(\beta)}$, in particular $i(\alpha) = i(\beta)$, then $\mu(\alpha) <_{\text{lex}} \mu(\beta)$.
4. If

$$\begin{cases} N_i = \emptyset \\ N_i = \mathbb{N}^r \\ N_i \neq \emptyset,\ \Gamma_i \neq \emptyset \end{cases} \quad \text{then} \quad \begin{cases} \Gamma_i = \mathbb{N}^r,\ (i,0) \in \Delta,\ S(i,0) = \{1, \cdots, r\} \\ \Gamma_i = \emptyset,\ (i,0) \in D \\ (i,0) \in \Delta,\ S(i,0) \neq \{1, \cdots, r\}. \end{cases}$$

3 The Normal Form and Hilbert Function of Multidimensional Behaviors

This section contains an introduction into the language of multidimensional systems and the notations which we are going to use [20, 23], *the computation of the Hilbert data of a system (Theorem* 3.6*) and the elimination of non-essential components (Theorem* 3.12*).* The letter F denotes a field and $\mathcal{D} := F[s] := F[s_1, \cdots, s_r]$ the polynomial algebra in r indeterminates over F. We will use the ring $\mathcal{D}$ as ring of differential or difference operators, hence the letter $\mathcal{D}$ for this ring. Let $K := F(s)$ denote the quotient field of $\mathcal{D}$ of rational functions.

It turns out and will be understood below that for the purposes of this paper we need matrices whose row and column indices belong to arbitrary finite index sets instead of sets $\{1, 2, \cdots, p\}$ etc. only. If M is a set and I and J are finite sets the elements $A = (A_{ij})_{i\in I, j\in J} \in M^{I\times J}$ are called $I \times J$-matrices with coefficients in M. The elements $f = (f_i)_{i\in I} \in M^{1\times I} := M^{\{1\}\times I}$ resp. $w = (w_i)_{i\in I} \in M^I := M^{I\times\{1\}}$ are called row resp. column vectors as usual. If M is an additive monoid $M^{I\times J}$ is the same with the usual componentwise addition. If M is a $\mathcal{D}$-module with the scalar multiplication $\circ$ and $R \in \mathcal{D}^{D\times I}$ and $w \in M^{I\times J}$ are matrices then

$$R \circ w := \left(\sum_{i\in I} R_{di} \circ w_{ij}\right)_{d\in D, j\in J} \in M^{D\times J}.$$

If M is a $\mathcal{D}$-module a matrix $E \in \mathcal{D}^{I\times J}$ induces the $\mathcal{D}$-linear maps

$$\circ E : M^{1\times I} \to M^{1\times J},\ x \mapsto x \circ E,\ \text{and } E\circ : M^J \to M^I,\ w \mapsto E \circ w. \tag{26}$$

Remark that

$$(E_2E_1)\circ = (E_2\circ)(E_1\circ),\ \text{but } \circ(E_2E_1) = (\circ E_1)(\circ E_2).$$

We identify

$$\begin{aligned} \mathcal{D}^{I\times J} &= \mathrm{Hom}_{\mathcal{D}}(\mathcal{D}^{1\times I}, \mathcal{D}^{1\times J}),\ E = \circ E \text{ and call}\\ \mathrm{row}(E) &:= \mathrm{im}(\circ E) = \mathcal{D}^{1\times I}E = \sum_{i\in I} \mathcal{D}E_{i-} \text{ resp.}\\ \mathrm{col}(E) &:= \mathrm{im}(E\circ) = E\mathcal{D}^J = \sum_{j\in J} \mathcal{D}E_{-j} \end{aligned} \tag{27}$$

the row- resp. column-module of the matrix E. Since $\mathcal{D}$ is noetherian every submodule of $\mathcal{D}^{1\times I}$ is a row module of some matrix. The standard basis of $\mathcal{D}^{1\times I}$ and of $\mathcal{D}^I$ is denoted by

$$\delta_i := (0, \cdots, 0, \overset{i}{1}, 0, \cdots, 0)^{(\top)}. \tag{28}$$

Let $\mathrm{mod}_{\mathcal{D}}$ denote the category of $\mathcal{D}$-modules and let $\mathcal{F}$ be any $\mathcal{D}$-module which in the main part of the paper is a $\mathcal{D}$-module of functions, hence the letter $\mathcal{F}$. We consider the duality functor (with respect to $\mathcal{F}$)

$$\begin{aligned} D &:= \mathrm{Hom}_{\mathcal{D}}(-, \mathcal{F}) : \quad \mathrm{mod}_{\mathcal{D}}^{\mathrm{op}} \longrightarrow \quad \mathrm{mod}_{\mathcal{D}},\ \text{and identify}\\ \mathcal{F}^I &= D(\mathcal{D}^{1\times I}) = \mathrm{Hom}_{\mathcal{D}}(\mathcal{D}^{1\times I}, \mathcal{F}),\ w = (w_i)_{i\in I} = (\delta_i \mapsto w_i),\ \text{and}\\ E\circ &= D(\circ E) = \mathrm{Hom}(\circ E, \mathcal{F}) :\ \mathcal{F}^J \to \mathcal{F}^I\\ \text{for } E &= \circ E \in \mathcal{D}^{I\times J} = \mathrm{Hom}_{\mathcal{D}}(\mathcal{D}^{1\times I}, \mathcal{D}^{1\times J}). \end{aligned} \tag{29}$$

The map

$$\mathcal{D}^{1\times I} \times \mathcal{F}^I \to \mathcal{F},\ ((f_i)_{i\in I}, (w_i)_{i\in I}) \mapsto f \circ w = \sum_{i\in I} f_i \circ w_i, \tag{30}$$

is $\mathcal{D}$-bilinear. Let $\mathbb{P}(M)$ denote the *projective geometry* of all $\mathcal{D}$-submodules of a $\mathcal{D}$-module M. The bilinear map (30) gives rise to the *Galois correspondence*

$$\begin{aligned} &\mathbb{P}(\mathcal{D}^{1\times I}) \rightleftarrows \mathbb{P}(\mathcal{F}^I),\\ &U \to U^\perp := \{w \in \mathcal{F}^I;\ \forall f \in U : f \circ w = 0\}\\ &\qquad \mathcal{B}^\perp := \{f \in \mathcal{D}^{1\times I};\ \forall w \in \mathcal{B} :\ f \circ w = 0\} \leftarrow \mathcal{B} \text{ with}\\ &\qquad U \subseteq U^{\perp\perp},\ \mathcal{B} \subseteq \mathcal{B}^{\perp\perp}. \end{aligned} \tag{31}$$

The submodules of $\mathcal{F}^I$ of the form $\mathcal{B} := U^\perp$ are called *behaviors* [25], specifically $\mathcal{F}$-behaviors, and satisfy $\mathcal{B}^{\perp\perp} = \mathcal{B}$. In [20, Def. 15, p. 21] they were called $\mathcal{F}$-systems. If

$$G \text{ is a finite subset of } \mathcal{D}^{1\times I},\ \ R \in \mathcal{D}^{G\times I} \text{ with } R_{g-} := g \text{ and}$$

$$U := \mathrm{row}(R) = \sum_{g\in G} \mathcal{D}g \text{ then}$$

$$\mathcal{B} := U^\perp = \{w \in \mathcal{F}^I;\ \forall g \in G : g \circ w = 0\} = \{w \in \mathcal{F}^I;\ \ R \circ w = 0\}. \quad (32)$$

The latter representation of $\mathcal{B}$ is called a *kernel representation* in the systems community [25]. Consider the factor module

$$M := \mathcal{D}^{1\times I}/U \text{ and its dual } \mathrm{D}(M) = \mathrm{Hom}_{\mathcal{D}}(\mathcal{D}^{1\times I}/U, \mathcal{F}). \quad (33)$$

The identification (29) induces the identification [20, p. 21]

$$U^\perp = \mathcal{B} = \mathrm{D}(M),\ \ w = (w_i)_{i\in I} = (\overline{\delta_i} \mapsto w_i), \quad (34)$$

where $\overline{\delta_i}$ denotes the residue class of δ_i in M. This simple, but important observation is originally due to B. Malgrange and one starting point of *Algebraic Analysis*. If $M_i = \mathcal{D}^{1\times I_i}/U_i,\ i = 1, 2$ are two finitely generated modules with the associated behaviors $\mathcal{B}_i = U_i^\perp$ a matrix $E \in \mathcal{D}^{I_1\times I_2}$ induces a linear map

$$(\circ E)_{\mathrm{ind}} : M_1 \to M_2,\ \overline{\delta_{i_1}} \mapsto \overline{E_{i_1-}} \text{ if and only } U_1 E \subset U_2, \quad (35)$$

and then we also identify

$$E\circ = \mathrm{D}((\circ E)_{\mathrm{ind}}) : \mathcal{B}_2 = \mathrm{D}(M_2) \to \mathcal{B}_1 = \mathrm{D}(M_1). \quad (36)$$

The latter maps are called *behavior or system morphisms*.

In general, the identity $U = U^{\perp\perp}$ does not hold. It holds, however, for injective cogenerators and important function modules $\mathcal{F}$ (see Example 3.1). The module $\mathcal{F}$ is called an *injective cogenerator* [20, p. 29] if the duality functor $D := \mathrm{Hom}_{\mathcal{D}}(-, \mathcal{F})$ preserves and reflects the exactness of sequences. The following examples were discussed in [20].

Example 3.1 The following function modules except the last one are injective cogenerators over $F[s]$.

1. The field F is that of real or complex numbers and $\mathcal{F} := C^\infty(\mathbb{R}^r)$ is the F-space of infinitely often differentiable, F-valued functions $y(z) = y(z_1, \cdots, z_r)$ on which $F[s]$ acts by partial differentiation via $s_\rho \circ y = \partial y/\partial z_\rho$.
2. The field F is that of real or complex numbers and $\mathcal{F} := \mathcal{D}'(\mathbb{R}^r)$ is the F-space of F-valued distributions on which $F[s]$ acts by partial differentiation.
3. The field is arbitrary and

$$\mathcal{F} := F^{\mathbb{N}^r} = F[[z]] = F[[z_1, \cdots, z_r]] \ni y = (y_\mu)_{\mu\in\mathbb{N}^r} = \sum_{\mu\in\mathbb{N}^r} y_\mu z^\mu$$

   is the space of multi-sequences or formal power series on which $F[s]$ acts by left shifts via $(s_\rho \circ y)_\mu = y_{\mu+\epsilon_\rho}$.
4. The field is of characteristic zero and $\mathcal{F} = F[[z]]$ is the space of power series on which $F[s]$ acts by partial differentiation.

5. $\mathcal{F}$ is one of the spaces from the preceding examples and

$$\mathcal{F}_{\mathrm{lf}} := \{y \in \mathcal{F};\ \dim_F(\mathcal{D}y) < \infty\}$$

is the injective cogenerator of the *locally finite* elements of $\mathcal{F}$ [21, 1.13, 1.14]. For instance, if $F = \mathbb{C}$, then [21, th. 6.6]

$$\mathcal{D}'(\mathbb{R})_{\mathrm{lf}} = C^\infty(\mathbb{R}^r)_{\mathrm{lf}} = \oplus_{\zeta \in \mathbb{C}^r} \mathbb{C}[z] \exp(\zeta \bullet z) \text{ with } \zeta \bullet z := \zeta_1 z_1 + \cdots + \zeta_r z_r$$

is the space of *polynomial-exponential* functions.
6. (Delay-differential equations, [20, p. 17]) Let F be the field of complex numbers, $\mathcal{D} := \mathbb{C}[s_1, \cdots, s_r, t_1, \cdots, t_n]$, $\mathcal{F} = C^\infty(\mathbb{R}^r)$ and $h^\nu \in \mathbb{R}^r,\ \nu = 1, \cdots, n$. We let $\mathcal{D}$ act on $\mathcal{F}$ via

$$s_\rho \circ y := \partial y / \partial z_\rho \text{ for } \rho = 1, \cdots, r \text{ (differentiation) and}$$
$$(t_\nu \circ y)(z) := y(z - h^\nu) \text{ for } \nu = 1, \cdots, n \text{ (translation).}$$

The resulting equations are called *delay-differential*. In this case the space $\mathcal{F}$ is neither injective nor a cogenerator [10]. However, the state representation theorem below is applicable to all $\mathcal{D}$-modules $\mathcal{F}$ and in particular to this situation. Again, one can replace the C^∞-functions by distributions where the translation operators are also defined.

Remark 3.2 The injective cogenerator property over $F[s]$ of the function module $\mathcal{F}$ is of paramount importance for systems theory since it implies that the correspondence $M \leftrightarrow \mathcal{B} := \mathrm{D}(M)$ is a categorical duality, i.e., in colloquial terms, M and $\mathrm{D}(M)$ contain the same information. In particular, a linear system of partial differential or difference equations with constant coefficients has a solution in $\mathcal{F}$ if and only if the necessary algebraic compatibility or integrability conditions are satisfied. For an engineer, for instance, the primary object is the behavior $\mathcal{B}$ with its trajectories whereas the module M is a mathematical tool to investigate $\mathcal{B}$.

For linear systems with variable coefficients no standard function module is an injective cogenerator, say over the Weyl algebra, except for the space of hyperfunctions in the continuous one-dimensional case [6]. Therefore the *algebraic* or *formal* theory of linear systems of partial differential equations, for instance in [28] and the references quoted there, is developed farther than its analytic counterpart which holds generically only [13, 19, 32]. The injective cogenerator property is explicitly and implicitly used at various places of the present paper.

Let now I be a finite index set and consider the data from above:

$$U \subseteq \mathcal{D}^{1\times I},\ M := \mathcal{D}^{1\times I}/U \text{ and } \mathcal{B} := U^\perp \subseteq \mathcal{F}^I. \tag{37}$$

We choose any *term order* $<$ [16, p. 54] on $I \times \mathbb{N}^r$, i.e., an order satisfying the following conditions:

1. The order is strict or total.
2. For $i \in I$ and $0 \neq \mu \in \mathbb{N}^r$ the relation $(i, 0) < (i, \mu)$ holds.
3. The order is compatible with the additive action of $\mathbb{N}^r$ on $I \times \mathbb{N}^r$, ie.

$$(i, \mu) < (j, \nu) \text{ and } \lambda \in \mathbb{N}^r \ \Rightarrow\ (i, \mu) + \lambda := (i, \mu + \lambda) < (j, \nu) + \lambda.$$

The term order is a *well-order* or artinian, ie. every non-empty subset of $I \times \mathbb{N}^r$ has a least element. Remark that $I \times \mathbb{N}^r$ is also artinian with respect to the cw-order.

Remark 3.3 For all practical purposes, in particular for MAPLE computations, we choose a term order in the following fashion [32, ch.VII, 104], [23, Assumption 2.1, p. 280], [16, Def. 1.4.11, p. 52]. Like MAPLE we use column vectors.

Consider I as a subset of $\mathbb{N}$, choose an additional indeterminate s_0 and identify

$$\alpha = (i, \mu) = \begin{pmatrix} i \\ \mu \end{pmatrix} = s^\mu \delta_i = s_0^i s^\mu = s_0^i * s_1^{\mu_1} \cdot * \cdots * s_r^{\mu_r}, \text{ hence}$$
$$F[s]^{1\times I} = \oplus_{i\in I} F[s]\delta_i = \oplus_{i\in I} F[s]s_0^i \subset F[s_0, s_1, \cdots, s_r]. \tag{38}$$

Choose a matrix

$$U = (U_{-0}U_{II}) = (U_{-0}U_{-1}\cdots U_{-r}) \in \mathbb{R}^{r'\times(1+r)},$$

with the following properties:

(a) The $1 + r$ columns of U are $\mathbb{Z}$-linearly independent so that the $\mathbb{Z}$-linear map

$$U : \mathbb{Z}^{1+r} \to \mathbb{R}^{r'},\ \alpha = \begin{pmatrix} i \\ \mu \end{pmatrix} \mapsto U\alpha = U_{-0}i + U_{II}\mu$$
$$= U_{-0}i + U_{-1}\mu_1 + \cdots + U_{-r}\mu_r, \tag{39}$$

is injective. MAPLE permits the choice of $U \in \mathbb{Z}^{r'\times(1+r)}$ of rank $1 + r$ only.

(b) The column space $\mathbb{R}^{r'}$ with the lexicographic order is a strictly ordered vector space and with respect to this order the columns $U_{-\rho},\ \rho = 1, \cdots, r$, are assumed to be greater than zero or in the positive cone. Then the order

$$\alpha < \beta :\Leftrightarrow\ U\alpha < U\beta \text{ in } \mathbb{R}^{r'}. \tag{40}$$

is a term order on $I \times \mathbb{N}^r$. If also $U_{-0} > 0$ then the order is even a term order on $\mathbb{N}^{1+r}$. In the sequel we assume for computational purposes that the term order on $I \times \mathbb{N}^r$ is given according to (40).

The term order (90) for the construction of the state representation in Section 5 is of this type and can therefore also be handled by MAPLE.

The degree of a non-zero vector $h = (h_i)_{i\in I} = \sum_{i,\mu} h_i(\mu)s^\mu\delta_i \in \mathcal{D}^{1\times I}$ is

$$d := (i(d), \mu(d)) := \max\{(i, \mu) \in I \times \mathbb{N}^r;\ h_i(\mu) \neq 0\}. \tag{41}$$

As usual we define $\deg(0) := -\infty$ and

$$N := \deg(U) := \{\deg(h);\ 0 \neq h \in U\} \text{ and } \Gamma := (I \times \mathbb{N}^r) \setminus N. \tag{42}$$

It is obvious that $N = \deg(U)$ is an order submodule of $I \times \mathbb{N}^r$ in the sense of Section 2. A simple transfinite induction argument implies

$$\mathcal{D}^{1\times I} = \oplus_{\gamma\in\Gamma} Fs^{\mu(\gamma)}\delta_{i(\gamma)} \oplus U \ni f = f_{\mathrm{nf}} + f_U,\ f_{\mathrm{nf}} = \sum_{\gamma\in\Gamma} f_\gamma s^{\mu(\gamma)}\delta_{i(\gamma)}, \tag{43}$$

where $f_{\rm nf}$ is called the *normal form* of f with respect to U and the chosen term order. Let

$$\Gamma = \uplus_{\alpha\in\Delta}(\alpha+\mathbb{N}^{S(\alpha)}) \text{ with}$$
$$D := \min_{\rm cw}(N),\ \Delta\subseteq\Gamma,\ S(\alpha)\subseteq[r] \text{ for } \alpha\in\Delta \tag{44}$$

be the disjoint decomposition according to Corollary 2.3 and define $s^{(\alpha)} := (s_\rho)_{\rho\in S(\alpha)}$ according to (13).

Corollary 3.4 The preceding considerations and data imply the direct sum decompositions

$$\begin{aligned}\mathcal{D}^{1\times I} &= \oplus_{\gamma\in\Gamma} F s^{\mu(\gamma)}\delta_{i(\gamma)} \oplus U\\ &= \oplus_{\alpha\in\Delta,\ \mu\in\mathbb{N}^{S(\alpha)}} F s^{\mu(\alpha)+\mu}\delta_{i(\alpha)} \oplus U = \oplus_{\alpha\in\Delta} F[s^{(\alpha)}]s^{\mu(\alpha)}\delta_{i(\alpha)} \oplus U \text{ and}\\ M = \mathcal{D}^{1\times I}/U &= \oplus_{\alpha\in\Delta,\ \mu\in\mathbb{N}^{S(\alpha)}} F s^{\mu(\alpha)+\mu}\overline{\delta_{i(\alpha)}} = \oplus_{\alpha\in\Delta} F[s^{(\alpha)}]s^{\mu(\alpha)}\overline{\delta_{i(\alpha)}}.\end{aligned}$$

In particular, each $f\in\mathcal{D}^{1\times I}$ admits a unique representation

$$\begin{aligned}f &= f_{\rm nf} + f_U \text{ with the normal form}\\ f_{\rm nf} &= \sum_{\alpha\in\Delta} f_\alpha s^{\mu(\alpha)}\delta_{i(\alpha)} = \sum_{\alpha\in\Delta,\ \mu\in\mathbb{N}^{S(\alpha)}} f_{\alpha,\mu}s^{\mu(\alpha)+\mu}\delta_{i(\alpha)} \text{ where}\\ & f_\alpha\in F[s^{(\alpha)}] \text{ and } f_{\alpha,\mu}\in F.\end{aligned}$$

The normal form of $f\in\mathcal{D}^{1\times I}$ depends on U and on the chosen term order on $I\times\mathbb{N}^r$.

The construction of a basis of M according to Corollary 3.4 is due Macaulay [38, Th. 1.1.1], [16, Th. 1.5.7]. The direct sum decomposition into $F[s^{(\alpha)}]$-modules is called a *Stanley decomposition* [38, Def. 1.4.1]. Its construction is a variant of [34, Th. 5.13], [38, Prop. 1.4.3] and [2]. The decomposition of M also permits to compute a suitable *Hilbert function, polynomial and series* in the following fashion. For $\mu\in\mathbb{N}^r$ let $|\ \mu\ | := \mu_1+\cdots+\mu_r$ and for each $m\in\mathbb{N}$ define

$$\begin{aligned}&\mathcal{D}^{1\times I}_m := \oplus_{(i,\mu)}\{Fs^\mu\delta_i;\ |\ \mu\ |= m\} \subseteq \mathcal{D}^{1\times I}_{\le m} := \oplus_{(i,\mu)}\{Fs^\mu\delta_i;\ |\ \mu\ |\le m\} = \oplus_{l=0}^m\mathcal{D}^{1\times I}_l\\ &U_{\le m} := U\cap\mathcal{D}^{1\times I}_{\le m},\ \ \widehat{M}_m := \mathcal{D}^{1\times I}_{\le m}/U_{\le m}\ni[f]_m := f+U_{\le m},\ \ f\in\mathcal{D}^{1\times I}_{\le m}\\ &\widehat{M} := \oplus_{m=0}^\infty\widehat{M}_m.\end{aligned}$$

The space $\mathcal{D}^{1\times I}_{\le m}$ is that of polynomial vectors of total degree at most m. Let s_0 be an additional indeterminate. The module $\widehat{M}$ is a graded $F[s_0,s]=F[s_0,s_1,\cdots,s_r]$-module with the scalar multiplication

$$s_\rho[f]_m := \begin{cases}[s_\rho f]_{m+1} & \text{if } 1\le\rho\le r\\ [f]_{m+1} & \text{if } \rho=0\end{cases},\quad f\in\mathcal{D}^{1\times I}_{\le m}.$$

For $|\ \mu\ |\le m$ we obtain $[s^\mu\delta_i]_m = s_0^{m-|\mu|}s^\mu[\delta_i]_0$. Hence $\widehat{M}$ is a finitely generated $F[s_0,s]$-module and therefore its *Hilbert function*

$$\mathrm{HF}_{\widehat{M}}:\mathbb{N}\to\mathbb{N},\ m\mapsto\dim_F(\widehat{M}_m)=\dim_F(\mathcal{D}^{1\times I}_{\le m}/U_{\le m}),$$

is a polynomial in m for large m, ie. there is the unique *Hilbert polynomial* $\mathrm{HP}_{\widehat{M}}$ of $\widehat{M}$ such that $\mathrm{HF}_{\widehat{M}}(m) = \mathrm{HP}_{\widehat{M}}(m)$ for almost all $m \in \mathbb{N}$. Let $d(M)$ denote its degree. The module $\widehat{M}$ coincides with the s_0-extension of M in the sense of [42, p. 4–5, Th. 4.2], up to isomorphism, and the numbers $\mathrm{HF}_{\widehat{M}}(m)$ are called the *complexity indices* of the behavior $\mathcal{B} = U^{\perp}$ in [42, Def. 4.1, Th. 4.2]. The map $M \mapsto \widehat{M}$ corresponds to projectivization in geometry.

Result 3.5 The number $d(M)$ is the Krull dimension $\dim(M)$ of M.

It is surprising that this basic result on the Krull dimension of a polynomial *module* is not *explicitly* contained in any of the standard references on Commutative Algebra. It is, however, a consequence of [42, results 7.4–7.6] and the literature quoted there. The graded $F[s_0, s]$-module $\widehat{M}$ gives also rise to its *Hilbert power series* $\mathrm{HS}_{\widehat{M}} := \sum_{m=0}^{\infty} H_{\widehat{M}}(m)t^m \in \mathbb{Z}[[t]] \cap \mathbb{Q}(t)$.

Theorem 3.6 *Data as before.*

1. *The module M is a torsion module if and only if*

$$\max_{\alpha\in\Delta} \mid S(\alpha) \mid < r = \dim(F[s_1, \cdots, s_r]).$$

2. *If the term order is graded then*

$$\dim(M) = \max_{\alpha\in\Delta} \mid S(\alpha) \mid$$

$$\mathrm{HF}_{\widehat{\mathrm{M}}}(\mathrm{m}) = \sum_{\alpha} \left\{ \binom{m- \mid \mu(\alpha) \mid + \mid S(\alpha) \mid}{\mid S(\alpha) \mid};\ \alpha = (i(\alpha), \mu(\alpha)) \in \Delta,\ \mid \mu(\alpha) \mid \leq m \right\}$$

$$\mathrm{HP}_{\widehat{\mathrm{M}}}(\mathrm{m}) = \sum_{\alpha\in\Delta} \binom{m- \mid \mu(\alpha) \mid + \mid S(\alpha) \mid}{\mid S(\alpha) \mid}$$

$$\mathrm{HS}_{\widehat{\mathrm{M}}} = \sum_{\alpha\in\Delta} t^{|\mu(\alpha)|}(1-t)^{-|S(\alpha)|-1}.$$

Here $\mid S \mid$ is the number of elements of the finite set S. The term order is graded if $\mid \mu \mid < \mid \nu \mid$ implies $(i, \mu) < (j, \nu)$ for all i, j.

Proof

1. [42, Th. 7.6]. The assertion also follows from item 2 resp. Result 3.7.
2. If $f \in \mathcal{D}^{1\times I}$ is a non-zero vector of degree $\deg(f) = \alpha = (i, \mu)$ and $\mid \mu \mid \leq m$ then $f \in \mathcal{D}^{1\times I}_{\leq m}$. This follows from the assumed gradedness of the term order. Consider the decomposition

$$\mathcal{D}^{1\times I} = V \oplus U,\ V := \oplus_{\gamma=(i,\mu)\in\Gamma} Fs^{\mu}\delta_i,\ \text{in particular}$$

$$s^{\nu}\delta_j = f + g,\ f := (s^{\nu}\delta_j)_{\mathrm{nf}} \text{ where } s^{\nu}\delta_j \in \mathcal{D}^{1\times I}_{\leq m},\ \mid \nu \mid \leq m, \tag{45}$$

is any basis vector of $\mathcal{D}^{1\times I}_{\le m}$. Since $\deg(f) \in \Gamma \uplus \{-\infty\}$ and $\deg(g) \in \deg(U) \uplus \{-\infty\}$ these degrees are distinct and hence

$$\begin{aligned}\alpha &:= (j,\nu) = \deg(s^{\nu}\delta_j) = \max(\deg(f),\deg(g)),\\ \text{say } \alpha &= \deg(f) > \deg(g) \text{ and hence } f, g \in \mathcal{D}^{1\times I}_{\le m}\\ \mathcal{D}^{1\times I}_{\le m} &= V_{\le m} \oplus U_{\le m} \text{ with } V_{\le m} := V \cap \mathcal{D}^{1\times I}_{\le m} = \oplus_{\gamma=(i,\mu)\in\Gamma,\ |\mu|\le m} F s^{\mu}\delta_i\\ V_{\le m} &\cong \mathcal{D}^{1\times I}_{\le m}/U_{\le m},\ \ \mathrm{HF}_{\widehat{M}}(m) = \dim_F(V_{\le m}).\end{aligned}$$

From

$$\begin{aligned}V &= \oplus_{(i,\mu)\in\Gamma} F s^{\mu}\delta_i = \oplus_{\alpha=(i(\alpha),\mu(\alpha))\in\Delta} F[s^{(\alpha)}] s^{\mu(\alpha)}\delta_{i(\alpha)} \text{ we infer}\\ V_{\le m} &= \oplus_{\alpha\in\Delta,\ |\mu(\alpha)|\le m,}\ \oplus_{\mu\in\mathbb{N}^{S(\alpha)}} \{F s^{\mu(\alpha)+\mu}\delta_{i(\alpha)};\ |\ \mu\ |\le m- |\ \mu(\alpha)\ |\} \text{ and}\\ \mathrm{HF}_{\widehat{M}}(m) &= \sum_{\alpha\in\Delta,\ |\mu(\alpha)|\le m,} |\ \{\mu\in\mathbb{N}^{S(\alpha)};\ |\ \mu\ |\le m- |\ \mu(\alpha)\ |\}\ |\ .\end{aligned}$$

For $s, k \in \mathbb{N}$ the standard formulas

$$|\ \{\mu\in\mathbb{N}^{s};\ |\ \mu\ |\le k\}\ |= \binom{k+s}{s} \text{ and } t^k(1-t)^{-s-1} = \sum_{m\ge k}\binom{m-k+s}{s} t^m$$

hold. For $|\ \mu(\alpha)\ |\le m$ these imply

$$\begin{aligned}&|\ \{\mu\in\mathbb{N}^{S(\alpha)};\ |\ \mu\ |\le m- |\ \mu(\alpha)\ |\}\ |= \binom{m- |\ \mu(\alpha)\ | + |\ S(\alpha)\ |}{|\ S(\alpha)\ |} \text{ and}\\ &\sum_{|\mu(\alpha)|\le m} |\ \{\mu\in\mathbb{N}^{S(\alpha)};\ |\ \mu\ |\le m- |\ \mu(\alpha)\ |\}\ |\ t^m = t^{|\mu(\alpha)|}(1-t)^{-|S(\alpha)|-1}\end{aligned}$$

and thus the expressions for the Hilbert function, polynomial and series. Since $\binom{k+s}{s}$ is a polynomial of degree s in k the Hilbert polynomial has the degree $\dim(M) = d(M) = \max_{\alpha\in\Delta} |\ S(\alpha)\ |$. □

Result 3.7 Let $\mathfrak{m} :=_{F[s]}< s_1, \cdots, s_r >$ denote the unique graded maximal ideal of the graded polynomial ring $F[s] = F[s_1,\cdots,s_r]$, $M_{\mathfrak{m}}$ the localization of M with respect to $\mathfrak{m}$ and $<$ any, not necessarily graded term order on $I\times\mathbb{N}^r$ The other objects are the same as those of the preceding theorem.

1.

$$\dim(M_{\mathfrak{m}}) \le \max_{\alpha\in\Delta} |\ S(\alpha)\ |\le \dim(M).$$

2. If M is itself a graded $F[s]$-module with the Hilbert polynomial $\mathrm{HP_M}$ then

$$\begin{aligned}&\dim(M_{\mathfrak{m}}) = \max_{\alpha\in\Delta} |\ S(\alpha)\ |= \dim(M) = \deg(\mathrm{HP_M}) + 1,\ \text{hence}\\ &\dim(M) = \deg(\mathrm{HP}_{\widehat{\mathrm{M}}}) = \deg(\mathrm{HP_M}) + 1.\end{aligned}$$

If U is a graded submodule of $\mathcal{D}^{1\times I}$ and $M = \mathcal{D}^{1\times I}/U$ carries the induced grading then

$$\widehat{M} = \oplus_{l=0}^{m} M_l \text{ and } \mathrm{HF}_{\widehat{\mathrm{M}}}(\mathrm{m}) = \sum_{\mathrm{l}=0}^{\mathrm{m}} \mathrm{HF_M(l)},\ \ \mathrm{HP}_{\widehat{\mathrm{M}}}(\mathrm{m}) = \sum_{\mathrm{l}=0}^{\mathrm{m}} \mathrm{HP_M(l)}.$$

I have only a somewhat lengthy proof and omit it since the result is not used in the present paper. For system theoretic purposes the Hilbert function $\mathrm{HF}_{\widehat{M}}$ is more interesting than HF_M.

Corollary and Definition 3.8 (Reduced Gröbner basis) [38, Def. 1.1.4],[16, Th. 2.4.13] For every $d = (i(d), \mu(d)) \in D := \min_{\mathrm{cw}}(\deg(U))$ there is a unique vector

$$f_d = s^{\mu(d)}\delta_{i(d)} - \sum_{\alpha\in\Delta} f_{d,\alpha}s^{\mu(\alpha)}\delta_{i(\alpha)}$$

$$= s^{\mu(d)}\delta_{i(d)} - \sum_{\alpha\in\Delta,\ \mu\in\mathbb{N}^{S(\alpha)}} f_{d,\alpha,\mu}s^{\mu(\alpha)+\mu}\delta_{i(\alpha)} \text{ in } U \text{ with}$$

$$\deg(f_d) = d,\ f_{d,\alpha} \in F[s^{(\alpha)}],\ f_{d,\alpha,\mu} \in F. \tag{46}$$

The family $(f_d)_{d\in D}$ of these vectors generates U as $\mathcal{D}$-module and is called the reduced Gröbner basis of U with respect to the given term order on $I \times \mathbb{N}^r$.

Proof The defining Equation (46) has the form

$$s^{\mu(d)}\delta_{i(d)} = \sum_{\alpha\in\Delta} f_{d,\alpha}s^{\mu(\alpha)}\delta_{i(\alpha)} + f_d \in \mathcal{D}^{1\times I} = \oplus_{\alpha\in\Delta}F[s^{(\alpha)}]s^{\mu(\alpha)}\delta_{i(\alpha)} \oplus U$$

$$\text{hence } (s^{\mu(d)}\delta_{i(d)})_{\mathrm{nf}} = \sum_{\alpha\in\Delta} f_{d,\alpha}s^{\mu(\alpha)}\delta_{i(\alpha)}.$$

□

We define the *canonical Gröbner matrix* of U or $\mathcal{B}$ by

$$R \in \mathcal{D}^{D\times I} \text{ with } R_{d-} := f_d; \tag{47}$$

its rows are the reduced Gröbner basis of U.

Corollary and Definition 3.9 (Normal form or canonical representation) The representations

$$U = \sum_{d\in D}\mathcal{D}f_d = \sum_{d\in D}\mathcal{D}(s^{\mu(d)}\delta_{i(d)} - \sum_{\alpha\in\Delta} f_{d,\alpha}s^{\mu(\alpha)}\delta_{i(\alpha)}) = \mathrm{row}(R) \text{ and}$$

$$\mathcal{B} = U^\perp = \{w = (w_i)_{i\in I} \in \mathcal{F}^I;\ \forall d \in D:\ \ f_d \circ w = 0\}$$

$$= \{w \in \mathcal{F}^I;\ \forall d \in D:\ s^{\mu(d)} \circ w_{i(d)} = \sum_{\alpha\in\Delta} f_{d,\alpha}s^{\mu(\alpha)} \circ w_{i(\alpha)}\}$$

$$= \{w \in \mathcal{F}^I;\ R \circ w = 0\}$$

are called the canonical kernel representation or normal form (compare [20, pp. 99–104]) of the submodule U resp. the $\mathcal{F}$-behavior $\mathcal{B}$. They are uniquely determined by U and the chosen term order. If $\mathcal{F}$ is an injective cogenerator the representation is also uniquely determined by $\mathcal{B}$ due to $U = \mathcal{B}^\perp$.

Buchberger's Gröbner basis algorithm (compare [16, Th. 2.5.5], [38, Alg. 1.2.2]) and its implementation in MAPLE permit to compute these normal forms of U or $\mathcal{B}$ from any finite system G of generators of U or from a kernel representation $\mathcal{B} = \{w \in \mathcal{F}^I;\ R' \circ w = 0\}$ with given matrix R'. Canonical normal forms or structure theorems also play a prominent part in one-dimensional systems theory [14, Section 6.7], [41, Th. 4.3.3].

Example 3.10 In the one-dimensional case ($r = 1$) let $R' \in F[s]^{g\times q}$ and $U := \mathrm{row}(R')$. On $\{1, \cdots, q\} \times \mathbb{N}$ there are only two essentially different term orders, viz.

$$(i, \mu) < (j, \nu) \;\Leftrightarrow\; i > j \text{ or } i = j \text{ and } \mu < \nu \tag{48}$$

and

$$(i, \mu) < (j, \nu) \;\Leftrightarrow\; \mu < \nu \text{ or } \mu = \nu \text{ and } i > j. \tag{49}$$

The canonical Gröbner matrix R of R' with respect to (48) is its unique *row Hermite form* [20, p. 102] and can be computed with the MAPLE command LinearAlgebra[HermiteForm] (compare [14, p. 375] for the standard use).

With respect to the graded order (49) one obtains a *canonical row-reduced* form of R' which can be computed with MAPLE's Groebner package, but not with the LinearAlgebra package, and is a variant of the polynomial-row echelon form of R' [14, Ch.6, p. 481].

We show next that every behavior is isomorphic to one without non-essential components w_i *(compare* [41, p. 142, foot-note]). *This is a necessary first reduction step for the construction of the state representation of a behavior in* Section 5.

For this purpose we decompose the index set I and the modules $\mathcal{D}^{1\times I}$ and $\mathcal{F}^I$ as

$$\begin{aligned}
I \;&=\; I^\Gamma \uplus I^N \text{ with}\\
I^\Gamma \;&:= \{i \in I;\;(i,0) \in \Gamma \text{ or } (i,0) \in \Delta\}\\
I^N \;&:= \{i \in I;\;(i,0) \in N \text{ or } (i,0) \in D \text{ or } N_i = \mathbb{N}^r\}\\
\mathcal{D}^{1\times I^\Gamma} \;&:= \mathcal{D}^{1\times I^\Gamma} \times \{0\} \subset \mathcal{D}^{1\times I} = \mathcal{D}^{1\times I^\Gamma} \times \mathcal{D}^{1\times I^N} \ni f = (f^\Gamma, f^N)\\
\mathcal{F}^{I^\Gamma} \;&=\; \mathcal{F}^{I^\Gamma} \times \{0\} \subset \mathcal{F}^I = \mathcal{F}^{I^\Gamma} \times \mathcal{F}^{I^N} \ni w = \begin{pmatrix} w^\Gamma \\ w^N \end{pmatrix}
\end{aligned} \tag{50}$$

where $N_i = \{\mu \in \mathbb{N}^r;\;(i,\mu) \in N\}$.

Lemma 3.11 If $i \in I$ and

$$\begin{cases} d = (i, \mu(d)) \in D,\;\; \mu(d) \neq 0,\;\text{ or} \\ \alpha = (i, \mu(\alpha)) \in \Delta \end{cases} \qquad \text{then } i \in I^\Gamma.$$

Proof Assume that $i \in I^N$ and hence $(i, 0) \in D \subset N$.

In the first case $(i, 0) \leq_{\mathrm{cw}} d = (i, \mu(d))$ and the cw-discreteness of D would imply $(i, 0) = d$ and $\mu(d) = 0$, a contradiction.

In the second case $\alpha = (i, 0) + \mu(\alpha)$ and $N = D + \mathbb{N}^r$ would imply $\alpha \in N$, a contradiction to $\Delta \cap N = \emptyset$. □

We define

$$\begin{aligned} D^{\Gamma} &:= D \setminus \{(i,0);\ i \in I^N\},\ U^{\Gamma} := \sum_{d \in D^{\Gamma}} \mathcal{D} f_d \subset U \cap \mathcal{D}^{1\times I^{\Gamma}} \\ M^{\Gamma} &= \mathcal{D}^{1\times I^{\Gamma}}/U^{\Gamma},\ \mathcal{B}^{\Gamma} := (U^{\Gamma})^{\perp} \\ &= \{w^{\Gamma} \in \mathcal{F}^{I^{\Gamma}};\ \forall d \in D^{\Gamma} : f_d \circ w^{\Gamma} = 0\} \\ &= \{w^{\Gamma} = (w_i)_{i\in I^{\Gamma}} \in \mathcal{F}^{I^{\Gamma}};\ \forall d \in D^{\Gamma} :\ s^{\mu(d)} \circ w_{i(d)} = \sum_{\alpha\in\Delta} f_{d,\alpha} s^{\mu(\alpha)} \circ w_{i(\alpha)}\} \end{aligned} \tag{51}$$

The inclusion $U^{\Gamma} \subset \mathcal{D}^{1\times I^{\Gamma}}$ follows from the preceding lemma. Let id denote the identity matrix of suitable size.

Theorem 3.12 *Under the preceding assumptions the canonical injection*

$$\mathrm{inj} := \circ(\mathrm{id}\, 0) : \mathcal{D}^{1\times I^{\Gamma}} \to \mathcal{D}^{1\times I},\ \delta_i \mapsto \delta_i \textit{ for } i \in I^{\Gamma},$$

and the linear map

$$\Psi : \mathcal{D}^{1\times I} \to \mathcal{D}^{1\times I^{\Gamma}},\ \delta_i \mapsto \begin{cases} \delta_i & \textit{for } i \in I^{\Gamma} \\ \sum_{\alpha\in\Delta} f_{(i,0),\alpha} s^{\mu(\alpha)} \delta_{i(\alpha)} & \textit{for } i \in I^N \end{cases}$$

with $\Psi \circ \mathrm{inj} = \mathrm{id}_{\mathcal{D}^{1\times I^{\Gamma}}}$ *induce mutually inverse* $\mathcal{D}$*-isomorphisms*

$$M^{\Gamma} = \mathcal{D}^{1\times I^{\Gamma}}/U^{\Gamma} \underset{\Psi_{\mathrm{ind}}}{\overset{\mathrm{inj}_{\mathrm{ind}}}{\rightleftarrows}} M = \mathcal{D}^{1\times I}/U \tag{52}$$

and inverse behavior isomorphisms

$$\begin{aligned} &\mathcal{B}^{\Gamma} \underset{(\mathrm{id}\,0)\circ}{\overset{\Psi_{\mathrm{ind}}}{\rightleftarrows}} \mathcal{B},\ w^{\Gamma} \leftrightarrow w, \\ &w_i^{\Gamma} = w_i \textit{ for } i \in I^{\Gamma},\ w_i = \sum_{\alpha\in\Delta} f_{(i,0),\alpha} s^{\mu(\alpha)} \circ w_{i(\alpha)}^{\Gamma} \textit{ for } i \in I^N. \end{aligned} \tag{53}$$

Proof Since $U^{\Gamma} \subset U$ according to (51) the injection $\circ(\mathrm{id}\,0)$ induces a map from M^{Γ} to M.

On the other hand, Ψ maps $U = \sum_{d\in D} \mathcal{D} f_d$ into $U^{\Gamma} = \sum_{d\in D^{\Gamma}} \mathcal{D} f_d$ and thus induces a map $M \to M^{\Gamma}$: Indeed, if $d \in D^{\Gamma}$ then $f_d \in \mathcal{D}^{1\times I^{\Gamma}}$ and hence $\Psi(f_d) = f_d \in U^{\Gamma}$. If $d = (i,0),\ i \in I^N$ then

$$\begin{aligned} &f_{(i,0)} = \delta_i - \sum_{\alpha\in\Delta} f_{(i,0),\alpha} s^{\mu(\alpha)} \delta_{i(\alpha)} \text{ and} \\ &\Psi(f_{(i,0)}) = \sum_{\alpha\in\Delta} f_{(i,0),\alpha} s^{\mu(\alpha)} \delta_{i(\alpha)} - \sum_{\alpha\in\Delta} f_{(i,0),\alpha} s^{\mu(\alpha)} \delta_{i(\alpha)} = 0 \in U^{\Gamma}. \end{aligned}$$

The relation $\Psi \circ \mathrm{inj} = \mathrm{id}_{\mathcal{D}^{1\times I^\Gamma}}$ implies $\Psi_{\mathrm{ind}} \circ \mathrm{inj}_{\mathrm{ind}} = \mathrm{id}_{M^\Gamma}$ and thus that $\mathrm{inj}_{\mathrm{ind}}$ is a monomorphism. But

$$0 = \overline{f_{(i,0)}} = \overline{\delta_i} - \sum_{\alpha\in\Delta} f_{(i,0),\alpha} s^{\mu(\alpha)} \overline{\delta_{i(\alpha)}} \in M \text{ for } i \in I^N \text{ implies}$$

$$\overline{\delta_i} \in \sum_{j\in I^\Gamma} \mathcal{D}\overline{\delta_j} \text{ and hence } M = \sum_{j\in I^\Gamma} \mathcal{D}\overline{\delta_j}.$$

This means that $\mathrm{inj}_{\mathrm{ind}}$ is surjective too, hence an isomorphism and of course inverse to Ψ_{ind} as asserted.

The duality functor $D = \mathrm{Hom}_{\mathcal{D}}(-, \mathcal{F})$ transforms the preceding isomorphisms into the given behavior isomorphisms. □

Definition 3.13 We call the module $M = \mathcal{D}^{1\times I}/U$ and the behavior $\mathcal{B}$ minimally embedded if $I^N = \emptyset$, ie. if $N_i \neq \mathbb{N}^r$ or $(i, 0) \in \Delta$ for all $i \in I$.

We furnish $I^\Gamma \times \mathbb{N}^r$ with the induced term order of that on $I \times \mathbb{N}^r$.

Corollary 3.14 The behavior $\mathcal{B}^\Gamma$ of the preceding theorem is minimally embedded. Hence an arbitrary behavior can always be replaced by an isomorphic minimally embedded one. This is the first reduction step which is used for the construction of the canonical state representation of a behavior in Section 5.

Proof From $U^\Gamma \subset U$ we infer $\deg(U^\Gamma) \subset (I^\Gamma \times \mathbb{N}^r) \cap N$. This inclusion and $(i, 0) \in \deg(U^\Gamma),\ i \in I^\Gamma$, would imply $i \in I^N$ in contradiction to $I^\Gamma \cap I^N = \emptyset$. □

Corollary 3.15 In the situation of Theorem 3.12 assume that a term order on $\mathbb{N}^r$ and a strict order of I are given and that $I \times \mathbb{N}^r$ carries the lexicographic term order

$$(i, \mu) < (j, \nu) \ \Leftrightarrow\ \mu < \nu \text{ or } \mu = \nu \text{ and } i < j.$$

Then the vectors $f_{(i,0)},\ i \in I^N$, of the reduced Gröbner basis have the simple form

$$f_{(i,0)} = \delta_i - \sum_{j\in I^\Gamma} f_{i,j}\delta_j,\ \ f_{i,j} \in F,\ \text{for all } i \in I^N, \tag{54}$$

and the behavior isomorphisms are given as

$$\mathcal{B}^\Gamma \cong \mathcal{B},\ w^\Gamma \leftrightarrow w,\ \text{where}$$

$$w_i = \begin{cases} w_i^\Gamma & \text{if } i \in I^\Gamma \\ \sum_{j\in I^\Gamma} f_{i,j} w_j & \text{if } i \in I^N,\ \ f_{i,j} \in F. \end{cases}$$

Hence the $w_i,\ i \in I^N$, are F-linear combinations of the $w_j = w_j^\Gamma,\ j \in I^\Gamma$, and obviously non-essential in the description of the behavior.

Proof The vector $f_{(i,0)},\ i \in I^N$ has the form

$$f_{(i,0)} = \delta_i - \sum_{\alpha\in\Delta,\ \mu\in\mathbb{N}^{S(\alpha)}} f_{(i,0),\alpha,\mu} s^{\mu(\alpha)+\mu}\delta_{i(\alpha)},\ f_{(i,0),\alpha,\mu} \in F,\ \text{with}$$

$$(i(\alpha), \mu(\alpha)+\mu) < (i,0) \text{ if } f_{(i,0),\alpha,\mu} \neq 0.$$

The chosen term order implies $\mu(\alpha)+\mu = 0$ for $f_{(i,0),\alpha,\mu} \neq 0$, and hence the representation (54). □

4 The Canonical and Normal Input/Output-Form of a Behavior

An input/output (IO) *partition* [25, Def. 3.3.1] or *structure* [20, Th. 2.69] is an important additional attribute of a behavior

$$\mathcal{B} := \{w \in \mathcal{F}^I;\ R' \circ w = 0\},\ R' \in \mathcal{D}^{\bullet\times I}.$$

In this section we show that any chosen term order on $I \times \mathbb{N}^r$ gives rise to a canonical *IO-structure of $\mathcal{B}$ (Theorem 4.3) and to a* normal *IO-representation $\mathcal{B}_{\text{nf}}$ (Theorem 4.5) and that, conversely, each such IO-structure is the canonical one for a suitable term order. This last result is important since, on the one hand, a special IO-structure is often implied by the modeling process and since, on the other hand, the state representation of $\mathcal{B}$ established in the last section of this paper depends on the choice of a term order. In other terms, the latter can be adapted to the modeling needs.*

The notations from the last section remain in force. At first we recall the notion of an IO-structure. We assume that

$$U = \text{row}(R') = \text{row}(R),\ M := \mathcal{D}^{1\times I}/U \text{ and } \mathcal{B} := U^{\perp}$$

where $R' \in \mathcal{D}^{\bullet\times I}$ is any matrix, whereas $R \in \mathcal{D}^{D\times I}$ is the canonical one from (47) whose rows are the reduced Gröbner basis of U. We do not assume that the behavior is minimally embedded.

Let $I = J^0 \uplus J^{\text{free}}$ be any disjoint decomposition of I. As usual it induces the corresponding direct product decompositions

$$\mathcal{D}^{1\times I} = \mathcal{D}^{1\times J^0} \times \mathcal{D}^{1\times J^{\text{free}}},\ \mathcal{F}^I = \mathcal{F}^{J^0} \times \mathcal{F}^{J^{\text{free}}} \ni w = \begin{pmatrix} y \\ u \end{pmatrix},$$

$$R' = (P', -Q') \in \mathcal{D}^{\bullet\times I} = \mathcal{D}^{\bullet\times J^0} \times \mathcal{D}^{\bullet\times J^{\text{free}}},\ \text{hence}$$

$$\mathcal{B} = \{w = \begin{pmatrix} y \\ u \end{pmatrix} \in \mathcal{F}^I;\ P' \circ y = Q' \circ u\}. \tag{55}$$

The choice of the letters y, u, R, P, Q in this context is customary in systems theory [25, p. 82]. Recall the ranks

$$m := \text{rank}(M) := \text{rank}(\mathcal{B}) := \dim_{F(s)}(F(s) \otimes_{F[s]} M) \text{ and}$$

$$p := \text{rank}(R') = | I | - \text{rank}(M) \tag{56}$$

which were called the input- resp. the output-dimension of the behavior $\mathcal{B}$ in [20, p. 37] and are customarily denoted with the letters m and p since Kalman's classical state systems. Consider the exact sequence

$$0 \longrightarrow \mathcal{D}^{1\times J^{\text{free}}} \xrightarrow{\text{inj}} \mathcal{D}^{1\times I} = \mathcal{D}^{1\times J^0} \times \mathcal{D}^{1\times J^{\text{free}}} \xrightarrow{\text{proj}} \mathcal{D}^{1\times J^0} \longrightarrow 0$$
$$\text{with } \text{inj} = \circ(0\,\text{id}) : \mathcal{D}^{1\times J^{\text{free}}} \to \mathcal{D}^{1\times I} \text{ and}$$
$$\text{proj} = \text{can} := \circ\begin{pmatrix}\text{id}\\ 0\end{pmatrix} : \mathcal{D}^{1\times I} \to \mathcal{D}^{1\times J^0} \underset{identification}{=} \mathcal{D}^{1\times I}/\mathcal{D}^{1\times J^{\text{free}}} \tag{57}$$

where id denotes the identity matrix of suitable size, and also the modules

$$U^0 := \text{proj}(U) = \text{row}(P') \in \mathcal{D}^{1\times J^0},\ M^0 := \mathcal{D}^{1\times J^0}/U^0$$
$$\text{with the associated behavior } \mathcal{B}^0 := (U^0)^\perp \subset \mathcal{F}^{J^0} \text{ and}$$
$$\text{proj}_{\text{ind}} = \circ\begin{pmatrix}\text{id}\\ 0\end{pmatrix} : M = \mathcal{D}^{1\times I}/U \to M^0 = \mathcal{D}^{1\times J^0}/U^0. \tag{58}$$

Lemma and Definition 4.1 (IO-structure and transfer matrix [20, Th. 65, Th. 94]) The following assertions are equivalent:

1. $\text{rank}(R') = \text{rank}(P') = | J^0 |$.
2. The module M^0 is a torsion module and the sequence
$$0 \longrightarrow \mathcal{D}^{1\times J^{\text{free}}} \xrightarrow{\text{inj}_{\text{ind}}} M \xrightarrow{\text{proj}_{\text{ind}}} M^0 \longrightarrow 0$$
is exact.
3. The map
$$F(s) \otimes \text{inj} : F(s)^{1\times J^{\text{free}}} \cong F(s) \otimes_{F[s]} M$$
is an $F(s)$-isomorphism.
4. If $\mathcal{F}$ is an injective cogenerator: The behavior sequence
$$0 \longrightarrow \mathcal{B}^0 \xrightarrow{\binom{\text{id}}{0}\circ} \mathcal{B} \xrightarrow{(0\,\text{id})\circ} \mathcal{F}^{J^{\text{free}}} \longrightarrow 0$$
$$y \mapsto \begin{pmatrix}y\\ 0\end{pmatrix},\ w = \begin{pmatrix}y\\ u\end{pmatrix} \mapsto u$$
is exact and $\mathcal{B}^0$ is autonomous.

If these conditions are satisfied the decomposition $I = J^0 \uplus J^{\text{free}}$ is called an *input/output- or IO-structure* of M or $\mathcal{B}$ and the components y resp. u of $w = \binom{y}{u}$ are called the *output* resp. the *input* of $\mathcal{B}$ with respect to this IO-structure. Moreover there is a unique rational matrix $H \in F(s)^{J^0\times J^{\text{free}}}$ such that $P'H = Q'$, and $\mathcal{D}^{1\times J^{\text{free}}}$ is a free submodule of M of maximal dimension, up to isomorphism. The matrix H is called the *transfer matrix* of $\mathcal{B}$ with respect to this IO-structure and depends on $\mathcal{B}$ and the IO-structure only.

Recall that a module M over the integral domain $\mathcal{D} = F[s]$ is called a *torsion module* if for each $x \in M$ there is an $f \neq 0$ such that $fx = 0$ or, in other terms, the annihilator ideal $\mathrm{ann}_{\mathcal{D}}(x) := \{f \in \mathcal{D};\ \ fx = 0\}$ is not zero. If M is finitely generated as always in this paper this is also equivalent with $\mathrm{ann}_{\mathcal{D}}(M) := \{f \in \mathcal{D};\ \ fM = 0\} \neq 0$. If $M = \mathcal{D}^{1\times I}/U$ is a torsion module the behavior $\mathcal{B}{=}U^{\perp}$ is called *autonomous*.

With these preparations we are now going to construct the *canonical* or naturally defined IO-structure of $M = \mathcal{D}^{1\times I}/U$ and $\mathcal{B} = U^{\perp}$ *induced from the given term order* on $I \times \mathbb{N}^r$. With the data from (25) define

$$I = I^0 \uplus I^{\text{free}} \text{ with } I^{\text{free}} := \{i \in I;\ (i,0) \in \Delta,\ S(i,0) = [r] := \{1, \cdots, r\}\}. \tag{59}$$

Lemma and Definition 4.2 The preceding decomposition of I induces the decomposition $\Delta = \Delta^0 \uplus \Delta^{\text{free}}$ with

$$\Delta^{\text{free}} := \{\alpha \in \Delta;\ S(\alpha) = [r]\} = \{(i,0); i \in I^{\text{free}}\}$$
$$\Delta^0 := \{\alpha \in \Delta;\ S(\alpha) \neq [r]\} = \{\alpha = (i(\alpha), \mu(\alpha)) \in \Delta;\ i(\alpha) \in I^0\}.$$

If $\alpha = (i, \mu) \in N = \deg(U)$ then also $i \in I^0$.

Proof Let $\alpha = (i, \mu(\alpha)) \in \Delta(i) \subset \Delta$.

1. The condition $S(\alpha) = [r]$ and the construction of these data according to (25) imply that $\deg(U)_i = \emptyset$, $\Delta(i) = \{(i,0)\}$ and $S(i,0) = [r]$, hence $i \in I^{\text{free}}$.
2. Assume that $\alpha \in \Delta^0$, but $i \in I^{\text{free}}$. This would imply
$$(i,0) \in \Delta,\ S(i,0) = [r] \text{ and } \alpha \in (i,0) + \mathbb{N}^{S(i,0)}.$$
From the disjoint decomposition $\Gamma = \uplus_{\beta\in\Delta}(\beta + \mathbb{N}^{S(\beta)})$ we then infer $(i,0) = \alpha$ and $S(\alpha) = S(i,0) = [r]$, a contradiction to $S(\alpha) \neq [r]$ by definition of Δ^0.
3. Assume $\alpha = (i, \mu) \in N$ and $i \in I^{\text{free}}$. As in 2. we obtain
$$(i,0) \in \Delta,\ S(i,0) = [r] \text{ and } \alpha \in (i,0) + \mathbb{N}^{S(i,0)} \subset \Gamma := (I \times \mathbb{N}^r) \setminus N,$$
a contradiction to $\alpha \in N$. □

Let $J^0 := I^0$ and $J^{\text{free}} := I^{\text{free}}$ in equations (55) to (58). By the preceding lemma the canonical decompositions of the matrix R from (47) and of $\mathcal{B}$ are

$$\begin{aligned}
R &= (P, -Q) \in \mathcal{D}^{D\times I} = \mathcal{D}^{D\times I^0} \times \mathcal{D}^{D\times I^{\text{free}}} \text{ with} \\
P_{d-} &= s^{\mu(d)}\delta_{i(d)} - \sum_{\alpha\in\Delta^0} f_{d,\alpha}s^{\mu(\alpha)}\delta_{i(\alpha)},\ \ f_{d,\alpha} \in F[s^{(\alpha)}],\ \ S(\alpha) \neq [r] \text{ and} \\
Q_{d-} &= \sum_{i\in I^{\text{free}}} f_{d,(i,0)}\delta_i,\ \ f_{d,(i,0)} \in F[s] \\
M &= \mathcal{D}^{1\times I}/U = \mathcal{D}^{1\times I}/\operatorname{row}(P, -Q) \\
\mathcal{B} &:= \left\{\begin{pmatrix} y \\ u \end{pmatrix} \in \mathcal{F}^I = \mathcal{F}^{I^0} \times \mathcal{F}^{I^{\text{free}}};\ \ P \circ y = Q \circ u\right\} = \\
&= \left\{\begin{pmatrix} y \\ u \end{pmatrix} \in \mathcal{F}^I;\ \forall d \in D : s^{\mu(d)} \circ y_{i(d)} - \sum_{\alpha\in\Delta^0} f_{d,\alpha}s^{\mu(\alpha)} \circ y_{i(\alpha)} = \sum_{i\in I^{\text{free}}} f_{d,(i,0)} \circ u_i\right\}
\end{aligned} \tag{60}$$

Further we need the modules and the behavior

$$U^0 := \mathrm{proj}(\mathrm{row}(R)) = \mathrm{row}(P),\ \ M^0 := \mathcal{D}^{1\times I^0}/U^0 \text{ and}$$

$$\mathcal{B}^0 := (U^0)^\perp = \{y \in \mathcal{F}^{I^0};\ \ P \circ y = 0\}$$

$$= \left\{ y \in \mathcal{F}^{I^0};\ \forall d \in D:\ s^{\mu(d)} \circ y_{i(d)} - \sum_{\alpha\in\Delta^0} f_{d,\alpha} s^{\mu(\alpha)} \circ y_{i(\alpha)} = 0 \right\} \tag{61}$$

Theorem 4.3 (Canonical IO-structure) *Data as introduced in* (60) *and* (61)*. The decomposition* $I = I^0 \uplus I^{\mathrm{free}}$ *is an IO-structure of* M *or* $\mathcal{B}$ *and called* canonical *with respect to the given term order on* $I \times \mathbb{N}^r$*.*

Proof From Corollary 3.4 we obtain the direct sum decompositions

$$\mathcal{D}^{1\times I} = \oplus_{\alpha\in\Delta} F[s^{(\alpha)}]s^{\mu(\alpha)}\delta_{i(\alpha)} \oplus U \text{ and}$$

$$M = \mathcal{D}^{1\times I}/U = \oplus_{\alpha\in\Delta} F[s^{(\alpha)}]s^{\mu(\alpha)}\overline{\delta_{i(\alpha)}}.$$

With $\Delta = \Delta^0 \uplus \Delta^{\mathrm{free}}$ from Lemma 4.2 and $\mathcal{D}^{1\times I^{\mathrm{free}}} = \oplus_{i\in I^{\mathrm{free}}} F[s]\delta_i$ we infer

$$\begin{aligned}\mathcal{D}^{1\times I} &= \oplus_{\alpha\in\Delta^0} F[s^{(\alpha)}]s^{\mu(\alpha)}\delta_{i(\alpha)} \oplus \oplus_{i\in I^{\mathrm{free}}} F[s]\delta_i \oplus U\\ &= \oplus_{\alpha\in\Delta^0} F[s^{(\alpha)}]s^{\mu(\alpha)}\delta_{i(\alpha)} \oplus \mathcal{D}^{1\times I^{\mathrm{free}}} \oplus U\\ &= \oplus_{\alpha\in\Delta^0} F[s^{(\alpha)}]s^{\mu(\alpha)}\delta_{i(\alpha)} \oplus V \text{ with } V := \mathcal{D}^{1\times I^{\mathrm{free}}} \oplus U.\end{aligned}$$

There results the exact sequence

$$0 \longrightarrow \mathcal{D}^{1\times I^{\mathrm{free}}} \overset{\circ(0\,\mathrm{id})_{\mathrm{ind}}}{\longrightarrow} M = \mathcal{D}^{1\times I}/U \overset{\mathrm{can}}{\longrightarrow} \mathcal{D}^{1\times I}/V \longrightarrow 0 \text{ with}$$

$$\begin{aligned}M = \mathcal{D}^{1\times I}/U &= \oplus_{\alpha\in\Delta} F[s^{(\alpha)}]s^{\mu(\alpha)}\overline{\delta_{i(\alpha)}}\\ &= \oplus_{\alpha\in\Delta^0} F[s^{(\alpha)}]s^{\mu(\alpha)}\overline{\delta_{i(\alpha)}} \oplus \oplus_{i\in I^{\mathrm{free}}} F[s]\overline{\delta_i} \text{ and}\end{aligned}$$

$$\mathcal{D}^{1\times I}/V = \oplus_{\alpha\in\Delta^0} F[s^{(\alpha)}]s^{\mu(\alpha)}\overline{\delta_{i(\alpha)}} \tag{62}$$

Moreover the projection $\mathrm{proj} = \mathrm{can} : \mathcal{D}^{1\times I} \to \mathcal{D}^{1\times I^0}$ from (57) maps $V = \mathcal{D}^{1\times I^{\mathrm{free}}} \oplus U$ into $U^0 = \mathrm{proj}(U)$ and induces the epimorphism

$$\mathrm{proj}_{\mathrm{ind}} : \mathcal{D}^{1\times I}/V \to M^0 = \mathcal{D}^{1\times I^0}/U^0.$$

But this latter epimorphism is actually an isomorphism. Indeed, the relations

$$U^0 = \sum_{d\in D} \mathcal{D}P_{d-} \text{ and } R_{d-} = (P_{d-}, -Q_{d-}) \underset{identification}{=} P_{d-} - Q_{d-} \text{ imply}$$

$$P_{d-} = Q_{d-} + R_{d-} = \sum_{i\in I^{\mathrm{free}}} f_{d,(i,0)}\delta_i + R_{d-} \in \mathcal{D}^{1\times I^{\mathrm{free}}} \oplus U = V.$$

Therefore the injection

$$\circ(\mathrm{id}\,0) : \mathcal{D}^{1\times I^0} \to \mathcal{D}^{1\times I},\ \delta_i \mapsto \delta_i \text{ for } i \in I^0,$$

with $\mathrm{proj} \circ \mathrm{inj} = \mathrm{id}_{\mathcal{D}^{1\times I^0}}$ maps U^0 into V and hence induces a map

$$\mathrm{inj}_{\mathrm{ind}} = \circ(\mathrm{id}\,0)_{\mathrm{ind}} : M^0 = \mathcal{D}^{1\times I^0}/U^0 \to \mathcal{D}^{1\times I}/V \text{ with } \mathrm{proj}_{\mathrm{ind}} \circ \mathrm{inj}_{\mathrm{ind}} = \mathrm{id}_{M^0}.$$

In particular, $\mathrm{inj}_{\mathrm{ind}}$ is a monomorphism. Additionally, the relations $\mathcal{D}^{1\times I}/V = \sum_{i\in I} \mathcal{D}\overline{\delta_i}$ and $\delta_i \in V$ for $i \in I^{\mathrm{free}}$ imply $\mathcal{D}^{1\times I}/V = \sum_{i\in I^0} \mathcal{D}\overline{\delta_i}$ and that $\mathrm{inj}_{\mathrm{ind}}$ is also surjective. Thus $\mathrm{inj}_{\mathrm{ind}}$ and $\mathrm{proj}_{\mathrm{ind}}$ are isomorphisms, inverse of each other.

Hence the exact sequence (62) gives rise to the exact sequence

$$0 \longrightarrow \mathcal{D}^{1\times I^{\mathrm{free}}} \overset{\mathrm{inj}_{\mathrm{ind}}}{\longrightarrow} M = \mathcal{D}^{1\times I}/U \overset{\mathrm{proj}_{\mathrm{ind}}}{\longrightarrow} M^0 = \mathcal{D}^{1\times I^0}/U^0 \longrightarrow 0. \tag{63}$$

But $M^0 = \oplus_{\alpha\in\Delta^0} F[s^{(\alpha)}]s^{\mu(\alpha)}\overline{\delta_{i(\alpha)}}$ is a torsion module according to Theorem 3.6. We have thus completed the proof that $I = I^0 \uplus I^{\mathrm{free}}$ is an IO-structure of M or $\mathcal{B}$. □

The next theorem shows that any IO-structure is a canonical one with respect to a suitably chosen term order. So assume that with the notations from (55)–(58) $I := J^0 \uplus J^{\mathrm{free}}$ is an IO-structure of $U \subset \mathcal{D}^{1\times I}$ according to Lemma 4.1. Let $<$ be an ‘elimination’ term order on $I \times \mathbb{N}^r$ with the property that

$$(i,\mu), (j,\nu) \in I \times \mathbb{N}^r \text{ and } i \in J^{\mathrm{free}},\ j \in J^0 \text{ implies } (i,\mu) < (j,\nu). \tag{64}$$

Such term orders can obviously be defined in various ways.

Theorem 4.4 *If $I := I^0 \uplus I^{\mathrm{free}}$ is the canonical IO-structure of U with respect to a term order with property (64) then*

$$J^0 = I^0 \textit{ and } J^{\mathrm{free}} = I^{\mathrm{free}}.$$

Hence any IO-structure of M or $\mathcal{B}$ is the canonical one with respect to a suitable term order.

Proof We use the data of Lemma 4.1. The elements of $U = \mathrm{row}(R')$ have the form

$$\begin{aligned}(\xi,\eta) &= \zeta R' = \zeta(P', -Q') = (\zeta P', -\zeta P' H)\\ &= (\xi, -\xi H) \in \mathcal{D}^{1\times I} = \mathcal{D}^{1\times J^0} \times \mathcal{D}^{1\times J^{\mathrm{free}}},\ \text{hence } (\xi,\eta) \neq 0 \ \Leftrightarrow\ \xi \neq 0.\end{aligned}$$

Due to (64) this implies $\deg(\xi,\eta) = \deg(\xi) \in J^0 \times \mathbb{N}^r$ for each non-zero vector in U and hence $N := \deg(U) \subset J^0 \times \mathbb{N}^r$.

We show that $I^{\mathrm{free}} \subseteq J^{\mathrm{free}}$: Assume that $i \in I^{\mathrm{free}} \cap J^0$. By definition of I^{free} in (59) this implies $(i,0) + \mathbb{N}^r \subset \Gamma$.

On the other hand, since $i \in J^0$ and M^0 is a torsion module there is $0 \neq f \in \mathcal{D}$ such that $f\overline{\delta_i} = 0$ in M^0 or

$$0 \neq \xi := f\delta_i \in U^0 = \mathrm{row}(P') \subset \mathcal{D}^{1\times J^0}.$$

But then again

$$(\xi, -\xi H) \in U \text{ and by (64) } (i,\mu) := \deg(f\delta_i) = \deg(\xi, -\xi H) \in N = \deg(U)$$

in contradiction to $(i,\mu) \in (i,0) + \mathbb{N}^r \subset \Gamma$.

Since both $I = J^0 \uplus J^{\text{free}}$ and $I = I^0 \uplus I^{\text{free}}$ are IO-structures of U by assumption resp. by Theorem 4.3 the Lemma 4.1 implies

$$| I^0 | = | J^0 |, \text{ hence } | I^{\text{free}} | = | J^{\text{free}} | .$$

The just proven inclusion $I^{\text{free}} \subset J^{\text{free}}$ finally implies the asserted identity of the IO-structures (I^0, I^{free}) and (J^0, J^{free}). □

The canonical IO-representation of a behavior according to Theorem 4.3 and (60) can be further simplified if one replaces M or $\mathcal{B}$ by an isomorphic module or behavior. Recall that D is the set of row indices of R, P and Q so that their columns are contained in $\mathcal{D}^D$. Choose a term order on $D \times \mathbb{N}^r$, consider the column module $U^D := \text{col}(P) \subset \mathcal{D}^D$ and define

$$N^D := \deg(U^D) \text{ and } \Gamma^D := (D \times \mathbb{N}^r) \setminus N^D. \tag{65}$$

From Corollary 3.4 we derive the direct sum decomposition

$$\begin{aligned} \mathcal{D}^D &= \oplus_{\gamma=(d,\mu)\in\Gamma^D} F s^\mu \delta_d \oplus U^D \ni h = h_{\text{nf}} + P x_h \\ \text{with } h_{\text{nf}} &= \sum_{\gamma=(d,\mu)\in\Gamma^D} h_{d,\mu} s^\mu \delta_d, \; h_{d,\mu} \in F, \; x_h \in \mathcal{D}^{I^0} \end{aligned} \tag{66}$$

where h_{nf} is called the normal form of h with respect to P and the chosen term order on $D \times \mathbb{N}^r$. Since $\text{rank}(P) = | I^0 |$ the columns of P are linearly independent over $F[s]$ and $F(s)$ and therefore the coordinate column x_h is uniquely determined. Applying these considerations to all columns of an arbitrary matrix $A \in \mathcal{D}^{D\times\bullet}$ furnishes a unique product decomposition

$$A = P X_A + A_{\text{nf}} \text{ with } X_A \in \mathcal{D}^{I^0\times\bullet} \text{ and } A_{\text{nf}} \in \mathcal{D}^{D\times\bullet}. \tag{67}$$

This is the generalization of euclidean division from polynomials in one indeterminate to matrices of polynomials in several indeterminates. The preceding division especially furnishes

$$Q = P X_Q + Q_{\text{nf}} \in \mathcal{D}^{D\times I^{\text{free}}} \text{ with } X_Q \in \mathcal{D}^{I^0\times I^{\text{free}}} \text{ and } Q_{\text{nf}} \in \mathcal{D}^{D\times I^{\text{free}}}. \tag{68}$$

Theorem and Definition 4.5 *(Normal IO-behavior) Under the assumptions of Theorem 4.3 and with the preceding data define the* normal IO-behavior

$$\mathcal{B}_{\text{nf}} := \left\{ w = \begin{pmatrix} y \\ u \end{pmatrix} \in \mathcal{F}^I = \mathcal{F}^{I^0} \times \mathcal{F}^{I^{\text{free}}}; \; P \circ y = Q_{\text{nf}} \circ u \right\}.$$

Then the matrix $\begin{pmatrix} \text{id} & -X_Q \\ 0 & \text{id} \end{pmatrix} \in \mathcal{D}^{(I^0\uplus I^{\text{free}})\times(I^0\uplus I^{\text{free}})}$ *is invertible with inverse* $\begin{pmatrix} \text{id} & X_Q \\ 0 & \text{id} \end{pmatrix}$ *and induces the mutually inverse behavior isomorphisms*

$$\begin{pmatrix} \text{id} & -X_Q \\ 0 & \text{id} \end{pmatrix} : \mathcal{B} \cong \mathcal{B}_{\text{nf}}, \; \begin{pmatrix} y \\ u \end{pmatrix} \mapsto \begin{pmatrix} y - X_Q \circ u \\ u \end{pmatrix} \text{ and}$$

$$\begin{pmatrix} \text{id} & X_Q \\ 0 & \text{id} \end{pmatrix} : \mathcal{B}_{\text{nf}} \cong \mathcal{B}, \; \begin{pmatrix} y \\ u \end{pmatrix} \mapsto \begin{pmatrix} y + X_Q \circ u \\ u \end{pmatrix}$$

The two behaviors $\mathcal{B}$ and $\mathcal{B}_{\rm nf}$ have the same autonomous part $\mathcal{B}^0 = \{y \in \mathcal{F}^{J^0};\ P \circ y = 0\}$, but their transfer matrices are H resp. $H_{\rm nf} := H - X_Q$.

Proof The first assertion is obvious, and the second follows from

$$\begin{pmatrix} y \\ u \end{pmatrix} \in \mathcal{B} \ \Leftrightarrow\ P \circ y = Q \circ u = (PX_Q + Q_{\rm nf}) \circ u \ \Leftrightarrow$$

$$P \circ (y - X_Q \circ u) = Q_{\rm nf} \circ u \ \Leftrightarrow\ \begin{pmatrix} y - X_Q \circ u \\ u \end{pmatrix} = \begin{pmatrix} \mathrm{id} & -X_Q \\ 0 & \mathrm{id} \end{pmatrix} \circ \begin{pmatrix} y \\ u \end{pmatrix} \in \mathcal{B}_{\rm nf}.$$

The last assertion is a consequence of

$$Q = PH = PX_Q + Q_{\rm nf}, \text{ hence } P(H - X_Q) = Q_{\rm nf}. \qquad \square$$

Corollary 4.6 In the situation of the preceding theorem the invertible matrices induce mutually inverse isomorphisms

$$\circ \begin{pmatrix} \mathrm{id} & -X_Q \\ 0 & \mathrm{id} \end{pmatrix}_{\rm ind} : M_{\rm nf} := \mathcal{D}^{1\times I}/\,\mathrm{row}(P, -Q_{\rm nf}) \cong M = \mathcal{D}^{1\times I}/\,\mathrm{row}(P, -Q) \textit{ and}$$

$$\circ \begin{pmatrix} \mathrm{id} & X_Q \\ 0 & \mathrm{id} \end{pmatrix}_{\rm ind} : M = \mathcal{D}^{1\times I}/\,\mathrm{row}(P, -Q) \cong M_{\rm nf} := \mathcal{D}^{1\times I}/\,\mathrm{row}(P, -Q_{\rm nf})$$

The next result shows that non-autonomous multidimensional behaviors with finite-dimensional autonomous part are not very interesting in contrast to finite-dimensional autonomous behaviors [17, 21] and to the one-dimensional situation. We use the data of equations (55) to (58) and of Lemma 4.1, in particular the transfer matrix H with $P'H = Q'$.

Theorem 4.7 *If the module M^0 of the autonomous part $\mathcal{B}^0$ of the multidimensional IO-behavior ($r > 1$) is finite-dimensional over F then the transfer matrix H is polynomial. This implies the direct sum decomposition*

$$\mathcal{B} = \left\{ \begin{pmatrix} y \\ u \end{pmatrix} \in \mathcal{F}^{J^0} \times \mathcal{F}^{J^{\rm free}};\ P' \circ y = Q' \circ u \right\} = \mathcal{B}^0 \oplus \left\{ \begin{pmatrix} H \circ u \\ u \end{pmatrix};\ u \in \mathcal{F}^{J^{\rm free}} \right\}$$

$$\textit{where } \mathcal{B}^0 = \{y \in \mathcal{F}^{J^0};\ P' \circ y = 0\} \underset{\textit{identification}}{=} \mathcal{B} \cap (\mathcal{F}^{J^0} \times \{0\}).$$

Thus the output $y = H \circ u$, up to its autonomous part, is obtained from the input u by differentiation (in the continuous case).

Proof

1. Since M^0 is finite-dimensional over F the annihilator $\mathrm{ann}_{F[s_\rho]}(M^0)$ is non-zero for each indeterminate s_ρ and hence

$$\mathrm{ann}_{F[s_\rho]}(M^0) = F[s_\rho] f_\rho(s_\rho),\ \ f_\rho(s_\rho) \neq 0 \text{ for } \rho = 1, \cdots, r.$$

2. Since $f_\rho(s_\rho)$ annihilates M^0 there is a matrix X_ρ such that

$$f_\rho(s_\rho)\,\mathrm{id}_{J^0} = X_\rho P', \text{ hence } f_\rho H = f_\rho\,\mathrm{id}\, H = X_\rho P' H = X_\rho Q'$$

is polynomial. Therefore $f_\rho H_{ij}$ is a polynomial for each index ρ and each entry H_{ij} of H.

3. Let, more generally, $h \in F(s)$ be any rational function such that $g_\rho = f_\rho(s_\rho)h$ is a polynomial for each $\rho = 1, \cdots, r$. Such a rational function is necessarily a polynomial. Indeed, define

$$d_\rho = \gcd(f_\rho(s_\rho), g_\rho),\ \ f'_\rho := f_\rho/d_\rho,\ \ g'_\rho := g_\rho/d_\rho,\ \text{hence}$$

$$\gcd(f_\rho(s_\rho)', g'_\rho) = 1,\ \ h = g'_\rho/f'_\rho \text{ and } g'_\rho f'_\sigma = g'_\sigma f'_\rho \text{ for all } \rho \neq \sigma = 1, \cdots, r.$$

Here gcd denotes the greatest common divisor. The last identity implies that $f_\rho(s_\rho)'$ divides $f_\sigma(s_\sigma)'$ for each $\rho \neq \sigma$ and hence all $f_\rho(s_\rho)'$ are units in F and $h = g'_\rho/f'_\rho$ is a polynomial as asserted. With this and 2. the theorem is proved. □

5 The Canonical State Representation of a Behavior

In the first part of this principal section we define the notion of a state behavior *and establish its properties (Theorem* 5.2*, Result* 5.6*), in the second part we prove that every multidimensional behavior is* canonically isomorphic *to a state behavior with respect to a chosen term order (Theorems* 5.12, 5.14*). Moreover we demonstrate the algorithms for a general one-dimensional and an instructive two-dimensional example (Examples* 5.15, 5.16*).*

In the first part we assume that Δ is an arbitrary finite index set, $<$ on $\Delta \times \mathbb{N}^r$ is a given term order and $U \subset \mathcal{D}^{1\times\Delta}$ is a submodule with the degree set $N := \deg(U)$, its cw-minimal set $D := \min_{\mathrm{cw}}(N)$ and its complement $\Gamma := (\Delta \times \mathbb{N}^r) \setminus N$. Moreover let $\mathcal{F}$ be any $\mathcal{D}$-module, $M := \mathcal{D}^{1\times\Delta}/U$ and $\mathcal{B} := U^\perp = \mathrm{Hom}_{\mathcal{D}}(M, \mathcal{F}) \subset \mathcal{F}^\Delta$ the derived $\mathcal{F}$-behavior. If $\mathcal{F}$ is not an injective cogenerator the relation $U = \mathcal{B}^\perp$ is not necessarily true so that U determines $\mathcal{B}$ uniquely, but not conversely.

Definition 5.1 Data as just introduced. We say that U, M and $\mathcal{B}$ are in *state* form with respect to the given term order on $\Delta \times \mathbb{N}^r$ if the degree set N is of the form $N = D + \mathbb{N}^r$ where all members of D have the form (α, ϵ_ρ) with $\alpha \in \Delta$ and $\rho \in [r] = \{1, \cdots, r\}$. We also call M resp. $\mathcal{B}$ a *state module resp. behavior.*

In the sequel we assume that U is in state form. For each $\alpha \in \Delta$ define

$$S(\alpha) := \{\rho \in [r];\ (\alpha, \epsilon_\rho) \notin N = \deg(U)\},\ \text{hence}$$

$$\begin{cases} (\alpha, \epsilon_\rho) \in N \\ (\alpha, \epsilon_\rho) \notin N \end{cases} \Leftrightarrow \begin{cases} \rho \notin S(\alpha) \\ \rho \in S(\alpha) \end{cases} \tag{69}$$

By assumption no $(\alpha, 0)$ is contained in D or N, and hence the state module M is minimally embedded according to Definition 3.13:

$$\text{For all } \alpha \in \Delta:\ N_\alpha \neq \mathbb{N}^r,\ \text{ie. } (\alpha, 0) \in \Delta \times \mathbb{N}^r \text{ is cw-minimal in } \Gamma. \tag{70}$$

Theorem 5.2 *Data from above. Assume that the module U is in state form. Then*

1. $D = \min_{\mathrm{cw}}(N) = \{(\alpha, \epsilon_\rho);\ \alpha \in \Delta,\ \rho \notin S(\alpha)\}$.
2. *The decomposition of* $\Gamma := (\Delta \times \mathbb{N}^r) \setminus N$ *according to* (25) *is given by*

$$\Gamma = \uplus_{\alpha \in \Delta}((\alpha, 0) + \mathbb{N}^{S(\alpha)}) = \uplus_{\alpha \in \Delta}\{\alpha\} \times \mathbb{N}^{S(\alpha)}.$$

In particular, the set $\Delta \times \{0\} \subset \Delta \times \mathbb{N}^r$ *is the set* Δ *of* U *according to* (25) *with* $S(\alpha, 0) = S(\alpha)$, *and*

$$\mathcal{D}^{1\times\Delta} = \oplus_{\alpha\in\Delta} F[s^{(\alpha)}]\delta_\alpha \oplus U = \oplus_{\alpha\in\Delta,\mu\in\mathbb{N}^{S(\alpha)})} F s^\mu \delta_\alpha \oplus U$$
$$M = \oplus_{\alpha\in\Delta} F[s^{(\alpha)}]\overline{\delta_\alpha} = \oplus_{\alpha\in\Delta,\mu\in\mathbb{N}^{S(\alpha)}} F s^\mu \overline{\delta_\alpha}$$

where $s^{(\alpha)} := (s_\rho)_{\rho\in S(\alpha)}$ *according to* (13) *and* $\overline{\delta_\alpha}$ *denotes the residue class of* δ_α *in* M. *We identify* $\Delta = \Delta \times \{0\},\ (\alpha, 0) = \alpha$.

3. *The unique reduced Gröbner basis of U with respect to the given term order has the form*

$$g_{\alpha,\rho} := g_{(\alpha,\epsilon_\rho)} := s_\rho\delta_\alpha - \sum_{\beta\in\Delta} f_{\alpha,\rho,\beta}\delta_\beta = s_\rho\delta_\alpha - \sum_{\beta\in\Delta,\nu\in\mathbb{N}^{S(\beta)}} f_{\alpha,\rho,\beta,\nu} s^\nu \delta_\beta$$
$$\text{where } \alpha \in \Delta,\ \rho \notin S(\alpha),\ \deg(g_{\alpha,\rho}) = (\alpha, \epsilon_\rho) \in D,$$
$$f_{\alpha,\rho,\beta} \in F[s^{(\beta)}] \text{ and } f_{\alpha,\rho,\beta,\nu} \in F.$$

4. *The state behavior* $\mathcal{B}$ *is thus given by*

$$\mathcal{B} = \left\{ w = (w_\alpha)_{\alpha\in\Delta} \in \mathcal{F}^\Delta;\ \forall(\alpha, \epsilon_\rho) \in D:\ s_\rho \circ w_\alpha = \sum_{\beta\in\Delta} f_{\alpha,\rho,\beta} \circ w_\beta \right\}.$$

Proof

1. The subset $\{(\alpha, \epsilon_\rho);\ \alpha \in \Delta,\ \rho \in [r]\}$ of $\Delta \times \mathbb{N}^r$ is cw-discrete and hence so is D. This implies $D = \min_{\mathrm{cw}}(D + \mathbb{N}^r)$. Moreover

$$(\alpha, \epsilon_\rho) \in N \ \Leftrightarrow\ (\alpha, \epsilon_\rho) \in D \text{ and thus}$$
$$D = \{(\alpha, \epsilon_\rho);\ (\alpha, \epsilon_\rho) \in N\} = \{(\alpha, \epsilon_\rho);\ \alpha \in \Delta,\ \rho \notin S(\alpha)\}.$$

2. From

$$N = D + \mathbb{N}^r = \cup_{d\in D}(d + \mathbb{N}^r) = \cup_{\alpha\in\Delta,\rho\notin S(\alpha)}((\alpha, \epsilon_\rho) + \mathbb{N}^r) =$$
$$= \cup_{\alpha\in\Delta}\{\alpha\} \times N_\alpha \text{ with}$$
$$N_\alpha = \cup_{\rho\notin S(\alpha)}(\epsilon_\rho + \mathbb{N}^r) = \{\mu \in \mathbb{N}^r;\ \exists\rho \notin S(\alpha) \text{ with } \mu_\rho > 0\}$$

we infer

$$\Gamma_\alpha = \mathbb{N}^r \setminus N_\alpha = \mathbb{N}^{S(\alpha)},\ \Delta_\alpha = \{0\} \text{ and } \Gamma = \uplus_{\alpha\in\Delta}((\alpha, 0) + \mathbb{N}^{S(\alpha)})$$

as asserted.

3. The assertions 3 and 4 follow directly from 1 and 2. □

As a special case of Theorem 4.3 a state behavior also admits a canonical IO-structure. Indeed, define

$$\Delta^0 := \{\alpha \in \Delta;\ N_\alpha \neq \emptyset \text{ or } S(\alpha) \neq [r]\} \text{ and } \Delta^{\text{free}} := \Delta \setminus \Delta^0$$
$$\mathcal{F}^{\Delta} := \mathcal{F}^{\Delta^0} \times \mathcal{F}^{\Delta^{\text{free}}} \ni w = \begin{pmatrix} x \\ u \end{pmatrix}. \tag{71}$$

As usual we call $x := (w_\alpha)_{\alpha\in\Delta^0}$ the *state* and $u := (w_\alpha)_{\alpha\in\Delta^{\text{free}}}$ the *input* of the state behavior and define

$$\mathcal{B}^0 := \{x \in \mathcal{F}^{\Delta^0};\ \begin{pmatrix} x \\ 0 \end{pmatrix} \in \mathcal{B}\}. \tag{72}$$

If $(\alpha, \epsilon_\rho) \in D$ then $\rho \notin S(\alpha)$ and therefore $\alpha \in \Delta^0$ and $w_\alpha = x_\alpha$.

Corollary 5.3

1. The $\mathcal{F}$-systems $\mathcal{B}$ and $\mathcal{B}^0$ are given by the equations

$$\mathcal{B} := \left\{ w = \begin{pmatrix} x \\ u \end{pmatrix} \in \mathcal{F}^{\Delta};\ \forall\alpha \in \Delta\forall\rho \notin S(\alpha),\ ie.\ \forall(\alpha, \epsilon_\rho) \in D : \right.$$
$$\left. s_\rho \circ x_\alpha = \sum_{\beta\in\Delta^0} f_{\alpha,\rho,\beta} \circ x_\beta + \sum_{\beta\in\Delta^{\text{free}}} f_{\alpha,\rho,\beta} \circ u_\beta \right\}$$
$$\mathcal{B}^0 := \left\{ x \in \mathcal{F}^{\Delta^0};\ \forall(\alpha, \epsilon_\rho) \in D :\ s_\rho \circ x_\alpha = \sum_{\beta\in\Delta^0} f_{\alpha,\rho,\beta} \circ x_\beta \right\}$$

2. If $\mathcal{F}$ is an injective cogenerator the behavior $\mathcal{B}^0$ is autonomous and the behavior sequence

$$0 \longrightarrow \mathcal{B}^0 \xrightarrow{\binom{\text{id}}{0}\circ} \mathcal{B} \xrightarrow{(0\,\text{id})\circ} \mathcal{F}^{\Delta^{\text{free}}} \longrightarrow 0$$

is exact, ie. for every input $u \in \mathcal{F}^{\Delta^{\text{free}}}$ there is a state vector $x \in \mathcal{F}^{\Delta^0}$ such that $\binom{x}{u} \in \mathcal{B}$.

Remark 5.4 A polynomial matrix and the behavior defined by it are called *of the first order* if the entries of the matrix are polynomials of total degree at most one, i.e., linear with possibly non-zero constant term. Remark that, in general, the polynomials $f_{\alpha,\rho,\beta}$ are not linear or of the first order, but can have arbitrary total degree and the same applies to the polynomial vectors $g_{\alpha,\rho}$ of the reduced Gröbner basis. However, the following theorem from [23] justifies the state space terminology introduced above.

In the following theorem we use the space

$$F^{\mathbb{N}^r} = F[[z]] = F[[z_1, \cdots, z_r]] \ni y = (y_\mu)_{\mu\in\mathbb{N}^r} = \sum_{\mu\in\mathbb{N}^r} y_\mu z^\mu.$$

of multi-sequences or formal power series. For $\alpha \in \Delta$ we denote $z^{(\alpha)} = (z_\rho)_{\rho \in S(\alpha)}$ according to (13) and identify $F[[z^{(\alpha)}]]$ as $F[s^{(\alpha)}]$-submodule of $F[[z]]$. For $y \in F[[z]]$ and $\alpha \in \Delta$ we define $y(z^{(\alpha)}, 0) := \sum_{\mu \in \mathbb{N}^{S(\alpha)}} y_\mu z^\mu$. If $F = \mathbb{R}$ or $F = \mathbb{C}$ and if y is a convergent power series and therefore an analytic function on $\mathbb{R}^r$ resp. $\mathbb{C}^r$ in the neighborhood of zero the power series $y(z^{(\alpha)}, 0)$ is exactly the restriction of y to $\mathbb{R}^{S(\alpha)} \subset \mathbb{R}^r$ resp. $\mathbb{C}^{S(\alpha)}$.

Assumption 5.5 We assume that the function $\mathcal{D}$-module $\mathcal{F}$ is a subspace of $F[[z]]$ on which $\mathcal{D} = F[s]$ acts either by left shifts, ie. $(s_\rho \circ y)_\mu = y_{\mu + \epsilon_\rho}$, or by partial differentiation, i.e., $s_\rho \circ y = \partial y / \partial z_\rho$, and consider the following five cases which were treated in [20] and more comprehensively in [23].

1. F is a field of arbitrary characteristic and $\mathcal{F} := F[[z]]$ with the action by left shifts [20, Th. 5.63], [23, Th. 15].
2. F is a field of characteristic zero and $\mathcal{F} = F[[z]]$ with the action by partial differentiation [23, Th. 16].
3. F is the field of real or complex numbers and $\mathcal{F} = F\langle z \rangle$ is the algebra of *(locally) convergent* power series with the action by left shifts [23, Th. 24]. A power series $y = \sum_{\mu \in \mathbb{N}^r} y_\mu z^\mu$ is convergent if and only if its coefficients satisfy a growth condition

$$| y_\mu | \leq C a_1^{\mu_1} * \cdots * a_r^{\mu_r} \text{ for all } \mu \in \mathbb{N}^r$$

   where C and the a_ρ, $\rho = 1, \cdots, r$, are positive real numbers.
4. F is the field of real or complex numbers and $\mathcal{F} = F\langle z \rangle$ is the algebra of convergent power series with the action by partial differentiation [23, Th. 29] and the term order is graded (compare Theorem 3.6). The solution of the Cauchy problem in this case is the main result in Riquier's book [32, p. 254, Théorème d'Existence], even for certain, so-called orthonomic non-linear systems. The Cauchy-Kovalevskaya theorem was a predecessor of Riquier's results. A variant of Riquier's theorem was also reproven by Pommaret [26, Th. 4.3] in a different language and was attributed to Cartan and Kähler. Gerdt also treated the problem in [9, Th. 29, Th. 31], but did not prove the *existence* of analytic solutions. Malgrange has recently written a survey with historical comments on this subject [19].
5. $F := \mathbb{C}$ is the field of complex numbers and $\mathcal{F} := O(\mathbb{C}^r, \exp) \in \mathbb{C}\langle z \rangle$ is the algebra of *entire holomorphic* functions of *exponential type* with the action by partial differentiation [23, Th. 26]. An entire holomorphic function y is called of exponential type if it satisfies a growth condition

$$| y(z) | \leq C \exp(a_1 | z_1 | + \cdots + a_r | z_r |) \text{ for all } z \in \mathbb{C}^r$$

   where C and the a_ρ, $\rho = 1, \cdots, r$, are positive real numbers.

If S is a subset of $[r]$ and if $\mathcal{F}$ is any of the function spaces of the preceding assumption let $\mathcal{F}(S) = F[[(z_\rho)_{\rho \in S}]]$ denote the corresponding function space in the indeterminates or variables z_ρ, $\rho \in S$.

Result 5.6 (The Cauchy problem [23]) *Assume the data of Theorem 5.2 and that $\mathcal{F}$ is one of the function spaces from Assumption 5.5. Then for arbitrary functions*

$v_\alpha \in \mathcal{F}(S(\alpha))$, $\alpha \in \Delta$, *there is a unique trajectory* $w \in \mathcal{B}$ *such that* $w_\alpha(z^{(\alpha)}, 0) = v_\alpha$ *for all* $\alpha \in \Delta$. *In other terms, the map*

$$\mathcal{B} \to \prod_{\alpha\in\Delta^0} \mathcal{F}(S(\alpha)) \times \mathcal{F}^{\Delta^{\text{free}}},\ w = \begin{pmatrix} x \\ u \end{pmatrix} \mapsto \begin{pmatrix} x_\alpha(z^{(\alpha)}, 0)_{\alpha\in\Delta^0} \\ u \end{pmatrix}, \tag{73}$$

is an isomorphism.

The component v_α *is called an* input *resp. an* initial condition *if* $\alpha \in \Delta$ *belongs to* Δ^{free} *resp. to* Δ^0. *The vectors* $x = (w_\alpha)_{\alpha\in\Delta^0}$ *resp.* $(x_\alpha(z^{(\alpha)}, 0))_{\alpha\in\Delta^0}$ *are called the* global *resp. the* local *state of the system* $\mathcal{B}$. *Hence the input and the initial condition or local state give rise to a unique trajectory of the system.*

Proof The proof of this theorem is a special case of the theorems in [23], the exact references being given in Assumption 5.5. □

Remark 5.7

1. The preceding state definition generalizes the state property of x in the Kalman system $\overset{\bullet}{x} = Ax + Bu$ as explained in the introduction. It also satisfies Pommaret's requirement from [27, p. 311, lines 1^+ ff.]: "*Indeed* state *is what must be given in order to integrate the control system whenever the input is given.*" Integration of a differential system is the traditional term for solving it. In contrast to the *Spencer form* of a system which generalizes the one-dimensional Kalman form in Pommaret's approach to state representations [29, 30] and which is a system of the first order the state systems of the present paper are not generally of the first order.
2. The following remarks are taken from the literature and concern the Cauchy problem for linear systems of partial differential equations with constant coefficients for function spaces like C^∞-functions or distributions which do no consist of formal or convergent power series as assumed in Result 5.6. The general behaviors of the present paper have not been treated in this context. The preceding theorem does not hold in general, but requires additional conditions on the system or the initial data. However, the theorem is applicable to entire functions of exponential type and to convergent power series with a graded term order and therefore suggests the formulation of the Cauchy problem also for more general function spaces. One has to distinguish between *uniqueness* and *correctness* results [8, ch. II, ch. III]. Correctness or well-posedness of the Cauchy problem signifies that it is uniquely solvable and that the solution depends continuously on the initial data. In general, uniqueness results require initial data with bounded growth at ∞ [8, Th.1, p. 42], [24, ch.VI, Section 6.2]. If the function space is an injective cogenerator it suffices to solve the Cauchy problem for autonomous systems as for instance in [8].
3. Palamodov has proven a quite general uniqueness result for weakly hypoelliptic autonomous systems $R \circ w = 0$ [24, ch.VI, Section 6.2]. In this case the sets $S(\alpha)$ do not depend on $\alpha \in \Delta^0$.

4. Most systems treated in the literature [8, ch.II, Section 3, (1)], [36, (15.1)] are of or can be easily reduced to the form

$$(\mathrm{id}_n\, \partial/\partial z_1 + P_{II}(i\partial/\partial z_{II})x = 0,\ \ P_{II} \in \mathbb{C}[s_{II}]^{n\times n},\ \ x \in \mathcal{F}^n,$$
$$s_{II} = (s_2, \cdots, s_r),\ \ z_{II} = (z_2, \cdots, z_r). \tag{74}$$

The distinguished variable z_1 is usually interpreted as time. The constant $i := \sqrt{-1}$ in front of $\partial/\partial z_{II}$ is just a convention in connection with the Fourier transform in the z_{II}-space. For the lexicographic order $s_1 > s_2 > \cdots > s_r$ the matrix $P := s_1\,\mathrm{id}_n + P_{II}(is_{II})$ is a Gröbner matrix since the S-polynomials of the rows

$$P_{j-} = s_1\delta_j + P_{II,j-}(is_{II}) \in \mathbb{C}[s]^{1\times n}$$

are zero by definition. In particular, the autonomous system $P \circ x = 0$ and its inhomogeneous counter-part $P \circ x = u$ are in state space form and the isomorphism

$$\left\{\begin{pmatrix} x \\ u \end{pmatrix} \in \mathcal{F}^{n+n};\ P \circ x = u\right\} \cong \mathcal{F}(z_{II})^n \times \mathcal{F}^n,\ \begin{pmatrix} x \\ u \end{pmatrix} \mapsto \begin{pmatrix} x(0, z_{II}) \\ u \end{pmatrix},$$

holds for $\mathcal{F} = O(\mathbb{C}^r, \exp)$. In general, the isomorphism does not hold for *locally* convergent power series since the pure lexicographic term order is not graded. A counter-example [4, ch.1, Section 2.2] is given by the heat equation

$$y_{z_1} - y_{z_2z_2} = 0,\ \ y(0, z_2) = (1 - z_2)^{-1}$$

whose unique formal power series solution y satisfies $y(z_1, 0) = \sum_{k=0}^{\infty}(2k)!z_1^k$ and is not convergent.

5. For the heat equation $(s_1 - s_2^2) \circ y = y_{z_1} - y_{z_2z_2} = 0$, $\mathcal{F} := C^\infty(\mathbb{R}^2)$ and the lexicographic term order $s_1 > s_2$ the map (73) $y \mapsto y(0, z_2)$ is not injective. Indeed, there is a solution y whose support is exactly the half-space $\{z \in \mathbb{R}^2;\ z_1 = z \bullet \delta_1 \le 0\}$ [11, Th. 8.6.7]. The reason is that $\{z_1 = z \bullet \delta_1 = 0\}$ is a *characteristic* hyper-plane (line).
For $\mathcal{F} := C^\infty(\mathbb{R}^2)$ and indeed any $\mathbb{C}[s]$-module the system isomorphism

$$\{y \in \mathcal{F};\ (s_1 - s_2^2) \circ y = 0\} \cong \left\{x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathcal{F}^2;\ \begin{pmatrix} s_2 & -1 \\ s_1 & -s_2 \end{pmatrix} \circ x = 0\right\}$$
$$y = x_1 \leftrightarrow x = \begin{pmatrix} y \\ s_2 \circ y \end{pmatrix}$$

holds. The preceding non-uniqueness also applies to this first order autonomous system with a square-matrix. These examples underline the necessity of growth conditions on the function spaces for uniqueness results as mentioned in item 2.

6. We specialize the systems of (74) to first order ones, i.e., we consider systems [36, ch.15, (15.1)]

$$P\circ x=u \text{ with } P:=s_1\,\mathrm{id}_n-\sum_{\rho=2}^{r}s_\rho A_\rho-A_0,\ \ A_\sigma\in\mathbb{C}^{n\times n},\ \ x,u\in\mathcal{F}^n$$

$$\mathcal{B}:=\left\{\begin{pmatrix}x\\u\end{pmatrix}\in\mathcal{F}^{n+n};\ \ P\circ x=u\right\},\ \ \mathcal{B}^0:=\{x\in\mathcal{F}^n;\ \ P\circ x=0\}. \tag{75}$$

The determinant of P is the characteristic polynomial of $\sum_{\rho=2}^{r}s_\rho A_\rho+A_0$ and of the form

$$\chi=\chi_n+\cdots+\chi_0,\ \ \chi_n=s_1^n+\cdots$$

where χ_k is the homogeneous component of degree k of χ. Any $x\in\mathcal{B}^0$ also satisfies $\chi\circ x=0$. The vector $\delta_1:=(1,0,\cdots,0)$ is non-characteristic for χ, i.e., $\chi_n(\delta_1)=1^n\neq 0$. Holmgren's theorem and its consequence [11, Cor. 8.6.9] imply that any distributional solution of $\chi\circ y=0$ whose support lies in the right half-space $\{z\in\mathbb{R}^r;\ \ z_1=z\bullet\delta_1\geq 0\}$ is indeed zero. If x is a C^1-solution of $P\circ x=0$ with initial condition $x(0,z_{II})=0$ the continuous function x_+ defined by

$$x_+(z):=\begin{cases}x(z) & \text{if } z_1\geq 0\\ 0 & \text{if } z_1\leq 0\end{cases}\ \text{satisfies } P\circ x_+=0 \text{ in } \mathcal{D}'(\mathbb{R}^r)$$

and is zero by the preceding remarks. This argument uses the first order property of P; I learned it from my colleague Peter Wagner. Hence the map

$$\{x\in C^1(\mathbb{R}^r)^n;\ \ P\circ x=0\}\to C^1(\mathbb{R}^{r-1})^n,\ \ x\mapsto x(0,z_{II}),$$

is injective [4, Ch.1, pp. 34–36]. This argument can be extended to distributional solutions in $C^1(\mathbb{R},\mathcal{D}'(\mathbb{R}^{r-1}))\subset\mathcal{D}'(\mathbb{R}^r)$.

Correctness of the Cauchy problem in the present context characterizes *hyperbolicity*. Indeed, according to [36, ch.15, Def. 15.1, Th. 15.1, Th. 15.2] the following assertions are equivalent:

a. For $\mathcal{F}:=C^\infty(\mathbb{R}^r)$ the map (73)

$$\mathcal{B}\to C^\infty(\mathbb{R}^{r-1})^n\times\mathcal{F}^n,\ \ w=\begin{pmatrix}x\\u\end{pmatrix}\mapsto\begin{pmatrix}x(0,z_{II})\\u\end{pmatrix}$$

is an isomorphism.

b. The system (75) is *hyperbolic*. This signifies that there is constant $C>0$ such that for the matrix $A:=A_2\xi_2+\cdots A_r\xi_r-iA_0,\ \xi\in\mathbb{R}^r,\ i:=\sqrt{-1}$, and every eigenvalue λ of A the inequality $|\operatorname{im}(\lambda)|\leq C$ holds.

Hyperbolicity for single equations is thoroughly discussed in [12, ch.XII]. Hyberbolic systems of the form (74) are defined and investigated in [8, Ch. III, Section 3]. To my knowledge hyperbolicity has not been defined for the general behaviors of the present paper.

Remark 5.8 (*First order representations*) In [43, Section 6.1] Zerz proves that any polynomial matrix $R \in F[s_1, \cdots, s_r]^{k\times l}$ admits an LFT- representation (*linear fractional transformation*)

$$\begin{aligned} R &= D + C\Delta(s)(\mathrm{id}_n - A\Delta(s))^{-1}B \text{ where} \\ n &:= n_1 + \cdots + n_r,\ \ \Delta(s) = \mathrm{diag}(s_1\,\mathrm{id}_{n_1}, \cdots, s_r\,\mathrm{id}_{n_r}) \\ &A \in F^{n\times n},\ B \in F^{n\times l},\ C \in F^{k\times n},\ D \in F^{k\times l} \\ &\det(\mathrm{id}_n - A\Delta(s)) = 1. \end{aligned} \tag{76}$$

The latter condition is automatically satisfied if A is a strictly lower triangular matrix, and hence representations (76) can be randomly generated for experimental purposes as is done in Example 5.16 below. The polynomial matrix $\mathrm{id}_n - A\Delta(s)$ is invertible and gives rise to the mutually inverse system isomorphisms

$$\begin{aligned} \mathcal{B}_{\mathrm{lin}} &:= \left\{\begin{pmatrix} x \\ w\end{pmatrix} \in \mathcal{F}^{n+l};\ R_{\mathrm{lin}} \circ \begin{pmatrix} x \\ w\end{pmatrix} = 0\right\} \cong \left\{w \in \mathcal{F}^l;\ R\circ w = 0\right\} \\ \begin{pmatrix} x \\ w\end{pmatrix} &= \begin{pmatrix} (\mathrm{id}_n - A\Delta(s))^{-1}\circ Bw \\ w\end{pmatrix} \leftrightarrow w \\ \text{where } R_{\mathrm{lin}} &:= \begin{pmatrix} \mathrm{id}_n - A\Delta(s) & -B \\ C\Delta(s) & D\end{pmatrix} = (M R_0) \text{ with } M = \begin{pmatrix} \mathrm{id}_n - A\Delta(s) \\ C\Delta(s)\end{pmatrix},\ R_0 := \begin{pmatrix} -B \\ D\end{pmatrix}, \\ \text{hence } \mathcal{B}_{\mathrm{lin}} &= \left\{\begin{pmatrix} x \\ w\end{pmatrix} \in \mathcal{F}^{n+l};\ M\circ x + R_0 w = 0\right\}. \end{aligned} \tag{77}$$

The system $\mathcal{B}_{\mathrm{lin}}$ is a first order system and a state system with state x and manifest variables w in the sense of Willems [40, problem 1.12,p. 56], [33], but not in the sense of the present paper. No analogue of Result 5.6 holds in this situation. For the solution of the Cauchy problem of $R\circ w = 0$ the isomorphism $\mathcal{B}_{\mathrm{lin}} \cong \mathcal{B}$ is of little use only since, in general and as exemplified in Example 5.16, the *canonical* state representation of $\mathcal{B}_{\mathrm{lin}}$ is not of the first order although the matrix R_{lin} is.

In the second part of this section we are going to construct the canonical state representation of a behavior with respect to a chosen term order.

For this purpose we make the following assumptions. Let I be a finite set, $<$ a term order on $I \times \mathbb{N}^r$, U a $\mathcal{D}$-submodule of $\mathcal{D}^{1\times I}$, $M := \mathcal{D}^{1\times I}/U$ and $N := \deg(U)$ where the degree is taken with respect to the given term order $<$. The module U gives rise to the $\mathcal{F}$-behavior

$$\mathcal{B} := U^{\perp} = \left\{w = (w_i)_{i\in I} \in \mathcal{F}^I;\ \forall f = (f_i)_{i\in I} \in U : f\circ w = \sum_{i\in I} f_i \circ w_i = 0\right\}$$

where $\mathcal{F}$ is any $\mathcal{D}$-module. We use the canonical data

$$I \times \mathbb{N}^r = \Gamma \uplus N,\ \ D := \min_{\mathrm{cw}}(N),\ \ \Gamma = \uplus_{\alpha\in\Delta}(\alpha + \mathbb{N}^{S(\alpha)}), \tag{78}$$

from (25) which depend on U and the chosen term order $<$ only. Then

$$\begin{aligned} \mathcal{D}^{1\times I} &= \oplus_{\gamma\in\Gamma} F s^{\mu(\gamma)}\delta_{i(\gamma)} \oplus U = \oplus_{\alpha\in\Delta} F[s^{(\alpha)}] s^{\mu(\alpha)}\delta_{i(\alpha)} \oplus U \\ M &= \oplus_{\gamma\in\Gamma} F s^{\mu(\gamma)}\overline{\delta_{i(\gamma)}} = \oplus_{\alpha\in\Delta} F[s^{(\alpha)}] s^{\mu(\alpha)}\overline{\delta_{i(\alpha)}} \end{aligned} \tag{79}$$

as in Corollary 3.4 where $\overline{x} := x + U$ denotes the residue class of $x \in \mathcal{D}^{1\times I}$ in M. As in Corollary 3.8 let

$$\begin{aligned} f_d &= s^{\mu(d)}\delta_{i(d)} - \sum_{\alpha\in\Delta} f_{d,\alpha} s^{\mu(\alpha)}\delta_{i(\alpha)} \\ &= s^{\mu(d)}\delta_{i(d)} - \sum_{\alpha\in\Delta,\ \mu\in\mathbb{N}^{S(\alpha)}} f_{d,\alpha,\mu} s^{\mu(\alpha)+\mu}\delta_{i(\alpha)} \text{ with } d\in D, \\ \deg(f_d) &= d = (i(d),\mu(d)),\ \ f_{d,\alpha}\in F[s^{(\alpha)}],\ \ f_{d,\alpha,\mu}\in F \end{aligned} \tag{80}$$

denote the unique reduced Gröbner basis of U. According to Corollary 3.14 we assume without loss of generality that *the module M is minimally embedded*, i.e.,

$$\text{For each } i\in I:\ \ (i,0)\in\Delta \text{ or } N_i\neq\mathbb{N}^r. \tag{81}$$

According to Theorem 4.3 the module U has a canonical IO-structure which is defined by (see (59) and Lemma 4.2)

$$\begin{aligned} I^0 &:= \{i\in I;\ \ (i,0)\in\Delta,\ \ S(i,0)\neq[r] \text{ or } N_i\neq\emptyset\} \\ I^{\text{free}} &:= \{i\in I;\ \ (i,0)\in\Delta,\ \ S(i,0)=[r] \text{ or } N_i=\emptyset\},\ \ \text{hence} \\ I &= I^0 \uplus I^{\text{free}} \text{ and } i(d)\in I^0 \text{ for all } d\in D,\ \text{and} \\ \Delta^0 &:= \{\alpha=(i,\mu)\in\Delta;\ \ i\in I^0 \text{ or } S(\alpha)\neq[r]\}, \\ \Delta^{\text{free}} &:= \{(i,0)\in\Delta;\ \ i\in I^{\text{free}}\} \underset{\text{identification}}{=} I^{\text{free}},\ \ (i,0)=i, \\ \Delta &= \Delta^0\uplus\Delta^{\text{free}} \text{ and } \mathcal{F}^I=\mathcal{F}^{I^0}\times\mathcal{F}^{I^{\text{free}}}\ni w=\begin{pmatrix} y\\ u\end{pmatrix} \end{aligned} \tag{82}$$

where $y := (w_i)_{i\in I^0}$ is the output and $u := (w_i)_{i\in I^{\text{free}}}$ is the input of $\mathcal{B}$. With these notations we have

$$\begin{aligned} \mathcal{B} &:= \left\{ w\in\mathcal{F}^I;\ \forall d\in D: s^{\mu(d)}\circ w_{i(d)} - \sum_{\alpha\in\Delta} f_{d,\alpha}s^{\mu(\alpha)}\circ w_{i(\alpha)} = 0 \right\} \\ &= \Bigg\{ w=\begin{pmatrix} y\\ u\end{pmatrix}\in\mathcal{F}^{I^0}\times\mathcal{F}^{I^{\text{free}}};\forall d\in D: \\ &\qquad s^{\mu(d)}\circ y_{i(d)} - \sum_{\alpha\in\Delta^0} f_{d,\alpha}s^{\mu(\alpha)}\circ y_{i(\alpha)} = \sum_{i\in I^{\text{free}}} f_{d,(i,0)}\circ u_i \Bigg\} \\ \mathcal{B}^0 &:= \left\{ y\in\mathcal{F}^{I^0};\ \begin{pmatrix} y\\ 0\end{pmatrix}\in\mathcal{B} \right\} \\ &= \left\{ y\in\mathcal{F}^{I^0};\ \forall d\in D:\ s^{\mu(d)}\circ y_{i(d)} - \sum_{\alpha\in\Delta^0} f_{d,\alpha}s^{\mu(\alpha)}\circ y_{i(\alpha)} = 0 \right\}. \end{aligned} \tag{83}$$

If $\mathcal{F}$ is an injective cogenerator the behavior $\mathcal{B}^0$ is the autonomous part of $\mathcal{B}$, and for every input $u\in\mathcal{F}^{I^{\text{free}}}$ there is an output $y\in\mathcal{F}^{I^0}$ such that $\binom{y}{u}\in\mathcal{B}$.

With the preceding data we further define matrices

$$C\in F^{I\times\Delta}\subset\mathcal{D}^{I\times\Delta}\text{ by }C_{i-}:=\delta_{(i,0)}\in F^{1\times\Delta}\text{ or}$$

$$C_{i\alpha}:=\delta_{(i,0)\alpha}=\begin{cases}1 & \text{if }(i,0)=\alpha\\ 0 & \text{otherwise}\end{cases},$$

$$E\in\mathcal{D}^{\Delta\times I}\text{ by }E_{\alpha-}:=s^{\mu(\alpha)}\delta_{i(\alpha)}\in\mathcal{D}^{1\times I}\text{ or}$$

$$E_{\alpha i}:=s^{\mu(\alpha)}\delta_{i(\alpha)i}.$$

$$\text{Then }CE=\mathrm{id}_I \tag{84}$$

where id_I denotes the $I\times I$-identity matrix. Due to $I=I^0\uplus I^{\mathrm{free}}$ and $\Delta=\Delta^0\uplus\Delta^{\mathrm{free}}$ and with the identification $I^{\mathrm{free}}=\Delta^{\mathrm{free}}$ these matrices can be written in block form as

$$C=\begin{pmatrix}C^0 & 0\\ 0 & \mathrm{id}_{I^{\mathrm{free}}}\end{pmatrix}\in F^{(I^0\uplus I^{\mathrm{free}})\times(\Delta^0\uplus\Delta^{\mathrm{free}})}\text{ and}$$

$$E=\begin{pmatrix}E^0 & 0\\ 0 & \mathrm{id}_{I^{\mathrm{free}}}\end{pmatrix}\in\mathcal{D}^{(\Delta^0\uplus\Delta^{\mathrm{free}})\times(I^0\uplus I^{\mathrm{free}})}. \tag{85}$$

Remark that C is well-defined only due to assumption (81). The letter C is chosen in analogy to the matrix C in Kalman-state systems $s\circ x=Ax+Bu,\ \ y=Cx+Du$ in the one-dimensional case. The matrices E and C are the analogues of A_0^0 and C_0^0 in [14, ch.6,(32a),(32b))] and induce the $\mathcal{D}$-linear maps

$$\circ C:\mathcal{D}^{1\times I}\to\mathcal{D}^{1\times\Delta},\ \delta_i\mapsto\delta_iC=C_{i-}=\delta_{(i,0)},$$
$$\circ E:\mathcal{D}^{1\times\Delta}\to\mathcal{D}^{1\times I},\ \delta_\alpha\mapsto\delta_\alpha E=E_{\alpha-}=s^{\mu(\alpha)}\delta_{i(\alpha)}. \tag{86}$$

Define the module

$$U^s:=(\circ E)^{-1}(U):=\left\{f=(f_\alpha)_{\alpha\in\Delta}\in\mathcal{D}^{1\times\Delta};\ \sum_{\alpha\in\Delta}f_\alpha s^{\mu(\alpha)}\delta_{i(\alpha)}\in U\right\}$$
$$=\left\{f=(f_\alpha)_{\alpha\in\Delta}\in\mathcal{D}^{1\times\Delta};\ \sum_{\alpha\in\Delta}f_\alpha s^{\mu(\alpha)}\overline{\delta_{i(\alpha)}}=0\text{ in }M\right\}\subset\mathcal{D}^{1\times\Delta},$$

its factor module $M^s:=\mathcal{D}^{I\times\Delta}/U^s$ and its associated $\mathcal{F}$-behavior

$$\mathcal{B}^s:=(U^s)^{\perp}:=\left\{w^s=(w^s_\alpha)_{\alpha\in\Delta}\in\mathcal{F}^\Delta;\ \forall(f_\alpha)_{\alpha\in\Delta}\in U^s:\ \sum_\alpha f_\alpha\circ w^s_\alpha=0\right\}. \tag{87}$$

We are going to show that $\mathcal{B}^s$ is the canonical state behavior isomorphic to $\mathcal{B}$.

Theorem 5.9 *Data as just introduced. The matrices C and E induce inverse $\mathcal{D}$-isomorphisms*

$$M^s\ \underset{(\circ C)_{\mathrm{ind}}}{\overset{(\circ E)_{\mathrm{ind}}}{\rightleftarrows}}\ M,\quad \overline{(f_\alpha)_\alpha}\to\sum_{\alpha\in\Delta}f_\alpha s^{\mu(\alpha)}\overline{\delta_{i(\alpha)}},\quad \sum_{i\in I}f_i\overline{\delta_{(i,0)}}\leftarrow\overline{(f_i)_i}. \tag{88}$$

and behavior isomorphisms

$$\mathcal{B}^s \underset{E\circ}{\overset{C\circ}{\rightleftarrows}} \mathcal{B}, \quad w^s \to (w^s_{(i,0)})_{i\in I}, \quad (s^{\mu(\alpha)} \circ w_{i(\alpha)})_{\alpha\in\Delta} \leftarrow w. \tag{89}$$

Proof By definition of U^s and M^s the matrix E induces a *monomorphism* $(\circ E)_{\mathrm{ind}} : M^s \to M$. Since

$$(i,0) \in \Delta \text{ for all } i \in I,\ \ M = \sum_{i\in I} \mathcal{D}\overline{\delta_i} \text{ and } (\circ E)_{\mathrm{ind}}(\overline{\delta_{(i,0)}}) = \overline{\delta_i}$$

this induced map $(\circ E)_{\mathrm{ind}}$ is also surjective and therefore an isomorphism which, by duality, ie. application of the functor $\mathrm{Hom}_{\mathcal{D}}(-,\mathcal{F})$, implies the isomorphism $E\circ : \mathcal{B} \cong \mathcal{B}^s$.

The matrix identity $CE = \mathrm{id}_I$ implies $UCE = U$, hence that $UC \subset U^s = (\circ E)^{-1}(U)$ and that the induced map $(\circ C)_{\mathrm{ind}} = M \to M^s$ is well-defined. Since $(\circ E)_{\mathrm{ind}}$ is an isomorphism and $CE = \mathrm{id}$ we infer that $(\circ C)_{\mathrm{ind}}$ resp. $C\circ : \mathcal{B}^s \cong \mathcal{B}$ are the inverse isomorphisms of $(\circ E)_{\mathrm{ind}} : M^s \cong M$ resp. $E\circ : \mathcal{B} \cong \mathcal{B}^s$. □

It remains to show that U^s and $\mathcal{B}^s$ are a state module resp. behavior with respect to a canonical term order which we are going to define instantly.

We use the given term order on $I \times \mathbb{N}^r$ and the lexicographic term order on $\mathbb{N}^r$ with $\epsilon_1 > \epsilon_2 > \cdots$. On $\Delta \times \mathbb{N}^r$ we define the order

$$\begin{aligned}(\beta,\nu) < (\alpha,\mu) = (i(\alpha),\mu(\alpha),\mu) :\Leftrightarrow\ & \beta+\nu < \alpha+\mu = (i(\alpha),\mu(\alpha)+\mu) \text{ or}\\ & \beta+\nu = \alpha+\mu \text{ and } \mu(\alpha) <_{\mathrm{lex}} \mu(\beta).\end{aligned} \tag{90}$$

Remark that $\beta+\nu = \alpha+\mu$ implies $i(\alpha) = i(\beta)$.

Lemma 5.10 The well-order from (90) is a term order.

Proof We give the proof for the computationally essential case that the order on $I \times \mathbb{N}^r \subset \mathbb{N}^{1+r}$ is given by

$$\alpha = \begin{pmatrix} i \\ \mu \end{pmatrix} < \beta = \begin{pmatrix} j \\ \nu \end{pmatrix} :\Leftrightarrow\ \ U\alpha = U_{-0}i + U_{II}\mu$$

according to (40) and show additionally that the new order admits the same construction and therefore the Gröbner basis computation with MAPLE. The $\mathbb{Z}$-linear monomorphism

$$\begin{pmatrix} \alpha \\ \mu \end{pmatrix} = \begin{pmatrix} i(\alpha) \\ \mu(\alpha) \\ \mu \end{pmatrix} \mapsto \begin{pmatrix} U(\alpha+\mu) \\ -\mu(\alpha) \end{pmatrix} = \begin{pmatrix} U_{-0} & U_{II} & U_{II} \\ 0 & -\mathrm{id}_r & 0 \end{pmatrix} \begin{pmatrix} i(\alpha) \\ \mu(\alpha) \\ \mu \end{pmatrix} \tag{91}$$

on $\Delta \times \mathbb{N}^r \subset \mathbb{N}^{1+r+r}$ induces the term order from (90). □

Remark 5.11 For the use of the preceding order (91) with MAPLE we introduce additional indeterminates $d_0, \cdots, d_r$ and $d := (d_0, \cdots, d_r)$ and identify

$$(\alpha, \mu) = \begin{pmatrix} \alpha \\ \mu \end{pmatrix} = s^\mu \delta_\alpha = s^\mu d^\alpha = s^{\mu_1} * \cdots * s^{\mu_r} * d_0^{i(\alpha)} * d_1^{\mu_1(\alpha)} * \cdots * d_r^{\mu_r(\alpha)}$$

$$\text{and } F[s]^{1\times\Delta} = \oplus_{\alpha\in\Delta} F[s]\delta_\alpha = \oplus_{\alpha\in\Delta} F[s]d^\alpha \subset F[s_1, \cdots, s_r][d_0, \cdots, d_r].$$

With respect to the term order just introduced and according to (25) we define

$$N^s := \deg(U^s) \subset \Delta \times \mathbb{N}^r,\ D^s := \min_{\mathrm{cw}}(N^s),\ \Gamma^s := (\Delta \times \mathbb{N}^r) \setminus N^s,\ \Delta^s \subset \Gamma^s,$$
$$\text{and for all } \alpha^s \in \Delta^s \text{ the set } S(\alpha^s) \subset [r] \text{ with } \Gamma^s = \uplus_{\alpha^s\in\Delta^s}(\alpha^s + \mathbb{N}^{S(\alpha^s)}). \tag{92}$$

Finally we define the discrete subset

$$D^{s,1} := \{(\alpha, \epsilon_\rho) \in \Delta \times \mathbb{N}^r;\ \alpha \in \Delta,\ \rho \notin S(\alpha)\} \subset \Delta \times \mathbb{N}^r \tag{93}$$

whose coincidence with D^s will be shown in the next theorem.

Due to Corollary 3.4 and (79) there is, for each $(\alpha, \epsilon_\rho) \in D^{s,1}$ or for each $\alpha \in \Delta$ and $\rho \notin S(\alpha)$, a unique basis representation

$$s_\rho s^{\mu(\alpha)}\overline{\delta_{i(\alpha)}} = \sum_{\beta\in\Delta} f_{\alpha,\rho,\beta} s^{\mu(\beta)}\overline{\delta_{i(\beta)}} \in M \text{ with } f_{\alpha,\rho,\beta} \in F[s^{(\beta)}] \text{ or}$$
$$(s_\rho s^{\mu(\alpha)}\delta_{i(\alpha)})_{\mathrm{nf}} = \sum_{\beta\in\Delta} f_{\rho,\alpha,\beta} s^{\mu(\beta)}\delta_{i(\beta)} \in \mathcal{D}^{1\times I}. \tag{94}$$

For each $(\alpha, \epsilon_\rho) \in D^{s,1}$ define

$$g_{(\alpha,\epsilon_\rho)} := s_\rho\delta_\alpha - \sum_{\beta\in\Delta} f_{\alpha,\rho,\beta}\delta_\beta \in \mathcal{D}^{1\times\Delta}. \tag{95}$$

Theorem and Definition 5.12 (The canonical state representation) *Assume that* $<$ *on* $I \times \mathbb{N}^r$ *is a term order, that* $U \subset \mathcal{D}^{1\times I}$ *is a submodule with its associated* $\mathcal{F}$*-behavior* $\mathcal{B} := U^\perp \subset \mathcal{F}^I$ *where* $\mathcal{F}$ *is any* $\mathcal{D}$*-module and the other data just introduced.*

1. *The module* U^s *and the behavior* $\mathcal{B}^s$ *have state form with*

$$D^s := \min_{\mathrm{cw}}(\deg(U^s)) = \{(\alpha, \epsilon_\rho) \in \Delta \times \mathbb{N}^r;\ \alpha \in \Delta,\ \rho \notin S(\alpha)\},$$
$$\Delta^s = \Delta \times \{0\} \text{ and } S(\alpha, 0) = S(\alpha) \text{ for all } \alpha \in \Delta.$$

2. *The vectors* $g_{(\alpha,\epsilon_\rho)} \in \mathcal{D}^{1\times\Delta}$ *from* (95) *are the unique reduced Gröbner basis of* U^s *with* $\deg(g_{(\alpha,\epsilon_\rho)}) = (\alpha, \epsilon_\rho)$ *for all* $(\alpha, \epsilon_\rho) \in D^s$.
3. *The maps*

$$\mathcal{B}^s \underset{E\circ}{\overset{C\circ}{\rightleftarrows}} \mathcal{B},\quad w^s \to (w^s_{(i,0)})_{i\in I},\quad (s^{\mu(\alpha)} \circ w_{i(\alpha)})_{\alpha\in\Delta} \leftarrow w.$$

are behavior isomorphisms inverse of each other.

4. *The preceding isomorphisms respect the IO-structures from* (71) *and* (83), *i.e., the maps*

$$\mathcal{B}^{s,0} \underset{E^0\circ}{\overset{C^0\circ}{\rightleftarrows}} \mathcal{B}^0,\quad x \to (x_{(i,0)})_{i\in I^0},\quad (s^{\mu(\alpha)} \circ y_{i(\alpha)})_{\alpha\in\Delta^0} \leftarrow y,$$

are inverse isomorphisms and the diagram

$$\begin{array}{ccccccc}
0 \longrightarrow & \mathcal{B}^{s,0} & \stackrel{\binom{\mathrm{id}}{0}\circ}{\longrightarrow} & \mathcal{B}^{s} & \stackrel{(0,\mathrm{id})\circ}{\longrightarrow} & \mathcal{F}^{\Delta^{\mathrm{free}}} & \longrightarrow 0\\
 & E^0\circ \uparrow\downarrow\ C^0\circ & & E\circ \uparrow\downarrow\ C\circ & & \uparrow \mathrm{id} \downarrow & \\
0 \longrightarrow & \mathcal{B}^{0} & \stackrel{\binom{\mathrm{id}}{0}\circ}{\longrightarrow} & \mathcal{B} & \stackrel{(0,\mathrm{id})\circ}{\longrightarrow} & \mathcal{F}^{I^{\mathrm{free}}} & \longrightarrow 0\\
 & x & \mapsto & \binom{x}{0},\ \binom{x}{u} & \mapsto & u & \\
 & \downarrow & & \downarrow & & & \\
 & C^0\circ x & \mapsto & \binom{C^0\circ x}{0},\ \binom{C^0\circ x}{u} = C\circ\binom{x}{u} & \mapsto & u &
\end{array}$$

with vertical isomorphisms is commutative. If $\mathcal{F}$ is injective the rows are exact, otherwise they are exact only up to the module $\mathcal{F}^{I^{\mathrm{free}}} = \mathcal{F}^{\Delta^{\mathrm{free}}}$ without the zero module on the right.

The system $\mathcal{B}^s$ is controllable, i.e., M is torsionfree, if and only if $\mathcal{B}$ has this property. The constructions are valid for arbitrary M and $\mathcal{F}$.

The module M^s and the behavior $\mathcal{B}^s$ with the two isomorphisms from (88) to (89) are called the canonical state representation *of $M = \mathcal{D}^{1\times I}/U$ resp. $\mathcal{B}$. They depend on U and the chosen term order $<$ on $I \times \mathbb{N}^r$ only.*

Remark 5.13

1. The mutually inverse maps $C\circ$ and $E\circ$ are of utmost simplicity. The isomorphism

$$y \mapsto x = E^0 \circ y = (s^{\mu(\alpha)} \circ y_{i(\alpha)})_{\alpha\in\Delta^0}$$

expresses the state components as *suitable* derivatives of the output components (in the continuous case) and is called a *state map* in [31, (4.2), Section 6] in the one-dimensional continuous case. The proto type of this representation is the replacement of a d-th order differential equation $y^{(d)} + a_{d-1}y^{(d-1)} + \cdots + a_0 y = u$ by the equivalent matrix equation

$$s \circ (y, y^{(1)}, \cdots, y^{(d-1)})^\top = Ax + (0, 0, \cdots, 0, u)^\top$$

where A denotes the companion matrix of the polynomial $s^d + a_{d-1}s^{d-1} + \cdots + a_0$. The injectivity of the maps

$$\begin{pmatrix} x \\ u \end{pmatrix} \mapsto C \circ \begin{pmatrix} x \\ u \end{pmatrix} = \begin{pmatrix} (x_{(i,0)})_{i\in I^0} \\ u \end{pmatrix} \text{ and } x \mapsto C^0 \circ x = (x_{(i,0)})_{i\in I^0}$$

signifies that the state x is *observable* from the input u and the output y (compare [20, (62)]).
2. The algorithm of the theorem has been programmed in MAPLE for the two-dimensional situation whereas that of Theorem 5.14 and the higher dimensional case have not yet been implemented.

Proof

1. We show first that no $(\alpha, 0)$, $\alpha \in \Delta$, is contained in $N^s = \deg(U^s)$, ie. that M^s is minimally embedded. Indeed, assume that $(\alpha, 0) \in \deg(U^s)$. This signifies that there is a vector

$$f = \delta_\alpha - \sum_{\beta\in\Delta,\nu\in\mathbb{N}^r} f_{\beta,\nu} s^\nu \delta_\beta \in U^s,\ f_{\beta,\nu} \in F,\ \text{with}$$

$$(\beta, \nu) < (\alpha, 0) \text{ for all } (\beta, \nu) \text{ with } f_{\beta,\nu} \neq 0.$$

By definition (90) the relation

$$(\beta, \nu) < (\alpha, 0) \text{ implies } \beta + \nu < \alpha + 0 = \alpha \text{ or } \beta + \nu = \alpha \text{ and } \mu(\alpha) <_{\text{lex}} \mu(\beta).$$

In the second case this would mean $\mu(\alpha) = \mu(\beta) + \nu <_{\text{lex}} \mu(\beta)$. Since $<_{\text{lex}}$ is also a term order this is a contradiction. Hence the second case cannot occur and therefore

$$\beta + \nu = (i(\beta), \mu(\beta) + \nu) < \alpha = (i(\alpha), \mu(\alpha)) \text{ for all } (\beta, \nu) \text{ with } f_{\beta,\nu} \neq 0.$$

Since $U^s E \subset U$ by construction we obtain

$$fE = s^{\mu(\alpha)}\delta_{i(\alpha)} - \sum_{\beta\in\Delta,\nu\in\mathbb{N}^r} f_{\beta,\nu} s^{\mu(\beta)+\nu}\delta_{i(\beta)} \in U \text{ with}$$

$$(i(\beta), \mu(\beta) + \nu) < (i(\alpha), \mu(\alpha)) \text{ for all } (\beta, \nu) \text{ with } f_{\beta,\nu} \neq 0,\ \text{hence}$$

$$\alpha = (i(\alpha), \mu(\alpha)) = \deg(fE) \in N = \deg(U) \text{ in contradiction to } \alpha \in \Delta.$$

Therefore $(\alpha, 0) \notin N^s$ for all $\alpha \in \Delta$.

2. We show that $D^{s,1} = \{(\alpha, \epsilon_\rho); \alpha \in \Delta, \rho \notin S(\alpha))\}$ is contained in N^s.

1. Case: $\alpha + \epsilon_\rho = \beta + \mu \in \Gamma$ with $\beta \in \Delta$, $\mu \in \mathbb{N}^{S(\beta)}$ and $i(\alpha) = i(\beta)$: From Corollary 2.3,(3) , we infer $\mu(\alpha) <_{\text{lex}} \mu(\beta)$. The relation

$$s_\rho s^{\mu(\alpha)}\overline{\delta_{i(\alpha)}} = s^{\mu(\beta)+\mu}\overline{\delta_{i(\beta)}} \in M \text{ implies } g_{(\alpha,\epsilon_\rho)} = s_\rho\delta_\alpha - s^\mu\delta_\beta.$$

The relations $\alpha + \epsilon_\rho = \beta + \mu$ and $\mu(\alpha) <_{\text{lex}} \mu(\beta)$ imply $(\beta, \mu) < (\alpha, \epsilon_\rho)$ with respect to the term order (90), and hence $(\alpha, \epsilon_\rho) = \deg(g_{(\alpha,\epsilon_\rho)}) \in N^s = \deg(U^s)$.

2. Case: $\alpha + \epsilon_\rho \in N$: Choose a representation $\alpha + \epsilon_\rho = d + \mu$ with $d \in D$, $\mu \in \mathbb{N}^r$ and $i(d) = i(\alpha)$. Equation (80) yields

$$f_d = s^{\mu(d)}\delta_{i(d)} - \sum_{\beta\in\Delta,\ \nu\in\mathbb{N}^{S(\beta)}} f_{d,\beta,\nu} s^{\mu(\beta)+\nu}\delta_{i(\beta)} \in U,\ \ f_{d,\beta,\nu} \in F,$$

with $\deg(f_d) = d$ and hence $\beta + \nu < d$ for all (β, ν) with $f_{d,\beta,\nu} \neq 0$. Multiplication of f_d with s^μ gives

$$s^\mu f_d = s_\rho s^{\mu(\alpha)}\delta_{i(\alpha)} - \sum_{\beta\in\Delta,\ \nu\in\mathbb{N}^{S(\beta)}} f_{d,\beta,\nu} s^{\mu(\beta)+\nu+\mu}\delta_{i(\beta)} \in U$$

with $\deg(s^\mu f_d) = d + \mu = \alpha + \epsilon_\rho$ and $\beta + \nu + \mu < d + \mu = \alpha + \epsilon_\rho$ for all (β, ν) with $f_{d,\beta,\nu} \neq 0$. By definition of U^s the vector

$$g := s_\rho \delta_\alpha - \sum_{\beta \in \Delta,\ \nu \in \mathbb{N}^{S(\beta)}} f_{d,\beta,\nu} s^{\nu+\mu} \delta_\beta \text{ is in } U^s \text{ and } \deg(g) = (\alpha, \epsilon_\rho)$$

since $\beta + \nu + \mu < d + \mu = \alpha + \epsilon_\rho$ implies $(\beta, \nu + \mu) < (\alpha, \epsilon_\rho)$ with respect to the term order (90). This proves $(\alpha, \epsilon_\rho) \in N^s = \deg(U^s)$ in the second case and therefore $D^{s,1} \subset N^s$.

3. From 1. and 2. we conclude that $D^{s,1}$ is a cw-discrete set of cw-minimal elements of $N^s = \deg(U^s)$, hence

$$D^{s,1} \subset D^s,\ N^{s,1} := D^{s,1} + \mathbb{N}^r \subset N^s = D^s + \mathbb{N}^r,$$
$$\Gamma^s := (\Delta \times \mathbb{N}^r) \setminus N^s \subset \Gamma^{s,1} := (\Delta \times \mathbb{N}^r) \setminus N^{s,1} = \cup_{\alpha \in \Delta}((\alpha, 0) + \mathbb{N}^{S(\alpha)})$$

where the form of $\Gamma^{s,1}$ follows as in Theorem 5.2, (2). According to Corollary 3.4 the vectors

$$s^\mu \overline{\delta_\alpha},\ (\alpha, \mu) \in \Gamma^s, \text{ are an } F\text{-basis of } M^s.$$

On the other hand, the module M has the F-basis $s^\mu s^{\mu(\alpha)} \overline{\delta_{i(\alpha)}},\ \alpha \in \Delta,\ \mu \in \mathbb{N}^{S(\alpha)}$, which, by means of the isomorphism

$$(\circ E)_{\text{ind}} : M^s \cong M, \text{ yields the } F\text{-basis } s^\mu \overline{\delta_\alpha},\ (\alpha, \mu) \in \Gamma^{s,1}, \text{ of } M^s.$$

Since $\Gamma^s \subset \Gamma^{s,1}$ we infer that indeed

$$\Gamma^{s,1} = \Gamma^s,\ N^{s,1} = N^s,\ D^{s,1} = D^s \text{ and}$$
$$\Delta^s = \Delta \times \{0\} \text{ with } S(\alpha, 0) = S(\alpha).$$

Hence U^s is indeed a state module to which all considerations of Theorem 5.2 are applicable.

4. As in 5.2, (3), let

$$h_{(\alpha,\epsilon_\rho)} := s_\rho \delta_\alpha - \sum_{\beta \in \Delta} h_{\alpha,\rho,\beta} \delta_\beta \in \mathcal{D}^{1 \times \Delta},\ (\alpha, \epsilon_\rho) \in D^s,\ h_{\alpha,\rho,\beta} \in F[s^{(\beta)}],$$

be the unique reduced Gröbner basis of U^s. Multiplication with E on the right and taking residue classes in M gives

$$0 = \overline{h_{(\alpha,\epsilon_\rho)} E} = s_\rho s^{\mu(\alpha)} \overline{\delta_{i(\alpha)}} - \sum_{\beta \in \Delta} h_{\alpha,\rho,\beta} s^{\mu(\beta)} \overline{\delta_{i(\beta)}} \text{ or}$$
$$s_\rho s^{\mu(\alpha)} \overline{\delta_{i(\alpha)}} = \sum_{\beta \in \Delta} h_{\alpha,\rho,\beta} s^{\mu(\beta)} \overline{\delta_{i(\beta)}}.$$

Comparison with (94) and (95) furnishes $h_{\rho,\alpha,\beta} = f_{\rho,\alpha,\beta}$ and $h_{(\alpha,\epsilon_\rho)} = g_{(\alpha,\epsilon_\rho)}$ for all $(\alpha, \epsilon_\rho) \in D^s$.

5. The third assertion of the theorem was already proved in Theorem 5.9.
6. The last assertion concerning the IO-structures follows from the block form of the matrices C and E according to (85). The IO-decompositions of the vectors w^s resp. w are

$$w = \begin{pmatrix} y \\ u \end{pmatrix} = C \circ w^s = \begin{pmatrix} C^0 \circ x \\ u^s \end{pmatrix}, \text{ hence } u^s = u.$$

The input of $\mathcal{B}^s$ and $\mathcal{B}$ coincide as they should. □

We finally construct the normal state representation of $\mathcal{B}$ by combining the preceding theorem with Theorem 4.5. For this purpose we define the canonical matrices $R = (P, -Q) \in \mathcal{D}^{D\times(I^0 \uplus I^{\text{free}})}$ of $\mathcal{B}$ and its transfer matrix $H \in F(s)^{I^0\times I^{\text{free}}}$ with $PH = Q$ according to (60) and do the same for the state behavior $\mathcal{B}^s$ where we always identify $\Delta^{\text{free}} = I^{\text{free}}$, $(i, 0) = i$:

$$\begin{aligned}
R^s &= (P^s, -Q^s) \in F[s]^{D^s\times(\Delta^0 \uplus I^{\text{free}})} \text{ where for all } (\alpha, \epsilon_\rho) \in D^s \\
P^s_{(\alpha,\epsilon_\rho)-} &:= s_\rho \delta_\alpha - \sum_{\beta\in\Delta^0} f_{\rho,\alpha,\beta}\delta_\beta \in \mathcal{D}^{1\times\Delta^0}, \;\; f_{\rho,\alpha,\beta} \in F[s^{(\beta)}], \\
P^s_{(\alpha,\epsilon_\rho),\beta} &:= s_\rho \delta_{\alpha,\beta} - f_{\rho,\alpha,\beta} \text{ for all } \beta \in \Delta^0 \\
Q^s_{(\alpha,\epsilon_\rho)-} &:= \sum_{i\in I^{\text{free}}} f_{\alpha,\rho,(i,0)}\delta_{(i,0)} \in \mathcal{D}^{1\times I^{\text{free}}}, \;\; f_{\alpha,\rho,(i,0)} \in F[s], \\
Q^s_{(\alpha,\epsilon_\rho),i} &:= f_{\alpha,\rho,(i,0)} \text{ for all } i \in I^{\text{free}} H^s \in F(s)^{\Delta^0\times I^{\text{free}}} \text{ with } P^s H^s = Q^s.
\end{aligned} \tag{96}$$

The transfer matrix $H^s \in F(s)^{\Delta^0\times I^{\text{free}}}$ of $\mathcal{B}^s$ is uniquely defined by the equation $P^s H^s = Q^s$ and satisfies $C^0 H^s = H$ according to [20, Th. 82].

As in Theorem 4.5 and (68) we choose a term order on $D^s \times \mathbb{N}^r$ and obtain the unique normal form decomposition

$$Q^s = P^s X + B \text{ with } X := X_{Q^s} \in \mathcal{D}^{\Delta^0\times I^{\text{free}}} \text{ and } B := (Q^s)_{\text{nf}} \in \mathcal{D}^{D^s\times I^{\text{free}}}. \tag{97}$$

Theorem and Definition 5.14 (The normal state representation) *With the preceding data define the IO-behavior*

$$\mathcal{B}^s_{\text{nf}} := \left\{ \begin{pmatrix} x \\ u \end{pmatrix} \in \mathcal{F}^{\Delta^0 \uplus I^{\text{free}}};\;\; P^s \circ x = B \circ u \right\}. \textit{ Then}$$

$$\begin{pmatrix} C^0 & D^0 \\ 0 & \text{id} \end{pmatrix} \circ : \mathcal{B}^s_{\text{nf}} \to \mathcal{B}, \; \begin{pmatrix} x \\ u \end{pmatrix} \mapsto \begin{pmatrix} C^0 \circ x + D^0 \circ u \\ u \end{pmatrix}, \textit{ with } D^0 := C^0 X$$

is an isomorphism of IO-behaviors. The inverse of this isomorphism, called state map in [31, (4.2)], *is given by* $\binom{y}{u} \mapsto \binom{E^0\circ y - X\circ u}{u}$. *The behavior* $\mathcal{B}^s_{\text{nf}}$ *is called the* normal state representation *of the behavior* $\mathcal{B}$. *This representation depends on the module* U *of* $\mathcal{B}$ *and the two chosen term orders on* $I \times \mathbb{N}^r$ *and on* $D^s \times \mathbb{N}^r$.

Proof The asserted isomorphism is the composition of the isomorphisms from Theorems 4.5 and 5.12:

$$\begin{pmatrix} C^0 & D^0 \\ 0 & \mathrm{id} \end{pmatrix} \circ = \begin{pmatrix} C^0 & C^0 X \\ 0 & \mathrm{id} \end{pmatrix} \circ : \mathcal{B}^s_{\mathrm{nf}} \overset{\begin{pmatrix} \mathrm{id} & X \\ 0 & \mathrm{id} \end{pmatrix}\circ}{\longrightarrow} \mathcal{B}^s \overset{\begin{pmatrix} C^0 & 0 \\ 0 & \mathrm{id} \end{pmatrix}\circ}{\longrightarrow} \mathcal{B}.$$

□

Example 5.15 (*The one-dimensional case*) Let $\mathcal{D} := F[s]$ be the one-dimensional polynomial algebra, $U \subset \mathcal{D}^{1\times I}$ a submodule such that $M := \mathcal{D}^{1\times I}/U$ is minimally embedded in the sense of Definition 3.13 and $\mathcal{B} := \{w \in \mathcal{F}^I;\ U \circ w = 0\}$ the associated one-dimensional behavior with the transfer matrix H. We choose one of the term orders from (48) or (49). Gröbner basis computations yield

$$\begin{aligned} &\deg(U) = \uplus_{i\in I^0}\{i\} \times (d(i) + \mathbb{N}),\ I = I^0 \uplus I^{\mathrm{free}}, \\ &\quad \Gamma = \Delta^0 \uplus (\Delta^{\mathrm{free}} + \mathbb{N}) \text{ where } \Delta^0 := \{(i,\mu);\ i \in I^0,\ 0 \le \mu \le d(i)-1\} \text{ and} \\ &\quad I^{\mathrm{free}} \cong \Delta^{\mathrm{free}},\ i \mapsto (i,0),\ \text{and the reduced Gröbner basis} \\ &\quad f_i := f_{(i,d(i))} = s^{d(i)}\delta_i - \sum_{(j,\nu)\in\Delta^0} f_{i,j,\nu} s^\nu \delta_j - \sum_{j\in I^{\mathrm{free}}} f_{i,j}\delta_j,\ i \in I^0, \\ &\qquad \text{with } f_{i,j,\nu} \in F,\ \ f_{i,j} \in F[s]. \end{aligned} \tag{98}$$

The assumed minimal embeddedness implies

$$d(i) > 0 \text{ for all } i \in I^0 \text{ or } (i,0) \in \Delta \text{ for all } i \in I.$$

We identify $I^{\mathrm{free}} = \Delta^{\mathrm{free}},\ i = (i,0)$. The canonical IO-structure for the chosen term order is

$$I = I^0 \uplus I^{\mathrm{free}} \text{ or } w_i =: \begin{cases} y_i = \text{output} & \text{if } i \in I^0 \\ u_i = \text{input} & \text{if } i \in I^{\mathrm{free}} \end{cases}.$$

In addition to being a Gröbner basis of U the $f_i,\ i \in I^0$, are an $F[s]$-basis of U, i.e., $F[s]$-linearly independent.

According to Theorem 5.12 and equation (95) the state vector resp. the state system $\mathcal{B}^s$ are given as

$$\begin{aligned} x &= (x_{j,\nu})_{(j,\nu)\in\Delta^0} \in \mathcal{F}^{\Delta^0} \\ \mathcal{B}^s &:= \left\{ \begin{pmatrix} x \\ u \end{pmatrix} \in \mathcal{F}^{\Delta^0 \uplus I^{\mathrm{free}}};\ \begin{pmatrix} x \\ u \end{pmatrix} \text{ satisfies} \right. \\ s \circ x_{i,\mu} &= \begin{cases} x_{i,\mu+1} \\ \sum_{(j,\nu)\in\Delta^0} f_{i,j,\nu} x_{j,\nu} + \sum_{j\in I^{\mathrm{free}}} f_{i,j} \circ u_j \end{cases} \quad \text{if} \\ &\qquad \left. \begin{cases} i \in I^0, \mu < d(i)-1 \\ i \in I^0, \mu = d(i)-1 \end{cases} \right\}. \end{aligned} \tag{99}$$

We define the matrices $A \in F^{\Delta^0 \times \Delta^0}$, $Q^s \in F[s]^{\Delta^0 \times I^{\mathrm{free}}}$ by

$$A_{\alpha,\beta} := \begin{cases} \delta_{\alpha+1,\beta} \\ f_{i,j,\nu} \end{cases} \quad \text{if } \beta = (j,\nu) \in \Delta^0 \text{ and } \alpha = \begin{cases} (i,\mu),\ i \in I^0,\ \mu < d(i)-1 \\ (i,d(i)-1),\ i \in I^0 \end{cases}$$

$$Q^s_{\alpha j} := \begin{cases} 0 \\ f_{i,j} \end{cases} \quad \text{if } j \in I^{\mathrm{free}} \text{ and } \alpha = \begin{cases} (i,\mu),\ i \in I^0,\ \mu < d(i)-1 \\ (i,d(i)-1),\ i \in I^0 \end{cases}. \tag{100}$$

With these matrices the representation (99) gets the form

$$\mathcal{B}^s = U^{\perp}_{\mathcal{F}} = \left\{ \begin{pmatrix} x \\ u \end{pmatrix} \in \mathcal{F}^{\Delta^0 \uplus I^{\mathrm{free}}};\ P^s \circ x = Q^s \circ u \right\}$$

$$\text{with } P^s := s\,\mathrm{id}_{\Delta^0} - A \in F[s]^{\Delta^0 \times \Delta^0},\ Q^s \in F[s]^{\Delta^0 \times I^{\mathrm{free}}},$$

$$\mathcal{B}^s \cong \mathcal{B},\ \begin{pmatrix} x \\ u \end{pmatrix} \leftrightarrow \begin{pmatrix} y \\ u \end{pmatrix},\ x = E^0 y,\ y := C^0 x,\ \text{where}$$

$$y_i := x_{i,0},\ x_{i,\mu} = s^{\mu} \circ y_i,\ C^0 \in F[s]^{I^0 \times \Delta^0},\ E^0 \in F[s]^{\Delta^0 \times I^0}$$

$$C^0_{i,(j,\nu)} = \begin{cases} 1 & \text{if } i = j,\ \nu = 0 \\ 0 & \text{otherwise} \end{cases},\quad E^0_{(i,\mu),j} = \begin{cases} s^{\mu} & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}. \tag{101}$$

As in Theorem 5.14 we finally choose the graded term order on $\Delta^0 \times \mathbb{N}$ according to the second possibility of (49) and apply this to the free *column* module $F[s]^{\Delta^0}$. This choice implies that the column $P^s_{-\beta} = s\delta_\beta - A_{-\beta},\ \beta \in \Delta^0$, has the leading term $s\delta_\beta$ and the degree $(\beta, 1)$. According to [16, Th. 2.55] the columns of $P^s = s\,\mathrm{id}_{\Delta^0} - A$ are the unique reduced Gröbner basis of the column module $P^s F[s]^{\Delta^0}$ since no S-vectors have to be formed. This implies

$$\deg(P^s F[s]^{\Delta^0}) = \Delta^0 \times (1 + \mathbb{N}) \text{ and}$$

$$F[s]^{\Delta^0} = F^{\Delta^0} \oplus P^s F[s]^{\Delta^0} = \oplus_{\beta \in \Delta^0} F\delta_\beta \oplus P^s F[s]^{\Delta^0}. \tag{102}$$

In particular, each normal form vector with respect to $P^s = s\,\mathrm{id}_{\Delta^0} - A$ has entries in the base field F. According to (97) we obtain

$$Q^s = P^s X + B \text{ with } X \in F[s]^{\Delta^0 \times I^{\mathrm{free}}} \text{ and}$$

$$B := (Q^s)_{\mathrm{nf}} \in F^{\Delta^0 \times I^{\mathrm{free}}},\ X = (s\,\mathrm{id} - A)^{-1}(Q^s - B). \tag{103}$$

The normal ISO-representation of $\mathcal{B}$ is thus given by the Kalman state system

$$\mathcal{B}^s_{\mathrm{nf}} := \left\{ \begin{pmatrix} x \\ u \end{pmatrix} \in \mathcal{F}^{\Delta^0 \uplus I^{\mathrm{free}}};\ s \circ x = Ax + Bu \right\}$$

and the isomorphism, compare (3),

$$\begin{pmatrix} C^0 & D^0 \\ 0 & \mathrm{id} \end{pmatrix} \circ : \mathcal{B}^s_{\mathrm{nf}} \cong \mathcal{B},\ \begin{pmatrix} x \\ u \end{pmatrix} \leftrightarrow \begin{pmatrix} y \\ u \end{pmatrix},\ \text{with } D^0 := C^0 X,$$

$$y = C^0 x + D^0 \circ u,\ x = E^0 \circ y - X \circ u. \tag{104}$$

The map $\binom{y}{u} \mapsto x = E^0 \circ y - X \circ u$ is again the *state map* according to [31, (4.2)]. Due to the system isomorphism (104) the Kalman system

$$s \circ x = Ax + Bu, \ \ y = C^0 x + D^0 \circ u$$

is observable. It is also controllable if and only if $\mathcal{B}$ has this property.

The transfer matrix of $\mathcal{B}$ is $H = D^0 + C^0(s\,\mathrm{id} - A)^{-1}B$. Since the first summand of H is the polynomial matrix $D^0 = C^0 X$ and the second is strictly proper this representation of H is the unique one coming from the direct sum decomposition $F(s) = F[s] \oplus F(s)_{\mathrm{spr}}$ where $F(s)_{\mathrm{spr}}$ denotes the strictly proper rational functions $\frac{a(s)}{b(s)}$ with $\deg(a) < \deg(b)$. In particular, the matrix D^0 has entries in F if and only if H is proper. In contrast to Kalman who considered proper transfer matrices only Wolovich [41, p. 137] and Vardulakis [37, Def. 2.8] consider the system isomorphism (104) a state space representation also for non-proper H or non-constant D^0.

All data in (104) follow from the Gröbner basis computations of the f_i and of $B = Q^s_{\mathrm{nf}}$. It seems to us that the computations of the present example compare very well with those of [14, section 6.7], [25, Th. 6.4.2], [31, section 6], [37, section 2.2.1], [41, Th. 5.2.14].

Example 5.16 (*The two-dimensional case*) The following MAPLE-generated example has been constructed according to Remark 5.8 over the polynomial algebra $\mathcal{D} := F[s_1, s_2]$ where $F := \mathbb{Z}/\mathbb{Z}3 = \{0, 1, 2\}$ is the small finite field with three elements. It applies to all $\mathcal{D}$-modules $\mathcal{F}$, not only to injective cogenerators or special function modules. The small field has been chosen to keep the numerical constants small. The same example has been tested over the finite field $\mathbb{Z}/\mathbb{Z}97$ and over the rational number field $\mathbb{Q}$ and it turns out that the general properties of the behaviors are the same whereas the numerical values of the derived matrices are, of course, more complicated. In particular, the example can be considered as a system of partial differential equations.
The matrices

$$A := \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \ B := \begin{bmatrix} 2 & 0 & 2 \\ 2 & 0 & 1 \end{bmatrix}, \ C := \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, \ D := \begin{bmatrix} 2 & 2 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

and $\Delta(s) := \begin{pmatrix} s_1 & 0 \\ 0 & s_2 \end{pmatrix}$ give rise to the second resp. first order polynomial matrices

$$\begin{aligned} R &:= D + C\Delta(s)(\mathrm{id} - A\Delta(s))^{-1}B \\ &= \begin{bmatrix} 2 + 2s_1 + 2s_1s_2 + 2s_2 & 2 & 2 + 2s_1 + 2s_1s_2 + s_2 \\ 2s_1 + s_1s_2 + s_2 & 0 & 2s_1 + 2s_2 + s_1s_2 \end{bmatrix} \end{aligned}$$

and

$$R_{\mathrm{lin}} := \begin{pmatrix} \mathrm{id} - A\Delta(s) & -B \\ C\Delta(s) & D \end{pmatrix} = \begin{bmatrix} 1 & 0 & -2 & & -2 \\ -s_1 & 1 & -2 & 0 & -1 \\ s_1 & s_2 & 2 & 2 & 2 \\ s_1 & 2s_2 & 0 & 0 & 0 \end{bmatrix},$$

the behaviors

$$\mathcal{B}_{\text{lin}} := \{w_{\text{lin}} = (w_{\text{lin},1}, w_{\text{lin},2}, w_{\text{lin},3}, w_{\text{lin},4}, w_{\text{lin},5})^\top \in \mathcal{F}^5;\ R_{\text{lin}} \circ w_{\text{lin}} = 0\}$$
$$\mathcal{B} := \{w = (w_1, w_2, w_3)^\top \in \mathcal{F}^3;\ R \circ w = 0\}$$

and the mutually inverse behavior isomorphisms

$$\text{proj} = (0_2\ \text{id}_3)\circ : \mathcal{B}_{\text{lin}} \to \mathcal{B},\ w_{\text{lin}} \mapsto w := (w_{\text{lin},3}, w_{\text{lin},4}, w_{\text{lin},5})^\top \text{ and}$$
$$\begin{pmatrix} (\text{id}_2 - A\Delta(s))^{-1}B \\ \text{id}_3 \end{pmatrix} \circ : \mathcal{B} \mapsto \mathcal{B}_{\text{lin}}$$
$$w \mapsto (2w_1 + 2w_3, (2s_1 + 1) \circ w_1 + (2s_1 + 1) \circ w_3, w_1, w_2, w_3)^\top$$
$$\text{where } (\text{id}_2 - A\Delta(s))^{-1}B = \begin{bmatrix} 2 & 0 & 2 \\ 2s_1 + 2 & 0 & 2s_1 + 1 \end{bmatrix}.$$

We have applied the algorithm of Theorem 5.12 to several thousand computer generated examples and in particular to the simple behaviors $\mathcal{B}$ and $\mathcal{B}_{\text{lin}}$. That of Theorem 5.14 is not yet implemented, but will bring a further improvement. We have used the six MAPLE term orders

$$D1 := \textit{poly\_algebra}\,(s_0, s_1, s_2, \textit{characteristic} = 3)\,;$$
$$T_1 := \textit{termorder}\,(D1, \textit{lexdeg}\,([s_1, s_2], [s_0])\,, [s_0])\,;$$
$$T_2 := \textit{termorder}\,(D1, \textit{lexdeg}\,([s_2, s_1], [s_0])\,, [s_0])\,;$$
$$T_3 := \textit{termorder}\,(D1, \textit{plex}\,(s_1, s_2, s_0)\,, [s_0])\,;$$
$$T_4 := \textit{termorder}\,(D1, \textit{plex}\,(s_2, s_1, s_0)\,, [s_0])\,;$$
$$T_5 := \textit{termorder}\,(D1, \textit{plex}\,(s_0, s_1, s_2)\,, [s_0])\,;$$
$$T_6 := \textit{termorder}\,(D1, \textit{plex}\,(s_0, s_2, s_1)\,, [s_0])$$

in order to test the dependence of the algorithms on the term orders. In most cases and as expected the matrices of lowest degree and therefore the simplest state representations are obtained for the graded term orders T_1 and T_2 whereas T_5 and T_6 often lead to state representations of very high polynomial degree.

The state system $\mathcal{B}^s$ of $\mathcal{B}$ with respect to the graded term-order T_1 is computed below. In this case we use the identification

$$\delta_1 = (1, 0, 0) = s_0^2,\ \delta_2 = (0, 1, 0) = s_0, \delta_3 = (0, 0, 1) = s_0^0 \in F[s_1, s_2]^{1\times 3} \subset F[s_0, s_1, s_2].$$

The reduced Gröbner basis of the row module of R with respect to T_1 consists of the three vectors

$$1 + s_0^2 + 2s_0s_2 + s_0^2s_2 + s_0 = (s_2 + 1, 2s_2 + 1, 1)$$
$$2 + s_1 + 2s_0^2 + 2s_0 + s_0^2s_1 = (s_1 + 2, 2, s_1 + 2)$$
$$s_1s_2 + s_0s_1s_2 + 2s_0s_1 + s_0s_2 + 2 = (0, s_1s_2 + 2s_1 + s_2, s_1s_2 + 2)$$
$$\text{of degrees } (1, 0, 1), (1, 1, 0), (2, 1, 1) \in \{1, 2, 3\} \times \mathbb{N}^2.$$

The algorithm of equations (24), (25) easily yields the table

$$\Delta = \{(1,0,0), (2,0,0), (2,1,0), (3,0,0)\}$$

α	$(1,0,0)$	$(2,0,0)$	$(2,1,0)$	$(3,0,0)$
$S(\alpha)$	$\emptyset$	$\{2\}$	$\{1\}$	$\{1,2\}$
$\{1,2\}\setminus S(\alpha)$	$\{1,2\}$	$\{1\}$	$\{1\}$	$\emptyset$

The canonical IO-structure according to Theorem 4.3, the set I^Γ after (50), the output y and input u of $\mathcal{B}$ and the state vector $x = (x_\alpha)_{\alpha\in\Delta}$ of $\mathcal{B}^s$ according to Theorem 5.12, (4), are thus given as

$$I := \{1,2,3\} = I^0 \uplus I^{\text{free}} = \{1,2\} \uplus \{3\},\ \ \Delta^{\text{free}} = \{(3,0,0)\} \text{ and } I^\Gamma = I$$

$$y := (w_1, w_2)^\top,\ u := w_3,\ x = (x_{(1,0,0)}, x_{(2,0,0)}, x_{(2,1,0)}, x_{(3,0,0)})^\top.$$

In particular, the system $\mathcal{B}$ is minimally embedded and the state representation Theorem 5.12 is directly applicable. The matrices C_R and E_R according to (84) are

$$C_R := \begin{bmatrix} 0 & x_{1,0,0} & x_{2,0,0} & x_{3,0,0} & x_{2,1,0} \\ w_1 & 1 & 0 & 0 & 0 \\ w_2 & 0 & 1 & 0 & 0 \\ w_3 & 0 & 0 & 1 & 0 \end{bmatrix},\ E_R := \begin{bmatrix} 0 & w_1 & w_2 & w_3 \\ x_{1,0,0} & 1 & 0 & 0 \\ x_{2,0,0} & 0 & 1 & 0 \\ x_{3,0,0} & 0 & 0 & 1 \\ x_{2,1,0} & 0 & s_1 & 0 \end{bmatrix}$$

The first row and column of these matrices do not represent matrix entries proper, but the column resp. row indices which explain on which component of w resp. x the corresponding coefficient acts. The mutually inverse isomorphisms $\mathcal{B}^s \underset{E_R\circ}{\overset{C_R\circ}{\rightleftarrows}} \mathcal{B}$ imply

$$x_{(1,0,0)} = w_1 = y_1,\ x_{(2,0,0)} = w_2 = y_2$$

$$x_{(2,1,0)} = s_1 \circ w_2 = s_1 \circ y_2,\ x_{(3,0,0)} = w_3 = u.$$

The following matrices

$$R_1^s := \begin{bmatrix} 0 & x_{1,0,0} & x_{2,0,0} & x_{3,0,0} & x_{2,1,0} \\ s_1x_{1,0,0} & s_1 & 0 & 0 & 0 \\ s_2x_{1,0,0} & s_2 & 0 & 0 & 0 \\ s_1x_{2,0,0} & 0 & s_1 & 0 & 0 \\ s_2x_{2,1,0} & 0 & 0 & 0 & s_2 \end{bmatrix}$$

$$R_2^s := \begin{bmatrix} 0 & x_{1,0,0} & x_{2,0,0} & x_{3,0,0} & x_{2,1,0} \\ s_1x_{1,0,0} & 1 & 1 & 2s_1+1 & 0 \\ s_2x_{1,0,0} & 2 & s_2+2 & 2 & 0 \\ s_1x_{2,0,0} & 0 & 0 & 0 & 1 \\ s_2x_{2,1,0} & 0 & 2s_2 & 2s_1s_2+s_2 & 1 \end{bmatrix}$$

are computed according to (95) where again the first row and column represent the column resp. row indices and where $s_\rho x_\alpha$ stands for the index of $g_{(\alpha,\rho)}$ in (95). According to Theorem 5.12, (2), the rows of $R_1^s - R_2^s$ are the unique reduced Gröbner basis of the state system $\mathcal{B}^s$, in particular

$$\mathcal{B}^s = \{x \in \mathcal{F}^\Delta;\ (R_1^s - R_2^s)\circ x = 0\}.$$

If the $F[s_1, s_2]$-module $\mathcal{F}$ consists of power series in two variables z_1, z_2 as in result 5.6 the isomorphism (73) of the Cauchy problem for $\mathcal{B}^s$ has the form

$$\mathcal{B}^s \cong F \times \mathcal{F}(z_2) \times \mathcal{F}(z_1) \times \mathcal{F},\, x \mapsto (x_{(1,0,0)}(0,0), x_{(2,0,0)}(0, z_2), x_{(2,1,0)}(z_1, 0), x_{(3,0,0)}).$$

The first three components of the latter vector are the local states whereas $u = x_{(3,0,0)}$ is the input. These data gives rise to a unique $x \in \mathcal{B}^s$ and then to a unique output $y = (x_{(1,0,0)}, x_{(2,0,0)})^\top$.

The algorithm of Theorem 3.6 gives the Hilbert function and polynomial

$$\mathrm{HF} := \mathrm{HF}_{\widehat{M}} := [3, 7],\ \ \mathrm{HP} := \mathrm{HP}_{\widehat{M}} := \frac{1}{2}t^2 + \frac{7}{2}t + 3 \text{ where } M := \mathcal{D}^{1\times 3}/\mathcal{D}^{1\times 2}R,$$

$\mathrm{HF}(0) = 3$ and $\mathrm{HF}(n) = \mathrm{HP}(n)$ for $n \geq 1$. In this simple case HF and HP coincide for all arguments $n \in \mathbb{N}$. Since $\mathcal{B}$ has the free component $w_3 = u$ its Krull dimension and degree of HP are 2.

In this particular example the calculations show that the state representation $\mathcal{B}^s_{\mathrm{lin}}$ of $\mathcal{B}_{\mathrm{lin}}$ coincides with $\mathcal{B}^s$ up to a renaming of the components. Remark that, in general, $\mathcal{B}$, $\mathcal{B}^s$, $\mathcal{B}_{\mathrm{lin}}$ and $\mathcal{B}^s_{\mathrm{lin}}$ are isomorphic behaviors. The unique reduced Gröbner basis of $\mathcal{D}^{1\times 4}R_{\mathrm{lin}}$ or $\mathcal{B}_{\mathrm{lin}}$ with respect to the term order T_1 is

$$(0,1,2,1,0),\ \ (1,0,1,0,1),\ \ (0,0,s_2+1, 2s_2+1, 1),$$
$$(0,0,s_1+2, 2, s_1+2),\ \ (0,0,0, s_1s_2+2s_1+s_2, s_1s_2+2s_2).$$

The last vector contains quadratic terms and this property also holds for the coefficient fields $\mathbb{Z}/\mathbb{Z}97$ and $\mathbb{Q}$ and all term orders T_1 to T_6. This is in contrast to the fact that the Gröbner basis with respect to a pure lexicographic term order of an *ideal* generated by first order or linear polynomials is again generated by linear polynomials [16, ex.9, p. 78]. Likewise, for all these fields and term orders the Gröbner matrix of $\mathcal{B}^s_{\mathrm{lin}}$ contains terms of degree at least two. This shows that, in general, the canonical state representation of a first order system need not again be of first order or, in colloquial terms, first order systems may be more complicated than they appear. On the other hand computer experiments have shown that for most first order systems there is at least one term order, most often T_1 or T_2, for which the canonical state representation is again of first order.

Conjecture 5.17 For every behavior $\mathcal{B}$ there is a term order such that the order of (the Gröbner matrix of) its state representation $\mathcal{B}^s$ exceeds that of $\mathcal{B}$ by at most one. The conjecture has been verified in many computer experiments.

Acknowledgements I thank two (for me non-anonymous) referees for their critical reading of this article. The appendix of the talk [22] contains historical comments in context with the judgements of one of the referees.

References

1. Apel, J.: The theory of involutive divisions and an application to Hilbert function computations. J. Symbolic Comput. **25**, 683–704 (1998)
2. Apel, J.: On a conjecture of R.P. Stanley II: quotients modulo monomoial ideals. J. Algebraic Combin. **17**, 57–74 (2003)

3. Blondel, V.D., Megretski, A. (eds.): Unsolved Problems in Mathematical Systems and Control Theory. Princeton University Press, New Jersey (2004)
4. Egorov, Y.V., Shubin, M.A. (eds.): Partial Differential Equations I. Springer, Berlin Heidelberg New York (1992)
5. Egorov, Y.V., Shubin, M.A. (eds.): Partial Differential Equations III. Springer, Berlin Heidelberg New York (1991)
6. Fröhler, S., Oberst, U.: Continuous time-varying linear systems. Systems Control Lett. **35**, 97–110 (1998)
7. Ganzha, V.G., Mayr, E.W., Vorozhtsov, E.V. (eds.): Computer Algebra in Scientific Computing. Springer, Berlin Heidelberg New York (1999)
8. Gelfand, I.M., Shilov, G.E.: Generalized Functions III. Theory of Differential Equations. Academic Press, New York (1967)
9. Gerdt, V.P.: Completion of linear differential systems to involution. In: Ganzha, V.G., Mayr, E.W., Vorozhtsov, E.V. (eds.) Computer Algebra in Scientific Computing, pp. 115–137. Springer, Berlin Heidelberg New York (1999)
10. Gluesing-Luerssen, H.: Linear Delay-Differential Systems with Commensurate Delays: An Algebraic Approach. Lecture Notes in Mathematics 1770, Springer, Berlin Heidelberg New York (2002)
11. Hörmander, L.: The Analysis of Linear Partial Differential Operators I. Springer, Berlin Heidelberg New York (1983)
12. Hörmander, L.: The Analysis of Linear Partial Differential Operators II. Springer, Berlin Heidelberg New York (1983)
13. Janet, M.: Sur les systèmes d'équations aux dérivées partielles. J. Math. Pures Appl. **8**, 65–151 (1920)
14. Kailath, T.: Linear Systems. Prentice-Hall, Englewood Cliffs, New Jersey (1980)
15. Kalman, R.E.: Mathematical description of linear dynamical systems. SIAM J. Control Optim. **1**, 152–192 (1963)
16. Kreuzer, M., Robbiano, L.: Computational Commutative Algebra 1. Springer, Berlin Heidelberg New York (2000)
17. Lu, P., Liu, M., Oberst, U.: Linear recurring arrays, linear systems and multidimensional cyclic codes over Quasi–Frobenius rings. Acta Appl. Math. **80**, 175–198 (2004)
18. Macaulay, F.S.: The Algebraic Theory of Modular Systems. Cambridge University Press, UK (1916)
19. Malgrange, B.: Systèmes Différentiels Involutifs. Prépublication 636, Institut Fourier Grenoble, France (2004)
20. Oberst, U.: Multidimensional constant linear systems. Acta Appl. Math. **20**, 1–175 (1990)
21. Oberst, U.: Variations on the fundamental principle for linear systems of partial differential and difference equations with constant coefficients. AAECC **6**, 211–243 (1995)
22. Oberst, U.: Canonical State Representations and Hilbert Functions of Multidimensional Systems (with historical comments and a discussion of Professor Pommaret's remarks). Talk at the conference D3, Linz, Austria (May 2006)(homepage of the Gröbner-Semester)
23. Oberst, U., Pauer, F.: The constructive solution of linear systems of partial difference and differential equations with constant coefficients. Multidimens. Systems Signal Process. **12**, 253–308 (2001)
24. Palamodov, V.P.: Linear Differential Operators. Springer, Berlin Heidelberg New York (1970)
25. Polderman, J.W., Willems, J.C.: Introduction to Mathematical Systems Theory. Springer, Berlin Heidelberg New York (1998)
26. Pommaret, J.-F.: Systems of Partial Differential Equations and Lie Pseudogroups. Gordon and Breach, New York (1978)
27. Pommaret, J.-F.: New perspectives in control theory for partial differential equations. IMA J. Math. Control Inform. **9**, 305–330 (1992)
28. Pommaret, J.-F.: Partial Differential Control Theory, Volume I: Mathematical Tools, Volume II: Control Systems. Kluwer, Dordrecht (2001)
29. Pommaret, J.-F.: Localization and transfer matrix computation for linear multidimensional control systems. Proceedings MTNS, Leuven, Belgium (2004)
30. Pommaret, J.-F.: Lecture Notes of the mini-course on algebraic analysis of control systems defined by partial differential equations, pp. 1–45. Gröbner Semester Linz, Sections D2 and D3, RISC, Hagenberg, and RICAM, Linz, Austria (May 2006)
31. Rapisarda, P., Willems, J.C.: State maps for linear systems. SIAM J. Control Optim. **35**, 1053–1091 (1997)

32. Riquier, C.: Les Systèmes d'Équations aux Dérivées Partielles. Gauthiers-Villars, Paris, France (1910)
33. Rocha, P., Willems, J.C.: Markov properties for systems described by PDEs and first-order representations. Systems Control Lett. **55**, 538–542 (2006)
34. Stanley, R.P.: Combinatorics and Commutative Algebra, 2nd edn., Birkhäuser, Boston (1996)
35. Sturmfels, B.: Solving Systems of Polynomial Equations. American Mathematical Society, Providence, Rhode Island (2002)
36. Treves, F.: Basic Linear Partial Differential Equations. Academic Press, New York (1975)
37. Vardulakis, A.I.G.: Linear Multivariable Control. Wiley, Chichester, West Sussex, England (1991)
38. Vasconcelos, W.V.: Computational Methods in Commutative Algebra and Algebraic Geometry. Springer, Berlin Heidelberg New York (1998)
39. Willems, J.C.: Paradigms and puzzles in the theory of dynamical systems. IEEE Trans. Automat. Control **36**, 259–294 (1991)
40. Willems, J.C.: State and first order representations. In: Blondel, V. D. and Megretski, A. (eds.) Unsolved Problems in Mathematical Systems and Control Theory, pp. 54–57. Princeton University Press, New Jersey (2004)
41. Wolovich, W.A.: Linear Multivariable Systems. Springer, Berlin Heidelberg New York (1974)
42. Wood, J., Rocha, P., Rogers, E., Owens, D.H.: Structure indices for multidimensional systems. IMA J. Math. Control Inform. **17**, 227–256 (2000)
43. Zerz, E.: Topics in Multidimensional Linear Systems Theory. Lecture Notes in Control and Information Sciences 256, Springer, Berlin Heidelberg New York (2000)