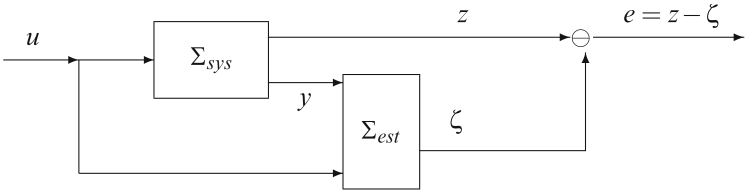
Chapter 7
Observer Theory

Observer theory is one of the most basic, and important, aspects of linear systems theory. The problem addressed in this chapter is that of indirect observation, or partial state estimation. It arises from the fact that in a control system Σ_{sys}, the observed variables are not necessarily the variables one needs to estimate for control, or other, purposes. A standard situation often encountered is that of partial state estimation, where a few, or all, state variables are to be estimated from the output variables. Of course, if one can estimate the state, then one automatically has the ability to estimate a function of the state. However, especially in a large and complex system, estimating the full state may be a daunting task and more than what is needed. The task of state estimation is also instrumental for practical implementations of state feedback control using estimates of the unknown state functions. More generally, our aim is to find a mechanism, called an **observer**, that allows us to use information on observed variables y and the inputs u in order to estimate linear functions z of the state variables. Loosely speaking, an observer for the system is itself a linear system Σ_{est}, which is driven by the variables u and y and whose output is the desired estimate ζ of z, with the estimation error being $e = z - \zeta$. The following diagram describes the observation process:



The error trajectory depends on the system transfer function, the observer transfer function, and on the initial conditions of both Σ_{sys} and Σ_{est}. There is great freedom in the choice of the observer, the only constraint being the compatibility with

P.A. Fuhrmann, U. Helmke, *The Mathematics of Networks of Linear Systems*, Universitext, DOI 10.1007/978-3-319-16646-9_7

the signals u, y, z. Of course, if the observer is not chosen appropriately, the error trajectory may be large, which makes the estimate useless. Our aim is to characterize the properties of the observer in terms of these transfer functions. Even of greater importance is, whenever possible, the construction of observers having desired properties.

The issues of observation and estimation play a crucial role in analyzing observation processes for networks and are therefore of paramount importance for all questions concerning fault detection, monitoring, and measurement processes. Observer theory has a long history laden with vagueness, imprecision, and incomplete or even false proofs; see Trumpf (2013) for a short list of these. The principal reason for the difficulty in clarifying the structure theory of observers, functional observers in particular, seems to be that a full understanding of the problems requires the ability to integrate many topics and viewpoints that cover most of algebraic systems theory. These include state-space theory (including the dual Brunovsky form, realizations and partial realizations, Sylvester equations, and some old results of Roth and Halmos), polynomial and rational model theory, geometric control (conditioned invariant subspaces, as well as detectability and outer observability subspaces), and Hankel matrices. Another interesting point of view on observers is the behavioral approach, as developed and presented in the papers by Valcher and Willems (1999) and Fuhrmann (2008). However, to keep the exposition within reasonable limits, this direction will not be pursued. We will draw heavily on Fuhrmann and Helmke (2001a) and Trumpf (2002, 2013). Of course, the theory of observers depends strongly on the interpretation of what a linear system is and how it is represented. The state-space representation of a finite-dimensional, time-invariant linear system is chosen as our starting point. In addition, we focus on discrete-time systems because this simplifies matters when comparing trajectory-based formulations with statements for rational functions. Moreover, this enables us to state several results for systems over fields more general than the field $\mathbb{R}$ of real numbers. Because most of the statements remain true for continuous-time systems (and the field of real numbers), this restriction to discrete-time systems presents no real loss of generality.

7.1 Classical State Observers

In the preceding chapter, we discussed how to design state feedback control laws $u_t = -Kx_t + v_t$ for linear discrete-time systems of the form

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t, \\ y_t &= Cx_t, \end{aligned} \tag{7.1}$$

where $A \in \mathbb{F}^{n\times n}, B \in \mathbb{F}^{n\times m}, C \in \mathbb{F}^{p\times n}$, and $\mathbb{F}$ is a field. Obviously, to implement control laws such as $u_t = -Kx_t + v_t$, one needs to know the state x_t, or at least an approximation to the state. State observers are designed precisely to fulfill such a purpose and are thus indispensable for practical applications of state feedback control. In this section, the classical construction of a full state observer, due to Luenberger (1964), is described; see also Luenberger (1971). To investigate standard stability properties of observers, it will be assumed that $\mathbb{F}$ is a subfield of the field $\mathbb{C}$ of complex numbers.

A **state observer** for system (7.1) is an n-dimensional linear control system

$$z_{t+1} = Fz_t + Gy_t + Hu_t, \tag{7.2}$$

with matrices $F \in \mathbb{F}^{n\times n}, G \in \mathbb{F}^{n\times p}, H \in \mathbb{F}^{n\times m}$, such that, for each initial condition $x_0, z_0 \in \mathbb{F}^n$ and every input sequence $u = (u_t)$, $\lim_{t\to\infty}(x_t - z_t) = 0$. A state observer is therefore a dynamical system that is driven by the input and output of (7.1) and whose state vectors z_t will asymptotically converge to the state vectors of (7.1).

How can one construct such an observer? Luenberger's ingenious idea was to consider systems with $F = A - LC$, $G = L$, $H = B$, i.e.,

$$\begin{aligned} z_{t+1} &= Az_t + Bu_t + L(y_t - \hat{y}_t), \\ \hat{y}_t &= Cz_t, \end{aligned} \tag{7.3}$$

with an **observer gain matrix** $L \in \mathbb{F}^{n\times p}$. Thus (7.3) consists of an identical copy of a system that is driven by the **innovations** $y_t - \hat{y}_t$. System (7.3) is often called the **Luenberger observer**. It has only one free design parameter, i.e., the observer gain matrix. To see how to choose the observer gain in order to achieve a state observer, one must consider the evolution of the **estimation error**

$$e_t = x_t - z_t, \quad t \in \mathbb{N}.$$

The dynamics of the estimation error is

$$e_{t+1} = (A - LC)e_t.$$

Thus the estimation error converges to zero if and only if L is chosen such that $A - LC$ has all its eigenvalues in the open unit disc. This leads to the following classical result on state observers.

Theorem 7.1. *The Luenberger observer (7.3) is a state observer for system (7.1) if and only if the observer gain $L \in \mathbb{F}^{n\times p}$ is such that $A - LC$ is asymptotically stable. Such an observer gain exists if and only if (C, A) is detectable.*

Proof. The observer condition for (7.3) is equivalent to $\lim_{t\to\infty} e_t = 0$ for all initial conditions $e_0 \in \mathbb{F}^n$. Thus (7.3) defines a state observer if and only if $A - LC$ is asymptotically stable, i.e., has all eigenvalues in the open unit disc. There exists such a stabilizing observer gain L if and only if (C, A) is detectable. ∎

Having a state observer at hand, how can one use it for the purpose of state feedback control? Here we consider the closed-loop control system of the form

$$\begin{aligned} x_{t+1} &= (A - B\mathscr{F})x_t + Bu_t, \\ y_t &= Cx_t, \end{aligned} \tag{7.4}$$

where $\mathscr{F} \in \mathbb{F}^{m\times n}$ is a desired state feedback gain. For example, $\mathscr{F}$ may be chosen such that the closed-loop characteristic polynomial $\det(zI - A + B\mathscr{F})$ is a prescribed monic polynomial of degree n. How must one choose the observer gain L? A beautiful simple result, the **separation principle**, provides a solution. It states that the designs of the state feedback and observer gain matrices can be done separately. But even then there is a problem because implementing (7.4) requires knowledge of the feedback term $-\mathscr{F}x_t$. This can be resolved by replacing $-\mathscr{F}x_t$ with the observer estimate $-\mathscr{F}z_t$. This then leads to the composed controller/observer dynamics with joint state variables $\xi = \operatorname{col}(\hat{x}, z)$:

$$\begin{aligned} \hat{x}_{t+1} &= A\hat{x}_t - B\mathscr{F}z_t + Bu_t, \\ z_{t+1} &= Az_t - B\mathscr{F}z_t + Bu_t + L(C\hat{x}_t - Cz_t), \\ y_t &= C\hat{x}_t. \end{aligned} \tag{7.5}$$

Written in matrix form we obtain

$$\xi_{t+1} = \mathscr{A}_c\xi_t + \mathscr{B}_c u_t,$$

with

$$\mathscr{A}_c = \begin{pmatrix} A & -B\mathscr{F} \\ LC & A - LC - B\mathscr{F} \end{pmatrix}, \quad \mathscr{B}_c = \begin{pmatrix} B \\ B \end{pmatrix}. \tag{7.6}$$

The fundamental result for a combined controller and observer design is stated next.

Theorem 7.2 (Separation Principle). *Let $\mathscr{A}_c$ be defined by (7.6).*

1. The identity

$$\det(zI - \mathscr{A}_c) = \det(zI - A + B\mathscr{F})\det(zI - A + LC)$$

is true. In particular, for each state feedback gain $\mathscr{F}$ and every output injection gain $L \in \mathbb{F}^{n\times p}$ such that $A - LC$ is asymptotically stable, the composed controller/observer dynamics (7.5) satisfies

$$\lim_{t\to\infty}(z_t - \hat{x}_t) = 0$$

for arbitrary initial states $\hat{x}_0, z_0$ and input sequences (u_t).

2. *The transfer function from u to y of (7.5) is*

$$G(z) = C(zI - A + B\mathscr{F})^{-1}B.$$

More generally, the Z-transforms of (u_t) and (y_t) are related as

$$\begin{aligned} y(z) &= C(zI - A + B\mathscr{F})^{-1}Bu(z) \\ &\quad + C(zI - A + B\mathscr{F})^{-1}\hat{x}_0 + C(zI - A + B\mathscr{F})^{-1}B\mathscr{F}(zI - A + LC)^{-1}(z_0 - \hat{x}_0). \end{aligned}$$

3. *Assuming that $A - B\mathscr{F}$ is stable, the system*

$$\begin{aligned} x_{t+1} &= (A - B\mathscr{F})x_t + Bu_t, \\ y_t &= Cx_t \end{aligned}$$

satisfies

$$\lim_{t\to\infty}(x_t - \hat{x}_t) = 0$$

for all initial states $\hat{x}_0, z_0$ and input sequences (u_t).

Proof. For the invertible matrix

$$S = \begin{pmatrix} I & 0 \\ -I & I \end{pmatrix}$$

we compute

$$S\mathscr{A}_c S^{-1} = \begin{pmatrix} A - B\mathscr{F} & -B\mathscr{F} \\ 0 & A - LC \end{pmatrix}, \quad S\mathscr{B}_c = \begin{pmatrix} B \\ 0 \end{pmatrix}, \quad \begin{pmatrix} C & 0 \end{pmatrix} S^{-1} = \begin{pmatrix} C & 0 \end{pmatrix}.$$

The transfer function of (7.5) is thus equal to $C(zI - A + B\mathscr{F})^{-1}B$. For the remaining parts, we proceed to consider the dynamics of the error term $\varepsilon_t := z_t - \hat{x}_t$, which is given as

$$\varepsilon_{t+1} = (A - LC)\varepsilon_t.$$

Our assumption on L implies $\lim_{t\to\infty}(z_t - \hat{x}_t) = 0$. For the last claim, consider the error sequence $e_t := x_t - \hat{x}_t$ with associated error dynamics. It satisfies

$$e_{t+1} = (A - B\mathscr{F})e_t + B\mathscr{F}(z_t - \hat{x}_t).$$

Since $A - B\mathscr{F}$ is stable and $\lim_{t\to\infty}(z_t - \hat{x}_t) = 0$, we conclude that $\lim_{t\to\infty}(x_t - \hat{x}_t) = 0$. This completes the proof. ∎

Of course, the preceding results on full state observers are only the starting point for a deeper theory of observers that enables one to estimate a finite number of linear state functionals. This more general observer theory is developed in subsequent sections.

7.2 Observation Properties

One of Kalman's major achievements has been the introduction of the concepts of reachability and observability, as distinct from compensator or observer design. This separation is reflected in Valcher and Willems (1999), where the observability or detectability of one set of system variables from another is studied before observer design is attempted. Such an approach is adopted in this section. Clearly, observers depend on the observability properties of a system, and a few gradations of observability will be introduced. Naturally, one expects that the stronger the observation properties of a system are, the better behaved should be the corresponding observers. How the observation properties of the system are reflected in the corresponding observers will be examined in Subsection 7.3.

To state the subsequent definitions and results over a field $\mathbb{F}$, the meaning of convergence in a finite-dimensional vector space $\mathbb{F}^n$ must be clarified. Here we proceed as in Chapter 6 with respect to the dual situation, i.e., that of state feedback stabilization. A general field $\mathbb{F}$ is endowed with the discrete topology, i.e., the unique topology on $\mathbb{F}$ whose open (and closed) subsets are defined by subsets of $\mathbb{F}$. In contrast, for a subfield $\mathbb{F} \subset \mathbb{C}$ we introduce the Euclidean topology on $\mathbb{F}$, which is defined by the standard Euclidean distance $|x-y|$ of complex numbers $x,y \in \mathbb{C}$. In either case, a discrete-time dynamical system $x_{t+1} = Ax_t$ on $\mathbb{F}^n$ is called stable whenever the sequence $x_t = A^t x_0$ converges to zero for all initial conditions $x_0 \in \mathbb{F}^n$. Thus a discrete-time dynamical system $x_{t+1} = Ax_t$ is asymptotically stable if and only if either A is Schur stable, for the Euclidean topology on $\mathbb{F} \subset \mathbb{C}$, or if A is nilpotent, for the discrete topology on $\mathbb{F}$. Let (e_t) denote a sequence of points in $\mathbb{F}^m$ that defines a proper rational function

$$e(z) = \sum_{t=0}^{\infty} e_t z^{-t} \in z^{-1}\mathbb{F}[[z^{-1}]]^m,$$

and let $(A,B,C,D) \in \mathbb{F}^{n\times n} \times \mathbb{F}^{n\times 1} \times \mathbb{F}^{m\times n} \times \mathbb{F}^m$ denote a minimal realization of

$$e(z) = D + C(zI - A)^{-1}B.$$

Thus $e_0 = D$, $e_t = CA^{t-1}B$, $t \geq 1$, is the sequence of Markov parameters of $e(z)$.

Proposition 7.3. *Let $e(z) \in \mathbb{F}(z)^m$ be proper rational. The sequence of Markov parameters (e_t) satisfies* $\lim_{t\to\infty} e_t = 0$ *if and only if:*

1. *All poles of $e(z)$ have absolute value < 1 or, equivalently, A is Schur stable. This assumes that $\mathbb{F} \subset \mathbb{C}$ carries the Euclidean topology;*
2. *All poles of $e(z)$ are in $z = 0$ or, equivalently, A is nilpotent. This assumes that $\mathbb{F}$ carries the discrete topology.*

Proof. Since $e_t = CA^{t-1}B$ for $t \geq 1$, and, by the minimality of (A,B,C), the observability and reachability matrices

$$\begin{pmatrix} C \\ \vdots \\ CA^{n-1} \end{pmatrix}, \quad \begin{pmatrix} B \cdots A^{n-1}B \end{pmatrix}$$

have full column rank and full row rank, respectively. Thus the sequence of Hankel matrices

$$\begin{pmatrix} e_{nt+1} & \cdots & e_{n(t+1)} \\ \vdots & & \vdots \\ e_{n(t+1)} & \cdots & e_{n(t+2)-1} \end{pmatrix} = \begin{pmatrix} C \\ \vdots \\ CA^{n-1} \end{pmatrix} A^{tn} \begin{pmatrix} B \cdots A^{n-1}B \end{pmatrix}, \quad t \in \mathbb{N},$$

converges to zero if and only if $\lim_{t\to\infty} A^{tn} = 0$. If the field $\mathbb{F}$ carries the discrete topology, this is equivalent to A being nilpotent, while for the Euclidean topology on $\mathbb{F} \subset \mathbb{C}$ this is equivalent to the eigenvalues of A being in the open complex unit disc. ■

The next characterization will be useful later on.

Proposition 7.4. *Let $Q(z) \in \mathbb{F}[z]^{m\times m}$ be nonsingular, and let X^Q denote the associated rational model.*

1. *Let $\mathbb{F} \subset \mathbb{C}$ be endowed with the Euclidean topology. The following statements are equivalent:*
   (a) *All elements $h(z) \in X^Q$ are stable, i.e., the coefficients h_t of $h(z)$ satisfy* $\lim_{t\to\infty} h_t = 0$.
   (b) $\det Q(z)$ *is a Schur polynomial, i.e., all its roots are in the open unit disc.*
2. *Let $\mathbb{F}$ be endowed with the discrete topology. The following statements are equivalent:*
   (a) *All elements $h(z) \in X^Q$ are stable, i.e., the coefficients h_t of $h(z)$ satisfy $h_t = 0$ for t sufficiently large.*
   (b) $\det Q(z)$ *is a monomial, i.e., all its roots are equal to* 0.

Proof. Choosing a polynomial basis matrix $P(z) \in \mathbb{F}[z]^{m\times n}$ for the finite-dimensional polynomial model X_Q implies $Q(z)^{-1}P(z)$ is a basis matrix for the rational model X^Q, and therefore the elements of X^Q are of the form $h(z) = Q(z)^{-1}P(z)\xi$ for unique vectors $\xi \in \mathbb{F}^n$. By Proposition 4.36, there exists an observable pair $(C,A) \in \mathbb{F}^{m\times n} \times \mathbb{F}^{n\times n}$, with

$$C(zI-A)^{-1} = Q(z)^{-1}P(z).$$

Thus the coefficients of $Q(z)^{-1}P(z)\xi$ converge to zero for all choices of ξ if and only if $\lim_{t\to\infty} CA^t = 0$. This is equivalent to (C,A) being detectable. By Proposition 6.34, and applying the observability of (C,A), this is equivalent to $\det(zI-A)$ being a Schur polynomial, i.e., to all eigenvalues of A being in the open unit disc. Part (1) follows by observing the identity $\det Q(z) = \det(zI-A)$. Part (2) is proven similarly.

Next, some of the important observation properties of a system are introduced, e.g., the extent to which the observed variables y determine the relevant, or to-be-estimated, variables z.

Definition 7.5. Let Σ_{sys} be a linear system with the state-space representation

$$\Sigma_{sys} := \begin{cases} x_{t+1} = Ax_t + Bu_t, \\ y_t = Cx_t, \\ z_t = Kx_t, \end{cases} \tag{7.7}$$

with x_t, y_t, u_t, z_t taking values in $\mathbb{F}^n, \mathbb{F}^p, \mathbb{F}^m, \mathbb{F}^k$, respectively.

1. The variable z is ***T*-trackable** from (y,u) if there exists a nonnegative integer T such that for every two solutions $(x,y,u,z), (\bar{x},y,u,\bar{z})$ of (7.7) the condition $\bar{z}_t = z_t$ for $0 \le t \le T$ implies $\bar{z} = z$. The smallest such T is called the **tracking index** and is denoted by τ.
2. The variable z is **detectable** from (y,u) if each pair of solutions (x,y,u,z) and $(\bar{x},y,u,\bar{z})$ of (7.7) satisfies $\lim_{t\to\infty}(z_t - \bar{z}_t) = 0$.
3. The variable z is **reconstructible** from (y,u) if for each pair of solutions (x,y,u,z) and $(\bar{x},y,u,\bar{z})$ of (7.7) there exists a nonnegative integer T such that $z_t - \bar{z}_t = 0$ for $t > T$. The smallest such T is called the **reconstructibility index**.
4. The variable z is **observable** from (y,u) if each pair of solutions (x,y,u,z) and $(\bar{x},y,u,\bar{z})$ of (7.7) satisfies $\bar{z} = z$.

One says that a system Σ_{sys}, given by (7.7), is ***T*-trackable** if z is ***T*-trackable** from y. We similarly define **detectability**, **reconstructibility**, and **observability**.

In view of the Cayley–Hamilton theorem, it is obvious that every linear state function $z = Kx$ is T-trackable from the output y with $u = 0$ of an n-dimensional linear system

$$\begin{aligned} x_{t+1} &= Ax_t \\ y_t &= Cx_t, \end{aligned}$$

provided $T \ge n$. Thus a finite-dimensional linear system is always trackable, unless one requires an a priori bound on the tracking index. Therefore, defining trackability without imposing a constraint on the tracking index is meaningless. For reconstructibility, the situation becomes slightly different because not every linear system is reconstructible. Of course, the definition of detectability strongly depends on the topology of the field $\mathbb{F}$. If the field $\mathbb{F}$ is finite, or more generally if $\mathbb{F}$ carries the discrete topology, detectability is equivalent to reconstructibility. However, for a subfield $\mathbb{F} \subset \mathbb{C}$ with the standard Euclidean topology, this is no longer true. In fact, for the standard Euclidean topology on a subfield $\mathbb{F} \subset \mathbb{C}$, the detectability of z from y is equivalent to the condition that for an unobservable state x_0 the rational function $K(zI - A)^{-1}x_0$ has only poles in the open unit disc. Intuitively, it is clear that, since perfect knowledge of the system is assumed, the effect of the input variable on the estimate can be removed without affecting the observation properties. The following simple proposition is stated, with its trivial proof omitted.

Proposition 7.6. *Let Σ be a linear system with the state-space representation (7.7), together with the associated system Σ' given by*

$$(\Sigma') \quad \begin{cases} x_{t+1} = Ax_t, \\ y_t = Cx_t, \\ z_t = Kx_t. \end{cases} \tag{7.8}$$

Then:

1. *The following conditions are equivalent:*
   (a) *z is T-trackable from (y,u) with respect to Σ,*
   (b) *z is T-trackable from y with respect to Σ',*
   (c) *For all initial conditions x_0 such that $y_t = 0$ and $u_t = 0$ are satisfied for all t, the condition $z_0 = \cdots = z_T = 0$ implies $z_t = 0$ for all t;*
2. *The following conditions are equivalent:*
   (a) *z is detectable from (y,u) with respect to Σ,*
   (b) *z is detectable from y with respect to Σ',*
   (c) *For all initial conditions x_0 such that $y_t = 0$ and $u_t = 0$ are satisfied for all t, $\lim_{t\to\infty} z_t = 0$;*
3. *The following conditions are equivalent:*
   (a) *z is reconstructible from (y,u) with respect to Σ,*
   (b) *z is reconstructible from y with respect to Σ',*
   (c) *For all initial conditions x_0 that satisfy $y_t = 0$ and $u_t = 0$ for all t, then $z_t = 0$ for all $t > T$;*
4. *The following conditions are equivalent:*
   (a) *z is observable from (y,u) with respect to Σ,*
   (b) *z is observable from y with respect to Σ',*
   (c) *For all initial conditions x_0 such that $y_t = 0$ and $u_t = 0$ are satisfied for all t, $z_t = 0$ for all t.*

Next, the invariance of the preceding notions with respect to **output injection equivalence** is explored.

Proposition 7.7. *Let $S \in GL_n(\mathbb{F}), L \in \mathbb{F}^{n\times p}, R \in GL_p(\mathbb{F}), U \in GL_m(\mathbb{F})$ be output injection transformations. A system (A,B,C,K) is T-trackable, reconstructible, detectable, or observable, respectively, if and only if the output injection equivalent system $(S(A+LC)S^{-1}, SBU^{-1}, RCS^{-1}, KS^{-1})$ is.*

Proof. It is obvious that state-space similarity transformations and invertible coordinate changes in the input and output spaces, respectively, do not change the aforementioned properties. Thus it suffices to prove the result for output injection transformations. But the invariance of the notions under output injection $(A,B,C,K) \mapsto (A+LC,B,C,K)$ is obvious from Proposition 7.6. ■

It is convenient, for our further analysis, to first transform the system into a simple normal form by state-space similarity. This is done next and depends on an observability condition for the pair $\left(\begin{pmatrix} C \\ K \end{pmatrix}, A\right)$. A priori, there is no reason to assume that the pair $\left(\begin{pmatrix} C \\ K \end{pmatrix}, A\right)$ is observable, but the following proposition shows that this entails no great loss of generality.

Proposition 7.8. *1. Every linear system (7.8) can be reduced to the case that the pair $\left(\begin{pmatrix} C \\ K \end{pmatrix}, A\right)$ is observable.*

2. If $(\begin{pmatrix} C \\ K \end{pmatrix}, A)$ is observable but (C,A) is not, then the system Σ' has a state-space equivalent representation of the form

$$
\begin{aligned}
A &= \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix}, \\
C &= \begin{pmatrix} C_1 & 0 \end{pmatrix}, \\
K &= \begin{pmatrix} K_1 & K_2 \end{pmatrix},
\end{aligned}
\tag{7.9}
$$

with both pairs (C_1, A_{11}) and (K_2, A_{22}) observable.

Proof. 1. If the pair $\left(\begin{pmatrix} C \\ K \end{pmatrix}, A\right)$ is not observable, the system can be reduced to an observable one. Letting $\mathscr{V} = \bigcap_{j\geq 0} \operatorname{Ker} \begin{pmatrix} C \\ K \end{pmatrix} A^j$ be the unobservable subspace for the pair $(\begin{pmatrix} C \\ K \end{pmatrix}, A)$ and $\mathscr{W}$ be a complementary subspace leads to the direct sum decomposition of the state space into $\mathscr{X} = \mathscr{W} \oplus \mathscr{V}$. Writing $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, with $x_1 \in \mathscr{W}$ and $x_2 \in \mathscr{V}$, implies the following block matrix representations:

$$
A = \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix}, \qquad \begin{pmatrix} C \\ K \end{pmatrix} = \begin{pmatrix} C_1 & 0 \\ K_1 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}.
$$

Necessarily, $(\begin{pmatrix} C_1 \\ K_1 \end{pmatrix}, A_{11})$ is an observable pair and (7.7) can be replaced by

$$\begin{aligned} x_{t+1} &= A_{11}x_t + B_1u_t, \\ y_t &= C_1x_t, \\ z_t &= K_1x_t, \end{aligned}$$

since C_2 and K_2 are both zero, and hence x_2 plays no role.

2. If (C, A) is not an observable pair, then let $\mathscr{O}_* = \mathscr{O}_*(C, A) = \bigcap_{j \geq 0} \operatorname{Ker} CA^j \subset \mathbb{F}^n$ be the unobservable subspace of (C, A). Let $\mathscr{W}$ be a complementary subspace to $\mathscr{O}_*$. With respect to the direct sum decomposition

$$\mathbb{F}^n = \mathscr{W} \oplus \mathscr{O}_*, \tag{7.10}$$

we obtain the block matrix representation (7.9). By construction, the pair (C_1, A_{11}) is observable. Also, A_{22} is similar to $A|\mathscr{O}_*$. Our assumption that the pair $\left(\begin{pmatrix} C \\ K \end{pmatrix}, A\right)$ is observable implies that the pair (K_2, A_{22}) is also observable.

∎

Coprime factorizations are the most effective tool in bridging the gap between frequency-domain and state-space methods. This is done using the shift realization. The following results, split into two separate theorems, examine the corresponding functional characterizations. Consider a state-space system

$$\Sigma_{sys} := \begin{cases} x_{t+1} = Ax_t + Bu_t, \\ y_t = Cx_t, \\ z_t = Kx_t, \end{cases}$$

with $A \in \mathbb{F}^{n \times n}, B \in \mathbb{F}^{n \times m}, C \in \mathbb{F}^{p \times n}, K \in \mathbb{F}^{k \times n}$.

Theorem 7.9. *Assume that* $\left(\begin{pmatrix} C \\ K \end{pmatrix}, A\right)$ *is observable and has the representation (7.9).*

1. *There exists a left coprime factorization of the state-to-output transfer function of* Σ_{sys} *of the form*

$$\begin{pmatrix} C_1 & 0 \\ K_1 & K_2 \end{pmatrix} \begin{pmatrix} zI - A_{11} & 0 \\ -A_{21} & zI - A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} D_{11}(z) & 0 \\ D_{21}(z) & D_{22}(z) \end{pmatrix}^{-1} \begin{pmatrix} \Theta_{11}(z) & 0 \\ \Theta_{21}(z) & \Theta_{22}(z) \end{pmatrix}, \tag{7.11}$$

with $D_{11}(z) \in \mathbb{F}[z]^{p \times p}$ *and* $D_{22}(z) \in \mathbb{F}[z]^{k \times k}$ *nonsingular,* $D_{21}(z) \in \mathbb{F}[z]^{k \times p}, \Theta_{11}(z) \in \mathbb{F}[z]^{p \times (n-r)}, \Theta_{21}(z) \in \mathbb{F}[z]^{k \times (n-r)}, \Theta_{22}(z) \in \mathbb{F}[z]^{k \times r}$, *for which:*

(a) $D_{11}(z)^{-1}\Theta_{11}(z)$ *and* $D_{22}(z)^{-1}\Theta_{22}(z)$ *are left coprime factorizations of the transfer functions* $C_1(zI-A_{11})^{-1}$ *and* $K_2(zI-A_{22})^{-1}$*, respectively.* $D_{21}(z)$ *and* $\Theta_{21}(z)$ *satisfy the polynomial equation*

$$D_{22}(z)K_1+\Theta_{22}(z)A_{21}=-D_{21}(z)C_1+\Theta_{21}(z)(zI-A_{11}); \tag{7.12}$$

(b) $D_{11}(z)$ *and* $D_{22}(z)$ *are row proper;*
(c) $D_{21}(z)D_{11}(z)^{-1}$ *and* $D_{22}(z)^{-1}D_{21}(z)D_{11}(z)^{-1}$ *are strictly proper.*

One refers to the coprime factorization (7.11), satisfying parts (a)–(c), as an ***adapted coprime factorization****.*

2. *Assume that (7.11) is an adapted coprime factorization. The following properties are in force:*

(a) $n=\deg\det(zI-A)=\deg\det D_{11}(z)+\deg\det D_{22}(z)$.
(b) $D_{22}(z)$ *is a nonsingular polynomial matrix.*
(c) The linear map

$$\psi:\mathscr{O}_*(C,A)\longrightarrow X_{D_{22}},\quad \psi(x)=\Theta_{22}x,$$

is bijective, satisfying

$$\psi A_{22}=S_{D_{22}}\psi. \tag{7.13}$$

This implies the isomorphism

$$S_{D_{22}}\simeq A_{22}=A|_{\mathscr{O}_*(C,A)}. \tag{7.14}$$

Proof. The first claim of statement 1 is proved first. Applying Proposition 7.8, with respect to the direct sum decomposition (7.10), yields the block matrix representation (7.9), with the pairs (C_1,A_{11}) and (K_2,A_{22}) observable. Let $D_{11}(z)^{-1}\Theta_{11}(z)$ and $D_{22}(z)^{-1}\Theta_{22}(z)$ be left coprime factorizations of $C_1(zI-A_{11})^{-1}$ and $K_2(zI-A_{22})^{-1}$, respectively. Since a left coprime factorization is unique only up to a common left unimodular factor, we will assume, without loss of generality, that $D_{11}(z)$ and $D_{22}(z)$ are both row proper. So (b) holds by construction. Thus the (11)- and (22)-terms on both sides of (7.11) are equal. Comparing the (21)-terms of both sides of equation (7.11), multiplying by $D_{22}(z)$ on the left and by $(zI-A_{11})$ on the right, we obtain (7.12). Thus $D_{21}(z),\Theta_{21}(z)$ fulfill (7.11) if and only if (7.12) is satisfied. By the observability of the pair (C_1,A_{11}), the existence of a polynomial solution $X(z),Y(z)$ of the Bezout equation $X(z)C_1+Y(z)(zI-A_{11})=I$ follows. Consequently, taking into consideration the general, polynomial, solution of the homogeneous equation, we obtain the parameterization

$$\begin{aligned} D_{21}(z) &= -(D_{22}(z)K_1+\Theta_{22}(z)A_{21})X(z)-Q(z)D_{11}(z),\\ \Theta_{21}(z) &= (D_{22}(z)K_1+\Theta_{22}(z)A_{21})Y(z)+Q(z)\Theta_{11}(z), \end{aligned}$$

where $Q(z) \in \mathbb{F}[z]^{k\times p}$. Choosing $Q(z) = -\pi_+ \left((D_{22}K_1 + \Theta_{22}A_{21})XD_{11}^{-1}\right)$ guarantees that $D_{21}(z)D_{11}(z)^{-1}$ is strictly proper. This does not change the row properness of $D_{11}(z)$ and $D_{22}(z)$. That $D_{22}(z)^{-1}D_{21}(z)D_{11}(z)^{-1}$ is strictly proper follows from the strict properness of $D_{21}(z)D_{11}(z)^{-1}$ and the fact that the nonsingular, row proper polynomial matrix $D_{22}(z)$ has a proper inverse.

Proof of 2. From the observability assumption on the pair $\left(\begin{pmatrix} C \\ K \end{pmatrix}, A\right)$ and the left coprime factorization (7.11), we conclude that

$$\begin{aligned} n &= \deg\det(zI - A) = \deg\det \begin{pmatrix} D_{11}(z) & 0 \\ D_{21}(z) & D_{22}(z) \end{pmatrix} \\ &= \deg\det D_{11}(z) + \deg\det D_{22}(z), \end{aligned}$$

which proves the first claim. The nonsingularity of $D_{22}(z)$ follows from the nonsingularity of $\begin{pmatrix} D_{11}(z) & 0 \\ D_{21}(z) & D_{22}(z) \end{pmatrix}$.
To prove the last claim, note that the coprime factorization

$$D_{22}(z)^{-1}\Theta_{22}(z) = K_2(zI - A_{22})^{-1} \tag{7.15}$$

is equivalent to the intertwining relation

$$\Theta_{22}(z)(zI - A_{22}) = D_{22}(z)K_2.$$

Applying Theorems 3.20 and 3.21 proves the intertwining relation (7.13) as well as the invertibility of the map ψ defined by (7.23). The isomorphism (7.14) follows from (7.13) and the invertibility of ψ. ■

Theorem 7.10. *Assume that (7.11) is an adapted coprime factorization.*

1. *Define strictly proper rational matrices $Z_K(z), Z_C(z)$ by*

$$\begin{aligned} Z_C(z) &= C(zI - A)^{-1}, \\ Z_K(z) &= K(zI - A)^{-1}. \end{aligned} \tag{7.16}$$

The general rational solutions of the equation

$$Z_K(z) = Z_1(z)Z_C(z) + Z_2(z) \tag{7.17}$$

are

$$\begin{aligned} Z_1(z) &= -D_{22}(z)^{-1}D_{21}(z) + W(z)D_{11}(z), \\ Z_2(z) &= \left(D_{22}(z)^{-1}\Theta_{21}(z) - W(z)\Theta_{11}(z),\ D_{22}(z)^{-1}\Theta_{22}(z)\right), \end{aligned} \tag{7.18}$$

where $W(z) \in \mathbb{F}(z)^{k\times p}$ is a rational function. Let $Z_1(z)$ be proper rational. Using an adapted left coprime factorization (7.11), there exists a proper rational $W(z)$ such that $Z_1(z)$ is given by (7.18) and $Z_2(z)$ is strictly proper.

2. (a) *$D_{22}(z)$ is a stable matrix if and only if the pair (C,A) is detectable.*
   (b) *$D_{22}(z)$ is a unimodular matrix if and only if the pair (C,A) is observable. In this case, that is, where $\mathscr{O}_*(C,A) = \{0\}$, the coprime factorization (7.11) reduces to*

$$\begin{pmatrix} C \\ K \end{pmatrix} (zI-A)^{-1} = \begin{pmatrix} D_{11}(z) & 0 \\ D_{21}(z) & I \end{pmatrix}^{-1} \begin{pmatrix} \Theta_{11}(z) \\ \Theta_{21}(z) \end{pmatrix}, \tag{7.19}$$

with $D_{11}^{-1}\Theta_{11}$ a left coprime factorization of $C(zI-A)^{-1}$ and $D_{21}(z), \Theta_{21}(z)$ determined from the equation

$$K = -D_{21}(z)C + \Theta_{21}(z)(zI-A). \tag{7.20}$$

In this case, the general solution of equation (7.17) is given by

$$\begin{aligned} Z_1(z) &= -D_{21}(z) + W(z)D_{11}(z), \\ Z_2(z) &= \Theta_{21}(z) - W(z)\Theta_{11}(z). \end{aligned} \tag{7.21}$$

3. (a) *Let $\phi : \mathbb{F}^n \longrightarrow X^{zI-A}$ be defined by $\phi(x) = (zI-A)^{-1}x$. Then ϕ is injective, with*

$$\phi(\mathscr{O}_*) = \{0\} \oplus X^{zI-A_{22}}, \tag{7.22}$$

and thus induces an isomorphism $\phi : \mathscr{O}_* \longrightarrow X^{zI-A_{22}}$.
   (b) *The map $\psi : \mathscr{O}_* \longrightarrow X_{D_{22}}$, defined by*

$$\psi(x) = \Theta_{22}(z)x, \tag{7.23}$$

is an isomorphism satisfying $\psi(Ax) = S_{D_{22}}\psi(x)$.
   (c) *The map $\Psi : \mathscr{O}_* \longrightarrow X^{D_{22}}$ defined by $x \mapsto D_{22}(z)^{-1}\Theta_{22}(z)x$ is an isomorphism satisfying $\Psi(Ax) = S^{D_{22}}\Psi(x)$, i.e., $A|\mathscr{O}_* \simeq S^{D_{22}}$.*

Proof. Proof of 1. Using (7.11) and computing

$$\begin{aligned}
Z_K(z) &= \begin{pmatrix} K_1 & K_2 \end{pmatrix} \begin{pmatrix} (zI-A_{11})^{-1} & 0 \\ (zI-A_{22})^{-1}A_{21}(zI-A_{11})^{-1} & (zI-A_{22})^{-1} \end{pmatrix} \\
&= \left(K_1(zI-A_{11})^{-1} + K_2(zI-A_{22})^{-1}A_{21}(zI-A_{11})^{-1},\ K_2(zI-A_{22})^{-1} \right) \\
&= \left(-D_{22}(z)^{-1}D_{21}(z)D_{11}(z)^{-1}\Theta_{11}(z) + D_{22}(z)^{-1}\Theta_{21}(z),\ D_{22}(z)^{-1}\Theta_{22}(z) \right) \\
&= Z_1(z)Z_C(z) + Z_2(z)
\end{aligned} \tag{7.24}$$

leads to a particular solution of (7.17), i.e.,

$$\begin{aligned}
Z_1(z) &= -D_{22}(z)^{-1}D_{21}(z), \\
Z_2(z) &= \left(D_{22}(z)^{-1}\Theta_{21}(z),\ D_{22}(z)^{-1}\Theta_{22}(z) \right).
\end{aligned}$$

To obtain the general rational solution, one needs to add to (Z_1, Z_2) the general rational solution (Y_1, Y_2) of the homogeneous equation $Y_1(z)Z_C(z) + Y_2(z) = 0$ or, equivalently, $Y_1(z)C + Y_2(z)(zI-A) = 0$. Noting that $Z_C(z) = \left(D_{11}(z)^{-1}\Theta_{11}(z)\ 0 \right)$ and writing $Y_2(z) = \left(Y_2'(z),\ Y_2''(z) \right)$ implies $Y_1(z) = W(z)D_{11}(z)$, $Y_2'(z) = -W(z)\Theta_{11}(z))$, and $Y_2''(z) = 0$, with $W(z)$ a free, rational parameter. This proves (7.18).

For the second claim, choose

$$\begin{aligned}
W(z) &= (Z_1(z) + D_{22}(z)^{-1}D_{21}(z))D_{11}(z)^{-1} \\
&= Z_1(z)D_{11}(z)^{-1} + D_{22}(z)^{-1}D_{21}(z)D_{11}(z)^{-1},
\end{aligned}$$

and note that the properness of $W(z)$ follows from the assumed properness of $Z_1(z)$ and the assumption that (7.11) is an adapted coprime factorization.

Proof of 2. The pair (C, A) is detectable if and only if $A_{22} \simeq A|\bigcap_{i=0}^{\infty} \operatorname{Ker} CA^i$ is stable. By the isomorphism (7.14), this is equivalent to the stability of $D_{22}(z)$, which proves the first claim. The pair (C, A) is observable if and only if the equality of degrees $n = \deg\det(zI-A) = \deg\det D_{11}(z)$ is satisfied. This is equivalent to $\deg\det D_{22}(z) = 0$, i.e., to $D_{22}(z)$ being unimodular. Equation (7.20) is a special case of (7.12). Similarly, the parameterization (7.21) is a special case of (7.18).

Proof of 3. Since (C, A) is assumed to be in Kalman decomposition form, $x \in \mathscr{O}_*$ if and only if $x = \operatorname{col}(0, x_2)$. Thus $\phi(\mathscr{O}_*) = \{(zI-A)^{-1}x | x \in \mathscr{O}_*\} = \{0\} \times X^{zI-A_{22}}$, which proves (7.22). The coprime factorizations $D_{22}(z)^{-1}\Theta_{22}(z) = K_2(zI-A_{22})^{-1}$ yields the intertwining relation

$$\Theta_{22}(z)(zI-A_{22}) = D_{22}(z)K_2.$$

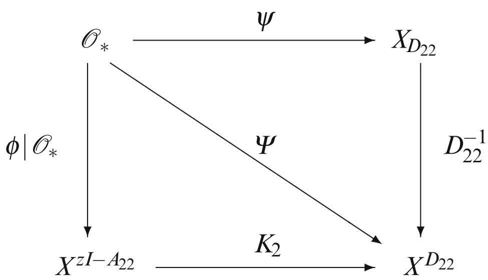
Thus the map $\psi : X_{zI-A_{22}} \longrightarrow X_{D_{22}}$ defined by $\psi(x) = \pi_{D_{22}}\Theta_{22}x = \Theta_{22}x$ is an isomorphism. Note that $\{0\} \times X_{zI-A_{22}} = \mathscr{O}_*$.

Since autonomous behaviors are equal to rational models (Theorem 3.36), we conclude that $X^{D_{22}} = \mathrm{Ker}\, D_{22}(\sigma)$ is true for the backward shift operator σ. Now the multiplication map $D_{22}^{-1} : X_{D_{22}} \longrightarrow X^{D_{22}}$ is an $\mathbb{F}[z]$-module isomorphism, and therefore the composed map $\Psi = D_{22}^{-1}\psi$ is also an $\mathbb{F}[z]$-module isomorphism from $\mathscr{O}_*$ onto $X^{D_{22}}$. This proves the last claim. ■

For unobservable states $x \in \mathscr{O}_*$, using (7.15), one computes

$$K_2\phi(x) = K_2(zI - A_{22})^{-1}x = D_{22}(z)^{-1}\Theta_{22}(z)x = \Psi x = D_{22}(z)^{-1}\psi(x),$$

which implies that the following diagram is a commutative diagram of $\mathbb{F}$-vector space isomorphisms:



It is of principal interest to find characterizations of the observation properties introduced in Definition 7.5. This depends very much on the functional relation between the observed variables y and the to-be-estimated variables z. Of course, in the state-space representation (7.8) of the system Σ', this relation is indirect. To get a direct relation, one needs to eliminate the state variable x from (7.8). This is best done in a behavioral setting but is avoided here.

Thus, avoiding the explicit use of behaviors, we proceed by characterizing the tracking index of linear systems. First let us note, as an immediate consequence of the definition, that the trackability of (7.7) with tracking index τ is satisfied if and only if τ is the smallest number such that, for every initial state $x_0 \in \mathbb{F}^n$ with $C(zI - A)^{-1}x_0 = 0$, the implication

$$Kx_0 = \cdots = Kx_\tau = 0 \implies Kx_t = 0, \quad \forall t \geq 0,$$

follows. Note that if (C, A) is observable, then the tracking index of an output functional $z = Kx$ is $\tau = 0$.

Proposition 7.11. *A linear system (7.7) with $(\begin{pmatrix} C \\ K \end{pmatrix}, A)$ observable and having the representation (7.9), has a tracking index τ if and only if the largest observability index of (K_2, A_{22}) is equal to τ. In particular, T-trackability is fulfilled for every T that is greater than or equal to the degree of the minimal polynomial of A.*

Proof. Using the representation of states with respect to the direct sum (7.10), let an initial state have the representation $x_0 = \begin{pmatrix} u \\ v \end{pmatrix}$. The initial condition x_0 is unobservable, that is, $C(zI-A)^{-1}x_0 = 0$ if and only if $u = 0$. For each such x_0 one has $KA^i x_0 = K_2 A_{22}^i v$ for all $i \geq 0$, implying the system is T-trackable if and only if the implication $K_2 A_{22}^i v = 0, \quad i = 0, \ldots, T \implies K_2 A_{22}^i v = 0 \quad \forall i \geq 0$ is valid for all vectors $v \in \mathbb{F}^{n_2}$. Equivalently, this says that

$$\operatorname{Ker} \begin{pmatrix} K_2 \\ \vdots \\ K_2 A_{22}^T \end{pmatrix} \subset \operatorname{Ker} \begin{pmatrix} K_2 \\ \vdots \\ K_2 A_{22}^{n-1} \end{pmatrix}.$$

In turn, this is equivalent to saying that all observability indices of (K_2, A_{22}) are less than or equal to T, which implies the result. ■

The following lemma will be needed.

Lemma 7.12. *Let $Q(z) \in \mathbb{F}[z]^{r\times r}$ be nonsingular with degree ℓ, i.e., $Q(z) = Q_0 + Q_1 z + \cdots + Q_\ell z^\ell$. Assuming $h(z) = \sum_{j=1}^{\infty} \frac{h_j}{z^j} \in X^Q$ and $h_1 = \cdots = h_\ell = 0$ implies $h(z) = 0$.*

Proof. Since $X^Q = \operatorname{Im} \pi^Q$, it follows that $h \in X^Q$ if and only if $h = \pi^Q h$. However, under our assumptions, $\pi^Q h = \pi_- Q^{-1} \pi_+(Qh) = \pi_- Q^{-1} \pi_+(Qz^{-\ell} z^\ell h)$. Clearly, $Q(z)z^{-\ell}$ is proper, whereas $z^\ell h$ is strictly proper, so the product is strictly proper with $\pi_+(Qh) = 0$. This implies $h(z) = 0$. ■

Lemma 7.12 leads us to a simple polynomial characterization of the tracking index.

Proposition 7.13. *Assume that (7.7) has the representation (7.9) such that the pair $(\begin{pmatrix} C \\ K \end{pmatrix}, A)$ is observable. Letting $K_2(zI - A_{22})^{-1} = D_{22}(z)^{-1}\Theta_{22}(z)$ be a left coprime factorization, with $D_{22}(z)$ row proper, implies the degree of $D_{22}(z)$ coincides with the tracking index τ of (7.7). If $D_{22}(z)$ is not row proper, the degree of $D_{22}(z)$ provides an upper bound for the tracking index.*

Proof. By Proposition 7.11, the minimal tracking index is equal to the maximal observability index of (K_2, A_{22}). In view of Corollary 6.9, the observability indices of (K_2, A_{22}) coincide with the right Wiener–Hopf indices of $D_{22}(z)$. Since $D_{22}(z)$ is assumed to be in row proper form, the row indices of $D_{22}(z)$ coincide with the right Wiener–Hopf indices. In particular, the degree of $D_{22}(z)$ is the largest right Wiener–Hopf index, i.e., it is equal to the largest observability index of (K_2, A_{22}). If $D_{22}(z)$ is not row proper, Lemma 7.12 can be applied to see that the degree of $D_{22}(z)$ gives an upper bound for the tracking index. ■

The preceding analysis leads to the following explicit characterizations of the different observation properties.

Theorem 7.14. *Assume that (7.7) has the representation (7.9) such that* $(\begin{pmatrix} C \\ K \end{pmatrix}, A)$ *is observable. Let*

$$K_2(zI - A_{22})^{-1} = D_{22}(z)^{-1}\Theta_{22}(z),$$

with $D_{22}(z), \Theta_{22}(z)$ *left coprime.*

1. *Suppose* $D_{22}(z)$ *is row proper. The following conditions are equivalent:*
   (a) *(7.7) has tracking index* τ.
   (b) *The largest observability index of* (K_2, A_{22}) *is equal to* τ.
   (c) *The degree of* $D_{22}(z)$ *is equal to* τ.
2. *The following conditions are equivalent:*
   (a) z *is detectable from* (y,u) *in (7.7).*
   (b) A_{22} *is stable.*
   (c) $\det D_{22}(z)$ *is a Schur polynomial.*
   (d) *All elements of* $X^{D_{22}}$ *have their poles in the open unit disc.*
3. *The following conditions are equivalent:*
   (a) z *is reconstructible from* (y,u) *in (7.7).*
   (b) A_{22} *is nilpotent.*
   (c) $D_{22}(z)$ *is a monomic polynomial matrix.*
   (d) *All elements of* $X^{D_{22}}$ *have their poles in zero.*
4. *The following conditions are equivalent:*
   (a) z *is observable from* (y,u) *in (7.7).*
   (b) $D_{22}(z)$ *is unimodular.*
   (c) (C,A) *is observable.*

Proof. Part 1 has already been shown.

Part 2. By Proposition 7.6, detectability is satisfied if and only if $\lim_{t\to\infty} z_t = 0$, whenever $(y_t) = 0$ and $(u_t) = 0$. The Z-transform of (z_t) is equal to $K(zI-A)^{-1}x_0$, where x_0 is in the unobservable subspace $\mathcal{O}_*$. By Theorem 7.10,

$$\{K(zI-A)^{-1}x_0 \mid x_0 \in \mathcal{O}_*\} = K_2 X^{zI-A_{22}} = X^{D_{22}}$$

is an autonomous behavior. Applying Proposition 7.4, we conclude that detectability is satisfied if and only if $\det D_{22}(z)$ is a Schur polynomial. Since $\det D_{22}(z)$ is equal to $\det(zI - A_{22})$, this is equivalent to A_{22} being stable.

For part 3 one can argue similarly. In fact, reconstructibility is equivalent to all elements of $X^{D_{22}}$ being stable for the discrete topology of $\mathbb{F}$, which just says that all elements of $X^{D_{22}}$ are of the form $z^{-N}p(z)$ for polynomials of degree $< N$ or, equivalently, that $D_{22}(z)$ is monomic. Applying Proposition 7.4, one sees that this in turn is equivalent to $\det D_{22}(z) = cz^{n_2}$ being a monomial or, equivalently, that all roots of $\det D_{22}(z) = 0$ are equal to zero, or that A_{22} is nilpotent.

Finally, the observability of z from (y,u) is fulfilled if and only if $(y_t) = 0$ and $(u_t) = 0$ implies $(z_t) = 0$. This is equivalent to

$$\{K(zI-A)^{-1}x_0 \mid x_0 \in \mathscr{O}_*\} = K_2 X^{zI-A_{22}} = \{0\}.$$

By the observability of (K_2, A_{22}), this is possible only if (C,A) is observable. This completes the proof. ■

It may be surprising to note that the characterizations of detectability, reconstructibility, and observability for a linear functional K in Theorem 7.14 are identical with the corresponding ones for $K = I_n$. This is stated as a corollary.

Corollary 7.15. *The same assumptions are used as in Theorem 7.14. For a linear system (7.7), the output $z = Kx$ is detectable, reconstructible, or observable from y if and only if the pair (C,A) is detectable, reconstructible, or observable, respectively.*

7.3 Functional State Observers

When dealing with complex systems, one may want to track only a sample of the state variables. However, these variables of interest may be impossible to observe directly, so one must have recourse to estimation procedures that utilize only the available observations. In certain cases this can be achieved even if the system is not completely observable. As indicated in the introduction to this chapter and emphasized by the observer diagram appearing there, an observer is itself a dynamical system driven by inputs and observations and whose output is an estimate ζ of the relevant variable z. This leads us to the following definition of functional state observers, which broadly extends the class of full state Luenberger observers. In the sequel, the principal results will be stated and proved only for discrete-time systems; these results hold, mutatis mutandis, also in the continuous-time case.

Let

$$\Sigma_{sys} := \begin{cases} x_{t+1} = Ax_t + Bu_t \\ y_t = Cx_t \\ z_t = Kx_t \end{cases} \tag{7.25}$$

be a linear system, with A,B,C,K in $\mathbb{F}^{n\times n}, \mathbb{F}^{n\times m}, \mathbb{F}^{p\times n}, \mathbb{F}^{k\times n}$, respectively. Let another system,

$$\Sigma_{est} := \begin{cases} \xi_{t+1} = F\xi_t + Gy_t + Hu_t, \\ \zeta_t = J\xi_t + Ey_t, \end{cases} \tag{7.26}$$

be given with system matrices F,G,H,J,E in $\mathbb{F}^{q\times q},\mathbb{F}^{q\times p},\mathbb{F}^{q\times m},\mathbb{F}^{k\times q},\mathbb{F}^{k\times p}$ respectively, and driven by the input u and output y of (7.25). It will always be assumed that J is of full row rank, which presents no restriction of generality, as well as that both C and K have full row rank and that the pair $(\begin{pmatrix} C \\ K \end{pmatrix}, A)$ is observable.

Define the **estimation error** or **error trajectory** e by

$$e_t = z_t - \zeta_t = Kx_t - J\xi_t - Ey_t = Kx_t - \begin{pmatrix} J & E \end{pmatrix} \begin{pmatrix} x_t \\ y_t \end{pmatrix}.$$

The error trajectory defines the strictly proper power series

$$e = e(z) = \sum_{t=0}^{\infty} e_t z^{-t-1}. \tag{7.27}$$

As usual, the error trajectory will often be identified with the equivalent formal power series expansion (7.27) that it defines. We will refer to (7.26) as a **functional observer** because it is designed to estimate a function of the state rather than the state itself.

Definition 7.16. Consider the linear system (7.25). The system Σ_{est} defined by (7.26) will be called

1. a **finitely determined observer** for K if there exists a $T \in \mathbb{N}$ such that $e_t = 0$ for $t < T$ implies $e = 0$;
2. a **tracking observer** for K if for every $x_0 \in \mathbb{F}^n$ there exists a $\xi_0 \in \mathbb{F}^q$ such that, for all input functions u, the solutions x_t and ξ_t of (7.25) and (7.26), respectively, satisfy $e_t = z_t - \zeta_t = 0$ for all $t \geq 0$;
3. a **strongly tracking observer** for K if $e_0 = z_0 - \zeta_0 = 0$ implies $e_t = 0$ for all input functions u and $t \geq 0$;
4. an **asymptotic observer** for K if, for all initial conditions of the states x and ξ and all inputs u, $\lim_{t\to\infty} e_t = \lim_{t\to\infty}(z_t - \zeta_t) = 0$; an observer is called an **asymptotic tracking observer** for K if it is both a tracking observer and an asymptotic observer;
5. **spectrally assignable** if, given a polynomial $p(z)$ of degree q, there exists an observer in the family for which the characteristic polynomial of F is $p(z)$.

In all cases, q will be called the **order** of the observer.

Naturally, there are two fundamental problems that present themselves, namely, given system (7.25), how does one obtain a characterization of observers and how does one show the existence of observers of the various types, together with a computational procedure for observer construction? Note further that, in general, the initial value of the state of a system is not known, which is at the core of the estimation/observation problem. Even if a tracking observer exists, there will be a nonzero tracking error whenever the initialization of the observer is incorrect. This points out the importance of asymptotic observers as well as, even more, spectrally assignable observers where one also controls the rate of convergence to zero of the error. Several further remarks are in order.

1. Incorporating a feedthrough term in the observer, as in (7.26), improves our ability to construct reduced-order functional observers. An example of this is the case of observing Kx, where $\operatorname{Ker} C \subset \operatorname{Ker} K$. This implies the existence of a map E for which $K = EC$, which leads to a zero-order, i.e., nondynamic, observer given by $\zeta_t = Ey_t$. Note that incorporating a feedthrough term in the observer is not new; it already appeared in Luenberger (1971) in the construction of an observer for a single functional of the state.
2. The definition of a tracking observer clearly implies that the set of the trajectories to be estimated is included in the set of outputs of the tracking observer.
3. A strongly tracking observer is at the same time a tracking observer. This follows from our assumption that J has full row rank. Thus $e_0 = Kx_0 - J\xi_0$ can always be made zero by an appropriate choice of ξ_0. Note also that a strongly tracking observer is finitely determined, with $T = 1$.

It was already observed that the trackability of an output of a finite-dimensional linear system is always satisfied. In the same vein, it is always possible to construct tracking observers by inspection. For instance, taking the copy of a system as

$$\begin{aligned} \xi_{t+1} &= A\xi_t + Bu_t, \\ \zeta_t &= K\xi_t \end{aligned}$$

obviously leads to a tracking observer of system (7.25). Note that this observer has the same dimension n as (7.25). More generally, for a matrix $L \in \mathbb{F}^{n\times p}$, the system

$$\begin{aligned} \xi_{t+1} &= (A - LC)\xi_t + LCx_t + Bu_t, \\ \zeta_t &= K\xi_t \end{aligned}$$

is a tracking observer for (7.25). Therefore, the main issue is not the existence of tracking observers (they always exist), but whether or not tracking observers with prescribed dimension $q \leq n$ exist or, even better, whether a minimal-order observer can be constructed. This problem will be addressed in Theorem 7.27, but first, a characterization of functional observers is derived.

Our starting point is the derivation of a state-space characterization, in terms of matrix Sylvester equations, for the classes of observers introduced in Definition 7.16. Thus we consider a linear system Σ_{sys}:

$$\Sigma_{sys} := \begin{cases} x_{t+1} = Ax_t + Bu_t, \\ y_t = Cx_t, \\ z_t = Kx_t, \end{cases} \tag{7.28}$$

with state space $\mathbb{F}^n$ and the estimator system

$$\Sigma_{est} := \begin{cases} \xi_{t+1} = F\xi_t + Gy_t + Hu_t, \\ \zeta_t = J\xi_t + Ey_t, \end{cases} \tag{7.29}$$

in the state space $\mathbb{F}^q$.

Theorem 7.17. *Assume that both* $(\begin{pmatrix} C \\ K \end{pmatrix}, A)$ *and* (J,F) *are observable.*

1. *System (7.29) is a tracking observer for K if and only if there exists a solution* $Z \in \mathbb{F}^{q\times n}$ *of the* ***observer Sylvester equations***

$$\begin{aligned} ZA &= \begin{pmatrix} F & G \end{pmatrix}\begin{pmatrix} Z \\ C \end{pmatrix}, \\ H &= ZB, \\ K &= \begin{pmatrix} J & E \end{pmatrix}\begin{pmatrix} Z \\ C \end{pmatrix}. \end{aligned} \tag{7.30}$$

   The solution Z of (7.30) is uniquely determined.

2. *Let Z be the unique solution to the observer Sylvester equations (7.30). Defining* $Z_K(z) = K(zI-A)^{-1}$ *and* $Z_C(z) = C(zI-A)^{-1}$*, the equation*

$$Z_K(z) = Z_1(z)Z_C(z) + Z_2(z) \tag{7.31}$$

   is solvable with

$$Z_1(z) = E + J(zI-F)^{-1}G, \quad Z_2(z) = J(zI-F)^{-1}Z. \tag{7.32}$$

3. *Defining an auxiliary variable* ε *by*

$$\varepsilon = Zx - \xi, \tag{7.33}$$

   the observer error dynamics with the initial condition $\varepsilon_0 = Zx_0 - \xi_0$ *are*

$$\begin{aligned} \varepsilon_{t+1} &= F\varepsilon_t, \\ e_t &= J\varepsilon_t, \end{aligned} \tag{7.34}$$

   i.e., the error trajectory is the output of an autonomous linear system. The set $\mathscr{B}_{err}$ *of all error trajectories is an autonomous behavior of the form*

$$\mathscr{B}_{err} = X^Q, \tag{7.35}$$

where

$$Q(z)^{-1}P(z) = J(zI-F)^{-1} \tag{7.36}$$

are coprime factorizations.

4. *The following conditions are equivalent:*
   (a) *System (7.29) is an asymptotic tracking observer for K.*
   (b) *There exists a linear transformations Z, with F stable, such that (7.30) holds.*
5. *The following conditions are equivalent:*
   (a) *System (7.29) represents a family of spectrally assignable tracking observers for K.*
   (b) *With the characteristic polynomial of F preassigned, there exists a linear transformation Z that satisfies the observer Sylvester equations (7.30).*

Proof. 1.

For initial conditions x_0 for Σ_{sys} and ξ_0 for Σ_{est}, the Z-transforms of the solutions to equations (7.28), (7.29), and (7.34) are given by

$$\begin{aligned}
x &= (zI-A)^{-1}x_0 + (zI-A)^{-1}Bu,\\
y &= C(zI-A)^{-1}x_0 + C(zI-A)^{-1}Bu,\\
z &= K(zI-A)^{-1}x_0 + K(zI-A)^{-1}Bu,\\
\xi &= (zI-F)^{-1}\xi_0 + (zI-F)^{-1}GC(zI-A)^{-1}x_0 + (zI-F)^{-1}(H+GC(zI-A)^{-1}B)u,\\
\zeta &= J(zI-F)^{-1}\xi_0 + J(zI-F)^{-1}GC(zI-A)^{-1}x_0 + EC(zI-A)^{-1}x_0\\
&\quad + J(zI-F)^{-1}(H+GC(zI-A)^{-1}B)u + EC(zI-A)^{-1}x_0 + EC(zI-A)^{-1}Bu,\\
e &= z-\zeta = \left[K - J(zI-F)^{-1}GC - EC\right](zI-A)^{-1}x_0 - J(zI-F)^{-1}\xi_0,\\
&\quad + \left[(K-EC)(zI-A)^{-1}B - J(zI-F)^{-1}(H+GC(zI-A)^{-1}B)\right]u.
\end{aligned} \tag{7.37}$$

To begin, one takes $u = 0$. The trackability assumption translates into the following statement. For each vector $x_0 \in \mathbb{F}^n$, there exists a vector $\xi_0 \in \mathbb{F}^q$ such that

$$J(zI-F)^{-1}\xi_0 + J(zI-F)^{-1}GC(zI-A)^{-1}x_0 + EC(zI-A)^{-1}x_0 - K(zI-A)^{-1}x_0 = 0.$$

This implies that ξ_0 is a linear function of x_0. Because x_0 is unrestricted, this means that there exists a $Z \in \mathbb{F}^{q\times n}$ for which $\xi_0 = Zx_0$. This leads to the identity

$$J(zI-F)^{-1}Z + J(zI-F)^{-1}GC(zI-A)^{-1} + EC(zI-A)^{-1} - K(zI-A)^{-1} = 0.$$

Equating residues, one obtains

$$K = JZ + EC = \begin{pmatrix} J & E \end{pmatrix} \begin{pmatrix} Z \\ C \end{pmatrix}. \tag{7.38}$$

Using this identity one computes

$$\begin{aligned}
0 &= J(zI-F)^{-1}Z + J(zI-F)^{-1}GC(zI-A)^{-1} + EC(zI-A)^{-1} - K(zI-A)^{-1} \\
&= J(zI-F)^{-1}Z + J(zI-F)^{-1}GC(zI-A)^{-1} + (K-JZ)(zI-A)^{-1} - K(zI-A)^{-1} \\
&= J(zI-F)^{-1}[Z(zI-A) + GC - (zI-F)Z](zI-A)^{-1} \\
&= J(zI-F)^{-1}[-ZA + GC + FZ](zI-A)^{-1}.
\end{aligned}$$

The nonsingularity of $(zI-A)$ and the observability of the pair (J,F) imply the identity

$$ZA - FZ = GC, \tag{7.39}$$

which can be rewritten as

$$ZA = \begin{pmatrix} F & G \end{pmatrix} \begin{pmatrix} Z \\ C \end{pmatrix}.$$

By inserting identities (7.38) and (7.39), together with $\xi_0 = Zx_0$, back into the representation of e in (7.37), one gets

$$\begin{aligned}
0 &= J\left[Z - (zI-F)^{-1}GC\right](zI-A)^{-1}Bu - J(zI-F)^{-1}Hu \\
&= J(zI-F)^{-1}\left[(zI-F)Z - GC\right](zI-A)^{-1}Bu - J(zI-F)^{-1}Hu \\
&= J(zI-F)^{-1}\left[Z(zI-A)\right](zI-A)^{-1}Bu - J(zI-F)^{-1}Hu \\
&= J(zI-F)^{-1}(ZB-H)u.
\end{aligned}$$

Choosing constant inputs and using the observability of (J,F), this implies $H = ZB$. Thus the observer Sylvester equations (7.30) hold.

To show the uniqueness of the solution to the observer Sylvester equations (7.30), assume there exist two maps Z', Z'' satisfying them. Setting $Z = Z'' - Z'$ yields $ZA = FZ$ and $JZ = 0$. The intertwining relation implies that $ZA^k = F^kZ$, for all $k \geq 0$, and hence $JF^kZ = JZA^k = 0$, i.e., $\operatorname{Im} Z \subset \bigcap_{k\geq 0} \operatorname{Ker} JF^k$. The observability of the pair (J,F) implies now $Z = 0$, i.e., $Z'' = Z'$.

Conversely, assume the observer Sylvester equations (7.30) are satisfied. For a control u and an initial condition x_0 for Σ', we choose $\xi_0 = Zx_0$. By (7.37), using the Sylvester equations, the error trajectory is given by

$$\begin{aligned}
e &= \left[K - J(zI-F)^{-1}GC\right](zI-A)^{-1}x_0 - J(zI-F)^{-1}Zx_0 \\
&\quad + \left[JZ(zI-A)^{-1}B - J(zI-F)^{-1}(H + GC(zI-A)^{-1}B)\right]u \\
&= J(zI-F)^{-1}\left[(zI-F)Z - Z(zI-A) - GC\right](zI-A)^{-1}x_0 \\
&\quad + J(zI-F)^{-1}\left[(zI-F)Z - Z(zI-A) - GC\right](zI-A)^{-1}Bu(z) = 0.
\end{aligned}$$

This shows that Σ_{est} is a tracking observer for Σ_{sys}.

2. From equation (7.39) it follows that

$$Z(zI-A) - (zI-F)Z = -GC,$$

and hence

$$(zI-F)^{-1}Z - Z(zI-A)^{-1} = -(zI-F)^{-1}GC(zI-A)^{-1}.$$

Using (7.38), this leads to

$$J(zI-F)^{-1}Z + J(zI-F)^{-1}GC(zI-A)^{-1} = JZ(zI-A)^{-1} = (K-EC)(zI-A)^{-1},$$

which proves the statement.

3. To determine the error dynamics, one computes, using the observer Sylvester equations (7.30),

$$\begin{aligned}
\varepsilon_{t+1} &= Zx_{t+1} - \xi_{t+1} \\
&= ZAx_t + ZBu_t - F\xi_t - GCx_t - Hu_t \\
&= ZAx_t + ZBu_t - [F(Zx_t - \varepsilon_t) + GCx_t - ZBu_t] \\
&= F\varepsilon_t, \\
e_t &= Kx_t - J\xi_t = J\varepsilon_t.
\end{aligned}$$

This proves (7.34). The error behavior, i.e., the space of error trajectories, is given by $\mathscr{B}_{err} = \{J(zI-F)^{-1}\xi \mid \xi \in \mathbb{F}^q\}$. Applying Proposition 4.36 to the coprime factorizations (7.36) leads to the representation $\{J(zI-F)^{-1}\xi \mid \xi \in \mathbb{F}^q\} = X^Q$.

4. $(a) \Leftrightarrow (b)$.

Assume that (7.29) is an asymptotic tracking observer for (7.28). By part 1, there exists a uniquely determined linear transformation Z that satisfy the Sylvester equations (7.30). Since the error dynamics is given by (7.34), and (J,F) is an observable pair by assumption, the convergence $e_t \to 0$ always implies $\varepsilon_t \to 0$. Thus the error dynamics (7.34) are stable, which shows that F is stable.

Conversely, the existence of a map Z solving the Sylvester equations (7.30) implies Σ_{est} is a tracking observer. The assumed stability of F implies, using the error dynamics (7.34), that Σ_{est} is actually an asymptotic tracking observer.

5. $(a) \Leftrightarrow (b)$.
   Follows directly from the definition of a spectrally assignable family of tracking observers, together with part 1.

■

It is of interest to relate the error trajectories to the proper, rational solutions $Z_1(z), Z_2(z)$ of the equation $Z_K(z) = Z_1(z)Z_C(z) + Z_2(z)$. Here $Z_K(z)$ and $Z_C(z)$ are defined by (7.16) and $Z_1(z)$ and $Z_2(z)$ by (7.32). Note that $Z_1(z)$ is the transfer function of the observer from y to ζ, while $Z_2(z)$ is related to the error estimate. Choosing the initial condition of the observer as $\xi_0 = Zx_0$ would make the error trajectory zero. However, the initial state x_0 is unknown to the observer and, in the absence of that initial state information, the challenge is to obtain an error estimate. This is provided by the following proposition.

Proposition 7.18. *Let the initial conditions for system (7.25) and the observable tracking observer (7.26) be x_0 and ξ_0, respectively. The error trajectory is given by*

$$e = J(zI - F)^{-1}(Zx_0 - \xi_0) = Z_2(z)x_0 - J(zI - F)^{-1}\xi_0. \tag{7.40}$$

In particular, e does not depend on the input u.

Proof. Computing, using equation (7.37), together with the observer Sylvester equations (7.30) and (7.32),

$$\begin{aligned}
e &= \left[K - J(zI-F)^{-1}GC - EC\right](zI-A)^{-1}x_0 - J(zI-F)^{-1}\xi_0 \\
&\quad + \left[(K-EC)(zI-A)^{-1}B - J(zI-F)^{-1}(H + GC(zI-A)^{-1}B)\right]u \\
&= [Z_K(z) - Z_1(z)Z_C(z)]\,x_0 - J(zI-F)^{-1}\xi_0 \\
&\quad + [Z_K - Z_1(z)Z_C(z) - Z_2(z)]\,Bu \\
&= Z_2(z)x_0 - J(zI-F)^{-1}\xi_0
\end{aligned}$$

completes the proof. ■

How the existence of tracking observers is preserved under the action of the output injection group $\mathscr{G}$ is shown in the next proposition. This is done by showing how the observer Sylvester equations (7.30) transform under the same group action.

Proposition 7.19. *Let the linear system*

$$\Sigma_{sys} := \begin{cases} x_{t+1} = Ax_t + Bu_t, \\ y_t = Cx_t, \\ z_t = Kx_t \end{cases} \tag{7.41}$$

act in the state space $\mathbb{F}^n$*. It is assumed that* C *and* K *have full row rank and that* $\left(\begin{pmatrix} C \\ K \end{pmatrix}, A\right)$ *is observable. Assume that*

$$\Sigma_{est} := \begin{cases} \xi_{t+1} = F\xi_t + Gy_t + Hu_t, \\ \zeta_t = J\xi_t + Ey_t \end{cases}$$

is an observable tracking observer with state space $\mathbb{F}^q$ *that satisfies the observer Sylvester equations*

$$\begin{aligned} ZA &= \begin{pmatrix} F & G \end{pmatrix} \begin{pmatrix} Z \\ C \end{pmatrix}, \\ H &= ZB, \\ K &= \begin{pmatrix} J & E \end{pmatrix} \begin{pmatrix} Z \\ C \end{pmatrix}. \end{aligned} \tag{7.42}$$

Equations (6.3) are extended to the action of the output injection group $\mathscr{G}$ *on quadruples* A,B,C,K *by*

$$\begin{pmatrix} \overline{A} & \overline{B} \\ \overline{C} & 0 \\ \overline{K} & 0 \end{pmatrix} = \begin{pmatrix} R & L & 0 \\ 0 & S & 0 \\ 0 & 0 & I \end{pmatrix} \begin{pmatrix} A & B \\ C & 0 \\ K & 0 \end{pmatrix} \begin{pmatrix} R & 0 \\ 0 & I \end{pmatrix}^{-1}.$$

Under this action, the observer Sylvester equations transform as follows:

$$\begin{aligned} \overline{Z}\overline{A} &= \begin{pmatrix} \overline{F} & \overline{G} \end{pmatrix} \begin{pmatrix} \overline{Z} \\ \overline{C} \end{pmatrix}, \\ \overline{H} &= \overline{Z}\overline{B}, \\ \overline{K} &= \begin{pmatrix} \overline{J} & \overline{E} \end{pmatrix} \begin{pmatrix} \overline{Z} \\ \overline{C} \end{pmatrix}, \end{aligned} \tag{7.43}$$

where

$$\overline{A} = (RA - LC)R^{-1}, \quad \overline{B} = RB, \quad \overline{C} = SCR^{-1}, \quad \overline{K} = KR^{-1},$$

$$\overline{Z} = ZR^{-1}, \quad \overline{L} = LS^{-1},$$

$$\overline{F} = F, \quad \overline{G} = G - \overline{Z}\overline{L}, \quad \overline{H} = H, \quad \overline{J} = J, \quad \overline{E} = ES^{-1}.$$

Proof. Equation (7.43) can be rewritten as

$$\begin{aligned} (ZR^{-1})(R(A - LS^{-1}C)R^{-1}) &= \begin{pmatrix} F & (G - ZLS^{-1}) \end{pmatrix} \begin{pmatrix} ZR^{-1} \\ CR^{-1} \end{pmatrix}, \\ H &= (ZR^{-1})(RB), \\ KR^{-1} &= \begin{pmatrix} J & E \end{pmatrix} \begin{pmatrix} ZR^{-1} \\ CR^{-1} \end{pmatrix}. \end{aligned}$$

This in turn is equivalent to the Sylvester equations (7.30).

Corollary 7.20. *Consider system (7.28). The observer (7.29) is finitely determined if and only if it is tracking.*

Proof. Follows from the error dynamics given by (7.35) and Lemma 7.12. ∎

The matrix Z that solves the observer Sylvester equations is not necessarily of full rank. Next, it is shown how the maps F,G,H,J,E can always be modified in such a way that a full rank solution $\overline{Z}$ to the Sylvester equations of the modified observer exists.

Proposition 7.21. *Suppose there exists a q-dimensional observable tracking observer for K. Then there exists a tracking observer $\overline{F},\overline{G},\overline{H},\overline{J},\overline{E}$ of dimension $\overline{q} \le q$ together with a full row rank solution $\overline{Z}$ to the observer Sylvester equations*

$$\begin{aligned} \overline{Z}A - \overline{FZ} &= \overline{G}C, \\ \overline{H} &= \overline{Z}B, \\ K &= \overline{JZ} + \overline{E}C. \end{aligned} \tag{7.44}$$

Furthermore, the pair $(\overline{J},\overline{F})$ can be chosen to be observable.

Proof. Let F,G,H,J,E denote an observable tracking observer for system (7.25), with Z the solution to the observer Sylvester equation (7.30). If Z is not surjective, this implies that, in an appropriate basis, one has $Z = \begin{pmatrix} \overline{Z} \\ 0 \end{pmatrix}$, with $\overline{Z}$ surjective, i.e., of full row rank. The corresponding representations are as follows:

$$F = \begin{pmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{pmatrix}, \quad G = \begin{pmatrix} G_1 \\ G_2 \end{pmatrix}, \quad H = \begin{pmatrix} H_1 \\ H_2 \end{pmatrix}, \quad J = \begin{pmatrix} J_1 & J_2 \end{pmatrix}.$$

Equations (7.44) can now be rewritten as

$$\begin{pmatrix} \overline{Z} \\ 0 \end{pmatrix} A = \begin{pmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{pmatrix} \begin{pmatrix} \overline{Z} \\ 0 \end{pmatrix} + \begin{pmatrix} G_1 \\ G_2 \end{pmatrix} C,$$

$$\begin{pmatrix} H_1 \\ H_2 \end{pmatrix} = \begin{pmatrix} \overline{Z} \\ 0 \end{pmatrix} B,$$

$$K = \begin{pmatrix} J_1 & J_2 \end{pmatrix} \begin{pmatrix} \overline{Z} \\ 0 \end{pmatrix} + EC.$$

This in turn implies

$$\overline{Z}A - F_{11}\overline{Z} = G_1 C,$$

$$H_1 = \overline{Z}B,$$

$$K = J_1\overline{Z} + EC.$$

The observability of the pair (J,F) implies the observability of the pair (J_1,F_{11}). Therefore,

$$\begin{aligned}\xi_{t+1} &= F_{11}\xi_t + G_1 y_t + H_1 u_t,\\ \zeta_t &= J_1\xi_t + EC\end{aligned}$$

is a tracking observer for (7.28). ∎

Before proceeding to establish the basic connections between observation properties and observer constructions, the following simple lemma is stated.

Lemma 7.22. *Let (F,G,H,J,E) be a tracking observer for (A,B,C,K).*

1. *If Z is a solution to the observer Sylvester equations, then, for each output injection map L, $(F,G-ZL,H,J,E)$ is a tracking observer for $(A-LC,B,C,K)$.*
2. *If Z is a solution to the observer Sylvester equations and P is nonsingular, then PZ solves the Sylvester equations for the tracking observer defined by $(PFP^{-1},PG,PH,JP^{-1},E)$.*

Proof. 1. Theorem 7.17 is applied to conclude that there exists a solution Z to the Sylvester equations (7.30). The first equation, $ZA = FZ + GC$, implies $Z(A - LC) = FZ + (G - ZL)C$, while the other two equations remain untouched. This implies that the observer defined by $(F,G-ZL,H,J,E)$ is a tracking observer for the system defined by $(A-LC,B,H,C,K)$.

2. Using the observer Sylvester equations (7.30), one computes

$$\begin{aligned}(PZ)A &= (PFP^{-1})(PZ) + (PG)C,\\ K &= (JP^{-1})(PZ) + EC.\end{aligned}$$

The identity of transfer functions

$$\left[\begin{array}{c|cc} PFP^{-1} & PG & PZ \\ \hline JP^{-1} & E & 0 \end{array}\right] = \left[\begin{array}{c|cc} F & G & Z \\ \hline J & E & 0 \end{array}\right]$$

is easily checked.

∎

An interesting question is to analyze the extent of our control over the error dynamics. In particular, one might want to clarify the following question: under what conditions can we preassign the error dynamics? In one direction, this is easily resolved using observability subspaces. This is studied in Fuhrmann and Trumpf (2006). For further studies, one would first have to extend Definition 7.16 to the notion of spectral assignability and derive results characterizing the existence of such observers.

7.4 Existence of Observers

Having studied the observation properties of linear systems in Section 7.2 and introduced several classes of observers in Section 7.3, it will come as no great surprise that there is a natural correspondence between observation properties and observers of linear systems. This correspondence is addressed next.

Theorem 7.23. *Let the linear system*

$$\Sigma_{sys} := \begin{cases} x_{t+1} = Ax_t + Bu_t, \\ y_t = Cx_t, \\ z_t = Kx_t \end{cases} \tag{7.45}$$

act in the state space $\mathbb{F}^n$. *It is assumed that* $(\begin{pmatrix} C \\ K \end{pmatrix}, A)$ *is observable and is of the form (7.9).*

1. *Let* Σ_{sys} *be trackable with minimal tracking index* τ. *Then the maximal observability index of every tracking observer for* K *is greater than or equal to* τ.
2. *The following conditions are equivalent:*
   (a) *There exists an asymptotic tracking observer for* K.
   (b) *The pair* (C,A) *is detectable.*
3. *The following conditions are equivalent:*
   (a) *There exists a spectrally assignable family of tracking observers for* K.
   (b) *The pair* (C,A) *is observable.*

Proof. Part 1.

Let Σ_{est} be a tracking observer (7.26) with maximal observability index equal to τ_*. Choose an initial condition $x_0 \in \mathbb{F}^n$ with $CA^t x_0 = 0$ for all t. Assume $Kx_0 = \cdots = Kx_{\tau_*} = 0$. Since (7.26) is a tracking observer, there exists ξ_0 such that $Kx_t = J\xi_t$ is true for all t. Since $y_t = CA^t x_0 = 0$ for all t, one concludes $J\xi_t = JF^t\xi_0$ for all t. In particular, $J\xi_0 = \cdots = JF^{\tau_*-1}\xi_0 = 0$. By the observability of J and F, we have that $\xi_0 = 0$, and thus $Kx_t = 0$ for all t. This implies the bound $\tau \le \tau_*$ for the minimal tracking index.

Part 2. $(a) \Leftrightarrow (b)$

Assume that an asymptotic tracking observer

$$\begin{aligned} \xi_{t+1} &= F\xi_t + Gy_t + Hu_t, \\ \zeta_t &= J\xi_t + Ey_t \end{aligned} \tag{7.46}$$

for z exists; then F is necessarily stable. One must show that all unobservable modes of (C,A) are stable. Choose $u = 0$, and pick an unobservable initial state $x_0 \in \mathscr{O}_*$. Then $\zeta_t = JF^t\xi_0$ is true for all t and all ξ_0. The stability of F implies

$\lim_{t\to\infty} JF^t\xi_0 = 0$. Since (7.46) is a tracking observer, there exists ξ_0 such that $\zeta_t = z_t$ for all $t \geq 0$. Thus $\lim_{t\to\infty} KA^t x_0 = \lim_{t\to\infty} z_t = 0$ for all $x_0 \in \mathscr{O}_*$. This is equivalent to $\lim_{t\to\infty} K_2 A_{22}^t = 0$ and, therefore, by the observability of (K_2, A_{22}), to the stability of A_{22}. Thus (C,A) is detectable.

Conversely, assume (C,A) is detectable. By Theorem 7.14, the unobservable modes of A are stable and there exists an output injection matrix L such that $A - LC$ is stable. Therefore, the full state Luenberger observer

$$\begin{aligned} \xi_{t+1} &= (A - LC)\xi_t + Ly_t + Bu_t, \\ \zeta_t &= K\xi_t \end{aligned} \tag{7.47}$$

is an asymptotic tracking observer for K.

Part 3. $(a) \Leftrightarrow (b)$

If (C,A) is observable, then one can find L such that $A - LC$ has a preassigned characteristic polynomial. Thus the Luenberger observer (7.47) yields a spectrally assignable tracking observer for K. This shows that $(b) \Rightarrow (a)$. For the converse, assume (7.46) is a spectrally assignable tracking observer. Thus F can be chosen with a preassigned characteristic polynomial. To show that (C,A) is observable, suppose, to obtain a contradiction, that there exists a nonzero unobservable state $x_0 \in \mathscr{O}_*$. Choose $u = 0$ and ξ_0 such that $z_t = \zeta_t$ for all t. Then, for each v, there exists ξ_0 with $z_t = K_2 A_{22}^t v = JF^t\xi_0$, or, equivalently, there exists a matrix Z that satisfies

$$K_2 A_{22}^t = JF^t Z$$

for all $t \geq 0$. Equivalently,

$$K_2(zI - A_{22})^{-1} = J(zI - F)^{-1}Z. \tag{7.48}$$

Since (7.46) is spectrally assignable, one can choose F so that the minimal polynomial $b(z)$ of F is coprime to the minimal polynomial $a(z)$ of A_{22}. Then the poles of the rational functions on both sides of (7.48) are disjoint. Therefore, both sides must be zero, i.e.,

$$K_2(zI - A_{22})^{-1} = 0 = J(zI - F)^{-1}Z,$$

in contradiction to the observability of (K_2, A_{22}). This completes the proof. ∎

Our next objective is the characterization of several classes of observers

$$\Sigma_{est} = \begin{cases} \xi_{t+1} &= F\xi_t + Gy_t + Hu_t, \\ \zeta_t &= J\xi_t + Ey_t, \end{cases} \tag{7.49}$$

defined in the state space $\mathbb{F}^q$. This includes the characterization of minimal-order observers as a special case, relating, for example, minimal-order q tracking observers (7.49) with maximal conditioned invariant subspaces of codimension q. It is tacitly assumed throughout the subsequent arguments that (7.49) is observable. Our analysis is done in essentially two ways, geometrically and functionally. For the case of detectability, it will always be assumed that the underlying field is a subfield of $\mathbb{C}$. Our results remain in force for all fields $\mathbb{F}$ endowed with the discrete topology. In that case, detectability is equivalent to reconstructibility.

Theorem 7.24. *Let the linear system*

$$\Sigma_{sys} := \begin{cases} x_{t+1} = Ax_t + Bu_t, \\ y_t = Cx_t, \\ z_t = Kx_t \end{cases} \tag{7.50}$$

act in the state space $\mathbb{F}^n$. It is assumed that the pair $\left(\begin{pmatrix} C \\ K \end{pmatrix}, A\right)$ is observable and C and K are both of full row rank. Let $Z_C(z), Z_K(z)$ be defined by (7.16).

1. *The following conditions are equivalent:*
   (a) *There exists a tracking observer for K of order q.*
   (b) *There exists a conditioned invariant subspace $\mathscr{V} \subset \mathbb{F}^n$, of codimension q, satisfying*
   $$\mathscr{V} \cap \operatorname{Ker} C \subset \operatorname{Ker} K. \tag{7.51}$$
   (c) *There exist proper rational functions $Z_1(z), Z_2(z)$, with McMillan degree of $\begin{pmatrix} Z_1(z) & Z_2(z) \end{pmatrix}$ less than or equal to q, that solve*
   $$\begin{pmatrix} Z_1(z) & Z_2(z) \end{pmatrix} \begin{pmatrix} C \\ zI - A \end{pmatrix} = K \tag{7.52}$$
   or the equivalent equation
   $$Z_K(z) = Z_1(z)Z_C(z) + Z_2(z). \tag{7.53}$$
2. *The following statements are equivalent:*
   (a) *There exists an order q asymptotic tracking observer for K.*
   (b) *There exists an outer detectable subspace, with $\operatorname{codim} \mathscr{D} = q$, satisfying*
   $$\mathscr{D} \cap \operatorname{Ker} C \subset \operatorname{Ker} K. \tag{7.54}$$
   (c) *There exist strictly proper, stable rational functions $Z_1(z), Z_2(z)$, with McMillan degree of $\begin{pmatrix} Z_1(z) & Z_2(z) \end{pmatrix}$ equal to q, that solve (7.53).*

3. *The following conditions are equivalent:*
   (a) *There exists an order-q, spectrally assignable family of tracking observers for K.*
   (b) *There exists an outer observability subspace $\mathscr{O}$ of codimension q satisfying*

$$\mathscr{O} \cap \operatorname{Ker} C \subset \operatorname{Ker} K. \tag{7.55}$$

   (c) *There exist polynomial matrices $P_1(z)$ and $P_2(z)$ that solve (7.53).*

Proof. Part 1. $(a) \Leftrightarrow (b)$.

Assume that a q-dimensional tracking observer exists and is given by (7.49). By Theorem 7.17, there exists a solution to the observer Sylvester equations (7.30). By Proposition 7.21, it can be assumed without loss of generality that Z is of full row rank. For minimal tracking observers this is automatically satisfied. Otherwise, the observer order could be further reduced, contradicting the assumption that the observer has minimal order. Define now $\mathscr{V} = \operatorname{Ker} Z$. The equation

$$ZA = \begin{pmatrix} F & G \end{pmatrix} \begin{pmatrix} Z \\ C \end{pmatrix}$$

implies

$$A \operatorname{Ker} \begin{pmatrix} Z \\ C \end{pmatrix} \subset \operatorname{Ker} Z$$

or, equivalently,

$$A(\mathscr{V} \cap \operatorname{Ker} C) \subset \mathscr{V},$$

which shows that $\mathscr{V}$ is a conditioned invariant subspace. By Theorem 7.17, there exist F and Z that satisfy the Sylvester equations (7.30). In particular, $ZA = FZ + GC$ shows that $\mathscr{V} = \operatorname{Ker} Z$ is conditioned invariant. Moreover, the equation $K = JZ + EC$ implies the inclusion

$$\mathscr{V} \cap \operatorname{Ker} C = \operatorname{Ker} \begin{pmatrix} Z \\ C \end{pmatrix} \subset \operatorname{Ker} K.$$

Z having row rank implies $\operatorname{rank} Z = q$. The equality $\dim \operatorname{Ker} Z = n - \operatorname{rank} Z$ implies $\operatorname{codim} \mathscr{V} = q$.

To prove the converse, assume there exists a conditioned invariant subspace $\mathscr{V}$ of codimension q satisfying the inclusion (7.51). Let $Z \in \mathbb{F}^{q \times n}$ be of full row rank q such that $\operatorname{Ker} Z = \mathscr{V}$. The inclusion (7.51) implies the factorization

$$K = \begin{pmatrix} J & E \end{pmatrix} \begin{pmatrix} Z \\ C \end{pmatrix}.$$

Since $\mathscr{V}$ is a conditioned invariant subspace, there exists an output injection L such that $(A - LC)\mathrm{Ker}\,Z \subset \mathrm{Ker}\,Z$. This inclusion implies the existence of $F \in \mathbb{F}^{q\times q}$ for which $Z(A-LC) = FZ$. Defining $G = ZL$ and $H = ZB$, the Sylvester equations (7.30) and an order q observer are obtained.

$(a) \Leftrightarrow (c)$.

Assume that Σ_{est}, given by (7.49), is an order-q tracking observer. Define the transfer function

$$\begin{pmatrix} Z_1(z) & Z_2(z) \end{pmatrix} = \left[\begin{array}{c|cc} F & G & Z \\ \hline J & E & 0 \end{array}\right], \tag{7.56}$$

that is,

$$\begin{aligned} Z_1(z) &= J(zI-F)^{-1}G + E, \\ Z_2(z) &= J(zI-F)^{-1}Z. \end{aligned}$$

Using the Sylvester equations (7.30), we compute $Z_1(z)Z_C(z) + Z_2(z) =$

$$\begin{aligned} &= (J(zI-F)^{-1}G+E)C(zI-A)^{-1} + J(zI-F)^{-1}Z \\ &= J(zI-F)^{-1}GC(zI-A)^{-1} + (K-JZ)(zI-A)^{-1} + J(zI-F)^{-1}Z \\ &= K(zI-A)^{-1} + J(zI-F)^{-1}[GC - (zI-F)Z + Z(zI-A)](zI-A)^{-1} \\ &= K(zI-A)^{-1} + J(zI-F)^{-1}[GC + FZ - ZA](zI-A)^{-1} \\ &= K(zI-A)^{-1} = Z_K(z), \end{aligned}$$

i.e., we obtain a proper solution of (7.53), of McMillan degree q. Note that the equivalence of the solvability of equations (7.52) and (7.53) is trivial.

Conversely, assume that, with $Z_C(z), Z_K(z)$ defined in (7.16), $Z_1(z), Z_2(z)$ is a proper solution of equation (7.53), of McMillan degree q. Note that, since $Z_K(z)$ and $Z_1(z)Z_C(z)$ are both strictly proper, necessarily $Z_2(z)$ is strictly proper, too. Therefore, a minimal realization of $\begin{pmatrix} Z_1(z) & Z_2(z) \end{pmatrix}$ has the form $\left(\begin{array}{c|cc} F & G & Z \\ \hline J & E & 0 \end{array}\right)$, which has dimension q. Then

$$0 = K(zI-A)^{-1} - (J(zI-F)^{-1}G+E)C(zI-A)^{-1} - J(zI-F)^{-1}Z.$$

By inspection of the residue term, this implies $K = JZ + EC$. Substituting this back into the previous equation, we compute

$$\begin{aligned}
0 &= (JZ+EC)(zI-A)^{-1} - (J(zI-F)^{-1}G+E)C(zI-A)^{-1} - J(zI-F)^{-1}Z \\
&= JZ(zI-A)^{-1} - J(zI-F)^{-1}GC(zI-A)^{-1} - J(zI-F)^{-1}Z \\
&= J(zI-F)^{-1}[(zI-F)Z - GC - Z(zI-A)](zI-A)^{-1} \\
&= J(zI-F)^{-1}[-FZ-GC+ZA](zI-A)^{-1}.
\end{aligned}$$

By the nonsingularity of $zI-A$ and the observability of the pair (J,F), we conclude that $ZA-FZ-GC=0$. Defining $H=ZB$, it follows that (7.29) is a tracking observer of dimension at most q.

Proof of the equivalence $(a) \Leftrightarrow (b)$ in part 4.

Assuming there exists an outer detectable subspace $\mathscr{D}$ that satisfies (7.54), there exists an output injection map L for which $(A-LC)\mathscr{D} \subset \mathscr{D}$ and the induced map $(A-LC)|_{\mathscr{X}/\mathscr{D}}$ is stable. As in part 1, we set $\mathscr{D} = \operatorname{Ker} Z$ for some surjective linear transformation Z; then, by Lemma 3.8, there exists a map F that satisfies $Z(A-LC) = FZ$, i.e., $ZA-FZ=GC$ with $G=ZL$, as well as maps J and E, with $K=JZ+EC$. The stability of the induced map $(A-LC)|_{\mathscr{X}/\mathscr{D}}$ and the isomorphism $F \simeq (A-LC)|_{\mathscr{X}/\mathscr{D}}$ imply the stability of F. Finally, we define $H=ZB$. Thus, equations (7.30) have been derived.

Conversely, assume Σ_{est} is an asymptotic tracking observer for Σ_{sys}. By Proposition 7.21, there exists a reduced-order observer for which the Sylvester equations (7.30) are satisfied, with Z of full row rank $\overline{q} \le q$. By the surjectivity of Z, there exists an L for which $G=ZL$, and hence $Z(A-LC)=FZ$ holds. Moreover, since $K=JZ+EC$, $\mathscr{D} = \operatorname{Ker} Z$ is a conditioned invariant subspace of codimension $\overline{q} \le q$ that satisfies (7.54). By the surjectivity of Z, the map $(A-LC)|_{\mathbb{F}^n/\mathscr{D}}$ is isomorphic to F. Since F is stable, $\mathscr{D}$ is an outer detectability subspace of codimension $\overline{q} \le q$.

$(a) \Leftrightarrow (c)$

Assume the Sylvester equations (7.30) hold, with F stable. Then $Z_1(z)$ and $Z_2(z)$, as defined in (7.56), are necessarily proper and stable. Conversely, assume equation (7.52) is solvable with strictly proper and stable $Z_1(z)$ and $Z_2(z)$. Choose a minimal realization

$$\left(\begin{array}{c|cc} F & G & Z \\ \hline J & E & 0 \end{array}\right)$$

of $\begin{pmatrix} Z_1(z) & Z_2(z) \end{pmatrix}$. Necessarily, F is stable. By part 1, equations (7.30) are satisfied with F stable.

Proof of the equivalence $(a) \Leftrightarrow (b)$ in part 5.

Assume there exists a spectrally assignable family of observers (7.29). The subspace $\mathscr{O} = \operatorname{Ker} Z$ is a conditioned invariant subspace such that, for each polynomial $f(z)$ of degree equal to $\operatorname{codim} \mathscr{O}$, there exists a friend L of $\mathscr{O}$ for which the characteristic polynomial of F is $f(z)$. Necessarily, $\mathscr{O}$ is an outer observability subspace.

Conversely, assume that $\mathscr{O}$ is an observability subspace of codimension q that satisfies (7.55). Let $\mathscr{O}$ have the kernel representation $\mathscr{O} = \operatorname{Ker} Z$, with Z surjective. By the definition of observability subspaces, for each polynomial $f(z)$ of degree q there exists a friend L of $\mathscr{O}$ for which $Z(A-LC) = FZ$ and F has $f(z)$ as its characteristic polynomial. From the inclusion (7.55) it follows that there exist unique J and E for which $K = JZ + EC$. Finally, by defining $G = ZL$, the required family of observers is obtained.

$(c) \Leftrightarrow (b)$

Let $\mathscr{O}_*$ denote the unobservability subspace, i.e., the smallest outer observability subspace of (C,A). Assume that there exists an outer observability subspace $\mathscr{O}$ that is contained in the kernel of K. Thus $\mathscr{O}_* \subset \operatorname{Ker} K$, i.e., with respect to the representation (7.9), one has $K_2 = 0$. Since (C_1, A_{11}) is observable, there exist polynomial matrices $P_1(z)$ and $P_2(z)$, with $P_1(z)C_1 + P_2(z)(zI - A_{11}) = K_1$. Thus

$$\begin{pmatrix} P_1(z) & P_2(z) & 0 \end{pmatrix} \begin{pmatrix} C_1 & 0 \\ zI - A_{11} & 0 \\ -A_{21} & zI - A_{22} \end{pmatrix} = \begin{pmatrix} K_1 & 0 \end{pmatrix},$$

and a polynomial solution of (7.53) was constructed.

Conversely, let $P_1(z), P_2(z)$ denote a polynomial solution to (7.53). Then

$$P_1 C(zI - A)^{-1} + P_2(z) = K(zI - A)^{-1}.$$

Choose an unobservable state $x \in \mathscr{O}_*$. Thus $C(zI-A)^{-1}x = 0$, and therefore

$$P_2(z)x = K(zI - A)^{-1}x.$$

Since the left-hand side is polynomial and the right-hand side is strictly proper, we conclude $K(zI-A)^{-1}x = 0$ and $P_2(z)x = 0$ for all $x \in \mathscr{O}_*$. This implies $Kx = 0$, i.e., $\mathscr{O}_* \subset \operatorname{Ker} K$. This proves the converse. Moreover, $x = \operatorname{col}(u, v)$, with $u = 0$ and $K_2(zI - A_{22})^{-1}v = K(zI-A)^{-1}x = 0$. Thus the observability of (K_2, A_{22}) implies $\mathscr{O}_* = \{0\}$. This shows the observability of (C,A). ∎

Remarks: 1. It has already been noted that trackability is a weak concept. Therefore, one expects that a tracking observer for Σ_{sys}, given by (7.45), should always exist. This is indeed the case. One can define the observer as

$$\begin{aligned} \xi_{t+1} &= A\xi_t + Bu_t, \\ \zeta_t &= K\xi_t, \end{aligned} \tag{7.57}$$

and check that it is a tracking observer. Also, note that one strictly proper solution of (7.52) is given by $\begin{pmatrix} Z_1(z) & Z_2(z) \end{pmatrix} = \begin{pmatrix} 0 & K(zI-A)^{-1} \end{pmatrix}$. This also leads to the observer (7.57). Finally, the zero subspace is a conditioned invariant subspace for (C,A) and is contained in $\operatorname{Ker} K$. This allows us to take $Z = I$ and, hence, from the Sylvester equations, show that $F = A$ and $J = K$. So, once again, we are back

to the observer (7.57). Such an observer is of course totally useless because it disregards all the observed data y.
2. Note that the existence of fixed-order tracking observers with preassignable spectra does not necessarily imply the existence of a suitable observability subspace, not even in the minimal-order case.
3. Clearly, the existence of a spectrally assignable family of observers implies the existence of an asymptotic observer. In particular, part 3 (c) of Theorem 7.24 should imply part 2 (c), and in the same way, part 3 (d) should imply part 2 (d). This can be verified directly using partial realization theory, a topic that will not be discussed in this book.

It may be of interest to understand the conditions under which the observer equations (7.49) can be simplified to the form $\zeta_{t+1} = F\zeta_t + Gy_t + Hu_t$. The following proposition addresses this question and gives a geometric characterization for the existence of strongly tracking observers.

Proposition 7.25. *Consider the system*

$$\Sigma_{sys} := \begin{cases} x_{t+1} = Ax_t + Bu_t, \\ y_t = Cx_t, \\ z_t = Kx_t. \end{cases}$$

It is assumed that both C and K have full row rank and that $\left(\begin{pmatrix} C \\ K \end{pmatrix}, A\right)$ is observable.

1. A tracking observer

$$\Sigma_{est} := \begin{cases} \xi_{t+1} = F\xi_t + Gy_t + Hu_t, \\ \zeta_t = J\xi_t, \end{cases} \tag{7.58}$$

with (J,F) observable, is a strongly tracking observer if and only if J is nonsingular. In that case, we may assume without loss of generality that the observer is given by

$$\zeta_{t+1} = F\zeta_t + Gy_t + Hu_t. \tag{7.59}$$

2. A strongly tracking observer of the form (7.59) exists if and only if $\operatorname{Ker} K$ is a conditioned invariant subspace. In this case, the error dynamics are given by

$$e_{t+1} = Fe_t. \tag{7.60}$$

Proof. 1. Assume J in (7.58) is nonsingular. The error dynamics are given by (7.34), and hence

$$e_{t+1} = J\varepsilon_{t+1} = JF\varepsilon_t = JFJ^{-1}J\varepsilon_t = JFJ^{-1}e_t.$$

This shows that $e_t = (JFJ^{-1})^{t-1}e_0$, and hence $e_0 = 0$ implies $e_t = 0$, i.e., Σ_{est} is a strongly tracking observer.

Conversely, assume that Σ_{est} is a strongly tracking observer. The error dynamics are given by (7.34), and hence $e_t = JF^{t-1}\varepsilon_0$. By the property of strong tracking, $e_0 = 0$ implies $e_t = 0$ for all $t \geq 0$, i.e., $\varepsilon_0 \in \bigcap \operatorname{Ker} JF^{t-1}$. By the observability of the pair (J,F), we conclude that $e_0 = 0$ implies $\varepsilon_0 = 0$. This shows that J is injective and, hence, since it was assumed that J has full row rank, actually invertible. Substituting $\xi = J^{-1}\zeta$ into the observer equation and multiplying from the left by J, $\zeta_{t+1} = (JFJ^{-1})\zeta_t + (JG)y_t + (JH)u_t$ follows. Modifying appropriately the definitions of F,G,H, equation (7.59) is proved.

2. Assume (7.59) is a strongly tracking observer. By Theorem 7.17, there exists a map Z satisfying the following Sylvester equations:

$$\begin{aligned} ZA &= FZ + GC, \\ H &= ZB, \\ K &= Z. \end{aligned}$$

Letting $x \in \operatorname{Ker} K \cap \operatorname{Ker} C$ implies $K(Ax) = 0$, i.e., $Ax \in \operatorname{Ker} K$, so $\operatorname{Ker} K$ is a conditioned invariant subspace.

Conversely, assume $\operatorname{Ker} K$ is a conditioned invariant subspace. Letting $Z = K$, there exists a map L such that $(A - LC)\operatorname{Ker} K \subset \operatorname{Ker} K$, and using, once again Lemma 3.8, we infer that $K(A - LC) = FK$ for some L. Thus $KA - FK = GC$, with $G = KL$. Setting $J = I$ and defining $H = KB$, we are done. That the error dynamics are given by (7.60) follows from (7.34) and the fact that $J = I$.

■

This section ends with some ideas on state-space constructions of functional observers. Theorem 7.24 contains equivalent characterizations for minimal-order observers but does not give the minimal order of a tracking observer or a way of computing such an observer. From a practical point of view, it is important to have a systematic way of constructing observers, and, for computational efficiency, it is important to have the order of the observer minimal. These questions are addressed next. To this end, assume Σ_{est}, defined by (7.49), is a minimal-order tracking observer for system (7.50). By Theorem 7.17, there exists a solution Z of the observer Sylvester equations (7.30), and, by Proposition 7.21, it may be assumed without loss of generality that Z has full row rank. Define now $\mathscr{V} = \operatorname{Ker} Z$. The first equation in (7.30) means that $A(\mathscr{V} \cap \operatorname{Ker} C) \subset \mathscr{V}$, i.e., that $\mathscr{V}$ is a conditioned invariant subspace, whereas the last equation in (7.30) means that $\mathscr{V} \cap \operatorname{Ker} C \subset \operatorname{Ker} K$. In view of the geometric characterizations given in Theorem 7.24, to find minimal-order observers for system (7.28), one must find all maximal dimensional conditioned invariant subspaces $\mathscr{V}$ that satisfy $\mathscr{V} \cap \operatorname{Ker} C \subset \operatorname{Ker} K$. Since the set of all conditioned invariant subspaces is closed under intersections but not under sums, one must approach the minimality question differently. In spirit, we follow Michelangelo's dictum: "Carving is easy, you just go down to the skin and stop."

Therefore, to get a minimal-order tracking observer, one must choose a minimal rank extension $\begin{pmatrix} z \\ c \end{pmatrix}$ for which $\mathscr{V} = \operatorname{Ker} Z$ is conditioned invariant and

$$\mathscr{V} \cap \operatorname{Ker} C = \operatorname{Ker} \begin{pmatrix} Z \\ C \end{pmatrix} \subset \operatorname{Ker} K. \tag{7.61}$$

By the minimality of such an extension, Z necessarily has full row rank. It is easy to fulfill the second requirement, simply by choosing $Z = K$, which implies

$$\operatorname{Ker} \begin{pmatrix} Z \\ C \end{pmatrix} = \operatorname{Ker} \begin{pmatrix} K \\ C \end{pmatrix} \subset \operatorname{Ker} K.$$

There are two problems with this choice, of which the first one is minor. Since, in the case where $\operatorname{Ker} K$ and $\operatorname{Ker} C$ have a nontrivial intersection, K can be reduced modulo C to get a lower rank extension K'. The second problem is due to the fact that, in general, there is no reason why $\mathscr{V} = \operatorname{Ker} K$ should be conditioned invariant. (But, as we shall see in Example 7.32, it may.) One way to overcome this is to add, if necessary, additional terms to Z. An easy way to do this is to set

$$Z = \begin{pmatrix} KA^{n-1} \\ \vdots \\ KA \\ K \end{pmatrix}.$$

Clearly, $\operatorname{Ker} Z$ is not only a conditioned invariant subspace but actually an invariant one. However, in general, the constructed Z does not have full row rank and thus would lead to a nonminimal observer. The remedy to these two problems is to maintain a fine balance between increasing the rank of Z sufficiently so that $\operatorname{Ker} Z$ is conditioned invariant and (7.61) being satisfied, but small enough to preserve the maximality of the dimension of $\operatorname{Ker} Z$ and, hence, leading to a minimal-order functional observer. This will be treated in Theorem 7.27.

To formalize the reduction process, the following proposition, which is of interest on its own, is stated and proved.

Proposition 7.26. *1. Let $S \in \mathbb{F}^{n\times q}$ and $T \in \mathbb{F}^{n\times p}$ be of full column rank. Then there exists $S' \in \mathbb{F}^{n\times r}$, of full column rank, such that*

$$\operatorname{Im} S' \subset \operatorname{Im} S,$$
$$\operatorname{Im} S' + \operatorname{Im} T = \operatorname{Im} S + \operatorname{Im} T,$$
$$\operatorname{Im} S' \cap \operatorname{Im} T = \{0\},$$

implying the direct sum representation $\operatorname{Im} S + \operatorname{Im} T = \operatorname{Im} S' \oplus \operatorname{Im} T$.

2. *Let $K \in \mathbb{F}^{q\times n}$ and $C \in \mathbb{F}^{p\times n}$ be of full row rank. Then there exists $K' \in \mathbb{F}^{r\times n}$, of full row rank, such that*

$$\operatorname{Ker} K' \supset \operatorname{Ker} K,$$
$$\operatorname{Ker} K' \cap \operatorname{Ker} C = \operatorname{Ker} K \cap \operatorname{Ker} C,$$
$$\operatorname{Ker} K' + \operatorname{Ker} C = \mathbb{F}^n,$$

i.e., $\operatorname{Ker} K \cap \operatorname{Ker} C$ *is the transversal intersection of* $\operatorname{Ker} K'$ *and* $\operatorname{Ker} C$*, which implies the direct sum representation*

$$\mathbb{F}^n/(\operatorname{Ker} K' \cap \operatorname{Ker} C) = \operatorname{Ker} K'/(\operatorname{Ker} K' \cap \operatorname{Ker} C) \oplus \operatorname{Ker} C/(\operatorname{Ker} K' \cap \operatorname{Ker} C).$$

Moreover,

$$r = \operatorname{rank} K' = \operatorname{codim} \operatorname{Ker} K'.$$

Proof. 1. Noting that S is assumed to be of full column rank, it follows that, with S_i being the columns of S, the set $\mathscr{B} = \{S_1, \dots, S_q\}$ is a basis for $\operatorname{Im} S$. Let $\{R_1, \dots, R_r\}$ be a basis for $\operatorname{Im} S \cap \operatorname{Im} T$. By the basis exchange theorem, there exist r elements of $\mathscr{B}$, which without loss of generality one can take to be the first r, for which $\{R_1, \dots, R_r, S_{r+1}, \dots, S_q\}$ is a basis for $\operatorname{Im} S$. Defining $S' = \begin{pmatrix} S_{r+1} & \dots & S_q \end{pmatrix}$, we are done.

2. The first assertion follows from the first part by duality considerations, while the second one follows from the identity $\dim \operatorname{Ker} K' + \dim \operatorname{Im} K' = n$.

∎

Let

$$\Sigma_{sys} := \begin{cases} x_{t+1} &= Ax_t + Bu_t, \\ y_t &= Cx_t, \\ z_t &= Kx_t \end{cases}$$

be a linear system acting in the state space $\mathbb{F}^n$. Assume that $(\begin{pmatrix} C \\ K \end{pmatrix}, A)$ is observable and that both C and K are of full row rank. Define the **ith partial observability matrix** $\mathscr{O}_i(K, A)$ by

$$\mathscr{O}_i(K,A) := \begin{pmatrix} KA^{i-1} \\ \vdots \\ KA \\ K \end{pmatrix}.$$

Define, inductively, a sequence of full row rank matrices $\{Z_i\}$ as follows:
Set $Z_0 = 0$, and proceed inductively. Assume $Z_0, \dots, Z_i$ are constructed. If $\operatorname{Ker} Z_i \cap \operatorname{Ker} C$ is conditioned invariant, then one sets $Z = Z_i$ and stops. Otherwise, Proposition 7.26 is applied to construct Z_{i+1}, which satisfies

$$\begin{cases} \operatorname{Ker} Z_{i+1} \supset \operatorname{Ker} \mathscr{O}_{i+1}(K,A), \\ \operatorname{Ker} \begin{pmatrix} Z_{i+1} \\ C \end{pmatrix} = \operatorname{Ker} \begin{pmatrix} KA^i \\ Z_i \\ C \end{pmatrix} = \operatorname{Ker} \begin{pmatrix} \mathscr{O}_{i+1}(K,A), \\ C \end{pmatrix} \\ \operatorname{Ker} Z_{i+1} + \operatorname{Ker} C = \mathbb{F}^n. \end{cases} \tag{7.62}$$

Theorem 7.27. *Let Z_i be constructed as above. The following assertions are true:*

1. *For all i,* $\operatorname{Ker} \begin{pmatrix} Z_{i+1} \\ C \end{pmatrix} \subset \operatorname{Ker} \begin{pmatrix} Z_i \\ C \end{pmatrix}$.
2. *Let ν be the smallest index with*

$$\operatorname{Ker} \begin{pmatrix} Z_{i+1} \\ C \end{pmatrix} = \operatorname{Ker} \begin{pmatrix} Z_i \\ C \end{pmatrix}. \tag{7.63}$$

Setting $Z = Z_\nu$, the subspace $\mathscr{V} = \operatorname{Ker} Z$ is a maximal conditioned invariant subspace that satisfies (7.61).

3. *There exist matrices $F \in \mathbb{F}^{q\times q}$, $G \in \mathbb{F}^{q\times p}$, $J \in \mathbb{F}^{k\times q}$, and $E \in \mathbb{F}^{k\times p}$ for which the following Sylvester equations are satisfied:*

$$ZA = FZ + GC,$$
$$K = JZ + EC.$$

Then the system

$$\Sigma_{est} := \begin{cases} \xi_{t+1} = F\xi_t + Gy_t + Hu_t, \\ \zeta_t = J\xi_t + Ey_t \end{cases}$$

is an order-q tracking observer for K, where

$$q = \operatorname{rank} Z = \operatorname{codim} \operatorname{Ker} Z. \tag{7.64}$$

Proof. 1. Follows from the equality $\operatorname{Ker} \begin{pmatrix} Z_{i+1} \\ C \end{pmatrix} = \operatorname{Ker} \begin{pmatrix} KA^i \\ Z_i \\ C \end{pmatrix}$.

2. The equalities (7.62) and (7.63) imply the equality

$$\mathrm{Ker}\begin{pmatrix}\mathscr{O}_i(K,A)\\ C\end{pmatrix} = \mathrm{Ker}\begin{pmatrix}\mathscr{O}_{i+1}(K,A)\\ C\end{pmatrix},$$

which in turn implies the inclusion

$$A\mathrm{Ker}\begin{pmatrix}\mathscr{O}_i(K,A)\\ C\end{pmatrix} \subset \mathrm{Ker}\begin{pmatrix}\mathscr{O}_{i+1}(K,A)\\ C\end{pmatrix}.$$

Using this, one computes

$$\begin{aligned} A(\mathscr{V}\cap \mathrm{Ker}\, C) &= A\mathrm{Ker}\begin{pmatrix}Z_\nu\\ C\end{pmatrix} = A\mathrm{Ker}\begin{pmatrix}\mathscr{O}_\nu(K,A)\\ C\end{pmatrix} \subset \mathrm{Ker}\begin{pmatrix}\mathscr{O}_{\nu+1}(K,A)\\ C\end{pmatrix}\\ &= \mathrm{Ker}\begin{pmatrix}Z_{\nu+1}\\ C\end{pmatrix} = \mathrm{Ker}\begin{pmatrix}Z_\nu\\ C\end{pmatrix} = \mathscr{V},\end{aligned}$$

showing that $\mathscr{V}$ is conditioned invariant. Since, for $i \geq 1$, one has $\mathrm{Ker}\,\mathscr{O}_i(K,A) \subset \mathrm{Ker}\,K$, it follows that

$$\mathscr{V}\cap \mathrm{Ker}\, C = \mathrm{Ker}\begin{pmatrix}Z_\nu\\ C\end{pmatrix} = \mathrm{Ker}\begin{pmatrix}\mathscr{O}_\nu(K,A)\\ C\end{pmatrix} \subset \mathrm{Ker}\, K,$$

i.e., (7.61) holds.

3. Follows from Theorem 7.24.

■

The decision whether or not q, defined in (7.64), is the minimal order of tracking observers is left as an open problem. Next, we consider a few special cases of Theorem 7.27.

Case I: $\mathrm{Ker}\, C \subset \mathrm{Ker}\, K$. This means that unobserved states are not to be estimated, or, equivalently stated, states to be estimated are directly observed. From the inclusion $\mathrm{Ker}\, C \subset \mathrm{Ker}\, K$ one deduces that there exists an E for which $K = EC$. This implies $Z = 0$, $F = 0$, $G = 0$, and $J = 0$. Thus, as intuitively expected, the existence of a zero-order or, equivalently, a nondynamic, observer for K is obtained.

Case II: $K = I$, i.e., tracking the state. To this end, one needs a map Z for which $\mathrm{Ker}\begin{pmatrix}Z\\ C\end{pmatrix} \subset \mathrm{Ker}\, K = \{0\}$. The easiest choice, though not necessarily the minimal one, is to take $Z = I$ and $E = 0$ and $G = 0$. This implies $J = I$ and $F = A$. Thus, a tracking observer is given by

$$\Sigma_{est} := \begin{cases} \sigma\xi = A\xi, \\ \zeta = \xi, \end{cases}$$

and the Luenberger state observer has been rederived. To obtain a reduced-order state observer, Z is chosen so that $\begin{pmatrix}Z\\ C\end{pmatrix}$ is nonsingular.

Case III: $C = 0$, i.e., no observations are available. By our assumption of the observability of the pair $(\begin{pmatrix} C \\ K \end{pmatrix}, A)$, (K, A) is necessarily an observable pair. Let $K_1, \dots, K_q$ be the rows of K. Since K is assumed to be of full row rank, the rows are linearly independent. The **dual Kronecker indices** are defined by the following deletion process:

Starting from the top of the observability matrix

$$\mathrm{col}\,(K_1, \dots, K_q, K_1A, \dots, K_qA, \dots, K_1A^{n-1}, \dots, K_qA^{n-1}),$$

delete all row vectors that are linearly dependent on the set of preceding ones. In this way one obtains, up to a permutation of the rows, a matrix of the form

$$Z = \mathrm{col}\,(K_1, K_1A, \dots, K_1A^{\nu_1-1}, \dots, K_q, \dots, K_qA^{\nu_q-1}).$$

The observability of (K, A) implies $\sum_{i=1}^{q} \nu_i = n$ and, hence, the invertibility of Z. Defining $F = ZAZ^{-1}$, $J = KZ^{-1}$, then

$$\Sigma_{est} := \begin{cases} \sigma\xi = F\xi, \\ \zeta = J\xi \end{cases}$$

is a tracking observer. For tracking, given an initial condition x_0 of the state system, one chooses $\xi_0 = Zx_0$.

Example 7.28. This example is taken from Fernando, Trinh, Hieu and Jennings (2010). One takes A, C, K as follows:

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 1 & 0 \end{pmatrix}, \quad K = \begin{pmatrix} 1 & 2 & 0 \end{pmatrix}.$$

Note that the rows of C and K are linearly independent, so $Z_1 = K$. Computing $KA = \begin{pmatrix} -1 & -2 & 0 \end{pmatrix}$, which is linearly dependent on K, we conclude that $Z = Z_1 = K$. Since $\mathrm{Ker}\,Z + \mathrm{Ker}\,C = \mathbb{F}^3$, $\mathscr{V} = \mathrm{Ker}\,Z$ is a tight conditioned invariant subspace. It is easily checked that the observer Sylvester equations have a unique solution given by $F = -1, G = 0, J = 1, E = 0$, which gives an asymptotic observer.

Example 7.29. Let A, C, K be as follows:

$$A = \left(\begin{array}{ccc|cc} -1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right), \quad C = \left(\begin{array}{ccc|cc} 0 & 0 & 1 & 0 & 0 \end{array}\right), \quad K = \left(\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 1 \end{array}\right).$$

In this example, the pair (C,A) is not observable, but $\left(\begin{pmatrix} C \\ K \end{pmatrix}, A\right)$ is. Since C and K are linearly independent, one gets $Z_1 = K$. Computing

$$\mathscr{O}_5(K,A) = \begin{pmatrix} KA^4 \\ KA^3 \\ KA^2 \\ KA \\ K \end{pmatrix} = \begin{pmatrix} -4 & 1 & 0 & 4 & 1 \\ 3 & -1 & 0 & 3 & 1 \\ -2 & 1 & 0 & 2 & 1 \\ 1 & -1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix},$$

it is seen that the bottom four rows are linearly independent, but the top row depends linearly on them. Therefore, $Z = \mathscr{O}_4(K,A)$ and $\operatorname{Ker} Z = \{\operatorname{col}(0,0,\gamma,0,0) | \gamma \in \mathbb{R}\}$. Since

$$A \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \\ 0 \end{pmatrix},$$

$\operatorname{Ker} Z$ is actually an A-invariant subspace. It is easily checked that $\operatorname{Ker} Z + \operatorname{Ker} C = \mathbb{R}^5$, hence $\mathscr{V} = \operatorname{Ker} Z$ is a tight conditioned invariant subspace, which means that the corresponding tracking observer is uniquely determined up to similarity. Using the observer Sylvester equations, one obtains

$$F = \begin{pmatrix} 0 & 2 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad G = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad J = \begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix}, \quad E = \begin{pmatrix} 0 \end{pmatrix}.$$

7.5 Construction of Functional Observers

In Theorem 7.24, characterizations of various classes of observers in terms of conditioned invariant, outer detectability, and outer observability subspaces appear. Recalling that these subspaces have nice functional representations, given in Proposition 6.67 and Theorem 6.72, it is only natural to attempt observer construction using these representations. Key ingredients in the analysis are the coprime factorization (7.19), the parameterizations (7.18) and (7.21), and the shift realization (4.20).

Consider the system

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t, \\ y_t &= Cx_t, \\ z_t &= Kx_t, \end{aligned} \tag{7.65}$$

with $A \in \mathbb{F}^{n\times n}, B \in \mathbb{F}^{n\times m}, C \in \mathbb{F}^{p\times n}, K \in \mathbb{F}^{k\times n}$, assuming that $\left(\begin{pmatrix} C \\ K \end{pmatrix}, A\right)$ is observable. Let C, K, A have the representation (7.9), with respect to the direct sum representation (7.10), and let

$$\begin{pmatrix} C_1 & 0 \\ K_1 & K_2 \end{pmatrix} \begin{pmatrix} zI - A_{11} & 0 \\ -A_{21} & zI - A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} D_{11}(z) & 0 \\ D_{21}(z) & D_{22}(z) \end{pmatrix}^{-1} \begin{pmatrix} \Theta_{11}(z) & 0 \\ \Theta_{21}(z) & \Theta_{22}(z) \end{pmatrix}$$

be an adapted coprime factorization, i.e., satisfying conditions (1a)–(1c) of Theorem 7.9. Under these assumptions the following assertions are true.

Theorem 7.30. *1. The map*

$$\begin{aligned} \Theta : X_{\begin{pmatrix} zI - A_{11} & 0 \\ -A_{21} & zI - A_{22} \end{pmatrix}} &\longrightarrow X_{\begin{pmatrix} D_{11}(z) & 0 \\ D_{21}(z) & D_{22}(z) \end{pmatrix}} \\ \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \Theta \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} &= \begin{pmatrix} \Theta_{11}(z) & 0 \\ \Theta_{21}(z) & \Theta_{22}(z) \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \end{aligned} \tag{7.66}$$

is an $\mathbb{F}[z]$-isomorphism. Defining, via the shift realization,

$$\overline{A} := S_{\begin{pmatrix} D_{11} & 0 \\ D_{21} & D_{22} \end{pmatrix}}, \quad \overline{C}\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} := \left(\begin{pmatrix} D_{11}^{-1} & 0 \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \right)_{-1} = (D_{11}^{-1} f_1)_{-1},$$

$$\overline{K}\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} := \left(\begin{pmatrix} -D_{22}^{-1} D_{21} D_{11}^{-1} & D_{22}^{-1} \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \right)_{-1} = \left(-D_{22}^{-1} D_{21} D_{11}^{-1} f_1 + D_{22}^{-1} f_2 \right)_{-1}, \tag{7.67}$$

the intertwining relations

$$\Theta A = \overline{A}\Theta, \quad C = \overline{C}\Theta, \quad K = \overline{K}\Theta \tag{7.68}$$

are satisfied.

2. *For the parameterization (7.18) of the set of rational solutions of (7.17), we can choose a proper rational matrix $W(z)$, with left coprime factorization*

$$W(z) = T(z)^{-1} L(z), \tag{7.69}$$

such that $Z_1(z)$ is proper. Let $Q(z)$ be the l.c.l.m.$(T(z), D_{22}(z))$, and let $R_D(z)$ and $R_L(z)$ be left coprime polynomial matrices with

$$Q(z) = R_D(z)D_{22}(z) = R_T(z)T(z). \tag{7.70}$$

Then the parameterization (7.18) can be rewritten as

$$\begin{aligned} Z_1(z) &= Q(z)^{-1}(-R_D(z)D_{21}(z) + R_T(z)L(z)D_{11}(z)), \\ Z_2(z) &= Q(z)^{-1}\left(-R_T(z)L(z)\Theta_{11}(z) + R_D(z)\Theta_{21}(z) \;\; R_D(z)\Theta_{22}(z).\right) \end{aligned} \tag{7.71}$$

3. *Define maps*

$$Z : X_{\begin{pmatrix} D_{11} & 0 \\ D_{21} & D_{22} \end{pmatrix}} \longrightarrow X_Q, \quad F : X_Q \longrightarrow X_Q, \quad G : \mathbb{F}^p \longrightarrow X_Q,$$

$$H : \mathbb{F}^m \longrightarrow X_Q, \quad J : X_Q \longrightarrow \mathbb{F}^k, \quad E : \mathbb{F}^p \longrightarrow \mathbb{F}^k$$

by

$$\begin{aligned} Z &= \pi_Q \left(-R_T L \;\; R_D, \right) | X_{\begin{pmatrix} D_{11} & 0 \\ D_{21} & D_{22} \end{pmatrix}} \\ F &= S_Q, \\ G &= -\pi_Q \left(\left(-R_T L \;\; R_D, \right) \begin{pmatrix} D_{11} \\ D_{21} \end{pmatrix} (\cdot) \right) \\ H &= ZB, \\ J &= \left(Q^{-1}(\cdot) \right)_{-1}, \\ E &= \pi_+ Q^{-1} (-R_D D_{21} + R_T L D_{11}) \;. \end{aligned} \tag{7.72}$$

Then the system

$$\Sigma_{est} = \begin{cases} \xi_{t+1} = F\xi_t + Gy_t + Hu_t, \\ \quad \zeta_t = J\xi_t + Ey_t, \end{cases} \tag{7.73}$$

defined in the state space X_Q by these maps, is an observable tracking observer for K with transfer functions

$$Z_1(z) = \left[\begin{array}{c|c} F & G \\ \hline J & E \end{array} \right], \quad Z_2(z) = \left[\begin{array}{c|c} F & Z \\ \hline J & 0 \end{array} \right]$$

of (F, G, J, E) and $(F, Z, J, 0)$, respectively. The realization (F, G, J, E) of $Z_1(z)$ is always observable. It is reachable if and only if the polynomial matrices $Q(z)$ and $-R_T(z)L(z)D_{11}(z) + R_D(z)D_{21}(z)$ are left coprime.

Proof. The adapted coprime factorization leads to the intertwining relation

$$\begin{pmatrix} \Theta_{11}(z) & 0 \\ \Theta_{21}(z) & \Theta_{22}(z) \end{pmatrix} \begin{pmatrix} zI - A_{11} & 0 \\ -A_{21} & zI - A_{22} \end{pmatrix} = \begin{pmatrix} D_{11}(z) & 0 \\ D_{21}(z) & D_{22}(z) \end{pmatrix} \begin{pmatrix} C_1 & 0 \\ K_1 & K_2 \end{pmatrix}.$$

Applying Theorem 3.21 shows that the map Θ, defined by (7.68), is an $\mathbb{F}[z]$-module isomorphism. The intertwining relations (7.66) are easily verified. This proves part 1.

Part 2 follows from a straightforward substitution of (7.69) in the parameterization (7.18), using the identities $D_{22}(z)^{-1} = Q(z)^{-1}R_D(z)$ and $T(z)^{-1} = Q(z)^{-1}R_T(z)$.
Proof of part 3. Recall from Theorem 3.30 that

$$S_{\begin{pmatrix} D_{11} & 0 \\ D_{21} & D_{22} \end{pmatrix}} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = z \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} - \begin{pmatrix} D_{11} & 0 \\ D_{21} & D_{22} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix},$$

where $\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \left(\begin{pmatrix} D_{11} & 0 \\ D_{21} & D_{22} \end{pmatrix}^{-1} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \right)_{-1}$. For $f = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \in X_{\begin{pmatrix} D_{11} & 0 \\ D_{21} & D_{22} \end{pmatrix}}$, using (7.67), one computes $(Z\overline{A} - FZ)f =$

$$\begin{aligned}
&= \pi_Q\left(-R_T L\, R_D\right) S_{\begin{pmatrix} D_{11} & 0 \\ D_{21} & D_{22} \end{pmatrix}} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} - S_Q \pi_Q\left(-R_T L\, R_D\right) \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \\
&= \pi_Q\left(-R_T L\, R_D\right) \left(z \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} - \begin{pmatrix} D_{11} & 0 \\ D_{21} & D_{22} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}\right) - \pi_Q\left(-R_T L\, R_D\right) z \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \\
&= -\pi_Q\left(-R_T L\, R_D\right) \begin{pmatrix} D_{11} & 0 \\ D_{21} & D_{22} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = -\pi_Q\left(-R_T L\, R_D\right) \begin{pmatrix} D_{11} \\ D_{21} \end{pmatrix} \xi_1.
\end{aligned}$$

It was shown that, with $\overline{C}, \overline{A}$ defined by (7.67) and Z, F, G by (7.72), the observer Sylvester equation $Z\overline{A} = FZ + G\overline{C}$ is satisfied. The equation $\overline{K} = JZ + E\overline{C}$ can be verified similarly. This completes the proof. ■

A few remarks are in order.

1. The construction of the tracking observer (7.73) works for every left coprime pair of polynomials $T(z), L(z)$, with $T(z)$ nonsingular. In particular, one does not need to impose a properness assumption on

$$Z_1 = -(TD_{22})^{-1}(TD_{21} - LD_{11}).$$

However, if one wants to obtain the observer from a shift realization of $Z_1(z)$, then $W(z) = T(z)^{-1}L(z)$ must be chosen such that Z_1 is proper.

2. The unobservable subspace $\mathscr{O}_*(C,A)$ of system (7.65) has the representation

$$\mathscr{O}_*(C,A) = \begin{pmatrix} zI - A_{11} & 0 \\ -A_{21} & I \end{pmatrix} X_{\begin{pmatrix} I & 0 \\ 0 & zI - A_{22} \end{pmatrix}}$$

and dimension $\deg\det(zI - A_{22})$. Its image under Θ is given by

$$\begin{pmatrix} D_{11}(z) & 0 \\ D_{21}(z) & I \end{pmatrix} X_{\begin{pmatrix} I & 0 \\ 0 & D_{22}(z) \end{pmatrix}},$$

which has dimension $\deg\det D_{22} = \deg\det(zI - A_{22})$.

3. The constructed tracking observer

$$Z_1(z) = \left[\begin{array}{c|c} F & G \\ \hline J & E \end{array}\right]$$

can be written as in (7.71), with the shift realization defined in the state space $X_{TD_{22}}$. Clearly, $\dim X_{TD_{22}} = \deg\det(T) + \deg\det(D_{22})$, and the term $\deg\det(D_{22})$ is the price of tracking the unobservable subspace $\mathscr{O}_*(C,A)$.

4. The choice of the rational matrix $W(z)$ in Theorem 7.30 is closely related to partial realizations; however, we will not follow this path in this book and instead refer the reader to Fuhrmann (2008) for some of the details.

From Theorem 7.30 we deduce several special cases as corollaries.

Corollary 7.31. *Consider system (7.65), with* $\left(\begin{pmatrix} C \\ K \end{pmatrix}, A\right)$ *observable.*

1. *Assume that* (C,A) *is observable.*

   (a) *The coprime factorization (7.5) reduces to*

$$\begin{pmatrix} C \\ K \end{pmatrix} (zI - A)^{-1} = \begin{pmatrix} D_{11}(z) & 0 \\ D_{21}(z) & I \end{pmatrix}^{-1} \begin{pmatrix} \Theta_{11}(z) \\ \Theta_{21}(z) \end{pmatrix}.$$

   (b) *Choose a proper rational matrix* $W(z)$ *with left coprime factorization* $W(z) = T(z)^{-1}L(z)$ *so that* $Z_1(z)$ *in (7.18) is proper. Define maps*

$$Z : X_{D_{11}} \longrightarrow X_T, \quad F : X_T \longrightarrow X_T, \quad G : \mathbb{F}^p \longrightarrow X_T,$$
$$H : \mathbb{F}^m \longrightarrow X_T, \quad J : X_T \longrightarrow \mathbb{F}^k, \quad E : \mathbb{F}^p \longrightarrow \mathbb{F}^k$$

by

$$\begin{aligned} Z &= -\pi_T L | X_{D_{11}}, \\ F &= S_T, \\ G &= -\pi_T L D_{11}(\cdot), \\ H &= ZB, \\ J &= (T^{-1}(\cdot))_{-1}, \\ E &= \pi_+(-D_{21} + T^{-1} L D_{11}). \end{aligned} \tag{7.74}$$

Then system (7.73), defined by F, G, H, J, E, is a tracking observer for K.

2. *Assume $C = 0$, i.e., there are no observations at all.*
   (a) *The pair (K, A) is observable, and the coprime factorization (7.5) reduces to*

   $$K(zI - A)^{-1} = D_{22}(z)^{-1}\Theta_{22}(z).$$

   (b) *The map $Z : X_{zI-A} \longrightarrow X_{D_{22}}$, defined as $Z\xi = \Theta_{22}(z)\xi$, is an isomorphism, and the Sylvester equation (7.42) reduces to $ZA = FZ$, with Z invertible and $K = JZ$.*
   (c) *The pair (J, F), defined in the state space $X_{D_{22}}$ by*

   $$\begin{aligned} F &= S_{D_{22}}, \\ Jg &= (D_{22}^{-1} g)_{-1}, \end{aligned}$$

   is a tracking observer.
3. *Assume the pair (C, A) is detectable.*
   (a) *In the coprime factorization (7.5), the polynomial matrix $D_{22}(z)$ is stable.*
   (b) *In the parameterization (7.18) of the set of rational solutions of (7.17), we can choose the rational matrix $W(z)$ with coprime factorization $W(z) = T(z)^{-1}L(z)$ so that $Z_1(z)$ is proper and stable.*
   (c) *System (7.73), defined by the F, G, H, J, E in (7.72), is an asymptotic tracking observer for K.*

Proof. Basically, this follows from Theorem 7.30. However, we add a few remarks. A simple computation yields

$$X_{\begin{pmatrix} D_{11} & 0 \\ D_{21} & I \end{pmatrix}} = \left\{ \begin{pmatrix} f(z) \\ 0 \end{pmatrix} | f(z) \in X_{D_{11}} \right\}.$$

Note that the intertwining relation

$$\begin{pmatrix} I & 0 \end{pmatrix} \begin{pmatrix} D_{11}(z) & 0 \\ D_{21}(z) & I \end{pmatrix} = D_{11}(z) \begin{pmatrix} I & 0 \end{pmatrix},$$

taken together with the trivial associated coprimeness relations, implies that the map $\Phi : X_{\begin{pmatrix} D_{11} & 0 \\ D_{21} & I \end{pmatrix}} \longrightarrow X_{D_{11}}$, defined by $\begin{pmatrix} f(z) \\ 0 \end{pmatrix} \mapsto f(z)$, is an $\mathbb{F}[z]$-isomorphism. Using this, the maps in (7.72) have the simplified form (7.74). ∎

It should be pointed out that, to compensate for the total absence of observations, i.e., for the case $C = 0$, the order of a minimal tracking observer is necessarily equal to the dimension of the state space of the system.

To clarify the issues of observer characterization and construction, we present an example from two points of view: the state-space formulation on the one hand and the functional model formulation on the other. Each formulation has its own advantages and insights.

Example 7.32. Let A, C, K be as follows:

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}, \quad K = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}.$$

Computing $Z_1 = K = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$, $KA = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}$, clearly,

$$\operatorname{Ker} \begin{pmatrix} K \\ C \end{pmatrix} = \operatorname{Ker} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \subset \operatorname{Ker} \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} = \operatorname{Ker} KA,$$

which shows that $\mathscr{V} = \operatorname{Ker} K$ is a conditioned invariant subspace satisfying $\mathscr{V} \cap \operatorname{Ker} C \subset \operatorname{Ker} K$. Obviously, it is a maximal such subspace. The Sylvester equations lead to an observer defined by

$$\begin{aligned} F &= (0), \quad G = (1) \\ J &= (1), \quad E = (0). \end{aligned}$$

Since the matrix $\begin{pmatrix} K \\ C \end{pmatrix}$ has full row rank, this representation is uniquely determined.

Computing further $KA^2 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}$, and setting $Z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, one has $ZA = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$, and the equation $ZA = FZ + GC$ implies

$$F=\begin{pmatrix}0&0&-t_0\\1&0&-t_1\\0&1&-t_2\end{pmatrix},G=\begin{pmatrix}t_0\\t_1\\t_2\end{pmatrix}. \tag{7.75}$$

The nonuniqueness of the constructed minimal-order observer is a consequence of the inequality $\dim(\operatorname{Ker} Z+\operatorname{Ker} C)=2<3=\dim\mathbb{F}^3$.

From a polynomial point of view, one has $\mathscr{X}=X_{z^3}$ and $A=[S_{z^3}]$. Since $\operatorname{Ker} K=\{\xi_0+\xi_1 z\}$, the maximal conditioned invariant subspace $\mathscr{V}$ contained in $\operatorname{Ker} K$ is the zero subspace; hence, $\mathscr{V}=X_{z^3}\cap t(z)\mathbb{F}[z]$, with $t(z)$ a polynomial of degree greater than or equal to 3. To obtain minimal-order observers, we take $\deg t(z)=3$. Setting $t(z)=t_0+t_1z+t_2z^2+z^3$, we compute $F=[\pi_t|X_{z^3}]$ and $G=-[\pi_t z^3]$. Simple computations lead to (7.75).

7.6 Exercises

1. Consider the undamped harmonic oscillator

$$\begin{aligned}\dot{x}_1(t)&=x_2(t),\\ \dot{x}_2(t)&=-\omega^2x_1(t)+u(t),\\ y(t)&=x_2(t).\end{aligned}$$

   (a) Determine an observer

$$\begin{aligned}\dot{z}(t)&=(A-LC)z(t)+Ly(t)+Bu(t),\\ u(t)&=Fz(t)+v(t),\\ y(t)&=x_2(t),\end{aligned}$$

   such that the eigenvalues of the closed-loop system

$$\begin{pmatrix}A & B\mathscr{F}\\ LC & A-LC+B\mathscr{F}\end{pmatrix}$$

   are $-\omega\pm\sqrt{-1}\omega,-\omega,-\omega$.

   (b) Determine a one-dimensional reduced observer.

2. Consider the third-order system $\dot{x}=Ax+bu, y=cx$, with $\alpha<0,\beta\neq 0,k\neq 0$, and

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & -\beta & 1 \\ 0 & 0 & \alpha \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ k \\ 0 \end{pmatrix}, \quad c = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}.$$

(a) Prove that the system is stabilizable and observable.
(b) Find an observer of order 3 such that the eigenvalues of the closed-loop system are $-1, -2, -3, -4, -5, \alpha$.

3. Let $(A, B) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m}$. Assume that the unreachable modes $\lambda \in \mathbb{C}$ of

$$x_{t+1} = Ax_t + Bu_t$$

are unstable, i.e., satisfy $|\lambda| \geq 1$. Then, for every $C \in \mathbb{R}^{p\times n}$ with $C \neq 0$, there exists an initial state x_0 and an input sequence (u_t) such that $Cx_t \not\to 0$ for $t \to \infty$.
4. Assume that $G(z) = D + C(zI - A)^{-1}B$ is a $p \times m$ proper rational function such that $G(z)u(z)$ is stable for all proper rational functions $u(z) \in \mathbb{F}(z)^m$. Show that $D = 0$ and $C(zI - A)^{-1}B = 0$.
5. Assume that the output sequence $\{y_t\}$ of the linear system

$$x_{t+1} = Ax_t + Bu_t,$$
$$y_t = Cx_t + Du_t$$

satisfies $\lim_{t\to\infty} y_t = 0$ for all initial conditions $x_0 \in \mathbb{F}^n$ and all input sequences $u = (u_t)$. Assume further that all unreachable modes of A are unstable. Show that $C = 0$ and $D = 0$.
6. Let (J, F) be an observable pair, and let (A, B) be reachable. Prove the following result from Fuhrmann and Helmke (2001a): for matrices M and N one has

$$J(zI - F)^{-1}(zM + N)(zI - A)^{-1}B = 0$$

if and only if there exist constant matrices X and Y with $JY = 0$ and $XB = 0$ satisfying

$$zM + N = X(zI - A) - (zI - F)Y.$$

7.7 Notes and References

Probably the first application of modern control theory was that of system stabilization (Chapter 6), and this was done by state feedback. Since the state of a system is hardly ever available, the immediate question arises of how to estimate the state from measurements. This question immediately leads to the fields of optimal filtering and observer theory. Although attempts at state estimation were made earlier, it is generally accepted that the origin of observer theory can be traced to Luenberger (1964). What is surprising is that, over the years, the analysis and synthesis of functional observers, i.e., observer theory, did not attract the appropriate attention from the control community that it so rightly deserves. Moreover, in the system literature, there are several gaps, faulty proofs, and lack of insights that only now are beginning to be filled in. Refer to Trumpf (2013) for details on these gaps in the development of observer theory. In this connection we mention also, for example, Fuhrmann and Helmke (2001) for a fairly complete account of asymptotic observers and to the Ph.D. thesis by Trumpf (2002), which focuses on certain geometric properties that relate to observer theory.

In recent years, the behavioral approach, an approach that avoids the input/output point of view, has been initiated and developed by Willems (1986, 1991) and coworkers. For a study of observers in the behavioral context we refer to the work by Valcher and Willems (1999) and Trumpf, Trentelmann and Willems (2014); see also Fuhrmann (2008), who pointed out how conventional state observer theory fits into the behavioral framework. A full study of the connections between conventional and behavioral observer theories has not yet been undertaken.

The concept of reconstructibility is important for cases dealing with dead-beat observers, a case that will not be addressed in this book. For a treatment of dead-beat obsevers, see Bisiacco, Valcher and Willems (2006) and Fuhrmann and Trumpf (2006). The parameterization results, given by (7.18) when the pair (C,A) is not observable and by (7.21) when it is, relate to partial realizations; they are also reminiscent of the Youla–Kucera parameterization as outlined in Chapter 6. In fact, from the first equation of (7.21) it follows that $W = -D_{21}D_{11}^{-1} - Z_1D_{11}^{-1}$, which shows that $W(z)$ is a solution to a nice partial realization problem induced by $-D_{21}D_{11}^{-1}$. Minimal McMillan degree solutions to (7.17) can be obtained from minimal McMillan degree solutions of the partial realization problem. We refer the reader to Fuhrmann (2008) for a full analysis of the observable case.

The linear equation $Z_K(z) = Z_1(z)Z_C(z) + Z_2(z)$ in proper rational functions $Z_1(z), Z_2(z)$ plays a central role in our approach to functional observers. Since the space of proper rational functions is a valuation ring, and hence a local ring, this task amounts to studying linear matrix equations over a local ring. This fact may be useful in developing solution algorithms for $Z_K(z) = Z_1(z)Z_C(z) + Z_2(z)$.

In Definition 7.16, we made a distinction between asymptotic observers and asymptotic tracking observers. Theorem 7.17 left open the question of how to characterize asymptotic observers. A natural question arises as to how to specify extra conditions such that an asymptotic observer is also tracking. This issue has been addressed by Trumpf (2013) for continuous-time systems. Some preparatory results from Trumpf (2013) appear as Exercises 4–6. In the absence of a full proof, this is stated as a conjecture.

Conjecture 7.33. Let the linear system (7.28) act in the state space $\mathbb{R}^n$. Assuming that all unreachable modes of the system are unstable, the system, defined in the state space $\mathbb{R}^q$ by

$$\begin{aligned}\xi_{t+1} &= F\xi_t + Gy_t + Hu_t,\\ \zeta_t &= J\xi_t + Ey_t,\end{aligned}$$

is an observable, asymptotic observer for K if and only if F has all its eigenvalues in the open unit disc, (J,F) is observable, and there exists a matrix $Z \in \mathbb{R}^{q\times n}$ such that

$$ZA - FZ = GC, \quad H = ZB, \quad K = JZ + EC.$$

In particular, under such an assumption, asymptotic observers are automatically asymptotic tracking observers.

Under the assumption that the pair (A,B) is reachable, this was proved in Fuhrmann and Helmke (2001a). See also Fernando, Jennings and Trinh (2011) for a claim toward Conjecture 7.33. For a characterization of asymptotic observers in the behavioral framework, see Trumpf, Trentelmann and Willems (2014).