Chapter 7

Introduction to Dynamical Models

With this chapter, we start to consider modeling problems for linear time-invariant (LTI) systems. First, we give an introduction to the LTI model class using the behavioral language. As in the static case, a key question is the representation of the model, i.e., how it is described by equations. Again, the kernel, image, and input/output representations play an important role, but other representations that bring additional structure into evidence are used as well.

Dynamical systems are much richer in properties than static systems. In the dynamic case, the memory of the system is central, i.e., the fact that the past can affect the future. The intuitive notion of memory is formalized in the definition of state. In addition, a key role is played by the controllability property of the system. Every linear static system has an image representation. In the dynamic case this is no longer true. A necessary and sufficient condition for existence of an image representation is controllability.

7.1 Linear Time-Invariant Systems

Dynamical systems describe variables that are functions of one independent variable, referred to as "time". In Chapter 2, a system was defined as a subset $\mathscr{B}$ of a universum set $\mathscr{U}$. In the context of dynamical systems, $\mathscr{U}$ is a set of functions $w : \mathbb{T} \to \mathbb{W}$, denoted by $\mathbb{W}^{\mathbb{T}}$. The sets $\mathbb{W}$ and $\mathbb{T} \subseteq \mathbb{R}$ are called, respectively, signal space and time axis. The signal space is the set where the system variables take on their values and the time axis is the set where the time variable takes on its values. We use the following definition of a dynamical system [Wil86a].

A dynamical system Σ is a 3-tuple $\Sigma = (\mathbb{T}, \mathbb{W}, \mathscr{B})$, with $\mathbb{T} \subseteq \mathbb{R}$ the time axis, $\mathbb{W}$ the signal space, and $\mathscr{B} \subseteq \mathbb{W}^{\mathbb{T}}$ the behavior.

The behavior $\mathscr{B} \subseteq \mathbb{W}^{\mathbb{T}}$ is the set of all legitimate functions, according to the system Σ, from the universum set $\mathscr{U} = \mathbb{W}^{\mathbb{T}}$. When the time axis and the signal space are understood from the context, as is often the case, we may identify the system $\Sigma = (\mathbb{T}, \mathbb{W}, \mathscr{B})$ with its behavior $\mathscr{B}$. As with any set, the behavior can be described in a number of ways. In the context of dynamical systems, most often used are representations by equations

$f : \mathbb{W}^{\mathbb{T}} \to \mathbb{R}^g$, i.e., $\mathscr{B} = \{ w \in \mathbb{W}^{\mathbb{T}} \mid f(w) = 0 \}$. The equations $f(w) = 0$ are called annihilating behavioral equations.

Of interest are systems with special properties. In the behavioral setting,

> a property of the system Σ is always defined in terms of the behavior and then translated to equivalent statements in terms of particular representations.

Similarly, the statement that w is a trajectory of Σ, i.e., $w \in \mathscr{B}$, is translated to more convenient characterizations for numerical verification in terms of representations of Σ.

Note 7.1 (Classical vs. behavioral theory) In the classical theory, system properties are often defined on the representation level; i.e., a property of the system is defined as a property of a particular representation. (Think, for example, of controllability, which is defined as a property of a state space representation.) This has the drawback that such a definition might be representation dependent and therefore not a genuine property of the system itself. (For example, a controllable system (see Section 7.5 for definition) may have uncontrollable state representation.)

It is more natural to work instead the other way around.

1. Define the property in terms of the behavior $\mathscr{B}$.
2. Find the implications of that property on the parameters of the system in particular representations. On this level, algorithms for verification of the property are derived.

The way of developing system theory as a sequence of steps 1 and 2 is characteristic for the behavioral approach.

A static system $(\mathscr{U}, \mathscr{B})$ is *linear* when the universum set $\mathscr{U}$ is a vector space and the behavior $\mathscr{B}$ is a linear subspace. Analogously, a dynamical system $\Sigma = (\mathbb{T}, \mathbb{W}, \mathscr{B})$ is linear when the signal space $\mathbb{W}$ is a vector space and $\mathscr{B}$ is a linear subspace of $\mathbb{W}^{\mathbb{T}}$ (viewed as a vector space in the natural way).

The universum set $\mathbb{W}^{\mathbb{T}}$ of a dynamical system has special structure that is not present in the static case. For this reason dynamical systems are richer in properties than static systems. Next, we restrict ourselves to the case when the time axis is either $\mathbb{T} = \mathbb{N}$ or $\mathbb{T} = \mathbb{Z}$ and define two properties—time-invariance and completeness. In keeping with tradition, we call a function $w \in \mathbb{W}^{\mathbb{T}}$ a time series.

A system $\Sigma = (\mathbb{N}, \mathbb{W}, \mathscr{B})$ is *time-invariant* if $\mathscr{B} \subseteq \sigma\mathscr{B}$, where σ is the backward shift operator $(\sigma w)(t) := w(t+1)$ and $\sigma\mathscr{B} := \{ \sigma w \mid w \in \mathscr{B} \}$. In the case $\mathbb{T} = \mathbb{Z}$, a system $\Sigma = (\mathbb{Z}, \mathbb{W}, \mathscr{B})$ is time-invariant if $\mathscr{B} = \sigma\mathscr{B}$. Time-invariance requires that if a time series w is a trajectory of a time-invariant system, then all its backward shifts $\sigma^t w$, $t > 0$, are also trajectories of that system.

The restriction of the behavior $\mathscr{B} \subseteq (\mathbb{R}^{\mathtt{w}})^{\mathbb{T}}$ to the time interval $[t_1, t_2]$, where $t_1, t_2 \in \mathbb{T}$ and $t_1 < t_2$, is denoted by

$$\mathscr{B}|_{[t_1,t_2]} := \{ w \in (\mathbb{R}^{\mathtt{w}})^{t_2 - t_1 + 1} \mid \text{ there are } w_- \text{ and } w_+ \text{ such that } \operatorname{col}(w_-, w, w_+) \in \mathscr{B} \}.$$

A system $\Sigma = (\mathbb{T}, \mathbb{W}, \mathscr{B})$ is *complete* if

$$w|_{[t_0,t_1]} \in \mathscr{B}|_{[t_0,t_1]} \quad \text{for all } t_0, t_1 \in \mathbb{T},\ t_0 \le t_1 \qquad \Longrightarrow \qquad w \in \mathscr{B};$$

i.e., by looking at the time series w through a window of finite width $t_1 - t_0$, one can decide if it is in the behavior or not. Moreover, if the window can be taken to have a fixed width $t_1 - t_0 = l$, then the system is called l-complete. It turns out that a system is complete if and only if its behavior is closed in the topology of pointwise convergence, i.e., if $w_i \in \mathscr{B}$ for $i \in \mathbb{N}$ and $w_i(t) \to w(t)$, for all $t \in \mathbb{T}$, implies $w \in \mathscr{B}$. Also, a system is l-complete if and only if there is a difference equation representation of that system with l time shifts. For LTI systems, the completeness property is also called *finite dimensionality*.

We consider the class of discrete-time complete LTI systems. Our generic notation for the signal space is $\mathbb{W} = \mathbb{R}^{\mathtt{w}}$.

The class of all complete LTI systems with $\mathtt{w}$ variables is denoted by $\mathscr{L}^{\mathtt{w}}$.

Next, we discuss representations of the class $\mathscr{L}^{\mathtt{w}}$.

7.2 Kernel Representation

Consider the difference equation

$$R_0 w(t) + R_1 w(t+1) + \cdots + R_l w(t+l) = 0, \qquad \text{where} \quad R_\tau \in \mathbb{R}^{g \times \mathtt{w}}. \tag{DE}$$

It shows the dependence among consecutive samples of the time series w. Assuming that $R_l \neq 0$, the maximum number of shifts is l. The integer l is called the *lag* of the equation. Since in general (DE) is a vector equation, l is the largest lag among the lags $l_1, \ldots, l_g$ of all scalar equations.

Obviously, (DE) induces a dynamical system via the representation

$$\mathscr{B} = \{\, w \in (\mathbb{R}^{\mathtt{w}})^{\mathbb{Z}} \mid \text{(DE) holds} \,\}.$$

One can analyze $\mathscr{B}$ using the difference equation. It turns out, however, that it is more convenient to use polynomial matrix algebra for this purpose. (DE) is compactly written in terms of the polynomial matrix

$$R(z) := R_0 + R_1 z^1 + R_2 z^2 + \cdots + R_l z^l \in \mathbb{R}^{g \times \mathtt{w}}[z]$$

as $R(\sigma)w = 0$. Consequently, operations on the system of difference equations are represented by operations on the polynomial matrix R. The system induced by (DE) is

$$\ker\big(R(\sigma)\big) := \{\, w \in (\mathbb{R}^{\mathtt{w}})^{\mathbb{N}} \mid R(\sigma)w = 0 \,\}. \tag{KER repr}$$

We call (KER repr) a kernel representation of the system $\mathscr{B} := \ker\big(R(\sigma)\big)$.

The following theorem summarizes the representation-free characterization of the class of complete LTI systems, explained in the previous section, and states that

without loss of generality one can assume the existence of a kernel representation $\mathscr{B} = \ker\big(R(\sigma)\big)$ of a system $\mathscr{B} \in \mathscr{L}^{\mathtt{w}}$.

Theorem 7.2 (Willems [Wil86a]). *The following are equivalent:*

(i) $\Sigma = (\mathbb{Z}, \mathbb{R}^{\mathtt{w}}, \mathscr{B})$ *is linear, time invariant, and complete.*

(ii) $\mathscr{B}$ *is linear, shift-invariant, and closed in the topology of pointwise convergence.*

(iii) *There is a polynomial matrix* $R \in \mathbb{R}^{\bullet \times \mathtt{w}}[z]$, *such that* $\mathscr{B} = \ker\big(R(\sigma)\big)$.

The linearity of the system induced by (DE) follows from the linearity of (DE) with respect to w. The shift-invariance follows from the time-invariance of the coefficients $R_0, \dots, R_l$, and the completeness follows from the fact that (DE) involves a finite number l of shifts of the time series. Thus (iii) $\Longrightarrow$ (i) is immediate. The reverse implication, (i) $\Longrightarrow$ (iii), on the other hand, requires proof; see [Wil86a, Theorem 5].

A kernel representation associated with a given $\mathscr{B} \in \mathscr{L}^{\mathtt{w}}$ is not unique. The nonuniqueness is due to

1. linearly dependent equations (which refers to R not being full row rank) and
2. equivalence of the representations $\ker\big(R(\sigma)\big) = 0$ and $\ker\big(U(\sigma)R(\sigma)\big) = 0$, where $U \in \mathbb{R}^{g \times g}[z]$ is a unimodular matrix.

A square polynomial matrix U is *unimodular* if it has a polynomial inverse. A necessary and sufficient condition for U to be unimodular is its determinant to be a nonzero constant. Two kernel representations of the same behavior are called equivalent.

> Premultiplication of R with a unimodular matrix is a convenient way to represent a sequence of equivalence transformations on the system of difference equations (DE).

For a given system $\mathscr{B} \in \mathscr{L}^{\mathtt{w}}$, there always exists a kernel representation in which the polynomial matrix R has full row rank [Wil91, Proposition III.3]. Such a kernel representation is called a *minimal kernel representation*. In a minimal kernel representation, the number of equations $\mathtt{p} := \operatorname{row\,dim}(R)$ is minimal among all possible kernel representations of $\mathscr{B}$. All minimal kernel representations of a given system are in fact unimodularly equivalent; i.e., if $R'(\sigma) = 0$ and $R''(\sigma) = 0$ are both minimal, then there is a unimodular matrix U, such that $R' = UR''$.

There exists a minimal kernel representation $\mathscr{B} = \ker\big(R(\sigma)\big)$, in which the number of equations $\mathtt{p} = \operatorname{row\,dim}(R)$, the maximum lag $\mathtt{l} := \max_{i=1,\dots,\mathtt{p}} l_i$, and the total lag $\mathtt{n} := \sum_{i=1}^{\mathtt{p}} l_i$ are simultaneously all minimal over all possible kernel representations [Wil86a, Theorem 6]. Such a kernel representation is called *shortest lag representation*. A kernel representation $\mathscr{B} = \ker\big(R(\sigma)\big)$ is a shortest lag representation if and only if $R(z)$ is row proper. The polynomial matrix $R = \begin{bmatrix} r_1 & \cdots & r_{\mathtt{p}} \end{bmatrix}^\top$, $\deg(r_i) =: \mathtt{l}_i$ is *row proper* if the leading row coefficient matrix (i.e., the matrix of which the (i, j)th entry is the coefficient of the term with power $\mathtt{l}_i$ of $R_{ij}(z)$) is full row rank. It can be shown that the $\mathtt{l}_i$'s are the observability indices of the system.

> The minimal and shortest lag kernel representations correspond to special properties of the R matrix: in a minimal representation, R is full row rank, and in a shortest lag representation, R is row proper.

A shortest lag representation is a special minimal representation, because a row proper matrix is necessarily full row rank. A shortest lag representation, however, is still not unique.

The minimal number of equations $\mathtt{p}$, the lag $\mathtt{l}$, and the total lag $\mathtt{n}$ are invariants of $\mathscr{B}$. It turns out that $\mathtt{p}$ is equal to the number of outputs, called output cardinality, in an input/output representation. Correspondingly, the integer $\mathtt{m} := \mathtt{w} - \mathtt{p}$ is also an invariant of $\mathscr{B}$ and is called the input cardinality. It is equal to the number of inputs in an input/output representation. The total lag $\mathtt{n}$ is equal to the state dimension in a minimal state space representation of $\mathscr{B}$. We use the following notation:

$\mathbf{m}(\mathscr{B})$ for the input cardinality of $\mathscr{B}$,

$\mathbf{p}(\mathscr{B})$ for the output cardinality of $\mathscr{B}$,

$\mathbf{n}(\mathscr{B})$ for the minimal state dimension of $\mathscr{B}$, and

$\mathbf{l}(\mathscr{B})$ for the lag of $\mathscr{B}$.

7.3 Inputs, Outputs, and Input/Output Representation

Consider a projection operator $\Pi \in \mathbb{R}^{\mathtt{w}\times\mathtt{w}}$ and a partitioning of the time series $w \in (\mathbb{R}^{\mathtt{w}})^{\mathbb{Z}}$ into time series u and y as follows:

$$\begin{bmatrix} u \\ y \end{bmatrix} := \Pi^\top w, \qquad \text{where} \quad \dim\big(u(t)\big) =: \mathtt{m}, \quad \dim\big(y(t)\big) =: \mathtt{p}, \quad \text{with} \quad \mathtt{m}+\mathtt{p}=\mathtt{w}.$$

Respectively a behavior $\mathscr{B} \in \mathscr{L}^{\mathtt{w}}$ is partitioned into two subbehaviors $\mathscr{B}_u$ and $\mathscr{B}_y$. The variables in u are called free if $\mathscr{B}_u = (\mathbb{R}^{\mathtt{m}})^{\mathbb{Z}}$. If, in addition, any other partitioning results in no more free variables, then $\mathscr{B}_u$ is called maximally free in $\mathscr{B}$. A partitioning in which $\mathscr{B}_u$ is maximally free is called an input/output partitioning with u an input and y an output.

There always exists an input/output partitioning of the variables of $\mathscr{B} \in \mathscr{L}^{\mathtt{w}}$, in fact a componentwise one; see [Wil86a, Theorem 2]. It is not unique, but the number of free variables $\mathtt{m}$ and the number of dependent variables $\mathtt{p}$ are equal to, respectively, the input cardinality and the output cardinality of $\mathscr{B}$ and are invariant. In a minimal kernel representation $\ker\big(R(\sigma)\big) = \mathscr{B}$, the choice of such a partitioning amounts to the selection of a full-rank square submatrix of R. The variables corresponding to the columns of R that form the full-rank submatrix are dependent variables and the other variables are free.

The inputs together with the initial conditions determine the outputs. This property is called *processing* [Wil91, Definition VIII.2]. Also, the inputs can be chosen so that they are not anticipated by the outputs. *Nonanticipation* is also called causality [Wil91, Definition VIII.4].

Let $\ker\big(R(\sigma)\big)$ be a minimal kernel representation of $\mathscr{B} \in \mathscr{L}^{\mathtt{w}}$. One can always find a permutation matrix $\Pi \in \mathbb{R}^{\mathtt{w}\times\mathtt{w}}$, such that $P \in \mathbb{R}^{\mathtt{p}\times\mathtt{p}}[z]$, defined by $R\Pi =: \begin{bmatrix} Q & -P \end{bmatrix}$, has a nonzero determinant and the rational polynomial matrix

$$G(z) := P^{-1}(z)Q(z) \in \mathbb{R}^{\mathtt{p}\times\mathtt{m}}(z) \tag{TF}$$

is proper. This requires selecting a submatrix P among all full-rank square submatrices of R that has determinant of maximal degree. Then the corresponding partitioning of w,

$\operatorname{col}(u, y) := \Pi^\top w$, is an input/output partitioning. G being proper implies that u is not anticipated by y; see [Wil91, Theorem VIII.7].

The difference equation

$$P(\sigma)y = Q(\sigma)u \qquad \text{(I/O eqn)}$$

with an input/output partitioning Π is called an input/output equation, and the matrix G, defined in (TF), is called the transfer function of the system $\mathscr{B} := \ker\big(R(\sigma)\big)$.

> The class of LTI complete systems with $\mathtt{w}$ variables and at most $\mathtt{m}$ inputs is denoted by $\mathscr{L}^{\mathtt{w}}_{\mathtt{m}}$.

The system $\mathscr{B} \in \mathscr{L}^{\mathtt{w}}$ induced by an input/output equation with parameters (P, Q) (and input/output partitioning defined by Π) is

$$\mathscr{B}_{\mathrm{i/o}}(P, Q, \Pi) := \{\, w := \Pi \operatorname{col}(u, y) \in (\mathbb{R}^{\mathtt{w}})^{\mathbb{N}} \mid P(\sigma)y = Q(\sigma)u \,\}. \qquad \text{(I/O repr)}$$

(I/O repr) is called an input/output representation of the system $\mathscr{B} := \mathscr{B}_{\mathrm{i/o}}(P, Q, \Pi)$. If Π is the identity matrix $I_{\mathtt{w}}$, it is skipped in the notation of the input/output representation.

7.4 Latent Variables, State Variables, and State Space Representations

Modeling from first principles invariably requires the addition to the model of other variables apart from the ones that the model aims to describe. Such variables are called latent, and we denote them by l (not to be confused with the lag of a difference equation). The variables w that the model aims to describe are called manifest variables in order to distinguish them from the latent variables.

An important result, called the elimination theorem [Wil86a, Theorem 1], states that the behavior

$$\mathscr{B}(R, M) := \big\{\, w \in (\mathbb{R}^{\mathtt{w}})^{\mathbb{N}} \;\big|\; \exists\, l \in (\mathbb{R}^{\mathtt{l}})^{\mathbb{N}}, \text{ such that } R(\sigma)w = M(\sigma)l \,\big\} \qquad \text{(LV repr)}$$

induced by the latent variable equation

$$R(\sigma)w = M(\sigma)l \qquad \text{(LV eqn)}$$

is LTI. The behavior $\mathscr{B}(R, M)$ is called manifest behavior of the latent variable system. The behavior of the manifest and latent variables together is called the full behavior of the system. The elimination theorem states that if the full behavior is LTI, then the manifest behavior is LTI; i.e., by eliminating the latent variables, the resulting system is still LTI.

A latent variable system is *observable* if there is a map $w \mapsto l$, i.e., if the latent variables can be inferred from the knowledge of the system and the manifest variables. The kernel representation is a special case of the latent variable representation for $R = I$.

State variables are special latent variables that specify the memory of the system. More precisely, latent variables x are called state variables if they satisfy the following axiom of state [Wil91, Definition VII.1]:

$$(w_1, x_1), (w_2, x_2) \in \mathscr{B},\; t \in \mathbb{N}, \text{ and } x_1(t) = x_2(t) \quad \Longrightarrow \quad (w, x) \in \mathscr{B},$$

where

$$\big(w(\tau), x(\tau)\big) := \begin{cases} \big(w_1(\tau), x_1(\tau)\big) & \text{for } \tau < t, \\ \big(w_2(\tau), x_2(\tau)\big) & \text{for } \tau \geq t. \end{cases}$$

A latent variable representation of the system is a state variable representation if there exists an equivalent representation whose behavioral equations are first order in the latent variables and zeroth order in the manifest variables. For example, the equation

$$\sigma x = A'x + B'v, \qquad w = C'x + D'v$$

defines a state representation. It is called state representation with a driving input because v acts like the input: v is free and, together with the initial conditions, determines a trajectory $w \in \mathscr{B}$. The system induced by the parameters (A', B', C', D') is

$$\mathscr{B}_{\text{ss}}(A', B', C', D') := \big\{ w \in (\mathbb{R}^{\mathtt{w}})^{\mathbb{N}} \mid \exists\, v \in (\mathbb{R}^{\mathtt{v}})^{\mathbb{N}} \text{ and } x \in (\mathbb{R}^{\mathtt{n}})^{\mathbb{N}}, \\ \text{such that } \sigma x = A'x + B'v,\ w = C'x + D'v \big\}.$$

Any LTI system $\mathscr{B} \in \mathscr{L}^{\mathtt{w}}$ admits a representation by an input/state/output equation

$$\sigma x = Ax + Bu, \qquad y = Cx + Du, \qquad w = \Pi \operatorname{col}(u, y), \tag{I/S/O eqn}$$

in which both the input/output and the state structure of the system are explicitly displayed [Wil86a, Theorem 3]. The system $\mathscr{B}$, induced by an input/state/output equation with parameters (A, B, C, D) and Π, is

$$\mathscr{B}_{\text{i/s/o}}(A, B, C, D, \Pi) := \{ w := \Pi \operatorname{col}(u, y) \in (\mathbb{R}^{\mathtt{w}})^{\mathbb{N}} \mid \exists\, x \in (\mathbb{R}^{\mathtt{n}})^{\mathbb{N}}, \\ \text{such that } \sigma x = Ax + Bu,\ y = Cx + Du \}. \tag{I/S/O repr}$$

(I/S/O repr) is called an input/state/output representation of the system $\mathscr{B} := \mathscr{B}_{\text{i/s/o}}(A, B, C, D, \Pi)$. Again, Π is skipped whenever it is I.

An input/state/output representation is not unique. The minimal state dimension $\mathtt{n} = \dim(x)$ among all input/state/output representations of $\mathscr{B}$, however, is invariant (denoted by $\mathbf{n}(\mathscr{B})$).

We denote the class of LTI systems with $\mathtt{w}$ variables, at most $\mathtt{m}$ inputs, and minimal state dimension at most $\mathtt{n}$ by $\mathscr{L}_{\mathtt{m}}^{\mathtt{w},\mathtt{n}}$.

7.5 Autonomous and Controllable Systems

A system $\mathscr{B}$ is *autonomous* if for any trajectory $w \in \mathscr{B}$ the past

$$w_- := \big(\ldots, w(-2), w(-1)\big)$$

of w completely determines its future

$$w_+ := \big(w(0), w(1), \ldots\big).$$

A system $\mathscr{B}$ is autonomous if and only if its input cardinality $\mathbf{m}(\mathscr{B})$ equals 0. Therefore, an autonomous LTI system is parameterized by the pair of matrices A and C via the state space representation

$$\sigma x = Ax, \qquad y = Cx, \qquad w = y. \tag{AUT}$$

The system induced by the state space representation with parameters (A, C) is

$$\mathscr{B}_{\mathrm{i/s/o}}(A, C) := \{\, w \in (\mathbb{R}^{\mathrm{p}})^{\mathbb{N}} \mid \exists\, x \in (\mathbb{R}^{\mathrm{n}})^{\mathbb{N}}, \text{ such that } \sigma x = Ax,\ w = Cx \,\}.$$

The behavior of an autonomous system is finite dimensional; in fact, $\dim(\mathscr{B}) = \mathbf{n}(\mathscr{B})$. Alternatively, an autonomous LTI system is parameterized in a minimal kernel representation $\mathscr{B} = \ker\big(R(\sigma)\big)$ by a square and nonsingular matrix R, i.e., $R \in \mathbb{R}^{\mathrm{p}\times\mathrm{p}}[z]$, $\det(R) \neq 0$.

The system $\mathscr{B}$ is *controllable* if for any two trajectories $w_1, w_2 \in \mathscr{B}$, there is a third trajectory $w \in \mathscr{B}$, such that $w_1(t) = w(t)$, for all $t < 0$, and $w_2(t) = w(t)$, for all $t \geq 0$. The subset of controllable systems contained in the set $\mathscr{L}^{\mathrm{w}}$ is denoted by $\mathscr{L}^{\mathrm{w}}_{\mathrm{ctrb}}$. A noncontrollable system $\mathscr{B}$ can be represented [Wil91, Proposition V.8] as $\mathscr{B} = \mathscr{B}_{\mathrm{ctrb}} \oplus \mathscr{B}_{\mathrm{aut}}$, where $\mathscr{B}_{\mathrm{ctrb}}$ is the largest controllable subsystem in $\mathscr{B}$ and $\mathscr{B}_{\mathrm{aut}}$ is a (nonunique) autonomous subsystem.

A test for controllability of the system $\mathscr{B}$ in terms of the parameter $R \in \mathbb{R}^{g\times\mathrm{w}}[z]$ in a kernel representation $\mathscr{B} = \ker\big(R(\sigma)\big)$ is given in [Wil91, Theorem V.2]: $\mathscr{B}$ is controllable if and only if the matrix $R(z)$ has a constant rank for all $z \in \mathbb{C}$. Equivalently, $\mathscr{B}$ is controllable if and only if a matrix R that defines a minimal kernel representation of $\mathscr{B}$ is left prime. In terms of the input/output representation $\mathscr{B} = \mathscr{B}_{\mathrm{i/o}}(P, Q)$, $\mathscr{B}$ being controllable is equivalent to P and Q being left coprime.

The controllable subsystem $\mathscr{B}_{\mathrm{ctrb}}$ of $\mathscr{B}$ can be found via the factorization $R = FR'$, where $F \in \mathbb{R}^{g\times g}[z]$ and R' is prime: $\mathscr{B}_{\mathrm{ctrb}} = \ker\big(R'(\sigma)\big)$. In general, left multiplication of R with a nonsingular polynomial matrix changes the behavior: it amounts to adding an autonomous subbehavior. Only left multiplication with a unimodular matrix does not alter the behavior because it adds the trivial autonomous behavior $\{0\}$.

7.6 Representations for Controllable Systems

The transfer function G parameterizes the controllable subsystem of $\mathscr{B}_{\mathrm{i/o}}(P, Q)$. Let $\mathscr{Z}$ be the Z-transform

$$\mathscr{Z}(w) = w(0) + w(1)z^{-1} + w(2)z^{-2} + \cdots$$

and consider the input/output equation

$$\mathscr{Z}(y) = G(z)\mathscr{Z}(u). \tag{TF eqn}$$

(TF eqn) is known as a frequency domain equation because $G(e^{j\omega})$ describes how the sinusoidal input $u(t) = \sin(\omega t)$ is "processed" by the system:

$$y(t) = |G(e^{j\omega})| \sin\big(\omega t + \angle G(e^{j\omega})\big).$$

The system induced by G (with an input/output partition defined by Π) is

$$\mathscr{B}_{\mathrm{i/o}}(G, \Pi) := \Big\{\, w = \Pi \operatorname{col}(u, y) \in (\mathbb{R}^{\mathrm{w}})^{\mathbb{N}} \;\Big|\; y = \mathscr{Z}^{-1}\big(G(z)\mathscr{Z}(u)\big) \,\Big\}. \tag{TF repr}$$

(TF repr) is called a transfer function representation of the system $\mathscr{B} := \mathscr{B}_{\text{i/o}}(G, \Pi)$. In terms of the parameters of the input/state/output representation $\mathscr{B}_{\text{i/s/o}}(A, B, C, D) = \mathscr{B}_{\text{i/o}}(G)$, the transfer function is

$$G(z) = C(Iz - A)^{-1}B + D. \tag{TF$\leftarrow$I/S/O}$$

Define the matrix valued time series $H \in (\mathbb{R}^{\mathtt{p}\times\mathtt{m}})^{\mathbb{N}}$ by $H := \mathscr{Z}^{-1}(G)$, i.e.,

$$G(z) = H(0) + H(1)z^{-1} + H(2)z^{-2} + \cdots . \tag{TF$\leftarrow$CONV}$$

The time series H is a parameter in an alternative, time-domain representation of the system $\mathscr{B}_{\text{i/o}}(G, \Pi)$. Let $\star$ be the convolution operator. Then

$$y(t) := (H \star u)(t) = \sum_{\tau=0}^{t-1} H(\tau)u(t-\tau). \tag{CONV eqn}$$

The system induced by H (with an input/output partition defined by Π) is

$$\mathscr{B}_{\text{i/o}}(H, \Pi) := \left\{\, w = \Pi \operatorname{col}(u, y) \in (\mathbb{R}^{\mathtt{w}})^{\mathbb{N}} \;\middle|\; y = H \star u \,\right\}. \tag{CONV repr}$$

(CONV repr) is called a convolution representation of the system $\mathscr{B} := \mathscr{B}_{\text{i/o}}(H, \Pi)$.

The matrices $H(t)$, $t \geq 0$, are called Markov parameters of the representation $\mathscr{B}_{\text{i/o}}(H)$. In terms of the parameters of the state space representation $\mathscr{B}_{\text{i/s/o}}(A, B, C, D) = \mathscr{B}_{\text{i/o}}(H)$, the Markov parameters are

$$H(0) = D, \qquad H(t) = CA^{t-1}B, \quad t \geq 1. \tag{CONV$\leftarrow$I/S/O}$$

In addition to the transfer function (TF repr) and convolution (CONV repr) representations, a controllable system $\mathscr{B} \in \mathscr{L}^{\mathtt{w}}$ allows an image representation [Wil91, Theorem V.3]; i.e., there is a polynomial matrix $M \in \mathbb{R}^{\mathtt{w}\times g}[z]$, such that $\mathscr{B} = \operatorname{image}\big(M(\sigma)\big)$, where

$$\operatorname{image}\big(M(\sigma)\big) := \{\, w \in (\mathbb{R}^{\mathtt{w}})^{\mathbb{N}} \mid \exists\, l \in (\mathbb{R}^{\mathtt{w}})^{\mathbb{N}}, \text{ such that } w = M(\sigma)l \,\}. \tag{IMG repr}$$

The image representation is minimal if the number $\mathtt{l}$ of latent variables is minimal; i.e., there are no extra external variables in the representation than necessary. The image representation $\operatorname{image}\big(M(\sigma)\big)$ of $\mathscr{B}$ is minimal if and only if M is full column rank.

7.7 Representation Theorem

The following theorem summarizes the results presented in the previous sections of this chapter.

Theorem 7.3 (LTI system representations). *The following statements are equivalent:*

(i) *$\mathscr{B}$ is a complete LTI system with $\mathtt{w}$ variables, $\mathtt{m}$ inputs, and $\mathtt{p} := \mathtt{w} - \mathtt{m}$ outputs, i.e., $\mathscr{B} \in \mathscr{L}^{\mathtt{w}}$ and $\mathbf{m}(\mathscr{B}) = \mathtt{m}$;*

(ii) *there is a (full row rank) polynomial matrix $R \in \mathbb{R}^{\mathtt{p}\times\mathtt{w}}[z]$, such that $\mathscr{B} = \ker\big(R(\sigma)\big)$;*

(iii) *there are polynomial matrices* $Q \in \mathbb{R}^{\mathtt{p}\times\mathtt{m}}[z]$ *and* $P \in \mathbb{R}^{\mathtt{p}\times\mathtt{p}}[z]$, $\det(P) \neq 0$, $P^{-1}Q$ *proper, and a permutation matrix* $\Pi \in \mathbb{R}^{\mathtt{w}\times\mathtt{w}}$, *such that* $\mathscr{B} = \mathscr{B}_{\mathrm{i/o}}(P, Q, \Pi)$;

(iv) *there is a natural number* $\mathtt{n}$, *matrices* $A \in \mathbb{R}^{\mathtt{n}\times\mathtt{n}}$, $B \in \mathbb{R}^{\mathtt{n}\times\mathtt{m}}$, $C \in \mathbb{R}^{\mathtt{p}\times\mathtt{n}}$, *and* $D \in \mathbb{R}^{\mathtt{p}\times\mathtt{m}}$, *and a permutation matrix* $\Pi \in \mathbb{R}^{\mathtt{w}\times\mathtt{w}}$, *such that* $\mathscr{B} = \mathscr{B}_{\mathrm{i/s/o}}(A, B, C, D, \Pi)$;

(v) *there is a natural number* $\mathtt{l} \in \mathbb{N}$ *and polynomial matrices* $R \in \mathbb{R}^{\mathtt{p}\times\mathtt{m}}[z]$ *and* $M \in \mathbb{R}^{\mathtt{p}\times\mathtt{l}}[z]$, *such that* $\mathscr{B} = \mathscr{B}(R, M)$;

(vi) *there is a natural number* $\mathtt{l} \in \mathbb{N}$ *and matrices* $A' \in \mathbb{R}^{\mathtt{n}\times\mathtt{n}}$, $B' \in \mathbb{R}^{\mathtt{n}\times\mathtt{m}}$, $C' \in \mathbb{R}^{\mathtt{p}\times\mathtt{n}}$, *and* $D' \in \mathbb{R}^{\mathtt{p}\times\mathtt{m}}$, *such that* $\mathscr{B} = \mathscr{B}_{\mathrm{ss}}(A', B', C', D')$.

If in addition $\mathscr{B}$ *is controllable, then the following statement is equivalent to* (i)–(vi):

(vii) *there is a full column rank matrix* $M \in \mathbb{R}^{\mathtt{w}\times\mathtt{m}}[z]$, *such that* $\mathscr{B} = \operatorname{image}\big(M(\sigma)\big)$.

A controllable system $\mathscr{B}$ *has transfer function* $\mathscr{B}_{\mathrm{i/o}}(G, \Pi)$ *and convolution* $\mathscr{B}_{\mathrm{i/o}}(H, \Pi)$ *representations. These representations are unique when an input/output partitioning of the variables is fixed.*

The proofs of most of the implications of Theorem 7.3 can be found in [Wil86a] and [Wil91]. These proofs are constructive and give explicit algorithms for passing from one representation to another.

Figure 7.1 shows schematically the representations discussed up to now. To the left of the vertical line are representations that have no explicit input/output separation of the variables and to the right of the vertical line are representations with input/output separation of the variables. In the first row are state space representations. The representations below the second horizontal line exist only for controllable systems.

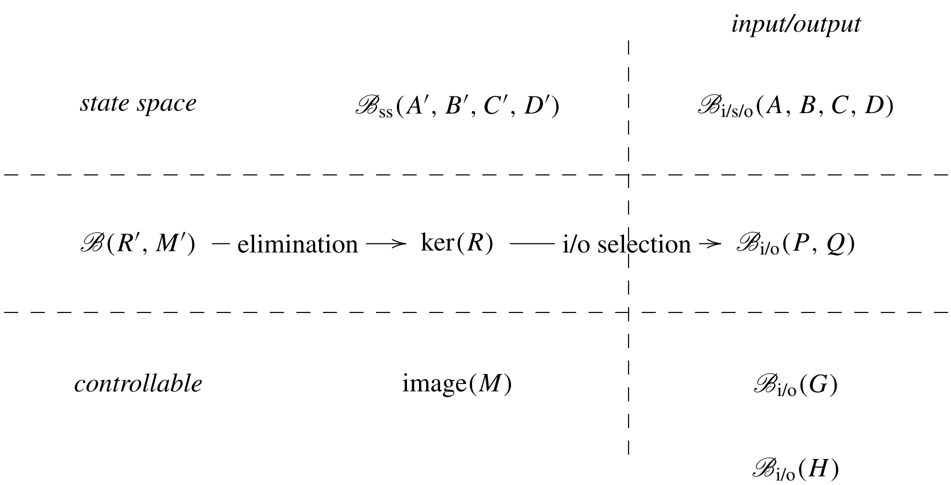

Figure 7.1. *Representations by categories: state space, input/output, and controllable.*

Transition from a latent variable representation to a representation without latent variables, for example $\mathscr{B}(R', M') \to \ker(R)$, involves elimination. Transition from a representation without an input/output separation to a representation with such a separation, for example $\ker(R) \to \mathscr{B}_{\text{i/o}}(P, Q)$, involves input/output selection. Transitions from a representation in the second or third rows to a representation in the first row is a realization problem.

In principle, all transitions from one type of representation to another are of interest (and imply algorithms that implement them). Moreover, all representations have special forms such as the controller canonical form, the observer canonical form, balanced representation, etc. Making the graph in Figure 7.1 connected suffices in order to be able to derive any representation, starting from any other one. Having a specialized algorithm that does not derive intermediate representations, however, has advantages from a computational point of view.

7.8 Parameterization of a Trajectory

A trajectory w of $\mathscr{B} \in \mathscr{L}^{\mathrm{w}}$ is parameterized by

1. a corresponding input u and

2. initial conditions x_{ini}.

If $\mathscr{B}$ is given in an input/state/output representation $\mathscr{B} = \mathscr{B}_{\text{i/s/o}}(A, B, C, D)$, then an input u is given and the initial conditions can be chosen as the initial state $x(1)$. The variation of constants formula

$$w = \operatorname{col}(u, y), \qquad y(t) = CA^{t-1}x_{\text{ini}} + \sum_{\tau=1}^{t-1} \underbrace{CA^{t-\tau-1}B}_{H(t-\tau)} u(\tau) + Du(t), \qquad t \geq 1, \quad \text{(VC)}$$

gives a parameterization of w. Note that the second term in the expression for y is the convolution of H and u. It alone gives the zero initial conditions response. The ith column of the impulse response H is the zero initial conditions response of the system to input $u = e_i\delta$, where e_i is the ith unit vector.

For a given pair of matrices (A, B), $A \in \mathbb{R}^{\mathrm{n}\times\mathrm{n}}$, $B \in \mathbb{R}^{\mathrm{n}\times\mathrm{m}}$, and $t \in \mathbb{N}$, define the extended controllability matrix (with t block columns)

$$\mathscr{C}_t(A, B) := \begin{bmatrix} B & AB & \cdots & A^{t-1}B \end{bmatrix} \qquad (\mathscr{C})$$

and let $\mathscr{C}(A, B) := \mathscr{C}_\infty(A, B)$. The pair (A, B) is controllable if $\mathscr{C}(A, B)$ is full row rank. By the Cayley–Hamilton theorem [Bro70, page 72], $\operatorname{rank}\big(\mathscr{C}(A, B)\big) = \operatorname{rank}\big(\mathscr{C}_{\mathrm{n}}(A, B)\big)$, so that it suffices to check the rank of the finite matrix $\mathscr{C}_{\mathrm{n}}(A, B)$. The smallest natural number i, for which $\mathscr{C}_i(A, B)$ is full row rank, is denoted by $\nu(A, B)$ and is called the controllability index of the pair (A, B). The controllability index is an invariant under state transformation; i.e., $\nu(A, B) = \nu(SAS^{-1}, SB)$ for any nonsingular matrix S. In fact, $\nu(A, B)$ is an invariant of any system $\mathscr{B}_{\text{i/s/o}}(A, B, \bullet, \bullet)$, so that it is legitimate to use the notation $\nu(\mathscr{B})$ for $\mathscr{B} \in \mathscr{L}^{\mathrm{w}}$. Clearly, $\nu(\mathscr{B}) \leq \mathbf{n}(\mathscr{B})$.

Similarly, for a given pair of matrices (A, C), $A \in \mathbb{R}^{\mathtt{n}\times\mathtt{n}}$, $C \in \mathbb{R}^{\mathtt{p}\times\mathtt{n}}$, and $t \in \mathbb{N}$, define the extended observability matrix (with t block rows)

$$\mathscr{O}_t(A, C) := \operatorname{col}(C, CA, \ldots, CA^{t-1}) \tag{$\mathscr{O}$}$$

and let $\mathscr{O}(A, C) := \mathscr{O}_\infty(A, C)$. The pair (A, C) is observable if $\mathscr{O}(A, C)$ is full column rank. Again, $\operatorname{rank}\big(\mathscr{O}(A, C)\big) = \operatorname{rank}\big(\mathscr{O}_{\mathtt{n}}(A, C)\big)$, so that it suffices to check the rank of $\mathscr{O}_{\mathtt{n}}(A, C)$. The smallest natural number i, for which $\mathscr{O}_i(A, C)$ is full row rank, is denoted by $\mu(A, C)$ and is called the observability index of the pair (A, C). The observability index is an invariant under state transformation; i.e., $\mu(A, C) = \mu(SAS^{-1}, CS^{-1})$ for any nonsingular matrix S. In fact, $\nu(A, C)$ is equal to the lag of any system $\mathscr{B}_{\text{i/s/o}}(A, \bullet, C, \bullet)$, so that it is invariant and it is legitimate to use the notation $\nu(\mathscr{B})$ for $\mathscr{B} \in \mathscr{L}^{\mathtt{w}}$. Clearly, $\mathbf{l}(\mathscr{B}) = \nu(\mathscr{B}) \leq \mathbf{n}(\mathscr{B})$.

If the pairs (A, B) and (A, C) are understood from the context, they are skipped in the notation of the extended controllability and observability matrices.

We define also the lower triangular block-Toeplitz matrix

$$\mathscr{T}_{t+1}(H) := \begin{bmatrix} H(0) & & & & \\ H(1) & H(0) & & & \\ H(2) & H(1) & H(0) & & \\ \vdots & \vdots & \ddots & \ddots & \\ H(t) & H(t-1) & \cdots & H(1) & H(0) \end{bmatrix} \tag{$\mathscr{T}$}$$

and let $\mathscr{T}(H) = \mathscr{T}_\infty(H)$. With this notation, equation (VC) can be written compactly as

$$\begin{bmatrix} u \\ y \end{bmatrix} = \begin{bmatrix} 0 & I \\ \mathscr{O}(A, C) & \mathscr{T}(H) \end{bmatrix} \begin{bmatrix} x_{\text{ini}} \\ u \end{bmatrix}. \tag{VC'}$$

If the behavior $\mathscr{B}$ is not given by an input/state/output representation, then the parameterization of a trajectory $w \in \mathscr{B}$ is more involved. For example, in an input/output representation $\mathscr{B} = \mathscr{B}_{\text{i/o}}(P, Q)$, w can be parameterized by the input u and the $\mathtt{l} = \deg(P)$ values of the time series $w_{\text{ini}} := \big(w(-\mathtt{l}+1), \ldots, w(0)\big)$ preceding w as follows:

$$y = \mathscr{O}_{\text{i/o}} w_{\text{ini}} + \mathscr{T}(H)u. \tag{VC i/o}$$

Here $\mathscr{O}_{\text{i/o}}$ is a matrix that induces a mapping from w_{ini} to the corresponding initial conditions response. Let $\mathscr{B}_{\text{i/s/o}}(A, B, C, D) = \mathscr{B}_{\text{i/o}}(P, Q)$. Comparing (VC') and (VC), we see that the matrix $\mathscr{O}_{\text{i/o}}$ can be factored as $\mathscr{O}_{\text{i/o}} = \mathscr{O}(A, C)X$, where X is a matrix that induces the map $w_{\text{ini}} \mapsto x_{\text{ini}}$, called a state map [RW97].

The graph in Figure 7.2 illustrates the two representations introduced in this section for a trajectory w of the system $\mathscr{B} \in \mathscr{L}^{\mathtt{w},\mathtt{n}}_{\mathtt{m}}$.

7.9 Complexity of a Linear Time-Invariant System

In Chapter 2, we introduced the complexity of a linear system $\mathscr{B}$ as the dimension of $\mathscr{B}$ as a subspace of the universum set. For an LTI system $\mathscr{B} \in \mathscr{L}^{\mathtt{w}}$ and for $T \geq \mathbf{l}(\mathscr{B})$,

$$\dim(\mathscr{B}|_{[1,T]}) = \mathbf{m}(\mathscr{B})T + \mathbf{n}(\mathscr{B}), \tag{dim $\mathscr{B}$}$$

$$\begin{tikzcd}
 & w \in \mathscr{B} \in \mathscr{L}_{\mathtt{m}}^{\mathtt{w},\mathtt{n}} \arrow[dl, leftrightarrow] \arrow[dr, leftrightarrow] & \\
\mathscr{B} = \mathscr{B}_{\mathrm{i/o}}(P, Q, \Pi),\ w_{\mathrm{ini}} \arrow[rr, leftrightarrow] & & \mathscr{B} = \mathscr{B}_{\mathrm{i/s/o}}(A, B, C, D, \Pi),\ x_{\mathrm{ini}}
\end{tikzcd}$$

Figure 7.2. *Links among $w \in \mathscr{B} \in \mathscr{L}_{\mathtt{m}}^{\mathtt{w},\mathtt{n}}$ and its parameterizations in input/output and input/state/output form.*

which shows that the pair of natural numbers $\big(\mathbf{m}(\mathscr{B}), \mathbf{n}(\mathscr{B})\big)$ (the input cardinality and the total lag) specifies the complexity of the system. The model class $\mathscr{L}_{\mathtt{m}}^{\mathtt{w},\mathtt{n}}$ contains LTI systems of complexity bounded by the pair $(\mathtt{m}, \mathtt{n})$.

In the context of system identification problems, aiming at a kernel representation of the model, we need an alternative specification of the complexity by the input cardinality $\mathbf{m}(\mathscr{B})$ and the lag $\mathbf{l}(\mathscr{B})$. In general,

$$\big(\mathbf{l}(\mathscr{B}) - 1\big)\mathbf{p}(\mathscr{B}) < \mathbf{n}(\mathscr{B}) \leq \mathbf{l}(\mathscr{B})\mathbf{p}(\mathscr{B}),$$

so that

$$\dim(\mathscr{B}|_{[1,T]}) \leq \mathbf{m}(\mathscr{B})T + \mathbf{l}(\mathscr{B})\mathbf{p}(\mathscr{B})$$

and the pair $\big(\mathbf{m}(\mathscr{B}), \mathbf{l}(\mathscr{B})\big)$ bounds the complexity of the system $\mathscr{B}$.

The class of LTI systems with $\mathtt{w}$ variables, at most $\mathtt{m}$ inputs, and lag at most $\mathtt{l}$ is denoted by $\mathscr{L}_{\mathtt{m},\mathtt{l}}^{\mathtt{w}}$.

This class specifies a set of LTI systems of a bounded complexity.

7.10 The Module of Annihilators of the Behavior*

Define the set of annihilators of the system $\mathscr{B} \in \mathscr{L}^{\mathtt{w}}$ as

$$\mathscr{N}_{\mathscr{B}} := \{\, r \in \mathbb{R}^{\mathtt{w}}[z] \mid r^{\top}(\sigma)\mathscr{B} = 0 \,\}$$

and the set of annihilators with length less than or equal to l as

$$\mathscr{N}_{\mathscr{B}}^{l} := \{\, r \in \mathscr{N}_{\mathscr{B}} \mid \deg(r) < l \,\}.$$

The sets $\mathscr{N}_{\mathscr{B}}$ and $\mathscr{N}_{\mathscr{B}}^{l}$ are defined as subsets of $\mathbb{R}^{\mathtt{w}}[z]$. With some abuse of notation, we also consider the annihilators as vectors; i.e., for $r(z) =: r_0 + r_1 z + \cdots + r_l z^l \in \mathscr{N}_{\mathscr{B}}$, we also write $\operatorname{col}(r_0, r_1, \ldots, r_l) \in \mathscr{N}_{\mathscr{B}}$.

Lemma 7.4. *Let $r(z) = r_0 + r_1 z + \cdots + r_{l-1} z^{l-1}$. Then $r \in \mathscr{N}_{\mathscr{B}}^{l}$ if and only if*

$$\operatorname{col}^{\top}(r_0, r_1, \ldots, r_{l-1})\mathscr{B}|_{[1,l]} = 0.$$

The set of annihilators $\mathscr{N}_{\mathscr{B}}$ is the dual $\mathscr{B}^{\perp}$ of the behavior $\mathscr{B}$.

The proof of the following facts can be found in [Wil86a]. The structure of $\mathscr{N}_{\mathscr{B}}$ is that of the module of $\mathbb{R}[z]$ generated by $\mathtt{p}$ polynomial vectors, say $r^{(1)}, \dots, r^{(\mathtt{p})}$. The polynomial matrix $R := [r^{(1)} \ \dots \ r^{(\mathtt{p})}]^{\top}$ yields a kernel representation of the behavior $\mathscr{B}$, i.e., $\mathscr{B} = \ker\big(R(\sigma)\big)$.

Without loss of generality, assume that R is row proper; i.e., $\ker\big(R(\sigma)\big)$ is a shortest lag kernel representation. By the row properness of R, the set of annihilators $\mathscr{N}_{\mathscr{B}}^{l}$ can be constructed from the $r^{(k)}$'s and their shifts

$$\mathscr{N}_{\mathscr{B}}^{l} = \operatorname{image}\big(r^{(1)}(z), zr^{(1)}(z), \dots, z^{l-\mu_1-1}r^{(1)}(z)\ ;\ \dots\ ;\\ r^{(\mathtt{p})}(z), zr^{(\mathtt{p})}(z), \dots, z^{l-\mu_{\mathtt{p}}-1}r^{(\mathtt{p})}(z)\big).$$

The dimension of $\mathscr{N}_{\mathscr{B}}^{l}$ is $l-\mu_1+l-\mu_2+\cdots+l-\mu_p = \mathtt{p}l-\mathtt{n}$.

In the proof of the fundamental lemma (see Appendix A.3), we need the following simple fact.

Lemma 7.5. *Let $r^{(1)}, \dots, r^{(\mathtt{p})}$, where $\deg(r_i) =: \mu_i$, be independent over the ring of polynomials. Then*

$$r^{(1)}(z), zr^{(1)}(z), \dots, z^{l-\mu_1-1}r^{(1)}(z)\ ;\ \dots\ ;\ r^{(\mathtt{p})}(z), zr^{(\mathtt{p})}(z), \dots, z^{l-\mu_{\mathtt{p}}-1}r^{(\mathtt{p})}(z)$$

are independent over the field of reals.