IMA Journal of Mathematical Control & Information (1985) **2**, 87–107

On the Inherent Integration Structure of Nonlinear Systems

H. NIJMEIJER

Department of Applied Mathematics, Twente University of Technology, P.O. Box 217, 7500 AE Enschede, The Netherlands

AND

J. M. SCHUMACHER

Centre for Mathematics and Computer Science, Kruislaan 413, 1098 SJ Amsterdam, The Netherlands

[Received 4 January 1985]

In this paper we characterize the inherent integration structure of affine nonlinear systems through a set of indices called—in analogy with existing terminology for linear systems—the orders of the zeros at infinity. We show that our definition encompasses earlier characterizations due to Hirschorn and Isidori. The discussion is entirely local in nature, so that we are able to use recent results in the 'geometric approach' to nonlinear system theory.

1. Introduction

PERHAPS a suitable way to describe the basic difference between the fields known as 'mathematical system theory' and 'dynamical systems' would be the following. Both fields are concerned with sets of differential (or difference) equations. However, in the theory of dynamical systems one is interested in the relations that these equations specify between parameters on one hand and functions on the other; whereas the system theorist is interested primarily in the relations between functions among each other.

In this view, the most elementary object of study in mathematical system theory is the relation between two real-valued time functions, one of which is the derivative of the other. It is customary to call the first function the 'output' and the second function the 'input', so that the relation between the two functions is taken in the direction in which it has a smoothing effect. The single-step integration can be turned into a multiple-step integration by putting several stages in series, and after applying feedback and adding interconnections with other such systems (perhaps nonlinear, and with the introduction of constraints) one may be left with a highly complex model. The question arises, to what extent the basic building block of a single-step integration is still recognizable in a complex system. Is there a way to relate a given system uniquely to an uncoupled set of series connections of simple integrators?

The problem of defining an 'inherent integration structure' for the class of linear systems has been studied by Hautus in [8]. For this class of systems one can profitably use the Laplace transform. The single-step integrator is transformed into the function s^{-1}, which is characterized by the fact that it has a first-order zero at infinity. More generally, it was shown in [8] that classical results from algebra allow the definition of a set of zeros at infinity, each with a specified order, for any rational matrix. It is then reasonable to identify this 'zero structure at infinity' with the sought-after inherent integration structure of the system. Because the analysis in [8] is based on the Laplace transform, it is not evident how to extend it to nonlinear systems. However, in a study of the invertibility problem for nonlinear systems [11], Hirschorn used a nonlinear version of Silverman's 'Structure Algorithm' [27] to define a set of indices which can be viewed as a representation of the inherent integration structure. Isidori [12, 13] has been able to define a 'formal zero structure at infinity' based on a certain formal power series representation for nonlinear systems. Isidori's definition, however, is tied to the class of 'input-output linearizable systems'. The class of systems considered by Hirschorn is somewhat larger, but he still needs a restrictive assumption to make sure that his indices are well defined. (Details of this are discussed in Section 4.)

In the present paper, we propose a definition of 'zeros at infinity' which is applicable to the full class of affine nonlinear systems. It would be too much to expect that the inherent integration structure of nonlinear systems could be expressed globally by a set of constants, and so we shall restrict ourselves to open subsets of the state manifold where certain regularity assumptions hold. We shall show that our definition leads to the same inherent integration structure as follows from Hirschorn's indices in the cases where these indices are well-defined. Since Isidori's definition of integration structure is the same as Hirschorn's (see Remark 5.4 in [12]) it will follow that our definition also encompasses the one given by Isidori.

The same definition has also been introduced in [23] where it was applied to give constructive necessary and sufficient conditions for the solvability of an input–output decoupling problem for nonlinear systems. It has also been used in [2] where the application was to a model-matching problem.

The structure of the present paper is as follows. We first review the linear situation in Section 2 where we also introduce an apparently new and convenient definition for the zeros at infinity of a linear system. Then in Section 3 the necessary material is given which leads to our definition of zeros at infinity for nonlinear systems. The connection with Hirschorn's indices is made in Section 4, with some technical arguments being placed in an Appendix. Conclusions follow in Section 5.

2. Zeros at infinity for linear systems

For single-input single-output systems, a quantity of obvious importance is the difference of the degrees of the denominator and the numerator of the transfer function, which is called the 'order of the zero at infinity' in the theory of rational functions. For instance, this quantity determines the number of poles that go off to infinity with increasing gain, according to root–locus analysis. In model-following problems it is clear that there will exist a proper function $g_1(s)$ such that $g_2(s)g_1(s) = g_3(s)$ only if the order of the zero at infinity of $g_3(s)$ equals or exceeds that of $g_2(s)$. It is also quite obvious that the inverse of a transfer function $g(s)$ will have a polynomial part of degree equal to the order of the zero at infinity of $g(s)$.

It is of interest to extend this concept to the multivariable case. Classically, the finite zeros of a rational matrix are defined through the Smith–McMillan canonical form (used in [15] to define the orders of the poles of a rational matrix, and in [25] to define the orders of the zeros). It is then possible to define zeros at infinity via an arbitrary Möbius transformation taking the point at infinity to some finite point in the complex plane (this was done in [15] for poles, and in [25] also for zeros). So, one says that a transfer matrix $G(s)$ has l zeros at infinity of respective orders $n_1, \ldots, n_l$ if the matrix $\tilde{G}(s) := G(1/s)$ has l zeros of orders $n_1, \ldots, n_l$ at 0, i.e. if the elements $s^{n_1}, \ldots, s^{n_l}$ appear in the numerators on the diagonal in the Smith–McMillan canonical form of $\tilde{G}(s)$.

Another direction was taken by Hautus [8] and Morse [19]. They showed that a canonical form can be given for rational matrices under left and right transformation by 'bicausal' matrices (i.e. proper matrices that have a proper inverse), although they did not use this term, which has become standard since [9]. This canonical form is determined by the dimensions of the given matrix and by a set of indices called the 'order indices' by Hautus, who was explicitly motivated by the idea of extending the scalar concept of 'order' (difference of denominator and numerator degrees) to the multivariable case. Morse ([19], p. 68) noted that these canonical indices also occur as one of the lists of invariants under a certain group of transformations of linear systems, which he had identified in earlier work [18]. Hautus [8] showed that the 'order indices' are related in a simple way to the indices obtained by Silverman in his study [27] of the inversion problem for linear multivariable systems, and he also recovered a result of Singh [31] which relates these indices, in turn, to the ranks of certain Toeplitz matrices defined by Sain and Massey in their treatment [26] of the inversion problem. The relation with the zeros at infinity as defined by Rosenbrock [25], however, was not made explicit in [8] and [19]. The definition of zeros at infinity through bicausal matrices is now accepted generally. It reads as follows.

DEFINITION A transfer matrix $G(s)$ is said to have l *zeros at infinity*, of *orders* $n_1, \ldots, n_l$ respectively, if there exist bicausal matrices $M(s)$ and $N(s)$ such that

$$M(s)G(s)N(s) = \begin{bmatrix} D(s) & 0 \\ 0 & 0 \end{bmatrix} \tag{2.1}$$

with

$$D(s) = \operatorname{diag}(s^{-n_1}, \ldots, s^{-n_l}). \tag{2.2}$$

Of course, one has to show that such bicausal matrices always exist, and that the indices $n_1, \ldots, n_l$ are uniquely defined. A simple direct proof of this can be found in [8]. The relation between this definition and the one given by Rosenbrock was established perhaps for the first time in the work of Verghese [36]; see also [37] for a survey of various ways to extract the pole–zero structure of a rational matrix at an arbitrary point in the extended complex plane. The pole–zero structure extraction problem had been studied using Toeplitz matrices by Vandewalle and Dewilde [33], who defined, in that paper, a zero of an invertible rational matrix simply as a pole of its inverse. The procedure can be generalized to singular matrices and applied to the point at infinity, in which case the Sain–Massey form is recovered; see [34].

Yet other equivalent definitions can be given of zeros at infinity for linear systems. There is close connection between the zeros at infinity of $G(s) = C(sI-A)^{-1}B+D$ and the infinite elementary divisors of the associated singular matrix pencil

$$P(s) = \begin{bmatrix} sI-A & -B \\ C & D \end{bmatrix} \tag{2.3}$$

as defined by Kronecker (see [7], Ch. XII). In fact, due to insufficient coordination of history, the orders of the infinite elementary divisors are equal to the orders of the infinite zeros plus one. The relation with the Kronecker theory was noted by Thorp [32]; cf. also [18]. It was already proposed by Rosenbrock ([25], p. 132) to say that a rational matrix $G(s)$ has a zero at infinity if every minor of a given order k tends to zero as $s \to \infty$. For ways of finding the zeros at infinity from the minors of the transfer matrix or of the associated system matrix, see [36, 37, 38, 3]. As to applications of the concept of zeros at infinity, the importance of Morse's 'list I_4' of invariants for multivariable root-locus theory was noted by Owens in [24]. The relevance of the concept for system inversion is quite obvious from the development sketched above. In particular, if a system is invertible, then the orders of its zeros at infinity are precisely the degrees of the differentiators that are needed to construct the inverse [36]. Infinite zeros are also important in the typically multivariable problem of input–output decoupling [35, 4, 5, 6]. For use of zeros at infinity in the model-following problem, see [16] and [17].

We shall now present still another way of defining the zeros at infinity, which, although straightforward, seems to be new. The main virtue of this definition for our purposes is that it leads naturally to a state-space characterization, which will be the basis for our generalization of the concept to nonlinear systems. Let $\mathbb{R}_0^m(s)$ denote the set of proper rational m-vector functions with real coefficients. For $f \in \mathbb{R}_0^m(s)$, we can define its value $\phi(f)$ at infinity:

$$\phi(f) = \lim_{s \to \infty} f(s) \quad (f \in \mathbb{R}_0^m(s)). \tag{2.4}$$

The set $\mathbb{R}_0^m(s)$ can be considered as a module over the ring of proper rational functions $\mathbb{R}_0(s)$. Note that every submodule of $\mathbb{R}_0^m(s)$ is mapped by ϕ onto a linear subspace of $\mathbb{R}^m$. In particular, it is easily verified that, for any $p \times m$ rational matrix $G(s)$ and any integer k, the following set is a submodule of $\mathbb{R}_0^m(s)$:

$$\mathcal{M}_k(G) := \{f \in \mathbb{R}_0^m(s) : s^k G(s) f(s) \in \mathbb{R}_0^p(s)\}. \tag{2.5}$$

So we can define a sequence of nonnegative integers p^k associated with G by writing

$$p^k = \dim \phi[\mathcal{M}_k(G)]. \tag{2.6}$$

For instance, p^1 is the number of independent functions in $\mathbb{R}_0^m(s)$ that are mapped by G to functions that have a zero of first or higher order at infinity, 'independent' being taken in the sense of independent first elements in the Taylor series development around infinity. We shall call p^k 'the number of zeros at infinity of $G(s)$ that have order $\geq k$'. This leads us to define, for a given rational matrix $G(s)$, the 'number of zeros of order k' which we shall denote by ζ_k, as follows:

$$\zeta_k = p^k - p^{k+1}. \tag{2.7}$$

It is not hard to see that this definition leads to the same result as the standard one, which was given above. Indeed, the numbers p^k are clearly invariant under left and right multiplication of G by bicausal matrices. So it is sufficient to verify the equivalence of the two definitions for matrices that have the form of the right-hand side of (2.1), and, for such matrices, the equivalence is seen by inspection.

An advantage of the definition that we just proposed is that it does not depend on the calculation of a special form. Also, it shows that the orders of the zeros at infinity are related to the dimensions of certain subspaces of $\mathbb{R}^m$ (the input space), and we shall now proceed to give the precise form of this relation in state-space terms. So, suppose that $G(s)$ is a strictly proper rational matrix, appearing as the transfer matrix of a given linear system $\Sigma(A, B, C)$:

$$G(s) = C(sI - A)^{-1}B \tag{2.8}$$

(minimality of the state-space representation is not assumed). The expansion of $G(s)$ around infinity is then given by

$$G(s) = CBs^{-1} + CABs^{-2} + \cdots. \tag{2.9}$$

Any $f \in \mathbb{R}_0^m(s)$ can also be expanded around infinity:

$$f(s) = f_0 + f_1 s^{-1} + f_2 s^{-2} + \cdots. \tag{2.10}$$

One will have $s^k G(s) f(s) \in \mathbb{R}_0^p(s)$ for $k \geq 1$, if and only if

$$\begin{bmatrix} 0 & & & 0 \\ CB & 0 & & \\ \vdots & \ddots & \ddots & \\ CA^{k-2}B & \dots & CB & 0 \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_{k-1} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}. \tag{2.11}$$

We can write this more compactly by introducing

$$D_k = \begin{bmatrix} 0 & & 0 \\ CB & \ddots & \\ \vdots & \ddots & \\ CA^{k-2}B & \dots\ CB & 0 \end{bmatrix}, \qquad C_k = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{k-1} \end{bmatrix}, \tag{2.12}$$

$$\tilde{f}_k = \begin{bmatrix} f_1 \\ \vdots \\ f_k \end{bmatrix}. \tag{2.13}$$

Using this notation, we get

$$\begin{aligned} p^k &= \dim\left\{f_0 \in \mathbb{R}^m : \exists \tilde{f}_{k-1} \text{ s.t. } \begin{bmatrix} 0 & 0 \\ C_{k-1}B & D_{k-1} \end{bmatrix}\begin{bmatrix} f_0 \\ \tilde{f}_{k-1} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}\right\} \\ &= \dim\{f_0 \in \mathbb{R}^m : C_{k-1}Bf_0 \in \operatorname{im} D_{k-1}\} \\ &= \dim[C_{k-1}^{-1}(\operatorname{im} D_{k-1}) \cap \operatorname{im} B] + \dim\ker B. \end{aligned} \tag{2.14}$$

The quantity $\dim\ker B$ is of little importance; it is often assumed that inputs are independent, and in this case one would have $\dim\ker B = 0$. In any case, it is an amount which does not depend on k so that its effect on ζ_k will be nil. The linear space $C_{k-1}^{-1}(\operatorname{im} D_{k-1})$ is a subspace of the state space $\mathscr{X}$. We did not actually define $C_0^{-1}(\operatorname{im} D_0)$ in (2.12), but it is seen from (2.14) that to be consistent one should take $C_0^{-1}(\operatorname{im} D_0) = \mathscr{X}$. From (2.12), one sees that the matrices C_k and D_k could be defined recursively via

$$C_{k+1} = \begin{bmatrix} C \\ C_kA \end{bmatrix}, \qquad D_{k+1} = \begin{bmatrix} 0 & 0 \\ C_kB & D_k \end{bmatrix}, \tag{2.15}$$

and this strongly suggests that it should also be possible to give a recursive definition for the sequence of subspaces $C_k^{-1}(\operatorname{im} D_k)$. In fact, a result of Silverman ([28], p. 356) shows how to do this. We state the result and include a short proof for completeness.

LEMMA 2.1 *Let $G(s)$ be a strictly proper rational matrix represented by* (2.8). *Define a sequence of subspaces $\mathscr{V}^k$ of the state space $\mathscr{X}$ in the following way:*

$$\mathscr{V}^0 = \mathscr{X} \tag{2.16}$$

$$\mathscr{V}^{k+1} = A^{-1}(\mathscr{V}^k + \operatorname{im} B) \cap \ker C. \tag{2.17}$$

Under these conditions, one has

$$C_k^{-1}(\operatorname{im} D_k) = \mathscr{V}^k \tag{2.18}$$

for all $k \geq 0$, where the left-hand side is defined by (2.12) *for $k \geq 1$, and equality holds for $k = 0$ by definition.*

Proof. The proof is by induction, and the first step has already been taken.

Suppose now that (2.18) holds for a given k. Then

$$\begin{aligned}\mathcal{V}^{k+1} &= \{x \in \mathcal{X} : Cx = \mathbf{0} \text{ and } \exists u \text{ s.t. } Ax + Bu \in \mathcal{V}^k\} \\ &= \left\{x \in \mathcal{X} : \exists u, \tilde{u}_k \text{ s.t. } \begin{bmatrix} C \\ C_k A \end{bmatrix} x + \begin{bmatrix} 0 & 0 \\ C_k B & D_k \end{bmatrix} \begin{bmatrix} u \\ \tilde{u}_k \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \right\} \\ &= C_{k+1}^{-1}(\operatorname{im} D_{k+1}). \end{aligned} \tag{2.19}$$

Of course, the algorithm (2.16–17) is the well known '$\mathcal{V}^*$-algorithm' ([39], p. 91). It produces a nonincreasing sequence of subspaces which must converge in a finite number of steps to a limit subspace denoted by $\mathcal{V}^*$. Combining (2.7), (2.14), and (2.18) now leads immediately to a formula for the number of zeros at infinity of a given order $k \geq 1$ in terms of the subspaces $\mathcal{V}^k$:

$$\zeta_k = \dim(\mathcal{V}^{k-1} \cap \operatorname{im} B) - \dim(\mathcal{V}^k \cap \operatorname{im} B). \tag{2.20}$$

This formula will be extended to the nonlinear situation in the next section. The connection (2.20) between the sequence $(\dim(\mathcal{V}^k \cap \operatorname{im} B))$ and the orders of the zeros at infinity has been given earlier, in a less direct way, by Malabre [16]. It is straightforward to extend the characterization we have given for a system $\Sigma(A, B, C, D)$ with direct feedthrough. One then uses the generalized form of the $\mathcal{V}^*$-algorithm as presented by Anderson in [1], and the role of im B is taken over by $B(\ker D)$, but with these modifications the end result (2.20) is still the same.

3. Zeros at infinity for nonlinear systems

We consider an affine nonlinear system given locally by

$$\dot{x}(t) = \mathbf{A}[x(t)] + \sum_{i=1}^{m} \mathbf{B}_i[x(t)]u_i(t), \tag{3.1}$$

$$y(t) = \mathbf{C}[x(t)], \tag{3.2}$$

where the vector x denotes local coordinates of a smooth n-dimensional manifold $\mathcal{M}$. The smooth vector fields $\mathbf{A}, \mathbf{B}_1, \ldots, \mathbf{B}_m$ are defined on $\mathcal{M}$, $u_i : \mathbb{R}_+ \to \mathbb{R}$ are piecewise smooth input functions ($i = 1, \ldots, m$), and $\mathbf{C}$ is a surjective submersion of $\mathcal{M}$ onto a smooth p-dimensional manifold $\mathcal{N}$. A sequence of distributions on $\mathcal{M}$ is defined in the following way (cf. [14, 20]).

$$\mathcal{V}^0 = \mathrm{T}\mathcal{M}, \tag{3.3}$$

$$\mathcal{V}^{k+1} = \ker \mathbf{C}_* \cap \Delta^{-1}(\Delta_0 + \mathcal{V}^k). \tag{3.4}$$

Here we denote

$$\Delta_0 = \operatorname{span}(\mathbf{B}_1, \ldots, \mathbf{B}_m), \qquad \Delta = \operatorname{span}(\mathbf{A}, \mathbf{B}, \ldots, \mathbf{B}_m), \tag{3.5}$$

$$\Delta^{-1}\mathcal{D} = \{\mathbf{X} \in \mathrm{V}(\mathcal{M}) : [\Delta, \mathbf{X}] \subset \mathcal{D}\} \tag{3.6}$$

for any distribution $\mathcal{D}$, with

$$[\Delta, \mathbf{X}] = \{[\mathbf{Y}, \mathbf{X}] : \mathbf{Y} \in \Delta\}. \tag{3.7}$$

The following are basic properties of the distributions $\mathcal{V}^k$.

PROPOSITION 3.1 *The distributions $\mathcal{V}^k$ $(k \geq 0)$ defined by (3.3–4) satisfy the following conditions.*

(a) $\mathcal{V}^{k+1} \subset \mathcal{V}^k$ *for all* $k \geq 0$;

(b) $\mathcal{V}^i$ *is involutive for all* $k \geq 0$.

Proof. (a) Clearly, $\mathcal{V}^1 = \ker \boldsymbol{C}_* \subset \mathrm{T}\mathcal{M} = \mathcal{V}^0$. Now suppose that $\mathcal{V}^k \subset \mathcal{V}^{k-1}$; then

$$\begin{aligned}\mathcal{V}^{k+1} &= \ker \boldsymbol{C}_* \cap \Delta^{-1}(\Delta_0 + \mathcal{V}^k)\\ &\subset \ker \boldsymbol{C}_* \cap \Delta^{-1}(\Delta_0 + \mathcal{V}^{k-1}) = \mathcal{V}^k.\end{aligned} \tag{3.8}$$

(b) This part of the proof also uses induction. First, $\mathcal{V}^0 = \mathrm{T}\mathcal{M}$ is clearly involutive. Suppose that $\mathcal{V}^{k-1}$ is involutive. By part (a) we may write, instead of (3.4),

$$\mathcal{V}^k = \mathcal{V}^{k-1} \cap \Delta^{-1}(\Delta_0 + \mathcal{V}^{k-1}). \tag{3.9}$$

Now, choose arbitrary vector fields $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ from $\mathcal{V}^k$. Then it is clear that $[\boldsymbol{X}_1, \boldsymbol{X}_2] \in \mathcal{V}^{k-1}$ because both $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ belong to $\mathcal{V}^{k-1}$, and $\mathcal{V}^{k-1}$ is involutive by assumption. It remains to show that $[\Delta, [\boldsymbol{X}_1, \boldsymbol{X}_2]] \subset \Delta_0 + \mathcal{V}^{k-1}$, or equivalently $[\boldsymbol{A}, [\boldsymbol{X}_1, \boldsymbol{X}_2]] \in \Delta_0 + \mathcal{V}^{k+1}$ and $[\boldsymbol{B}_i, [\boldsymbol{X}_1, \boldsymbol{X}_2]] \in \Delta_0 + \mathcal{V}^{k-1}$ for $i = 1, \ldots, m$. To this end, we use the Jacobi identity:

$$\begin{aligned}[\boldsymbol{A}, [\boldsymbol{X}_1, \boldsymbol{X}_2]] &= -[\boldsymbol{X}_1, [\boldsymbol{X}_2, \boldsymbol{A}]] - [\boldsymbol{X}_2, [\boldsymbol{A}, \boldsymbol{X}_1]]\\ &\in [\boldsymbol{X}_1, \Delta_0 + \mathcal{V}^{k-1}] + [\boldsymbol{X}_2, \Delta_0 + \mathcal{V}^{k-1}]\\ &\subset [\boldsymbol{X}_1, \Delta] + [\boldsymbol{X}_1, \mathcal{V}^{k-1}] + [\boldsymbol{X}_2, \mathcal{V}^{k-1}]\\ &\subset \Delta_0 + \mathcal{V}^{k-1}.\end{aligned} \tag{3.10}$$

In the final step, we again used the assumed involutivity of $\mathcal{V}^{k-1}$, plus the result of part (a). Finally, we have likewise

$$\begin{aligned}[\boldsymbol{B}_i, [\boldsymbol{X}_1, \boldsymbol{X}_2]] &= -[\boldsymbol{X}_1, [\boldsymbol{X}_2, \boldsymbol{B}_i]] - [\boldsymbol{X}_2, [\boldsymbol{B}_i, \boldsymbol{X}_1]]\\ &\subset \Delta_0 + \mathcal{V}^{k+1} \quad (i = 1, \ldots, m). \qquad \square\end{aligned} \tag{3.11}$$

Remark. Part (b) of the above proposition is an improvement of Theorem 4.1 of [20].

When specialized to linear systems, the algorithm (3.3–4) coincides with the one given in (2.16–17). This would suggest a definition of zeros at infinity for the nonlinear system (3.1–2) in analogy with (2.20), but of course there is in general no reason to expect that the distributions $\mathcal{V}^k \cap \Delta_0$, although they are defined throughout $\mathcal{M}$, will have constant dimension on $\mathcal{M}$. As a first step, it seems best to satisfy ourselves with assumptions which are at least valid on open parts of $\mathcal{M}$, and so we shall assume that

$$\dim(\mathcal{V}^k \cap \Delta_0) = \text{constant} \tag{3.12}$$

for all $k \geq 0$. (For analytic systems, this assumption will even hold on an open and dense submanifold of $\mathcal{M}$.) Then we define, as in [23], the zero structure at infinity in complete analogy with (2.20): the system (3.1–2) is said to have ζ_k zeros at infinity of order k (≥ 1) where

$$\zeta_k = \dim(\mathcal{V}^{k-1} \cap \Delta_0) - \dim(\mathcal{V}^k \cap \Delta_0). \tag{3.13}$$

Of course, the similarity to the linear definition is in itself no guarantee that the zeros at infinity as we just defined them really express the 'inherent integration structure' of a nonlinear system. The relevance of the concept can only be established through showing its usefulness in applications and its relation to other concepts. As to applications, we refer to [23] where the above definition is used for a study of the input–output decoupling problem, and to [2] where the concept is applied to a model-matching problem. In this paper, we shall concentrate on the relation between our definition of 'zeros at infinity' and Hirschorn's definition [11] of 'invertibility indices'. The two concepts turn out to be very closely related, which is no surprise since Hirschorn's result is based on a nonlinear version of Silverman's 'Structure Algorithm', which is well known to be closely related to the zero structure at infinity [29]. It should be emphasized that Hirschorn's definition applies to a smaller class of systems than does ours; the difference will be made clear in the next section. A still smaller class of nonlinear systems is considered by Isidori in [12]. Very interestingly, it turns out that for this class of systems the zeros at infinity can be obtained from a certain formal power series, taking the place of the transfer function in the linear analysis (cf. [8]). The approach in [12] again draws heavily on the nonlinear version of Silverman's algorithm, and, as noted in the paper itself, the structure at infinity defined there does not deviate from Hirschorn's definition.

Finally let us remark that for a general nonlinear system given locally by $\dot{x}=f(x, u)$, $y=g(x, u)$, one can still use the above definition by passing to the so-called 'extended system' [22]. This is explained in [23].

4. Relation to the Hirschorn algorithm

The 'structure algorithm' of Silverman [27] has been generalized by Hirschorn [11] (cf. also [30]) to the class of affine nonlinear systems. Essentially, the algorithm consists of a series of transformations; in each step, the outputs are rearranged in such a way that as many of them as possible do not depend explicitly on any of the inputs, and these outputs are then differentiated, after which the procedure is repeated. In the nonlinear case, the transformations that have to be performed on the outputs will in general be state-dependent, and this leads to considerable complications. We shall now describe the algorithm more precisely. Our version of the algorithm leaves more freedom in the output transformations than the original version of [11]; the difference will be explained as we proceed. Let an affine control system be given as in Section 3:

$$\dot{x}(t)=A[x(t)]+\sum_{i=1}^{m} B_i[x(t)]u_i(t), \qquad x(0)=x_0\in\mathcal{M}, \tag{4.1}$$
$$y(t)=C[x(t)]$$

We call this 'system (0)', and we write

$$\hat{z}_0(t)=y(t), \qquad \hat{C}^0=C. \tag{4.2}$$

In order to proceed recursively, let k be a nonnegative integer. We let $\hat{R}_k(x)$ and

$\bar{R}_k(\boldsymbol{x})$ be matrices, depending smoothly on $\boldsymbol{x}$, such that

$$\ker R_k(\boldsymbol{x}) = \operatorname{span}\left\{\begin{bmatrix} \bar{\boldsymbol{D}}_i^k(\boldsymbol{x}) \\ \mathrm{L}_{\boldsymbol{B}_i}\hat{\boldsymbol{C}}^k(\boldsymbol{x}) \end{bmatrix} : i = 1, \ldots, m\right\}, \tag{4.3}$$

where $\bar{\boldsymbol{D}}_i^k(\boldsymbol{x})$ and $\hat{\boldsymbol{C}}^k(\boldsymbol{x})$ are defined recursively by (4.5)–(4.7) below, and the matrix $R_k(\boldsymbol{x})$ is defined by

$$R_k(\boldsymbol{x}) = \begin{bmatrix} \bar{R}_k(\boldsymbol{x}) \\ \hat{R}_k(\boldsymbol{x}) \end{bmatrix} \tag{4.4}$$

is invertible for all x. (It is assumed that such matrices exist, i.e. that the distribution appearing on the right-hand side in (4.3) has constant dimension. See the remarks in Section 3; cf. also [11].) Now, we define

$$\bar{\boldsymbol{C}}^{k+1}(\boldsymbol{x}) = \bar{R}_k(\boldsymbol{x}) \begin{bmatrix} \bar{\boldsymbol{C}}^k(\boldsymbol{x}) \\ \mathrm{L}_{\mathbf{A}}\hat{\boldsymbol{C}}^k(\boldsymbol{x}) \end{bmatrix}, \tag{4.5}$$

$$\bar{\boldsymbol{D}}_i^{k+1}(\boldsymbol{x}) = \bar{R}_k(\boldsymbol{x}) \begin{bmatrix} \bar{\boldsymbol{D}}_i^k(\boldsymbol{x}) \\ \hline \mathrm{L}_{\boldsymbol{B}_i}\hat{\boldsymbol{C}}^k(\boldsymbol{x}) \end{bmatrix} \quad (i = 1, \ldots, m), \tag{4.6}$$

$$\hat{\boldsymbol{C}}^{k+1}(\boldsymbol{x}) = R_k(\boldsymbol{x}) \begin{bmatrix} \bar{\boldsymbol{C}}^k(\boldsymbol{x}) \\ \mathrm{L}_{\mathbf{A}}\hat{\boldsymbol{C}}^k(\boldsymbol{x}) \end{bmatrix}. \tag{4.7}$$

Finally, we define 'system $k+1$' by

$$\dot{\boldsymbol{x}}(t) = \mathbf{A}[\boldsymbol{x}(t)] + \sum_{i=1}^{m} \boldsymbol{B}_i[\boldsymbol{x}(t)]u_i(t), \tag{4.8}$$

$$\begin{bmatrix} \bar{\boldsymbol{z}}_{k+1}(t) \\ \hat{\boldsymbol{z}}_{k+1}(t) \end{bmatrix} = \begin{bmatrix} \bar{\boldsymbol{C}}^{k+1}[\boldsymbol{x}(t)] \\ \hat{\boldsymbol{C}}^{k+1}[\boldsymbol{x}(t)] \end{bmatrix} + \sum_{i=1}^{m} \begin{bmatrix} \bar{\boldsymbol{D}}_i^{k+1}[\boldsymbol{x}(t)] \\ \mathbf{0} \end{bmatrix} u_i(t).$$

4.1 *Remarks*

(i) It is understood that some of the vectors and matrices defined here may be empty, i.e. do not appear in the above expressions. In particular, this is always true for $\bar{\boldsymbol{C}}^0$ and $\bar{\boldsymbol{D}}_i^0$ $(i = 1, \ldots, m)$. Note that the algorithm becomes ineffective when $\hat{R}_k$ has become empty.

(ii) One way of obtaining the matrices $\hat{R}_k(\boldsymbol{x})$ and $\bar{R}_k(\boldsymbol{x})$ is the following. Form the matrix $M^k(\boldsymbol{x})$:

$$M^k(\boldsymbol{x}) = \begin{bmatrix} \boldsymbol{D}_1^k(\boldsymbol{x}) & \ldots & \bar{\boldsymbol{D}}_m^k(\boldsymbol{x}) \\ \mathrm{L}_{\boldsymbol{B}_1}\hat{\boldsymbol{C}}^k(\boldsymbol{x}) & \ldots & \mathrm{L}_{\boldsymbol{B}_m}\hat{\boldsymbol{C}}^k(\boldsymbol{x}) \end{bmatrix}. \tag{4.9}$$

Under the assumptions we have made, the rank of $M^k(\boldsymbol{x})$ is a constant, say r_{k+1}. We also assume that it is possible to select r_{k+1} rows from $M^k(\boldsymbol{x})$ that are linearly independent for all $\boldsymbol{x}$. Let E_k be an elementary matrix which brings these rows to the first r_{k+1} positions:

$$E_k M^k(\boldsymbol{x}) = \begin{bmatrix} M_1^k(\boldsymbol{x}) \\ M_2^k(\boldsymbol{x}) \end{bmatrix} \tag{4.10}$$

where $M_1^k(\boldsymbol{x})$ is $r_{k+1}\times m$ and has full row rank. The rows of $M_2^k(\boldsymbol{x})$ depend linearly on those of $M_1^k(\boldsymbol{x})$, so there exists a matrix $F_k(\boldsymbol{x})$ (depending smoothly on $\boldsymbol{x}$) such that

$$F_k(\boldsymbol{x})M_1^k(\boldsymbol{x})+M_2^k(\boldsymbol{x})=0. \tag{4.11}$$

We can now take

$$\bar{R}_k(\boldsymbol{x})=[I \quad 0]E_k, \qquad \hat{R}_k(\boldsymbol{x})=[F_k(\boldsymbol{x}) \quad I]E_k. \tag{4.12}$$

(iii) In [11], Hirschorn allows only matrices $R_k(\boldsymbol{x})$ that are constructed as in the previous remark. For example, if an affine system is given (on a suitable open subset of $\mathbb{R}^3$) by the equations

$$\begin{gathered}\dot{x}_1(t)=x_3(t)u_1(t)+u_2(t), \qquad \dot{x}_2(t)=x_2(t)+u_2(t), \qquad \dot{x}_3(t)=-x_1(t)u_1(t),\\ y_1(t)=x_1(t)^2+x_3(t)^2, \qquad y_2(t)=x_2(t)^2,\end{gathered} \tag{4.13}$$

then the matrix $M^0(\boldsymbol{x})$ is found to be

$$M^0(\boldsymbol{x})=\begin{bmatrix}0 & 2x_1\\ 0 & 2x_2\end{bmatrix}. \tag{4.14}$$

The procedure of Remark (ii) leads to either one of the following two possibilities:

$$R_0(\boldsymbol{x})=\begin{bmatrix}1 & 0\\ -x_2/x_1 & 1\end{bmatrix} \quad \text{or} \quad R_0(\boldsymbol{x})=\begin{bmatrix}0 & 1\\ 1 & -x_1/x_2\end{bmatrix}. \tag{4.15}$$

On the other hand, we also allow, for instance,

$$R_0(\boldsymbol{x})=\begin{bmatrix}1 & 0\\ -x_2 & x_1\end{bmatrix}. \tag{4.16}$$

4.2 *Discussion of the Algorithm* (4.1–8)

For the rest of Section 4 we consider the situation in which a sequence of systems has been produced by the algorithm. The sequence is determined by the original system (4.1) and by the consecutive row-transformation matrices $R_0, R_1, \ldots, R_{\alpha-1}$. We introduce a derived sequence of distributions $\mathscr{E}^k$ ($k=1,\ldots,\alpha$):

$$\mathscr{E}^k=\{\mathbf{Z}\in \mathrm{V}(\mathscr{M}): \mathrm{L}_{\mathbf{Z}}\mathrm{L}_{\mathbf{A}}^{j}R_l(\boldsymbol{x})=0,\ l\in\mathbb{N}_0,\ j\in\mathbb{N}_0,\ 0\leq l+j\leq k-2\} \tag{4.17}$$

(where we understand $\mathscr{E}^1=\mathrm{T}\mathscr{M}$). Note that the distributions $\mathscr{E}^k$ are involutive and that they form a nonincreasing sequence. Another simple property is the following Lemma.

LEMMA 4.1 *If* $\mathbf{Z}\in\mathscr{E}^{k+1}$ *then* $[\mathbf{Z},\mathbf{A}]\in\mathscr{E}^k$.

Proof. For $0\leq l+j\leq k-2$ we compute:

$$\mathrm{L}_{[\mathbf{Z},\mathbf{A}]}\mathrm{L}_{\mathbf{A}}^{j}R_l(\boldsymbol{x})=\mathrm{L}_{\mathbf{Z}}\mathrm{L}_{\mathbf{A}}^{j+1}R_l(\boldsymbol{x})-\mathrm{L}_{\mathbf{A}}\mathrm{L}_{\mathbf{Z}}\mathrm{L}_{\mathbf{A}}^{j}R_l(\boldsymbol{x})=0. \quad \square \tag{4.18}$$

The following result will turn out to be of crucial importance. Recall the algorithm (3.3–4).

THEOREM 4.2 *Suppose that $\Delta_0 \subset \mathscr{E}^{\alpha-1}$. Then, for all $k \in \{1, \ldots, \alpha\}$,*

$$\mathscr{E}^k \cap \mathscr{V}^k = \mathscr{E}^k \cap \bigcap_{i=0}^{k-1} \ker \hat{\boldsymbol{C}}_*^i. \tag{4.19}$$

The proof of this theorem is somewhat technical, and we relegate it to the Appendix. For linear systems and for single-output nonlinear systems, the transformation matrices R_k can be chosen constant, and then we have $\mathscr{E}^k = \mathrm{T}\mathscr{M}$ for all k. So, in these cases, one obtains

$$\mathscr{V}^k = \bigcap_{i=0}^{k-1} \ker \hat{\boldsymbol{C}}_*^i \tag{4.20}$$

which gives us a convenient way of computing $\mathscr{V}^k$. Unfortunately, simple examples show that (4.20) does not necessarily hold for multi-output nonlinear systems. For linear systems, the relation (4.20) appears (in dual form) in [10] (p. 387). An easy consequence of Theorem 4.2 is the following.

COROLLARY 4.3 *Suppose that $\Delta_0 \subset \mathscr{E}^\alpha$. Then, for all $k \in \{1, \ldots, \alpha\}$,*

$$\Delta_0 \cap \mathscr{V}^k = \Delta_0 \cap \bigcap_{i=0}^{k-1} \ker \hat{\boldsymbol{C}}_*^i. \tag{4.21}$$

Proof. If $\Delta_0 \subset \mathscr{E}^\alpha$, then $\Delta_0 \subset \mathscr{E}^k$ for all $k \in \{1, \ldots, \alpha\}$ since the sequence $(\mathscr{E}^k)$ is nonincreasing. So we obtain (4.21) by intersecting both sides of (4.19) with Δ_0. □

This allows us to establish the correspondence between the zeros at infinity, as defined in the previous section, and the 'invertibility indices' defined in [11] by

$$r_k = \dim \operatorname{span} (\bar{\boldsymbol{D}}_1^k, \ldots, \bar{\boldsymbol{D}}_m^k) \quad (k = 1, \ldots, \alpha). \tag{4.22}$$

We shall use another characterization of these indices. We still work with respect to a fixed sequence of row-transformation matrices.

LEMMA 4.4 For $k \in \{1, \ldots, \alpha\}$, we have

$$r_k = m - \dim \left(\Delta_0 \cap \bigcap_{i=0}^{k-1} \ker \hat{\boldsymbol{C}}_*^i \right). \tag{4.23}$$

Proof. We shall show that the relation

$$\sum_{i=1}^m \bar{\boldsymbol{D}}_i^k(\boldsymbol{x}) \gamma_i(\boldsymbol{x}) = \boldsymbol{0} \tag{4.24}$$

holds, for smooth functions $\gamma_i(\boldsymbol{x})$, if and only if

$$\sum_{i=1}^m \boldsymbol{B}_i(\boldsymbol{x}) \gamma_i(\boldsymbol{x}) \in \bigcap_{i=0}^{k-1} \ker \hat{\boldsymbol{C}}_*^i(\boldsymbol{x}). \tag{4.25}$$

This is equivalent to the statement of the lemma. The proof is by induction on k.

For $k=1$, we have $\bar{\boldsymbol{D}}_i^1 = \mathrm{L}_{\boldsymbol{B}_i}\boldsymbol{C}$ $(i=1,\ldots,m)$, and the equivalence of (4.24) and (4.25) is trivial. Now suppose that the equivalence holds for some given $l \in \{1,\ldots,\alpha-1\}$. First, let $\gamma_i(\boldsymbol{x})$ $(i=1,\ldots,m)$ be smooth functions such that (4.24) is satisfied with $k=l+1$. By (4.6), this means that

$$\bar{R}_l(\boldsymbol{x}) \sum_{i=1}^{m} \begin{bmatrix} \bar{\boldsymbol{D}}_i^l(\boldsymbol{x}) \\ \mathrm{L}_{\boldsymbol{B}_i}\hat{\boldsymbol{C}}^l(\boldsymbol{x}) \end{bmatrix} \gamma_i(\boldsymbol{x}) = \mathbf{0}. \tag{4.26}$$

But we also have (see (4.3))

$$\hat{R}_l(\boldsymbol{x}) \sum_{i=1}^{m} \begin{bmatrix} \bar{D}_i^l(\boldsymbol{x}) \\ \mathrm{L}_{\boldsymbol{B}_i}\hat{\boldsymbol{C}}^l(\boldsymbol{x}) \end{bmatrix} \gamma_i(\boldsymbol{x}) = \mathbf{0}. \tag{4.27}$$

Using the invertibility of $R_k(\boldsymbol{x})$, we obtain two results:

$$\sum_{i=1}^{m} \bar{\boldsymbol{D}}_i^l(\boldsymbol{x})\gamma_i(\boldsymbol{x}) = \mathbf{0}, \tag{4.28}$$

$$\sum_{i=1}^{m} \boldsymbol{B}_i(\boldsymbol{x})\gamma_i(\boldsymbol{x}) \in \ker \hat{\boldsymbol{C}}_*^l(\boldsymbol{x}). \tag{4.29}$$

An appeal to the induction assumption now completes the first half of the proof. Next, suppose that $\gamma_i(\boldsymbol{x})$ $(i=1,\ldots,m)$ are smooth functions such that (4.25) is satisfied with $k=l+1$. Then it follows from the induction assumption that

$$\sum_{i=1}^{m} \begin{bmatrix} \bar{\boldsymbol{D}}_i^k(\boldsymbol{x}) \\ \mathrm{L}_{\boldsymbol{B}_i}\hat{\boldsymbol{C}}^k(\boldsymbol{x}) \end{bmatrix} \gamma_i(\boldsymbol{x}) = \mathbf{0}. \tag{4.30}$$

The desired result is now obtained by letting $\bar{R}_k$ act on this, and by using (4.6) □

As they are defined in (4.22), the invertibility indices r_k depend on the selection of row-transformation matrices R_k. It is not difficult to find examples where different choices of the matrices R_k lead to different values of the indices r_k. However, it is shown in [11] that the r_k are determined uniquely if the condition

$$\Delta_0 \subset \mathscr{C}^\alpha \tag{4.31}$$

holds. Indeed, we see immediately from Corollary 4.3 and Lemma 4.4 that this conclusion remains true even under the relaxed restrictions that we have put on the matrices R_k, and that we have the following explicit expression for the invertibility indices r_k in terms of the original system parameters.

COROLLARY 4.5 *Suppose that* $\Delta_0 \subset \mathscr{C}^\alpha$. *Then the indices* r_k *defined in* (4.22) *satisfy*

$$r_k = m - \dim(\Delta_0 \subset \mathscr{V}^k). \tag{4.32}$$

It is shown in [11] that the system (4.1) is left invertible if there exists an $\alpha \in \mathbb{N}$ such that $\Delta_0 \subset \mathscr{C}^\alpha$ and $r_\alpha = \dim \Delta_0$. In [21], it was shown that (4.1) is left invertible if and only if the largest local controllability distribution contained in ker $\boldsymbol{C}_*$ is the zero distribution. Corollary 4.5 enables us to clarify, to a certain extent, the relation between these two invertibility results. The problem with the criterion of [21] is that there is no procedure available to verify whether the largest local controllability distribution in ker C_* is zero. We do have a method, however, to find out if the largest *regular* local controllability distribution in ker C_* is zero; in fact, this is equivalent to the condition $\Delta_0 \cap \mathscr{V}^* = 0$ (see [21]). Let α be such that $\mathscr{V}^\alpha = \mathscr{V}^*$; then it follows from Hirschorn's result and from Corollary 4.5 that this condition is not only necessary but also sufficient if matrices R_k can be found (in Hirschorn's way, as described in Remark (ii) of Section 4.1) such that $\Delta_0 \subset \mathscr{E}^\alpha$. Apparently, in this situation it suffices to consider only regular (rather than arbitrary) local controllability distributions. This observation leaves room for several conjectures, but we shall not pursue these now. Finally, we formulate the explicit connection between the zeros at infinity, as defined in Section 3, and the invertibility indices, as defined in [11]. The proof is immediate from Corollary 4.5.

THEOREM 4.6 *Suppose that the system* (4.1) *has been transformed α times through the algorithm* (4.2–8), *and that the corresponding row-transformation matrices R_k are such that $\Delta_0 \subset \mathscr{E}^\alpha$ (cf.* (4.17)). *Let r_k be defined by* (4.22) *for $k \geq 1$, and set $r_0 = 0$. Then, for each $k \geq 1$, the number of zeros at infinity of order k is given by*

$$\zeta_k = r_k - r_{k-1}. \tag{4.33}$$

Conversely, one also has

$$r_k = \sum_{i=1}^{k} \zeta_i. \tag{4.34}$$

5. Conclusions

In the effort to untangle the inherent integration structure of nonlinear systems, we have made some progress by proposing a definition of zeros at infinity for a large class of nonlinear systems, and by showing that this definition is compatible with the ones given earlier by Hirschorn and Isidori. The main advantage of our definition is its large scope, and the fact that it immediately shows that the orders of the zeros at infinity are feedback invariants. The advantage of Isidori's definition is that it enables one to make a connection with alternative system descriptions, notably via a formal power series which corresponds to the expansion around infinity of the transfer matrix in the linear case. However, this only goes for a restricted class of systems. What one would like to see is an interpretation of the inherent integration structure in more direct dynamical terms, at the level of generality that was used in this paper. This remains work for the future. An intriguing question in another direction is whether it is possible to say anything useful about zeros at infinity, not in local but in global terms.

REFERENCES

1. ANDERSON, B. D. O. 1975 Output-nulling invariant and controllability subspaces, *Preprints, Sixth Triennial World Congress IFAC, Boston/Cambridge, Mass.*, Part 1B, paper no. 43.6.
2. DI BENEDETTO, M. D., & ISIDORI, A. 1984 The matching of nonlinear models via dynamic state feedback, to appear in *Proceedings of the International Summer School on Geometric Methods in Nonlinear System Theory, Bierutowice (Poland).*
3. BOSGRA, O. H., & VAN DER WEIDEN, A. J. J. 1981 Realizations in generalized

state-space form for polynomial system matrices and the definitions of poles, zeros and decoupling zeros at infinity. *Int. J. Control* **33**, 391–411.

4. Descusse, J., & Dion, J. M. 1982 On the structure at infinity of linear square decoupled systems. *IEEE Trans. Autom. Control* AC-**27**, 971–974.
5. Descusse, J., Lafay, J. F., & Malabre, M. 1983 On the structure at infinity of linear block-decoupable systems: the general case. *IEEE Trans. Autom. Control*, AC-**28**, 1115–1118.
6. Dion, J. M. 1983 Feedback block decoupling and infinite structure of linear systems. *Int. J. Control* **37**, 521–533.
7. Gantmacher, F. R. 1959 *Matrix Theory*. New York: Chelsea.
8. Hautus, M. L. J. 1976 The formal Laplace transform for smooth linear systems, In: Marchesini, G., & Mitter, S. K. (eds.) *Mathematical Systems Theory* (Proceedings of the International Symposium, Udine, Italy, 1975), Lecture Notes in Economics and Mathematical Systems, vol 131. New York: Springer pp. 29–47.
9. Hautus, M. L. J., & Heymann, M. 1978 Linear feedback—an algebraic approach. *SIAM J. Control Optimiz.* **16**, 83–105.
10. Hautus, M. L. J., & Silverman, L. M. 1983 System structure and singular control. *Linear Algebra Applics.* **50**, 369–402.
11. Hirschorn, R. M. 1979 Invertibility of multivariable Nonlinear Control Systems. *IEEE Trans. Autom. Control* AC-**24**, 855–865.
12. Isidori, A. 1983 The matching of a prescribed linear input–output behavior in a nonlinear system. Dip. Sist. Inf., Universita di Roma.
13. Isidori, A. 1984 Nonlinear feedback, structure at infinity and the input–output linearization problem, In: Fuhrmann, P. A. (ed.). *Mathematical Theory of Networks and Systems*, Proceedings of the MTNS-83 International Symposium, Beer Sheva (Israel), 1983; Lecture Notes in Control and Information Science. Berlin: Springer-Verlag pp. 473–493.
14. Isidori, A., Krener, A. J., Gori-Giorgi, C., & Monaco, S. 1981 Nonlinear decoupling via feedback: a differential-geometric approach. *IEEE Trans. Automat. Control* AC-**26**, 331–345.
15. Macmillan, B. 1952 Introduction to formal realization theory II. *Bell Syst. Tech. J.* **31**, 541–600.
16. Malabre, M. 1982 Structure à l'infini des triplets invariants. Application à la poursuite parfaite de modèle, pp. 43–53 in: Bensoussan A., & Lions, J. L. (eds.) *Analysis and Optimization of Systems*, Lecture Notes in Control and Information Sciences, vol. 44. New York: Springer.
17. Malabre, M., & Kucera, V. 1983 Infinite structure and exact model matching problem: polynomial and geometric approaches, Rapport interne 0183, ENSM, Université de Nantes.
18. Morse, A. S. 1973 Structural invariants of linear multivariable systems. *SIAM J. Control*, **11**, 446–465.
19. Morse, A. S. 1976 System invariants under feedback and cascade control, pp. 61–74 in: Marchesini, G., & Mitter, S. K. (eds.) *Mathematical Systems Theory* (Proceedings of the International Symposium, Udine, Italy, 1975), Lecture Notes in Economics and Mathematical Systems, vol. 131. New York: Springer.
20. Nijmeijer, H. 1981 Controlled invariance for affine control systems. *Int. J. Control*, **34**, 824–833.
21. Nijmeijer, H. 1982 Invertibility of affine nonlinear control systems: a geometric approach. *Systems Control Lett.*, **2**, 122–129.
22. Nijmeijer, H. & van der Schaft, A. J. Controlled invariance for nonlinear systems. *IEEE Trans. Autom. Control*, AC-**27**, 904–914, 1982.
23. Nijmeijer, H., & Schumacher, J. M. 1985 Zeros at infinity for affine nonlinear control systems. *IEEE Trans. Autom. Control*, AC-**30**, 566–573.
24. Owens, D. H. 1978 On structural invariants and the root-loci of linear multivariable systems. *Int. J. Control*, **28**, 187–196.

25. ROSENBROCK, H. H. 1970 *State-space and Multivariable Theory*, Wiley, New York.
26. SAIN, M. K., & MASSEY, J. L. 1969 Invertibility of linear time-invariant dynamical systems. *IEEE Trans. Autom. Control*, AC-**14**, 141–149.
27. SILVERMAN, L. M. 1969 Inversion of multivariable linear systems. *IEEE Trans. Autom. Control*, AC-**14**, 270–276.
28. SILVERMAN, L. M. 1976 Discrete Riccati equations: alternative algorithms, asymptotic properties, and system theory interpretations. In: Leondes, C. T. (ed.). *Control and Dynamic Systems*, **12**, 313–386.
29. SILVERMAN, L. M., & KITAPCI, A. 1983 System structure at infinity. *Syst. Control Lett.*, **3**, 123–131.
30. SINGH, S. N. 1981 A modified algorithm for invertibility in nonlinear systems. *IEEE Trans. Automat. Control*, AC-**26**, 595–598.
31. SINGH, S. P. 1970 A note on inversion of linear systems. *IEEE Trans. Automat. Control*, AC-**15**, 492–493.
32. THORP, J. S. 1973 The singular pencil of a linear dynamical system. *Int. J. Control* **18**, 577–596.
33. VANDEWALLE, J., & DEWILDE, P. 1978 A local I/O structure theory for multivariable systems and its application to minimal cascade realization. *IEEE Trans. Circuits Syst.* CAS-**25**, 279–289.
34. VAN DOOREN, P. M., DEWILDE, P., & VANDEWALLE, J. 1979 On the determination of the Smith-Macmillan form of a rational matrix from its Laurent expansion. *IEEE Trans. Circuits Syst.*, CAS-**26**, 180–189.
35. VARDULAKIS, A. I. G. 1980 On infinite zeros. *Int. J. Contr.*, **32**, 849–866.
36. VERGHESE, G. C. 1978 Infinite-frequency behavior in generalized dynamical systems, Ph.D. dissertation, Dept. Elec. Eng., Stanford University.
37. VERGHESE, G. C., & KAILATH, T. 1981 Rational matrix Structure, *IEEE Trans. Automat., Control* AC-**26**, 434–439.
38. VAN DER WEIDEN, A. J. J., & BOSGRA, O. H. 1979 The determination of structural properties of a linear multivariable system by operations of system similarity. 1. Strictly proper systems. *Int. J. Control* **29**, 835–860.
39. WONHAM, W. M., 1979 *Linear Multivariable Control: a Geometric Approach*. New York: Springer.

Appendix

Our aim in this appendix is to prove Theorem 4.2. To keep the notation compact, we shall suppress indication of the x-dependence of matrices, functions, etc. We assume the convention that matrix multiplication has priority over differentiation, i.e. if $\mathbf{Z}$ is a vector field and R and S are matrices of compatible sizes, then the expression $\mathrm{L}_{\mathbf{Z}}RS$ is always read as $\mathrm{L}_{\mathbf{Z}}(RS)$ (and not as $(\mathrm{L}_{\mathbf{Z}}R)S$). If $\mathbf{Z}$ and R are such that $\mathrm{L}_{\mathbf{Z}}R = 0$, the product rule of differentiation becomes

$$\mathrm{L}_{\mathbf{Z}}RS = R\mathrm{L}_{\mathbf{Z}}S \tag{A.0}$$

where the right-hand side is read as $R(\mathrm{L}_{\mathbf{Z}}S)$, of course. This 'commuting' property will be used often. Finally we emphasize once more that we work with respect to a fixed sequence of row-transformation matrices $R_0, \ldots, R_{\alpha-1}$, which have been formed according to the rules in the algorithm. We start by proving a lemma.

LEMMA A.1 *Let $k \in \{1, \ldots, \alpha\}$. For vector fields Z satisfying*

$$\mathrm{L}_{\mathbf{Z}}R_l = 0 \quad (0 \leqslant l \leqslant k-2) \tag{A.1}$$

and

$$\mathrm{L}_{\mathbf{Z}}\hat{\boldsymbol{C}}^l = \mathbf{0} \quad (0 \le l \le k-1), \tag{A.2}$$

the following holds true. Smooth functions $\gamma_1, \ldots, \gamma_m$ *satisfy*

$$[\boldsymbol{Z}, \mathbf{A}] - \sum_{i=1}^{m} \boldsymbol{B}_i\gamma_i \in \bigcap_{i=0}^{k-1} \ker \hat{\boldsymbol{C}}_*^i \tag{A.3}$$

if and only if

$$\mathrm{L}_{\mathbf{Z}}\begin{bmatrix} \bar{\boldsymbol{C}}^{k-1} \\ \mathrm{L}_{\mathbf{A}}\hat{\boldsymbol{C}}^{k-1} \end{bmatrix} = \sum_{i=1}^{m} \begin{bmatrix} \bar{\boldsymbol{D}}^{k-1} \\ \mathrm{L}_{\mathbf{B}_i}\hat{\boldsymbol{C}}^{k-1} \end{bmatrix}\gamma_i. \tag{A.4}$$

Moreover, for $i \in \{1, \ldots, m\}$, *smooth functions* $\gamma_1, \ldots, \gamma_m$ *satisfy*

$$[\boldsymbol{Z}, \boldsymbol{B}_i] - \sum_{i=1}^{m} \boldsymbol{B}_i\gamma_i \in \bigcap_{i=0}^{k-1} \ker \hat{\boldsymbol{C}}_*^i \tag{A.5}$$

if and only if

$$\mathrm{L}_{\mathbf{Z}}\begin{bmatrix} \bar{\boldsymbol{D}}_i^{k-1} \\ \mathrm{L}_{\mathbf{B}_i}\hat{\boldsymbol{C}}^{k-1} \end{bmatrix} = \sum_{i=1}^{m} \begin{bmatrix} \bar{\boldsymbol{D}}_i^{k-1} \\ \mathrm{L}_{\mathbf{B}_i}\hat{\boldsymbol{C}}^{k-1} \end{bmatrix}\gamma_i. \tag{A.6}$$

Proof. The proof is by induction. For $k = 1$, the first part of the statement says that if $\boldsymbol{Z} \in \ker \boldsymbol{C}_*^0$ $(= \ker \boldsymbol{C}_*)$, then smooth functions $\gamma_1, \ldots, \gamma_m$ satisfy

$$[\boldsymbol{Z}, \mathbf{A}] - \sum_{i=1}^{m} \boldsymbol{B}_i\gamma_i \in \ker \boldsymbol{C}_* \tag{A.7}$$

if and only if

$$\mathrm{L}_{\mathbf{Z}}\mathrm{L}_{\mathbf{A}}\boldsymbol{C} = \sum_{i=1}^{m} \gamma_i \mathrm{L}_{\mathbf{B}_i}\boldsymbol{C}. \tag{A.8}$$

This is obviously true, if we note that

$$\mathrm{L}_{[\mathbf{Z}, \mathbf{A}]}\boldsymbol{C} = \mathrm{L}_{\mathbf{Z}}\mathrm{L}_{\mathbf{A}}\boldsymbol{C} - \mathrm{L}_{\mathbf{A}}\mathrm{L}_{\mathbf{Z}}\boldsymbol{C} = \mathrm{L}_{\mathbf{Z}}\mathrm{L}_{\mathbf{A}}\boldsymbol{C} \tag{A.9}$$

and

$$\sum_{i=1}^{m} \gamma_i \mathrm{L}_{\mathbf{B}_i}\boldsymbol{C} = \mathrm{L}_{\sum \mathbf{B}_i\gamma_i}\boldsymbol{C}. \tag{A.10}$$

Now suppose that the first part of the statement is true for some given $k \in \{1, \ldots, \alpha-1\}$. Let $\boldsymbol{Z}$ satisfy $\mathrm{L}_{\mathbf{Z}}R_l = 0$ $(0 \le l \le k-1)$ and $\mathrm{L}_{\mathbf{Z}}\hat{\boldsymbol{C}}^l = \mathbf{0}$ $(0 \le l \le k)$, and assume first that $\gamma_1, \ldots, \gamma_m$ are such that

$$\mathrm{L}_{\mathbf{Z}}\begin{bmatrix} \bar{\boldsymbol{C}}^{k} \\ \mathrm{L}_{\mathbf{A}}\hat{\boldsymbol{C}}^{l} \end{bmatrix} = \sum_{i=1}^{m} \begin{bmatrix} \bar{\boldsymbol{D}}_i^{k} \\ \mathrm{L}_{\mathbf{B}_i}\hat{\boldsymbol{C}}^{k} \end{bmatrix}\gamma_i. \tag{A.11}$$

Inserting (4.5) and (4.6) in the top line of (A.11) we obtain

$$\mathrm{L}_{\mathbf{Z}}\bar{R}_{k-1}\begin{bmatrix} \bar{\boldsymbol{C}}^{k-1} \\ \mathrm{L}_{\mathbf{A}}\hat{\boldsymbol{C}}^{k-1} \end{bmatrix} = \bar{R}_{k-1}\sum_{i=1}^{m} \begin{bmatrix} \bar{\boldsymbol{D}}_i^{k-1} \\ \mathrm{L}_{\mathbf{B}_i}\hat{\boldsymbol{C}}^{k-1} \end{bmatrix}\gamma_i. \tag{A.12}$$

Note that

$$\mathrm{L}_{\mathbf{Z}}\hat{R}_{k-1}\begin{bmatrix}\bar{\boldsymbol{C}}^{k-1}\\ \mathrm{L}_{\mathbf{A}}\hat{\boldsymbol{C}}^{k-1}\end{bmatrix}=\mathrm{L}_{\mathbf{Z}}\hat{\boldsymbol{C}}^{k}=\mathbf{0}. \tag{A.13}$$

From (4.3) we have

$$\hat{R}_{k-1}\sum_{i=1}^{m}\begin{bmatrix}\bar{\boldsymbol{D}}^{k-1}\\ \mathrm{L}_{\mathbf{B}_i}\hat{\boldsymbol{C}}^{k+1}\end{bmatrix}\gamma_i=\mathbf{0}. \tag{A.14}$$

Therefore we can write (cf. (A.11))

$$\mathrm{L}_{\mathbf{Z}}R_{k-1}\begin{bmatrix}\bar{\boldsymbol{C}}^{k-1}\\ \mathrm{L}_{\mathbf{A}}\hat{\boldsymbol{C}}^{k-1}\end{bmatrix}=R_{k-1}\sum_{i=1}^{m}\begin{bmatrix}\bar{\boldsymbol{D}}^{k-1}\\ \mathrm{L}_{\mathbf{B}_i}\hat{\boldsymbol{C}}^{k-1}\end{bmatrix}\gamma_i. \tag{A.15}$$

Because $\mathrm{L}_{\mathbf{Z}}R_{k-1}=0$ this can be re-written as

$$R_{k-1}\mathrm{L}_{\mathbf{Z}}\begin{bmatrix}\bar{\boldsymbol{C}}^{k-1}\\ \mathrm{L}_{\mathbf{A}}\hat{\boldsymbol{C}}^{k-1}\end{bmatrix}=R_{k-1}\sum_{i=1}^{m}\begin{bmatrix}\bar{\boldsymbol{D}}_i^{k-1}\\ \mathrm{L}_{\mathbf{B}_i}\hat{\boldsymbol{C}}^{k-1}\end{bmatrix}\gamma_i. \tag{A.16}$$

Since R_{k-1} is invertible we may cancel it on both sides. The induction assumption then shows that

$$[\mathbf{Z},\mathbf{A}]-\sum_{i=1}^{m}\boldsymbol{B}_i\gamma_i\in\bigcap_{i=0}^{k-1}\ker\hat{\boldsymbol{C}}_*^{i}. \tag{A.17}$$

Now, the bottom line in (A.11) reads

$$\mathrm{L}_{\mathbf{Z}}\mathrm{L}_{\mathbf{A}}\hat{\boldsymbol{C}}^{k}=\sum_{i=1}^{m}\gamma_i\mathrm{L}_{\mathbf{B}_i}\hat{\boldsymbol{C}}^{k}. \tag{A.18}$$

Using the fact that $\mathrm{L}_{\mathbf{Z}}\hat{\boldsymbol{C}}^{k}=\mathbf{0}$ this gives us

$$[\mathbf{Z},\mathbf{A}]-\sum_{i=1}^{m}\boldsymbol{B}_i\gamma_i\in\ker\hat{\boldsymbol{C}}_*^{k} \tag{A.19}$$

which is just what we needed to complete this part of the proof. Now assume that $\gamma_i,\ldots,\gamma_m$ are such that

$$[\mathbf{Z},\mathbf{A}]-\sum_{i=1}^{m}\boldsymbol{B}_i\gamma_i\in\bigcap_{i=1}^{m}\ker\hat{\boldsymbol{C}}_*^{i}. \tag{A.20}$$

Then

$$\begin{aligned}\mathrm{L}_{\mathbf{Z}}\bar{\boldsymbol{C}}^{k}&=\mathrm{L}_{\mathbf{Z}}\bar{R}_{k-1}\begin{bmatrix}\bar{\boldsymbol{C}}^{k-1}\\ \mathrm{L}_{\mathbf{A}}\hat{\boldsymbol{C}}^{k-1}\end{bmatrix}=\bar{R}_{k-1}\mathrm{L}_{\mathbf{Z}}\begin{bmatrix}\bar{\boldsymbol{C}}^{k-1}\\ \mathrm{L}_{\mathbf{A}}\hat{\boldsymbol{C}}^{k-1}\end{bmatrix}\\ &=\bar{R}_{k-1}\sum_{i=1}^{m}\begin{bmatrix}\bar{\boldsymbol{D}}_i^{k-1}\\ \mathrm{L}_{\mathbf{B}_i}\hat{\boldsymbol{C}}^{k-1}\end{bmatrix}\gamma_i=\sum_{i=1}^{m}\bar{\boldsymbol{D}}_i^{k}\gamma_i\end{aligned} \tag{A.21}$$

where we used (4.5), the equality $\mathrm{L}_{\mathbf{Z}}R_{k-1}=0$ (which implies, of course,

$L_Z\bar{R}_{k-1}=0$), the induction assumption, and (4.6). Moreover

$$L_Z L_A \hat{C}^k = L_{[Z,A]}\hat{C}^k = \sum_{i=1}^{m} \gamma_i L_{B_i}\hat{C}^k \tag{A.22}$$

since we have given that $L_Z\hat{C}^k=0$ and (A.20) holds. Combining (A.21) and (A.22) we see that the first statement of the lemma has been proved. The proof of the second part of the lemma is similar. □

We can now proceed to the proof of the theorem.

Proof (of Theorem 4.2). We first show that

$$\mathscr{E}^k \cap \bigcap_{i=0}^{k-1} \ker \hat{C}_*^i \subset \mathscr{V}^k \tag{A.23}$$

for $k=1,\ldots,\alpha$. To do this we apply induction. For $k=1$ the statement is true by definition. Now suppose that (A.23) holds for some given $k\in\{1,\ldots,\alpha-1\}$. Let $Z\in V(\mathcal{M})$ be such that

$$Z \in \mathscr{E}^{k+1} \cap \bigcap_{i=0}^{k} \ker \hat{C}_*^i. \tag{A.24}$$

From the induction assumption it follows immediately that $Z\in\mathscr{V}^k$, and so it remains to show that $[Z,A]\in\mathscr{V}^k+\Delta_0$ and $[Z,B_i]\in\mathscr{V}^k+\Delta_0$ for each $i\in\{1,\ldots,m\}$. First note that

$$\hat{R}_{k-1}L_Z\begin{bmatrix}\bar{C}^{k-1}\\ L_A\hat{C}^{k-1}\end{bmatrix} = L_Z\hat{R}_{k-1}\begin{bmatrix}\bar{C}^{k-1}\\ L_A\hat{C}^{k-1}\end{bmatrix} = L_A\hat{C}^k = \mathbf{0}. \tag{A.25}$$

By (4.3) this means that there exist smooth functions γ_i $(i=1,\ldots,m)$ such that

$$L_Z\begin{bmatrix}\bar{C}^{k-1}\\ L_A\hat{C}^{k-1}\end{bmatrix} = \sum_{i=1}^{m}\begin{bmatrix}\bar{D}_i^{k-1}\\ L_{B_i}\hat{C}^{k-1}\end{bmatrix}\gamma_i. \tag{A.26}$$

By Lemma A.1 this is equivalent to

$$[Z,A]-\sum_{i=1}^{m} B_i\gamma_i \in \bigcap_{i=0}^{k-1}\ker \hat{C}_*^i. \tag{A.27}$$

As noted in Lemma 4.1 it follows from $Z\in\mathscr{E}^{k+1}$ that $[Z,A]\in\mathscr{E}^k$. Also the assumption $\Delta_0\subset\mathscr{E}^{\alpha-1}$ implies

$$\sum_{i=1}^{m} B_i\gamma_i \in \mathscr{E}^k. \tag{A.28}$$

Therefore we can apply the induction assumption to the vector field $[Z,A]-\sum_{i=1}^{m} B_i\gamma_i$, to the effect that

$$[Z,A]-\sum_{i=1}^{m} B_i\gamma_i \in \mathscr{V}^k. \tag{A.29}$$

Of course, this implies that

$$[\mathbf{Z}, \mathbf{A}] \in \mathscr{V}^k + \Delta_0. \tag{A.30}$$

Next, we also have

$$\hat{R}_{k-1}\mathrm{L}_{\mathbf{Z}}\begin{bmatrix} \bar{\boldsymbol{D}}_i^{k-1} \\ \mathrm{L}_{\boldsymbol{B}_i}\hat{\boldsymbol{C}}^{k-1} \end{bmatrix} = \mathrm{L}_{\mathbf{Z}}\hat{R}_{k-1}\begin{bmatrix} \bar{\boldsymbol{D}}_i^{k-1} \\ \mathrm{L}_{\boldsymbol{B}_i}\hat{\boldsymbol{C}}^{k-1} \end{bmatrix} = \mathbf{0}. \tag{A.31}$$

So again there exist functions γ_i $(i = 1, \ldots, m)$ such that

$$\begin{bmatrix} \bar{\boldsymbol{D}}_i^{k-1} \\ \mathrm{L}_{\boldsymbol{B}_i}\hat{\boldsymbol{C}}^{k-1} \end{bmatrix} = \sum_{i=1}^{m} \begin{bmatrix} \bar{\boldsymbol{D}}_i^{k-1} \\ \mathrm{L}_{\boldsymbol{B}_i}\hat{\boldsymbol{C}}^{k-1} \end{bmatrix} \gamma_i \tag{A.32}$$

which means that

$$[\mathbf{Z}, \boldsymbol{B}_i] - \sum_{i=1}^{m} \boldsymbol{B}_i\gamma_i \in \bigcap_{i=0}^{k-1} \ker \hat{\boldsymbol{C}}_*^i. \tag{A.33}$$

Because $\mathbf{Z} \in \mathscr{C}^{k+1} \subset \mathscr{C}^k$, $\Delta_0 \subset \mathscr{C}^k$, and $\mathscr{C}^k$ is involutive, we can apply the induction assumption and conclude that

$$[\mathbf{Z}, \boldsymbol{B}_i] \in \mathscr{V}^k + \Delta_0. \tag{A.34}$$

Since this holds for all $i \in \{1, \ldots, m\}$ we have shown that $\mathbf{Z} \in \mathscr{V}^{k+1}$, and the first part of the proof is completed.

Next we have to prove that

$$\mathscr{C}^k \cap \mathscr{V}^k \subset \bigcap_{i=0}^{k-1} \ker \hat{\boldsymbol{C}}_*^i \tag{A.35}$$

for all $k \in \{1, \ldots, \alpha\}$. Again we proceed by induction. For $k = 1$, the statement is true by definition. Now suppose that (A.35) holds for some given $k \in \{1, \ldots, \alpha - 1\}$, and let $\mathbf{Z}$ be such that

$$\mathbf{Z} \in \mathscr{C}^{k+1} \cap \mathscr{V}^{k+1}. \tag{A.36}$$

Because $\mathscr{C}^{k+1} \cap \mathscr{V}^{k+1} \subset \mathscr{C}^k \cap \mathscr{V}^k$ it follows immediately from the induction assumption that

$$\mathbf{Z} \in \bigcap_{i=0}^{k-1} \ker \hat{\boldsymbol{C}}_*^i \tag{A.37}$$

and so it remains to show that $\mathrm{L}_{\mathbf{Z}}\hat{\boldsymbol{C}}^k = 0$. Because $\mathbf{Z} \in \mathscr{V}^{k+1}$ we have $[\mathbf{Z}, \mathbf{A}] \in \mathscr{V}^k + \Delta_0$, and so there exist functions γ_i $(i = 1, \ldots, m)$ such that

$$[\mathbf{Z}, \mathbf{A}] - \sum_{i=1}^{m} \boldsymbol{B}_i\gamma_i \in \mathscr{V}^k. \tag{A.38}$$

By the assumptions, we also have

$$[\mathbf{Z}, \mathbf{A}] - \sum_{i=1}^{m} \boldsymbol{B}_i\gamma_i \in \mathscr{C}^k. \tag{A.39}$$

So we can apply the induction assumption and conclude that

$$[\mathbf{Z}, \mathbf{A}] - \sum_{i=0}^{m} \boldsymbol{B}_i \gamma_i \in \bigcap_{i=0}^{k-1} \ker \hat{\boldsymbol{C}}_*^i. \tag{A.40}$$

By Lemma A.1 this is equivalent to

$$\mathrm{L}_{\mathbf{Z}} \begin{bmatrix} \bar{\boldsymbol{C}}_i^{k-1} \\ L_{\mathbf{A}} \hat{\boldsymbol{C}}^{k-1} \end{bmatrix} = \sum_{i=1}^{m} \begin{bmatrix} \bar{\boldsymbol{D}}_i^{k-1} \\ \mathrm{L}_{\boldsymbol{B}_i} \hat{\boldsymbol{C}}^{k-1} \end{bmatrix} \gamma_i. \tag{A.41}$$

We now compute:

$$\mathrm{L}_{\mathbf{Z}} \hat{\boldsymbol{C}}^k = \mathrm{L}_{\mathbf{Z}} \hat{\mathrm{R}}_{k-1} \begin{bmatrix} \bar{\boldsymbol{C}}^{k-1} \\ \mathrm{L}_{\mathbf{A}} \hat{\boldsymbol{C}}^{k-1} \end{bmatrix} = \hat{\mathrm{R}}_{k-1} \mathrm{L}_{\mathbf{Z}} \begin{bmatrix} \bar{\boldsymbol{C}}^{k-1} \\ \mathrm{L}_{\mathbf{A}} \hat{\boldsymbol{C}}^{k-1} \end{bmatrix} = \hat{\mathrm{R}}_{k-1} \sum_{i=1}^{m} \begin{bmatrix} \bar{\boldsymbol{D}}_i^{k-1} \\ \mathrm{L}_{\boldsymbol{B}_i} \hat{\boldsymbol{C}}^{k-1} \end{bmatrix} \gamma_i = \mathbf{0}. \tag{A.42}$$

The proof is complete. □