6.5   Realization Theory for Linear Systems

Motivated by the discussion in Sections 3.3 and 6.2, it becomes of interest to study *triples of matrices* (A, B, C) over a field $\mathbb{K}$,

$$A \in \mathbb{K}^{n \times n}, \quad B \in \mathbb{K}^{n \times m}, \quad C \in \mathbb{K}^{p \times n},$$

corresponding to linear time-invariant systems with outputs, and it is unambiguous to call such a triple "controllable" or "observable," meaning that the corresponding discrete-time and (for $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$) continuous-time systems have that property. The *dimension* of the triple is the number n, and m and p are the *number of inputs* and *outputs*, respectively. We denote by

$$S_{n,m,p}$$

the set of all such (A, B, C)'s, and we use

$$S_{n,m,p}^{c,o}$$

for the set of controllable and observable ones. Note that for $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, $S_{n,m,p}^{c,o}$ is an open dense subset of $S_{n,m,p}$, in the topology for

$$\mathbb{K}^{n \times n} \times \mathbb{K}^{n \times m} \times \mathbb{K}^{p \times n}$$

obtained when identifying this space with $\mathbb{K}^{n(n+m+p)}$.

Assume that (A, B, C) is so that rank $\mathbf{O}(A, C) = r$ and pick a decomposition as in Equation (6.8), where $T \in GL(n)$. Write

$$\tilde{B} := T^{-1}B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}$$

and consider the pair (A_1, B_1). If this is not controllable, we can decompose it via Lemma 3.3.3. Let $S_1 \in GL(r)$ be so that

$$S_1^{-1} A_1 S_1 = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} \quad \text{and} \quad S_1^{-1} B_1 = \begin{pmatrix} B_{11} \\ 0 \end{pmatrix}$$

and pick $S := \begin{pmatrix} S_1 & 0 \\ 0 & I \end{pmatrix} \in GL(n)$. Letting $Q := TS$, the following useful decomposition results:

Lemma 6.5.1 For any triple (A, B, C) there exists some $Q \in GL(n)$ so that

$$Q^{-1}AQ = \begin{pmatrix} A_{11} & A_{12} & 0 \\ 0 & A_{22} & 0 \\ A_{31} & A_{32} & A_{33} \end{pmatrix}$$

as well as

$$Q^{-1}B = \begin{pmatrix} B_{11} \\ 0 \\ B_{31} \end{pmatrix} \quad \text{and} \quad CQ = (C_{11} \quad C_{12} \quad 0)$$

---

for some matrices $A_{11}, \ldots$, and the triple

$$(A_{11}, B_{11}, C_{11})$$

is controllable and observable.

Proof. The pair (A_{11}, B_{11}) is controllable by construction, so it only remains to establish observability. Since

$$\mathbf{O}_r\left(\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}, (C_{11} \quad C_{12}) \right) = (\mathbf{O}_r(A_{11}, C_{11}) \quad *),$$

one only needs to show that the matrix on the left has rank r. By construction, however, the pair there is similar to the pair (A_1, C_1), which was already known to be observable, so the result follows. ∎

Lemma/Exercise 6.5.2 With the notations in Lemma 6.5.1,

$$CA^iB = C_{11}A_{11}^i B_{11}$$

for all $i = 0, 1, 2, \ldots$.   □

Definition 6.5.3 *The triple (A, B, C) is* **canonical** *if it is controllable and observable.*

For instance, the triple (A_{11}, B_{11}, C_{11}) in the above construction is canonical. One also calls a finite dimensional time-invariant discrete-time or continuous-time system Σ a *canonical system* if the corresponding triple is.

A sequence

$$\mathcal{A} = \{ \mathcal{A}_i, \ i = 1, 2, \ldots \},$$

where $\mathcal{A}_i \in \mathbb{K}^{p \times m}$, $i = 1, 2, \ldots$ for some fixed p and m, is a **Markov sequence**, and the $\mathcal{A}_i$'s are its **Markov parameters**. From Lemmas 2.4.6 and 2.7.13, the problem of realizing a discrete-time time-invariant linear behavior or a continuous-time time-invariant integral behavior is equivalent to the algebraic problem of factoring a Markov sequence as in (2.18), that is

$$\mathcal{A}_i = CA^{i-1}B \quad \text{for all } i > 0.$$

We say that a triple (A, B, C) for which this factorization holds is a *realization* of $\mathcal{A}$; if there is a realization, $\mathcal{A}$ is *realizable*.

From Lemmas 6.5.1 and 6.5.2 it follows that: □

Corollary 6.5.4 The Markov sequence $\mathcal{A}$ is realizable if and only if there exists a canonical triple realizing it.

Thus, for any (continuous-time or discrete-time finite dimensional) time-invariant linear system Σ, there exists a canonical system Σ_c having the same behavior as Σ. Next we study when a given sequence $\mathcal{A}$ is realizable, and we show that canonical realizations are essentially unique.

Realizability

Given a Markov sequence $\mathcal{A}$ and any pair of positive integers s, t, the (s,t)-th block **Hankel Matrix** associated to $\mathcal{A}$ is the matrix over $\mathbb{K}$:

$$\mathcal{H}_{s,t}(\mathcal{A}) := \begin{pmatrix} \mathcal{A}_1 & \mathcal{A}_2 & \cdots & \mathcal{A}_t \\ \mathcal{A}_2 & \mathcal{A}_3 & \cdots & \mathcal{A}_{t+1} \\ \vdots & \vdots & \cdots & \vdots \\ \mathcal{A}_s & \mathcal{A}_{s+1} & \cdots & \mathcal{A}_{s+t-1} \end{pmatrix} \quad (6.18)$$

of size $ps \times mt$ made up of st blocks, whose i,jth block is $\mathcal{A}_{i+j-1}$.

For any triple (A,B,C) and each k we consider the matrices $\mathbf{O}_k$ as in equation (6.6) as well as the reachability matrices

$$\mathbf{R}_k(A,B) = [B, AB, \ldots, A^{k-1}B]$$

(so that $\mathbf{R}_n = \mathbf{R}$). The main relations between these matrices and Hankel matrices are as follows:

Lemma 6.5.5 The triple (A,B,C) realizes $\mathcal{A}$ if and only if

$$\mathbf{O}_s(A,C)\,\mathbf{R}_t(A,B) = \mathcal{H}_{s,t}(\mathcal{A}) \quad (6.19)$$

for all s and t. □

Corollary 6.5.6 If (A,B,C) realizes $\mathcal{A}$, then

$$\operatorname{rank} \mathcal{H}_{s,t}(\mathcal{A}) \leq \max\{\operatorname{rank} \mathbf{O}_s(A,C), \operatorname{rank} \mathbf{R}_t(A,B)\} \leq n$$

for all s, t. □

Note that if the sequence $\mathcal{A}$ is realizable by the triple (A,B,C), then also

$$\mathbf{O}_s(A,C)\,A\mathbf{R}_t(A,B) = \mathcal{H}^1_{s+1,t}, \quad (6.20)$$

the submatrix of $\mathcal{H}_{s+1,t}$ formed by dropping the first block row.

A triple is canonical precisely if both $\mathbf{O}(A,C) = \mathbf{O}_n(A,C)$ and $\mathbf{R}(A,B)$ have rank n. In that case, there exist two matrices $\mathbf{O}^\#(A,C)$ and $\mathbf{R}^\#(A,B)$ such that

$$\mathbf{O}^\#(A,C)\,\mathbf{O}(A,C) = \mathbf{R}(A,B)\,\mathbf{R}^\#(A,B) = I. \quad (6.21)$$

We use the notation "#" to emphasize that in the particular case of $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ one could use the respective pseudoinverses, though many other one-sided inverses exist.

Corollary 6.5.7 If (A,B,C) realizes $\mathcal{A}$ and (A,B,C) is a canonical triple, then

$$\operatorname{rank} \mathcal{H}_{s,t}(\mathcal{A}) = n$$

whenever $s, t \geq n$.

Proof. Using one-sided inverses satisfying (6.21), $\mathbf{O}(A,C)\mathcal{H}_{n,n}(\mathcal{A})\mathbf{R}(A,B) = \mathcal{H}_{n,n}(\mathcal{A})$ implies that

$$I = \mathbf{O}^\#(A,C)\mathcal{H}_{n,n}(\mathcal{A})\mathbf{R}^\#(A,B),$$

so $\mathcal{H}_{n,n}(\mathcal{A})$ must have rank $\geq n$, and the conclusion follows. ∎

Definition 6.5.8 Let (A,B,C) and $(\widetilde{A}, \widetilde{B}, \widetilde{C})$ be two triples in $S_{n,m,p}$. Then (A,B,C) **is similar to** $(\widetilde{A}, \widetilde{B}, \widetilde{C})$, denoted

$$(A,B,C) \sim (\widetilde{A}, \widetilde{B}, \widetilde{C})$$

if

$$T^{-1}AT = \widetilde{A}, \quad T^{-1}B = \widetilde{B}, \quad \text{and} \quad CT = \widetilde{C}$$

for some $T \in GL(n)$. □

This is an equivalence relation, and the Markov sequences realized by any two similar triples are the same. The main result given below implies that, under a minimality assumption, the converse also holds; namely, if two minimal triples realize the same Markov sequence, then they must be similar.

Equivalence corresponds to a change of variables in the state space. Note that Lemma 6.5.1 provides a particularly important decomposition under similarity.

Definition 6.5.9 The n-dimensional triple (A,B,C) is **minimal** if any other triple $(\widetilde{A}, \widetilde{B}, \widetilde{C})$ realizing the same Markov sequence $\mathcal{A}$ must have dimension at least n. □

Minimality is in principle hard to check, since it involves comparisons with all other possible realizations of the same Markov sequence. The next result shows that this property is equivalent to being canonical and hence can be checked directly in terms of the data describing the system; we use the statement to summarize all relevant properties of realizations.

Theorem 27 Assume that $\mathcal{A}$ is a Markov sequence. Then the following properties hold:

1. If there is a realization of $\mathcal{A}$, then there is also a canonical realization.
2. A realization of $\mathcal{A}$ is minimal if and only if it is canonical.
3. Any two minimal realizations of $\mathcal{A}$ must be similar.

Proof. Statement 1 was given earlier as Corollary 6.5.4. We now prove statement 2.

Let (A,B,C) be a canonical triple of dimension n. By Corollary 6.5.4, $\operatorname{rank} \mathcal{H}_{n,n}(\mathcal{A}) = n$. It follows from Corollary 6.5.6 that every other realization has dimension at least n, so the triple is minimal.

On the other hand, if (A, B, C) is minimal, then it must be canonical, since otherwise from Lemma 6.5.2 and the construction in Lemma 6.5.1 it would follow that the triple (A_{11}, B_{11}, C_{11}) has lower dimension and realizes the same Markov sequence.

Next we establish statement 3. Assume that (A, B, C) and $(\widetilde{A}, \widetilde{B}, \widetilde{C})$ both realize $\mathcal{A}$ and are minimal, so in particular they must have the same dimension. Denote

$$\mathbf{R} = \mathbf{R}(A, B), \ \widetilde{\mathbf{R}} = \mathbf{R}(\widetilde{A}, \widetilde{B}), \ \mathbf{O} = \mathbf{O}(A, C), \ \widetilde{\mathbf{O}} = \mathbf{O}(\widetilde{A}, \widetilde{C})$$

and note that

$$\mathbf{OR} = \widetilde{\mathbf{O}}\widetilde{\mathbf{R}} \quad \text{and} \quad \mathbf{OAR} = \widetilde{\mathbf{O}}\widetilde{A}\widetilde{\mathbf{R}}$$

because of Equations (6.19) and (6.20) applied with $s = t = n$. Let $\widetilde{\mathbf{R}}^\#, \ldots,$ be one-sided inverses as in Equation (6.21), for each of the triples. Then with

$$T := \widetilde{\mathbf{R}}\mathbf{R}^\# = \left(\widetilde{\mathbf{O}}^\#\mathbf{O}\right)^{-1}$$

it follows that $T^{-1}AT = \widetilde{A}$. Applying Equation (6.19) with $s = n, t = 1$, and observing that $\mathbf{R}_1 = B, \widetilde{\mathbf{R}}_1 = \widetilde{B}$, we also obtain that $T^{-1}B = \widetilde{B}$. Finally, applying Equation (6.19) with $s = 1, t = n$ we obtain the remaining equality $CT = \widetilde{C}$. ∎

In terms of realizations of discrete-time time-invariant linear behaviors and continuous-time time-invariant integral behaviors, the Theorem asserts that canonical realizations exist if the behavior is realizable, and they are unique up to a change of basis in the state space.

Remark 6.5.10 If two controllable triples are similar, then the similarity must be given by the formulas obtained in the proof of Theorem 27, so in particular similarities between canonical systems are unique. This is because if T is as in Definition 6.5.8, then necessarily

$$T^{-1}A^iB = \widetilde{A}^i\widetilde{B}$$

for all i, so it must hold that $T^{-1}\mathbf{R} = \widetilde{\mathbf{R}}$, and therefore

$$T = \widetilde{\mathbf{R}}\mathbf{R}^\#.$$

In particular, this means that the only similarity between a canonical system and itself is the identity. In the terminology of group theory, this says that the action of $GL(n)$ on triples is free.

Theorem 27 leaves open the question of deciding when a realization exists; this is addressed next.

Definition 6.5.11 The rank of the Markov sequence $\mathcal{A}$ is

$$\sup_{s,t} \operatorname{rank} \mathcal{H}_{s,t}(\mathcal{A}).$$

Remark 6.5.12 In terms of the infinite Hankel matrix which is expressed in block form as

$$\mathcal{H}(\mathcal{A}) = \begin{pmatrix} A_1 & A_2 & \cdots & A_t & \cdots \\ A_2 & A_3 & \cdots & A_{t+1} & \cdots \\ \vdots & \vdots & & \vdots & \cdots \\ A_s & A_{s+1} & \cdots & A_{s+t-1} & \cdots \\ \vdots & \vdots & \cdots & \vdots & \cdots \end{pmatrix}$$

one may restate the definition of rank as follows.

The rank of an infinite matrix such as $\mathcal{H}(\mathcal{A})$ is, by definition, the dimension of the column space of $\mathcal{H}(\mathcal{A})$, which is seen as a subspace of the space $\mathbb{K}^\infty$ consisting of infinite column vectors $(x_1, x_2, x_3, \ldots)'$ with entries over $\mathbb{K}$, with pointwise operations.

When this rank is finite and less than n, all columns are linear combinations of at most $n - 1$ columns, and therefore all submatrices of $\mathcal{H}(\mathcal{A})$ must have rank less than n; this implies that $\mathcal{A}$ must have rank less than n. Conversely, we claim that if $\mathcal{A}$ has rank less than n, then the rank of the infinite matrix $\mathcal{H}(\mathcal{A})$ is less than n. If this were not to be the case, there would be a set of n independent columns $c_1, \ldots, c_n$ of $\mathcal{H}(\mathcal{A})$. If so, let M_i be the $i \times n$ matrix obtained by truncating each column c_j at the first i rows. Then some M_i has rank n: Consider the nonincreasing sequence of subspaces

$$Q_i := \ker M_i \subseteq \mathbb{K}^n$$

and let $Q := \cap Q_i$; by dimensionality, there is some k so that $Q_k = Q$, and if $x \in Q$ then $M_i x = 0$ for all i means that x is a linear relation between the c_j's, so $x = 0$. This means that $Q_k = 0$ and so M_k has rank n as desired. Let s, t be such that M_k is a submatrix of $\mathcal{H}_{s,t}$. Then

$$\operatorname{rank} \mathcal{A} \geq \operatorname{rank} \mathcal{H}_{s,t} \geq \operatorname{rank} M_k = n,$$

contradicting rank $\mathcal{A} < n$.

The conclusion is that the rank of $\mathcal{A}$ is equal to the rank (possibly infinite) of $\mathcal{H}(\mathcal{A})$. □

Theorem 28 The Markov sequence $\mathcal{A}$ is realizable if and only if it has finite rank n. In addition, if this holds, then:

(a) There is a canonical realization of dimension n.

(b) rank $\mathcal{H}_{n,n} = n$.

Proof. If there is a realization of dimension k, then it follows from Corollary 6.5.6 that $\mathcal{A}$ has rank at most k, and in particular the necessity statement is obtained. We show next that if $\mathcal{A}$ has rank n, then there is some rank n canonical

realization; this will establish property (a) as well as sufficiency. Property (b) is then a consequence of Corollary 6.5.7.

For the construction of the canonical realization, it is convenient to first generalize the notion of realizability to allow for infinite dimensional triples. In general, we consider objects $(\mathcal{X}, A, B, C)$ consisting of a vector space $\mathcal{X}$ and linear maps $A : \mathcal{X} \to \mathcal{X}$, $B : \mathbb{K}^m \to \mathcal{X}$, and $C : \mathcal{X} \to \mathbb{K}^p$, and say that $\mathcal{A}$ is realizable by $(\mathcal{X}, A, B, C)$ if

$$C \circ A^{i-1} \circ B = \mathcal{A}_i$$

for all i. These objects can be identified to discrete-time time-invariant linear systems; when $\mathcal{X}$ is finite dimensional, choosing a basis on $\mathcal{X}$ provides a triple realizing $\mathcal{A}$.

Now given any Markov sequence $\mathcal{A}$, we let $\mathcal{X}_0$ denote the space $\mathbb{K}^\infty$ introduced in Remark 6.5.12. Note that the *shift operator*

$$\sigma : \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \end{pmatrix} \mapsto \begin{pmatrix} x_2 \\ x_3 \\ x_4 \\ \vdots \end{pmatrix}$$

is linear on $\mathcal{X}_0$. Let A be defined as σ^p, the shift by p positions. Let $B : \mathbb{K}^m \to \mathcal{X}_0$ be defined on the natural basis as follows:

$$Be_j := j\text{th column of } \mathcal{H}(\mathcal{A})$$

and let $C : \mathcal{X}_0 \to \mathbb{K}^p$ be the projection on the first p coordinates.

We claim that $(\mathcal{X}_0, A, B, C)$ is a realization of $\mathcal{A}$. Consider any $i \geq 0$ and any $j = 1, \ldots, m$. Then $CA^i Be_j$ is the vector consisting of the entries in positions $ip + 1, \ldots, ip + p$ of the jth column of $\mathcal{H}(\mathcal{A})$, that is, the jth column of $\mathcal{A}_{i+1}$, as wanted.

The abstract realization just obtained is observable, in the sense that

$$\bigcap_{i=0}^{\infty} \ker CA^i = 0,$$

since for each i the elements of $\ker CA^i$ are precisely those vectors whose entries in positions $ip + 1, \ldots, ip + p$ all vanish. The desired canonical realization will be obtained by restricting to an A-invariant subspace of $\mathcal{X}_0$; this restriction will also be observable, since the kernels of the restrictions still must intersect at zero.

Let $\mathcal{X}$ be the subspace of $\mathcal{X}_0$ spanned by all the iterates $A^i Be_j$, over all i, j. This is A-invariant by definition; we denote the restriction of A to $\mathcal{X}$ again as A. Moreover, $\mathcal{X}$ contains the image of B, so we may consider B as a map from $\mathbb{K}^m$ into $\mathcal{X}$; similarly, we restrict C to $\mathcal{X}$ and denote it also by C. We claim next that this space equals the column space of $\mathcal{H}(\mathcal{A})$. Here the Hankel pattern

becomes essential: Every column of $\mathcal{H}(\mathcal{A})$ is of the form $A^i Be_j$ for some i and j; namely, the $im + j$th column ($1 \leq j \leq m$) is the same as $A^i Be_j$, for each $i \geq 0$.

We conclude that $(\mathcal{X}, A, B, C)$ realizes $\mathcal{A}$, where $\mathcal{X}$ is the column space of $\mathcal{H}(\mathcal{A})$ and therefore is n-dimensional. ∎

The criterion in Theorem 28 is of interest because it allows us to check realizability, in the sense that one may look at Hankel matrices of increasing size and if a realization exists it will eventually be found. If a realization does not exist, however, there is no way of knowing so by means of this technique.

If one is interested merely in investigating the existence of realizations, as opposed to studying minimality, the problem can be reduced to the scalar case ($m = p = 1$), which in turn belongs to the classical theory of linear difference equations:

Exercise 6.5.13 Given any Markov sequence $\mathcal{A}$, introduce the *pm* sequences corresponding to each coordinate, that is,

$$\mathcal{A}^{ij} := (\mathcal{A}_1)_{ij}, (\mathcal{A}_2)_{ij}, (\mathcal{A}_3)_{ij}, \ldots$$

for each $i = 1, \ldots, p$ and $j = 1, \ldots, m$. Show, not using any of the results just developed, that $\mathcal{A}$ is realizable if and only if each $\mathcal{A}^{ij}$ is. □

Exercise 6.5.14 Calculate a canonical realization, and separately calculate the rank of the Hankel matrix, for each of these examples with $m = p = 1$:

1. The sequence of natural numbers $1, 2, 3, \ldots$.
2. The Fibonacci sequence $0, 1, 1, 2, 3, 5, 8, \ldots$. □

6.6 Recursion and Partial Realization

An alternative characterization of realizability is through the concept of recursive sequences. We shall say that $\mathcal{A}$ is a *recursive Markov sequence* if there exist a positive integer n and scalars $\alpha_1, \ldots, \alpha_n$ such that

$$\mathcal{A}_{k+n+1} = \alpha_n \mathcal{A}_{k+n} + \alpha_{n-1}\mathcal{A}_{k+n-1} + \ldots + \alpha_2 \mathcal{A}_{k+2} + \alpha_1 \mathcal{A}_{k+1} \qquad (6.22)$$

for all $k \geq 0$; in this case $\mathcal{A}$ is said to satisfy a recursion of order n.

Proposition 6.6.1 The Markov sequence $\mathcal{A}$ is realizable if and only if it is recursive.

Proof. If (A, B, C) realizes $\mathcal{A}$, let

$$\chi_A(s) = s^n - \alpha_n s^{n-1} - \ldots - \alpha_1$$

so that by the Cayley-Hamilton Theorem

$$A^n = \alpha_n A^{n-1} + \ldots + \alpha_1 I.$$

Multiplying this last equation by CA^k on the left and by B on the right, the recursion (6.22) results. Conversely, if there is a recursion of order n, all columns of the infinite Hankel matrix must be linear combinations of the first nm columns (that is, the columns appearing in the first n blocks), so the matrix has finite rank and therefore $\mathcal{A}$ is realizable. ∎

Corollary 6.6.2 If $\mathcal{A}$ is a Markov sequence of finite rank n, then $\mathcal{A} \equiv 0$ if and only if

$$\mathcal{A}_1 = \mathcal{A}_2 = \ldots = \mathcal{A}_n = 0.$$

Proof. By Theorem 28 there exists a realization of dimension n, and therefore, by the proof of the Proposition, there is a recursion of order n. Recursively, all $\mathcal{A}_i = 0$. ∎

Corollary 6.6.3 If $\mathcal{A}^1$ and $\mathcal{A}^2$ are two Markov sequences of finite ranks n_1 and n_2, respectively, then $\mathcal{A}^1 = \mathcal{A}^2$ if and only if

$$\mathcal{A}_i^1 = \mathcal{A}_i^2, \ i = 1, \ldots, n_1 + n_2.$$

Proof. The sequence $\mathcal{A} := \mathcal{A}^1 - \mathcal{A}^2 = \{\mathcal{A}_i^1 - \mathcal{A}_i^2\}$ has rank at most $n := n_1 + n_2$, as follows by considering its realization:

$$\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \begin{pmatrix} C_1 & -C_2 \end{pmatrix}$$

where (A_i, B_i, C_i) is a realization of $\mathcal{A}^i$ of dimension n_i. Then Corollary 6.6.2 gives the result. ∎

Thus, sequences of rank $\leq n$ are uniquely determined by their first $2n$ Markov parameters.

There is a direct construction of a realization from the recursion (6.22) and the Markov sequence. This is as follows. Let $\mathcal{H}_{n+k,n}^k$ be the matrix obtained by dropping the first k blocks of rows from $\mathcal{H}_{n+k,n}(\mathcal{A})$, that is, the matrix

$$\begin{pmatrix} \mathcal{A}_{k+1} & \cdots & \mathcal{A}_{k+n} \\ \vdots & \cdots & \vdots \\ \mathcal{A}_{k+n} & \cdots & \mathcal{A}_{k+2n-1} \end{pmatrix}$$

including the case $k = 0$, where this is just $\mathcal{H}_{n,n}(\mathcal{A})$. Then (6.22) implies

$$A \mathcal{H}_{n+k,n}^k = \mathcal{H}_{n+k+1,n}^{k+1}$$ (6.23)

for all $k \geq 0$, where

$$A := \begin{pmatrix} 0 & I & 0 & \cdots & 0 \\ 0 & 0 & I & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I \\ \alpha_1 I & \alpha_2 I & \alpha_3 I & \cdots & \alpha_n I \end{pmatrix}.$$ (6.24)

analogously to the controller form in Definition 5.1.5, except that this is now a block matrix. Each block has size $p \times p$, and the matrix A is of size $np \times np$. Equation (6.23) implies that

$$A^i \mathcal{H}_{n,n}(\mathcal{A}) = \mathcal{H}_{n+i,n}^i$$

for all $i \geq 0$, and so

$$\begin{pmatrix} I & 0 & \cdots & 0 \end{pmatrix} A^i \mathcal{H}_{n,n}(\mathcal{A}) \begin{pmatrix} I \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \mathcal{A}_{i+1}$$ (6.25)

for all $i \geq 0$. Letting

$$C := \begin{pmatrix} I & 0 & \cdots & 0 \end{pmatrix}$$ (6.26)

and

$$B := \mathcal{H}_{n,n}(\mathcal{A}) \begin{pmatrix} I \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \mathcal{A}_1 \\ \mathcal{A}_2 \\ \vdots \\ \mathcal{A}_n \end{pmatrix},$$ (6.27)

there results a realization of $\mathcal{A}$ of dimension np. Since $\mathbf{O}_n(A,C) = I$:

Lemma 6.6.4 The system (A, B, C) given by (6.24), (6.27), and (6.26) is an observable realization of $\mathcal{A}$.

This is often called an **observability form realization** of $\mathcal{A}$.

Remark 6.6.5 Let $\mathcal{A}$ be any recursive Markov sequence. Consider the transposed sequence

$$\mathcal{A}' := \mathcal{A}_1', \mathcal{A}_2', \ldots .$$

This satisfies a recursion with the same coefficients α_i's. For these coefficients we let (A, B, C) be the observability form realization of $\mathcal{A}'$. Since $CA^{i-1}B = \mathcal{A}_i'$ for all i, also $B'(A')^{i-1}C' = \mathcal{A}_i$ for all i, and the system (A', C', B') is a controllable realization of $\mathcal{A}$. We have obtained the **controllability form realization** of $\mathcal{A}$:

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & \alpha_1 I \\ I & 0 & \cdots & 0 & \alpha_2 I \\ 0 & I & \cdots & 0 & \alpha_3 I \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & \alpha_n I \end{pmatrix} \quad B = \begin{pmatrix} I \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad C = \begin{pmatrix} \mathcal{A}_1 & \mathcal{A}_2 & \cdots & \mathcal{A}_n \end{pmatrix}$$

of dimension mn.

$\square$

Corollary 6.6.6 If $p=1$ or $m=1$, the minimal possible order of a recursion equals the rank of $\mathcal{A}$. Furthermore, there is a unique such minimal order recursion, and its coefficients are those of the (common) characteristic polynomial of $\mathcal{A}$.

Proof. The proof of Proposition 6.6.1 shows that, for arbitary n,p, if there is a realization of dimension n, then there is a recursion of order n. Conversely, in the cases $p=1$ or $m=1$ there is always a realization of dimension equal to the order of any given recursion, namely the observability form or the controllability form realization, respectively. The second assertion follows from the constructions. ∎

Remark 6.6.7 Observability form realizations are in general not controllable, except if $p=1$. In this case, if $(\alpha_1,\ldots,\alpha_n)$ give a minimal recursion then by Corollary 6.6.6 this realization must be minimal. In fact,

$$\mathbf{R}_n(A,B) = \mathcal{H}_{n,n},$$

which has rank n. Of course, for $p>1$ it is possible to reduce any such realization to a controllable one using the Kalman decomposition. $\square$

Exercise 6.6.8 Show by providing a counterexample that the hypothesis that either $p=1$ or $m=1$ cannot be dropped in the first assertion on Corollary 6.6.6. $\square$

Given any finite sequence $\mathcal{A}_1,\mathcal{A}_2,\ldots,\mathcal{A}_r$, this is always part of a realizable Markov sequence, since

$$\mathcal{A}_1,\mathcal{A}_2,\ldots,\mathcal{A}_r,0,0,\ldots,0,\ldots$$

has finite rank. More interesting is the following fact. Given any finite sequence $\mathcal{A}_1,\mathcal{A}_2,\ldots,\mathcal{A}_r$, whenever $s+t\le r+1$ we can consider the Hankel matrices $\mathcal{H}_{s,t}$ obtained using formula (6.18).

Lemma 6.6.9 Assume that $\mathcal{A}_1,\mathcal{A}_2,\ldots,\mathcal{A}_{2n}$ are $2n$ matrices in $\mathbb{K}^{p\times m}$ for which

$$\text{rank}\,\mathcal{H}_{n,n} = \text{rank}\,\mathcal{H}_{n+1,n} = \text{rank}\,\mathcal{H}_{n,n+1} = n. \qquad (6.28)$$

Then there exists a (unique) Markov sequence $\mathcal{A}$ of rank n whose first $2n$ Markov parameters are $\mathcal{A}_1,\mathcal{A}_2,\ldots,\mathcal{A}_{2n}$.

Proof. From the equality $\text{rank}\,\mathcal{H}_{n,n} = \text{rank}\,\mathcal{H}_{n+1,n}$ it follows that the last p rows of $\mathcal{H}_{n+1,n}$ must be linear combinations of the rows of $\mathcal{H}_{n,n}$. This means that there must exist $p\times p$ matrices

$$C_i,\ i=1,\ldots,n$$

such that

$$\mathcal{A}_j = C_1\mathcal{A}_{j-1} + \ldots + C_n\mathcal{A}_{j-n} \qquad (6.29)$$

for each $j=n+1,\ldots,2n$. Similarly, $\text{rank}\,\mathcal{H}_{n,n} = \text{rank}\,\mathcal{H}_{n,n+1}$ implies that the last m columns of $\mathcal{H}_{n,n+1}$ must be linear combinations of the columns of $\mathcal{H}_{n,n}$. So there must also exist $m\times m$ matrices

$$D_i,\ i=1,\ldots,n$$

such that

$$\mathcal{A}_j = \mathcal{A}_{j-1}D_1 + \ldots + \mathcal{A}_{j-n}D_n \qquad (6.30)$$

for each $j=n+1,\ldots,2n$.

We now *define* $\mathcal{A}_j$ for $j>2n$ recursively using the formula (6.29), and let $\mathcal{A}$ be the Markov sequence so obtained. It follows from this definition that all rows of $\mathcal{H}(\mathcal{A})$ are linearly dependent on the first pn rows. So $\mathcal{A}$ has rank at most pn; we next show that its rank is in fact just n. For this, it is enough to establish that (6.30) holds for all $j>2n$ as well, since this will then imply that all columns depend linearly on the first nm columns, and therefore that the rank of $\mathcal{H}(\mathcal{A})$ is the same as the rank of $\mathcal{H}_{n,n}$, which is n. By induction on j,

$$\begin{aligned}
\mathcal{A}_{j+1} &= \sum_{i=1}^n C_i\mathcal{A}_{j+1-i} \\
&= \sum_{i=1}^n C_i \sum_{l=1}^n \mathcal{A}_{j+1-i-l}D_l \\
&= \sum_{l=1}^n \left(\sum_{i=1}^n C_i\mathcal{A}_{j+1-i-l}\right) D_l \\
&= \sum_{l=1}^n \mathcal{A}_{j+1-l}D_l,
\end{aligned}$$

as desired.

Uniqueness follows from the fact that any other extension would agree in its first $2n$ parameters and hence would have to be the same, because of Corollary 6.6.3. ∎

The previous result can be used to give an explicit description of the quotient space obtained when identifying triples up to similarity. Let $\mathcal{M}_{n,m,p}$ denote the set

$$\{(\mathcal{A}_1,\ldots,\mathcal{A}_{2n}) \mid \text{rank}\,\mathcal{H}_{n,n} = \text{rank}\,\mathcal{H}_{n+1,n} = \text{rank}\,\mathcal{H}_{n,n+1} = n\}.$$

Note that this can be thought of as a subset of $\mathbb{K}^{2nmp}$, and as such it is a set defined by polynomial equalities and inequalities. Let

$$\beta : S^{c,o}_{n,m,p} \to \mathcal{M}_{n,m,p}$$

be the "behavior" function

$$(A,B,C) \to (CB, \ldots, CA^{2n-1}B).$$

That β indeed maps into $\mathcal{M}_{n,m,p}$ follows from Theorem 28 (p. 288). Moreover, from Lemma 6.6.9 we know that β is onto. From the uniqueness result, we also know that, if

$$\beta(A,B,C) = \beta(\widetilde{A},\widetilde{B},\widetilde{C}),$$

then (A,B,C) and $(\widetilde{A},\widetilde{B},\widetilde{C})$ are two canonical triples realizing the same Markov sequence and so, by Theorem 27 (p. 286), Part 3, that these two triples are similar. This discussion can be summarized as follows.

Corollary 6.6.10 The map β induces a bijection between the quotient space $S_{n,m,p}^{c,o}/\sim$ and $\mathcal{M}_{n,m,p}$. □

Exercise 6.6.11 Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Let $\mathcal{M}_{n,m,p}$ have the topology induced by $\mathbb{K}^{2nmp}$, and let $S_{n,m,p}^{c,o}/\sim$ be given the quotient topology, when $S_{n,m,p}^{c,o}$ is thought of as a subspace of $\mathbb{K}^{n(n+p+m)}$. Show that β induces a homeomorphism on the quotient space. (*Hint:* Establish first that the realization given in the proof of Theorem 28 can be made to be locally continuous on the sequence. This is done by observing that, by Cramer's rule after choosing a set of n linearly independent columns of $\mathcal{H}_{n,n}$, the entries of the operator A can be taken to be rational on the coefficients of the Markov parameters. This provides a covering of $\mathcal{M}_{n,m,p}$ by open subsets V and continuous mappings $\rho : V \to S_{n,m,p}^{c,o}$ such that $\beta\rho$ is the identity.)

Remark 6.6.12 (This remark uses some concepts from differential geometry.) In the case when $\mathbb{K} = \mathbb{R}$ (or the complex case) one can show that the quotient space $S_{n,m,p}^{c,o}/\sim$ has a differentiable manifold structure under which the natural projection is smooth. This fact can be established constructively, by exhibiting an explicit set of charts for this quotient manifold, or one can use a general theorem on group actions, the quicker path which we choose here. Consider the action of $GL(n)$ on $S_{n,m,p}^{c,o}$,

$$T.(A,B,C) := (T^{-1}AT, T^{-1}B, CT),$$

seen as a smooth action of a Lie group on a manifold ($S_{n,m,p}^{c,o}$ is an open subset of $\mathbb{R}^{n(n+m+p)}$). This action is free (see Remark 6.5.10). According to Proposition 4.1.23 in [2], the quotient will have a differentiable manifold structure for which the quotient mapping $S_{n,m,p}^{c,o} \to S_{n,m,p}^{c,o}/\sim$ is a smooth submersion, provided that the graph of the similarity relation is closed and that the action is proper. Moreover, in this case the natural map

$$S_{n,m,p}^{c,o} \to S_{n,m,p}^{c,o}/\sim$$

defines a principal fibre bundle (same reference, Exercise 4.1M). Properness of the action means that the following property must hold: Whenever $\{\Sigma_i\}$ is a

convergent sequence of triples and $\{T_i\}$ is a sequence of elements of $GL(n)$ for which $\{T_i.\Sigma_i\}$ is a convergent sequence of triples, the sequence $\{T_i\}$ must have a convergent subsequence. So we must prove that the action is closed and proper.

Assume that the sequences $\{\Sigma_i\}$ and $\{T_i\}$ are as in the above paragraph, and let $\Sigma'_i := T_i.\Sigma_i$ for each i. By assumption, there are systems Σ and Σ' so that

$$\Sigma_i \to \Sigma \quad \text{and} \quad \Sigma'_i \to \Sigma'.$$

We use primes and subscripts to denote the matrices A, B, C for the various triples. The triples Σ and Σ' must realize the same Markov sequence, since this is true of the corresponding pairs (Σ_i, Σ'_i) and the sequence's elements depend continuously on the triple. Therefore the matrix

$$T = \mathbf{R}(\mathbf{R}')^{\#}$$

provides a similarity between these two triples. Observe that $\mathbf{R}_{(i)}$ (the n-step reachability matrix for the triple Σ_i) converges to $\mathbf{R}$. Moreover, and this is the critical observation, one may also assume that $(\mathbf{R}'_{(i)})^{\#}$, the one-sided inverse of the n-step reachability matrix for the triple Σ'_i, also converges to $(\mathbf{R}')^{\#}$. The reason for this latter fact is that one may *pick* such a one-sided inverse continuously about any given system: Just use Cramer's rule after choosing a set of n linearly independent columns of $\mathbf{R}'$ (these columns remain linearly independent for triples near the triple Σ'). We conclude that

$$T_i = \mathbf{R}_{(i)}(\mathbf{R}'_{(i)})^{\#} \to T$$

because of uniqueness of the similarity between two minimal systems (Remark 6.6.11). This establishes properness. (In fact, we proved that the sequence T_i itself is convergent, rather than merely a subsequence.)

The proof of closeness is even easier. We need to see that, if

$$\Sigma_i \sim \Sigma'_i \quad \text{for all } i$$

and

$$\Sigma_i \to \Sigma, \quad \Sigma'_i \to \Sigma',$$

then necessarily Σ and Σ' are similar. This is immediate from the uniqueness Theorem, because by continuity these two triples must give rise to the same Markov sequence. □

Corollary 6.6.10, Exercise 6.6.11, and Remark 6.6.12 provide only a very brief introduction to the topic of *families of systems* and more specifically to *moduli problems* for triples. Much more detailed results are known; For instance, $S_{n,m,p}^{c,o}/\sim$ has a natural structure of nonsingular algebraic variety consistent with the differentiable manifold structure given above, for $\mathbb{K} = \mathbb{R}, \mathbb{C}$.

6.7   Rationality and Realizability

We now characterize realizability of a Markov sequence in terms of the rationality of an associated power series.

Consider the set of all semi-infinite sequences

$$a_k, a_{k+1}, \ldots, a_0, a_1, a_2, a_3, \ldots, a_l, \ldots$$

formed out of elements of $\mathbb{K}$. The starting index k may be negative or positive but it is finite; however, the sequence may be infinite to the right. We think of these sequences as formal *Laurent series* in a variable s^{-1}:

$$\sum_{i=k}^{\infty} a_i s^{-i},$$

where $a_i \in \mathbb{K}$ for each i. If the sequence is not identically zero and k is the smallest integer so that $a_k \neq 0$, then the series has *order* k. If $k \geq 0$, this is just a formal power series in the variable s^{-1}. If all coefficients $a_i, i > 0$ vanish, this is simply a polynomial on s. Let $\mathcal{K}((s^{-1}))$ denote the set of all such Laurent series.

The set $\mathcal{K}((s^{-1}))$ has a natural vector space structure over $\mathbb{K}$, corresponding to coefficientwise operations:

$$\left(\sum a_i s^{-i}\right) + \left(\sum b_i s^{-i}\right) := \sum (a_i + b_i) s^{-i}, \qquad (6.31)$$

but the use of the power series notation also suggests the convolution product

$$\left(\sum a_i s^{-i}\right) \cdot \left(\sum b_i s^{-j}\right) := \sum c_l s^{-l}, \qquad (6.32)$$

where for each $l \in \mathbb{Z}$

$$c_l := \sum_{i+j=l} a_i b_j$$

is a finite sum because there are at most finitely many nonzero coefficients with negative i and j. With these operations, $\mathcal{K}((s^{-1}))$ forms a ring. It is an integral domain, that is, the product of two nonzero sequences cannot be zero, because if k_1 and k_2 are the orders of these sequences, then the product has order $k_1 + k_2$. (In fact, since $\mathbb{K}$ is a field, it can be shown that $\mathcal{K}((s^{-1}))$ is also a field, but this is not needed in what follows.)

More generally, we consider series whose coefficients are matrices of a fixed size:

$$\sum_{i=k}^{\infty} A_i s^{-i},$$

where $A_i \in \mathbb{K}^{p \times m}$ for all i. The sum and product given by formulas (6.31) and (6.32) are still well defined (assuming that sizes match), where now each A_i,

B_j and C_l is a matrix. The usual distributivity and associativity properties of matrix product hold. We let $\mathcal{K}^{p \times m}((s^{-1}))$ denote the set of all matrix Laurent series with fixed p and m.

A series $W(s) \in \mathcal{K}^{p \times m}((s^{-1}))$ will be said to be **rational** if there exist a monic polynomial

$$q(s) = s^n - \alpha_n s^{n-1} - \ldots - \alpha_1 \qquad (6.33)$$

and a matrix polynomial

$$P(s) = P_0 + P_1 s + \ldots + P_h s^h \qquad (6.34)$$

in $\mathbb{K}^{p \times m}[s]$ such that

$$qW = P$$

(which is also written as $W = q^{-1}P$, or P/q). For instance, any series having only finitely many terms is rational, since

$$A_k s^k + \ldots + A_0 + A_1 s^{-1} + \ldots + A_l s^{-l}$$

can be written as

$$\frac{1}{s^l}(A_k s^{k+l} + \ldots + A_0 s^l + A_1 s^{l-1} + \ldots + A_l).$$

For another example consider the scalar series

$$1 + s^{-1} + s^{-2} + \ldots + s^{-k} + \ldots,$$

which is rational since it equals $s/(s-1)$.

We associate to each Markov sequence $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \ldots)$ its *generating series*

$$W_{\mathcal{A}}(s) := \sum_{i=1}^{\infty} \mathcal{A}_i s^{-i}.$$

In terms of this we may state another criterion for realizability:

Proposition 6.7.1 A Markov sequence is realizable if and only if its generating series is rational.

Proof. Because of Proposition 6.6.1, we must simply show that $W_{\mathcal{A}}$ is rational if and only if $\mathcal{A}$ is recursive.

Note first that, in general, if $W = q^{-1}P$ is rational and has order ≥ 1 and q is as in (6.33), then the polynomial P in (6.34) can be taken to be of degree at most $n - 1$, since all terms in qW corresponding to s^h, $h \geq n$, must vanish. So rationality of $W_{\mathcal{A}}$ is equivalent to the existence of elements $\alpha_1, \ldots, \alpha_n$ in $\mathbb{K}$ and matrices $P_0, \ldots, P_{n-1}$ over $\mathbb{K}$ such that

$$(s^n - \alpha_n s^{n-1} - \ldots - \alpha_1)\left(\sum_{i=1}^{\infty} \mathcal{A}_i s^{-i}\right) = P_0 + P_1 s + \ldots + P_{n-1} s^{n-1}. \qquad (6.35)$$

If such an equation holds, then comparing coefficients of s^{-k-1} results in

$$\mathcal{A}_{k+n+1} = \alpha_n \mathcal{A}_{k+n} + \alpha_{n-1}\mathcal{A}_{k+n-1} + \cdots + \alpha_2 \mathcal{A}_{k+2} + \alpha_1 \mathcal{A}_{k+1}$$

for each $k \geq 0$, so the sequence is recursive. Conversely, if this equation holds one may just define the matrices P_i by

$$P_j := \mathcal{A}_{n-j} - \sum_{i=1}^{n-j-1} \alpha_{i+j+1}\mathcal{A}_i \quad j = 0,\ldots,n-1, \quad (6.36)$$

i.e., the equations that are derived from (6.35) by comparing the coefficients of $1, s, \ldots, s^{n-1}$.

The proof shows that the minimal degree of a possible denominator q equals the minimal order of a recursion satisfied by $\mathcal{A}$. ∎

Corollary 6.7.2 If $m = 1$ or $p = 1$ and $W_\mathcal{A} = P/q$ with q monic of degree equal to the rank of $\mathcal{A}$, then q is the (common) characteristic polynomial of the canonical realizations of $\mathcal{A}$.

Proof. We know from Corollary 6.6.6 that for m or p equal to 1 the coefficients of a minimal recursion are those of the characteristic polynomial. The q in the statement must be a denominator of minimal degree, since a lower degree polynomial would give rise to a lower order recursion and therefore to a lower dimensional realization. From the above construction, this polynomial then corresponds to a minimal recursion. ∎

Exercise 6.7.3 Show that, if $m = p = 1$ and if $W_\mathcal{A} = P/q$ with P of degree $\leq n-1$ and q of degree n, then $\mathcal{A}$ has rank n if and only if P and q are relatively prime. (*Hint:* Use Corollary 6.6.6 and the fact that $\mathcal{K}((s^{-1}))$ forms an integral domain.) □

Remark 6.7.4 One can show directly that realizability implies rationality, as follows. Elements of $\mathcal{K}^{p\times m}((s^{-1}))$ can be identified naturally with $p \times m$ matrices over the ring $\mathcal{K}((s^{-1}))$, and this identification preserves the convolution structure. Under this identification, rational elements are precisely those that correspond to matrices all whose entries are rational. In particular, the matrix $(sI - A)^{-1}$, the inverse over $\mathcal{K}((s^{-1}))^{n\times n}$, is rational. On the other hand, it holds that

$$(sI - A)^{-1} = \sum_{i=1}^{\infty} A^{i-1} s^{-i}$$

from which it follows that the series

$$C(sI - A)^{-1}B = \sum_{i=1}^{\infty} CA^{i-1}Bs^{-i} = \sum_{i=1}^{\infty} \mathcal{A}_i s^{-i}$$

is rational if realizability holds. The minimal dimension of a realization is also called the *McMillan degree* of the corresponding rational matrix. □

Remark 6.7.5 When $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, one may use complex variables techniques in order to study realizability. Take for simplicity the case $m = p = 1$ (otherwise one argues with each entry). If we can write $W_\mathcal{A}(s) = P(s)/q(s)$ with $\deg P < \deg q$, pick any positive real number λ that is greater than the magnitudes of all zeros of q. Then, $W_\mathcal{A}$ must be the Laurent expansion of the rational function P/q on the annulus $|s| > \lambda$. (This can be proved as follows: The formal equality $qW_\mathcal{A} = P$ implies that the Taylor series of P/q about $s = \infty$ equals $W_\mathcal{A}$, and the coefficients of this Taylor series are those of the Laurent series on $|s| > \lambda$. Equivalently, one could substitute $z := 1/s$ and let

$$\tilde{q}(z) := z^d q(1/z), \quad \tilde{P}(z) := z^d P(1/z),$$

with $d := \deg q(s)$; there results the equality $\tilde{q}(z)W(1/z) = \tilde{P}(z)$ of power series, with $\tilde{q}(0) \neq 0$, and this implies that $W(1/z)$ is the Taylor series of $\tilde{P}/\tilde{q}$ about 0. Observe that on any other annulus $\lambda_1 < |s| < \lambda_2$ where q has no roots the Laurent expansion will in general have terms in s^k, with $k > 0$, and will therefore be different from $W_\mathcal{A}$.) Thus, if there is any function g which is analytic on $|s| > \mu$ for some μ and is so that $W_\mathcal{A}$ is its Laurent expansion about $s = \infty$, realizability of $\mathcal{A}$ implies that g must be rational, since the Taylor expansion at infinity uniquely determines the function. Arguing in this manner it is easy to construct examples of nonrealizable Markov sequences. For instance,

$$\mathcal{A} = 1, \frac{1}{2}, \frac{1}{3!}, \frac{1}{4!}, \ldots$$

is unrealizable, since $W_\mathcal{A} = e^{1/s} - 1$ on $s \neq 0$. As another example, the sequence

$$\mathcal{A} = 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots,$$

cannot be realized because $W_\mathcal{A} = -\ln(1 - s^{-1})$ on $|s| > 1$. □

Input/Output Equations

Rationality and realizability can also be interpreted in terms of high-order differential or difference equations satisfied by input/output pairs corresponding to a given behavior.

The *input/output pairs* of the behavior Λ are by definition the possible pairs $(\omega, \bar{\lambda}(\omega))$. For each $\omega \in \mathcal{U}^{[\sigma,\tau]}$ in the domain of the behavior, this is a pair consisting of a function in $\mathcal{U}^{[\sigma,\tau]}$ and one in $\mathcal{Y}^{[\sigma,\tau]}$.

Assume from now on that Λ is a time-invariant continuous-time integral behavior with analytic kernel

$$K(t) = \sum_{k=0}^{\infty} \mathcal{A}_{k+1}\frac{t^k}{k!} \quad (6.37)$$

(assume K is entire, that is, the series converges for all t), where $\mathcal{A} = \mathcal{A}_1, \mathcal{A}_2, \ldots$ is a Markov sequence over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Observe that whenever ω is $r-1$-times

(continuously) differentiable the corresponding output function $\bar{\lambda}(\omega)$ is r-times differentiable; such a pair $(\omega, \bar{\lambda}(\omega))$ is of class $\mathcal{C}^r$.

The behavior Λ is said to satisfy the i/o equation

$$y^{(n)}(t) = \sum_{i=0}^{n-1} \alpha_{i+1} y^{(i)}(t) + \sum_{i=0}^{n-1} P_i u^{(i)}(t), \quad (6.38)$$

where $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ and $P_0, \ldots, P_{n-1} \in \mathbb{K}^{p \times m}$, if this equation is satisfied by every i/o pair $u = \omega, y = \omega = \bar{\lambda}(\omega)$ of class $\mathcal{C}^{n-1}$, for all $t \in [\sigma, \tau)$. The nonnegative integer n is called the order of the i/o equation. (For $n = 0$ the equation is $y(t) = 0$.)

It is a basic fact in linear system theory that the existence of an i/o equation is equivalent to rationality, and hence to realizability by a finite dimensional linear time-invariant system.

Proposition 6.7.6 The behavior Λ satisfies (6.38) if and only if $W_A = q^{-1}P$, where q and P are as in Equations (6.33) and (6.34), with the same α_i's and P_i's.

Proof. We first make some general observations. If K is as in (6.37), then

$$K^{(r)}(t) = \sum_{k=0}^{\infty} A_{k+r+1} \frac{t^k}{k!} \quad (6.39)$$

for each $r = 0, 1, \ldots$. If, in addition, A is known to satisfy the order-n recursion

$$A_{k+n+1} = \sum_{j=1}^{n} \alpha_j A_{k+j}$$

for all $k \geq 0$, then

$$
\begin{aligned}
K^{(n)}(t) &= \sum_{k=0}^{\infty} \sum_{j=1}^{n} \alpha_j A_{k+j} \frac{t^k}{k!} \\
&= \sum_{j=1}^{n} \alpha_j \sum_{k=0}^{\infty} A_{k+j} \frac{t^k}{k!} \\
&= \sum_{j=1}^{n} \alpha_j K^{(j-1)}(t)
\end{aligned}
$$

for all t.

On the other hand, for each i/o pair (ω, η) of class $\mathcal{C}^{n-1}$ one concludes by induction from

$$\eta(t) = \int_{\sigma}^{t} K(t - \mu)\omega(\mu) \, d\mu$$

that

$$\eta^{(r)}(t) = \sum_{i=1}^{r} A_i \omega^{(r-i)}(t) + \int_{\sigma}^{t} K^{(r)}(t-\mu)\omega(\mu) \, d\mu \quad (6.40)$$

for each $r = 0, \ldots, n$. In particular,

$$\eta^{(n)}(t) = \sum_{i=1}^{n} A_i \omega^{(n-i)}(t) + \sum_{j=1}^{n} \alpha_j \int_{\sigma}^{t} K^{(j-1)}(t-\mu)\omega(\mu) \, d\mu \quad (6.41)$$

if A satisfies the above recursion. Since also

$$\sum_{i=0}^{n-1} \alpha_{i+1}\eta^{(i)}(t) = \sum_{i=0}^{n-1} \alpha_{i+1} \left[\sum_{l=1}^{i} A_l \omega^{(i-l)}(t) + \int_{\sigma}^{t} K^{(i)}(t-\mu)\omega(\mu) \, d\mu \right],$$

it follows that when a recursion exists

$$
\begin{aligned}
\eta^{(n)}(t) &- \sum_{i=0}^{n-1} \alpha_{i+1}\eta^{(i)}(t) \\
&= \sum_{i=1}^{n} A_i \omega^{(n-i)}(t) - \sum_{i=0}^{n-1} \alpha_{i+1} \sum_{l=1}^{i} A_l \omega^{(i-l)}(t) \\
&= \sum_{j=0}^{n-1} \left(A_{n-j} - \sum_{i=1}^{n-j-1} \alpha_{i+j+1} A_i \right) \omega^{(j)}(t) \quad (6.42)
\end{aligned}
$$

for all t.

We conclude from Equation (6.42) that if $qW_A = P$, which implies that both Equations (6.39) and (6.36) hold, then also the i/o equation (6.38) is valid.

Conversely, assume that (6.38) is true for all i/o pairs of class $\mathcal{C}^{n-1}$. Pick any arbitrary vectors $v_0, \ldots, v_{n-1} \in \mathbb{K}^m$, and any interval $[\sigma, \tau]$, and for these let ω be the control on $[\sigma, \tau)$ defined by

$$\omega(t) := \sum_{i=0}^{n-1} v_i \frac{(t-\sigma)^i}{i!}.$$

Using again (6.40), the i/o equation at time $t = \sigma$ gives that

$$A_1 v_{n-1} + \ldots + A_n v_0 = \sum_{i=0}^{n-1} \alpha_{i+1}(A_i v_{i-1} + \ldots + A_i v_0) + \sum_{i=0}^{n-1} P_i v_i$$

must hold. Since these vectors were arbitrary, one may compare coefficients and equation (6.36) results. It also follows by differentiating (6.38) that

$$y^{(n+k+1)}(t) = \sum_{i=0}^{n-1} \alpha_{i+1} y^{(i+k+1)}(t) + \sum_{i=0}^{n-1} P_i u^{(i+k+1)}(t)$$

for all $k \geq 0$ and pairs of order $\mathcal{C}^{n+k}$. Applied in particular to the constant controls $\omega \equiv v_0$, v_0 arbitrary, and evaluating at $t = \sigma$, the recursion (6.39) results, and we know then that $qW_A = P$, as desired. ∎

A similar result holds for discrete-time systems. A linear time-invariant discrete-time behavior is said to satisfy the i/o equation of order n

$$y(t+n) = \sum_{i=0}^{n-1} \alpha_{i+1} y(t+i) + \sum_{i=0}^{n-1} P_i u(t+i) \tag{6.43}$$

if this equation holds for each i/o pair with $\omega \in \mathcal{U}^{[\sigma,\tau)}$ and each $\sigma \leq t \leq \tau - n$.

Lemma/Exercise 6.7.7 The discrete-time time-invariant behavior Λ satisfies (6.43) if and only if $W_A = q^{-1}P$, where q and P are as in Equations (6.33) and (6.34), with the same α_i's and P_i's. □

Exercise 6.7.8 Refer to Exercise 3.2.12. (Take all constants equal to one, for simplicity.)

(a) Using $y = h$ as the output, find an i/o equation of order 4 and the transfer function of the input/output behavior of Σ.

(b) Repeat the computation taking instead $x_3 = \dot\phi$ as the output; why is a transfer function with denominator of degree two obtained this time? Find a two-dimensional realization of the new i/o behavior. □

Exercise 6.7.9 Refer to Exercise 3.2.13, and take for simplicity $M = m = F = g = l = 1$.

(a) Using $y = x_1 = \delta$ as the output, find an i/o equation of order 4 and the transfer function of the input/output behavior of Σ.

(b) Repeat the computation taking instead $x_3 = \phi$ as the output; show that now there is an i/o equation of order three. Find a three-dimensional realization of this i/o behavior. □

6.8 Abstract Realization Theory*

There is an abstract theory of realization, which is of interest especially in the case of realizability by finite systems. We restrict attention here to *complete time-invariant* behaviors; unless otherwise stated, throughout this section Λ denotes a fixed such behavior. Accordingly, we wish to study realizations of Λ by time-invariant initialized complete systems (Σ, x^0). The realizability condition is that

$$\lambda_{\Sigma, x^0}^{0,t}(\omega) = \lambda^{0,t}(\omega)$$

_____
* This section can be skipped with no loss of continuity.

for all $t \in \mathcal{T}_+$ and each $\omega \in \mathcal{U}^{[0,t)}$. We drop the superscripts $0, t$ and/or the subscripts Σ and x^0 on the left when they are clear from the context.

The first observation is that, at this level of abstraction, *every* behavior is realizable. To see this, consider the set

$$\Omega := \bigcup_{T \in \mathcal{T}_+} \mathcal{U}^{[0,T)}$$

and the map

$$\phi_\Omega(\tau, \sigma, \nu, \omega) := \nu\omega^{T-\sigma}$$

corresponding to concatenation. That is, if

$$\omega \in \mathcal{U}^{[\sigma,\tau)} \quad \text{and} \quad \nu \in \mathcal{U}^{[0,T)},$$

then this is the control equal to ν on $[0,T)$ and equal to

$$\omega(t - T + \sigma)$$

if $t \in [T, T + \tau - \sigma)$.

Lemma/Exercise 6.8.1 The data $(\mathcal{T}, \Omega, \mathcal{U}, \phi)$ define a complete time-invariant system. □

We add the output function

$$h_\lambda : \Omega \to \mathcal{Y} : \omega \mapsto \lambda(\omega)$$

and the initial state $x^0 = \diamond$. Since $\phi_\Omega(t, 0, \diamond, \omega) = \omega$ for all ω defined on intervals of the form $[0,t)$, it follows that

$$h_\lambda(\phi_\Omega(t, 0, \diamond, \omega)) = h_\lambda(\omega) = \lambda(\omega),$$

and the following is established:

Lemma 6.8.2 The initialized system with outputs

$$\Sigma_{\Omega,\Lambda} := (\mathcal{T}, \Omega, \mathcal{U}, \phi_\Omega, \mathcal{Y}, h_\lambda, \diamond)$$

is a realization of λ. This system is complete, time-invariant, and reachable from the state $x^0 = \diamond$. □

Of course, the size of the state space of this system is huge, since it merely memorizes all inputs. A much more interesting realization results when one identifies indistinguishable states. The construction needed is more general, and it applies to arbitrary complete time-invariant systems, as follows.

Let $\Sigma = (\mathcal{T}, \mathcal{X}, \mathcal{U}, \phi, \mathcal{Y}, h)$ be any time-invariant complete system with outputs. Consider the space

$$\widetilde{\mathcal{X}} := \mathcal{X}/\sim$$

Given any behavior Λ, let Σ_Λ be the observability reduction of the system $\Sigma_{\Omega,\Lambda}$ in Lemma 6.8.2. This is again reachable (from $\diamond$), and by construction is also observable (as well as time-invariant and complete). We conclude as follows:

Theorem 29 *The initialized system with outputs* $(\Sigma_\Lambda, \diamond)$ *is a complete, time-invariant, canonical realization of* λ.   □

Our next objective is to show that canonical realizations are unique up to a relabeling of states. Assume given two initialized systems (Σ_i, x_i^0). A *system morphism*

$$T : (\Sigma_1, x_1^0) \to (\Sigma_2, x_2^0)$$

is by definition a map

$$T : \mathcal{X}_1 \to \mathcal{X}_2$$

such that $Tx_1^0 = x_2^0$,

$$T(\phi_1(\tau,\sigma,x,\omega)) = \phi_2(\tau,\sigma,T(x),\omega)$$

for all $\omega \in \mathcal{U}^{[\sigma,\tau)}$ and all $x \in \mathcal{X}$, and $h(x) = h(T(x))$ for all $x \in \mathcal{X}$. It is a *system isomorphism* if there exists another system morphism $S : (\Sigma_2, x_2^0) \to (\Sigma_1, x_1^0)$ such that the compositions $T \circ S$ and $S \circ T$ are the identity on $\mathcal{X}_2$ and $\mathcal{X}_1$, respectively. Since at this set-theoretic level a map is invertible if and only if it is one-to-one and onto, it follows easily that T is an isomorphism if and only if the underlying map $T : \mathcal{X}_1 \to \mathcal{X}_2$ is bijective.

Theorem 30 *Let* (Σ_1, x_1^0) *and* (Σ_2, x_2^0) *be two (complete, time-invariant) initialized systems with output. Assume that* Σ_1 *is reachable from* x_1^0, Σ_2 *is observable, and they realize the same behavior. Then there exists a system morphism* $T : (\Sigma_1, x_1^0) \to (\Sigma_2, x_2^0)$. *Furthermore:*

1. *There is a unique such morphism.*

2. *If* Σ_2 *is reachable from* x_2^0, *then* T *is onto.*

3. *If* Σ_1 *is observable, then* T *is one-to-one.*

4. *If both systems are canonical, then* T *is an isomorphism.*

Proof. Consider the set G consisting of all pairs $(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$ for which

$$h_1(\phi_1(\tau,\sigma,x_1,\omega)) = h_2(\phi_2(\tau,\sigma,x_2,\omega)) \ \forall \sigma \geq \tau, \forall \omega \in \mathcal{U}^{[\sigma,\tau)}$$

(indistinguishable but in different state spaces). We claim that G is the graph of a system morphism.

First note that, if (x_1, x_2) and (x_1, x_2') are both in G, then the definition of G forces x_2 and x_2' to be indistinguishable, so by observability it follows that $x_2 = x_2'$, which means that G is the graph of a partially defined map. Its domain

---

consisting of all equivalence classes under indistinguishability. For each $\sigma \leq \tau$ in $\mathcal{T}_+$, each equivalence class $[x]$, and each $\omega \in \mathcal{U}^{[\sigma,\tau)}$ let

$$\widetilde{\phi}(\tau,\sigma,[x],\omega) := [\phi(\tau,\sigma,x,\omega)],$$

that is, pick an arbitrary element in the equivalence class, apply the transitions of Σ, and then see in which equivalence class one lands. This map is well defined in the sense that the result is independent of the particular element $[x]$. In other words, for every $x, z \in \mathcal{X}$ so that $x \sim z$ and every ω, also $x_1 := \phi(\tau,\sigma,x,\omega)$ and $z_1 := \phi(\tau,\sigma,z,\omega)$ are indistinguishable. This is clear from the fact that, if ν is a control distinguishing x_1 and z_1, then the concatenation $\omega\nu$ distinguishes x and z. Similarly, the map

$$\widetilde{h}([x]) := h(x)$$

is well defined because indistinguishable states give rise in particular to identical instantaneous outputs.

This allows us to define a system

$$\widetilde{\Sigma} := (\mathcal{T}, \widetilde{\mathcal{X}}, \mathcal{U}, \widetilde{\phi}, \mathcal{Y}, \widetilde{h})$$

called the *observable reduction* of Σ. It satisfies that

$$\widetilde{h}(\widetilde{\phi}(\tau,\sigma,[x],\omega)) = \widetilde{h}([\phi(\tau,\sigma,x,\omega)]) = h(\phi(\tau,\sigma,x,\omega))$$

for all ω and x, so in particular with initial state

$$\widetilde{x^0} := [x^0]$$

it realizes the same behavior as (Σ, x^0), for each fixed $x^0 \in \mathcal{X}$. It is indeed observable, since $[x] \sim [z]$ implies

$$h(\phi(\tau,\sigma,x,\omega)) = \widetilde{h}(\widetilde{\phi}(\tau,\sigma,[x],\omega)) = \widetilde{h}(\widetilde{\phi}(\tau,\sigma,[z],\omega)) = h(\phi(\tau,\sigma,z,\omega))$$

for all ω, and therefore $[x] = [z]$. Finally, note that, if the original system Σ is reachable from a state x^0, then the observable reduction is reachable from $[x^0]$. Summarizing:

Lemma 6.8.3 For each system Σ and each $x^0 \in \mathcal{X}$:

1. $\widetilde{\Sigma}$ is observable.

2. If Σ is reachable from x^0, then $\widetilde{\Sigma}$ is reachable from $\widetilde{x^0}$.

3. The i/o behaviors of (Σ, x^0) and $(\widetilde{\Sigma}, \widetilde{x^0})$ coincide.   □

An initialized system (Σ, x^0) is a **canonical system** if it is reachable from x^0 and observable. This terminology is consistent with that used for linear systems, when applied with initial state $x^0 = 0$, since reachability from the origin is equivalent to complete controllability in that case.

is all of $\mathcal{X}_1$: by reachability, given any $x_1 \in \mathcal{X}_1$ there exists some control ω' so that

$$\phi_1(\sigma, 0, x_1^0, \omega') = x_1$$

and we define $x_2 := \phi_2(\sigma, 0, x_2^0, \omega')$; then for all $\omega \in \mathcal{U}^{[\sigma,\tau)}$,

$$h_1(\phi_1(\tau, \sigma, x_1, \omega)) = h_2(\phi_2(\tau, \sigma, x_2, \omega)),$$

which shows that $(x_1, x_2) \in G$. We have proved that G is the graph of a map $T : \mathcal{X}_1 \to \mathcal{X}_2$.

To see that T defines a morphism, observe that $(x_1^0, x_2^0) \in G$, by definition of G and the fact that the behaviors coincide, and also that for each $(x_1, x_2) \in G$ necessarily $h_1(x_1) = h_2(x_2)$ (from the case $\sigma = \tau$ in the definition of G). Finally, consider any $\omega \in \mathcal{U}^{[\sigma,\tau)}$ and let for $i = 1, 2$

$$\hat{x}_i := \phi_i(\tau, \sigma, x_i, \omega)$$

so we need to show that $(\hat{x}_1, \hat{x}_2) \in G$, too. This follows trivially from the definition of G and the semigroup property.

For any morphism T_0, if (x_1, x_2) is in the graph of T_0, then also

$$(\phi_1(\tau, \sigma, x_1, \omega), \phi_2(\tau, \sigma, x_2, \omega))$$

is in the graph, for all ω; it follows that also

$$h_1(\phi_1(\tau, \sigma, x_1, \omega)) = h_2(\phi_2(\tau, \sigma, x_2, \omega))$$

for all ω, so $(x_1, x_2) \in G$. This means that the graph of T_0 is included in G, which implies $T = T_0$, establishing the uniqueness claim.

Finally, if Σ_2 is reachable, then an argument totally analogous to that used to prove that T is defined everywhere shows that T must be onto, and an argument like the one used to show that T is single-valued gives that T is one-to-one if Σ_1 is observable. When both systems are canonical, the previous conclusions show that T must be bijective. Alternatively, the Theorem can be applied twice, resulting in a $T : \Sigma_1 \to \Sigma_2$ and an $S : \Sigma_2 \to \Sigma_1$; the compositions $T \circ S$ and $S \circ T$ are system morphisms, and since the identities $I : \Sigma_1 \to \Sigma_1$ and $I : \Sigma_2 \to \Sigma_2$ are morphisms, too, the uniqueness statements imply that the compositions equal the identity. ∎

Thus, just as in the linear case, one may conclude that every behavior is realizable by a canonical system, and canonical realizations are unique. No finiteness statements have been made yet, however.

Remark 6.8.4 One can in fact obtain many of the results for the linear theory as consequences of the above abstract considerations. For instance, take the proof of the fact that canonical realizations of linear behaviors must be unique up to a linear isomorphism. Assume that two systems are given, with the same spaces $\mathcal{U}, \mathcal{Y}$, which are assumed to be vector spaces over a field $\mathbb{K}$, that the state

spaces $\mathcal{X}_1, \mathcal{X}_2$ are also vector spaces, and that the maps ϕ_i and h_i are linear in (x, ω) and x, respectively. Furthermore, we assume that the initial states are zero. Then we claim that the set G is a linear subspace, which implies that the unique morphism T must correspond to a linear map $\mathcal{X}_1 \to \mathcal{X}_2$. Indeed, if (x_1, x_2) and (z_1, z_2) are in G and if $k \in \mathbb{K}$, then linearity gives that

$$
\begin{aligned}
h(\phi_1(\tau, \sigma, x_1 + kz_1, \omega)) &= h(\phi_1(\tau, \sigma, x_1, \mathbf{0})) + kh(\phi_1(\tau, \sigma, z_1, \omega)) \\
&= h(\phi_2(\tau, \sigma, x_2, \mathbf{0})) + kh(\phi_2(\tau, \sigma, z_2, \omega)) \\
&= h(\phi_2(\tau, \sigma, x_2 + kz_2, \omega)),
\end{aligned}
$$

which implies that $(x_1 + kz_1, x_2 + kz_2) \in G$. In particular, if both systems are canonical, then G is the graph of a linear isomorphism.

Arbitrary linear isomorphisms are not very interesting when considering infinite dimensional linear systems. For example, in the context of systems whose state spaces $\mathcal{X}_i$ are Banach spaces, the output value set $\mathcal{Y}$ is a normed space, and the maps ϕ_i and h_i are assumed to be bounded (continuous) linear operators, one may want to conclude that the morphism T given by the Theorem is bounded, too. This is an easy consequence of the above proof: It is only necessary to notice that G must be closed (because of the assumed continuities), so by the Closed Graph Theorem (see, for instance, [399], Theorem 4.2-I) the operator T must indeed be continuous. If both systems are canonical, T^{-1} is also bounded, by the same argument. □

We now turn to minimality. A *minimal system* will be one for which

$$\text{card } X = n < \infty$$

and with the property that any other system realizing the same behavior must have a state space of cardinality at least n.

Lemma 6.8.5 A (time-invariant, complete, initialized) system is minimal if and only if it is canonical.

Proof. Assume that Σ is minimal. If it is not observable, then its observable reduction (cf. Lemma 6.8.3) realizes the same behavior and has fewer states, a contradiction. If it is not reachable from its initial state x^0, then the restriction to $\mathcal{R}(x^0)$ gives a system with the same behavior and fewer states.

Conversely, assume that Σ is canonical. If it is not minimal, then there exists another realization Σ' of lower cardinality. Reducing if necessary by observability or reachability as in the above paragraph, we may assume that Σ' is canonical. By the isomorphism Theorem given above, Σ and Σ' must be isomorphic, contradicting the cardinality assertion. ∎

Note that the proof also shows that, if there exists any realization with finite cardinality, then there is a canonical one of finite cardinality.

Example 6.8.6 We consider again the parity check example discussed in Example 2.3.3. In particular, we shall see how to prove, using the above results, the last two claims in Exercise 2.3.4. The behavior to be realized is $\lambda(\tau, 0, \omega) =$

$$\begin{cases} 1 & \text{if } \omega(\tau-3) + \omega(\tau-2) + \omega(\tau-1) \text{ is odd and 3 divides } \tau > 0 \\ 0 & \text{otherwise} \end{cases}$$

and we take the system with

$$\mathcal{X} := \{0,1,2\} \times \{0,1\}$$

and transitions

$$\mathcal{P}((i,j),l) := (i+1 \bmod 3, j+l \bmod 2)$$

for $i = 1, 2$ and

$$\mathcal{P}((0,j),l) := (1,l).$$

The initial state is taken to be $(0,0)$, and the output map has $h(i,j) = 1$ if $i = 0$ and $j = 1$ and zero otherwise. (The interpretation is that $(k, 0)$ stands for the state "t is of the form $3s + k$ and the sum until now is even," while states of the type $(k, 1)$ correspond to odd sums.) This is clearly a realization, with 6 states. To prove that there is no possible (time-invariant, complete) realization with less states, it is sufficient to show that it is reachable and observable.

Reachability follows from the fact that any state of the form $(0,j)$ can be obtained with an input sequence $j00$, while states of the type $(1,j)$ are reached from x^0 using input j (of length one) and states $(2,j)$ using input $j0$.

Observability can be shown through consideration of the following controls ω_{ij}, for each (i,j):

$$\omega_{01} := \diamond, \ \omega_{00} := 100, \ \omega_{10} := 10, \ \omega_{11} := 00, \ \omega_{21} := 0, \ \omega_{20} := 0.$$

Then, ω_{01} separates $(0,1)$ from every other state, while for all other pairs $(i,j) \neq (0,1)$,

$$\lambda_{(i,j)}(\omega_{\alpha\beta}) = 1$$

if and only if $(i,j) = (\alpha,\beta)$. □

Exercise 6.8.7 Let Λ be any time-invariant complete behavior and let $\sim$ be the following equivalence relation on Ω:

$$\omega \sim \omega' \Leftrightarrow \lambda(\omega\nu) = \lambda(\omega'\nu) \ \forall \nu$$

(to be more precise, one should write the translated version of ν). This is the *Nerode equivalence relation*. Prove that λ admits a finite-cardinality realization if and only if there are only finitely many equivalence classes under the Nerode relation. (*Hint:* It only takes a couple of lines, using previous results.) □

The result in the Exercise can be interpreted in terms of generalized Hankel matrices. Consider the matrix with rows and columns indexed by elements of Ω, and with $\lambda(\omega\nu)$ in position (ω, ν). Then finite realizability is equivalent to this matrix having only finitely many rows, and the number of different rows is equal to the cardinality of the state space of a minimal realization. This is in complete analogy to the linear case.

6.9 Notes and Comments

Basic Observability Notions

The concept of observability originates both from classical automata theory and from basic linear systems. The material on final-state observability can be generalized to certain types of infinite systems, which in the continuous-time case give, because of Proposition 6.1.9, results about observability (see, for instance, [358] and [384]).

The text [52] has an extensive discussion of the topic of multiple experiments, final-state determination, and control for finite systems. Also, [97] addresses the observability question for such systems. In the context of VLSI design a central problem is that of testing circuits for defects, and notions of observability appear; see, for instance, [148] and the references therein.

Observability of Time-Invariant Systems

The concept of observability, its duality with controllability for linear systems, and the notion of canonical realization, all arose during the early 1960s. Some early references are [155], [215], and [320].

In fact, Kalman's original filtering paper [214] explicitly mentions the duality between the optimization versions of feedback control (linear-quadratic problem) and observer construction (Kalman filtering). The term "canonical" was used already in [217], where one can find reachability and observability decompositions. For related results on duality of time-varying linear systems, see [435].

For more on the number of experiments needed for observability, and in particular Exercise 6.2.3, see, for instance, [97], [125], and especially [156]. See also [6] for continuous-time questions related to this problem. Results on observability of recurrent "neural" nets (the systems studied in Section 3.8, with linear output $y = Cx$) can be found in [14].

In the form of Exercise 6.2.12 (which, because of the simplifying assumption that the numbers ω_i are distinct, could also have been easily proved directly), one refers to the eigenvalue criterion for sampling as *Shannon's Theorem* or the *Sampling Theorem*. It gives a sufficient condition to allow reconstruction of the signal η and is one of the cornerstones of digital signal processing. The result can be generalized, using Fourier transform techniques, to more general signals η, containing an infinite number of frequency components.

It is possible to develop a large amount of the foundations of time-invariant systems based on the notion of *observables* associated to a system; see [373].

Linearization Principle for Observability

Often one refines Definition 6.4.1 to require that states be distinguishable without large excursions. One such possibility is to ask that for each neighborhood W of x^0 there be a neighborhood V so that every state in V is distinguishable from x^0 using a control that makes the resulting trajectory stay in W. This is more natural in the context of Lyapunov stability, and can be characterized elegantly for continuous-time smooth systems; see, for instance, [5], [7], [28], [139], and [365]. Other nonlinear observability references are, for instance, [185] and [365]. [310].

Realization Theory for Linear Systems

There are papers on realization of time-varying linear systems that use techniques close to those used for the time-invariant case. See, for instance, [235] and [434].

Recursion and Partial Realization

See, for instance, [180], [181], [192], [395], and the many references therein, as well as the early paper [225], for further results on families of systems.

Small perturbations of the Markov parameters will result in a nonrealizable sequence, since all of the determinants of the submatrices of $\mathcal{H}$ are generically nonzero. In this context it is of interest to look for *partial realizations*, in which only a finite part of the sequence is matched; this problem is closely related to classical mathematical problems of *Padé approximation*. Lemma 6.6.9 is one result along these lines; see, for instance, [19] and the references therein, for a detailed treatment of partial realization questions and relations to problems of rational interpolation. The problem can also be posed as one of optimal approximation of Hankel operators; see the by now classical paper [3]. In addition, the procedures that we described are numerically unstable, but various modifications render them stable; a reference in that regard is [110]. A recursive realization procedure is given in [18], which permits realizations for additional data to make use of previous realizations.

Since a Markov sequence is specified in terms of an infinite amount of data, one cannot expect to solve completely the question of realizability unless some sort of finite description is first imposed. It is known however that, for arbitrary "computable" descriptions, the problem of deciding realizability is undecidable in the sense of logic, that is, there is no possible computer program that will always correctly determine, given a description of a Markov sequence, whether a realization exists; see, for instance, [356].

Rationality and Realizability

The ideas of realization and relations to input/output equations go back at least to the nineteenth century, in the context of using integrators to solve algebraic differential equations; see [400] as well as the extensive discussion in [212], Chapter 2. The relations between rationality and finite Hankel rank are also classical (for the scalar case) and go back to the work of Kronecker (see [152]).

Algebraic techniques for studying realizations of linear systems were emphasized by Kalman; see, for instance, [220], [221], and [223], as well as [228] for relations to econometrics.

Proposition 6.7.6 could also be proved using Laplace transform techniques; we used a direct technique that in principle can be generalized to certain nonlinear systems. For these generalizations, see, for instance, [360] and [421], which establish that certain types of i/o behaviors satisfy polynomial equations

$$E(y(t+n), y(t+n-1), \ldots, y(t), u(t+n-1), \ldots, u(t)) = 0 \qquad (6.44)$$

(or $E(y^{(n)}(t), y^{(n-1)}(t), \ldots, y(t), u^{(n-1)}(t), \ldots, u(t)) = 0$) if and only if they are realizable, for discrete-time and continuous-time systems, respectively. In the nonlinear case, however, not every "causal" i/o equation necessarily gives rise to an i/o behavior, and this in turn motivates a large amount of research on such equations; see, for instance, [408]. Some authors, motivated by differential-algebraic techniques, have suggested that realizability should be *defined* in terms of i/o equations; see especially [141] and the references therein.

The reference [272] discusses i/o equations for nonlinear systems in the context of identification problems. Related material is also presented in [307], which uses i/o equations (6.44) in which E is not linear, nor polynomial, but instead is given by iterated compositions of a fixed scalar nonlinear function with linear maps. Numerical experience seems to suggest that such combinations are particularly easy to estimate using gradient descent techniques, and they are in principle implementable in parallel processors. The name "neural network" is used for this type of function E, because of the analogy with neural systems: the linear combinations correspond to dendritic integrations of signals, and the scalar nonlinear function corresponds to the "firing" response of each neuron, depending on the weighted input to it.

Realization and i/o equations can also be studied in a stochastic context, especially when $u(0), u(1), \ldots$ in (6.43) are random variables. In that case, especially if these variables are independent and identically distributed, (6.43) describes what is called an *ARMA* or *autoregressive moving average* model for the stochastic process, in the stochastic systems and statistics literature. (The outputs form a *time series*, and the realization is a Markov model for the series.) See, for instance, the "weak" and "strong" Gaussian stochastic realization problems studied, respectively, in [131], [132], [222], [409], and [10], [11], [280], and the related [344].