

# On the realization of nonlinear discrete-time systems

S. MONACO

*Dipartimento di Informatica e Sistemistica, University of Rome 'La Sapienza', Via Eudossiana 18, 00184 Rome, Italy*

D. NORMAND-CYROT

*Laboratoire des Signaux et Systèmes, CNRS-ESE, Plateau du Moulon, 91190 Gif-sur-Yvette, France*

Received 5 December 1983

Revised 13 August 1984

It is shown, on the basis of a compact expression of the kernels of the discrete Volterra series associated to a linear analytic discrete-time system, that the kernels satisfy a suitable property which enables their inductive characterization. We also give a first result on the realization of a given family of functions by a linear analytic discrete-time system with linear invertible drift term.

*Keywords:* Nonlinear discrete-time systems, Realization, Volterra kernels.

## 1. Introduction

The characterization of the structure of the Volterra kernels in the input–output response of a nonlinear continuous-time system was first given in [1]. On this basis it was pointed out in [2] that a suitable property of the kernels (which includes separation into functions of time characteristic of linear systems) enables to state a result on realization of a given family of kernels by means of a linear analytic control system. The property of the kernels there involved can be considered as a straightforward extension of the well known property of separability for the kernels of linear and bilinear systems.

If discrete-time linear and bilinear systems are considered it is well known that such a property of separability of the kernels holds as well [3]; this fact enables one to obtain results in the field of the realization problem which can be considered as translations of the continuous-time ones. From these considerations one might expect to get results in the nonlinear discrete-time realization theory which parallel the continuous-time ones; unfortunately, as will be clarified in the sequel, this is not the case when dealing with more general classes of discrete-time nonlinear systems.

In this paper we give first a characterization of the structure of the kernels which appear in the development of the input–output maps of a linear analytic discrete-time system; then we show that the kernels satisfy a suitable property which enables their inductive characterization. This property, which is at the basis of realization problems, is discussed in order to stress the analogies and differences with the equivalent property of the continuous-time case. Finally, we state a first result on a given family of kernels by a linear analytic discrete-time system with linear invertible drift term.

## 2. Input–output maps and Volterra kernels

In what follows we will consider a nonlinear discrete-time system of the form

$$\begin{aligned}x(t+1) &= x(t) + f(x(t)) + g(x(t))u(t), \\y(t) &= h(x(t)), \quad x(0) = x_0 \in \mathbb{R}^n,\end{aligned}\tag{1}$$

where  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $g: \mathbb{R}^n \rightarrow \mathbb{R}^n$  and  $h: \mathbb{R}^n \rightarrow \mathbb{R}$  are analytic functions on all  $\mathbb{R}^n$ ;  $u(\cdot) \in \mathbb{R}$ ,  $y(\cdot) \in \mathbb{R}$ . The single-input-single-output assumption is made for simplicity of notations only. The input-output response of (1) can be derived as follows. If  $\sigma: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $x \mapsto \sigma(x)$ , denotes any function, which will be assumed to be analytic on all  $\mathbb{R}^n$ , let us introduce the following differential operators:

$$L_\sigma^{\otimes 0} = \text{Id}, \quad L_\sigma^{\otimes k} \triangleq \sum_{i_1, \dots, i_k=1}^n \sigma_{i_1} \cdots \sigma_{i_k} \frac{\partial^{(k)}}{\partial x_{i_1} \cdots \partial x_{i_k}}, \quad k \geq 1, \quad (2)$$

$$\Delta_\sigma \triangleq \sum_{k \geq 0} \frac{1}{k!} L_\sigma^{\otimes k}, \quad (3)$$

where  $\sigma_j$ ,  $j = 1, \dots, n$ , and Id denote the  $j$ -th component of  $\sigma$  and the identity operator respectively. It follows from (2) and (3) that for any choice of  $k$  analytic functions  $\tau_i: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $i = 1, \dots, k$ , such that  $\sigma = \tau_1 + \tau_2 + \cdots + \tau_k$ , one has

$$\Delta_\sigma = \Delta_{\tau_1 + \dots + \tau_k} = \Delta_{\tau_1} \otimes \Delta_{\tau_2} \otimes \cdots \otimes \Delta_{\tau_k}; \quad (4)$$

moreover, denoting by  $\delta_i$  any differential operator of the form either (2) or (3),

$$\delta_1 \otimes (\delta_2 + \delta_3) = \delta_1 \otimes \delta_2 + \delta_1 \otimes \delta_3. \quad (5)$$

In particular in (4),

$$\Delta_{\tau_1} \otimes \Delta_{\tau_2} = \sum_{k_1, k_2 \geq 0} \frac{1}{k_1! k_2!} L_{\tau_1}^{\otimes k_1} \otimes L_{\tau_2}^{\otimes k_2}.$$

The differential operator defined by (3) can be used to express in a suitable manner the composition of functions. Namely if  $\lambda: \mathbb{R}^n \rightarrow \mathbb{R}$  is any analytic function on  $\mathbb{R}^n$  it is a matter of computation to verify that the composition, denoted by ' $\circ$ ', of  $\lambda$  with  $I + \sigma$ ,  $I$  the identity function, can be expressed as

$$\lambda \circ (I + \sigma)(x) = \lambda(x + \sigma(x)) = \Delta_\sigma \lambda|_x \quad (6)$$

where  $|_x$  denotes the evaluation at  $x$ . More in general, if  $\sigma_i: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $i = 1, \dots, k$ , are analytic functions on  $\mathbb{R}^n$ , the iterated composition of  $\lambda$  with  $I + \sigma_1, \dots, I + \sigma_k$  can be expressed by means of the iterated application of operators of the form (3) as

$$\begin{aligned} \lambda \circ (I + \sigma_k) \circ \cdots \circ (I + \sigma_1)(x) &= \lambda \circ (I + \sigma_k) \circ \cdots \circ (I + \sigma_2)(x + \sigma_1(x)) \\ &= \Delta_{\sigma_1} \lambda \circ (I + \sigma_k) \circ \cdots \circ (I + \sigma_2)|_{x + \sigma_1(x)} = \Delta_{\sigma_1} \circ \cdots \circ \Delta_{\sigma_k} \lambda|_x. \end{aligned} \quad (7)$$

A property to be taken into account follows from (6) and (7):

$$\lambda \circ (I + \sigma_k) \circ \cdots \circ (I + \sigma_2)(x + \sigma_1(x)) = \Delta_{\sigma_2} \circ \cdots \circ \Delta_{\sigma_k} \lambda|_{x + \sigma_1(x)}. \quad (8)$$

The tools introduced are suitable to represent the evolutions of the system (1) which involve compositions of functions. The state and output evolutions in one step can be obtained from (6), taking into account (4) and (5). The replacement of  $\sigma$  with  $f + u(t)g$ , and  $\lambda$  with  $h$  and  $I$  in (6) enables us to write

$$\begin{aligned} x(t+1) &= \Delta_{f+u(t)g} I|_{x(t)} = \Delta_f \otimes \left( \sum_{n \geq 0} \frac{u(t)^n}{n!} L_g^{\otimes n} I \right) |_{x(t)} = \sum_{n \geq 0} \frac{u(t)^n}{n!} \Delta_f \otimes L_g^{\otimes n} I|_{x(t)}, \\ y(t+1) &= \Delta_{f+u(t)g} h|_{x(t)} = \Delta_f \otimes \left( \sum_{n \geq 0} \frac{u(t)^n}{n!} L_g^{\otimes n} h \right) |_{x(t)} = \sum_{n \geq 0} \frac{u(t)^n}{n!} \Delta_f \otimes L_g^{\otimes n} h|_{x(t)}. \end{aligned}$$

The output at time  $t$  depending on  $x_0$ , assumed at time  $t_0 = 0$ , and the input sequence  $u(0), \dots, u(t-1)$

can now be expressed by means of (7) taking into account (5), (8):

$$\begin{aligned} y(t) &= \Delta_{f+u(0)g} \circ \cdots \circ \Delta_{f+u(t-1)g} h|_{x_0} \\ &= \sum_{n_0, \dots, n_{t-1} \geq 0} \frac{u(0)^{n_0} \cdots u(t-1)^{n_{t-1}}}{n_0! \cdots n_{t-1}!} \Delta_f \otimes L_g^{\otimes n_0} \circ \cdots \circ \Delta_f \otimes L_g^{\otimes n_{t-1}} h|_{x_0}. \end{aligned} \quad (9)$$

The input-output development (9) stresses the dependence of the output at time  $t$  on all the possible products of inputs until time  $t-1$ . This expression can be suitably used to stress the dependence of the output on the products of inputs at a fixed degree, say  $k$ . Denoting by  $w_k(t; \tau_1, \dots, \tau_k; x_0)$  such a dependence, with  $t > \tau_1 \geq \tau_2 \geq \cdots \geq \tau_k \geq 0$ , the output at time  $t$  can be expressed as

$$y(t) = w_0(t; x_0) + \sum_{k=1}^{\infty} \sum_{\tau_1 \geq \cdots \geq \tau_k}^{t-1} w_k(t; \tau_1, \dots, \tau_k; x_0) u(\tau_1) \cdots u(\tau_k) \quad (10)$$

which is the equivalent, with reference to the discrete-time system (1), of the Volterra series development for a linear analytic continuous-time control system. Because of this analogy  $w_k$  will be called 'kernel'. The expression of each kernel in terms of the differential operators (2) and (3) can be deduced by equating (9) and (10); by computation one has

$$\begin{aligned} k=0; \quad w_0(t; x_0) &= \Delta_f^t h|_{x_0}, \\ k=1; \quad w_1(t; \tau_1; x_0) &= \Delta_f^t \circ \Delta_f \otimes L_g \circ \Delta_f^{t-\tau_1-1} h|_{x_0}, \\ k=2; \quad w_2(t; \tau_1, \tau_2; x_0) &= \begin{cases} \Delta_f^{\tau_2} \cdot \Delta_f \otimes L_g \circ \Delta_f^{\tau_1-\tau_2-1} \circ \Delta_f \otimes L_g \circ \Delta_f^{t-\tau_1-1} h|_{x_0}, & \tau_1 > \tau_2, \\ \Delta_f^{\tau_1} \circ \Delta_f \otimes \frac{L_g^{\otimes 2}}{2!} \cdot \Delta_f^{t-\tau_1-1} h|_{x_0}, & \tau_1 = \tau_2. \end{cases} \end{aligned}$$

In order to set a compact expression for the kernel of order  $k$  let us denote by  $l_i$ ,  $i=0, 1, \dots, r$ , any increasing sequence of indices of length at most equal to  $k+1$  such that:

$$(i) \quad 0 = l_0 < l_1 < \cdots < l_r = k,$$

$$(ii) \quad \tau_\alpha = \tau_\beta \quad \forall \alpha, \beta \text{ s.t. } l_{\alpha-1} < \alpha, \beta < l_s, s \in \{1, \dots, r\};$$

moreover let  $\sigma_i$  be defined by  $\sigma_i = l_i - l_{i-1}$  ( $i=1, \dots, r$ ) and  $\sigma_0 = 1$ .

The typical kernel of order  $k$ , which specifies the dependence of the output at time  $t$  on the product of inputs  $u(\tau_1)^{\sigma_1} \cdot u(\tau_{l_1+1})^{\sigma_2} \cdots u(\tau_{l_{r-1}+1})^{\sigma_r}$  (product of inputs of degree  $k$  with  $\sum_{i=1}^r \sigma_i = l_r = k$ ), can be expressed as

$$w_k(t; \tau_1, \dots, \tau_k; x_0) = \Delta_f^{t-r+1} \circ \left[ \Delta_f \otimes \frac{1}{\sigma_s!} L_g^{\otimes \sigma_s} \circ \Delta_f^{\tau_{l_{s-1}} - \sigma_{s-1} + 1 - \tau_{l_{s-1}+1} + 1} \right]_{s=r}^1 \circ h|_{x_0} \quad (11)$$

where  $[\cdot]_{s=r}^1$  denotes the iterated application, from  $r$  until  $1$ , of the operator in parentheses.

The previous formula includes the well known expression of the kernels of linear and bilinear systems. In particular if  $I+f$ ,  $g$  and  $h$  are linear in  $x$  (bilinear system),

$$\Delta_f \otimes L_g^{\otimes n} \circ \Delta_f^k h|_{x_0} = 0, \quad n > 1, \quad \forall k, \quad (12)$$

hence  $w_k = 0$  for any choice of  $l_1, \dots, l_k$  of length less than  $k$  or, what is the same, the output is never affected by products of inputs at the same time instant [3]; the expression of the kernels can be easily deduced by performing the computations. When  $I+f$  and  $h$  are linear in  $x$  and  $g$  is a constant (linear system), (12) holds true; moreover any term of (11) corresponding to  $s > 1$  does not appear because

$$\Delta_f \otimes L_g \circ \Delta_f \otimes L_g \circ \Delta_f^k h|_{x_0} = 0 \quad \forall k.$$

Hence  $w_k = 0$  for any  $k > 1$  or, what is the same, the output is never affected by products of inputs.

**Remark.** It follows from the previous considerations that there are several kernels of order  $k$ , the expression of each of them depending on the sequence  $\sigma_1, \dots, \sigma_r$  where  $\sigma_i, i = 1, \dots, r$ , denotes the power of the inputs at the same time and  $\sigma_1 + \dots + \sigma_r = k$ . This fact, which does not occur in the continuous-time case, is one of the main reasons of the differences between discrete-time and continuous-time systems. When a discrete-time bilinear system is considered, because of the linearity in  $x$  there exists only one kernel of order  $k$  (associated to the sequence of inputs  $u(0) \dots u(k-1)$ ), which makes the situation similar to the continuous-time case.

### 3. Some basic results on the realization

With the positions stated previously for the indices  $l_i$  and  $\sigma_i$ , we can now give the following result.

**Proposition 1.** Assuming  $I + f$  in (1) to be invertible, the  $k$ -th kernel (11) associated to the sequence of indices  $l_1, \dots, l_r$  can be computed, from the  $(k - \sigma_r)$ -th kernel associated to the indices  $l_0, \dots, l_{r-1}$ , by

$$w_k(t; \tau_1, \dots, \tau_k; x_0) = \sum_{\substack{n_1, \dots, n_{\sigma_r} \geq 0 \\ n_1 + \dots + n_{\sigma_r} = \sigma_r}} \frac{1}{n_1! \dots n_{\sigma_r}!} \cdot L_{P_1(\tau_{l_{r-1}+1}, x_0)}^{\otimes n_1} \otimes \dots \otimes L_{P_{\sigma_r}(\tau_{l_{r-1}+1}, x_0)}^{\otimes n_{\sigma_r}} \circ w_{k-\sigma_r}(t; \tau_1, \dots, \tau_{k-\sigma_r}; \cdot)|_{x_0} \quad (13)$$

for a suitable choice of the analytic functions  $P_i: \mathbb{N} \times \mathbb{R}^n \rightarrow \mathbb{R}^n, i = 1, \dots, \sigma_r$ .

**Proof.** Because of the invertibility of the function  $I + f$  the differential operator  $\Delta_f^{-\tau}$  is well defined for any  $\tau \geq 0$ , i.e. the operator which has the property

$$\Delta_f^\tau \circ \Delta_f^{-\tau} = \Delta_f^{-\tau} \circ \Delta_f^\tau = \text{Id.}$$

Let us denote by  $P_j(\tau), j \geq 1, \tau \geq 0$ , the differential operator defined by

$$P_j(\tau) = \frac{1}{j!} \Delta_f^\tau \circ \Delta_f \otimes L_g^{\otimes j} \circ \Delta_f^{-1} \circ \Delta_f^{-\tau}. \quad (14)$$

The typical kernel (11) can now be expressed by

$$w_k(t; \tau_1, \dots, \tau_k; x_0) = \left[ \Delta_f^{\tau_{l_{r-1}+1}} \circ \Delta_f \otimes \frac{1}{\sigma_r!} L_g^{\otimes \sigma_r} \circ \Delta_f^{-1} \circ \Delta_f^{-\tau_{l_{r-1}+1}} \right]_{s=r}^1 \circ \Delta_f' h|_{x_0} \\ = P_{\sigma_r}(\tau_{l_{r-1}+1}) \circ \dots \circ P_{\sigma_1}(\tau_1) \circ \Delta_f' h|_{x_0}. \quad (15)$$

This formula enables one to express the typical kernel of order  $k$  associated to the indices  $l_1, \dots, l_r$  by means of the application of the operator  $P_{\sigma_r}(\tau_{l_{r-1}+1})$  to the function  $w_{k-\sigma_r}(t; \tau_1, \dots; \cdot)$  which yields, at  $x_0$ , the kernel of order  $k - \sigma_r$  associated to the indices  $l_1, \dots, l_{r-1}$ , i.e.

$$w_k(t; \tau_1, \dots, \tau_k; x_0) = P_{\sigma_r}(\tau_{l_{r-1}+1}) w_{k-\sigma_r}(t; \tau_1, \dots, \tau_{l_{r-2}}; \cdot)|_{x_0}. \quad (16)$$

A remarkable property of the operator (14) is that

$$P_j(\tau) = \sum_{\substack{n_1, \dots, n_j > 0 \\ n_1 + 2n_2 + \dots + jn_j = j}} \frac{1}{n_1! \dots n_j!} L_{P_1(\tau)(x)}^{\otimes n_1} \otimes \dots \otimes L_{P_j(\tau)(x)}^{\otimes n_j} \quad (17)$$

where  $P_i(\tau)(x)$  denotes the function obtained by applying the differential operator  $P_i(\tau)$  to the identity function and  $L_{P_1(\tau)(x)}^{\otimes n_1} \otimes \dots \otimes L_{P_j(\tau)(x)}^{\otimes n_j}$  is the differential operator defined by (2).

To show that (17) holds true, let us recall that given any analytic function  $\lambda: \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $\Delta_f \otimes L_g^{\otimes j} \lambda|_x$  represents the coefficient of power  $j$  in  $u$  in the Taylor series of the composite function  $\lambda \circ (I + f + gu)$  starting from the point  $(I + f)(x)$ , i.e. symbolically

$$\Delta_f \otimes L_g^{\otimes j} \lambda|_{x_0} = \frac{\delta^{(j)}(\lambda \circ (I + f + gu))}{\partial u^{(j)}} \Big|_{\substack{x=x_0 \\ u=0}} = \frac{\partial^{(j)} \lambda}{\partial x^{(j)}} \Big|_{(I+f)(x_0)} \times g^j|_{x_0}$$

where the bar denotes the evaluation at the point  $(x_0, 0)$  and  $(I + f)(x_0)$  respectively, and

$$\frac{\partial^{(j)} \lambda}{\partial x^{(j)}} \Big|_{(I+f)(x_0)} \times g^j|_{x_0} = \sum_{i_1, \dots, i_j=1}^n \frac{\partial^{(j)}}{\partial x_{i_1} \dots \partial x_{i_j}} \Big|_{(I+f)(x_0)} \times g_{i_1}(x_0) \times \dots \times g_{i_j}(x_0).$$

Taking into account (7) and (8) it is easily seen that

$$\Delta_f^\tau \circ \Delta_f \otimes L_g^{\otimes j} \circ \Delta_f^{-\tau-1} \circ \lambda|_{x_0} = \frac{\partial^{(j)}(\lambda \circ (I + f)^{\tau+1})}{\partial x^{(j)}} \Big|_{(I+f)^{\tau+1}(x_0)} \times g^j|_{(I+f)^\tau(x_0)} \tag{18}$$

where  $(I + f)^\tau$  denotes the  $\tau$ -times composition of the function  $I + f$ .

By performing the computations on the right-hand side of (18) it is not difficult to verify the following expression:

$$\begin{aligned} \Delta_f^\tau \circ \Delta_f \otimes L_g^{\otimes j} \Delta_f^{-\tau-1} \lambda|_{x_0} = & \sum_{\substack{n_1, \dots, n_j > 0 \\ n_1 + 2n_2 + \dots + jn_j = j}} \frac{j!}{n_1! \dots n_j! (1!)^{n_1} \dots (j!)^{n_j}} \\ & \cdot L_{\tilde{P}_1(\tau)(x_0)}^{\otimes n_1} \otimes \dots \otimes L_{\tilde{P}_j(\tau)(x_0)}^{\otimes n_j} \circ \lambda|_{x_0} \end{aligned} \tag{19}$$

where

$$\tilde{P}_r(\tau)(x_0) = \Delta_f^\tau \Delta_f \otimes L_g^{\otimes r} \Delta_f^{-\tau-1} I|_{x_0}, \quad r = 1, \dots, j.$$

By substituting in (19)

$$\tilde{P}_r(\tau)(x_0) = r! P_r(\tau)(x_0) = r! P_r(\tau, x_0), \tag{20}$$

(19), (17) and (16) show that (13) holds true.  $\square$

From (13) and (15), with the stated positions for the indices, we have

$$w_0(t; x) = \Delta_f^t h|_x, \tag{21}$$

$$\begin{aligned} w_k(t; \tau_1, \dots, \tau_k; x) = & \left[ \sum_{\substack{n_1, \dots, n_{\sigma_k} > 0 \\ n_1 + \dots + \sigma_k n_{\sigma_k} = \sigma_k}} \frac{1}{n_1! \dots n_{\sigma_k}!} L_{P_1(\tau_{i-1}, x)}^{\otimes n_1} \otimes \dots \otimes L_{P_{\sigma_k}(\tau_{i-1}, x)}^{\otimes n_{\sigma_k}} \right]_{s=r}^1 \circ \Delta_f^t h|_x, \\ & k \geq 1. \end{aligned} \tag{22}$$

Proposition 1 states the announced property of the kernels of a linear analytic control system which has invertible drift term. Moreover for  $\sigma_i = 1, i = 1, \dots, k$  (i.e.  $\tau_1 > \tau_2 > \dots > \tau_k$ ), (13) gives

$$\begin{aligned}
 w_k(t; \tau_1, \dots, \tau_k; x_0) &= L_{P_1(t, x_0)} \circ w_{k-1}(t; \tau_1, \dots, \tau_{k-1}; \cdot)|_{x_0} \\
 &= \Delta_f^{\tau_k} \circ L_{P_1(0, x)} \circ \Delta_{\bar{f}}^{-\tau_k} w_{k-1}(t; \tau_1, \dots, \tau_{k-1}; \cdot)|_{x_0} \quad \text{by (17)} \\
 &= \Delta_f^{\tau_k} \cdot L_{P_1(0, x)} w_{k-1}(t; \tau_1, \dots, \tau_{k-1}; \Delta_{\bar{f}}^{-\tau_k} x)|_{x_0} \\
 &= \frac{\partial w_{k-1}(t; \tau_1, \dots; \Delta_{\bar{f}}^{-\tau_k}(x))}{\partial x} \Big|_{(I+f)^{\tau_k}(x_0)} \times P_1(0, x) \Big|_{(I+f)^{\tau_k}(x_0)} \quad (23)
 \end{aligned}$$

(23) points out the strict analogy with the formula which expresses the link between two successive kernels of a linear analytic continuous-time system. To be more precise, denoting by  $\bar{f}$  and  $\bar{g}$  the vector fields which define such a system, the link between the kernels  $\bar{w}_k$  and  $\bar{w}_{k-1}$  is given by

$$\bar{w}_k(t; \tau_1, \dots, \tau_k; x_0) = \frac{\partial \bar{w}_{k-1} \circ \varphi_{-\tau_k}}{\partial x} \Big|_{\varphi_{\tau_k}(x_0)} \times \bar{g}(\varphi_{\tau_k}(x_0)) \quad (24)$$

where  $\varphi_{\tau} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  denotes the flow associated to  $\bar{f}$ . The analogy of (23) with (24) is now evident taking into account that  $(I+f)^{\tau} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  denotes the ‘flow’ associated to  $I+f$ . A difference which must be noted is that instead of  $g(x)$ ,  $P_1(0)(x)$  is present in (23); this correspondence turns out to be meaningful for a better understanding of the links between nonlinear continuous-time and discrete-time systems.

Let us now consider the usual nonlinear realization problem.

**Definition.** A family of functions  $\{\lambda_i(t; \tau_1, \dots, \tau_i)\}_{i=0}^{\infty}$ ,  $t \geq \tau_1 \geq \dots \geq \tau_i$ ,  $t$  and  $\tau_i$  belonging to  $\mathbb{N}$ , has a linear analytic discrete-time realization if there exists a system of the form (1) such that, denoting by  $w_k$  its kernels,

$$w_k(t; \tau_1, \dots, \tau_k; x_0) = \lambda_k(t; \tau_1, \dots, \tau_k) \quad \forall t \geq \tau_1 \geq \dots \geq \tau_k.$$

The quadrupule  $\{f, g, h, x_0\}$  is called a realization of the given family of functions.

The property stated in Proposition 1 must obviously be taken into account when dealing with a realization problem of a family of functions  $\{\lambda_i\}_{i=0}^{\infty}$ . Unfortunately in (13) the link between  $w_k$  and the previous ones depends on the functions  $P_i$ ,  $i \geq 1$ . This would make it possible to formulate the problem of realization by means of a linear analytic system in terms of the existence of a family of functions  $P_i : \mathbb{N} \times \mathbb{R}^n \rightarrow \mathbb{R}$  and  $Q : \mathbb{N} \times \mathbb{R}^n \rightarrow \mathbb{R}$  which satisfy suitable properties. Such a result does not seem to provide much insight into the problem itself. A nice result can be obtained if one restricts oneself to find a realization which has a linear invertible drift term; with  $f(x)$  in (1) given by

$$x + f(x) = Ax + d, \quad A \in \mathbb{R}^{n \times n}, \quad d \in \mathbb{R}^{n \times 1}, \quad (25)$$

with  $A$  invertible.

Under the position assumed for the indices  $l_i$  and  $\sigma_i$  in the previous section, we have:

**Theorem 1.** A family of functions  $\{\lambda_i(t; \tau_1, \dots, \tau_i)\}_{i=0}^{\infty}$  has a linear analytic realization of the form (1) with a linear invertible drift term if and only if there exist an integer  $n$ , two functions

$$P : \mathbb{N} \times \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad Q : \mathbb{N} \times \mathbb{R}^n \rightarrow \mathbb{R},$$

and an  $x_0 \in \mathbb{R}^n$  such that:

$$\text{(i) } \lambda_0(t) = Q(t, x_0) \quad \forall t,$$

$$\text{(ii) } \lambda_k(t; \tau_1, \dots, \tau_k) = \left[ \frac{1}{\sigma_s!} L_{P(\tau_{i-1}, \dots, \tau_1, x_0)}^{\otimes \sigma_s} \right]_{s=r}^1 \circ Q(t, x) \Big|_{x_0} \quad \forall t \geq \tau_1 \geq \dots \geq \tau_k, \quad \forall k \geq 1.$$

**Proof. Necessity:** Because of the linearity of  $I + f$ , the functions  $P_i(\tau, x)$  defined by (20) and (14) are identically zero for  $i > 1$ :

$$P_i(\tau, x) = 0, \quad i > 1.$$

Denoting

$$Q(t, x) := \Delta'_f h|_x \quad \forall t, \quad P(\tau, x) := P_1(\tau, x) \quad \forall \tau,$$

it is immediately verified by means of (22) that (i) and (ii) hold true.

*Sufficiency:* Let us define the discrete-time system

$$\begin{aligned} z(t+1) &= z + F(z) + G(z)u(t), \\ y(z) &= H(z), \quad z(0) = z_0, \end{aligned} \tag{26}$$

with

$$\begin{aligned} z &= (x_1, \dots, x_n, i)', \quad z_0(x'_0, 0)' \\ F &= (0, \dots, 0, 1), \quad G(z) = (p_1(z), \dots, p_n(z), 0)', \quad H(z) = Q(z), \end{aligned}$$

where ' denotes transposition and  $p_j(z)$ ,  $j = 1, \dots, n$ , are the components of the vector-valued function  $P$ . We will show that the Volterra kernels of the system (26) coincide with the functions on the right-hand side of (i) and (ii); i.e. the system (26) initialized at  $(x'_0, 0)'$  provides a realisation of the given family of functions. For, denoting by  $W_k$ ,  $k \geq 0$ , the Volterra kernels of the system (26), one has

$$W_0(t; z_0) = \Delta'_f H|_{z_0} = H \circ (I + F)'(z_0) = H((x'_0, t)') = Q(t, x_0) = \lambda_0(t).$$

As far as the other kernels are concerned, with the notations introduced in the previous section for the indices, we have

$$W_k(t; \tau_1, \dots, \tau_k; z_0) = \left[ \Delta_{F'}^{\tau_{k-1}+1} \Delta_{F'} \otimes \frac{L_G^{\otimes k}}{\sigma_s!} \circ \Delta_{F'}^{-1} \circ \Delta_{F'}^{-\tau_{k-1}-1} \right]_{s=r}^1 \circ \Delta'_f H|_{z_0}.$$

The proof will be achieved as soon as we show that the following equalities hold true:

$$\Delta_F^\tau \Delta_F \otimes L_G^{\otimes k} \Delta_F^{-1} \Delta_F^{-\tau} = L_{P(\tau, x)}^{\otimes k} \quad \forall k \geq 1. \tag{27}$$

For this purpose let us recall that from (21), (14) and (17), because of the structure of  $F$  (which is a constant in this case) it results that

$$\Delta_F^\tau \Delta_F \otimes L_G^{\otimes k} \Delta_F^{-1} \Delta_F^{-\tau} = L_{\Delta_F^\tau \Delta_F \otimes L_G \Delta_F^{-1} \Delta_F^{-\tau}(z)}$$

Hence (27) holds true if and only if

$$\Delta_F^\tau \Delta_F \otimes L_G \Delta_F^{-1} \Delta_F^{-\tau}(z) = P(\tau, x).$$

Recalling that (see (18))

$$\Delta_F^\tau \Delta_F \otimes L_G \Delta_F^{-\tau-1}(z) = \left. \frac{\partial(I+F)^{-\tau-1}}{\partial z} \right|_{(I+F)^{\tau+1}(z)} \times G|_{(I+F)^\tau(z)} \tag{28}$$

and that, because of its structure,

$$(I+F)^\tau(z) = (I+F)^\tau(x', i) = (x', i + \tau),$$

we obtain

$$\left. \frac{\partial(I+F)^{-\tau}}{\partial z} \right|_{(I+F)^\tau} = \text{Id}^0 \quad (\text{the identity matrix}) \quad \forall \tau,$$

which substituted into (28) gives

$$\Delta_F^r \Delta_F \otimes L_G \Delta_F^{-r-1} = \text{Id}^0 \cdot G \circ (I + F)^r(z) = \text{Id}^0 \cdot G(x, \tau) = P(x, \tau),$$

which concludes the proof.  $\square$

Particular assumptions on the structure of the function  $P$  of Theorem 1 include known results on the realization of discrete-time linear and bilinear systems [3,4]. In order to clarify this aspect we give the two following corollaries whose proof is straightforward.

**Corollary 1** (bilinear systems). *A family of functions  $\{\lambda_i(t; \tau_1, \dots, \tau_i)\}_{i=0}^\infty$ ,  $t > \tau_1 > \dots > \tau_k$ , has a bilinear realization with invertible drift term if and only if the conditions (i) and (ii) are satisfied with functions  $P$  and  $Q$  linear in  $x$ :*

$$P(t, x) = P(t)x, \quad Q(t, x) = Q(t)x \quad \forall t,$$

such that  $P(t)$  and  $Q(t)$ , ( $n \times n$ ) and ( $1 \times n$ ) matrices, satisfy the conditions

$$(a) \quad Q(t) = Q(t-1)X, \quad (b) \quad P(t) = X^{-1}P(t-1)X \quad \forall t > 0,$$

for a suitable ( $n \times n$ ) invertible matrix  $X$ .

**Corollary 2.** *Two functions  $\lambda_0(t)$ ,  $\lambda_1(t, \tau)$ ,  $t > \tau$ , have a linear realization with invertible drift term if and only if the conditions (i) and (ii) are verified with  $P$  constant and  $Q$  linear with respect to  $x$ :*

$$P(t, x) = P(t), \quad Q(t, x) = Q(t)x \quad \forall t,$$

and  $P(t)$  and  $Q(t)$  ( $n \times 1$ ) and ( $1 \times n$ ) matrices such that

$$(a) \quad Q(t) = Q(t-1)X, \quad (b) \quad P(t) = X^{-1}(P(t-1)) \quad \forall t > 0,$$

for a suitable invertible constant ( $n \times n$ ) matrix  $X$ .

For a linear realizations in the zero state,  $x_0 = 0$ , Corollary 2 includes the well known separability condition:

$$w(t; \tau) = Q(t)P(\tau).$$

## References

- [1] A.J. Krener and C. Lesiak, The existence and uniqueness of Volterra series for nonlinear systems, *IEEE Trans. Automat. Control* **23** (1978) 1090–1095.
- [2] A. Isidori and A. Ruberti, A separation property of realizable Volterra kernels, *Systems Control Lett.* **1** (1982) 309–311.
- [3] A. Isidori, Direct construction of minimal bilinear realizations from nonlinear input–output maps, *IEEE Trans. Automat. Control* **18** (6) (Dec. 1973).
- [4] M. Fliess, Sur la réalisation des systèmes dynamiques bilinéaires, *C.R. Acad. Sci. Paris A.* **277** (1973) 923–926.
- [5] L. Comtet, *Analyse Combinatoire*, tome I (Coll. Sup. Universitaires de France, Paris, 1970).
- [6] D. Normand-Cyrot, Théorie et pratique des systèmes non-linéaires en temps discret, Thèse d'Etat, Université Paris-Sud (1983).
- [7] P.E. Crouch, Polynomial system theory: a review, *IEEE Proc.* **127** (5) (Sept. 1980).
- [8] E.D. Sontag, Polynomial response maps, in: A.V. Balakrishnan and M. Thoma, Eds., *Lecture Notes in Control and Inform. Sci.*, Vol. 13 (Springer, Berlin–New York, 1979).