On Carleman Linearization of Linearly Observable Polynomial Systems

Dorota Mozyrska[1] and Zbigniew Bartosiewicz[2]

[1] Bialystok Technical University, Faculty of Computer Science, Poland
admoz@w.tkb.pl
[2] Bialystok Technical University, Faculty of Computer Science, Poland
bartos@pb.bialystok.pl

Summary. Carleman linearization is used to transform a polynomial control system with output, defined on n-dimensional space, into a linear or bilinear system evolving in the space of infinite sequences. Such a system is described by infinite matrices with special properties. Linear observability of the original system is studied. It means that all coordinate functions can be expressed as linear combinations of functions from the observation space. It is shown that this property is equivalent to a rank condition involving matrices that appear in the Carleman linearization. This condition is equivalent to observability of the first n coordinates of the linearized system.

1 Introduction

Carleman linearization is a procedure that allows to embed a finite dimensional system of differential equations, with analytic or polynomial data, into a system of linear differential equations on an infinite dimensional space. Thus we trade polynomials (or analytic functions) that describe the system for the infinite matrices of the Carleman linearization. The reader may consult [11], which gives a general introduction to the subject.

There were several attempts to apply Carleman linearization in control theory. Let us mention [17], where this technique was used for linear systems with polynomial output. We consider a more general situation, where also the system's dynamics is polynomial. Our goal is to relate observability of the original system and its Carleman linearization. We study linear observability of a polynomial system on $\mathbb{R}^n$, which means that all the coordinate functions can be expressed as linear combinations of functions from the observation space of the system. Two cases are studied. The simpler one concerns a system without control, which leads to a linear system with output, also without control. We show that the original system is linearly observable if and only if the first n coordinates of its Carleman linearized system are observable, and express this property by a rank condition involving the matrices of the linearized system. The other case, where the original system contains control, leads to a bilinear infinite dimensional system. A similar rank condition for linear observability is presented.

Checking observability of infinite dimensional linear systems is not easy as we have to deal with infinite matrices. Though the Kalman condition of observability holds for this class of systems, it must be expressed in a different way and there is no finite algorithm to check this. Here we have a simpler task, as only finitely many coordinates are to be observed. Several results on observability of systems described by infinite matrices were given in our papers [2, 3] and duality between observability and controllability was studied in [16]. In [2] we considered discrete-time systems, which are easier to study since existence and uniqueness of solutions is always guaranteed. The continuous-time case, examined in [3], is much harder as even for linear infinite systems of differential equations, solutions of initial value problems may not exist or be nonunique. Concerning this subject more can be found in [9, 12, 13]. In the next section we provide the reader with basic definitions and facts. The Banach space case, studied in [5, 6], is more regular. Observability of nonlinear infinite dimensional systems was studied in [14, 15]. Such systems may appear as infinite extensions of finite dimensional nonlinear control systems (see e.g. [10]).

2 Preliminaries

Let $\mathbb{R}^\infty$ denote the linear space of all real sequences denoted by infinite columns. Let $\Pi_n : \mathbb{R}^\infty \to \mathbb{R}^n$ denotes the projection on the first n coordinates, that is if $z = (z_1, \ldots, z_n, \ldots)^T \in \mathbb{R}^\infty$ then $\Pi_n(z) = (z_1, \ldots, z_n)^T$. We say that a function $f : \mathbb{R}^\infty \to \mathbb{R}$ is *finitely presented* on $\mathbb{R}^\infty$ if there is $n \in \mathbb{N}$ and a function $\tilde{f} : \mathbb{R}^n \to \mathbb{R}$ such that $f = \tilde{f} \circ \Pi_n$. If we consider $\mathbb{R}^\infty$ as a topological space we use the product (Tikhonov) topology. A basis of this topology consists of the sets $U = \prod_{i\in\mathbb{N}} U_i$, where U_i is the open subset of $\mathbb{R}$ and $U_i = \mathbb{R}$ for all i, but finite number values of i. It is the weakest topology for which projections Π_n are continuous.

From [1] we have the following:

Proposition 1. *Let $L(\mathbb{R}^\infty, \mathbb{R})$ be the space of linear and continuous functions on $\mathbb{R}^\infty$. If $f \in L(\mathbb{R}^\infty, \mathbb{R})$ then f is finitely presented and there are $n_f \in \mathbb{N}$ and $c_1, \ldots, c_{n_f} \in \mathbb{R}$ such that for all $z \in \mathbb{R}^\infty$, $f(z) = \sum_{i=1}^{n_f} c_i z_i$.*

We deal with differential systems described by infinite matrices which can be interpreted as functions from $\mathbb{R}^\infty \times \mathbb{R}^\infty$ to $\mathbb{R}$. We say that a matrix $A = (a_{ij})_{i,j\in\mathbb{N}}$ is *row-finite* if for each $i \in \mathbb{N}$ there is $n(i) \in \mathbb{N}$ such that for $j > n(i)$, $a_{ij} = 0$. The matrix is *upper-diagonal* if $a_{ij} = 0$ for $j < i$. The set of row-finite matrices forms an algebra over $\mathbb{R}$ with a unit $E = (\delta_{ij})_{i,j\in\mathbb{N}}$. Hence if A is a row-finite matrix, then for each $k \in \mathbb{N}$, A^k is well defined and it is a row-finite matrix as well.

Let A be an infinite row-finite matrix. Then the system of differential equations $\dot{z}(t) = \frac{dz}{dt} = Az(t)$ is called a row-finite system. If together with this system we consider the initial condition $z(0) = z^0 \in \mathbb{R}^\infty$ then the discussion of existence and uniqueness of solutions of the initial value problem can be found, e.g., in [3, 6, 14, 15]. In particular the concept of formal solutions is there presented.

Proposition 2. *Let* $\frac{dz}{dt} = Az, z(0) = z^0 \in \mathbb{R}^\infty$*, be the initial value problem with* A *being a row-finite matrix. Then it has the unique formal solution* $\Gamma_{z^0,A} := \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k z^0$.

We are concerned with the system with output:

$$(\Sigma) : \begin{array}{r} \dot{z}(t) = \ Az(t) \\ y(t) = Cz(t), \end{array} \tag{1}$$

where $z : [0,\infty) \to \mathbb{R}^\infty, y : [0,\infty) \to \mathbb{R}^p$, and $A \in \mathbb{R}^\infty \times \mathbb{R}^\infty$ and $C \in \mathbb{R}^p \times \mathbb{R}^\infty$ are row–finite. Let $z^0 \in \mathbb{R}^\infty$. Given a formal solution $\Gamma_{z^0,A}$ of the dynamical part of the system and corresponding to the initial condition z^0 we define the formal output: $\mathcal{Y}_{z^0} = C\Gamma_{z^0,A}$.

Definition 1. *We say that* $z^1, z^2 \in \mathbb{R}^\infty$ *are* indistinguishable *(with respect to* (Σ)*) if* $\mathcal{Y}_{z^1} = \mathcal{Y}_{z^2}$. *Otherwise* z^1, z^2 *are* distinguishable. *We say that the system* (Σ) *is* observable *if any two distinct points are distinguishable.*

Proposition 3. ([3])
The points $z^1, z^2 \in \mathbb{R}^\infty$ *are indistinguishable iff for all* $k \in \mathbb{N} \cup \{0\} : CA^k z^1 = CA^k z^2$.

Corollary 1. ([3])
System (Σ) *is observable if and only if* $\forall\ n \in \mathbb{N}\ \exists k \in \mathbb{N} \cup \{0\}$:

$$\operatorname{rank} \begin{pmatrix} C \\ \vdots \\ CA^k \end{pmatrix} = \operatorname{rank} \begin{pmatrix} C \\ \vdots \\ CA^k \\ E_n^T \end{pmatrix}, \tag{2}$$

where E_n^T *denotes the infinite row with* 1 *at the* n*-th position.*

Let $D := \begin{pmatrix} C \\ CA \\ \vdots \end{pmatrix}$. Since the rows of D correspond to derivatives of the output, one can characterize observability as possibility to compute every state variable as a linear combination of finitely many outputs and their derivatives.

As using formal solutions and formal outputs brought us to the characterization of observability based on matrices of the systems, we extend now this concept by defining some kind of observability of infinite bilinear systems.

Definition 2. *Let us consider an infinite bilinear system*

$$\begin{aligned} \dot{z} &= (A + uB)z, \\ y &= Cz, \end{aligned} \tag{3}$$

where $z : [0, \infty) \to \mathbb{R}^\infty, y : [0, \infty) \to \mathbb{R}^p$*,* $A, B \in \mathbb{R}^\infty \times \mathbb{R}^\infty$ *and* $C \in \mathbb{R}^p \times \mathbb{R}^\infty$ *are row–finite matrices. Let*

$$\Gamma(C, A, B)_k := col[C, CA, CB, CA^2, CAB, CBA, CB^2, \dots, CB^k].$$

System (3) is said to be formally observable *if* $\forall\ n \in \mathbb{N}\ \exists k \in \mathbb{N} \cup \{0\}$ *such that*

$$\operatorname{rank} \Gamma(C, A, B)_k = \operatorname{rank} \begin{pmatrix} \Gamma(C, A, B)_k \\ E_n^T \end{pmatrix}. \tag{4}$$

Remark 1. Condition (4) becomes the same as (2) for $B = 0$.

We will consider the situation when condition (4) is satisfied only for some number of variables.

Definition 3. *We say that system (3)* is observable with respect to the variable z_i *if for* $n = i$ *there is* k *such that (4) holds.*

Remark 2. System (3) is formally observable iff it is observable with respect to each variable.

3 Carleman linearization

By $\mathcal{M}(m, n)$ we denote the set of matrices of dimensions $m \times n$ with real elements.

Let the function $x : \mathbb{R} \supset J \to \mathbb{R}^n, x \in C^1(J)$, be a solution of the first order system of ordinary differential equations:

$$\Sigma : \frac{dx}{dt} = F(x), \tag{5}$$

where $F = (f_1, \dots, f_n)^T$ is a vector field whose components are polynomials without constant term, i.e. $F(0) = 0$. Then we can write $F(x) = \sum_{k=1}^m F_k(x)$, where each F_k is a vector of homogeneous polynomials of degree k and $m \in \mathbb{N}$, $m \geq \max\limits_i \deg f_i$.

For every integer $k \geq 1$, let H_k denote the space of homogeneous polynomials of degree k in n variables $x_1, x_2, \dots, x_n$. In H_k we choose the canonical bases $\{x^q = x_1^{q_1} \dots x_n^{q_n}\}$ with $q = (q_1, \dots, q_n)$, where $q_i \in \mathbb{N}$ and $|q| = \sum_{i=1}^n q_i = k$. We use the lexicographic order in the set of monomials x^q induced by the order $x_1 < x_2 < \dots < x_n$. We use the following notation:

$$\begin{aligned} &e_1^k = x_1^k, e_2^k = x_1^{k-1} x_2, \dots, e_{d_k}^k = x_n^k \\ &\zeta_k = \left(e_1^k, \dots, e_{d_k}^k\right)^T \end{aligned}, \tag{6}$$

where $d_k = \binom{n+k-1}{k} = \dim H_k$. Hence if $\varphi \in H_k$ then $\varphi(x) = \sum_{|q|=k} \beta_q x^q = \sum_{i=1}^{d_k} \alpha_i e_i^k$. Let H^∞ be the space of all polynomials (without constant term) in variables $x_1, \ldots, x_n$. Then H^∞ may be represented by the direct sum of the family $\{H_k\}_{k\in\mathbb{N}}$ of the spaces of homogeneous polynomials, i.e. $H^\infty := \bigoplus_{k\in\mathbb{N}} H_k$. Let us mention that the direct sum of an infinite family of modules is defined to be the set of all functions w with domain $\mathbb{N}$ such that $w(k) \in H_k$ for all $k \in \mathbb{N}$ and $w(k) = 0$ for all but finitely many indices k.

Let P be a polynomial of degree r with $P(0) = 0$. So $P \in H^\infty$ and there are homogeneous polynomials $\varphi_k \in H_k, k = 1, \ldots, m$, such that $P(x) = \sum_{k=1}^{r} \varphi_k(x)$. Using the above notation we can write that

$$P(x) = \sum_{k=1}^{r} \sum_{i=1}^{d_k} p_{ki} e_i^k = \sum_{k=1}^{r} (p_{k1}, \ldots, p_{kd_k}) \zeta_k. \tag{7}$$

Then system (5) can be written in the form

$$\frac{dx}{dt} = A_{11}\zeta_1 + \cdots + A_{1m}\zeta_m, \tag{8}$$

where $F_k(x) = A_{1k}\zeta_k$ and matrices $A_{1k} \in \mathcal{M}(n, d_k)$, $i = 1, \ldots, m$. Let us observe that $F_1(x) = A_{11}\zeta_1$, where the matrix $A_{11} \in M(n,n)$, forms the linear part of system (5). Additionally we can obtain matrices A_{1k} by the formula $A_{1k} = \left(\frac{1}{q_1! \cdots q_n!} \frac{\partial^k f_i}{\partial x_1^{q_1} \cdots \partial x_n^{q_n}}(0) \right)$, where $q_i \in \mathbb{N}$ and $\sum_{i=1}^{n} q_i = k$.

Let $\zeta = (\zeta_1^T, \zeta_2^T, \ldots)^T$ be the infinite vector of elements of the basis of H^∞. Then by (7)

$$P(x) = (p_{11}, \ldots, p_{1d_1}, \ldots, p_{r1}, \ldots, p_{rd_r}, 0, \ldots)\zeta. \tag{9}$$

The Lie derivative in the direction of the vector field F of system (5) defines the linear map $D_\Sigma : H^\infty \to H^\infty$ by $(D_\Sigma P)(x) = \nabla P(x) \cdot F(x)$. Let $P(x)$ be in the form (7). Then we have $D_\Sigma P(x) = \sum_{k=1}^{m} \sum_{i=1}^{d_k} p_{ki} D_\Sigma e_i^k = \sum_{k=1}^{m} \sum_{i=1}^{d_k} p_{ki} \nabla e_i^k \cdot F(x)$. As in particular $D_\Sigma x_i = f_i(x)$ then $\begin{pmatrix} D_\Sigma x_1 \\ \vdots \\ D_\Sigma x_n \end{pmatrix} = \sum_{j=1}^{m} A_{1j}\zeta_j$ and there are uniquely determined matrices $A_{kj} \in \mathcal{M}(d_k, d_j)$ such that

$$\begin{pmatrix} D_\Sigma e_1^k \\ \vdots \\ D_\Sigma e_{d_k}^k \end{pmatrix} = \sum_{j=k}^{m+k-1} A_{kj}\zeta_j. \tag{10}$$

Applying (10) to $D_\Sigma P(x)$ we obtain

$$D_\Sigma P(x) = \sum_{k=1}^{r} (p_{k1}, \ldots, p_{kd_k}) \sum_{j=k}^{m+k-1} A_{kj}\zeta_j. \tag{11}$$

This yields

$$D_\Sigma P(x) = (p_{11}, \ldots, p_{1d_1}, \ldots, p_{rd_r}, 0 \ldots) M_F \zeta, \tag{12}$$

where

$$M_F = \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1m} & 0 & 0 & \ldots \\ 0 & A_{22} & \ldots & A_{2m} & A_{2,m+1} & 0 & \ldots \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \end{pmatrix}. \tag{13}$$

In the following example we show how matrices A_{kj} could be determined for the case $n = 2$.

Example 1. We have that $A_{1k} = \left(\frac{1}{q_1! q_2!} \frac{\partial^k f_i}{\partial x_1^{q_1} \partial x_2^{q_2}}(0) \right)$, where $q_i \in \mathbb{N}$ and $q_1 + q_2 = k$. For calculating matrices $A_{2k}, k = 1, \ldots, m+1$ we can use the following. Let us observe that

$$\dot{\zeta}_2 = \begin{pmatrix} D_\Sigma x_1^2 \\ D_\Sigma x_1 x_2 \\ D_\Sigma x_2^2 \end{pmatrix} = \begin{pmatrix} 2x_1 & 0 \\ x_2 & x_1 \\ 0 & 2x_2 \end{pmatrix} \dot{\zeta}_1 = \begin{pmatrix} 2x_1 & 0 \\ x_2 & x_1 \\ 0 & 2x_2 \end{pmatrix} \sum_{k=1}^{m} A_{1k}\zeta_k.$$

The last equality can be written as

$$\dot{\zeta}_2 = \begin{pmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \sum_{k=1}^{m} A_{1k} x_1 \zeta_k + \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 2 \end{pmatrix} \sum_{k=1}^{m} A_{1k} x_2 \zeta_k = \sum_{k=2}^{m+1} A_{2k}\zeta_k,$$

where:

$$A_{22} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} A_{11} \left(I_3 \ \mathbf{0} \right) + \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 2 \end{pmatrix} A_{11} \left(\mathbf{0} \ I_3 \right),$$

where I_3 is the identity matrix of the degree 3. And for $k = 2, \ldots, m$:

$$A_{2,k+1} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} A_{1k} \left(I_{d_k} \ \mathbf{0} \right) + \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 2 \end{pmatrix} A_{1m} \left(\mathbf{0} \ I_{d_k} \right)$$

with I_{d_k} being identity matrices also. Next matrices can be produced in a similar way.

Definition 4. *Let M_F denotes matrix given by (13). The system of equations*

$$\frac{dz}{dt} = M_F z \tag{14}$$

for an infinite sequence of functions $z = (z_1, z_2, \ldots)^T$ with $z_i \in C^1(J), i \in \mathbb{N}$ and with J an open interval, is called the associated infinite linear system to the finite nonlinear system (5). The setting up the associated infinite linear system to a given finite system is called the Carleman linearization procedure (Carleman embedding).

Definition 5. *Let $s = \sum_{i=1}^{k} \dim H_i$. By the truncation of the order $s \geq 1$ of system (14) we mean the following finite dimensional system:*

$$\frac{d}{dt}\begin{pmatrix} z_1 \\ \vdots \\ z_s \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1k} \\ 0 & A_{22} & \dots & A_{2k} \\ & & \ddots & \vdots \\ & & & A_{kk} \end{pmatrix} \begin{pmatrix} z_1 \\ \vdots \\ z_s \end{pmatrix}. \tag{15}$$

Remark 3. [11] Each solution x of system (5) gives a solution of (14). And conversely each solution z of (14) gives $x = (z_1, \dots, z_n)$ as a solution of the system (8).

Remark 4. If $\frac{dx}{dt} = x + a_2x^2 + \dots + a_mx^m$, $a_2, \dots, a_m \in \mathbb{R}$, then the matrix M_F in the system (14) is triangular, with $m-1$ filled diagonal lines above the main diagonal.

Example 2. Let us consider one-dimensional system: $\frac{dx}{dt} = -x + x^2$ with the initial condition: $x(0) = c \in \mathbb{R}$. The solution of the initial value problem is the following: $x(t) = \frac{c}{c+(1-c)\exp(t)}$. Using the Carleman technique we take $z_1 := x, z_2 := x^2, \dots, z_n = x^n, \dots$. Then $\frac{dz_n}{dt} = \frac{dx^n}{dt} = nx^{n-1}\frac{dx}{dt}$ and $\frac{dz_n}{dt} = -nz_n + nz_{n+1}$. Hence $\frac{dz}{dt} = M_F z$, where the matrix $M_F = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 \dots \\ 0 & -2 & 2 & 0 & 0 & 0 \dots \\ 0 & 0 & -3 & 3 & 0 & 0 \dots \\ \vdots & \vdots & & \ddots & \ddots & \ddots \end{pmatrix}$ is infinite-dimensional, triangular and row-finite. Additionally $z_n(0) = c^n$ and $x(t) = z_1(t) = \frac{c}{c+(1-c)\exp(t)}$. We can look at truncated versions of the Carleman linearization, e.g. the system
$\frac{d}{dt}\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -2 & 2 \\ 0 & 0 & -3 \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}$ is an approximation of $\dot{x} = -x + x^2$.

Now the vector field of system (14) defines the map

$$D_{\tilde{\Sigma}} : L(\mathbb{R}^\infty, \mathbb{R}) \to L(\mathbb{R}^\infty, \mathbb{R})$$

by

$$D_{\tilde{\Sigma}} f(z) = CM_F z, \tag{16}$$

where $f(z) = Cz$ and $C = (c_1, c_2, \dots, c_{n_f}, 0, \dots)$.

Let $\zeta = (\zeta_1^T, \zeta_2^T, \dots)^T$ be a basis in H^∞. Let us consider $P \in H^\infty$ in the form (9). Then we define the map $\alpha : H^\infty \to L(\mathbb{R}^\infty, \mathbb{R})$ in the following way

$$\alpha(P)(z) = \sum_{s=1}^{l} c_s z_s = Cz, \tag{17}$$

where $l = \sum_{j=1}^{d_m} d_j$ and $c_s = p_{ki}$ for $s = i + \sum\limits_{j=1}^{k-1} d_j$.

Proposition 4. *The map α defined by (17) is a linear bijective mapping from H^∞ to $L(\mathbb{R}^\infty, \mathbb{R})$.*

Proposition 5. *Let α be the map defined by (17). Then*

$$D_{\tilde{\Sigma}} \circ \alpha = \alpha \circ D_\Sigma. \tag{18}$$

Proof. Let $P \in H^\infty$ be a polynomial of degree r in n-variables, in the form (9). Then by the definition (17) of the map α and the definition (16) of $D_{\tilde{\Sigma}}$ we get $(D_{\tilde{\Sigma}} \circ \alpha)(P)(z) = (p_{11}, \ldots, p_{rd_r}, 0, \ldots)M_F z$. Hence by (12) we get that (18) is true.

Let $D^0_\Sigma P := P$ and $D^k_\Sigma P := D_\Sigma(D^{k-1}_\Sigma P)$. Then by induction we conclude that

Corollary 2. $\alpha(D^k_\Sigma P)(z) = CM^k_F z$.

4 Linear observability of polynomial dynamical system

Let (Σ) be a polynomial system with output:

$$\dot{x} = F(x) \tag{19}$$

$$y = h(x), \tag{20}$$

where $x \in \mathbb{R}^n, y \in \mathbb{R}^p$ and $F = (f_1, f_2, \ldots, f_n)^T, h = (h_1, \ldots, h_p)^T$ are vectors of polynomials without constant terms. Let $F(x) = A_{11}\zeta_1 + \cdots + A_{1m}\zeta_m = \sum_{k=1}^m F_k(x)$, where $F_k(x) = A_{1k}\zeta_k$ and $h_j(x) = C_j\zeta$, where $C_j = (\beta^j_{11}, \ldots, \beta^j_{rd_r}, 0, \ldots)$.

Definition 6. *By $\mathcal{O}(\Sigma)$ we denote the smallest linear subspace of H^∞ containing $h_j, j = 1, \ldots, p$, the components of h, and invariant under the map D_Σ, the action of the vector field F of the system (Σ). The system (Σ) is said to be linearly observable if for each $i = 1, \ldots, n$: $x_i \in \mathcal{O}(\Sigma)$.*

Example 3. Let (Σ) be the following :

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} -x_2 - x_1x_2 \\ x_1 + x_1x_2 \end{pmatrix}$$
$$y = x_1 + x_2,$$

where the dynamics is the same as in the Lotka-Volterra model. Then $\mathcal{O}(\Sigma)(x) = \text{span}\{(D^k_\Sigma y)(x), k = 0, 1, \ldots\} = \text{span}\{x_1 + x_2, x_1 - x_2, \ldots\}$. Hence $x_1 = \frac{1}{2}(y + D_\Sigma y)$ and $x_2 = \frac{1}{2}(y - D_\Sigma y)$ and (Σ) is linearly observable.

By $(\tilde{\Sigma})$ we denote the associated with (Σ) (by Definition 4) infinite-dimensional system:

$$\dot{z} = M_F z \tag{21}$$
$$y = Cz, \tag{22}$$

where $z \in \mathbb{R}^\infty, y \in \mathbb{R}^p$ and the matrix $C = \begin{pmatrix} C_1 \\ \vdots \\ C_p \end{pmatrix}$.

Corollary 3.

1. *For each $k = 0, 1, \ldots$ the following holds: $\alpha(D_\Sigma^k h_j)(z) = C_j M_F^k z$, where for $j = 1, \ldots, p$ and $h_j(x) = C_j \zeta$ are components of the output of (Σ).*
2. *$\alpha\left(\mathcal{O}\left(\Sigma\right)\right) = \mathrm{span}\{C_j M_F^k z, j = 1, \ldots, p$ and $k = 0, 1, \ldots\}$.*

Proposition 6. *Let $E_i = (0, \ldots, 0, 1, 0, \ldots)^T$ be the vector from $\mathbb{R}^\infty$ with 1 at the i-th position. The finite-dimensional polynomial system (Σ) is linearly observable iff there is $k \in \mathbb{N} \cup \{0\}$ such that*

$$\mathrm{rank} \begin{pmatrix} C \\ CM_F \\ \vdots \\ CM_F^k \\ E_1^T \\ \vdots \\ E_n^T \end{pmatrix} = \mathrm{rank} \begin{pmatrix} C \\ CM_F \\ \vdots \\ CM_F^k \end{pmatrix}. \tag{23}$$

Proof. Let $i = 1, \ldots, n$. As, by Proposition 4, α is a bijective mapping from H^∞ to $L(\mathbb{R}^\infty, \mathbb{R})$, we have that $x_i \in \mathcal{O}\left(\Sigma\right) \Leftrightarrow \alpha(x_i) = E_i z \in \alpha\left(\mathcal{O}\left(\Sigma\right)\right)$. Moreover $E_i^T z \in \alpha(\mathcal{O}\left(\Sigma\right)) = \mathrm{span}\{C_j M_F^k z, j = 1, \ldots, p, k \geq 0\}$ iff $E_i^T \in \mathrm{span}\{C_j M_F^k, j = 1, \ldots, p, k = 0, 1, \ldots\}$. It holds for all $i = 1, \ldots, n$ iff the condition (23) is satisfied.

Remark 5. Using Definition 3 we can formulate Proposition 6 as follows: system (Σ) is linearly observable iff its Carleman linearization is observable with respect to variables $z_1, \ldots, z_n$.

Corollary 4. *If the system (Σ) is linearly observable then there is $k \geq \deg h$ such that the truncation of the order k of associated infinite linear system is observable on $\mathbb{R}^k$.*

5 Carleman bilinearization for polynomial control system

Let (Λ) denote the finite dimensional polynomial system with one-dimensional input u:

$$\dot{x} = F(x) + G(x)u \tag{24}$$
$$y = h(x), \tag{25}$$

where F, G are vectors of polynomials without constant term and $u \in \mathcal{U}$, where $\mathcal{U}$ denotes the set of piece wise constant functions $u : [0, T_u] \to \mathbb{R}$ and T_u depends on u, $T_u \geq 0$. Using the same description as for systems without control, equation (24) of (Λ) can be written in the following form: $\dot{x} = \sum_{j=1}^{m} (A_{1j} + uB_{1j})\, \zeta_j$, where $m = \max_{j=1}^{n} \deg(f_j, g_j)$ and $F = (f_1, \ldots, f_n)^T, G = (g_1, \ldots, g_n)^T$. Let D_Λ^u denote the derivation in the direction of the vector field $F + Gu$, for fixed $u \in \mathbb{R}$. Moreover let $D_\Lambda^u = D_0 + D_u$, where by D_0 we mean the derivation in the direction of the vector field F while the derivation D_u is defined by the vector field uG. Let $\mathcal{D}_\Lambda = \{D_\Lambda^u = D_0 + D_u : u \in \mathbb{R}\}$. Then the observability of this system depends on properties of the space $\mathcal{O}(\Lambda)$, which is the smallest subspace of H^∞ that contains all functions $h_j, j = 1, \ldots, p$, and is invariant under the action of the maps from $\mathcal{D}_\Lambda$. Hence, according to Definition 6, (Λ) is linearly observable if for each $i = 1, \ldots, n : x_i \in \mathcal{O}(\Lambda)$.

Similarly as in the case of system Σ, we consider the action of the derivation D_Λ^u on a polynomial $P \in H^\infty$ given by (7). By (10) and (12): $D_\Lambda^u P(x) = \sum\limits_{k=1}^{r} (p_{k1}, \ldots, p_{kd_k}) \sum\limits_{j=k}^{m+k-1} (A_{kj} + uB_{kj})\, \zeta_j$. Let

$$M_G = \begin{pmatrix} B_{11} & B_{12} & \ldots & B_{1m} & 0 & 0 & \ldots \\ 0 & B_{22} & \ldots & B_{2m} & B_{2,m+1} & 0 & \ldots \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \end{pmatrix}. \tag{26}$$

By $(\tilde{\Lambda})$ we denote the infinite system associated (by Carleman embedding) to the system (Λ).

Remark 6. $(\tilde{\Lambda})$ has the form:

$$\frac{dz}{dt} = (M_F + uM_G)z \tag{27}$$
$$y = Cz, \tag{28}$$

where $z \in \mathbb{R}^\infty, y \in \mathbb{R}^p$. $(\tilde{\Lambda})$ is a bilinear infinite-dimensional system.

Example 4. Let (Λ) be as follows: $\begin{cases} \dot{x} = -x + x^2 + xu \\ y = x^2 \end{cases}$.

Then let for $z_i = x^i, i \in \mathbb{N}$ we have $\frac{dz_i}{dt} = -iz_i + iz_{i+1} + iz_i u$.

Hence $(\tilde{\Lambda})$: $\frac{d}{dt} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \end{pmatrix} =$

$$= \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 \dots \\ 0 & -2 & 2 & 0 & 0 & 0 \dots \\ 0 & 0 & -3 & 3 & 0 & 0 \dots \\ \vdots & \vdots & & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \end{pmatrix} + +u \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \dots \\ 0 & 2 & 0 & 0 & 0 & 0 \dots \\ 0 & 0 & 3 & 0 & 0 & 0 \dots \\ \vdots & \vdots & & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \end{pmatrix},$$

$y = (0, 1, 0, \dots)z.$

Proposition 7.
Let $\Gamma(C, M_F, M_G)_k =$ $col[C, CM_F, CM_G, CM_F^2, CM_F M_G, CM_G M_F, CM_G^2, \dots, CM_G^k]$. *System* (Λ) *is linearly observable iff there is* $k \in \mathbb{N} \cup \{0\}$:

$$\operatorname{rank} \Gamma(C, M_F, M_G)_k = \operatorname{rank} \begin{pmatrix} \Gamma(C, M_F, M_G)_k \\ E_1^T \\ \vdots \\ E_n^T \end{pmatrix}. \tag{29}$$

Proof. Let us observe that similarly as in the proof of Proposition 6 we conclude the thesis by the fact that the map α is bijective and $\alpha(D_0^l D_u^k h_j)(z) = C_j M_F^l M_G^k z$.

Example 5. Let $(\Lambda) : \dot{x}_1 = -x_2 u,\ \dot{x}_2 = x_1 u,\ y = x_1^2 + x_2^2$. (Λ) is not observable. Let $z_1 = x_1, z_2 = x_2, z_3 = x_1^2, z_4 = x_1 x_2, z_5 = x_2^2$. Then from the truncation of the order $s = 5$: $\frac{d}{dt} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 2 & 0 \end{pmatrix} \tilde{z} u,$
$y = \begin{pmatrix} 0 & 0 & 1 & 0 & 1 \end{pmatrix} \tilde{z},\ \tilde{z} = (z_1, z_2, z_3, z_4, z_5)^T$ of the corresponding system $(\tilde{\Lambda})$ we can establish that $CM_F = CM_G = 0$. Hence the equation (29) is not satisfied.

Example 6. Let $(\Lambda) : \dot{x} = x^2 + xu, y = -x + x^2$. Then $(\tilde{\Lambda})$:
$\dot{z} = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 2 & 0 & \dots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix} z + \begin{pmatrix} 1 & 0 & \dots \\ 0 & 2 & \dots \\ & \ddots & \ddots \end{pmatrix} zu, y = (-1, 1, 0, \dots)z$. Hence it is enough to take $k = 1$ to have the equality (29) true.

Acknowledgment

This work was supported by the Bialystok Technical University grants:
Dorota Mozyrska: W/WI/7/07
Zbigniew Bartosiewicz: W/WI/1/07.

References

1. S. Banach, *Théorie des opérations linéaires*, Warsaw 1932.
2. Z. Bartosiewicz, D. Mozyrska, Observability of infinite–dimensional finitely presented discrete–time linear systems, *Zeszyty Naukowe Politechniki Białostockiej. Matematyka-Fizyka-Chemia* 20 (2001), 5–14.
3. Z. Bartosiewicz, D. Mozyrska, Observability of row–finite countable systems of linear differential equations, Proceedings of 16th IFAC Congress, 4-8 July 2005, Prague
4. G. Chen, J.D. Dora, Rational normal form for dynamical systems by Carleman linearization, Proceedings of the 1999 International Symposium on Symbolic and Algebraic Computation, Vancouver, British Columbia, Canada, 165 – 172, 1999.
5. R.F. Curtain, A.J. Pritchard, *Infinite Dimensional Linear Systems Theory*, Springer-Verlag, Berlin 1978.
6. R. Deimling, *Ordinary Differential Equations in Banach Spaces*, Lecture Notes in Mathematics, vol. 596, Springer-Verlag 1977.
7. D.L. Elliott, Bilinear Systems, Encyclopedia of Electrical Engineering, edited by John Webster; J. Wiley and Sons, 1999.
8. P. Sen, On the choice of input for observability in bilinear systems, IEEE Transactions on Automatic Control, vol. AC-26, no. 2, April 1981.
9. G. Herzog, On Lipschitz conditions for ordinary differential equations in Fréchet spaces, Czech. Math. J. 48 (1998), 95–103.
10. B. Jakubczyk, Remarks on equivalence and linearization of nonlinear systems, in: Proc. Nonlinear Control Systems Design Symposium, Bordeaux, France, 1992.
11. K. Kowalski, W.H. Steeb, *Nonlinear dynamical systems and Carleman linearization*, World Scientific Publishing Co. Pte. Ltd., Singapore 1991.
12. R. Lemmert, On ordinary differential equations in locally convex spaces, *Nonlinear Anal.* 10 (1986), 1385–1390.
13. K. Moszyński, A. Pokrzywa, Sur les systémes infinis d'équations différentielles ordinaires dans certain espaces de Fréchet, *Dissert. Math.* 115 (1974).
14. D. Mozyrska, Z. Bartosiewicz, Local observability of systems on $\mathbb{R}^\infty$, in: Proceedings of MTNS'2000, Perpignan, France, 2000.
15. D. Mozyrska, Local observability of infinitely-dimensional finitely presented dynamical systems with output (in Polish), Ph.D. thesis, Technical University of Warsaw, Poland, 2000.
16. D. Mozyrska, Z. Bartosiewicz, Dualities for linear control differential systems with infinite matrices, Control & Cybernetics vol. 35, 2006.
17. Y. Zhou, C. Martin, Carleman linearization of linear systems with polynomial output, Report 2002/2003.