Realization by Inspection

Joachim Rosenthal, J.M. Schumacher

Abstract— **We investigate which first-order representations can be obtained from high-order representations of linear systems 'by inspection', that is just by rearrangement of the data. Under quite weak conditions it is possible to obtain minimal realizations in the so-called pencil form; under stronger conditions one can obtain minimal realizations in standard state space form by inspection. The development is based on a reformulation of the realization problem as a problem of finding a complete set of basis vectors for the nullspace of a given constant matrix. Since no numerical computation is needed, the realization method is in particular suitable for situations in which some of the coefficients are symbolic rather than numerical.**

I. Introduction

As is well-known, the set of solutions of a higher-order linear differential equation in one variable

$$w^{(\ell)}(t) + p_{\ell-1}w^{(\ell-1)}(t) + \cdots + p_0 w(t) = 0 \tag{1}$$

may also be described in first-order form by

$$\dot{z}(t) = Fz(t), \quad w(t) = Hz(t)$$

where one can take for instance

$$F = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & 0 & 1 \\ -p_0 & -p_1 & \dots & \dots & -p_{\ell-1} \end{bmatrix}, \quad H = [1, 0, \dots, 0]. \tag{2}$$

The above equations give a 'realization' (in the behavioral sense, see [1]) of (1). There is a straightforward generalization of this for vector equations of the form $P(\frac{d}{dt})w(t) = 0$ when $P(s) \in \mathbb{R}^{p\times p}[s]$ is *monic*, i. e. $P(s) = \sum_{i=0}^{\ell} P_i s^i$ with $P_\ell = I$. In [2], [3] the term 'linearization' is used rather than 'realization'. The situation becomes more complicated if P_ℓ is singular or not even square. Indeed, assume that $P(s) = \sum_{i=0}^{\ell} P_i s^i$ is a $p \times (m+p)$ polynomial matrix. One readily verifies that the system $P(\frac{d}{dt})w = 0$ is represented by the first-order equations

$$G\dot{z}(t) = Fz(t), \quad w(t) = Hz(t) \tag{3}$$

if one chooses matrices

$$G = \begin{bmatrix} I_p & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & I_p & 0 \\ 0 & \dots & 0 & P_\ell \end{bmatrix}, \; F = \begin{bmatrix} 0 & \dots & \dots & \dots & -P_0 \\ I_p & 0 & & & -P_1 \\ 0 & I_p & \ddots & & \vdots \\ \vdots & & \ddots & 0 & \vdots \\ 0 & \dots & 0 & I_p & -P_{\ell-1} \end{bmatrix},$$

$$H = [0 \;\; | \;\; -I_{m+p}], \tag{4}$$

having size $p\ell \times (p\ell + m)$, $p\ell \times (p\ell + m)$ and $(m+p) \times (p\ell + m)$ respectively. However this may be rather crude since the obtained representation turns out to be minimal only if P_ℓ has full row rank (see Example 5.1 below). On the other hand, the realization (4) is easy to obtain since it only requires a reordering of the data and no numerical computation at all is involved; in other words, the realization is obtained from the data *by inspection*.

It is the purpose of the present paper to investigate more precisely which first-order representations can be obtained from a given polynomial representation by inspection, paying attention in particular to minimality properties. In general it is too much to ask that a standard state-space representation

$$\dot{x} = Ax + Bu, \quad y = Cx + Du, \qquad \begin{bmatrix} y \\ u \end{bmatrix} = w \tag{5}$$

can be obtained only by rearrangement of the data, but as we will demonstrate in this paper a representation in 'pencil' form (3), which is so-called completely observable (see Definition 2.4) can always be gotten by inspection. Pencil representations have recently been studied in [4], [5], [6], and we describe in Remark 3.6 below how standard state space representations can be obtained from them (in general at the cost of some numerical computation). Of course, realization theory has been studied extensively for several decades (see for instance [11]) and not surprisingly our algorithms show similarities to those that are already available in the literature. However, our purpose here is to determine to what extent realization algorithms survive when the constraint of no numerical computations is imposed.

The paper is built up as follows. In the next section we show that the realization problem can be reduced to a problem of finding a complete set of basis vectors for the nullset of a given constant matrix. Actually this reduction can be done in several ways, depending on the choice of what we call a 'polynomial basis matrix'.

In Section 3 we recall some characterizations of minimality properties. Minimality for realizations of the form (3) refers to minimality of the size of the matrices G and F among all representations of the same behavior.

In Section 4 we note that finding a basis for the nullset of a given matrix is under some conditions a problem that can be solved without calculations, and that we can in fact ensure that these conditions hold by making the use of the freedom we have in selecting a polynomial basis matrix. This leads immediately to a number of realization algorithms that are free of numerical computations.

In Section 5 we illustrate the realization algorithm presented in Section 4 by two examples. We conclude the paper with a table in Section 6 which summarizes the relations between the properties of high-order representations and of the corresponding first-order realizations that can be obtained with no computations, i. e. by inspection.

In connection with quantities that depend on a complex parameter s, we shall sometimes use the symbol $\equiv$ to denote equality for all $s \in \mathbb{C}$. A polynomial matrix $R(s)$ will be said to have *constant rank* if there exists an integer r such that $\operatorname{rank} R(s) \equiv r$.

II. Realization via a polynomial basis matrix

First let us briefly recall what is understood by realization in the behavioral sense; see for instance [1], [7], [8], [9] for a more extensive account. Given a polynomial matrix $P(s) \in \mathbb{R}^{p\times(m+p)}[s]$, the ($C^\infty$) behavior associated with $P(s)$ is defined by

$$\mathcal{B}(P) \; = \; \{w \in C^\infty(\mathbb{R}; \mathbb{R}^{m+p}) \mid P(\tfrac{d}{dt})w = 0\}. \tag{6}$$

This work was supported in part by NSF grant DMS-9400965.

J. Rosenthal is with the Department of Mathematics at the University of Notre Dame, Notre Dame, IN, USA 46556-5683.

J.M. Schumacher is with CWI, P.O. Box 94079, 1090 GB Amsterdam, and with the Department of Economics, Tilburg University, Tilburg, The Netherlands

Note that elementary row operations on $P(s)$ will not change the behavior. Such row operations correspond to premultiplication of $P(s)$ by a unimodular matrix $U(s)$. Moreover if both $P(s)$ and $\tilde{P}(s)$ are full row rank polynomial matrices then $\mathcal{B}(P) = \mathcal{B}(\tilde{P})$ if and only if there is a unimodular matrix $U(s)$ such that $\tilde{P}(s) = U(s)P(s)$ [10, Cor. 2.5].

Turning now to first-order representations, the behavior associated with a triple of matrices (F, G, H) (F and G in $\mathbb{R}^{n\times(n+m)}$, H in $\mathbb{R}^{(m+p)\times(n+m)}$) is given by

$$\mathcal{B}(F,G,H) = \{w \in C^\infty(\mathbb{R};\mathbb{R}^{m+p}) \mid \exists z \in C^\infty(\mathbb{R};\mathbb{R}^{n+m})\colon G\dot{z}=Fz,\ w=Hz\}.$$

The triple (F, G, H) is said to be a *realization* of the polynomial matrix $P(s)$ if $\mathcal{B}(F, G, H) = \mathcal{B}(P)$. Note that if (F, G, H) is a realization of $P(s)$ then so is $(SFT^{-1}, SGT^{-1}, HT^{-1})$, where S and T are nonsingular matrices. Triples that are related in this way will be said to be *isomorphic*.

The following basic lemma gives algebraic conditions for (F, G, H) to be a realization of $P(s)$. The lemma is a special case of [8, Lemma 4.1], although we do add a small extension. Since a large part of this paper is based on the lemma we outline the short proof.

Lemma 2.1: Let a polynomial matrix $P(s) \in \mathbb{R}^{p\times(m+p)}[s]$ and a triple of constant matrices (F, G, H) (F and G in $\mathbb{R}^{n\times(n+m)}$, H in $\mathbb{R}^{(m+p)\times(n+m)}$) be given. If there exists a polynomial matrix $X(s) \in \mathbb{R}^{p\times n}[s]$ such that $[X(s) \mid P(s)]$ has constant rank and the equality

$$\ker_{\mathbb{R}(s)}[X(s) \mid P(s)] \;=\; \operatorname{im}_{\mathbb{R}(s)}\begin{bmatrix} sG-F \\ H \end{bmatrix} \tag{7}$$

holds, then $\mathcal{B}(P) = \mathcal{B}(F, G, H)$, so (F, G, H) is a realization of $P(s)$.

Proof: There exists (see for instance [11, Thm. 6.3-2]) a unimodular matrix $U(s)$ such that

$$U(s)[X(s) \mid P(s)] \;=\; \begin{bmatrix} X_0(s) & P_0(s) \\ 0 & 0 \end{bmatrix}$$

where $[X_0(s) \mid P_0(s)]$ has full row rank as a rational matrix. By the assumption that $[X(s) \mid P(s)]$ has constant rank, we get that $[X_0(s) \mid P_0(s)]$ even has full row rank for all separate $s \in \mathbb{C}$. Moreover, it is obvious that $\mathcal{B}(P_0) = \mathcal{B}(P)$ and that $\ker_{\mathbb{R}(s)}[X_0(s) \mid P_0(s)] = \ker_{\mathbb{R}(s)}[X(s) \mid P(s)]$. So, replacing $P(s)$ by $P_0(s)$ and $X(s)$ by $X_0(s)$ if necessary, it is no restriction of generality to assume that $[X(s) \mid P(s)]$ has full row rank for all $s \in \mathbb{C}$. Then one can find (see for instance [11, Lemma 6.3-9]) polynomial matrices $U_1(s), U_2(s)$ such that

$$U(s) := \begin{bmatrix} U_1(s) & U_2(s) \\ X(s) & P(s) \end{bmatrix}$$

is a unimodular matrix. Let $T(s) := U_1(s)(sG - F) + U_2(s)H$. Because of (7) and the identity

$$\begin{bmatrix} U_1(s) & U_2(s) \\ X(s) & P(s) \end{bmatrix}\begin{bmatrix} sG-F \\ H \end{bmatrix} = \begin{bmatrix} T(s) \\ 0 \end{bmatrix}$$

it follows that the $(n + m) \times (n + m)$ polynomial matrix $T(s)$ is nonsingular. This implies (cf. [1, Prop. 3.3]) that the linear map

$$T : C^\infty(\mathbb{R};\mathbb{R}^{n+m}) \to C^\infty(\mathbb{R};\mathbb{R}^{n+m}), \quad z(t) \mapsto T(\tfrac{d}{dt})z(t)$$

is surjective. Note also that the differential equations

$$\begin{bmatrix} \frac{d}{dt}G-F \\ H \end{bmatrix} z(t) = \begin{bmatrix} 0 \\ w(t) \end{bmatrix}$$

and

$$\begin{bmatrix} T(\frac{d}{dt}) \\ 0 \end{bmatrix} z(t) = \begin{bmatrix} U_2(\frac{d}{dt}) \\ P(\frac{d}{dt}) \end{bmatrix} w(t)$$

describe the same smooth behavior. (Just transform the first equation by the unimodular matrix U.) By the surjectivity of $T(\frac{d}{dt})$, the latter equation describes exactly $\mathcal{B}(P)$. ∎

In the lemma the matrix $X(s)$ acts as a certification that the given triple (F, G, H) is indeed a realization of $P(s)$, but one may of course also reverse this: start with some chosen $X(s)$, then try to find a realization of $P(s)$ by looking for a triple (F, G, H) that satisfies (7). The question then is how to choose $X(s)$ so that this can indeed be done (easily) and that will be our main concern in this paper.

When looking for solutions of (7), one may restrict attention to triples (F, G, H) such that

$$\ker F \;\cap\; \ker G \;\cap\; \ker H \;=\; \{0\}. \tag{8}$$

Indeed, if (F, G, H) is a solution that doesn't satisfy (8), then there exists a nonsingular matrix T such that

$$\begin{bmatrix} F \\ G \\ H \end{bmatrix} T \;=\; \begin{bmatrix} F_1 & 0 \\ G_1 & 0 \\ H_1 & 0 \end{bmatrix}$$

and (F_1, G_1, H_1) satisfies both (8) and (7).

Definition 2.2: Let $P(s)$ and $X(s)$ be polynomial matrices such that $[X(s) \mid P(s)]$ has constant rank. A triple of constant matrices (F, G, H) is said to be a *realization of $P(s)$ associated to $X(s)$* if it satisfies both (7) and (8).

The following lemma shows that, for realizations associated to $X(s)$, the matrix $[sG^T - F^T \mid H^T]^T$ is guaranteed to have full column rank (even for all individual $s \in \mathbb{C}$ as well as at infinity) if $X(s)$ is chosen to have linearly independent columns.

Lemma 2.3: Let $P(s)$ and $X(s)$ be polynomial matrices, and suppose that the columns of $X(s)$ are linearly independent over $\mathbb{R}$ (i. e. if $X(s)z \equiv 0$ for some constant vector z, then $z = 0$). If (F, G, H) is a realization of $P(s)$ associated to $X(s)$, then the following holds true:

(i) $[\begin{smallmatrix} G \\ H \end{smallmatrix}]$ has full column rank

(ii) $[\begin{smallmatrix} sG-F \\ H \end{smallmatrix}]$ has full column rank for all $s \in \mathbb{C}$.

Proof: To prove part (i), suppose that $[\begin{smallmatrix} G \\ H \end{smallmatrix}]z = 0$ for some constant vector z. From the equation $X(s)(sG-F)+P(s)H \equiv 0$ it then follows that $X(s)Fz \equiv 0$. Because the columns of $X(s)$ are linearly independent over $\mathbb{C}$, this implies that $Fz = 0$. It now follows from (8) that $z = 0$. So we have proved that $[\begin{smallmatrix} G \\ H \end{smallmatrix}]$ has full column rank.

For part (ii), suppose that $[\begin{smallmatrix} \lambda G-F \\ H \end{smallmatrix}]z = 0$ for some $\lambda \in \mathbb{C}$ and some constant z. Since $sG - F \equiv (s - \lambda)G + (\lambda G - F)$, the equation $X(s)(sG - F) + P(s)H \equiv 0$ implies that $X(s)(s - \lambda)Gz \equiv 0$. From this it follows that $X(s)Gz \equiv 0$ and hence $Gz = 0$. But then, since $(\lambda G - F)z = 0$, we also have $Fz = 0$ and (8) implies that $z = 0$. It follows that $[\begin{smallmatrix} sG-F \\ H \end{smallmatrix}]$ has full column rank for all s. ∎

Following the terminology of [5], we shall define:

Definition 2.4: A triple (F, G, H) that satisfies conditions (i) and (ii) of the above lemma is called *completely observable*.

Condition (i) corresponds to 'observability at infinity' and condition (ii) characterizes the 'observability of the finite modes'. In connection with a particular interpretation of the dynamics associated to the triple (F, G, H), the term 'ex-in nulling' has also been used instead of 'completely observable' [12].

We now introduce a class of polynomial matrices from which we shall choose the matrix $X(s)$ on which our realization procedure is based.

Definition 2.5: Let $\nu = (\nu_1, \ldots, \nu_p)$ be a p-tuple of nonnegative integers. A polynomial matrix $X(s)$ is called a *polynomial basis matrix of type* ν or simply a *basis matrix* if every polynomial p-vector $\xi(s) \in \mathbb{R}^p[s]$ whose i-th component has degree at most $\nu_i - 1$ can uniquely be written as $\xi(s) = X(s)\alpha$ where α is a constant vector.

Remark 2.6: If $\nu_i = 0$ for some i then it is understood in the definition that the i-th component of $\xi(s)$ is 0. Note that one can identify the space of polynomials of degree at most $\nu_i - 1$ with the vector space $\mathbb{R}^{\nu_i}$. So a basis matrix of type $\nu = (\nu_1, \ldots, \nu_p)$ can be viewed as providing a basis for the vector space

$$\mathbb{R}^{\nu_1} \times \cdots \times \mathbb{R}^{\nu_p} \;\simeq\; \mathbb{R}^n$$

where $n = \sum_{i=1}^{p} \nu_i$. In particular it follows that a basis matrix must have size $p \times n$. It also follows that a basis matrix of a given type is determined uniquely up to right multiplication by a nonsingular constant matrix; more specifically, every basis matrix $X(s)$ can be written in the form $X(s) = X_\nu(s)S$ where S is a nonsingular constant matrix and $X_\nu(s)$ is the 'canonical' basis matrix of type $\nu = (\nu_1, \ldots, \nu_p)$ given by

$$X_\nu(s) = \begin{bmatrix} 1 & s & \cdots & s^{\nu_1-1} & 0 & \cdots & \cdots & & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & 0 & 1 & \cdots & s^{\nu_2-1} & 0 & \cdots & \cdots & 0 \\ \vdots & & & & & \ddots & \ddots & & & & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1 & \cdots & s^{\nu_p-1} \end{bmatrix}. \tag{9}$$

If some index ν_i is zero it is understood that the corresponding i-th row of $X_\nu(s)$ is zero.

We now arrive at the main result of this section. The realization method used in the proof will be the basis of the algorithms to be presented in Section 4.

Theorem 2.7: Let $P(s)$ be a $p \times (m+p)$ polynomial matrix whose i-th row degree is at most ν_i, and let $X(s)$ be a basis matrix of type $\nu = (\nu_1, \ldots, \nu_p)$. Under these conditions, the following holds.

(i) The matrix $[X(s) \mid P(s)]$ has constant rank.

(ii) There exist realizations of $P(s)$ associated to $X(s)$.

(iii) All realizations of $P(s)$ associated to $X(s)$ are completely observable.

(iv) If (F, G, H) and (F', G', H') are both realizations of $P(s)$ associated to $X(s)$, then there exists a nonsingular constant matrix T such that $F' = FT$, $G' = GT$, and $H' = HT$.

Proof: In order to prove the first part of the statement we will assume without loss of generality that $X(s)$ is the canonical basis matrix $X_\nu(s)$ and that the row degrees are ordered with $\nu_1 \geq \nu_2 \geq \cdots \geq \nu_j \geq 1$ and $\nu_i = 0$ for $i > j$. Under those assumptions we have

$$[X(s) \mid P(s)] = \begin{bmatrix} X_1(s) & P_1(s) \\ 0 & P_2 \end{bmatrix} \tag{10}$$

where $X_1(s)$ is the canonical basis matrix of type $(\nu_1, \ldots, \nu_j)$ and where by assumption P_2 is a constant matrix of size $(p-j) \times (p+m)$. Let the rank of P_2 be $p - j - r$. Note that the $j \times j$ submatrix of $X_1(s)$ consisting of the columns with indices $1, \nu_1 + 1, \nu_1 + \nu_2 + 1, \cdots, \nu_1 + \cdots + \nu_{j-1} + 1$ is in fact the identity matrix, so that $X_1(s)$ must have full row rank for all $s \in \mathbb{C}$. It follows that $[X(s) \mid P(s)]$ has constant rank $p - r$. This proves claim (i).

Since $p - r$ is of course also the rank of $[X(s) \mid P(s)]$ as a rational matrix and since the matrix $[X(s) \mid P(s)]$ has size $p \times (n + p + m)$ where $n = \sum_{i=1}^{p} \nu_i$, it follows that $\ker_{\mathbb{R}(s)}[X(s) \mid P(s)]$ has dimension $n+m+r$. In order to prove part (ii) identify the set of all polynomial vectors $\phi(s) \in \mathbb{R}^p[s]$ whose i-th component has degree at most ν_i with the vector space $\mathbb{R}^{n+p}$. Now consider the linear map

$$\begin{array}{rcl} \phi : \mathbb{R}^{2n+p+m} & \longrightarrow & \mathbb{R}^{n+p} \\ v & \longmapsto & [X(s) \mid sX(s) \mid P(s)]v. \end{array} \tag{11}$$

The dimension of the image of ϕ as a real vector space is given by the number of $\mathbb{R}$-linearly independent columns of the matrix

$$[X(s) \mid sX(s) \mid P(s)] = \begin{bmatrix} X_1(s) & sX_1(s) & P_1(s) \\ 0 & 0 & P_2 \end{bmatrix}.$$

Since all columns of $P_1(s)$ can be written as $\mathbb{R}$-linear combinations of the columns of $X_1(s)$ and $sX_1(s)$ (by the assumption that the row degrees of $P(s)$ are at most ν_i, and by the definition of a polynomial basis matrix), we get

$$\begin{array}{rcl} \dim \operatorname{im}_{\mathbb{R}} \phi & = & \operatorname{rank}_{\mathbb{R}} [X_1(s) \mid sX_1(s)] + \operatorname{rank}_{\mathbb{R}} P_2 \\ & = & (n + j) + (p - j - r) \;=\; n + p - r. \end{array}$$

From this we obtain $\dim \ker_{\mathbb{R}} \phi = n + m + r$. Choose constant matrices F, G, and H such that $[-F^T \mid G^T \mid H^T]^T$ is a basis matrix for $\ker_{\mathbb{R}} \phi$; of course these matrices must have $n+m+r$ columns. Then (8) certainly holds, and we have $X(s)(sG - F) + P(s)H = 0$ so that

$$\operatorname{im}_{\mathbb{R}(s)} \begin{bmatrix} sG - F \\ H \end{bmatrix} \subset \ker_{\mathbb{R}(s)}[X(s) \mid P(s)]. \tag{12}$$

The fact that actually equality holds in (12) follows from a dimension count: by Lemma 2.3, we have $\dim \operatorname{im}_{\mathbb{R}(s)}[sG^T - F^T \mid H^T] = n + m + r = \dim \ker_{\mathbb{R}(s)}[X(s) \mid P(s)]$.

Claim (iii) is immediate from Lemma 2.3. Finally, if a triple (F, G, H) satisfies (7) and (8), then the matrices F, G, and H must have $n + m + r$ columns, and $[-F^T \mid G^T \mid H^T]^T$ must be a basis matrix for $\ker_{\mathbb{R}} \phi$. All such matrices are related by nonsingular transformations as described in claim (iv). ∎

III. Minimality conditions

A pencil representation (F, G, H) with F and G in $\mathbb{R}^{n \times (n+m)}$ is said to be *minimal* if, whenever (F', G', H') with F' and G' in $\mathbb{R}^{n' \times (n'+m')}$ satisfies $\mathcal{B}(F', G', H') = \mathcal{B}(F, G, H)$, one has $n' \geq n$ and $n' + m' \geq n + m$. This means that both the number of auxiliary variables and the number of equations in those variables is minimal. For the relation between minimal pencil representations and standard input/state/output representations see Remark 3.6 below. The following algebraic conditions for minimality are well-known (see for instance [8, Prop. 1.1]).

Proposition 3.1: A pencil representation (F, G, H) is minimal (in the sense of smooth behaviors) if and only if it is completely observable and the matrix G has full row rank. Minimal realizations are unique up to isomorphism.

The full row rank condition on the matrix G corresponds to 'controllability at infinity'. Triples (F, G, H) can be used also for the representation of so-called impulsive-smooth behaviors [13], [12]. The definition of minimality is the same as above, with the smooth behaviors $\mathcal{B}(F, G, H)$ replaced by impulsive-smooth behaviors $\mathcal{B}_{\text{i-s}}(F, G, H)$. For this situation we have the following result [12, Thm. 4.1, 4.2].

Proposition 3.2: A pencil representation (F, G, H) is minimal in the sense of impulsive-smooth behaviors if and only if it is completely observable and $sG - F$ has full row rank as a rational matrix. Minimal realizations are unique up to isomorphism.

When we speak below about 'minimal' representations without further indication, we shall always mean minimality in the sense of smooth behaviors. The following lemma shows that minimality in the sense of impulsive-smooth behaviors is automatically obtained when $P(s)$ has full row rank.

Lemma 3.3: Let $P(s)$ be a $p \times (m+p)$ polynomial matrix whose i-th row degree is at most ν_i, and let $X(s)$ be a basis matrix of type $\nu = (\nu_1, \dots, \nu_p)$. Assume furthermore that $P(s)$ has full row rank as a rational matrix. If (F, G, H) is a realization associated to $X(s)$, then the matrix $sG - F$ has full row rank as a rational matrix.

Proof: We refer to the notation used in the proof of Thm. 2.7. Note that the full row rank assumption on $P(s)$ implies that $r = 0$, so that the matrix $sG - F$ has size $n \times (n+m)$. Now take any $\lambda \in \mathbb{C}$ such that $\operatorname{rank} P(\lambda) = p$. The equation $X(\lambda)(\lambda G - F) + P(\lambda)H = 0$ implies that H maps $\ker(\lambda G - F)$ into $\ker P(\lambda)$ and because of the observability of the triple (F, G, H) it does so in a one-to-one way. Therefore, we have

$$\dim \ker(\lambda G - F) \;\leq\; \dim \ker P(\lambda) \;=\; m. \qquad (13)$$

On the other hand, we also have $\dim \ker(\lambda G - F) \geq m$ since $\lambda G - F$ has size $n \times (n+m)$. It follows that $\dim \ker(\lambda G - F) = m$ and so $\operatorname{rank}(\lambda G - F) = n$. This implies that $sG - F$ has full row rank n as a rational matrix. ■

Remark 3.4: The proof actually shows that, for any $\lambda \in \mathbb{C}$, the matrix $\lambda G - F$ will have full row rank if $P(\lambda)$ has full row rank. In particular it follows that if the conditions of the lemma hold and $P(s)$ has constant full row rank p, then $sG - F$ has constant full row rank n. Recall that the first condition is the algebraic characterization of controllability of the behavior $\mathcal{B}(P)$ in the sense of Willems [9, Thm. V.2], whereas the second characterizes controllability of the system $G\dot{z} = Fz$.

We now consider the more specialized situation in which $P(s)$ is row proper and the type of the polynomial basis matrix $X(s)$ is matched to the row degrees of $P(s)$.

Lemma 3.5: Let $P(s)$ be a row proper polynomial matrix of size $p \times (m+p)$, with row degrees $\nu = (\nu_1, \dots, \nu_p)$. Let $X(s)$ be a basis matrix of type ν and let (F, G, H) be a realization associated with this basis matrix. Then the matrix G must have full row rank.

Proof: The statement follows from the previous lemma and [12, Lemma 3.3]. ■

Remark 3.6: From a minimal pencil representation, a standard state space representation can be obtained as follows. Since G has full row rank and $\left[\begin{smallmatrix} G \\ H \end{smallmatrix}\right]$ has full column rank, we can select a submatrix H' from H such that $\left[\begin{smallmatrix} G \\ H' \end{smallmatrix}\right]$ is an invertible matrix. After a permutation of the external variables and a transformation $T \in Gl_{m+n}$ of the internal variables the triple (F, G, H) appears in the following form:

$$F = [A \mid -B], \quad G = [I \mid 0], \quad H = \begin{bmatrix} C & D \\ 0 & I \end{bmatrix}. \qquad (14)$$

Denoting the two components of w by y and u respectively, we arrive at the familiar form $\dot{x} = Ax + Bu$, $y = Cx + Du$. For the particular pencil

$$\begin{bmatrix} sG - F \\ H \end{bmatrix} = \begin{bmatrix} sI - A & B \\ C & D \\ 0 & I \end{bmatrix}$$

the algebraic conditions for observability and controllability then reduce to the standard conditions. An algorithm to obtain a minimal pencil representation from an arbitrary one is given in [10]. For cases in which an input/output structure is given *a priori* and in such a way that the corresponding submatrix of $[G^T \mid H^T]^T$ is not invertible, see [4].

IV. Realization algorithms

In Section 2 we have seen that the problem of finding a realization can be reduced to the problem of finding a complete set of basis vectors for the nullset of a given matrix. Note now that in some cases this problem is rather easy, namely when the given matrix is of the form $[I \mid M]$. Obviously we can immediately write

$$\ker [I \mid M] \;=\; \operatorname{im} \begin{bmatrix} -M \\ I \end{bmatrix}$$

and no calculation is necessary. If the given matrix is a column permuted form of $[I \mid M]$, then some rearrangement will be needed but still no numerical calculations will be involved. By judicious choice of the polynomial basis matrix $X(s)$ (for instance the canonical basis matrix is suitable) we can in fact create such a situation. The following two theorems are based on this observation. The proofs are in both cases straightforward applications of Lemma 2.1, applied with the canonical basis matrix.

First we introduce some notation. For a given polynomial matrix $P(s)$ of size $p \times (m+p)$, let $f_i(s) \in \mathbb{R}^{m+p}[s]$ denote the i-th row of $P(s)$, and let $\tilde{\nu}_i$ be its degree. For $0 \leq k \leq \tilde{\nu}_i$ define vectors f_i^k through the expansion

$$f_i(s) = \sum_{k=0}^{\tilde{\nu}_i} f_i^k s^k, \quad f_i^k \in \mathbb{R}^{m+p}$$

and define $f_i^k = 0$ for $k > \tilde{\nu}_i$. Let $\nu = (\nu_1, \dots, \nu_p)$ be positive integers satisfying $\nu_i \geq \tilde{\nu}_i$. For $i = 1, \dots, p$ define matrices of sizes $\nu_i \times (\nu_i - 1)$ and $\nu_i \times (m+p)$ respectively:

$$\Phi_i(s) := \begin{bmatrix} s & 0 & \cdots & 0 \\ -1 & \ddots & & \vdots \\ 0 & \ddots & \ddots & 0 \\ \vdots & & \ddots & s \\ 0 & \cdots & 0 & -1 \end{bmatrix}, \; \Psi_i(s) := \begin{bmatrix} f_i^0 \\ f_i^1 \\ \vdots \\ f_i^{\nu_i - 2} \\ s f_i^{\nu_i} + f_i^{\nu_i - 1} \end{bmatrix}.$$

Theorem 4.1: Let $P(s)$ be given and let $\Phi_i(s)$, $\Psi_i(s)$ be defined as above. Then

$$\begin{aligned} sG - F \;&:=\; \begin{bmatrix} \Phi_1(s) & 0 & \cdots & 0 & \Psi_1(s) \\ 0 & \Phi_2(s) & & \vdots & \Psi_2(s) \\ \vdots & & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & \Phi_p(s) & \Psi_p(s) \end{bmatrix}, \\ H \;&:=\; [0 \mid -I_{m+p}] \end{aligned}$$

is a completely observable realization of $P(s)$.

Proof: Let $X_\nu(s)$ be the standard basis matrix as introduced in (9). A direct computation verifies that

$$X_\nu(s)\,[sG - F] = \left[0_{p \times (n-p)} \mid P(s)\right] = -P(s)H.$$

By a dimension count we find that (7) holds. Since (F, G, H) also satisfies (8) it follows from Thm. 2.7 that (F, G, H) is a completely observable realization of $P(s)$. ■

Remark 4.2: It follows from the lemmas 3.3 and 3.5 that the realization obtained above will be minimal in the sense of impulsive-smooth behaviors if $P(s)$ has full row rank as a rational matrix, and that it will be minimal if $P(s)$ is row proper and $\tilde{\nu}_i = \nu_i$ for all i. Note that the latter requirement implies that $P(s)$ can have no constant rows. So the following obstructions can exist to obtaining a minimal representation by inspection:

(i) $P(s)$ doesn't have full row rank, (ii) $P(s)$ is not row proper, (iii) $P(s)$ has some constant rows. All of these obstructions may be overcome at the cost of some computation, which one may choose to carry out on the polynomial level (before realization) or on the first-order level (after realization).

We now present a theorem that produces a standard state space representation by inspection for strictly proper systems. Naturally, this is only possible when $P(s)$ satisfies a rather special condition. Again we first introduce some notation. Assume that $P(s)$ is partitioned into $P(s) = [D(s) \mid N(s)]$ where $D(s)$ is a $p \times p$ polynomial matrix. We will assume that $P(s)$ is row proper with row degrees $\nu_1 \geq \cdots \geq \nu_p \geq 1$. For $i, j = 1, \ldots, p$ let

$$d_{i,j}(s) = \sum_{k=0}^{\nu_i} d_{i,j}^k \, s^k$$

denote the polynomial entries of $D(s)$. Similarly let

$$n_i(s) = \sum_{k=0}^{\nu_i} n_i^k \, s^k$$

denote the i-th row of $N(s)$. Define for $i = 1, \ldots, p$ matrices of sizes $\nu_i \times \nu_i$, $\nu_i \times m$ and $1 \times \nu_i$ respectively:

$$A_{i,i} := \begin{bmatrix} 0 & \ldots & \ldots & \ldots & -d_{i,i}^0 \\ 1 & 0 & & & -d_{i,i}^1 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & & \ddots & 0 & \vdots \\ 0 & \ldots & 0 & 1 & -d_{i,i}^{\nu_i - 1} \end{bmatrix}, \quad B_i := \begin{bmatrix} n_i^0 \\ n_i^1 \\ \vdots \\ n_i^{\nu_i - 1} \end{bmatrix},$$

$$C_i := [0, \ldots, -1].$$

Finally for $i, j = 1, \ldots, p$, $i \neq j$ define matrices of size $\nu_i \times \nu_j$:

$$A_{i,j} := \begin{bmatrix} 0 & \ldots & 0 & -d_{i,j}^0 \\ \vdots & & \vdots & -d_{i,j}^1 \\ \vdots & & \vdots & \vdots \\ 0 & \ldots & 0 & -d_{i,j}^{\nu_i - 1} \end{bmatrix}.$$

With these definitions we can state the following.

Theorem 4.3: If, in the situation discussed above, the high order row coefficient matrix P_∞ is of the form $P_\infty = [I_p \mid 0]$ then

$$\begin{aligned} \dot{x}(t) &= \begin{bmatrix} A_{1,1} & \cdots & A_{1,p} \\ \vdots & \ddots & \vdots \\ A_{p,1} & \cdots & A_{p,p} \end{bmatrix} x(t) + \begin{bmatrix} B_1 \\ \vdots \\ B_p \end{bmatrix} u(t), \\ y(t) &= \begin{bmatrix} C_1 & & 0 \\ & \ddots & \\ 0 & & C_p \end{bmatrix} x(t) \end{aligned} \tag{15}$$

represents a minimal state space realization of the system

$$D(\tfrac{d}{dt})y(t) + N(\tfrac{d}{dt})u(t) = 0. \tag{16}$$

Proof: As in the proof of Theorem 4.1 one readily verifies that

$$[X_\nu(s) \mid P(s)] \begin{bmatrix} sI - A & B \\ C & 0 \\ 0 & I \end{bmatrix} = 0.$$

Again a dimension count confirms that we do have a realization. Minimality (in the behavioral sense) is guaranteed by Thm. 2.7. ■

Remark 4.4: Because behavioral equivalence is an extension of transfer equivalence, we have in particular that

$$-D^{-1}(s)N(s) = C(sI - A)^{-1}B.$$

It follows from Remark 3.4 (see also Remark 3.6) that the obtained realization will be controllable if the matrix $P(s)$ has full row rank for all s, or in other words, if the pair $(D(s), N(s))$ is left coprime. So in this case we even have minimality in the transfer sense. See [14] for a review of the various notions of minimality.

Remark 4.5: The choice of the canonical basis matrix $X_\nu(s)$ introduced in (9) has produced a matrix A in a well known companion form as it can be found e. g. in [15, p. 82]. Of course other choices of basis matrices are possible and lead to various results, see for instance Example 5.1 below. There is clearly a connection here to canonical forms, and this is discussed in more detail in [14].

Remark 4.6: If the high-order coefficient matrix is of the form $[P_1 \mid P_2]$ with P_1 invertible then the situation of the theorem can be achieved (at the cost of some computation) by a linear transformation in the space of external variables. Reversion of this transformation after realization will lead to a realization in (A, B, C, D) form.

V. Examples

Example 5.1: Consider a $p \times (m+p)$ polynomial matrix of the form $P(s) := \sum_{i=0}^{\ell} P_i s^i \in \mathbb{R}[s]^{p \times (m+p)}$. Although we have worked with the canonical basis matrix $X_\nu(s)$ (as introduced in Section 2) throughout the main part of the paper other choices are quite possible. Consider for instance the basis matrix

$$X(s) := \left[I_p \mid sI_p \mid \cdots \mid s^{\ell-1} I_p \right].$$

Let (F, G, H) be the triple of matrices introduced in (4). One readily verifies that

$$X(s)\,[sG - F] = \left[0_{p \times ((\ell-1)p)} \mid P(s) \right] = -P(s)H.$$

By Theorem 2.7 (F, G, H) is a completely observable realization and by Proposition 3.1 this realization is minimal if and only if P_ℓ and therefore G has full row rank. Actually it is not difficult to derive these facts from first principles; the example shows however that also in the present approach the particular realization 4 appears as the result of making some simple choices. To compare this with Thm. 4.1, note that $P(s)$ is row proper whenever P_ℓ has full row rank, but not conversely.

Example 5.2: This example illustrates Thm. 4.1. We consider the situation of a 2×4 polynomial system $P(s)$ having row degrees $\nu_1 = 3$ and $\nu_2 = 2$. Using earlier notation $P(s)$ is of the form

$$P(s) = \begin{bmatrix} f_1(s) \\ f_2(s) \end{bmatrix} = \begin{bmatrix} f_{1,1}(s), \ldots, f_{1,4}(s) \\ f_{2,1}(s), \ldots, f_{2,4}(s) \end{bmatrix},$$

where

$$f_{1,j}(s) = \sum_{k=0}^{3} f_{1,j}^k s^k, \quad f_{2,j}(s) = \sum_{k=0}^{2} f_{2,j}^k s^k, \quad j = 1, \ldots, 4.$$

The canonical basis matrix of size $\nu = (3, 2)$ has the form

$$X_{(3,2)}(s) = \begin{bmatrix} 1 & s & s^2 & 0 & 0 \\ 0 & 0 & 0 & 1 & s \end{bmatrix}.$$

The computation of the kernel of

$$[X_{(3,2)}(s) \mid sX_{(3,2)}(s) \mid P(s)]$$

is equivalent to finding a complete set of basis vectors for the space determined by the equation

$$\left[\begin{array}{ccccc|ccccc|c} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & f_1^0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & f_1^1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & f_1^2 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & f_2^0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & f_2^1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & f_1^3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & f_2^2 \end{array}\right] x \;=\; 0.$$

Since the minor consisting of columns $1, 2, 3, 4, 5, 8, 10$ is just an identity matrix the kernel is found 'by inspection' and is given by (see Thm. 4.1)

$$\begin{bmatrix} -F \\ G \\ H \end{bmatrix} = \left[\begin{array}{cccc} 0 & 0 & 0 & -f_1^0 \\ -1 & 0 & 0 & -f_1^1 \\ 0 & -1 & 0 & -f_1^2 \\ 0 & 0 & 0 & -f_2^0 \\ 0 & 0 & -1 & -f_2^1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -f_1^3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -f_2^2 \\ \hline \vec{0} & \vec{0} & \vec{0} & I_4 \end{array}\right].$$

The realization is minimal if and only if the row vectors f_1^3 and f_2^2 are linearly independent.

VI. Conclusions

In this paper we showed that a linear system represented by a system of higher order differential equations of the form $P(\frac{d}{dt})w(t) = 0$ can always be realized in a generalized first-order pencil form by a simple rearrangement of the coefficients. Since no numerical computation is involved, the approach is suitable in particular in situations where some of the coefficients are symbolic parameters rather than actual numbers. The first-order realizations that are obtained by the methods of this paper will contain the same parameters, together with zeros and fixed constants. Genericity issues for such systems have been studied by Murota [16]. Another possibility that suggests itself is to allow for coefficients that come from a ring rather than from a field, but we shall not go into that here.

Whether the first-order form that is obtained by inspection can be made to have certain desirable properties depends on the data that one starts from. This is detailed in the following table.

Realization by inspection		
High-order form	First-order form	Reference
No special properties	Completely observable pencil form	Thm. 4.1
$P(s)$ of full generic row rank	Completely observable pencil form, minimal in the sense of impulsive-smooth behaviors	Thm. 4.1, Lemma 3.3
$P(s)$ row proper, no constant rows	Minimal pencil representation	Thm. 4.1, Lemma 3.5
$P(s) = [D(s) \mid N(s)]$, high-order coefficient matrix is $[I \mid 0]$, no constant rows	Observable standard state space representation	Thm. 4.3
The above plus coprimeness of $D(s)$ and $N(s)$	Observable and controllable standard state space representation	Thm. 4.3, Remark 4.4

References

[1] J. C. Willems, "Input–output and state–space representations of finite–dimensional linear time–invariant systems", *Linear Algebra Appl.*, vol. 50, pp. 581–608, 1983.

[2] I. Gohberg, P. Lancaster, and L. Rodman, *Matrix Polynomials*, Academic Press, New York, 1982.

[3] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, Academic Press, 1985.

[4] M. Kuijper, *First-Order Representations of Linear Systems*, Birkhäuser, Boston, 1994.

[5] V. G. Lomadze, "Finite-dimensional time-invariant linear dynamical systems: Algebraic theory", *Acta Appl. Math*, vol. 19, pp. 149–201, 1990.

[6] M. S. Ravi and J. Rosenthal, "A general realization theory for higher order linear differential equations", *Systems & Control Letters*, vol. 25, no. 5, pp. 351–360, 1995.

[7] J. C. Willems, "From time series to linear system. Part I: Finite dimensional linear time invariant systems", *Automatica*, vol. 22, pp. 561–580, 1986.

[8] M. Kuijper and J. M. Schumacher, "Realization of autoregressive equations in pencil and descriptor form", *SIAM J. Control Optim.*, vol. 28, no. 5, pp. 1162–1189, 1990.

[9] J. C. Willems, "Paradigms and puzzles in the theory of dynamical systems", *IEEE Trans. Automat. Control*, vol. AC-36, no. 3, pp. 259–294, 1991.

[10] J. M. Schumacher, "Transformations of linear systems under external equivalence", *Linear Algebra Appl.*, vol. 102, pp. 1–33, 1988.

[11] T. Kailath, *Linear Systems*, Prentice-Hall, Englewood Cliffs, N.J., 1980.

[12] A. H. W. Geerts and J. M. Schumacher, "Impulsive-smooth behavior in multimode systems. Part II: Minimality and equivalence", *Automatica*, vol. 32, no. 6, pp. 819–832, 1996.

[13] A. H. W. Geerts and J. M. Schumacher, "Impulsive-smooth behavior in multimode systems. Part I: State-space and polynomial representations", *Automatica*, vol. 32, no. 5, pp. 747–758, 1996.

[14] M. S. Ravi, J. Rosenthal, and J. M. Schumacher, "System equivalences and canonical forms from a behavioral point of view", in *Proc. of the 34th IEEE Conference on Decision and Control*, New Orleans, Louisiana, 1995, pp. 484–489.

[15] W. A. Wolovich, *Linear Multivariable Systems*, vol. 11 of *Appl. Math. Sc.*, Springer Verlag, New York, 1974.

[16] K. Murota, *Systems Analysis by Graphs and Matroids*, Springer, New York, 1987.