Pergamon

PII: S0005-1098(96)00007-6

Automatica, Vol. 32, No. 6, pp. 819–832, 1996

0005-1098/96 $15.00 + 0.00

Impulsive–Smooth Behavior in Multimode Systems Part II: Minimality and Equivalence*

A. H. W. (TON) GEERTS† and J. M. SCHUMACHER‡

The conditions are obtained for minimality of first-order representations of impulsive–smooth behaviors.

Key Words—System models; singular systems; switching.

Abstract—This is the second part of a two-part study of linear multimode systems. In the first part, it was argued that the behavior of such a system on an interval between switches should be described in a framework that allows for impulses at the switching instant, and both first-order and polynomial representations were introduced that satisfy this requirement. Here we determine the conditions under which first-order representations are minimal. We also show how two minimal representations of the same behavior are related; this leads in particular to an appropriate state-space isomorphism theorem. The minimality conditions are given a dynamic interpretation.

1. INTRODUCTION

In Part I of this study (Geerts and Schumacher, 1996), we motivated the study of a particular class of piecewise-linear systems, suitable for the description of multimode systems in particular when the operating modes correspond to systems that are not all of the same McMillan degree. It has been argued that the system description should allow for impulses at switching instants. We have therefore looked at behaviors (Willems, 1991) consisting of impulsive–smooth distributions (Hautus, 1976; Hautus and Silverman, 1983). We have shown that both state-space and polynomial representations can be given for such behaviors, and we have presented algorithms that transform a state-space to a polynomial description, and vice versa. In this part, we concentrate on minimality issues. Our aim is to answer the following two questions: under what conditions is a representation minimal, and how are minimal representations of the same behavior related?

We proceed as follows. In Section 2, we derive necessary conditions for minimality of state-space representations by showing that systems not satisfying these conditions can be reduced. Before showing that these conditions are also sufficient, it is mathematically convenient to turn first to polynomial representations, and we do this in Section 3. We determine to what extent polynomial representations are unique. Then we get back to first-order systems in Section 4, prove that the necessary conditions of Section 2 are actually also sufficient for minimality, and obtain a state-space isomorphism theorem. The reader may wonder about the dynamic interpretation of the minimality conditions that we find; an answer is supplied in Section 5. Finally, conclusions are given in Section 6.

The notational conventions of Part I remain in force here. For the convenience of the reader, let us briefly repeat the main points. We denote by $\mathscr{C}(t_{\mathrm{in}}, t_{\mathrm{out}})$ the set of restrictions of $C^\infty(\mathbb{R})$ functions to $(t_{\mathrm{in}}, t_{\mathrm{out}})$, with $-\infty < t_{\mathrm{in}} < t_{\mathrm{out}} \le +\infty$. The product space $\{\mathbb{R}[p] \times \mathscr{C}(t_{\mathrm{in}}, t_{\mathrm{out}})\}^k$ is denoted by $\mathscr{C}^k_{\mathrm{imp}}(t_{\mathrm{in}}, t_{\mathrm{out}})$; so elements u of this space consist of a polynomial part, which we shall refer to as the 'purely impulsive part' $u_{\text{p-imp}}$ (representing a pulse at time t_{in}), and a function part, which is called the 'smooth part' u_{sm}. We write $u = u_{\text{p-imp}} + u_{\mathrm{sm}}$; the summation is motivated by the fact that the elements of $\mathscr{C}^k_{\mathrm{imp}}(t_{\mathrm{in}}, t_{\mathrm{out}})$ may be identified with certain distributions (see Part I; Hautus, 1976). To describe dynamics, we use an operator p defined by

$$pu = pu_{\text{p-imp}} + u_{\mathrm{sm}}(t_{\mathrm{in}}^+) + \dot{u}_{\mathrm{sm}}. \tag{1}$$

* Received 15 October 1994; received in final form 27 October 1995. A preliminary version of this paper was presented at the 2nd IFAC Workshop on System Structure and Control, which was held in Prague, Czech Republic, during 3–5 September 1992. The Published Proceedings of this IFAC meeting may be ordered from Elsevier Science Limited, The Boulevard, Langford Lane, Kiklington, Oxford OX5 1GB, U.K. This paper was recommended for publication in revised form by Associate Editor Hassan Khalil under the direction of Editor Tamer Başar.
Corresponding author Professor J. M. Schumacher, Tel. +31 20 5924090; Fax +31 20 5924199; E-mail hans.schumacher@cwi.nl

† Marensedijk 18, 5398 KM Maren-Kessel, The Netherlands.

‡ CWI, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands, and Tilburg University, Center and Department of Economics, P.O. Box 90153, 5000 LE Tilburg, The Netherlands.

The interval $(t_{\text{in}}, t_{\text{out}})$ will be fixed in the discussion below, and so, in order to ease notation, we shall suppress reference to the specific instants t_{in} and t_{out} as much as possible. Impulsive–smooth behaviors (i.e. linear subspaces of $\mathscr{C}^q_{\text{imp}}$) can now be described in various ways. The *conventional representation* is the following.

Definition 1.1. For a matrix triple (F, G, H) $(F, G \in \mathbb{R}^{n\times(n+m)}, H \in \mathbb{R}^{q\times(n+m)})$, we define

$$\mathscr{B}_c(F, G, H) = \{w \in \mathscr{C}^q_{\text{imp}} \mid \exists z \in \mathscr{C}^{n+m}_{\text{imp}}, \quad z_0 \in \mathbb{R}^{n+m} \text{ s.t. } pGz = Fz + Gz_0,\ w = Hz\}. \quad (2)$$

The *pencil representation* given below treats initial data slightly differently.

Definition 1.2. For a matrix triple (F, G, H) $(F, G \in \mathbb{R}^{n\times(n+m)}, H \in \mathbb{R}^{q\times(n+m)})$, we define

$$\mathscr{B}(F, G, H) = \{w \in \mathscr{C}^q_{\text{imp}} \mid \exists z \in \mathscr{C}^{n+m}_{\text{imp}}, \quad x_0 \in \mathbb{R}^{n} \text{ s.t. } pGz = Fz + x_0,\ w = Hz\}. \quad (3)$$

Polynomial representations can be introduced as follows.

Definition 1.3. Let $R(s) \in \mathbb{R}^{p\times q}[s]$ and $V(s) \in \mathbb{R}^{p\times n}[s]$. We define

$$\mathscr{B}(R, V) = \{w \mid R(p)w \in \operatorname{span}_{\mathbb{R}} V(p)\}. \quad (4)$$

The relation between polynomial and first-order representations is as follows (Lemma 5.7 of Part I).

Lemma 1.4. If one has

$$\operatorname{im}\begin{bmatrix} sG - F \\ H \end{bmatrix} = \ker\,[-V(s) \quad R(s)] \quad (5)$$

as an equality between rational vector spaces then $\mathscr{B}(F, G, H) = \mathscr{B}(R, V)$.

To make sure that polynomial representations do have corresponding first-order representations, we consider only 'eligible pairs', which are defined as follows (Definition 5.8 and 5.11 in Part I).

Definition 1.5. For a polynomial matrix $R(s) \in \mathbb{R}^{p\times q}[s]$, define

$$X_R = \{f(s) \in \mathbb{R}^p[s] \mid f(s) = R(s)g(s) \text{ for some strictly proper } g(s)\}. \quad (6)$$

Definition 1.6. A pair of polynomial matrices $(R(s), V(s)) \in \mathbb{R}^{p\times q}[s] \times \mathbb{R}^{p\times n}[s]$ is called *eligible* if the following conditions hold:

(i) $[-V(s)\ \ R(s)]$ has full row rank for all $s \in \mathbb{C}$;

(ii) $R(s)$ has full row rank as a rational matrix;

(iii) the columns of $V(s)$ are linearly independent over $\mathbb{R}$;

(iv) $\operatorname{span}_{\mathbb{R}} V(s) = X_R + R(s)L(s)$, where $L(s)$ is an $\mathbb{R}$-linear subspace of $\mathbb{R}^k[s]$ that is *shift-invariant*, i.e. it is mapped into itself by the operation $\sigma: f(s) \mapsto [f(s) - f(0)]/s$.

The following standard lemma will be used on several occasions.

Lemma 1.7. Let A, B, X and Y be matrices such that $AX + BY = 0$. If $[A\ \ B]$ and X have full row rank then also B has full row rank.

Proof. Let η be a row vector such that $\eta B = 0$. It follows from $\eta(AX + BY) = 0$ that $\eta AX = 0$, and consequently $\eta A = 0$ because X has full row rank. But then we have $\eta[A\ \ B] = 0$, which implies that $\eta = 0$. □

2. REDUNDANCY IN FIRST-ORDER REPRESENTATIONS

A first-order representation of a behavior is called *minimal* if both the number of equations and the number of variables are minimal among all equivalent representations of the same form. Because there are two indices to be minimized, even the existence of minimal representations is not trivial. The following lemmas will be useful in proving necessary conditions for minimality. Recall that two matrix pencils $sG - F$ and $s\tilde{G} - \tilde{F}$ are called *strictly equivalent* if there exist nonsingular matrices S and T such that $\tilde{G} = SGT$ and $\tilde{F} = SFT$.

Lemma 2.1. For any matrix pencil $sG - F$, there exists a strictly equivalent pencil $s\tilde{G} - \tilde{F}$ of the form

$$s\tilde{G} - \tilde{F} = \begin{bmatrix} sG_1 - F_1 & 0 \\ 0 & sG_2 - F_2 \\ 0 & 0 \end{bmatrix}, \quad (7)$$

where $sG_1 - F_1$ has full row rank (as a rational matrix), $sG_2 - F_2$ has full column rank for all $s \in \mathbb{C}$, and G_2 has full column rank.

Proof. The decomposition (7) can be obtained as an 'aggregate' of the Kronecker canonical form for matrix pencils (Gantmacher, 1959, Chap. XII) by letting the zero block and the block $sG_2 - F_2$ correspond to the zero row minimal indices and the nonzero row minimal indices respectively, and the block $sG_1 - F_1$ to the remaining invariants. The claimed properties can then be read off from the canonical form. □

Lemma 2.2. Let F and G be matrices such that $sG - F$ has full column rank for all s, and moreover G has full column rank. Under these conditions, the equation $pGz = Fz + x_0$, where x_0 is a constant vector, can only be satisfied if $x_0 = 0$ and $z = 0$.

Proof. Let z and x_0 be such that $(pG - F)z = x_0$, and suppose that z is nonzero. Because $sG - F$ has a polynomial left inverse, it follows that z must be purely impulsive, say

$$z = z_k p^k + \ldots + z_1 p + z_0, \tag{8}$$

with $z_k \neq 0$. But the equation $pGz = Fz + x_0$ implies in particular that $Gz_k = 0$, contradicting the assumption that G has full column rank. Therefore z must be equal to zero, and this implies that x_0 is also zero. □

For pencil representations, we have the following necessary conditions for minimality.

Theorem 2.3. The following conditions are necessary for a triple (F, G, H) to be a minimal representation of its associated behavior $\mathscr{B}(F, G, H)$:

(i) $sG - F$ has full row rank as a rational matrix;

(ii) $\begin{bmatrix} G \\ H \end{bmatrix}$ has full column rank;

(iii) $\begin{bmatrix} sG - F \\ H \end{bmatrix}$ has full column rank for all $s \in \mathbb{C}$.

Proof. We have to show that a representation can be reduced if any of conditions (i)–(iii) is not satisfied. For condition (i), this follows from the two lemmas above (equations corresponding to the zero block and the block $sG_2 - F_2$ in (7) can be eliminated without changing the set of solutions). The other two conditions are shown to be necessary exactly as in the case of smooth behaviors (cf. e.g. Schumacher, 1989). □

Remark 2.4. In a more 'homogeneous' style, conditions (i)–(iii) may equivalently be formulated as follows:

(i) $sG - tF$ has full row rank for some $(s, t) \in \mathbb{C}^2 \backslash \{(0, 0)\}$;

(ii) $\begin{bmatrix} sG - tF \\ H \end{bmatrix}$ has full column rank for all $(s, t) \in \mathbb{C}^2 \backslash \{(0, 0)\}$.

This formulation shows clearly that the conditions are symmetric with respect to F and G.

Remark 2.5. Note in particular that 'controllability at infinity' (G full row rank, in the language of pencil representations) is *not* required in Theorem 2.3. This is due to the fact that we allow 'autonomous' impulsive behavior, that is, impulsive behavior that is due to initial conditions rather than to inputs. For the simplest example of this, consider the system defined by $F = 1$, $G = 0$, $H = 1$. The corresponding behavior as described by (3) is

$$\mathscr{B} = \{w \in \mathscr{C}_{\mathrm{imp}} \mid \exists z \in \mathscr{C}_{\mathrm{imp}}, \; x_0 \in \mathbb{R} \text{ s.t. } 0 = z + x_0, \; w = z\}, \tag{9}$$

which is the space of constant multiples of the delta distribution.

For conventional representations, the conditions of Theorem 2.3 are supplemented as follows.

Theorem 2.6. A conventional representation $\mathscr{B} = \mathscr{B}_c(F, G, H)$ is minimal only if the matrices F, G and H satisfy conditions (i)–(iii) of Theorem 2.3 and the additional condition

(iv) $F[\ker G] \subset \operatorname{im} G$.

Proof. The necessity of conditions (i)–(iii) is shown in the same way as for pencil representations. Suppose that (iv) does not hold. By change of coordinates, we may assume that F and G are in the form

$$F = \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & 0 & 0 \\ F_{31} & 0 & I \end{bmatrix}, \quad G = \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{10}$$

The equations $pGz = Fz + Gz_0$ then take the form

$$\begin{aligned} pz_1 &= F_{11}z_1 + F_{12}z_2 + F_{13}z_3 + z_0, \\ 0 &= F_{21}z_1, \\ 0 &= F_{31}z_1 + z_3. \end{aligned} \tag{11}$$

Obviously it is possible to reduce these equations by eliminating z_3. This means that the triple (F, G, H) can only be minimal if the identity block does not appear in the decomposition of F as in (10), which is the same as requiring that $F[\ker G] \subset \operatorname{Im} G$. □

3. EQUIVALENCE OF POLYNOMIAL REPRESENTATIONS

In this section, we show that the class of eligible polynomial pairs is large enough to represent all impulsive–smooth behaviors that can be obtained as in (3) (or (2)). In addition, we determine exactly to what extent representations by eligible pairs are unique. The following lemmas will be instrumental.

Lemma 3.1. Suppose that (5) holds for a triple of constant matrices (F, G, H) and a pair of polynomial matrices $(R(s), V(s))$. If $\begin{bmatrix} G \\ H \end{bmatrix}$ has full column rank and $\begin{bmatrix} sG - F \\ H \end{bmatrix}$ has full column rank for all s, then the columns of $V(s)$ are linearly independent over $\mathbb{R}$.

Proof. Let x be a constant vector such that $V(s)x = 0$. We have

$$\begin{bmatrix} x \\ 0 \end{bmatrix} \in \ker [-V(s) \quad R(s)] = \operatorname{im} \begin{bmatrix} sG - F \\ H \end{bmatrix}, \qquad (12)$$

so that

$$\begin{bmatrix} x \\ 0 \end{bmatrix} = \begin{bmatrix} sG - F \\ H \end{bmatrix} f(s) \qquad (13)$$

for some rational $f(s)$ (which is uniquely determined, since certainly $\begin{bmatrix} sG - F \\ H \end{bmatrix}$ has full column rank as a rational matrix). It follows from the assumptions that $\begin{bmatrix} sG - F \\ H \end{bmatrix}$ has a polynomial left inverse; with (13) it then follows that $f(s)$ is polynomial. On the other hand, we may also rewrite (13) as

$$\begin{bmatrix} s^{-1}x \\ 0 \end{bmatrix} = \begin{bmatrix} G - s^{-1}F \\ H \end{bmatrix} f(s). \qquad (14)$$

Because of the full-column-rank assumption on $\begin{bmatrix} G \\ H \end{bmatrix}$, the matrix $\begin{bmatrix} G - s^{-1}F \\ H \end{bmatrix}$ has a proper left inverse, and so it follows from (14) that $f(s)$ is strictly proper. Consequently, we must have $f(s) = 0$, which implies that $x = 0$. □

Lemma 3.2. Suppose that (5) holds for a triple of constant matrices (F, G, H) and a pair of polynomial matrices $(R(s), V(s))$. If G has full row rank then the columns of $V(s)$ all belong to X_R.

Proof. If G has full row rank then $G - s^{-1}F$ has a proper right inverse, so $sG - F$ has a strictly proper right inverse, say $(sG - F)Q(s) = I$. From the equation $V(s)(sG - F) = R(s)H$, it follows that $V(s) = R(s)HQ(s)$. Since $HQ(s)$ is strictly proper, this shows that the columns of $V(s)$ are in X_R. □

Lemma 3.3. Suppose that (5) holds for a triple of constant matrices (F, G, H) and a pair of polynomial matrices $(R(s), V(s))$. Suppose also that $sG - F$ has full row rank and $\begin{bmatrix} sG - F \\ H \end{bmatrix}$ full column rank (as rational matrices), and that $[-V(s) \;\; R(s)]$ has full row rank for all $s \in \mathbb{C}$. Under these conditions, the matrix $R(s)$ has full row rank, and we have the inequality

$$\dim X_R \le \operatorname{rank} G. \qquad (15)$$

Proof. Denote the number of rows of F and G by n, the number of columns of F and G by $n + m$, the number of rows of H by q, and the number of rows of $V(s)$ and $R(s)$ by p. Note that, under the stated conditions, m is necessarily nonnegative, and that we must have $p = q - m$ by the full-rank assumptions and the equality (5). The fact that $R(s)$ has full row rank follows from (5) and Lemma 1.7. We assumed that $[-V(s) \;\; R(s)]$ is right-unimodular, and so there exist polynomial matrices $U_1(s)$ and $U_2(s)$, of sizes $(n + m) \times n$ and $(n \times m) \times q$ respectively, such that the matrix

$$\begin{bmatrix} U_1(s) & U_2(s) \\ -V(s) & R(s) \end{bmatrix}$$

is unimodular. Define polynomial matrices $T(s)$ and $Z(s)$ by

$$T(s) := \begin{bmatrix} U_1(s) & U_2(s) \\ -V(s) & R(s) \end{bmatrix} \begin{bmatrix} sG - F & 0 \\ H & I \end{bmatrix}$$
$$:= \begin{bmatrix} Z(s) & U_2(s) \\ 0 & R(s) \end{bmatrix}. \qquad (16)$$

Note that $Z(s)$ has size $(n + m) \times (n + m)$ and that $\operatorname{rank} Z(s) = \operatorname{rank} \begin{bmatrix} sG - F \\ H \end{bmatrix} = n + m$; so $Z(s)$ is nonsingular. In order to establish the inequality (15), we shall bound the sum of the minimal row degrees of the polynomial matrix $T(s)$ in two ways. Denote the sum of the minimal row degrees of a given full-row-rank polynomial matrix $M(s)$ by $\nu(M)$. As is well known (cf. Verghese and Kailath, 1981; Willems, 1986; Kuijper and Schumacher, 1990), $\nu(M)$ is equal both to $\dim X_M$ and to the maximum of the degrees of the full-size minors of $M(s)$. For ease of notation, let us assume that $R(s) = [R_1(s)\ R_2(s)]$, where $R_1(s)$ has maximal degree among the $p \times p$ minors of $R(s)$. Then we have

$$\dim X_R = \deg R_1(s) \leq \deg \begin{bmatrix} Z(s) & U_{21}(s) \\ 0 & R_1(s) \end{bmatrix} \leq v(T), \quad (17)$$

since the matrix appearing between the inequality signs is one of the full-size minors of $T(s)$. On the other hand, since the sum of the minimal row degrees is invariant under left multiplication by unimodular matrices, we also have

$$v(T) \leq v\left(\begin{bmatrix} sG - F & 0 \\ H & I \end{bmatrix}\right) \leq \operatorname{rank} G, \quad (18)$$

just by counting the row degrees. This completes the proof. □

Corollary 3.4. Suppose that (5) holds for a triple of constant matrices (F, G, H) and a pair of polynomial matrices $(R(s), V(s))$. Suppose also that G and $\begin{bmatrix} G \\ H \end{bmatrix}$ have full row rank and full column rank respectively, and that $\begin{bmatrix} sG - F \\ H \end{bmatrix}$ and $[-V(s)\ R(s)]$ have full column rank and full row rank respectively for all $s \in \mathbb{C}$. Under these conditions, the columns of the matrix $V(s)$ form a basis for the space X_R.

Proof. It follows from Lemma 3.1 that the columns of $V(s)$ are linearly independent. When G has full row rank then certainly $sG - F$ has full row rank as a rational matrix, and so Lemma 3.3 shows that the number of columns of $V(s)$ is at least equal to $\dim X_R$. On the other hand, it follows from Lemma 3.2 that all columns of $V(s)$ are in X_R. Taking everything together, we get the desired conclusion. □

Remark 3.5. The above corollary is a generalization of a well-known result due to Hautus and Heymann (1978, Theorem 4.10): if (C, A) is an observable pair and $C(sI - A)^{-1} = D^{-1}(s)N(s)$, where $D(s)$ and $N(s)$ are left-coprime polynomial matrices, then the columns of $N(s)$ form a basis for X_D. To get this from the corollary above, take G equal to the identity.

We need a matrix decomposition lemma that is a continuation of Lemma 2.1 in the sense that we now further subdivide the left upper block appearing in the lemma.

Lemma 3.6. If F and G be matrices of size $n \times (n + m)$ such that $sG - F$ has full row rank as a rational matrix then there exists a strictly equivalent pencil $s\tilde{G} - \tilde{F}$ of the form

$$s\tilde{G} - \tilde{F} = \begin{bmatrix} sG_1 - F_1 & 0 \\ 0 & sG_2 - I \end{bmatrix}, \quad (19)$$

where G_1 has full row rank and G_2 is nilpotent.

Proof. Again the result can be read off from the Kronecker canonical form. Because of the full-row-rank assumption on $sG - F$, there are no column minimal indices. The claimed properties are obtained by letting the block $sG_1 - F_1$ correspond to the finite elementary divisors and the row minimal indices, and the block $sG_2 - I$ to the infinite elementary divisors. □

Theorem 3.7. For every matrix triple $(F, G, H) \in \mathbb{R}^{n\times(n+m)} \times \mathbb{R}^{n\times(n+m)} \times \mathbb{R}^{(p+m)\times(n+m)}$, there exists an eligible pair $(R(s), V(s))$ such that $\mathcal{B}(F, G, H) = \mathcal{B}(R, V)$.

Proof. We can construct (see e.g. Kailath, 1980, Theorem 6.3-2)) a unimodular matrix $U(s)$ such that

$$\begin{bmatrix} U_{11}(s) & U_{12}(s) \\ U_{21}(s) & U_{22}(s) \end{bmatrix} \begin{bmatrix} sG - F \\ H \end{bmatrix} = \begin{bmatrix} Z(s) \\ 0 \end{bmatrix}, \quad (20)$$

where $Z(s)$ has full row rank. Now define $V(s) = -U_{21}(s)$ and $R(s) = U_{22}(s)$. Then (5) holds, and $[-V(s)\ R(s)]$ has full row rank for all $s \in \mathbb{C}$.

By Theorem 2.3, we may assume that the triple (F, G, H) satisfies conditions (i)–(iii) of that theorem. Property (i) of Theorem 2.3 now implies that $R(s)$ has full row rank (use Lemma 1.7). The independence of the columns of $V(s)$ follows from Lemma 3.1 together with properties (ii) and (iii) of Theorem 2.3.

It remains to show property (iv) in the definition of eligibility. By the matrix decomposition Lemma 3.6, we may suppose that the pencil $sG - F$ has the form of (19), with G_1 of full row rank and G_2 nilpotent. Denote the size of G_2 by $n_{\text{p-imp}}$ and the number of rows of G_1 by n_{sm}; then $n = n_{\text{sm}} + n_{\text{p-imp}}$, and the number of columns of G_1 is $n_{\text{sm}} + m$. Accordingly, write $V(s) = [V_1(s)\ V_2(s)]$, where $V_1(s)$ has size $p \times n_{\text{sm}}$ and $V_2(s)$ has size $p \times n_{\text{p-imp}}$. We then have, with $H = [H_1\ H_2]$,

$$\operatorname{im} \begin{bmatrix} sG_1 - F_1 & 0 \\ 0 & sG_2 - I \\ H_1 & H_2 \end{bmatrix} = \ker [-V_1(s)\quad -V_2(s)\quad R(s)]. \quad (21)$$

Note that both $[-V_1(s)\ R(s)]$ and $[-V_2(s)\ R(s)]$ have full row rank, because $R(s)$ has full row rank. Also, because the matrix $\begin{bmatrix} sG-F \\ H \end{bmatrix}$ as a whole has full column rank, it is immediate that both $\begin{bmatrix} sG_1-F_1 \\ H_1 \end{bmatrix}$ and $\begin{bmatrix} sG_2-I \\ H_2 \end{bmatrix}$ have full column rank as well. Therefore we may conclude from (21) that

$$\operatorname{im}\begin{bmatrix} sG_1-F_1 \\ H_1 \end{bmatrix} = \ker[-V_1(s)\quad R(s)], \qquad (22)$$

$$\operatorname{im}\begin{bmatrix} sG_2-I \\ H_2 \end{bmatrix} = \ker[-V_2(s)\quad R(s)]. \qquad (23)$$

The equation (23) implies that $V_2(s) = R(s)H_2(sG_2-I)^{-1}$. Because G_2 is nilpotent, $(sG_2-I)^{-1}$ is polynomial. If we denote $T(s) = H_2(sG_2-I)^{-1}$ then

$$\frac{T(s)-T(0)}{s} = H_2\left[\frac{(sG_2-I)^{-1}+I}{s}\right]$$
$$= H_2(sG_2-I)^{-1}G_2 = T(s)G_2, \qquad (24)$$

so that $\operatorname{span}_{\mathbb{R}} T(s)$ is shift-invariant. Finally, it follows from (22) and Corollary 3.4 that the columns of $V_1(s)$ span the space X_R. Indeed, note that the column-rank assumptions in Corollary 3.4 are automatically satisfied as a consequence of the corresponding assumptions on $\begin{bmatrix} sG-F \\ H \end{bmatrix}$; moreover, we can write

$$[-V_1(s)\quad -V_2(s)\quad R(s)]$$
$$= [-V_1(s)\quad R(s)]\begin{bmatrix} I & 0 & 0 \\ 0 & -H_2(G_2-I)^{-1} & I \end{bmatrix}, \qquad (25)$$

which shows that $[-V_1(s)\ R(s)]$ must have full row rank for all s. This completes the proof of the theorem. □

For a given polynomial matrix $R(s) \in \mathbb{R}^{p\times q}[s]$, we define the *smooth* behavior determined by $R(s)$ (on a given interval (t_1, t_2)) by

$$\mathcal{B}_{\text{sm}}(R) = \left\{ w \in C^\infty(t_1, t_2; \mathbb{R}^q) \;\middle|\; R\left(\frac{d}{dt}\right)w = 0 \right\}. \qquad (26)$$

We have the following proposition, which is analogous to Proposition 4.7 of Part I.

Proposition 3.8. If $(R(s), V(s)) \in \mathbb{R}^{p\times q}[s] \times \mathbb{R}^{p\times n}[s]$ is an eligible pair then

$$\mathcal{B}_{\text{sm}}(R) = \mathcal{B}(R, V) \cap \mathscr{C}^q_{\text{sm}}. \qquad (27)$$

Proof. First, take $w \in \mathcal{B}(R, V) \cap \mathscr{C}^q_{\text{sm}}$. By taking smooth and impulsive parts in the equation $R(p)w = V(p)x$ (cf. (30) in Part I), we immediately see that $R(d/dt)w = 0$, so that $w \in \mathcal{B}_{\text{sm}}(R)$. In order to prove the reverse inclusion, construct a first-order realization (F, G, H) of $(R(s), V(s))$ as in the proof of the realization theorem of Part I (Theorem 5.15). With respect to the triple (F_1, G_1, H_1), it was shown in Kuijper and Schumacher (1990) that

$$\mathcal{B}_{\text{sm}}(R) = \mathcal{B}_{\text{sm}}(F_1, G_1, H_1)$$
$$:= \{w \in \mathscr{C}^q_{\text{sm}} \mid \exists z \in \mathscr{C}^{n_1+m}_{\text{sm}}$$
$$\text{s.t. } G_1\dot{z} = F_1 z,\ w = H_1 z\}. \qquad (28)$$

(Actually the proof in Kuijper and Schumacher (1990) was given for the case $t_1 = -\infty$, $t_2 = +\infty$, but the proof carries over verbatim to the general case.) Now, let $w \in \mathcal{B}_{\text{sm}}(R)$, and let z_1 be as in (28). We then have

$$p\begin{bmatrix} G_1 & 0 \\ 0 & G_2 \end{bmatrix}\begin{bmatrix} z_1 \\ 0 \end{bmatrix} = \begin{bmatrix} F_1 & 0 \\ 0 & F_2 \end{bmatrix}\begin{bmatrix} z_1 \\ 0 \end{bmatrix} + \begin{bmatrix} G_1 z(t_1^+) \\ 0 \end{bmatrix}, \qquad (29)$$

$$w = [H_1\quad H_2]\begin{bmatrix} z_1 \\ 0 \end{bmatrix}, \qquad (30)$$

so that $w \in \mathcal{B}^{s/s}(F, G, H) \subset \mathcal{B}(F, G, H) = \mathcal{B}(R, V)$. □

Another way to get $\mathcal{B}_{\text{sm}}(R)$ from $\mathcal{B}(R, V)$ is given in the proposition below. Recall the space of polynomial vectors X_R introduced in Definition 1.5. From this space, we can construct in an obvious way a space of purely impulsive distributions, which will be denoted by $X_R(p)$.

Proposition 3.9. If $R(s),\ V(s)) \in \mathbb{R}'^{p\times q}[s] \times \mathbb{R}^{p\times n}[s]$ is an eligible pair, then

$$\mathcal{B}_{\text{sm}}(R) = \{w \in \mathscr{C}^q_{\text{sm}} \mid R(p)w \in X_R(p)\}. \qquad (31)$$

Proof. If $w \in \mathscr{C}^q_{\text{sm}}$ and $R(p)w \in X_R(p)$ then, in particular, $R(p)w$ is polynomial, and so it follows (cf. (30) in Part I) that $R(d/dt)w = 0$. Conversely, suppose that $w \in \mathscr{C}^q_{\text{sm}}$ and $R(d/dt)w = 0$. Appealing to the proof of the realization theorem in Part I (Theorem 5.15), we find as in the proof above, that there exists a smooth z_1 such that

$$\begin{bmatrix} pG_1-F_1 \\ H_1 \end{bmatrix} z_1 = \begin{bmatrix} G_1 z_1(t_1^+) \\ w \end{bmatrix}. \qquad (32)$$

Still with the notation of the proof of the realization theorem, the above implies that

$$[-V_1(p)\quad R(p)]\begin{bmatrix} G_1 z_1(t_1^+) \\ w \end{bmatrix} = 0, \qquad (33)$$

or, in other words, $R(p)w = V_1(p)G_1z_1(t_1^+)$. Since the columns of $V_1(s)$ span the space X_R by construction, it follows that $R(p)w \in X_R(p)$. □

It is well known how to describe the non-uniqueness of polynomial representations of smooth behaviors: if $R_1(s)$ and $R_2(s)$ are both polynomial matrices of full row rank then $\mathscr{B}_{\mathrm{sm}}(R_1) = \mathscr{B}_{\mathrm{sm}}(R_2)$ if and only if there exists a unimodular matrix $U(s)$ such that $R_2(s) = U(s)R_1(s)$. (This is shown for $C^\infty(\mathbb{R})$ functions in Schumacher (1988, Lemma 2.4); again the proof carries over without change to the case in which the domain of definition is an arbitrary open interval.) A very similar result can be shown for eligible polynomial representations of impulsive–smooth behaviors.

Theorem 3.10. Let $(R_1(s), V_1(s))$ and $(R_2(s), V_2(s))$ be two eligible pairs. We then have $\mathscr{B}(R_1, V_1) = \mathscr{B}(R_2, V_2)$ if and only if there exist a unimodular matrix $U(s)$ and a nonsingular matrix S such that $R_2(s) = U(s)R_1(s)$ and $V_2(s) = U(s)V_1(s)S$.

Proof. The 'if' part was already proved in Lemma 5.12 of Part I, so it remains to prove the 'only if' part. By taking intersections with $\mathscr{C}^q_{\mathrm{sm}}$, we get $\mathscr{B}_{\mathrm{sm}}(R_1) = \mathscr{B}_{\mathrm{sm}}(R_2)$, and so it follows from Schumacher (1988, Lemma 2.4) that there exists a unimodular polynomial matrix $U(s)$ such that $R_2(s) = U(s)R_1(s)$. Hence we may assume for the rest of the proof that $R_1(s) = R_2(s) =: R(s)$. Since both $V_1(s)$ and $V_2(s)$ are supposed to have linearly independent columns, it will suffice to show that $\mathrm{span}_{\mathbb{R}}\, V_1(s) = \mathrm{span}_{\mathbb{R}}\, V_2(s)$.

By taking intersections with $\mathscr{C}^q_{\text{p-imp}}$, we see that we must have $\mathscr{P}_1 = \mathscr{P}_2$, where $\mathscr{P}_1$ and $\mathscr{P}_2$ are defined by

$$\mathscr{P}_i = \{f(s) \in \mathbb{R}^q[s] \mid R(s)f(s) \in \mathrm{span}_{\mathbb{R}}\, V_i(s)\} \qquad (i = 1,2). \quad (34)$$

Take $g(s) \in \mathrm{span}_{\mathbb{R}}\, V_1(s)$. Because $V_1(s)$ is, by definition, of the form $V_1(s) = R(s)Q(s)$ for some rational matrix $Q(s)$, there exists a rational vector $q(s)$ such that $g(s) = R(s)q(s)$. Denote by $q_+(s)$ and $q_-(s)$ the polynomial and strictly proper parts of $q(s)$ respectively. Note that $R(s)q_-(s) = g(s) - R(s)q_+(s)$ is polynomial, so $R(s)q_-(s) \in X_R \subset \mathrm{span}_{\mathbb{R}}\, V_1(s)$. It follows that $R(s)q_+(s) \in \mathrm{span}_{\mathbb{R}}\, V_1(s)$, so that $R(s)q_+(s) \in \mathscr{P}_1 = \mathscr{P}_2$. Therefore $R(s)q_+(s)$ belongs to $\mathrm{span}_{\mathbb{R}}\, V_2(s)$. Because we also have $R(s)q_-(s) \in X_R \subset \mathrm{span}_{\mathbb{R}}\, V_2(s)$, we get $g(s) = R(s)q(s) \in \mathrm{span}_{\mathbb{R}}\, V_2(s)$. This shows that $\mathrm{span}_{\mathbb{R}}\, V_1(s) \subset \mathrm{span}_{\mathbb{R}}\, V_2(s)$; the reverse inclusion follows by symmetry. □

Remark 3.11. Let $\mathscr{B}$ be an impulsive–smooth behavior, i.e. $\mathscr{B} = \mathscr{B}(F, G, H)$ for some triple of matrices (F, G, H). It follows from the above that if $\mathscr{B} = \mathscr{B}(R_1, V_1) = \mathscr{B}(R_2, V_2)$, where (R_1, V_1) and (R_2, V_2) are both eligible pairs, then the matrices $V_1(s)$ and $V_2(s)$ must have the same size, say $p \times n$. In other words, the numbers p and n are determined by the behavior rather than by the specific choice of a representation. It can be seen that p is the number of outputs in an input–output representation of $\mathscr{B}$. The number n might be called the *full dynamic order* of $\mathscr{B}$. The theorem above also shows, in particular, that all invariants of $R(s)$ under left multiplication by unimodular matrices are intrinsic characteristics of the behavior. One of these is, for instance, the sum of the minimal row indices of $R(s)$, which might be called the *smooth dynamic order* of the system, since it gives the dimension of the state space in a minimal state-space representation in $\mathscr{B}_{\mathrm{sm}}(R)$.

4. STATE-SPACE ISOMORPHISM

In this section, we obtain necessary and sufficient conditions for minimality of first-order representations of impulsive–smooth behaviors, and prove that minimal representations in pencil form are related by similarity transformations. It is actually convenient to prove the state-space isomorphism theorem first.

Theorem 4.1. If the matrix triples (F_1, G_1, H_1) and (F_2, G_2, H_2) both satisfy conditions (i)–(iii) of Theorem 2.3 then $\mathscr{B}(F_1, G_1, H_1) = \mathscr{B}(F_2, G_2, H_2)$ if and only if there exist nonsingular matrices S and T such that $F_1 = SF_2T^{-1}$, $G_1 = SG_2T^{-1}$ and $H_1 = H_2T^{-1}$.

Proof. The 'if' part was already shown in Lemma 4.2 of Part I, so it remains to show the 'only if' part. By Theorem 3.7, we can find eligible pairs (R_1, V_1) and (R_2, V_2) such that $\mathscr{B}(R_1, V_1) = \mathscr{B}(F_1, G_1, H_1) = \mathscr{B}(F_2, G_2, H_2) = \mathscr{B}(R_2, V_2)$. By the uniqueness theorem for polynomial representations, there exist a unimodular polynomial matrix $U(s)$ and a nonsingular constant matrix S such that $V_2(s) = U(s)V_1(s)S$ and $R_2(s) = U(s)R_1(s)$, so that

$$\ker\,[-V_2(s)S^{-1} \quad R_2(s)] = \ker\,[-V_1(s) \quad R_1(s)]. \quad (35)$$

Now using (5), which holds by construction of the proof of Theorem 3.7, we get

$$\mathrm{im}\begin{bmatrix} sSG_2 - SF_2 \\ H_2 \end{bmatrix} = \mathrm{im}\begin{bmatrix} sG_1 - F_1 \\ H_1 \end{bmatrix}. \quad (36)$$

This implies that there exists a nonsingular rational matrix $T(s)$ such that

$$\begin{bmatrix} sSG_2 - SF_2 \\ H_2 \end{bmatrix} = \begin{bmatrix} sG_1 - F_1 \\ H_1 \end{bmatrix} T(s). \qquad (37)$$

Note that the matrix $T(s)$ is uniquely determined, because $\begin{bmatrix} sG_1 - F_1 \\ H_1 \end{bmatrix}$ has full column rank. Since the matrix multiplying $T(s)$ has full column rank for all s, it has a polynomial left inverse, and so it follows that $T(s)$ must be polynomial. But, on the other hand, we can also write (37) as

$$\begin{bmatrix} SG_2 - s^{-1}SF_2 \\ H_2 \end{bmatrix} = \begin{bmatrix} G_1 - s^{-1}F_1 \\ H_1 \end{bmatrix} T(s). \qquad (38)$$

This time the matrix multiplying $T(s)$ is proper rational and has full column rank at infinity. Therefore it has a proper rational left inverse, and it follows that $T(s)$ must be proper rational. So we can conclude that $T(s)$ is in fact a constant matrix T, and the proof is complete. □

Theorem 4.2. A matrix triple $(F, G, H) \in \mathbb{R}^{n \times (n+m)} \times \mathbb{R}^{n \times (n+m)} \times \mathbb{R}^{q \times (n+m)}$ satisfies conditions (i)–(iii) of Theorem 2.3 if and only if it is a minimal representation of its associated behavior in the following sense: if $(\tilde{F}, \tilde{G}, \tilde{H}) \in \mathbb{R}^{\tilde{n} \times (\tilde{n}+\tilde{m})} \times \mathbb{R}^{\tilde{n} \times (\tilde{n}+\tilde{m})} \times \mathbb{R}^{q \times (\tilde{n}+\tilde{m})}$ represents the same behavior as (F, G, H) (in the sense of (3)) then $\tilde{n} \geq n$ and $\tilde{m} \geq m$.

Proof. Conditions (i)–(iii) were already shown to be necessary in Theorem 2.3. For the converse part of the proof, note that it has been shown in the proof of Theorem 2.3 that every triple (F, G, H) can be replaced by an equivalent one that satisfies conditions (i)–(iii), by operations that *reduce* the size of the representation. The claim now follows from the fact shown above that all representations satisfying conditions (i)–(iii) are of the same size. □

The above corollary identifies conditions (i)–(iii) of Theorem 2.3 as minimality conditions for pencil representations of impulsive–smooth behaviors. Likewise, conditions (i)–(iv) of Theorem 2.6 represent the minimality conditions for conventional representations. We show this by reduction to the case of pencil representations by means of the following lemma.

Lemma 4.3. Consider a matrix triple $(F, G, H) \in \mathbb{R}^{n \times (n+m)} \times \mathbb{R}^{n \times (n+m)} \times \mathbb{R}^{q \times (n+m)}$ that satisfies conditions (i)–(iv) of Theorem 2.6. If a triple $(\tilde{F}, \tilde{G}, \tilde{H})$ is constructed from (F, G, H) by the method of the proof of Proposition 4.4 of Part I then the resulting triple satisfies conditions (i)–(iii) of Theorem 2.3. Let $\tilde{F}$ and $\tilde{G}$ be of size $\tilde{n} \times (\tilde{n} + \tilde{m})$. Then $\tilde{m} = m$, and $n - \tilde{n}$ is equal to the number of infinite elementary divisors of the pencil $s\tilde{G} - \tilde{F}$.

Proof. We begin with the second claim. For this, it is convenient to bring the pencil $sG - F$ into Kronecker canonical form by a suitable change of bases (Gantmacher, 1959, Chap. XII). Because of the full-row-rank condition on $sG - F$, the canonical form will be block-diagonal, with only three types of blocks, namely those corresponding to column minimal indices, those corresponding to finite elementary divisors, and those corresponding to infinite elementary divisors. The condition $F(\ker G) \subset \operatorname{im} G$ means that the blocks corresponding to infinite elementary divisors have size at least two. The construction of the proof of the cited proposition applies to the blocks separately. As is easily verified, the construction leaves the blocks corresponding to finite elementary divisors and column minimal indices unaffected, whereas it reduces the size of the blocks corresponding to infinite elementary divisors by one. Since the blocks of this type have size at least two, the number of infinite elementary divisors of $s\tilde{G} - \tilde{F}$ is the same as that of $sG - F$. We conclude that both the number of rows and the number of columns of $\tilde{F}$ and $\tilde{G}$ are less than the corresponding numbers for F and G by an amount that is equal to the number of infinite elementary divisors of $sG - F$ or equivalently of $s\tilde{G} - \tilde{F}$, as claimed in the lemma.

It remains to show that the triple $(\tilde{F}, \tilde{G}, \tilde{H})$ satisfies conditions (i)–(iii) of Theorem 2.3. Concerning the full-row-rank condition, this follows because the types of the Kronecker blocks of $(\tilde{F}, \tilde{G}, \tilde{H})$ are the same as those of (F, G, H). The full-column-rank conditions are obviously satisfied, because the mappings $\tilde{F}$, $\tilde{G}$ and $\tilde{H}$ are constructed as restrictions of the corresponding mappings F, G and H. □

Theorem 4.4. A matrix triple $(F, G, H) \in \mathbb{R}^{n \times (n+m)} \times \mathbb{R}^{n \times (n+m)} \times \mathbb{R}^{q \times (n+m)}$ satisfies conditions (i)–(iv) of Theorem 2.6 if and only if it is a minimal representation of its associated behavior in the following sense: if $(\tilde{F}, \tilde{G}, \tilde{H}) \in \mathbb{R}^{\tilde{n} \times (\tilde{n}+\tilde{m})} \times \mathbb{R}^{\tilde{n} \times (\tilde{n}+\tilde{m})} \times \mathbb{R}^{q \times (\tilde{n}+\tilde{m})}$ represents the same behavior as (F, G, H) (in the sense of (2)) then $\tilde{n} \geq n$ and $\tilde{m} \geq m$.

Proof. As in the proof of Corollary 4.2, it suffices to show that all representations satisfying conditions (i)–(iv) are of the same size. This follows from the previous lemma by noting that minimal pencil representations are uniquely determined up to left and right similarity transformations (Theorem 4.1), which implies in particular that the number of infinite elementary divisors of the corresponding pencils is fixed. □

Remark 4.5. Unfortunately, one cannot expect a state-space isomorphism theorem analogous to Theorem 4.1 to be true for minimal conventional representations. The following example shows this. Define a triple (F, G, H) by

$$G = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad F = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \tag{39}$$

and a triple $(\tilde{F}, \tilde{G}, \tilde{H})$ by $\tilde{G} = G$, $\tilde{F} = F$ and

$$\tilde{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}. \tag{40}$$

It is clear that $\mathscr{B}_c(F, G, H) = \mathscr{B}_c(\tilde{F}, \tilde{G}, \tilde{H})$, and one easily verifies directly that both triples are minimal conventional representations. However, there do *not* exist nonsingular matrices S and T such that $\tilde{F} = SFT$, $\tilde{G} = SGT$ and $\tilde{H} = HT$. Indeed, if this held then we should have $S = T^{-1}$ and so $GT = T\tilde{G}$, which implies that

$$T = \begin{bmatrix} \alpha & \beta & \gamma \\ 0 & \alpha & \beta \\ 0 & 0 & \alpha \end{bmatrix}, \tag{41}$$

with $\alpha, \beta, \gamma \in \mathbb{R}$ and $\alpha \neq 0$. The final condition $HT = \tilde{H}$ then becomes

$$\begin{bmatrix} \alpha & \beta & \gamma \\ 0 & \alpha & \beta \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \tag{42}$$

and this cannot be satisfied.

Remark 4.6. For $n \geq 0$, $m \geq 0$ and $p > 0$, denote by $\tilde{R}_{n,m,p}$ the set of all matrix triples $(F, G, H) \in \mathbb{R}^{n\times(n+m)} \times \mathbb{R}^{n\times(n+m)} \times \mathbb{R}^{(p+m)\times(n+m)}$ that satisfy the minimality conditions (i)–(iii) of Theorem 2.3. We may equip this set with the topology that it inherits from its ambient Euclidean space. Denote by $R_{n,m,p}$ the quotient of $\tilde{R}_{n,m,p}$ with respect to the action of $\mathrm{GL}(n) \times \mathrm{GL}(n+m)$ defined by $(S, T)\colon (F, G, H) \mapsto (SFT^{-1}, SGT^{-1}, HT^{-1})$. According to Theorem 4.1, the quotient space is exactly the set of impulsive–smooth behaviors of a fixed dynamic order n, a fixed number of inputs m and a fixed number of outputs p. A natural topology to consider on $R_{n,m,p}$ is the quotient topology, and it is of interest to study the properties of the resulting topological space.

Triples (F, G, H) satisfying the minimality conditions of Theorem 2.3 are called *completely observable linear systems* by Lomadze (1990). He notes that this set (actually its complexification) can be given the structure of a compact algebraic variety by looking at it as a particular case of the so-called Quot schemes considered by Grothendieck. As pointed out recently (Ravi and Rosenthal, 1995), the same fact has already been established in the algebraic geometric literature by Strømme, who shows that the space $R_{n,m,p}$ provides a nonsingular projective compactification of the space of rational curves of degree n from the Riemann sphere into the Grassmannian space of m-dimensional subspaces of $(m+p)$-dimensional linear space (Strømme 1987, Theorem 2.1). From the characterization given by Strømme, it can be readily seen that this space is the same as the quotient space of homogeneous polynomial matrices that was independently shown by Ravi and Rosenthal (1994) to be a smooth and compact manifold. Starting from a state-space representation, Helmke and Shayman (1991) proved independently that $R_{n,0,p}$ is smooth and compact. In some other special cases, the structure of the space R can be seen easily; in particular, $R_{0,m,p}$ is just the Grassmannian manifold of m-dimensional subspaces of $(m+p)$-dimensional linear space, and $R_{1,m,1}$ is diffeomorphic to the projective space of dimension $2m+1$.

In this paper, an alternative representation of the space $R_{n,m,p}$ has been given in terms of eligible pairs $(R(s), V(s))$. We also have defined behaviors associated to completely observable triples; this problem was naturally not addressed by Strømme (1987), and Lomadze (1990, p. 153) has only a brief remark on it.

5. DYNAMIC INTERPRETATION

The minimality conditions can be interpreted in terms of left-invertibility properties and solvability conditions (cf. Geerts, 1993a,c). We shall do this first for the conventional representation

$$pGz = Fz + Gz_0, \tag{43}$$

$$w = Hz. \tag{44}$$

We begin with a number of definitions (cf. Geerts and Mehrmann, 1990, Geerts, 1993c). These are stated for general (not necessarily minimal) representations; in particular, the index m might be negative.

Definition 5.1. Consider the equation (43). The *solution set corresponding to a given point*

$z_0 \in \mathbb{R}^{n+m}$ is

$$\mathcal{S}(z_0) = \{z \in \mathscr{C}_{\mathrm{imp}}^{n+m} \mid (pG - F)z = Gz_0\}. \tag{45}$$

The *solution space* is the space $\mathcal{S} = \mathcal{S}(F, G)$ given by

$$\mathcal{S} = \{z \in \mathscr{C}_{\mathrm{imp}}^{n+m} \mid \exists z_0 \in \mathbb{R}^{n+m}: (pG - F)z = z_0\}$$

$$= \bigcup_{z_0 \in \mathbb{R}^{n+m}} \mathcal{S}(z_0). \tag{46}$$

Definition 5.2. Consider the equation (43). A point $z_0 \in \mathbb{R}^{n+m}$ is said to be *consistent* if there exists $z \in \mathcal{S}(z_0) \cap \mathscr{C}_{\mathrm{sm}}^{n+m}$ such that $z(t_{\mathrm{in}}^+) = z_0$, and *weakly consistent* if $\mathcal{S}(z_0) \cap \mathscr{C}_{\mathrm{sm}}^{n+m}$ is nonempty.

Definition 5.3. A linear system (43), (44) with $F, G \in \mathbb{R}^{n\times(n+m)}$ and $H \in \mathbb{R}^{q\times(n+m)}$, is called *ex-in smooth* if

$$w \in \mathscr{C}_{\mathrm{sm}}^{q} \Rightarrow z \in \mathscr{C}_{\mathrm{sm}}^{n+m}, \tag{47}$$

ex-in nulling if

$$w = 0 \Rightarrow z = 0, \tag{48}$$

and *left-invertible* if

$$w = 0,\ z_0 = 0 \Rightarrow z = 0. \tag{49}$$

The system equation (43) is *solvable in the distribution sense* if

$$\forall z_0 \in \mathbb{R}^{n+m}:\quad \mathcal{S}(z_0) \neq \varnothing, \tag{50}$$

solvable in the function sense if

$$\forall z_0 \in \mathbb{R}^{n+m}:\quad \mathcal{S}(z_0) \cap \mathscr{C}_{\mathrm{sm}}^{n+m} \neq \varnothing, \tag{51}$$

and *solvable in the sense of consistency* if

$$\forall z_0 \in \mathbb{R}^{n+m} \exists z \in \mathcal{S}(z_0) \cap \mathscr{C}_{\mathrm{sm}}^{n+m}:\quad z(t_{\mathrm{in}}^+) = z_0. \tag{52}$$

Remark 5.4. A system that is ex-in smooth might also be called *impulse-observable* or *observable in the sense of Verghese* (cf. Cobb, 1984). The third solvability definition comes down to requiring that every point be consistent, whereas the second one means that every point be weakly consistent. Note that, in the latter case, (43) might be called *impulse-controllable.*

Theorem 5.5. The system (43), (44) is

(i) ex-in smooth if and only if $\ker \begin{bmatrix} G \\ H \end{bmatrix} \cap F^{-1}[\operatorname{im} G] = \{0\}$;

(ii) ex-in nulling if and only if $\ker \begin{bmatrix} G \\ H \end{bmatrix} \cap F^{-1}[\operatorname{im} G] = \{0\}$ and $\begin{bmatrix} sG - F \\ H \end{bmatrix}$ is left-unimodular;

(iii) left-invertible if and only if $\begin{bmatrix} sG - F \\ H \end{bmatrix}$ is left-invertible as a rational matrix.

Now, assume that $[F\ G]$ is of full row rank. The equation (43) is

(iv) solvable in the distribution sense if and only if $sG - F$ is right-invertible;

(v) solvable in the function sense if and only if $\operatorname{im} G + F(\ker G) = \mathbb{R}^n$;

(vi) solvable in the sense of consistency if and only if $\operatorname{im} G = \mathbb{R}^n$.

Proof. The first claim follows from Geerts (1993b, Theorem 3.2). For the second claim, first suppose that the system is ex-in nulling. If $\begin{bmatrix} G \\ H \end{bmatrix} v = 0$ and $Fv = Gz_0$ for some $v, z_0 \in \mathbb{R}^{n+m}$ then $v \in \mathcal{S}(-z_0)$, whereas $w := Hv = 0$. Hence $v = 0$. To show that $S(s) := \begin{bmatrix} sG - F \\ H \end{bmatrix}$ is left-unimodular, suppose that for some $\lambda \in \mathbb{C}$ and $y \in \mathbb{C}^{n+m}$, $S(\lambda)y = 0$. Then the smooth function $\tilde{z}$ defined by $\tilde{z}(t) = e^{\lambda t}y$ satisfies $(p - \lambda)\tilde{z} = y$, and so

$$\begin{aligned} F\tilde{z} + Gy &= (p - \lambda)^{-1}Fy + Gy \\ &= (p - \lambda)^{-1}(Fy + pGy - \lambda Gy) \\ &= pG\tilde{z}, \end{aligned} \tag{53}$$

so that $\tilde{z} \in \mathcal{S}(y)$, and $H\tilde{z} = 0$. It follows that $\tilde{z} = 0$, and consequently $y = (p - \lambda)\tilde{z} = 0$. Conversely, suppose now that the conditions of claim (ii) hold. Assume that $S(p)z = \begin{bmatrix} G\bar{z}_0 \\ 0 \end{bmatrix}$ for some $\bar{z}_0 \in \mathbb{R}^{n+m}$, and suppose that $z \neq 0$. Since $S(s)$ has a polynomial left inverse, it follows that z is impulsive, say $z = z_k p^k + \ldots + z_1 p + z_0$, with $z_k \neq 0$. From the equations, we have $Fz_k = Gz_{k-1}$ if $k > 0$ and $Fz_k = -G\bar{z}_0$ if $k = 0$, and in both cases we also have $Gz_k = 0$ and $Hz_k = 0$. Because $\ker \begin{bmatrix} G \\ H \end{bmatrix} \cap F^{-1}[\operatorname{im} G] = \{0\}$, it follows that $z_k = 0$, so we obtain a contradiction. Therefore we must have $z = 0$.

If the matrix $\begin{bmatrix} sG - F \\ H \end{bmatrix}$ has full column rank over $\mathbb{R}(s)$ then the associated mapping $\begin{bmatrix} pG - F \\ H \end{bmatrix}$ is injective, as noted in Section 5 of Part I. This proves one implication of claim (iii).

For the other part, suppose that $S(s)$ is not left-invertible. Then $S(s)z(s)=0$ for some $z(s)\in\mathbb{R}^{n+m}[s]$, with $z(s)\neq 0$. It follows that $S(p)z(p)=0$ (cf. (30) in Part I), so that the corresponding system is not left-invertible.

Claims (iv) and (v) follow from Geerts (1993c, Corollary 3.6, Theorem 4.5). Finally, to prove the sixth claim, first assume that every point z_0 is consistent, and let $\eta\in\mathbb{R}^{1\times(n+m)}$ be such that $\eta G=0$. Take $z_0\in\mathbb{R}^{n+m}$ and let $z\in\mathscr{C}^{n+m}_{\text{sm}}$ be such that $pGz=Fz+Gz_0$ and $z(t^+_{\text{in}})=z_0$. We then have $\eta Fz=0$ and, in particular, $\eta Fz(t^+_{\text{in}})=\eta Fz_0=0$. Since z_0 was arbitrary, we find that $\eta F=0$, and it follows by the full-row-rank assumption on $[F\ G]$ that $\eta=0$. For the converse, we may assume without loss of generality that $G=[I\ 0]$ and $F=[A\ B]$. If z_0 and z are partitioned accordingly then (43) becomes $pz_1=Az_1+Bz_2+z_{01}$. This equation has, for instance, the following smooth solution: take $z_2=z_{02}p^{-1}$ (the smooth constant function with value z_{02}), and let z_1 be the solution to the ordinary differential equation $\dot z_1=Az_1+Bz_{02}$ with initial condition $z_1(0)=z_{01}$. Note that the initial values of z_1 and z_2 are z_{01} and z_{02} respectively; so indeed there exists a smooth solution with initial value z_0. □

Remark 5.6. Our definition of left invertibility for (43), (44) corresponds to left invertibility in the *strong* sense for *descriptor* systems, i.e. systems

$$p[E\ \ 0]z=[A\quad B]z+Ex_0,\quad \begin{bmatrix} y\\ u\end{bmatrix}=\begin{bmatrix} C & D\\ 0 & I\end{bmatrix}z, \tag{54}$$

with $z=\begin{bmatrix} x\\ u\end{bmatrix}$, where u and x denote input and state variables respectively, as defined in Geerts (1993a, Corollary 4.15). Since we do not separate inputs from states, we cannot make a distinction between invertibility in the weak and the strong sense as is done in Geerts (1993a). Observe that a system (43), (44) is ex-in smooth only if it is left-invertible (cf. Geerts, 1993b, Remark 3.7). On the other hand, left unimodularity alone is not sufficient for a system (43), (44) to be either ex-in smooth or ex-in nulling. This is clear from Corollary 5.7 below. Finally, observe that an external trajectory for an ex-in nulling system can be generated by only one internal trajectory, whereas the initial condition is unique modulo $\ker G$.

Corollary 5.7. Consider (43), (44). The polynomial matrix $S(s):=\begin{bmatrix} sG-F\\ H\end{bmatrix}$ is left-unimodular if and only if $\ker F\cap\ker H=\{0\}$ and $w=0$ with z smooth implies $z_0\in\ker G$.

Proof. If $Fv=0$, $Gv=0$ and $Hv=0$, then $v\in\ker S(s)$ and so $v=0$. Next, if $\begin{bmatrix} pG-F\\ H\end{bmatrix}z=\begin{bmatrix} Gz_0\\ 0\end{bmatrix}$ for some smooth z then z is also impulsive by left unimodularity of $S(s)$. Hence $z=0$. For the converse, let $S(\lambda)y=0$ with $\lambda\in\mathbb{C}$ and $y\in\mathbb{C}^{n+m}$. We must show that $y=0$. Indeed, $Gy=0$ by the proof of claim (ii) of Theorem 5.5. From $S(\lambda)y=0$, we then also get $Fy=0$ and $Hy=0$; so $y=0$ and $S(s)$ is left-unimodular. □

Example 5.8. Let $G=[1\ 0]$, $F=[0\ 1]$ and $H=[1\ 0]$. The matrix $S(s)$ is unimodular, but $\ker\begin{bmatrix} G\\ H\end{bmatrix}\cap F^{-1}[\operatorname{im} G]\neq 0$. The impulse $z=\begin{bmatrix} 0\\ 1\end{bmatrix}$ belongs to $\mathscr{S}\left(\begin{bmatrix} -1\\ 0\end{bmatrix}\right)$, and $w=0$. The system is not ex-in smooth, and hence not ex-in nulling.

Example 5.9. Let $G=[1\ 0]$, $F=[0\ 1]$ and $H=[0\ 1]$. We have $\ker\begin{bmatrix} G\\ H\end{bmatrix}\cap F^{-1}[\operatorname{im} G]=\{0\}$ and $S(s)=\begin{bmatrix} s & -1\\ 0 & 1\end{bmatrix}$, left-invertible but not unimodular. If $w=0$ then $pz_1=z_{01}$ and $z_2=0$. The system is ex-in smooth, but not ex-in nulling.

Definition 5.10. Consider the equation (43). A point $z_0\in\mathbb{R}^{n+m}$ is called *zeroth-order-impulse smoothing* if there exists a $z\in\mathscr{S}(z_0)$ such that $pz-z_0\in\mathscr{S}$. The space of these points is denoted by $\mathscr{K}=\mathscr{K}(F,G)$.

Observe that z in Definition 5.10 is smooth if the impulsive part of $\bar z:=pz-z_0$ is of order 0, i.e. if $\bar z_{\text{p-imp}}=\alpha=\alpha\delta$ with $\alpha\in\mathbb{R}^{n+m}$. If $\bar z_{\text{p-imp}}=0$ then z_0 is consistent. The space $\mathscr{K}$ is easily determined if (43) is solvable in the distribution sense.

Lemma 5.11. Let $\mathscr{S}(z_0)\neq\varnothing$ for every $z_0\in\mathbb{R}^{n+m}$. Then $\mathscr{K}=F^{-1}[\operatorname{im} G]$.

Proof. Let z_0 be such that $pGz=Fz+Gz_0$, with $z=p^{-1}(\bar z+z_0)$ and $pG\bar z=F\bar z+G\bar z_0$ for some $\bar z_0$. Then $pG\bar z=pFz=F\bar z+G\bar z_0=F\bar z+Fz_0$ and $z_0\in F^{-1}[\operatorname{im} G]$. Conversely let $Fz_0=G\bar z_0$ for some $\bar z_0$. There exists a $\bar z\in\mathscr{C}^{n+m}_{\text{imp}}$ such that $pG\bar z=F\bar z+G\bar z_0$. Hence if $z=p^{-1}(\bar z+z_0)$ then $pGz=G\bar z+Gz_0=p^{-1}(F\bar z+Fz_0)+Gz_0=Gz+Gz_0$. □

In Geerts (1993a, Section 3), the following subspaces of $\mathbb{R}^{n+m}$ are defined and characterized.

Definition 5.12. Consider the equations (43) and (44) and denote $\Sigma = (F, G, H)$. A point z_0 is called *weakly unobservable* if there exists a smooth $z \in \mathcal{S}(z_0)$ such that $w = 0$. If, moreover, $z(t_{\text{in}}^+) = z_0$ then z_0 is called weakly unobservable *in the sense of consistency*. The space of the latter points is denoted by $\mathcal{V}_C = \mathcal{V}_C(\Sigma)$ and the space of the former points by $\mathcal{V} = \mathcal{V}(\Sigma)$. If z_0 is such that for an impulsive $z \in \mathcal{S}(z_0)$, $w = 0$ then z_0 is called *strongly controllable*. The space of strongly controllable points is denoted by $\mathcal{W} = \mathcal{W}(\Sigma)$.

Lemma 5.13. The space $\mathcal{V}_C$ is the largest subspace $\mathcal{L}$ such that $\begin{bmatrix} F \\ H \end{bmatrix}\mathcal{L} \subset \begin{bmatrix} G \\ 0 \end{bmatrix}\mathcal{L}$. Also, $\mathcal{V}_C = \begin{bmatrix} F \\ H \end{bmatrix}^{-1}\begin{bmatrix} G\mathcal{V}_C \\ 0 \end{bmatrix}$. The space $\mathcal{W}$ is the smallest subspace $\mathcal{K}$ such that $G^{-1}F[\mathcal{K} \cap \ker H] \subset \mathcal{K}$. Also, $G^{-1}F[\mathcal{W} \cap \ker H] = \mathcal{W}$. Moreover, $\mathcal{V} = \mathcal{V}_C + \ker G$.

Proof. See Geerts (1993a). □

Corollary 5.14. There holds $G(\mathcal{V}_C \cap \mathcal{W}) = G(\mathcal{V}_C) \cap F[\mathcal{W} \cap \ker H] = F(\mathcal{V}_C \cap \mathcal{W})$.

Proof. Since $\mathcal{V}_C \subset \ker H$, the claim follows from Lemma 5.13. □

Theorem 5.15. Consider the conventional system (43), (44), with $[F\ \ G]$ of full row rank. The triple $\Sigma = (F, G, H)$ satisfies the minimality conditions of Theorem 2.6 if and only if

(i) Σ is ex-in nulling and solvable in the distribution sense and

(ii) every weakly consistent point is zeroth-order-impulse smoothing.

Proof. Denote by $\bar{\mathcal{V}}_C$ and $\bar{\mathcal{V}}$ the spaces of *consistent* and *weakly consistent* points respectively (see Definition 5.2). Then $F\bar{\mathcal{V}}_C \subset G\bar{\mathcal{V}}_C$ and $\bar{\mathcal{V}} = \bar{\mathcal{V}}_C + \ker G$ (apply Lemma 5.13 with $H = 0$). Hence the claims $\ker G \subset F^{-1}[\operatorname{im} G]$ and $\bar{\mathcal{V}} \subset F^{-1}[\operatorname{im} G]$ are equivalent. Now combine Theorem 5.5 with Lemma 5.11. □

Remark 5.16. If $\bar{\mathcal{V}}_C = \bar{\mathcal{V}}$ then $\ker G \subset F^{-1}[\operatorname{im} G]$. The converse is not true; if

$$G = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad F = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{55}$$

then $\bar{\mathcal{V}}_C = 0$.

Next we give dynamic interpretations of the minimality conditions for pencil representations. Again we begin with a number of definitions.

Definition 5.17. Consider a linear system

$$pGz = Fz + x_0, \tag{56}$$

$$w = Hz, \tag{57}$$

with $F, G \in \mathbb{R}^{n \times (n+m)}$, $H \in \mathbb{R}^{q \times (n+m)}$ and $x_0 \in \mathbb{R}^n$. For given x_0, define the *solution set* corresponding to (56):

$$\mathcal{S}(x_0) = \{z \in \mathcal{C}_{\text{imp}}^{n+m} \mid (pG - F)z = x_0\}. \tag{58}$$

The system (56), (57) is called *ex-in smooth* if

$$w \in \mathcal{C}_{\text{sm}}^q \Rightarrow z \in \mathcal{C}_{\text{sm}}^{n+m}, \tag{59}$$

ex-in nulling if

$$w = 0 \Rightarrow z = 0, \tag{60}$$

and *left-invertible* if

$$w = 0,\ x_0 = 0 \Rightarrow z = 0. \tag{61}$$

The system equation (56) is *solvable in the distribution sense* if

$$\forall x_0 \in \mathbb{R}^n:\quad \mathcal{S}(x_0) \neq \emptyset, \tag{62}$$

and *solvable in the function sense* if

$$\forall x_0 \in \mathbb{R}^n:\quad \mathcal{S}(x_0) \cap \mathcal{C}_{\text{sm}}^{n+m} \neq \emptyset. \tag{63}$$

Theorem 5.18. The system (56), (57) is

(i) ex-in smooth if and only if $\ker \begin{bmatrix} G \\ H \end{bmatrix} = \{0\}$;

(ii) ex-in nulling if and only if $\ker \begin{bmatrix} G \\ H \end{bmatrix} = \{0\}$ and $\begin{bmatrix} sG - F \\ H \end{bmatrix}$ is left-unimodular;

(iii) left invertible if and only if $\begin{bmatrix} sG - F \\ H \end{bmatrix}$ is left-invertible as a rational matrix.

The equation (43) is

(iv) solvable in the distribution sense if and only if $sG - F$ is right-invertible;

(v) solvable in the function sense if and only if $\operatorname{im} G = \mathbb{R}^n$.

Proof. To prove the first claim, suppose first that the system is ex-in smooth and let $z_0 \in \mathbb{R}^{n+m}$ be such that $Gz_0 = 0$ and $Hz_0 = 0$. Write $x_0 = -Fz_0$, and define $z \in \mathcal{C}_{\text{imp}}^{n+m}$ by $z = z_0$; then we have $pGz = Fz + x_0$ and $w = Hz = 0$, so that $z \in \mathcal{C}_{\text{sm}}$. This implies that $z_0 = 0$. For the converse, suppose that $\begin{bmatrix} G \\ H \end{bmatrix}$ has full column rank, and let

$pGz = Fz + x_0$ with $w = Hz \in \mathscr{C}_{\text{sm}}$. Suppose now that z is not smooth. Then we can write

$$z = \sum_{i=0}^{k} z_i p^i + z_{\text{sm}},$$

with $z_k \neq 0$. It follows from $pGz = Fx + x_0$ that $Gz_k = 0$; because $w \in \mathscr{C}_{\text{sm}}$, we also have $Hz_k = 0$. It follows that $z_k = 0$, and so we have a contradiction.

For the second claim, suppose that our system is ex-in nulling, and suppose that

$$\begin{bmatrix} \lambda G - F \\ H \end{bmatrix} z_0 = 0$$

for some $\lambda \in \mathbb{C}$ and some z_0. Define a smooth function z by $z(t) = e^{\lambda t} z_0$. We then have $\lambda Gz = Fz$, $Hz = 0$ and $pz = \lambda z + z_0$. Consequently, $pGz = \lambda Gz + Gz_0 = Fz + Gz_0$ and $w = Hz = 0$. It follows that $z = 0$, and hence that $z_0 = 0$. The full-column-rank property of $\begin{bmatrix} G \\ H \end{bmatrix}$ follows in a similar way as in the proof of the first claim. For the converse, we assume the rank conditions and we want to show that the system is ex-in nulling. If (56), (57) hold with $w = 0$ then, for some z,

$$\begin{bmatrix} pG - F \\ H \end{bmatrix} z = \begin{bmatrix} x_0 \\ 0 \end{bmatrix},$$

which implies that z is purely impulsive, because $\begin{bmatrix} sG - F \\ H \end{bmatrix}$ has a polynomial left inverse. Suppose that z is nonzero. Then we could write $z = \sum_{i=0}^{k} z_i p^i$, with $z_k \neq 0$. But then $Gz_k = 0$ and $Hz_k = 0$, so that $z_k = 0$ and we have a contradiction.

The third claim is proved just as for conventional representations, so we proceed to claim (iv). One part of this claim is immediate from the fact, noted in Section 5 of Part I, that $pG - F$ is surjective as a mapping from $\mathscr{C}_{\text{imp}}^{n+m}$ to $\mathscr{C}_{\text{imp}}^{n}$ if $sG - F$ is right-invertible. For the converse, suppose that $sG - F$ is not right-invertible. Then there exists a nonzero polynomial row vector $\eta(s)$ such that $\eta(s)(sG - F) = 0$. Now take any $x_0 \in \mathbb{R}^n$ and let $z \in \mathscr{C}_{\text{imp}}^{n+m}$ be such that $(pG - F)z = x_0$. Then $\eta(p)x_0 = \eta(p)(pG - F)z = 0$. Since this holds for any x_0, it follows that $\eta(s) = 0$, and we have a contradiction.

Finally, suppose that (56) is solvable in the function sense, and let $x_0 \in \mathbb{R}^n$ be given. Let z be a smooth solution to $pGz = Fz + x_0$. Then $Gz(t_{\text{in}}^{+}) = x_0$, so that x_0 is in $\operatorname{im} G$. Since x_0 was arbitrary, it follows that G is surjective. Conversely, if G has full row rank then we may choose coordinates such that $G = [I \; 0]$, and write $F = [A \; B]$ in the same coordinates. A smooth solution to the equation $pGz = Fz + x_0$ is then given, for instance, by $z_1(t) = e^{A(t - t_{\text{in}})} x_0$, $z_2(t) = 0$. □

As an immediate consequence, we have the following result.

Theorem 5.19. Consider the pencil representation (56), (57). The triple (F, G, H) satisfies the minimality conditions of Theorem 2.3 if and only if the system is ex-in nulling and solvable in the distribution sense.

6. CONCLUSIONS

We have obtained the conditions for minimality of first-order representations of impulsive-smooth behaviors, and have described the extent to which minimal representations are unique. We have also given dynamic interpretations for the minimality conditions. It turns out that for impulsive–smooth behaviors, the minimality conditions are weaker than for smooth behaviors; in particular, 'controllability at infinity' is not required. Rephrasing the statement of Theorem 5.19, we can say that a pencil representation for an impulsive–smooth behavior is minimal when the mapping from trajectories of auxiliary variables to trajectories of external variables is one-to-one, and the equation for the auxiliary trajectories allows solutions (possibly with an impulsive component) for all initial data. Matrix characterizations of the conditions for minimality have been given, so that these conditions are in principle straightforward to check. The minimality conditions that we have given here are suitable for application in a multimodal context; they show that certain aspects that can be neglected in situations in which mode changes do not occur become important for multimode systems, and must be kept in the system's description.

Throughout the paper, we have used both the 'conventional' representation $pGz = Fz + Gz_0$ and the 'pencil' representation $pGz = Fz + x_0$. As has already been pointed out in Part I, there is no distinction between these two representations from the point of view of descriptive power: every impulsive–smooth behavior that can be represented in one way can also be represented in the other way, although the conventional representation may require more equations and auxiliary variables to describe the same behavior. One reason to use the conventional equation $pGz = Fz + Gz_0$ is that, for smooth solutions z, one has $Gz(0^+) = Gz_0$ (this follow immediately from (1)), so that z_0 may be given the natural interpretation of an initial condition. This interpretation no longer holds when one considers impulsive–smooth solutions, since for such solutions of the form $z = \bar{z}_k p^k + \ldots + \bar{z}_0 + z_{sm}$ one has $Gz_{sm}(0^+) = Gz_0 + F\bar{z}_0$; still, the model $pGz = Fz + Gz(0^-)$ appears to adequately describe jumps in many situations of mode change. On the other hand, the analysis in this paper has shown that there are certain advantages to using the pencil representation rather than the conventional form. In particular, the pencil representation allows a natural state-space isomorphism theorem, whereas this is not true for the conventional representation. Also, the dynamic characterization of minimality comes out nicer for pencil representations than for conventional representations (compare Theorem 5.15 with Theorem 5.19). The price one has to pay is that the vector x_0 in the pencil representation cannot be viewed as an initial condition, but must be considered as 'initial data' resulting from a preceding mode.

It was noted in Remark 4.6 that the class of systems we have studied is compact in the quotient Euclidean topology. Intuitively speaking, this implies that 'one cannot run out of it by taking limits'. The class of standard state-space systems does not have this property. For example, consider the parametrized family of systems $(A_\varepsilon, B_\varepsilon, C_\varepsilon, D_\varepsilon) = (\varepsilon^{-1}, 1, 1, 0)$. As ε tends to zero, there is no convergence to some limit (A_0, B_0, C_0, D_0). On the other hand, the same family can be represented in pencil terms by

$$(F_\varepsilon, G_\varepsilon, H_\varepsilon) = \left([1 \quad 0], \quad [\varepsilon^{-1} \quad 1], \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right),$$

or equivalently

$$(F_\varepsilon, G_\varepsilon, H_\varepsilon) = \left([\varepsilon \quad 0], \quad [1 \quad \varepsilon], \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right).$$

If now ε tends to zero, we get a well-defined limit

$$(F_0, G_0, H_0) = \left([0 \quad 0], \quad [1 \quad 0], \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right).$$

The limit system can not be rewritten as a standard state-space system, but can still be given a dynamic interpretation as a representation of an impulsive–smooth behavior.

Acknowledgements—The second author would like to thank M. S. Ravi for his patient explanation of the contents of Lomadze (1990) and Strømme (1987). The research by Dr Geerts was supported in part by the Dutch Organisation for Scientific Research (NWO), CWI, and Tilburg University.

REFERENCES

Cobb, J. D. (1984). Controllability, observability, and duality in singular systems. *IEEE Trans. Autom. Control*, **AC-29**, 1076–1082.

Gantmacher, F. R. (1959). *The Theory of Matrices*, Vols 1 and 2. Chelsea, New York.

Geerts, T. (1993a). Invariant subspaces and invertibility properties for singular systems: the general case. *Lin. Alg. Applics* **183**, 61–88.

Geerts, T. (1993b). Regularity and singularity in linear-quadratic control subject to implicit continuous-time systems. *Circuits Syst. Sig. Process*, **13**, 19–30.

Geerts, T. (1993c). Solvability conditions, consistency and weak consistency for linear differential–algebraic equations and time-invariant linear systems: the general case. *Lin. Alg. Applics* **181**, 111–130.

Geerts, T. and V. Mehrmann (1990). Linear differential equations with constant coefficients: a distributional approach. Technical Report SFB 343 90-073, Bielefeld University.

Geerts, T. and J. M. Schumacher (1996). Impulsive–smooth behavior in multimode systems. Part I: State-space and polynomial representations. *Automatica*, **32**, – .

Hautus, M. L. J. (1976). The formal Laplace transform for smooth linear systems. In G. Marchesini and S. K. Mitter (Eds), *Mathematical Systems Theory*, pp. 29–47, Lecture Notes in Economics and Mathematical Systems, Vol. 131. Springer-Verlag, New York.

Hautus, M. L. J. and M. Heymann (1978). Linear feedback—an algebraic approach. *SIAM J. Control*, **16**, 83–105.

Hautus, M. L. J. and L. M. Silverman (1983). System structure and singular control. *Lin. Alg. Applics*, **50**, 369–402.

Helmke, U. and M. A. Shayman (1991). Topology of the orbit space of generalized linear systems. *Math. Control, Sig. and Syst.* **4**, 411–437.

Kailath, T. (1980). *Linear Systems*. Prentice-Hall, Englewood Cliffs, NJ.

Kuijper, M. and J. M. Schumacher (1990). Realization of autoregressive equations in pencil and descriptor form. *SIAM J. Control. Optim.*, **28**, 1162–1189.

Lomadze, V. G. (1990). Finite-dimensional time-invariant linear dynamical systems: algebraic theory. *Acta Appl. Math.*, **19**, 149–201.

Ravi, M. S. and J. Rosenthal (1994). A smooth compactification of the space of transfer functions with fixed McMillan degree. *Acta Appl. Math.*, **34**, 329–352.

Ravi, M. S. and J. Rosenthal (1995). A general realization theory for higher order linear differential equations. *Syst. Control Lett.*, **25**, 351–360.

Schumacher, J. M. (1988). Transformations of linear systems under external equivalence. *Lin. Algebra Applics*, **102**, 1–33.

Schumacher, J. M. (1989). Linear system representations. In H. Nijmeijer and J. M. Schumacher (Eds), *Three Decades of Mathematical System Theory. A Collection of Surveys at the Occasion of the 50th Birthday of Jan C. Willems*, pp. 382–408. Lecture Notes in Control and Information Science, Vol. 135. Springer-Verlag, Berlin.

Strømme, S. A. (1987). On parametrized rational curves in Grassmann varieties. In F. Ghione, C. Peskine and E. Sernesi (Eds), *Space Curves* (Proc. Conf on Curves in Projective Space, Rocca di Papa, Italy, June 1995), pp. 251–272. Lecture Notes in Mathematics, Vol. 1266, Springer-Verlag, Berlin.

Verghese, G. C. and T. Kailath (1981). Rational matrix structure. *IEEE Trans. Autom. Control*, **AC-26**, 434–438.

Willems, J. C. (1986). From time series to linear system. Part I: Finite dimensional linear time invariant systems. *Automatica*, **22**, 561–580.

Willems, J. C. (1991). Paradigms and puzzles in the theory of dynamical systems. *IEEE Trans. Autom. Control*, **AC-36**, 259–294.