Coalgebraic Foundations of Linear Systems

(An Exercise in Stream Calculus)

J.J.M.M. Rutten

CWI and Vrije Universiteit Amsterdam

Abstract. Viewing discrete-time causal linear systems as (Mealy) coalgebras, we describe their semantics, minimization and realisation as universal constructions, based on the final coalgebras of streams and causal stream functions.

1 Introduction

Linear systems are a fundamental mathematical structure with applications in control theory, signal processing, and telecommunications. In computer science, they are given but little attention. However, linear systems provide a mathematical model for various types of networks, including signal flow graphs and linear sequential Boolean circuits (see, for instance, [Koh78, Lah98]). Such networks are highly relevant for the foundations of computing, being elementary and beautiful examples of the combined occurrence of memory and feedback.

In this paper, we give a coalgebraic account of the semantics of the following elementary type of linear system: a (state-based) discrete-time (strongly) causal linear system consists of a vector space V of states; vector spaces I and O of inputs and outputs; and linear maps $F : V \to V$, describing the system's dynamics, and $G : I \to V$ and $H : V \to O$, describing the system's input and output. We shall model such a system as a Mealy automaton (V, Φ), defined by

$$\Phi : V \to (O \times V)^I \qquad \Phi(v)(i) = \langle H(v),\, F(v) + G(i)\rangle$$

Such Mealy automata, or (I, O)-*systems* as we shall call them here, are coalgebras of the functor $\mathcal{F} : \mathit{Set} \to \mathit{Set}$ defined by $\mathcal{F}(S) = (O \times S)^I$. The choice to model linear systems as Mealy automata or, in other words, as coalgebras of this particular choice of functor $\mathcal{F}$, is motivated by the following observation: In [Rut06], it is shown that the final coalgebra of $\mathcal{F}$, which is to serve as our semantic universe, is (isomorphic to) the set Γ of all *causal functions* from the set of input streams I^ω to the set of output streams O^ω. In system theory, the input-output behavior of a linear discrete-time causal system is often described in terms of precisely such a causal stream function (traditionally called the *transfer* function of the system).

Note that we work in the category of sets and functions rather than vector spaces and linear maps. Although the functor $\mathcal{F}$ can also be defined on vector spaces, the function Φ defined above will in general not be linear, even if F, G, and H are. However, linearity of these maps *does* play a role in the various

T. Mossakowski et al. (Eds.): CALCO 2007, LNCS 4624, pp. 425–446, 2007.

characterisations of the semantics of linear systems, as we shall see later. (And, of course, vector addition in V is used in the definition of Φ.)

Once the functor (that is, the type of our systems) has been fixed and its final coalgebra identified, a coalgebraic treatment of linear systems follows from general insights of universal coalgebra: the *behaviour* (or semantics) of a system is given by the unique homomorphism into the final coalgebra; the image of this homomorphism constitutes the system's *minimisation*; and systems can be *specified* by elements of the final coalgebra and then *realised* (synthesised) by the corresponding generated subsystems of the final coalgebra.

The exercise mentioned in the title then consists of working out the details of all this. We view the formulation and the carrying out of this exercise as the main contribution of the present paper. Technically, we had to extend our earlier work [Rut03, Rut05] a bit in order to deal with streams of linear transformations, in Section 3. After recalling the coalgebraic treatment of (I, O)-systems, in Section 4, the main technical contribution lies in Section 5. It will be based on the elementary but crucial observation that the function Φ above factors through three maps of the following type (see (21)):

$$\Phi:\ V \longrightarrow V \times V \longrightarrow O \times V \longrightarrow (O \times V)^I$$

This is the basis for Theorem 8, which presents the final behaviour of (V, Φ) as the composition of three corresponding final homomorphisms. This final semantics f assigns to each (initial) state $v \in V$ a causal function $f(v) : I^\omega \to O^\omega$ (called the *transfer function* in system theory). This leads then to characterisations of system minimization and realisation, in Sections 6 and 7. Surprisingly, the final semantics f turns out to be the composition of a (linear) final mapping $H \times \tilde{F} : V \to O^\omega$ followed by a (non-linear) *injection* $g : O^\omega \to \Gamma$. As a consequence, minimization and realisation can be simply described in terms of just output streams, ignoring the presence of input streams altogether.

From the perspective of the theory of coalgebra, the relevance of our contribution consists of the following points. (i) It adds one more basic but important example to the family of mathematical structures that can be treated naturally and fairly completely by coalgebraic means. Other well-known examples are streams, automata, formal power series, infinite data types etc. (ii) Technically, the interaction between algebra and coalgebra is interesting. In general, (I, O)-systems (Mealy automata) live in the category of *sets*. As we shall see, *linear* (I, O)-systems are completely determined by their underlying linear O-systems (in which input plays no role), and these do live in the category of vector spaces and linear maps. As a consequence, the final behaviour of linear (I, O)-systems, which itself is obtained in *Set*, can be pleasantly characterised in terms of the basic operations (of sum and convolution product) of stream calculus. (iii) It also follows that streams – which constitute the prototypical example of a final coalgebra – are essentially all that is needed for the modelling of linear systems, since O-systems can be completely described in terms of O-streams. (iv) The final behaviour of *finite dimensional* linear (I, O)-systems will be characterised in terms of *rational* streams, in essentially the same way as finite deterministic automata, which can be viewed as elementary non-linear (I, O)-systems, correspond to rational (regular) languages. (v) More generally, the present model shows that from a coalgebraic perspective, there is no essential difference between the treatment of linear and non-linear systems. This opens the way for future applications of coalgebraic techniques to non-linear phenomena in system theory.

Some of these points may also be of interest for system theory, where the semantics of the linear systems that we are considering is since long well understood (see, for instance, [Kai80]). In particular, our emphasis on the central role of (the final coalgebra of) streams leads to a very elementary treatment of system realisation, which – depending on taste and background – might be considered as a simpler alternative to Kalman's [Kal63, KFA69] classical construction using Hankel matrices. See the appendix for a further discussion of this.

We mention a few directions for further research. Since the semantics of both linear and non-linear systems is given by finality, it would be interesting to try and fit instances of non-linear systems from system theory (cf. [Son79]) into the coalgebraic framework. Also generalisations to *continuous* systems could be considered. Finally, one of the hallmarks of coalgebra is the notion of bisimulation, or observational equivalence, which comes along with every (functor) type of system. It should therefore be possible to study notions of equivalence for linear systems, including recently introduced ones such as in [Pap03] and [vdS04], from a coalgebraic perspective.

2 Preliminaries

We define the set of *streams* over a given set A by

$$A^\omega = \{\sigma \mid \sigma : \{0, 1, 2, \ldots\} \to A\}$$

We will denote elements $\sigma \in A^\omega$ by $\sigma = (\sigma(0), \sigma(1), \sigma(2), \ldots)$. We define the *stream derivative* of a stream σ by

$$\sigma' = (\sigma(1), \sigma(2), \sigma(3), \ldots)$$

and we call $\sigma(0)$ the *initial value* of σ. For $a \in A$ and $\sigma \in A^\omega$ we use the following notation:

$$a : \sigma = (a, \sigma(0), \sigma(1), \sigma(2), \ldots)$$

For instance, $\sigma = \sigma(0) : \sigma'$, for any $\sigma \in A^\omega$. Any function $f : A \to B$ induces a function

$$f^\omega : A^\omega \to B^\omega \qquad f^\omega(\sigma) = (f(\sigma(0)),\, f(\sigma(1)),\, f(\sigma(2)),\, \ldots) \tag{1}$$

Any function $f : A \to A$ induces a function

$$\tilde{f} : A \to A^\omega \qquad \tilde{f}(a) = (a,\, f(a),\, f^2(a),\, \ldots) \tag{2}$$

where $f^0 = 1$, the identity on A and $f^{n+1} = f \circ f^n$. If V is a set and W is a *vector space* (over some field k) then the set W^V of all functions

$$W^V = \{f \mid f : V \to W\}$$

is a vector space, with addition and scalar multiplication given, for $v \in V$ and $c \in k$, by

$$(f + g)(v) = f(v) + g(v) \qquad (c \cdot f)(v) = c \cdot f(v)$$

In particular, if V is a vector space over k then so is the set V^ω of all streams over V. Both the operations of initial value and derivative are linear transformations: for all $c, d \in k$, $\sigma, \tau \in V^\omega$,

$$(c \cdot \sigma + d \cdot \tau)(0) = c \cdot \sigma(0) + d \cdot \tau(0) \qquad (c \cdot \sigma + d \cdot \tau)' = c \cdot \sigma' + d \cdot \tau'$$

For any set A and $n \geq 1$, we denote the elements $v \in A^n$ by $v = (v_1, \ldots, v_n)$. It will sometimes be convenient to switch between streams of tuples and tuples of streams. We define the *transpose* as follows:

$$(-)^T : (A^n)^\omega \to (A^\omega)^n \qquad (\sigma^T)_i(j) = (\sigma(j))_i \tag{3}$$

This function is an isomorphism and has an inverse which we denote again by

$$(-)^T : (A^\omega)^n \to (A^n)^\omega$$

A *semi-ring* is a set R with a commutative operation of addition $c + d$; a (generally non-commutative) operation of multiplication $c \cdot d$ with $c \cdot (d + e) = (c \cdot d) + (c \cdot e)$ and $(d + e) \cdot c = (d \cdot c) + (e \cdot c)$; and with neutral elements 0 and 1 such that $c + 0 = c$, $1 \cdot c = c \cdot 1 = c$ and $c \cdot 0 = 0 \cdot c = 0$. If every $c \in R$ moreover has an additive inverse $-c$ (with $c + (-c) = 0$) then R is a *ring*.

Any field is a ring. The following example of a ring will be used later. Let V be a vector space (over some field k). The set $V \to_L V$ of linear maps $F : V \to V$ is a ring with addition and multiplication defined by

$$(F + G)(v) = F(v) + G(v) \qquad (F \times G)(v) = F(G(v))$$

and with the everywhere zero map and the identity map as neutral elements 0 and 1.

3 Stream Calculus

Let R be a ring. We define the following operators on the set R^ω of streams over R, for all $c \in R$, $\sigma, \tau \in R^\omega$, $n \geq 0$:

$$\begin{aligned}
[c] &= (c, 0, 0, 0, \ldots) \quad \text{(often simply denoted again by } c\text{)} \\
X &= (0, 1, 0, 0, 0, \ldots) \\
(\sigma + \tau)(n) &= \sigma(n) + \tau(n) \qquad \text{[sum]} \\
(\sigma \times \tau)(n) &= \sum_{i=0}^{n} \sigma(i) \cdot \tau(n - i) \qquad \text{[convolution product]}
\end{aligned}$$

(where $\cdot$ denotes multiplication in the ring R). A stream σ has a (unique) multiplicative inverse σ^{-1} in R^ω:

$$\sigma^{-1} \times \sigma = [1]$$

whenever its initial value $\sigma(0)$ has a multiplicative inverse $\sigma(0)^{-1}$ in R. As usual, we shall often write $1/\sigma$ for σ^{-1} and σ/τ for $\sigma \times \tau^{-1}$. Since $X^2 = (0, 0, 1, 0, 0, 0, \ldots)$, $X^3 = (0, 0, 0, 1, 0, 0, 0, \ldots)$ and so on, the following infinite sum is well defined, for all $\sigma \in R^\omega$:

$$\sigma = \sigma(0) + (\sigma(1) \times X) + (\sigma(2) \times X^2) + \cdots$$

(Note that we write $\sigma(i)$ for $[\sigma(i)]$; similarly below.) It shows that σ can be viewed as a formal power series in the indeterminate X (which here in fact is a constant stream). What distinguishes our approach from formal power series is a systematic use of the operation of stream derivative and the universal property of finality it induces (see Section 4). This leads to a somewhat non-standard algebraic calculus, which we call *stream calculus*. We mention a few identities which are helpful for the computation of stream derivatives. (Computing stream derivatives is crucial in our approach to system *realisation*, in Section 7).

Lemma 1 ([Rut03]). *Let R be a ring. For all $\sigma, \tau \in R^\omega$,*

$$\begin{aligned} (\sigma + \tau)' &= \sigma' + \tau' \\ (\sigma \times \tau)' &= (\sigma' \times \tau) \;+\; (\sigma(0) \times \tau') \\ (\sigma^{-1})' &= -\sigma(0)^{-1} \times \sigma' \times \sigma^{-1} \end{aligned}$$

and $(\sigma + \tau)(0) = \sigma(0) + \tau(0)$, $(\sigma \times \tau)(0) = \sigma(0) \cdot \tau(0)$, and $\sigma^{-1}(0) = \sigma(0)^{-1}$ (if the latter exists). Moreover, $\sigma = \sigma(0) + (X \times \sigma')$ and $X \times \sigma = \sigma \times X$. □

We call a stream *polynomial* if it is of the form

$$c_0 + (c_1 \times X) + (c_2 \times X^2) + \cdots + (c_k \times X^k)$$

A stream is *rational* if it is the quotient $\sigma/\tau = \sigma \times \tau^{-1}$ of two polynomial streams σ and τ for which $\tau(0)^{-1}$ exists. We denote the set of all rational streams over R by

$$Rat(R^\omega) = \{\sigma \in R^\omega \mid \sigma \text{ is rational}\}$$

A prototypical example of a rational stream in R^ω, for $c \in R$, is

$$\frac{1}{1 - (c \times X)} = (1,\, c,\, c^2,\, \ldots)$$

If we consider the ring $V \to_L V$, for a vector space V, then streams $\phi \in (V \to_L V)^\omega$ are infinite sequences $\phi = (\phi(0), \phi(1), \phi(2), \ldots)$ of linear transformations $\phi(i) : V \to V$. For a linear transformation $F \in (V \to_L V)$, the example above becomes

$$\frac{1}{1 - (F \times X)} = (1, F, F^2, \cdots) \tag{4}$$

which, under the isomorphism $(V \to V)^\omega \cong V \to V^\omega$, is equal to $\tilde{F}$ defined in (2) above.

We shall also use the following type of convolution product. Let V and W be vector spaces. For streams $\phi \in (V \to_L W)^\omega$ and $\sigma \in V^\omega$, we define $\phi \times \sigma \in W^\omega$ by

$$(\phi \times \sigma)(n) = \sum_{i=0}^{n} \phi(i) \times \sigma(n-i) \tag{5}$$

where on the right we write $\phi(i) \times \sigma(n-i)$ for $\phi(i)(\sigma(n-i))$. For a linear map $H : V \to W$, we have as a special case

$$[H] \times \sigma = (H, 0, 0, 0, \ldots) \times \sigma = (H(\sigma(0)),\ H(\sigma(1)),\ H(\sigma(2)), \ldots)$$

which equals $H^\omega(\sigma)$ defined in (1) above. Note that if $W = V$, the set of streams $(V \to_L V)^\omega$ has itself also an operation of convolution product, which interacts nicely with the product defined in (5). For example, for $\phi, \psi \in (V \to_L V)^\omega$ and $\sigma \in V^\omega$,

$$(\phi \times \psi) \times \sigma = \phi \times (\psi \times \sigma) \tag{6}$$

Let k be a field. A linear transformation $F : k^n \to k^m$ between *finite dimensional* vector spaces corresponds to an $m \times n$ matrix M_F with values F_{ij} in k:

$$F : k^n \to k^m \qquad\qquad M_F = \begin{pmatrix} F_{11} & F_{12} & \cdots & F_{1n} \\ F_{21} & F_{22} & \cdots & F_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ F_{m1} & F_{m2} & \cdots & F_{mn} \end{pmatrix}$$

Here and in what follows, the matrix is with respect to the standard basis

$$(1, 0, \ldots, 0),\ \ldots,\ (0, \ldots, 0, 1)$$

of k^n and k^m. Any stream $\phi = (\phi(0),\ \phi(1),\ \phi(2),\ \ldots)$ of linear transformations $\phi(i) : k^n \to k^m$ corresponds to a stream of matrices

$$(M_{\phi(0)},\ M_{\phi(1)},\ M_{\phi(2)},\ \ldots) = M_{\phi(0)} + (M_{\phi(1)} \times X) + (M_{\phi(2)} \times X^2) + \cdots$$

If we consider $M_{\phi(i)} \times X^i$ as an $m \times n$ matrix obtained from $M_{\phi(i)}$ by multiplying each of its entries by X^i, then the infinite sum on the right can itself be viewed as an $m \times n$ matrix M_ϕ with entries in k^ω:

$$(M_\phi)_{ij} = (M_{\phi(0)})_{ij} + ((M_{\phi(1)})_{ij} \times X) + ((M_{\phi(2)})_{ij} \times X^2) + \cdots \tag{7}$$

For the special case of $[H] = (H, 0, 0, 0, \ldots)$, for a linear transformation $H : k^n \to k^m$, we have

$$(M_{[H]})_{ij} = ((M_H)_{ij}, 0, 0, 0, \ldots) \tag{8}$$

We will let the context determine whether entries in k or k^ω are intended, and we shall simply write

$$M_{[H]} = M_H \tag{9}$$

The correspondence between ϕ and M_ϕ is given by the following commutative diagram:

$$\begin{tikzcd}
(k^n)^\omega \arrow[r, "\phi\times(-)"] \arrow[d, "(-)^T"'] & (k^m)^\omega \arrow[d, "(-)^T"] \\
(k^\omega)^n \arrow[r, "M_\phi\times(-)"'] & (k^\omega)^m
\end{tikzcd} \qquad (\phi \times \sigma)^T = M_\phi \times \sigma^T \tag{10}$$

(Recall the definition of $(-)^T$ from (3).) Here $\phi\times(-)$ denotes convolution product and $M_\phi \times (-)$ denotes matrix multiplication. Note that $M_1 = 1$, where 1 on the left denotes the stream $(1, 0, 0, 0, \ldots)$ (consisting of the identity map followed by zero maps), and 1 on the right denotes the identity matrix (having 1's on the diagonal and 0's everywhere else). Also note that

$$M_{\phi\times\psi} = M_\phi \times M_\psi \tag{11}$$

We have the following proposition.

Proposition 2. *Let $\rho \in (k^n \to_L k^n)^\omega$ be a stream of linear transformations $\rho(i) : k^n \to k^n$. If ρ is rational then M_ρ defined in (7) has entries in $Rat(k^\omega)$.*

Proof: Consider two polynomial streams $\phi, \psi \in (k^n \to_L k^n)^\omega$. The entries of the matrices M_ϕ and M_ψ are polynomial streams in k^ω. If ψ moreover has an inverse ψ^{-1} then $M_1 = 1$ and (11) imply $M_{\psi^{-1}} = (M_\psi)^{-1}$, which has values in $Rat(k^\omega)$. It follows that $M_{\phi\times\psi^{-1}} = M_\phi \times (M_\psi)^{-1}$ has values in $Rat(k^\omega)$. □

Example 3. Let $k = \mathbb{R}$ and let $F, G : \mathbb{R}^2 \to \mathbb{R}^2$ be linear transformations defined by

$$M_F = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \qquad M_G = \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix}$$

We compute the matrices of the rational streams $\tilde{F} = (\,1 - (F \times X)\,)^{-1}$ and $\tilde{G} = (\,1 - (G \times X)\,)^{-1}$:

$$M_{\tilde{F}} = (M_{1-(F\times X)})^{-1} = \begin{pmatrix} 1-X & -X \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{1-X} & \frac{X}{1-X} \\ 0 & 1 \end{pmatrix}$$

$$M_{\tilde{G}} = (M_{1-(G\times X)})^{-1} = \begin{pmatrix} 1 & X \\ -X & 1-2X \end{pmatrix}^{-1} = \frac{1}{(1-X)^2} \cdot \begin{pmatrix} 1-2X & -X \\ X & 1 \end{pmatrix}$$

□

4 Systems Coalgebraically

We recall the coalgebraic semantics of systems with input and output. States, inputs and outputs will be represented by plain sets, and homomorphisms will be simply functions between sets. In Section 5, we will look at the coalgebraic modelling of linear systems, involving vector spaces and linear maps.

A *system* (S, n) consists of a set S and a function $n : S \to S$, assigning to a state $s \in S$ its next state $n(s)$. We call the function n the *dynamics* of the system (S, n). A system $(S, \langle o, n \rangle)$ *with output in* a given set O (or simply O-system) consists of a set S of states, a function $n : S \to S$ and an *output* function $o : S \to O$. (Categorically speaking, an O-system is a *coalgebra* of the functor $O \times (-) : Set \to Set$.) A *homomorphism of O-systems* $(S, \langle o_S, n_S \rangle)$ and $(T, \langle o_T, n_T \rangle)$ is a function $h : S \to T$ such that $n_T \circ h = h \circ n_S$ and $o_T \circ h = o_S$; that is, such that the diagram below commutes:

$$\begin{tikzcd}
S \arrow[r, "h"] \arrow[d, "\langle o_S, n_S \rangle"'] & T \arrow[d, "\langle o_T, n_T \rangle"] \\
O \times S \arrow[r, "1 \times h"] & O \times T
\end{tikzcd}$$

Here and throughout the paper, we use 1 to denote the identity function. The set of all streams O^ω is an O-system $(O^\omega, \langle h, t \rangle)$ where

$$h : O^\omega \to O, \quad h(\sigma) = \sigma(0) \quad \text{and} \quad t : O^\omega \to O^\omega, \quad t(\sigma) = \sigma'$$

(Recall that $\sigma' = (\sigma(1), \sigma(2), \sigma(3), \ldots)$.) Initial value and derivative are often called *head* and *tail*, hence our choice of symbols. The O-system $(O^\omega, \langle h, t \rangle)$ has the following universal property, called *finality*: For every O-system $(S, \langle o, n \rangle)$ there exists a unique homomorphism $f : (S, \langle o, n \rangle) \to (O^\omega, \langle h, t \rangle)$, called the *final behaviour* of $(S, \langle o, n \rangle)$. It is given by

$$\begin{tikzcd}
S \arrow[r, dashed, "f"] \arrow[d, "\langle o, n \rangle"'] & O^\omega \arrow[d, "\langle h, t \rangle"] \\
O \times S \arrow[r, dashed, "1 \times f"] & O \times O^\omega
\end{tikzcd} \qquad f(s) = (o(s),\, o \circ n(s),\, o \circ n^2(s),\, \ldots)$$

where $n^0(s) = s$ and $n^{l+1}(s) = n(n^l(s))$.

Any system (S, n) (without output) is an S-system $(S, \langle 1, n \rangle)$ with output $1 : S \to S$ in S. We denote the corresponding final homomorphism by $\tilde{n}$:

$$\begin{tikzcd}
S \arrow[r, dashed, "\tilde{n}"] \arrow[d, "\langle 1, n \rangle"'] & S^\omega \arrow[d, "\langle h, t \rangle"] \\
S \times S \arrow[r, dashed, "1 \times \tilde{n}"] & S \times S^\omega
\end{tikzcd} \qquad \tilde{n}(s) = (s,\, n(s),\, n^2(s),\, \ldots) \tag{12}$$

We call $\tilde{n}$ the *fully observable* behaviour of (S, n). The final behaviour f of an O-system $(S, \langle o, n\rangle)$ factors through its fully observable behaviour $\tilde{n}$ as follows:

$$\begin{tikzcd}
S \arrow[r, "\tilde{n}"'] \arrow[rr, bend left, "f"] \arrow[d, "\langle 1,n\rangle"] \arrow[dd, bend right=60, "\langle o,n\rangle"'] & S^\omega \arrow[r, "o^\omega"'] \arrow[d, "\langle h,t\rangle"] & O^\omega \arrow[dd, "\langle h,t\rangle"] \\
S \times S \arrow[r, "1\times\tilde{n}"'] \arrow[d, "o\times 1"] & S \times S^\omega \arrow[d, "o\times 1"] & \\
O \times S \arrow[r, "1\times\tilde{n}"] \arrow[rr, bend right, "1\times f"'] & O \times S^\omega \arrow[r, "1\times o^\omega"] & O \times O^\omega
\end{tikzcd}
\qquad f = o^\omega \circ \tilde{n} \tag{13}$$

Next we consider systems with output *and* input. As before let O be a set of outputs. In addition, let I be an arbitrary set, the elements of which we call *inputs*. A system (S, ϕ) *with input in I and output in O* (or simply (I, O)-system) consists of a set S of states together with a function $\phi : S \to (O \times S)^I$. The function ϕ maps a state $s \in S$ to a function $\phi(s) : I \to O \times S$ that sends an input i to a pair $\phi(s)(i) \in O \times S$. We shall sometimes use the following notation:

$$s_1 \xrightarrow{i|o} s_2 \iff \phi(s_1)(i) = \langle o, s_2\rangle$$

which can be read as: in state s_1 and with input i the system changes to state s_2 while producing output o. Note that in general both the next state and the output depends on *both* the starting state and the input. Systems with input in I and output in O are also known in the literature as *Mealy machines* [Eil74]. Categorically, an (I, O)-system is a coalgebra of the functor $\mathcal{F} : Set \to Set$ defined by $\mathcal{F}(S) = (O \times S)^I$.

Let (S, ϕ_S) and (T, ϕ_T) be two (I, O)-systems. For $s_1 \in S$ and $i \in I$ let $\phi(s_1)(i) = \langle o, s_2\rangle$. A *homomorphism of (I, O)-systems* is a function $h : S \to T$ such that $\phi_T(h(s))(i) = \langle o, h(s_2)\rangle$, for all $s_1 \in S$ and $i \in I$. Equivalently, h should make the diagram below commute:

$$\begin{tikzcd}
S \arrow[r, "h"] \arrow[d, "\phi_S"'] & T \arrow[d, "\phi_T"] \\
(O \times S)^I \arrow[r, "(1\times h)^1"] & (O \times T)^I
\end{tikzcd}$$

A *final* (I, O)-system can be constructed as follows. We call a function $g : I^\omega \to O^\omega$ *causal* (aka synchronous or letter-to-letter) if for any $\sigma \in I^\omega$ the n-th element of $g(\sigma)$ depends on only the first n elements of the input σ; that is,

$$\sigma(0) = \tau(0), \ldots, \sigma(n) = \tau(n) \;\Rightarrow\; g(\sigma)(n) = g(\tau)(n)$$

for all $\sigma, \tau \in I^\omega$ and $n \geq 0$. We denote the set of all causal functions by

$$\Gamma = \{\, g : I^\omega \to O^\omega \mid g \text{ is causal} \,\} \tag{14}$$

Let $g : I^\omega \to O^\omega$ be causal and let $i \in I$. We define the *initial output* of g on input i by

$$g[i] = g(i : \sigma)(0) \tag{15}$$

where $\sigma \in I^\omega$ is arbitrary. Note that the value $g[i] \in O$ does not depend on σ, since g is causal. We define the *stream function derivative* of g on input i by

$$g_i : I^\omega \to O^\omega, \quad g_i(\sigma) = g(i : \sigma)' \tag{16}$$

We obtain an (I, O)-system $(\mathit{\Gamma}, \gamma : \mathit{\Gamma} \to (O \times \mathit{\Gamma})^I)$ by defining:

$$\gamma(g)\,(i) = \langle g[i],\, g_i \rangle$$

Proposition 4 ([Rut06, HCR06]). *The (I, O)-system $(\mathit{\Gamma}, \gamma)$ of causal functions is* final*: for every (I, O)-system (S, ϕ) there exists a unique homomorphism*

$$\begin{tikzcd} S \arrow[r, dashed, "f"] \arrow[d, "\phi"'] & \mathit{\Gamma} \arrow[d, "\gamma"] \\ (O \times S)^I \arrow[r, dashed, "(1\times f)^1"] & (O \times \mathit{\Gamma})^I \end{tikzcd} \qquad \text{final behaviour of } (S, \phi)$$

Proof: Let $s_0 \in S$, $\sigma \in I^\omega$ and $n \geq 0$, and define

$$f(s_0)\,(\sigma)\,(n) = o_n \qquad \text{where} \quad s_0 \xrightarrow{\sigma(0)|o_0} s_1 \xrightarrow{\sigma(1)|o_1} \cdots \xrightarrow{\sigma(n)|o_n} s_{n+1}$$

Then $f(s_0)$ is causal and f is the unique function making the diagram above commute. □

5 Linear Systems Coalgebraically

We will now model linear systems coalgebraically, by simply applying the results from Section 4, and taking into account the fact that linear systems are defined in terms of vector spaces and linear maps. As before, we shall first treat systems with only output. Next we deal with systems that have both input and output.

We call a system (V, F) *linear* if the state space V is a vector space (over a given field k) and the dynamics $F : V \to V$ is a linear transformation. A system $(V, \langle H, F \rangle)$ with output in O is linear if in addition O is a vector space (over the same field k) and $H : V \to O$ is a linear transformation. A *homomorphism* of *linear* O-systems $(V, \langle H_V, F_V \rangle)$ and $(W, \langle H_W, F_W \rangle)$ is a homomorphism of O-systems which is linear:

$$\begin{tikzcd} V \arrow[r, "h"] \arrow[d, "\langle H_V, F_V \rangle"'] & W \arrow[d, "\langle H_W, F_W \rangle"] \\ O \times V \arrow[r, "1\times h"] & O \times W \end{tikzcd} \qquad h \text{ is a linear transformation}$$

Recall from Section 4 that the O-system $(O^\omega, \langle h, t\rangle)$ is final among *all* (not necessarily linear) systems. We saw (in Section 2) that if O is a vector space then O^ω is also a vector space. Since initial value and derivative are linear transformations, $(O^\omega, \langle h, t\rangle)$ is a linear O-system. The final behaviour $f : V \to O^\omega$ of an O-system $(V, \langle H, F\rangle)$ is given, according to (13), by

$$\begin{tikzcd}
V \arrow[r, "\tilde{F}"'] \arrow[rr, bend left, "f"] & V^\omega \arrow[r, "H^\omega"'] & O^\omega
\end{tikzcd} \qquad f(v) = H^\omega \circ \tilde{F}(v)$$

This is equivalent, for all $v \in V$, to

$$\begin{aligned}
f(v) &= H^\omega \circ \tilde{F}(v) \\
&= (H(v),\, H \circ F(v),\, H \circ F^2(v), \ldots) \\
&= (H, 0, 0, 0, \ldots) \times (1, F, F^2, \ldots) \times (v, 0, 0, 0, \ldots) \quad \text{[using (5) and (6)]} \\
&= (H, 0, 0, 0, \ldots) \times \tilde{F} \times (v, 0, 0, 0, \ldots) \quad \text{[using } (1, F, F^2, \ldots) = \tilde{F}\text{, as in (4)]} \\
&= [H] \times \tilde{F} \times [v]
\end{aligned}$$

Thus:

$$\begin{tikzcd}
V \arrow[r, "\tilde{F}\times[-]"'] \arrow[rr, bend left, "f"] & V^\omega \arrow[r, "{[H]\times(-)}"'] & O^\omega
\end{tikzcd} \qquad f(v) = [H] \times \tilde{F} \times [v]$$

It follows that f is a linear transformation and that $(O^\omega, \langle h, t\rangle)$ is final in the family of all linear O-systems and linear homomorphisms between them.

The final behaviour of *finite dimensional* linear O-systems can be further characterised as follows. Let $n, m \geq 1$ and consider a system $(k^n, \langle H, F\rangle)$ with linear transformations $F : k^n \to k^n$ and $H : k^n \to k^m$. By (10), the following diagram commutes:

$$\begin{tikzcd}
(k^n)^\omega \arrow[r, "\tilde{F}\times(-)"] \arrow[d, "(-)^T"'] & (k^n)^\omega \arrow[r, "{[H]\times(-)}"] \arrow[d, "(-)^T"] & (k^m)^\omega \arrow[d, "(-)^T"] \\
(k^\omega)^n \arrow[r, "M_{\tilde{F}}\times(-)"'] & (k^\omega)^n \arrow[r, "M_H\times(-)"'] & (k^\omega)^m
\end{tikzcd} \qquad ([H]\times\tilde{F}\times(-))^T = M_H \times M_{\tilde{F}} \times (-)^T \tag{17}$$

(where we use the convention (9) of writing $M_{[H]} = M_H$). It follows that the final behaviour f satisfies

$$f(v)^T \;=\; ([H] \times \tilde{F} \times [v])^T \;=\; M_H \times M_{\tilde{F}} \times [v]^T \tag{18}$$

We saw in (4) that $\tilde{F} = (1 - (F \times X))^{-1}$ is a rational stream. By Proposition 2, the matrix $M_{\tilde{F}}$ has values in $Rat(k^\omega)$. And so we have proved the following.

Proposition 5. *For a finite dimensional system* $(k^n, \langle H, F\rangle)$ *with dynamics* $F : k^n \to k^n$ *and output* $H : k^n \to k^m$*, the final behaviour* $f : k^n \to k^m$ *satisfies, for all* $v \in k^n$*,*

$$f(v)^T \;=\; M_H \times M_{\tilde{F}} \times [v]^T$$

and thus is obtained from $[v]^T$ *by multiplication with an* $m \times n$ *matrix with values in* $Rat(k^\omega)$*.* □

Example 6. Let $k = \mathrm{I\!R}$ and consider the linear system $(\mathrm{I\!R}^2, \langle H, F\rangle)$ with output $H : \mathrm{I\!R}^2 \to \mathrm{I\!R}$ and dynamics $F : \mathrm{I\!R}^2 \to \mathrm{I\!R}^2$ given by

$$H = \begin{pmatrix} 1 \; 1 \end{pmatrix} \qquad F = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

The matrix $M_{\tilde{F}}$ corresponding to $\tilde{F}$ has been computed in Example 3:

$$M_{\tilde{F}} = \begin{pmatrix} \frac{1}{1-X} & \frac{X}{1-X} \\ 0 & 1 \end{pmatrix}$$

The final behaviour $f_{\langle H,F\rangle} : \mathrm{I\!R}^2 \to \mathrm{I\!R}^\omega$ of this system is given, for any $(a, b) \in \mathrm{I\!R}^2$, by

$$\begin{aligned} f_{\langle H,F\rangle}(a, b) &= \begin{pmatrix} 1 \; 1 \end{pmatrix} \times \begin{pmatrix} \frac{1}{1-X} & \frac{X}{1-X} \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} a \\ b \end{pmatrix} \\ &= \frac{a+b}{1-X} \end{aligned}$$

(omitting square brackets around a and b as usual). Repeating the example with a different output function $\bar{H}$ and the same dynamics F:

$$\bar{H} = \begin{pmatrix} 1 \; 2 \end{pmatrix} \qquad F = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

leads to the following final behaviour:

$$f_{\langle \bar{H},F\rangle}(a, b) \;=\; \begin{pmatrix} \frac{1}{1-X} & \frac{2-X}{1-X} \end{pmatrix} \times \begin{pmatrix} a \\ b \end{pmatrix} \;=\; \frac{a + 2b - bX}{1-X}$$

□

Next we discuss linear systems with *input and* output. We shall model them as (I, O)-systems, as defined in Section 4, and then study their final behaviour.

Let I, O and V be vector spaces over k, and let $F : V \to V$, $G : I \to V$ and $H : V \to O$ be linear transformations. We define the (I, O)-system $(V, \Phi_{\langle H,F,G\rangle})$ by

$$\Phi_{\langle H,F,G\rangle} : V \to (O \times V)^I \qquad \Phi_{\langle H,F,G\rangle}(v)(i) = \langle H(v), F(v) + G(i)\rangle \tag{19}$$

or equivalently, expressed in terms of transitions,

$$v \xrightarrow{i \,|\, H(v)} F(v) + G(i)$$

We call $(V, \Phi_{\langle H,F,G\rangle})$ a *linear* (I, O)-system because of the linearity of F, G, and H. However, note that Φ itself is not linear and likewise, homomorphisms of linear (I, O)-systems will generally not be linear. This is in contrast with the family of linear O-systems, where everything *is* linear.

For a linear(I, O)-system $(V, \Phi_{\langle H,F,G\rangle})$ we call $(V, \langle H, F\rangle)$ its *underlying* O-system. The key to the coalgebraic understanding of a linear (I, O)-system is the observation that its behaviour is in essence determined by that of its underlying O-system.

The following lemma will be helpful. Consider the final O-system $(O^\omega, \langle h, t\rangle)$ and an arbitrary linear transformation $\psi : I \to O^\omega$. This gives rise to a linear (I, O)-system $(O^\omega, \Phi_{\langle h,t,\psi\rangle})$, with $\Phi_{\langle h,t,\psi\rangle}$ defined as in (19). The lemma below describes its final behaviour $g : O^\omega \to \Gamma$, introduced in Proposition 4.

Lemma 7. *For all* $\alpha \in O^\omega$ *and* $\sigma \in I^\omega$,

$$\begin{tikzcd} O^\omega \arrow[r, dashed, "g"] \arrow[d, "\Phi_{\langle h,t,\psi\rangle}"'] & \Gamma \arrow[d, "\gamma"] \\ (O \times O^\omega)^I \arrow[r, dashed, "(1\times g)^1"'] & (O \times \Gamma)^I \end{tikzcd} \qquad g(\alpha)(\sigma) = \alpha + (\psi \times X \times \sigma)$$

(On the right, we read ψ as a stream of linear transformations $\psi \in (I \to_L O)^\omega \cong I \to_L O^\omega$.)

Proof: By finality of (Γ, γ), it is sufficient to show that the function g defined as above is a homomorphism of (I, O)-systems. By definition of γ, we have $\gamma(g(\alpha))(i) = \langle\, g(\alpha)[i],\, g(\alpha)_i\,\rangle$, for all $i \in I$. Now

$$g(\alpha)[i] = \,(\, g(\alpha)(i : \sigma)\,)\,(0) = \,\alpha(0)$$

and, for all $\sigma \in I^\omega$,

$$\begin{aligned} g(\alpha)(i : \sigma) &= g(\alpha)(i + (X \times \sigma)) \quad \text{[by Lemma 1, with } i = (i, 0, 0, 0, \ldots)\text{]} \\ &= \alpha + (\, X \times \psi \times (i + (X \times \sigma))\,) \\ &= \alpha + (X \times \psi \times i\,) + (X \times \psi \times X \times \sigma) \end{aligned} \tag{20}$$

This implies

$$\begin{aligned} g(\alpha)_i(\sigma) &= (\, g(\alpha)(i : \sigma)\,)' \quad \text{[definition stream function derivative (16)]} \\ &= (\, \alpha' + (\psi \times i)\,) + (\psi \times X \times \sigma) \quad \text{[using (20) and Lemma 1]} \\ &= g(\alpha' + \psi(i))(\sigma) \quad \text{[using } \psi \times i = \psi(i)\text{]} \end{aligned}$$

It follows that

$$\begin{aligned} \gamma(g(\alpha))(i) &= \langle\, \alpha(0),\, g(\alpha' + \psi(i))\,\rangle \\ &= (1 \times g)\,(\,\langle\, \alpha(0), \alpha' + \psi(i)\,\rangle\,) \\ &= \big(\,(1 \times g)^1 \circ \Phi_{\langle h,t,\psi\rangle}\,(\alpha)\,\big)\,(i) \quad \text{[definition } \Phi_{\langle h,t,\psi\rangle} \text{ (19)]} \end{aligned}$$

This shows that the diagram above commutes. Thus g is a homomorphism. □
Next we observe that for a linear (I,O)-system $(V, \Phi_{\langle H,F,G\rangle})$, with $\Phi_{\langle H,F,G\rangle}$ as in (19), the function $\Phi_{\langle H,F,G\rangle}$ can be decomposed as follows:

$$\begin{tikzcd}
V \arrow[r, "\langle 1,F\rangle"'] \arrow[rrr, bend left, "\Phi_{\langle H,F,G\rangle}"] & V \times V \arrow[r, "H\times 1"'] & O \times V \arrow[r, "G_+"'] & (O \times V)^I
\end{tikzcd} \tag{21}$$

where the function G_+ is defined, for all $o \in O$, $v \in V$, and $i \in I$, by

$$G_+(\langle o, v\rangle)(i) = \langle o,\, v + G(i)\rangle$$

Theorem 8.
The final behaviour[1] $f : V \to \Gamma$ *of a linear* (I,O)*-system* $(V, \Phi_{\langle H,F,G\rangle})$ *(as defined in (19)) satisfies, for all* $v \in V$ *and* $\sigma \in I^\omega$,

$$f(v)(\sigma) = \Big([H] \times \tilde{F} \times [v]^T\Big) + \Big([H] \times \tilde{F} \times [G] \times X \times \sigma\Big)$$

Proof: Let $\psi : I \to O^\omega$ be defined by $\psi = [H] \times \tilde{F} \times [G]$ and consider the following diagram:

$$\begin{tikzcd}
V \arrow[r, "\tilde{F}"'] \arrow[d, "\langle 1,F\rangle"] \arrow[rrr, bend left, "f"] \arrow[ddd, bend right=60, "\Phi_{\langle H,F,G\rangle}"'] & V^\omega \arrow[r, "H^\omega"'] \arrow[d, "\langle h,t\rangle"] & O^\omega \arrow[r, "g"'] \arrow[dd, "\langle h,t\rangle"'] & \Gamma \arrow[ddd, "\gamma"] \\
V \times V \arrow[r, "1\times\tilde{F}"'] \arrow[d, "H\times 1"] & V \times V^\omega \arrow[d, "H\times 1"] & & \\
O \times V \arrow[r, "1\times\tilde{F}"'] \arrow[d, "G_+"] & O \times V^\omega \arrow[r, "1\times H^\omega"'] \arrow[d, phantom, "(*)"] & O \times O^\omega \arrow[d, "\psi_+"'] & \\
(O \times V)^I \arrow[rr, "(1\times(H^\omega\circ\tilde{F}))^1"'] \arrow[rrr, bend right, "(1\times f)^1"'] & & (O \times O^\omega)^I \arrow[r, "(1\times g)^1"] & (O \times \Gamma)^I
\end{tikzcd}$$

Recall that $H^\omega = [H] \times (-)$, using the convolution product introduced in (5) and, consequently, $H^\omega \circ \tilde{F} = [H] \times \tilde{F}$. The function ψ has been defined precisely such that the rectangle $(*)$ above commutes. (Note that a proof of $(*)$ will use the linearity of $[H] \times \tilde{F}$.) The right hand pentagon commutes by Lemma 7. Everything else commutes by finality. □

The final behaviour of *finite dimensional* linear (I,O)-systems can be further characterised, similarly to the case of linear O-systems. First we define for any causal function $g : (k^l)^\omega \to (k^m)^\omega$ a function.

---

[1] We observe that the final behaviour $f(0)$ of the initial state 0 corresponds to what is known in system theory as the *transfer function* of the system, where $\tilde{F}$ is often expressed as $(zI - F)^{-1}$.

$$\langle g\rangle : (k^\omega)^l \to (k^\omega)^m \qquad \langle g\rangle(\sigma) = g(\sigma^T)^T$$

for all $\sigma \in (k^\omega)^l$, and denote the image of Γ under this operation by $\langle\Gamma\rangle$. (Note that $\Gamma \cong \langle\Gamma\rangle$.)

Proposition 9.
For a finite dimensional linear (I,O)-system $(k^n, \Phi_{\langle H,F,G\rangle})$ with dynamics $F : k^n \to k^n$, input $G : k^l \to k^n$, and output $H : k^n \to k^m$, the final behaviour $\langle f\rangle : k^n \to \langle\Gamma\rangle$ satisfies, for all $v \in k^n$ and $\sigma \in (k^\omega)^l$,

$$\langle f(v)\rangle(\sigma) = \big(M_H \times M_{\tilde{F}} \times [v]^T \big) + \big(M_H \times M_{\tilde{F}} \times M_G \times X \times \sigma \big)$$

where all these matrices have values in $Rat(k^\omega)$.

Proof: By (10), all squares below commute:

$$\begin{tikzcd}
(k^\omega)^l \arrow[r, "M_G\times(-)"] \arrow[d, "(-)^T"] & (k^\omega)^n \arrow[r, "M_{\tilde{F}}\times(-)"] & (k^\omega)^n \arrow[r, "M_H\times(-)"] & (k^\omega)^m \\
(k^l)^\omega \arrow[r, "G\times(-)"'] & (k^n)^\omega \arrow[r, "\tilde{F}\times(-)"'] \arrow[u, "(-)^T"] & (k^n)^\omega \arrow[r, "H\times(-)"'] \arrow[u, "(-)^T"] & (k^m)^\omega \arrow[u, "(-)^T"]
\end{tikzcd}$$

The proposition follows from this diagram and Theorem 8. As in Proposition 5, all matrices have values in $Rat(k^\omega)$. □

Example 10. (This is Example 6, continued.) Let $k = \mathbb{R}$ and consider the linear system (I,O)-system $(\mathbb{R}^2, \Phi_{\langle H,F,G\rangle})$ with $H : \mathbb{R}^2 \to \mathbb{R}$, $F : \mathbb{R}^2 \to \mathbb{R}^2$ and $G : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$H = \begin{pmatrix} 1 & 1 \end{pmatrix} \qquad F = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \qquad G = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$$

The final behaviour $\langle f(a,b)\rangle : (\mathbb{R}^\omega)^2 \to \mathbb{R}^\omega$ of a state $(a,b) \in \mathbb{R}^2$ is given, for all pairs of input streams $(\sigma_1, \sigma_2) \in (\mathbb{R}^\omega)^2$, by

$$\begin{aligned}
\langle f(a,b)\rangle(\sigma) &= \left(M_H \times M_{\tilde{F}} \times \begin{pmatrix} a \\ b \end{pmatrix} \right) + \left(M_H \times M_{\tilde{F}} \times M_G \times X \times \begin{pmatrix} \sigma_1 \\ \sigma_2 \end{pmatrix} \right) \\
&= \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{1-X} & \frac{X}{1-X} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \quad + \\
&\quad \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{1-X} & \frac{X}{1-X} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} X \times \sigma_1 \\ X \times \sigma_2 \end{pmatrix} \\
&= \frac{a + (2X \times \sigma_1)}{1-X} \quad + \quad \frac{b + (3X \times \sigma_2)}{1-X}
\end{aligned}$$

□

6 Minimization and Equivalence

Because O- and (I, O)-systems are coalgebras, the general definition of coalgebraic equivalence applies. Here we spell out these definitions together with the observation that the corresponding minimization of a system is given by the (image under) the final behaviour mapping. For *linear* (I, O)-systems, we shall see that minimization and equivalence are particularly simple, as they are entirely determined by their underlying O-systems.

Equivalence of (not necessarily linear) O-systems is defined as follows. A relation $R \subseteq S \times T$ is called an *O-bisimulation* between O-systems $(S, \langle o_S, n_S\rangle)$ and $(T, \langle o_T, n_T\rangle)$ if for all $s \in S$ and $t \in T$:

$$\langle s, t\rangle \in R \Rightarrow \begin{cases} o_S(s) = o_T(t) \text{ and} \\ \langle n_S(s), n_T(t)\rangle \in R \end{cases}$$

We say that s and t are *O-equivalent* and write $s \sim_O t$ if there exists an O-bisimulation R with $\langle s, t\rangle \in R$. The final behaviour $f : S \to O^\omega$ of an O-system $(S, \langle o, n\rangle)$ identifies precisely all O-equivalent states: $s_1 \sim_O s_2$ iff $f(s_1) = f(s_2)$, for all $s_1, s_2 \in S$. (For the elementary proof, see [Rut03].) As a consequence, the minimization of an O-system with respect to O-equivalence is given by the image of S under f, which is a subsystem $f(S) \subseteq O^\omega$ because f is a homomorphism. It follows that if the system is linear, then the greatest O-equivalence on S is given by the kernel $ker(f)$.

For (I, O)-systems there exists a corresponding notion of (I, O)-equivalence and, again, the final behaviour identifies precisely all (I, O)-equivalent states: see [Rut06] for details. For *linear* (I, O)-systems, things are much simpler since their behaviour is determined by their underlying O-system.

Proposition 11. *The minimization of a linear (I, O)-system $(V, \Phi_{\langle H,F,G\rangle})$ is isomorphic to the minimization of its underlying O-system $(V, \langle H, F\rangle)$.*

Proof: By the proof of Theorem 8, the final behaviour $f : V \to \Gamma$ satisfies $f(v) = g(H \times \tilde{F} \times v)$, for all $v \in V$. Here the function $g : O^\omega \to \Gamma$ is given, according to Lemma 7, by $g(\alpha)(\sigma) = \alpha + (H \times \tilde{F} \times G \times X \times \sigma)$, for $\alpha \in O^\omega$ and $\sigma \in I^\omega$. Taking $\sigma = 0$, we see that g is injective. Thus the image of $(V, \Phi_{\langle H,F,G\rangle})$ under the final behaviour map f is isomorphic to its image under $H \times \tilde{F}$. The underlying O-system of this image is the minimization of $(V, \langle H, F\rangle)$. □

Example 12. Recall the (I, O)-system $(\mathbb{R}^2, \Phi_{\langle H,F,G\rangle})$ from Example 10. Computing its image W under $H \times \tilde{F}$ yields

$$W = \left(H \times \tilde{F}\right)(\mathbb{R}^2) = \{\frac{a+b}{1-X} \mid (a, b) \in \mathbb{R}^2\} \subseteq \mathbb{R}^\omega$$

Output and dynamics on W are induced by $\langle h, t\rangle : \mathbb{R}^\omega \to (\mathbb{R} \times \mathbb{R}^\omega)$. The input map on W is given by (the corestriction of) $\psi = H \times \tilde{F} \times G : \mathbb{R}^2 \to \mathbb{R}^\omega$ and satisfies

$$H \times \tilde{F} \times G = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{1-X} & \frac{X}{1-X} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} \frac{2}{1-X} & \frac{3}{1-X} \end{pmatrix}$$

Choosing $1/1 - X$ as a basis for W, we find that the resulting minimization is isomorphic to $\mathbb{R}$, with output and dynamics both given by $1 : \mathbb{R} \to \mathbb{R}$, and with input $(2\ \ 3) : \mathbb{R}^2 \to \mathbb{R}$. □

7 Realisation

We discuss the realisation of linear and non-linear systems, first with only output and then with input and output.

A state $s \in S$ in a (not necessarily linear) O-system $(S, \langle o, n \rangle)$ *realises* a stream $\sigma \in O^\omega$ if the final behaviour of s satisfies $f(s) = \sigma$. If O is a *set* (and not necessarily a vector space), a minimal realisation for a stream $\sigma \in O^\omega$ is obtained by taking as state space the set

$$S_\sigma = \{\sigma^{(0)}, \sigma^{(1)}, \sigma^{(2)}, \ldots\} \tag{22}$$

with $\sigma^{(0)} = \sigma$ and $\sigma^{(n+1)} = t(\sigma^{(n)}) = (\sigma^{(n)})'$. As output function and dynamics, one simply takes the restrictions of $h : O^\omega \to O$ and $t : O^\omega \to O^\omega$ to S_σ. The set inclusion $S_\sigma \subseteq O^\omega$ is a homomorphism of O-systems. By finality of $(O^\omega, \langle h, t \rangle)$, this homomorphism is unique. It follows that $f(\sigma) = \sigma$ and hence that $(S_\sigma, \langle h, t \rangle)$ with initial state σ is a minimal realisation of σ.

If O is a vector space then O^ω is also a vector space and we will be interested in realisations that themselves are vector spaces again. Thus a minimal realisation for a stream $\sigma \in O^\omega$ will consist of the smallest *subspace* of O^ω that contains σ and is closed under the linear transformation $t : O^\omega \to O^\omega$. This (so-called t-cyclic) vector space $Z_\sigma \subseteq O^\omega$ is the subspace of O^ω that is spanned by the set S_σ of vectors in (22).

Of special interest are those $\sigma \in O^\omega$ such that, for some $n \geq 1$, all of $\sigma = \sigma^{(0)}$ through $\sigma^{(n-1)}$ are linearly independent and

$$\sigma^{(n)} + \left(c_{n-1} \times \sigma^{(n-1)}\right) + \cdots + \left(c_1 \times \sigma^{(1)}\right) + \left(c_0 \times \sigma^{(0)}\right) = 0$$

for some coefficients $c_0, \ldots, c_{n-1}$ in the base field k of O and O^ω. Then Z_σ is a vector space of dimension n. The linear transformation $F : Z_\sigma \to Z_\sigma$ induced by $t : O^\omega \to O^\omega$ is given, with respect to the (ordered) basis $\sigma^{(0)}, \ldots, \sigma^{(n-1)}$, by the $n \times n$ matrix

$$M_F = \begin{pmatrix} 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & \cdots & 0 & -c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -c_{n-1} \end{pmatrix}$$

(This matrix is in fact (a variation of) the *companion* matrix of the so-called t-order polynomial of σ; cf. [BM77, Thm.15, p.339].) The linear transformation

$H : Z_\sigma \to O$ induced by $h : O^\omega \to O$ is given, again with respect to the basis $\sigma^{(0)}, \ldots, \sigma^{(n-1)}$, by the matrix (of size $dim(O) \times n$)

$$M_H = \begin{pmatrix} \sigma^{(0)}(0) & \sigma^{(1)}(0) & \sigma^{(2)}(0) & \cdots & \sigma^{(n-1)}(0) \end{pmatrix}$$

Thus we have obtained a linear O-system $(Z_\sigma, \langle H, F\rangle)$ of dimension n. As before, the inclusion $Z_\sigma \subseteq O^\omega$ is a homomorphism. Thus $f(\tau) = \tau$, for all $\tau \in Z_\sigma$ and $(Z_\sigma, \langle H, F\rangle)$ with σ as initial state is a minimal realisation of σ.

Example 13. Let $O = \mathbb{R}$ and consider the stream $\sigma = 1/(1-X)^2 \in O^\omega$. Computing the successive stream derivatives of $\sigma = \sigma^{(0)}$, using Lemma 1, gives

$$\sigma^{(1)} = \frac{2-X}{(1-X)^2} \qquad \sigma^{(2)} = \frac{3-2X}{(1-X)^2} = -\sigma^{(0)} + (2 \times \sigma^{(1)})$$

Thus $\sigma^{(0)}$ and $\sigma^{(1)}$ form a basis for Z_σ. Because $\sigma^{(0)}(0) = 1$ and $\sigma^{(1)}(0) = 2$, we have

$$M_H = \begin{pmatrix} 1 & 2 \end{pmatrix} \qquad M_F = \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix}$$

Now σ is realised by $(Z_\sigma, \langle H, F\rangle)$, with σ as the initial state. Clearly, $\mathbb{R}^2 \cong Z_\sigma$. Note that the isomorphism can also be obtained by computing the final behaviour $f : \mathbb{R}^2 \to \mathbb{R}^\omega$ of the O-system $(\mathbb{R}^2, \langle H, F\rangle)$, using Proposition 5. This gives, for all $(a, b) \in \mathbb{R}^2$,

$$\begin{aligned} f(a,b) &= M_H \times M_{\tilde{F}} \times (a,b) \\ &= \begin{pmatrix} 1 & 2 \end{pmatrix} \times \begin{pmatrix} \frac{1-2X}{(1-X)^2} & \frac{-X}{(1-X)^2} \\ \frac{X}{(1-X)^2} & \frac{1}{(1-X)^2} \end{pmatrix} \times \begin{pmatrix} a \\ b \end{pmatrix} \end{aligned}$$

which satisfies, as expected, $f(1,0) = \sigma$ and $f(0,1) = \sigma^{(1)}$. □

Example 14. Let $O = \mathbb{R}^2$ and consider the pair $(\tau, \sigma) \in (\mathbb{R}^\omega)^2 \cong (\mathbb{R}^2)^\omega$, with $\tau = 1/(1-2X)$ and $\sigma = 1/(1-X)^2$. Computing (pairs of) stream derivatives

$$(\tau,\sigma)^{(1)} = \left(\frac{2}{1-2X}, \frac{2-X}{(1-X)^2} \right) \qquad (\tau,\sigma)^{(2)} = \left(\frac{2^2}{1-2X}, \frac{3-2X}{(1-X)^2} \right)$$

$$(\tau,\sigma)^{(3)} = \left(\frac{2^3}{1-2X}, \frac{4-3X}{(1-X)^2} \right) = 2 \times (\tau,\sigma)^{(0)} - 5 \times (\tau,\sigma)^{(1)} + 4 \times (\tau,\sigma)^{(2)}$$

we see that $Z_{(\tau,\sigma)}$ has dimension 3 with $H : Z_{(\tau,\sigma)} \to \mathbb{R}^2$ and $F : Z_{(\tau,\sigma)} \to Z_{(\tau,\sigma)}$ given by

$$M_H = \begin{pmatrix} 1 & 2 & 4 \\ 1 & 2 & 3 \end{pmatrix} \qquad M_F = \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & -5 \\ 0 & 1 & 4 \end{pmatrix}$$

□

Proposition 15. *Let k be a field and let $O = k^m$. A vector of streams $\sigma \in (k^\omega)^m \cong (k^m)^\omega$ is realisable by a linear k^m-system of finite dimension iff $\sigma \in (Rat(k^\omega))^m$.*

Proof: From left to right, this is Proposition 5. For the converse, it is sufficient to observe that the examples above generalise to arbitrary vectors of rational streams. This is immediate from the fact that for a rational stream $\sigma = \rho/\tau$, the dimension of Z_σ in the construction above is bounded by the maximum of the degrees of ρ and τ. □

Next we turn to systems with *input and output*. Let I and O be sets. A state s in a (not necessarily linear) (I, O)-system (S, ϕ) realises a causal stream function $g : I^\omega \to O^\omega$ if the final behaviour of s satisfies $f(s) = g$. For a given g, a (minimal) realisation is obtained by taking the smallest subsystem S of the final (I, O)-system (Γ, γ) containing g. The system S can be constructed by adding to the singleton set $\{g\}$ all successive stream function derivatives g_i, $(g_i)_j$, etc. (for $i, j, \ldots \in I$), and taking the restriction of γ to S. The inclusion $S \subseteq \Gamma$ is a homomorphism of (I, O)-systems and by finality we have $f(g) = g$. In [Rut06, HCR06], this approach is systematically applied to the realisation (synthesis) of various (non-linear) causal functions on bitstreams (with $I = O = \{0, 1\}$).

For infinite I and O, this construction will in general not be finitely computable. However, if both I and O are finite dimensional vector spaces then the realisation of linear causal stream functions can simply be reduced to the realisation problem of streams, which we have already solved above.

Proposition 16. *Let k be a field and let $I = k^l$ and $O = k^m$. Let $g : (k^\omega)^l \to (k^\omega)^m$ be given by $g(\tau) = M \times X \times \tau$, for an $m \times l$ matrix $M \in (k^\omega)^{m \times l}$. Then g is realisable by a linear (I, O)-system of finite dimension iff $M \in (Rat(k^\omega))^{m \times l}$.*

Proof: From left to right, this is Proposition 9. For the converse, we first consider the case that $l = 1$. So assume that $M \in (Rat(k^\omega))^m$. By Proposition 15, there exists a finite dimensional system $(V, \langle H, F\rangle)$ and $v \in V$ realising M; that is, $f(v) = M_H \times M_{\tilde{F}} \times v = M$. If we define $G : k \to V$ by the matrix $M_G = v$ then $(V, \Phi_{\langle H,F,G\rangle})$ with $0 \in V$ as initial state realises g since, for all $\tau \in k^\omega$,

$$\begin{aligned} f(0)(\tau) &= M_H \times M_{\tilde{F}} \times M_G \times X \times \tau \quad \text{[Proposition 9]} \\ &= M_H \times M_{\tilde{F}} \times v \times X \times \tau \\ &= M \times X \times \tau \\ &= g(\tau) \end{aligned}$$

For $l > 1$ we write M as a direct sum (product) $M = M_1 \oplus \cdots \oplus M_l$, with $M_i \in (Rat(k^\omega))^m$, for $i = 1, \ldots, l$. Then we construct realisations for each of $g_i = M_i \times X$. Their direct sum is a realisation for g. □

Acknowledgments. Discussions with Albert Benveniste – on [Rut05, Ben06] – were very instructive and are gratefully acknowledged. Thanks are also due to Jiri Adamek, H.Peter Gumm, Jan Komenda, Prakash Panangaden, Hans-E. Porst, and Jan van Schuppen, for discussions and pointers to the literature. I am also very grateful to the constructive comments of the three anonymous referees. Amongst others, their suggestions for improvements of the notation were very useful.

References

[AM74] Arbib, M.A., Manes, E.G.: Foundations of system theory: decomposable systems. Automatica 10, 285–302 (1974)

[AM75] Arbib, M.A., Manes, E.G.: Adjoint machines, state-behaviour machines, and duality. Journal of Pure and Applied Algebra 6, 313–344 (1975)

[Ben06] Benveniste, A.: A brief on realisation theory for linear systems (unpublished note)

[BM77] Birkhoff, G., MacLane, S.: A survey of modern algebra, 4th edn. MacMillan Publishing Co., Inc. (1977)

[Eil74] Eilenberg, S.: Automata, languages and machines. In: Pure and applied mathematics, vol. A, Academic Press, London (1974)

[Fuh96] Fuhrmann, P.A.: A polynomial approach to linear algebra. Springer, Berlin (1996)

[HCR06] Hansen, H., Costa, D., Rutten, J.J.M.M.: Synthesis of Mealy machines using derivatives. In: Proceedings of CMCS 2006. ENTCS, vol. 164(1), pp. 27–45. Elsevier, Amsterdam (2006)

[Kai80] Kailath, T.: Linear systems. Prentice-Hall, Englewood Cliffs (1980)

[Kal63] Kalman, R.E.: Mathematical description of linear dynamical systems. SIAM J. Control 1, 152–192 (1963)

[KFA69] Kalman, R.E., Falb, P.L., Arbib, M.A.: Topics in mathematical system theory. McGraw-Hill, New York (1969)

[Koh78] Kohavi, Z.: Switching and finite automata theory. McGraw-Hill, New York (1978)

[Lah98] Lahti, B.P.: Signal Processing & Linear Systems. Oxford University Press, Oxford (1998)

[Pap03] Pappas, G.J.: Bisimilar linear systems. Automatica 39, 2035–2047 (2003)

[Rut03] Rutten, J.J.M.M.: Behavioural differential equations: a coinductive calculus of streams, automata, and power series (Fundamental Study). Theoretical Computer Science 308(1), 1–53 (2003)

[Rut05] Rutten, J.J.M.M.: A tutorial on coinductive stream calculus and signal flow graphs. Theoretical Computer Science 343(3), 443–481 (2005)

[Rut06] Rutten, J.J.M.M.: Algebraic specification and coalgebraic synthesis of Mealy automata. In: Proceedings of FACS 2005. ENTCS, vol. 160, pp. 305–319. Elsevier Science Publishers, Amsterdam (2006)

[Son79] Sontag, E.D.: Polynomial response maps. Lecture Notes in Control and Information Sciences, vol. 13. Springer, Heidelberg (1979)

[vdS04] van der Schaft, A.J.: Equivalence of dynamical systems by bisimulation. IEEE Transactions on Automatic Control 49, 2160–2172 (2004)

Appendix: A Comparison with Algebraic System Theory

In the wide area of (linear) system theory, our coalgebraic treatment of linear systems is probably closest related to what sometimes is called algebraic system theory. Below we give a brief overview of the approach of Kalman, who was one of the early contributors, and compare it to the present model. Classical references are [Kal63, KFA69], but see also [Kai80, Fuh96, Ben06]. Here we rely on the more categorical account of Kalman's model described in [AM74, AM75].

Let

$$I^{(\omega)} = \{ \, \sigma \in I^{\omega} \mid \sigma = (i_0, i_1, \ldots, i_k, 0, 0, 0, \ldots) \text{ for some } k \geq 0, i_j \in I \, \}$$

and consider the following diagram of vector spaces and linear maps:

$$\begin{tikzcd}
I \arrow[d, "e"'] \arrow[dr, "G"] & & O \\
I^{(\omega)} \arrow[r, dashed, "r"] \arrow[d, "X\times(-)"'] & V \arrow[ur, "H"] \arrow[r, dashed, "o"] \arrow[d, "F"] & O^{\omega} \arrow[u, "h"'] \arrow[d, "t"] \\
I^{(\omega)} \arrow[r, dashed, "r"'] & V \arrow[r, dashed, "o"'] & O^{\omega}
\end{tikzcd}$$

(with $e(i) = (i, 0, 0, 0, \ldots)$.) The diagram can be viewed as a theorem stating that every choice of linear transformations G and F induces a unique *reachability* map r such that the left half of the diagram commutes, and similarly, every choice of F and H induces a unique *observability* map o fitting in the right half of the diagram. In this manner, every linear system (V, H, F, G) induces a unique map (called the *transfer function*) $o \circ r : I^{(\omega)} \to O^{\omega}$. It satisfies (in our notation)

$$o \circ r(\sigma) = \; H \times \tilde{F} \times r(\sigma) \tag{23}$$

where the state $r(\sigma)$ reached on input $\sigma = (i_0, i_1, \ldots, i_k, 0, 0, \ldots) \in I^{(\omega)}$ is given by

$$\begin{aligned} r(\sigma) &= \Big(\tilde{F} \times G \times \sigma \Big) (k) \\ &= G(i_0) \, + \, F \circ G(i_1) \, + \, F^2 \circ G(i_2) \, + \, \cdots \, + \, F^k \circ G(i_k) \end{aligned}$$

(Note that the operational interpretation is that i_k is the first input and i_0 is the last.) Comparing (23) with the final behaviour of V given in Theorem 8, we note the following differences: (i) The final behaviour allows arbitrary input streams, not only almost-everywhere-zero ones. (ii) The ordering of the inputs coincides with the input order. (iii) In (23), the behaviour of V is described in two steps: first r computes the state that is reached on finite input, then the (infinite) output stream is computed; in contrast, Theorem 8 describes the behaviour of an arbitrary initial state for all (infinite) streams of inputs.

Further differences between the two approaches can be noted regarding the way realisation is handled. In Kalman's approach, a realisation of a linear map g as in Proposition 16 is obtained by constructing its so-called (infinite) *Hankel* matrix H_g, viewing H_g as a linear transformation from $I^{(\omega)}$ to O^{ω}, and the observation that if H_g has finite rank then this linear transformation factors through a finite dimensional vector space V as in the diagram above. In contrast, Proposition 16 reduces realisation of linear maps to the realisation of streams, and the latter are simply given by the corresponding (t-cyclic) subspaces of O^{ω}.