2

An Introduction to State-Space and Input–Output Descriptions of Systems

2.1 Introduction

State-space representations provide detailed descriptions of the internal behavior of a system, whereas input–output descriptions of systems emphasize external behavior and a system's interaction with this behavior.

In this chapter we address the state-space description of systems, which is an internal description of systems, and the input–output description of systems, also called the external description of systems. We will address continuous-time systems described by ordinary differential equations and discrete-time systems described by ordinary difference equations. We will emphasize linear systems. For such systems, the input–output descriptions involve the convolution integral for the continuous-time case and the convolution sum for the discrete-time case.

This chapter is organized into three parts. In the first of these (Section 2.2), we develop the state-space description of continuous-time systems, whereas in the second part (Section 2.3), we present the state-space representation of discrete-time systems. In the third part (Section 2.4), we address the input–output description of both continuous-time and discrete-time systems. Required background material for this chapter includes certain essentials in ordinary differential equations and linear algebra. This material can be found in Chapter 1 and the appendix, respectively.

2.2 State-Space Description of Continuous-Time Systems

Let us consider once more systems described by equations of the form

$$\dot{x} = f(t, x, u), \tag{2.1a}$$

$$y = g(t, x, u), \tag{2.1b}$$

where $x \in R^n, y \in R^p, u \in R^m, f : R \times R^n \times R^m \to R^n$, and $g : R \times R^n \times R^m \to R^p$. Here t denotes time and u and y denote system *input* and system *output*, respectively. Equation (2.1a) is called the *state equation*, (2.1b) is called the *output equation*, and (2.1a) and (2.1b) constitute the *state-space description* of continuous-time finite-dimensional systems.

The system input may be a function of t only (i.e., $u : R \to R^m$), or as in the case of *feedback control systems*, it may be a function of t and x (i.e., $u : R \times R^n \to R^m$). In either case, for a *given* (i.e., specified) u, we let $f(t, x, u) = F(t, x)$ and rewrite (2.1a) as

$$\dot{x} = F(t, x). \tag{2.2}$$

Now according to Theorems 1.13 and 1.14, if $F \in C(R \times R^n, R^n)$ and if for any compact subinterval $J_0 \subset R$ there is a constant L_{J_0} such that $\| F(t, x) - F(t, \tilde{x}) \| \leq L_{J_0} \| x - \tilde{x} \|$ for all $t \in J_0$ and for all $x, \tilde{x} \in R^n$, then the following statements are true:

1. For any $(t_0, x_0) \in R \times R^n$, (2.2) has a unique solution $\phi(t, t_0, x_0)$ satisfying $\phi(t_0, t_0, x_0) = x_0$ that exists for all $t \in R$.
2. The solution ϕ is continuous in t, t_0, and x_0.
3. If F depends continuously on parameters (say, $\lambda \in R^l$) and if x_0 depends continuously on λ, the solution ϕ is continuous in λ as well.

Thus, if the above conditions are satisfied, then for a given t_0, x_0, and u, (2.1a) will have a unique solution that exists for $t \in R$. Therefore, as already discussed in Section 1.8, $\phi(t, t_0, x_0)$ characterizes the *state* of the system at time t. Moreover, under these conditions, the system will have a unique *response* for $t \in R$, determined by (2.1b). We usually assume that $g \in C(R \times R^n \times R^m, R^p)$ or that $g \in C^1(R \times R^n \times R^m, R^p)$.

An important special case of (2.1) is systems described by linear time-varying equations of the form

$$\dot{x} = A(t)x + B(t)u, \tag{2.3a}$$
$$y = C(t)x + D(t)u, \tag{2.3b}$$

where $A \in C(R, R^{n\times n}), B \in C(R, R^{n\times m}), C \in C(R, R^{p\times n})$, and $D \in C(R, R^{p\times m})$. Such equations may arise in the modeling process of a physical system, or they may be a consequence of a linearization process, as discussed in Section 1.6.

By applying the results of Section 1.7, we see that for every initial condition $x(t_0) = x_0$ and for every given input $u : R \to R^m$, system (2.3a) possesses a unique solution that exists for all $t \in R$ and that is continuous in (t, t_0, x_0). Moreover, if A and B depend continuously on parameters, say, $\lambda \in R^l$, then the solutions will be continuous in the parameters as well. Indeed, in accordance with (1.87), this solution is given by

$$\phi(t, t_0, x_0) = \Phi(t, t_0)x_0 + \int_{t_0}^{t} \Phi(t, s)B(s)u(s)ds, \tag{2.4}$$

where $\Phi(t, t_0)$ denotes the state transition matrix of the system of equations

$$\dot{x} = A(t)x. \tag{2.5}$$

By using (2.3b) and (2.4) we obtain the *system response* as

$$y(t) = C(t)\Phi(t, t_0)x_0 + C(t)\int_{t_0}^{t} \Phi(t, s)B(s)u(s)ds + D(t)u(t). \tag{2.6}$$

When in (2.3), $A(t) \equiv A, B(t) \equiv B, C(t) \equiv C$, and $D(t) \equiv D$, we have the important linear time-invariant case given by

$$\dot{x} = Ax + Bu, \tag{2.7a}$$

$$y = Cx + Du. \tag{2.7b}$$

In accordance with (1.84), (1.85), (1.87), and (2.4), the solution of (2.7a) is given by

$$\phi(t, t_0, x_0) = e^{A(t-t_0)}x_0 + \int_{t_0}^{t} e^{A(t-s)}Bu(s)ds \tag{2.8}$$

and the response of the system is given by

$$y(t) = Ce^{A(t-t_0)}x_0 + C\int_{t_0}^{t} e^{A(t-s)}Bu(s)ds + Du(t). \tag{2.9}$$

Linearity

We have referred to systems described by the linear equations (2.3) [resp., (2.7)] as *linear systems.* In the following discussion, we establish precisely in what sense this linearity is to be understood. To this end, for (2.3) we first let y_1 and y_2 denote system outputs that correspond to system inputs given by u_1 and u_2, respectively, *under the condition that* $x_0 = 0$. By invoking (2.6), it is clear that the system output corresponding to the system input $u = \alpha_1 u_1 + \alpha_2 u_2$, where α_1 and α_2 are real scalars, is given by $y = \alpha_1 y_1 + \alpha_2 y_2$; i.e.,

$$\begin{aligned} y(t) &= C(t)\int_{t_0}^{t} \Phi(t, s)B(s)[\alpha_1 u_1(s) + \alpha_2 u_2(s)]ds + D(t)[\alpha_1 u_1(t) + \alpha_2 u_2(t)] \\ &= \alpha_1 C(t)\int_{t_0}^{t} \Phi(t, s)B(s)u_1(s)ds + \alpha_2 C(t)\int_{t_0}^{t} \Phi(t, s)B(s)u_2(s)ds \\ &\quad + \alpha_1 D(t)u_1(t) + \alpha_2 D(t)u_2(t) \\ &= \alpha_1 y_1(t) + \alpha_2 y_2(t). \end{aligned} \tag{2.10}$$

Next, for (2.3) we let y_1 and y_2 denote system outputs that correspond to initial conditions $x_0^{(1)}$ and $x_0^{(2)}$, respectively, *under the condition that* $u(t) = 0$

for all $t \in R$. Again, by invoking (2.6), it is clear that the system output corresponding to the initial condition $x_0 = \alpha_1 x_0^{(1)} + \alpha_2 x_0^{(2)}$, where α_1 and α_2 are real scalars, is given by $y = \alpha_1 y_1 + \alpha_2 y_2$; i.e.,

$$\begin{aligned} y(t) &= C(t)\Phi(t,t_0)[\alpha_1 x_0^{(1)} + \alpha_2 x_0^{(2)}] \\ &= \alpha_1 C(t)\Phi(t,t_0)x_0^{(1)} + \alpha_2 C(t)\Phi(t,t_0)x_0^{(2)} \\ &= \alpha_1 y_1(t) + \alpha_2 y_2(t). \end{aligned} \tag{2.11}$$

Equations (2.10) and (2.11) show that for systems described by the linear equations (2.3) [and, hence, by (2.7)], a *superposition principle* holds in terms of the input u and the corresponding output y of the system under the assumption of zero initial conditions, and in terms of the initial conditions x_0 and the corresponding output y under the assumption of zero input. It is important to note, however, that such a superposition principle will in general not hold under conditions that combine nontrivial inputs and nontrivial initial conditions. For example, with $x_0 \neq 0$ given, and with inputs u_1 and u_2 resulting in corresponding outputs y_1 and y_2 in (2.3), it does not follow that the input $\alpha_1 u_1 + \alpha_2 u_2$ will result in an output $\alpha_1 y_1 + \alpha_2 y_2$.

2.3 State-Space Description of Discrete-Time Systems

State-Space Representation

The state-space description of discrete-time finite-dimensional dynamical systems is given by equations of the form

$$x_i(k+1) = f_i(k, x_1(k), \dots, x_n(k), u_1(k), \dots, u_m(k)) \quad i = 1, \dots, n, \tag{2.12a}$$

$$y_i(k) = g_i(k, x_1(k), \dots, x_n(k), u_1(k), \dots, u_m(k)) \quad i = 1, \dots, p, \tag{2.12b}$$

for $k = k_0, k_0+1, \dots$, where k_0 is an integer. (In the following discussion, we let Z denote the set of integers and we let Z^+ denote the set of nonnegative integers.) Letting $x(k)^T = (x_1(k), \dots, x_n(k)), f(\cdot)^T = (f_1(\cdot), \dots, f_n(\cdot)), u(k)^T = (u_1(k), \dots, u_m(k)), y(k)^T = (y_1(k), \dots, y_p(k))$, and $g(\cdot)^T = (g_1(\cdot), \dots, g_m(\cdot))$, we can rewrite (2.12) more compactly as

$$x(k+1) = f(k, x(k), u(k)), \tag{2.13a}$$

$$y(k) = g(k, x(k), u(k)). \tag{2.13b}$$

Throughout this section we will assume that $f : Z \times R^n \times R^m \to R^n$ and $g : Z \times R^n \times R^m \to R^p$.

Since f is a function, for given $k_0, x(k_0) = x_0$, and for given $u(k), k = k_0, k_0+1, \dots$, (2.13a) possesses a unique solution $x(k)$ that exists for all $k = k_0, k_0+1, \dots$. Furthermore, under these conditions, $y(k)$ is uniquely defined for $k = k_0, k_0+1, \dots$.

As in the case of continuous-time finite-dimensional systems [see (2.1)], k_0 denotes *initial time*, k denotes *time*, $u(k)$ denotes the system *input* (evaluated at time k), $y(k)$ denotes the system *output* or system *response* (evaluated at time k), $x(k)$ characterizes the *state* (evaluated at time k), $x_i(k)$, $i = 1, \dots, n$, denote the *state variables*, (2.13a) is called the *state equation*, and (2.13b) is called the *output equation*.

A moment's reflection should make it clear that in the case of discrete-time finite-dimensional dynamical systems described by (2.13), questions concerning existence, uniqueness, and continuation of solutions are not an issue, as was the case in continuous-time systems. Furthermore, continuity with respect to initial data $x(k_0) = x_0$, or with respect to system parameters, is not an issue either, provided that $f(\cdot)$ and $g(\cdot)$ have appropriate continuity properties.

In the case of continuous-time systems described by ordinary differential equations [see (2.1)], we allow time t to evolve "forward" and "backward." Note, however, that in the case of discrete-time systems described by (2.13), we restrict the evolution of time k in the forward direction to ensure uniqueness of solutions. (We will revisit this issue in more detail in Chapter 3.)

Special important cases of (2.13) are *linear time-varying systems* given by

$$x(k+1) = A(k)x(k) + B(k)u(k), \tag{2.14a}$$

$$y(k) = C(k)x(k) + D(k)u(k), \tag{2.14b}$$

where $A : Z \to R^{n\times n}$, $B : Z \to R^{n\times m}$, $C : Z \to R^{p\times n}$, and $D : Z \to R^{p\times m}$. When $A(k) \equiv A, B(k) \equiv B, C(k) \equiv C$, and $D(k) \equiv D$, we have *linear time-invariant systems* given by

$$x(k+1) = Ax(k) + Bu(k), \tag{2.15a}$$

$$y(k) = Cx(k) + Du(k). \tag{2.15b}$$

As in the case of continuous-time finite-dimensional dynamical systems, many qualitative properties of discrete-time finite-dimensional systems can be studied in terms of *initial-value problems* given by

$$x(k+1) = f(k, x(k)), \quad x(k_0) = x_0, \tag{2.16}$$

where $x \in R^n$, $f : Z \times R^n \to R^n, k_0 \in Z$, and $k = k_0, k_0 + 1, \cdots$. We call the equation

$$x(k+1) = f(k, x(k)), \tag{2.17}$$

a *system of first-order ordinary difference equations*. Special important cases of (2.17) include *autonomous systems* described by

$$x(k+1) = f(x(k)), \tag{2.18}$$

periodic systems given by

$$x(k+1) = f(k, x(k)) = f(k+K, x(k)) \tag{2.19}$$

for fixed $K \in Z^+$ and for all $k \in Z$, *linear homogeneous systems* given by

$$x(k+1) = A(k)x(k), \tag{2.20}$$

linear periodic systems characterized by

$$x(k+1) = A(k)x(k) = A(k+K)x(k) \tag{2.21}$$

for fixed $K \in Z^+$ and for all $k \in Z$, *linear nonhomogeneous systems*

$$x(k+1) = A(k)x(k) + g(k), \tag{2.22}$$

and *linear, autonomous, homogeneous systems* characterized by

$$x(k+1) = Ax(k). \tag{2.23}$$

In these equations all symbols used are defined in the obvious way by making reference to the corresponding systems of ordinary differential equations (see Subsection 1.3.2).

Difference Equations of Order n

Thus far we have addressed systems of first-order difference equations. As in the continuous-time case, it is also possible to characterize initial-value problems by nth-order ordinary difference equations, say,

$$y(k+n) = h(k, y(k), y(k+1), \ldots, y(k+n-1)), \tag{2.24}$$

where $h : Z \times R^n \to R, n \in Z^+, k = k_0, k_0+1, \ldots$. By specifying an *initial time* $k_0 \in Z$ and by specifying $y(k_0), y(k_0+1), \ldots, y(k_0+n-1)$, we again have an initial-value problem given by

$$\begin{aligned} y(k+n) &= h(k, y(k), y(k+1), \ldots, y(k+n-1)), \\ y(k_0) &= x_{10}, \ldots, y(k_0+n-1) = x_{n0}. \end{aligned} \tag{2.25}$$

We call (2.24) an *nth-order ordinary difference equation*, and we note once more that in the case of initial-value problems described by such equations, there are no difficult issues involving the existence, uniqueness, and continuation of solutions.

We can reduce the study of (2.25) to the study of initial-value problems determined by systems of first-order ordinary difference equations. To accomplish this, we let in (2.25) $y(k) = x_1(k), y(k+1) = x_2(k), \ldots, y(k+n-1) = x_n(k)$. We now obtain the system of first-order ordinary difference equations

$$\begin{aligned} x_1(k+1) &= x_2(k), \\ &\cdots \\ x_{n-1}(k+1) &= x_n(k), \\ x_n(k+1) &= h(k, x_1(k), \ldots, x_n(k)). \end{aligned} \tag{2.26}$$

Equations (2.26), together with the initial data $x_0^T = (x_{10}, \ldots, x_{n0})$, are equivalent to the initial-value problem (2.25) in the sense that these two problems will generate identical solutions [and in the sense that the transformation of (2.25) into (2.26) can be reversed unambiguously and uniquely].

As in the case of systems of first-order ordinary difference equations, we can point to several important special cases of nth-order ordinary difference equations, including equations of the form

$$y(k+n) + a_{n-1}(k)y(k+n-1) + \cdots + a_1(k)y(k+1) + a_0(k)y(k) = g(k), \quad (2.27)$$

$$y(k+n) + a_{n-1}(k)y(k+n-1) + \cdots + a_1(k)y(k+1) + a_0(k)y(k) = 0, \quad (2.28)$$

and

$$y(k+n) + a_{n-1}y(k+n-1) + \cdots + a_1 y(k+1) + a_0 y(k) = 0. \quad (2.29)$$

We call (2.27) a *linear nonhomogeneous ordinary difference equation of order n*, we call (2.28) a *linear homogeneous ordinary difference equation of order n*, and we call (2.29) a *linear, autonomous, homogeneous ordinary difference equation of order n*. As in the case of systems of first-order ordinary difference equations, we can define *periodic* and *linear periodic ordinary difference equations of order n* in the obvious way.

Solutions of State Equations

Returning now to linear homogeneous systems

$$x(k+1) = A(k)x(k), \quad (2.30)$$

we observe that

$$\begin{aligned} x(k+2) &= A(k+1)x(k+1) = A(k+1)A(k)x(k) \\ &\cdots \\ x(n) &= A(n-1)A(n-2)\cdots A(k+1)A(k)x(k) \\ &= \prod_{j=k}^{n-1} A(j)x(k); \end{aligned}$$

i.e., the state of the system at time n is related to the state at time k by means of the $n \times n$ matrix $\prod_{j=k}^{n-1} A(j)$ (as can easily be proved by induction). This suggests that the *state transition matrix* for (2.30) is given by

$$\Phi(n,k) = \prod_{j=k}^{n-1} A(j), \quad n > k, \quad (2.31)$$

and that

$$\Phi(k,k) = I. \quad (2.32)$$

As in the continuous-time case, the solution to the initial-value problem

$$\begin{aligned} x(k+1) &= A(k)x(k) \\ x(k_0) &= x_{k_0}, \quad k_0 \in Z, \end{aligned} \tag{2.33}$$

is now given by

$$x(n) = \Phi(n, k_0)x_{k_0} = \prod_{j=k_0}^{n-1} A(j)x_{k_0}, n > k_0. \tag{2.34}$$

Continuing, let us next consider initial-value problems determined by linear nonhomogeneous systems (2.22),

$$\begin{aligned} x(k+1) &= A(k)x(k) + g(k), \\ x(k_0) &= x_{k_0}. \end{aligned} \tag{2.35}$$

Then

$$\begin{aligned} x(k_0+1) &= A(k_0)x(k_0) + g(k_0), \\ x(k_0+2) &= A(k_0+1)x(k_0+1) + g(k_0+1) \\ &= A(k_0+1)A(k_0)x(k_0) + A(k_0+1)g(k_0) + g(k_0+1), \\ x(k_0+3) &= A(k_0+2)x(k_0+2) + g(k_0+2) \\ &= A(k_0+2)A(k_0+1)A(k_0)x(k_0) + A(k_0+2)A(k_0+1)g(k_0) \\ &\quad + A(k_0+2)g(k_0+1) + g(k_0+2) \\ &= \Phi(k_0+3, k_0)x_{k_0} + \Phi(k_0+3, k_0+1)g(k_0) \\ &\quad + \Phi(k_0+3, k_0+2)g(k_0+1) + \Phi(k_0+3, k_0+3)g(k_0+2), \end{aligned}$$

and so forth. For $k \geq k_0 + 1$, we easily obtain the expression for the solution of (2.35) as

$$x(k) = \Phi(k, k_0)x_{k_0} + \sum_{j=k_0}^{k-1} \Phi(k, j+1)g(j). \tag{2.36}$$

In the time-invariant case

$$\begin{aligned} x(k+1) &= Ax(k) + g(k), \\ x(k_0) &= x_{k_0}, \end{aligned} \tag{2.37}$$

the solution is again given by (2.36) where now the state transition matrix

$$\Phi(k, k_0) = A^{k-k_0}, \quad k \geq k_0, \tag{2.38}$$

in view of (2.31) and (2.32). The solution of (2.37) is then

$$x(k) = A^{k-k_0}x_{k_0} + \sum_{j=k_0}^{k-1} A^{k-(j+1)}g(j), \quad k > k_0. \tag{2.39}$$

We note that when $x_{k_0} = 0$, (2.36) reduces to

$$x_p(k) = \sum_{j=k_0}^{k-1} \Phi(k, j+1) g(j), \tag{2.40}$$

and when $x_{k_0} \neq 0$ but $g(k) \equiv 0$, then (2.36) reduces to

$$x_h(k) = \Phi(k, k_0) x_{k_0}. \tag{2.41}$$

Therefore, the *total solution* of (2.35) consists of the sum of its *particular solution*, $x_p(k)$, and its *homogeneous solution*, $x_h(k)$.

System Response

Finally, we observe that in view of (2.14b) and (2.36), the *system response* of the system (2.14), is of the form

$$\begin{aligned} y(k) = C(k)\Phi(k, k_0)x_{k_0} + C(k) \sum_{j=k_0}^{k-1} \Phi(k, j+1) B(j) u(j) \\ + D(k)u(k), \quad k > k_0, \end{aligned} \tag{2.42}$$

and

$$y(k_0) = C(k_0)x_{k_0} + D(k_0)u(k_0). \tag{2.43}$$

In the time-invariant case, in view of (2.39), the system response of the system (2.15) is

$$y(k) = CA^{k-k_0}x_{k_0} + C \sum_{j=k_0}^{k-1} A^{k-(j+1)} B(j) u(j) + Du(k), \quad k > k_0, \tag{2.44}$$

and

$$y(k_0) = Cx_{k_0} + Du(k_0). \tag{2.45}$$

Discrete-time systems, as discussed above, arise in several ways, including the *numerical solution* of ordinary differential equations (see, e.g., our discussion in Exercise 1.4 of *Euler's method*); the representation of *sampled-data systems* at discrete points in time (which will be discussed in further detail in Chapter 3); in the modeling process of systems that are defined only at discrete points in time (e.g., digital computer systems); and so forth.

As a specific example of a discrete-time system we consider a *second-order section digital filter* in *direct form*,

$$\begin{aligned} x_1(k+1) &= x_2(k), \\ x_2(k+1) &= ax_1(k) + bx_2(k) + u(k), \end{aligned} \tag{2.46a}$$

$$y(k) = x_1(k), \tag{2.46b}$$

$k \in Z^+$, where $x_1(k)$ and $x_2(k)$ denote the state variables, $u(k)$ denotes the input, and $y(k)$ denotes the output of the digital filter. We depict system (2.46) in block diagram form in Figure 2.1.

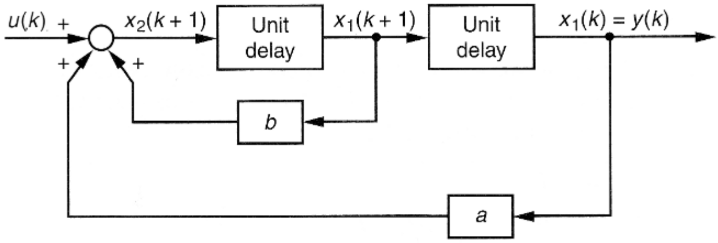

Figure 2.1. Second-order section digital filter in direct form

2.4 Input–Output Description of Systems

This section consists of four subsections. First we consider rather general aspects of the input–output description of systems. Because of their simplicity, we address the characterization of linear discrete-time systems next. In the third subsection we provide a foundation for the impulse response of linear continuous-time systems. Finally, we address the external description of linear continuous-time systems.

2.4.1 External Description of Systems: General Considerations

The state-space representation of systems presupposes knowledge of the *internal structure* of the system. When this structure is unknown, it may still be possible to arrive at a system description—an *external description*—that relates system inputs to system outputs. In linear system theory, a great deal of attention is given to relating the internal description of systems (the state representation) to the external description (the input–output description).

In the present context, we view *system inputs* and *system outputs* as elements of two real vector spaces U and Y, respectively, and we view a system as being represented by an operator T that relates elements of U to elements of Y. For $u \in U$ and $y \in Y$ we will assume that $u : R \rightarrow R^m$ and $y : R \rightarrow R^p$ in the case of *continuous-time systems*, and that $u : Z \rightarrow R^m$ and $y : Z \rightarrow R^p$ in the case of *discrete-time systems*. If $m = p = 1$, we speak of a *single-input/single-output (SISO) system*. Systems for which $m > 1$, $p > 1$, are called *multi-input/multi-output (MIMO) systems*. For continuous-time systems we define vector addition (on U) and multiplication of vectors by scalars (on U) as

$$(u_1 + u_2)(t) = u_1(t) + u_2(t) \tag{2.47}$$

and

$$(\alpha u)(t) = \alpha u(t) \tag{2.48}$$

for all $u_1, u_2 \in U, \alpha \in R$, and $t \in R$. We similarly define vector addition and multiplication of vectors by scalars on Y. Furthermore, for discrete-time systems we define these operations on U and Y analogously. In this case the elements of U and Y are real sequences that we denote, e.g., by $u = \{u_k\}$ or $u = \{u(k)\}$. (It is easily verified that under these rather general conditions, U and Y satisfy all the axioms of a vector space, both for the continuous-time case and the discrete-time case.) In the continuous-time case as well as in the discrete-time case the system is represented by $T : U \to Y$, and we write

$$y = T(u). \tag{2.49}$$

In the subsequent development, we will impose restrictions on the vector spaces U, Y, and on the operator T, as needed.

Linearity. If T is a linear operator, the system is called a *linear system.* In this case we have

$$\begin{aligned} y &= T(\alpha_1 u_1 + \alpha_2 u_2) \\ &= \alpha_1 T(u_1) + \alpha_2 T(u_2) \\ &= \alpha_1 y_1 + \alpha_2 y_2 \end{aligned} \tag{2.50}$$

for all $\alpha_1, \alpha_2 \in R$ and $u_1, u_2 \in U$ where $y_i = T(u_i) \in Y$, $i = 1, 2$, and $y \in Y$. Equation (2.50) represents the well-known *principle of superposition* of linear systems.

With or Without Memory. We say that a system is *memoryless*, or *without memory*, if its output for each value of the independent variable (t or k) is dependent only on the input evaluated at the same value of the independent variable [e.g., $y(t_1)$ depends only on $u(t_1)$ and $y(k_1)$ depends only on $u(k_1)$]. An example of such a system is the resistor circuit shown in Figure 2.2, where the current $i(t) = u(t)$ denotes the system input at time t and the voltage across the resistor, $v(t) = Ri(t) = y(t)$, denotes the system output at time t.

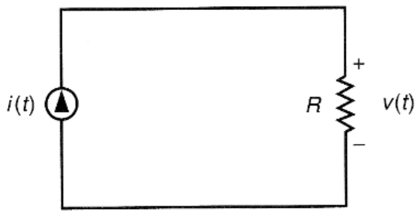

Figure 2.2. Resistor circuit

A system that is not memoryless is said to have memory. An example of a continuous-time *system with memory* is the capacitor circuit shown in Figure 2.3, where the current $i(t) = u(t)$ represents the system input at time t and the voltage across the capacitor,

$$y(t) = v(t) = \frac{1}{C}\int_{-\infty}^{t} i(\tau)d\tau,$$

denotes the system output at time t. Another example of a continuous-time system with memory is described by the scalar equation

$$y(t) = u(t-1), \quad t \in R,$$

and an example of a discrete-time system with memory is characterized by the scalar equation

$$y(n) = \sum_{k=-\infty}^{n} x(k), \quad n, k \in Z.$$

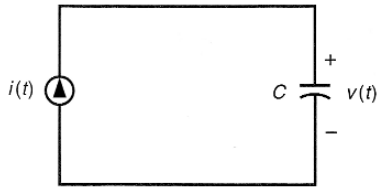

Figure 2.3. Capacitor circuit

Causality. A system is said to be *causal* if its output at any time, say t_1 (or k_1), depends only on values of the input evaluated for $t \le t_1$ (for $k \le k_1$). Thus, $y(t_1)$ depends only on $u(t), t \le t_1$ [or $y(k_1)$ depends only on $u(k), k \le k_1$]. Such a system is referred to as being *nonanticipative* since the system output does not anticipate future values of the input.

To make the above concept a bit more precise, we define the function $u_\tau : R \to R^m$ for $u \in U$ by

$$u_\tau(t) = \begin{cases} u(t), & t \le \tau, \\ 0, & t > \tau, \end{cases}$$

and we similarly define the function $y_\tau : R \to R^p$ for $y \in Y$. A system that is represented by the mapping $y = T(u)$ is said to be *causal* if and only if

$$(T(u))_\tau = (T(u_\tau))_\tau \quad \text{for all } \tau \in R, \text{ for all } u \in U.$$

Equivalently, this system is causal if and only if for $u, v \in U$ and $u_\tau = v_\tau$ it is true that

$$(T(u))_\tau = (T(v))_\tau \quad \text{for all } \tau \in R.$$

For example, the discrete-time system described by the scalar equation

$$y(n) = u(n) - u(n+1), \quad n \in Z,$$

is *not causal*. Neither is the continuous-time system characterized by the scalar equation

$$y(t) = x(t+1), \quad t \in R.$$

It should be pointed out that systems that are not causal are by no means useless. For example, causality is *not* of fundamental importance in image-processing applications where the independent variable is not time. Even when time is the independent variable, noncausal systems may play an important role. For example, in the processing of data that have been recorded (such as speech, meteorological data, demographic data, and stock market fluctuations), one is not constrained to processing the data causally. An example of this would be the smoothing of data over a time interval, say, by means of the system

$$y(n) = \frac{1}{2M+1} \sum_{k=-M}^{M} u(n-k).$$

Time-Invariance. A system is said to be *time-invariant* if a time shift in the input signal causes a corresponding time shift in the output signal. To make this concept more precise, for fixed $\alpha \in R$, we introduce the *shift operator* $Q_\alpha : U \to U$ as

$$Q_\alpha u(t) = u(t-\alpha), \quad u \in U, t \in R.$$

A system that is represented by the mapping $y = T(u)$ is said to be *time-invariant* if and only if

$$TQ_\alpha(u) = Q_\alpha(T(u)) = Q_\alpha(y)$$

for any $\alpha \in R$ and any $u \in U$. If a system is not time-invariant, it is said to be *time-varying*.

For example, a system described by the relation

$$y(t) = \cos u(t)$$

is time-invariant. To see this, consider the inputs $u_1(t)$ and $u_2(t) = u_1(t-t_0)$. Then

$$y_1(t) = \cos u_1(t), \quad y_2(t) = \cos u_2(t) = \cos u_1(t-t_0)$$

and

$$y_1(t-t_0) = \cos u_1(t-t_0) = y_2(t).$$

As a second example, consider a system described by the relation

$$y(n) = nu(n)$$

and consider two inputs $u_1(n)$ and $u_2(n) = u_1(n-n_0)$. Then

$$y_1(n) = nu_1(n) \quad \text{and} \quad y_2(n) = nu_2(n) = nu_1(n - n_0).$$

However, if we shift the output $y_1(n)$ by n_0, we obtain

$$y_1(n - n_0) = (n - n_0)u_1(n - n_0) \neq y_2(n).$$

Therefore, this system is not time-invariant.

2.4.2 Linear Discrete-Time Systems

In this subsection we investigate the representation of linear discrete-time systems. We begin our discussion by considering SISO systems.

In the following, we employ the *discrete-time impulse* (or *unit pulse* or *unit sample*), which is defined as

$$\delta(n) = \begin{cases} 0, & n \neq 0, n \in Z, \\ 1, & n = 0. \end{cases} \tag{2.51}$$

Note that if $\{p(n)\}$ denotes the *unit step sequence*, i.e.,

$$p(n) = \begin{cases} 1, & n \geq 0, n \in Z, \\ 0, & n < 0, n \in Z, \end{cases} \tag{2.52}$$

then

$$\delta(n) = p(n) - p(n-1)$$

and

$$p(n) = \begin{cases} \sum_{k=0}^{\infty} \delta(n-k), & n \geq 0, \\ 0, & n < 0. \end{cases} \tag{2.53}$$

Furthermore, note that an arbitrary sequence $\{x(n)\}$ can be expressed as

$$x(n) = \sum_{k=-\infty}^{\infty} x(k)\delta(n-k). \tag{2.54}$$

We can easily show that a transformation $T : U \to Y$ determined by the equation

$$y(n) = \sum_{k=-\infty}^{\infty} h(n,k)u(k), \tag{2.55}$$

where $y \triangleq \{y(k)\} \in Y$, $u \triangleq \{u(k)\} \in U$, and $h : Z \times Z \to R$, is a linear transformation. Also, we note that for (2.55) to make any sense, we need to impose restrictions on $\{h(n,k)\}$ and $\{u(k)\}$. For example, if for every fixed n, $\{h(n,k)\} \in l_2$ and $\{u(k)\} \in l_2 = U$, then it follows from the Hölder Inequality (resp., Schwarz Inequality), see Section A.7, that (2.55) is well defined. There are of course other conditions that one might want to impose on (2.55).

For example, if for every fixed n, $\sum_{k=-\infty}^{\infty} |h(n,k)| < \infty$ (i.e., for every fixed $n, \{h(n,k)\} \in l_1$) and if $\sup_{k\in Z} |u(k)| < \infty$ (i.e., $\{u(k)\} \in l_\infty$), then (2.55) is also well defined.

We shall now elaborate on the suitability of (2.55) to represent linear discrete-time systems. To this end, we will agree once and for all that, in the ensuing discussion, all assumptions on $\{h(n,k)\}$ and $\{u(k)\}$ are satisfied that ensure that (2.55) is well defined.

We will view $y \in Y$ and $u \in U$ as system outputs and system inputs, respectively, and we will let $T : U \to Y$ denote a linear transformation that relates u to y. We first consider the case when $u(k) = 0$ for $k < k_0$, k, $k_0 \in Z$. Also, we assume that for $k > n \geq k_0$, the inputs $u(k)$ do not contribute to the system output at time n (i.e., the system is *causal*). Under these assumptions, and in view of the linearity of T, and by invoking the representation of signals by (2.54), we obtain for $y = \{y(n)\}$, $n \in Z$, the expression $y(n) = T(\sum_{k=-\infty}^{\infty} u(k)\delta(n-k)) = T(\sum_{k=k_0}^{n} u(k)\delta(n-k)) = \sum_{k=k_0}^{n} u(k)T(\delta(n-k)) = \sum_{k=k_0}^{n} h(n,k)u(k)$, $n \geq k_0$, and $y(n) = 0$, $n < k_0$, where $T(\delta(n-k)) \triangleq (T\delta)(n-k) \triangleq h(n,k)$ represents the response of T to a unit pulse (resp., discrete-time impulse or unit sample) occurring at $n = k$.

When the assumptions in the preceding discussion are no longer valid, then a different argument than the one given above needs to be used to arrive at the system representation. Indeed, for *infinite sums*, the interchanging of the order of the summation operation $\sum$ with the linear transformation T is no longer valid. We refer the reader to a paper by I. W. Sandberg ("A Representation Theorem for Linear Systems," *IEEE Transactions on Circuits and Systems—I*, Vol. 45, No. 5, pp. 578–580, May 1998) for a derivation of the representation of general linear discrete-time systems. In that paper it is shown that an extra term needs to be added to the right-hand side of equation (2.55), even in the representation of *general*, linear, time-invariant, causal, discrete-time systems. [In the proof, the Hahn–Banach Theorem (which is concerned with the extension of bounded linear functionals) is employed and the extra required term is given by $\lim_{l\to\infty} T(\sum_{k=-\infty}^{-c_l-1} u(k)\delta(n-k) + \sum_{k=c_l+1}^{\infty} u(k)\delta(n-k)]$ with $c_l \to \infty$ as $l \to \infty$. For a statement and proof of the Hahn–Banach Theorem, refer, e.g., to A. N. Michel and C. J. Herget, *Applied Algebra and Functional Analysis*, Dover, New York, 1993, pp. 367–370.) In that paper it is also pointed out, however, that cases with such extra *nonzero* terms are not necessarily of importance in applications. In particular, if inputs and outputs are defined (to be nonzero) on just the non-negative integers, then for causal systems no additional term is needed (or more specifically, the extra term is zero), as seen in our earlier argument. In any event, *throughout this book we will concern ourselves with linear discrete-time systems that can be represented by equation (2.55)* for the single-input/single-output case (and appropriate generalizations for multi-input/multi-output cases).

Next, suppose that T represents a time-invariant system. This means that if $\{h(n,0)\}$ is the response to $\{\delta(n)\}$, then by time invariance, the response to $\{\delta(n-k)\}$ is simply $\{h(n-k,0)\}$. By a slight abuse of notation, we let $h(n-k,0) \triangleq h(n-k)$. Then (2.55) assumes the form

$$y(n) = \sum_{k=-\infty}^{\infty} u(k)h(n-k). \tag{2.56}$$

Expression (2.56) is called a *convolution sum* and is written more compactly as

$$y(n) = u(n) * h(n).$$

Now by a substitution of variables, we obtain for (2.56) the alternative expression

$$y(n) = \sum_{k=-\infty}^{\infty} h(k)u(n-k),$$

and therefore, we have

$$y(n) = u(n) * h(n) = h(n) * u(n);$$

i.e., the convolution operation $*$ commutes.

As a specific example, consider a linear, time-invariant, discrete-time system with unit impulse response given by

$$h(n) = \left\{ \begin{matrix} a^n, \, n \geq 0 \\ 0, \quad n < 0 \end{matrix} \right\} = a^n p(n), \quad 0 < a < 1,$$

where $p(n)$ is the unit step sequence given in (2.52). It is an easy matter to show that the response of this system to an input given by

$$u(n) = p(n) - p(n-N)$$

is

$$\begin{aligned} y(n) &= 0, n < 0, \\ y(n) &= \sum_{k=0}^{n} a^{n-k} = a^n \frac{1 - 1^{-(n+1)}}{1 - a^{-1}} = \frac{1 - a^{n+1}}{1 - a}, \quad 0 \leq n < N, \end{aligned}$$

and

$$y(n) = \sum_{k=0}^{N-1} a^{n-k} = a^n \frac{1 - a^{-N}}{1 - a^{-1}} = \frac{a^{n-N+1} - a^{n+1}}{1 - a}, \quad N \leq n.$$

Proceeding, with reference to (2.55) we note that $h(n,k)$ represents the system output at time n due to a δ-function input applied at time k. Now if system (2.55) is *causal*, then its output will be identically zero before an input is applied. Hence, a *linear system (2.55) is causal if and only if*

$$h(n,k) = 0 \quad \text{for all} \quad n < k.$$

Therefore, when the system (2.55) is causal, we have in fact

$$y(n) = \sum_{k=-\infty}^{n} h(n,k)u(k). \tag{2.57a}$$

We can rewrite (2.57a) as

$$\begin{aligned} y(n) &= \sum_{k=-\infty}^{k_0-1} h(n,k)u(k) + \sum_{k=k_0}^{n} h(n,k)u(k) \\ &\triangleq y(k_0-1) + \sum_{k=k_0}^{n} h(n,k)u(k). \end{aligned} \tag{2.57b}$$

We say that the discrete-time system described by (2.55) is *at rest* at $k = k_0 \in Z$ if $u(k) = 0$ for $k \geq k_0$ implies that $y(k) = 0$ for $k \geq k_0$. Accordingly, if system (2.55) is known to be at rest at $k = k_0$, we have

$$y(n) = \sum_{k=k_0}^{\infty} h(n,k)u(k).$$

Furthermore, if system (2.55) is known to be causal and at rest at $k = k_0$, its input–output description assumes the form [in view of (2.57b)]

$$y(n) = \sum_{k=k_0}^{n} h(n,k)u(k). \tag{2.58}$$

If now, in addition, system (2.55) is also time-invariant, (2.58) becomes

$$y(n) = \sum_{k=k_0}^{n} h(n-k)u(k) = \sum_{k=k_0}^{n} h(k)u(n-k), \tag{2.59}$$

which is a convolution sum. [Note that in (2.59) we have slightly abused the notation for $h(\cdot)$, namely that $h(n-k) = h(n-k,0)(= h(n,k).]$

Next, turning to linear, discrete-time, *MIMO systems*, we can generalize (2.55) to

$$y(n) = \sum_{k=-\infty}^{\infty} H(n,k)u(k), \tag{2.60}$$

where $y : Z \to R^p$, $u : Z \to R^m$, and

$$H(n,k) = \begin{bmatrix} h_{11}(n,k) & h_{12}(n,k) & \cdots & h_{1m}(n,k) \\ h_{21}(n,k) & h_{22}(n,k) & \cdots & h_{2m}(n,k) \\ \cdots & \cdots & \cdots & \cdots \\ h_{p1}(n,k) & h_{p2}(n,k) & \cdots & h_{pm}(n,k) \end{bmatrix}, \tag{2.61}$$

where $h_{ij}(n,k)$ represents the system response at time n of the ith component of y due to a discrete-time impulse δ applied at time k at the jth component of u, whereas the inputs at all other components of u are being held zero. The matrix H is called the *discrete-time unit impulse response matrix* of the system.

Similarly, it follows that the system (2.60) is *causal* if and only if

$$H(n,k) = 0 \quad \text{for all} \quad n < k,$$

and that the input–output description of linear, discrete-time, causal systems is given by

$$y(n) = \sum_{k=-\infty}^{n} H(n,k)u(k). \tag{2.62}$$

A discrete-time system described by (2.60) is said to be *at rest at* $k = k_0 \in Z$ if $u(k) = 0$ for $k \geq k_0$ implies that $y(k) = 0$ for $k \geq k_0$. Accordingly, if system (2.60) is known to be at rest at $k = k_0$, we have

$$y(n) = \sum_{k=k_0}^{\infty} H(n,k)u(k). \tag{2.63}$$

Moreover, if a linear discrete-time system that is at rest at k_0 is known to be causal, then its input–output description reduces to

$$y(n) = \sum_{k=k_0}^{n} H(n,k)u(k). \tag{2.64}$$

Finally, as in (2.56), it is easily shown that the unit impulse response $H(n,k)$ of a linear, *time-invariant*, discrete-time MIMO system depends only on the difference of n and k; i.e., by a slight abuse of notation we can write

$$H(n,k) = H(n-k,0) \triangleq H(n-k) \tag{2.65}$$

for all n and k. Accordingly, linear, time-invariant, causal, discrete-time MIMO systems that are at rest at $k = k_0$ are described by equations of the form

$$y(n) = \sum_{k=k_0}^{n} H(n-k)u(k). \tag{2.66}$$

We conclude by supposing that the system on hand is described by (2.14) under the assumption that $x(k_0) = 0$; i.e., the system is at rest at $k = k_0$. Then, according to (2.42) and (2.43), we obtain

$$H(n,k) = \begin{cases} C(n)\Phi(n,k+1)B(k), & n > k, \\ D(n), & n = k, \\ 0, & n < k. \end{cases} \tag{2.67}$$

Furthermore, for the time-invariant case, we obtain

$$H(n-k) = \begin{cases} CA^{n-(k+1)}B, & n > k, \\ D, & n = k, \\ 0, & n < k. \end{cases} \tag{2.68}$$

2.4.3 The Dirac Delta Distribution

For any linear time-invariant operator P from $C(R, R)$ to itself, we say that P admits an *integral representation* if there exists an integrable function (in the Riemann or Lebesgue sense), $g_p : R \to R$, such that for any $f \in C(R, R)$,

$$(Pf)(x) = (f * g_p)(x) \triangleq \int_{-\infty}^{\infty} f(\tau) g_p(x - \tau) d\tau.$$

We call g_p a *kernel of the integral representation of* P.

For the identity operator I [defined by $If = f$ for any $f \in C(R, R)$] an integral representation for which g_p is a function in the usual sense does not exist (see, e.g., Z. Szmydt, *Fourier Transformation and Linear Differential Equations*, D. Reidel Publishing Company, Boston, 1977). However, there exists a sequence of functions $\{\phi_n\}$ such that for any $f \in C(R, R)$,

$$(If)(x) = f(x) = \lim_{n \to \infty} (f * \phi_n)(x). \tag{2.69}$$

To establish (2.69) we make use of functions $\{\phi_n\}$ given by

$$\phi_n(x) = \begin{cases} n(1 - n|x|), & \text{if } |x| \leq \frac{1}{n}, \\ 0, & \text{if } |x| > \frac{1}{n}, \end{cases}$$

$n = 1, 2, 3, \ldots$. A plot of ϕ_n is depicted in Figure 2.4. In Antsaklis and Michel [1], the following useful property of ϕ_n is proved.

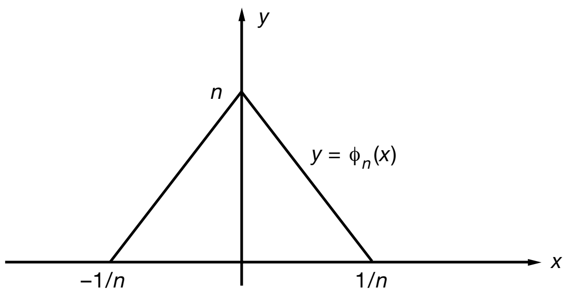

Figure 2.4. Generation of n delta distribution

Lemma 2.1. *Let f be a continuous real-valued function defined on R, and let ϕ_n be defined as above (Figure 2.4). Then for any $a \in R$,*

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} f(\tau)\phi_n(a-\tau)d\tau = f(a). \tag{2.70}$$

■

The above result, when applied to (2.69), now allows us to *define* a *generalized function* δ (also called a *distribution*) as the kernel of a *formal* or *symbolic* integral representation of the identity operator I; i.e.,

$$f(x) = \lim_{n\to\infty} \int_{-\infty}^{\infty} f(\tau)\phi_n(x-\tau)d\tau \tag{2.71}$$

$$\triangleq \int_{-\infty}^{\infty} f(\tau)\delta(x-\tau)d\tau \tag{2.72}$$

$$= f * \delta(x). \tag{2.73}$$

It is emphasized that the expression (2.72) *is not an integral at all* (in the Riemann or Lebesgue sense) *but only a symbolic representation.* The *generalized function* δ is called the *unit impulse* or the *Dirac delta distribution.*

In applications we frequently encounter functions $f \in C(R^+, R)$. If we extend f to be defined on all of R by letting $f(x) = 0$ for $x < 0$, then (2.70) becomes

$$\lim_{n\to\infty} \int_{0}^{\infty} f(\tau)\phi_n(a-\tau)d\tau = f(a) \tag{2.74}$$

for any $a > 0$, where we have used the fact that in the proof of Lemma 2.1, we need f to be continuous only in a neighborhood of a (refer to [1]). Therefore, for $f \in C(R^+, R)$, (2.71) to (2.74) yield

$$\lim_{n\to\infty} \int_{0}^{\infty} f(\tau)\phi_n(t-\tau)d\tau \triangleq \int_{0}^{\infty} f(\tau)\delta(t-\tau)d\tau = f(t) \tag{2.75}$$

for any $t > 0$. Since the ϕ_n are even functions, we have $\phi_n(t-\tau) = \phi_n(\tau - t)$, which allows for the representation $\delta(t-\tau) = \delta(\tau - t)$. We obtain from (2.75) that

$$\lim_{n\to\infty} \int_{0}^{\infty} f(\tau)\phi_n(\tau-t)d\tau \triangleq \int_{0}^{\infty} f(\tau)\delta(\tau-t)d\tau = f(t)$$

for any $t > 0$. Changing the variable $\tau' = \tau - t$, we obtain

$$\lim_{n\to\infty} \int_{-t}^{\infty} f(\tau'+t)\phi_n(\tau')d\tau' \triangleq \int_{-t}^{\infty} f(\tau'+t)\delta(\tau')d\tau' = f(t)$$

for any $t > 0$. Taking the limit $t \to 0^+$, we obtain

$$\lim_{n\to\infty} \int_{0^-}^{\infty} f(\tau'+t)\phi_n(\tau')d\tau' \triangleq \int_{0^-}^{\infty} f(\tau')\delta(\tau')d\tau' = f(0), \tag{2.76}$$

where $\int_{0-}^{\infty} f(\tau')\delta(\tau')d\tau'$ is not an integral but a symbolic representation of $\lim_{n\to\infty}\int_{0-}^{\infty} f(\tau'+t)\phi_n(\tau')d\tau'$.

Now let s denote a complex variable. If in (2.75) and (2.76) we let $f(\tau) = e^{-s\tau}, \tau > 0$, then we obtain the *Laplace transform*

$$\lim_{n\to\infty}\int_{0-}^{\infty} e^{-s\tau}\phi_n(\tau)d\tau \triangleq \int_{0-}^{\infty} e^{-s\tau}\delta(\tau)d\tau = 1. \tag{2.77}$$

Symbolically we denote (2.77) by

$$\mathcal{L}(\delta) = 1, \tag{2.78}$$

and we say that the Laplace transform of the unit impulse function or the Dirac delta distribution is equal to one.

Next, we point out another important property of δ. Consider a (time-invariant) operator P and assume that P admits an integral representation with kernel g_P. If in (2.75) we let $f = g_P$, we have

$$\lim_{n\to\infty}(P\phi_n)(t) = g_P(t), \tag{2.79}$$

and we write this (symbolically) as

$$P\delta = g_P. \tag{2.80}$$

This shows that the impulse response of a linear, time-invariant, continuous-time system with integral representation is equal to the kernel of the integral representation of the system.

Next, for any linear time-varying operator P from $C(R,R)$ to itself, we say that P admits an *integral representation* if there exists an integrable function (in the Riemann or Lebesgue sense), $g_P : R\times R \to R$, such that for any $f \in C(R,R)$,

$$(Pf)(\eta) = \int_{-\infty}^{\infty} f(\tau)g_P(\eta,\tau)d\tau. \tag{2.81}$$

Again, we call g_P a *kernel of the integral representation of* P. It turns out that *the impulse response of a linear, time-varying, continuous-time system with integral representation is again equal to the kernel of the integral representation of the system.* To see this, we first observe that if $h \in C(R\times R, R)$, and if in Lemma 2.1 we replace $f \in C(R,R)$ by h, then all the ensuing relationships still hold, with obvious modifications. In particular, as in (2.71), we have for all $t \in R$,

$$\lim_{n\to\infty}\int_{-\infty}^{\infty} h(t,\tau)\phi_n(\eta-\tau)d\tau \triangleq \int_{-\infty}^{\infty} h(t,\tau)\delta(\eta-\tau)d\tau = h(t,\eta). \tag{2.82}$$

Also, as in (2.75), we have

$$\lim_{n\to\infty}\int_{0}^{\infty} h(t,\tau)\phi_n(\eta-\tau)d\tau \triangleq \int_{0}^{\infty} h(t,\tau)\delta(\eta-\tau)d\tau = h(t,\eta) \tag{2.83}$$

for $\eta > 0$.

Now let $h(t,\tau) = g_P(t,\tau)$. Then (2.82) yields

$$\lim_{n\to\infty}\int_{-\infty}^{\infty} g_P(t,\tau)\phi_n(\eta-\tau)d\tau \triangleq \int_{-\infty}^{\infty} g_P(t,\tau)\delta(\eta-\tau)d\tau = g_P(t,\eta), \quad (2.84)$$

which establishes our assertion. The common interpretation of (2.84) is that $g_P(t,\eta)$ represents the response of the system at time t due to an impulse applied at time η.

2.4.4 Linear Continuous-Time Systems

We let P denote a linear time-varying operator from $C(R,R^m) \triangleq U$ to $C(R,R^p) = Y$, and we assume that P admits an *integral representation* given by

$$y(t) = (Pu)(t) = \int_{-\infty}^{\infty} H_P(t,\tau)u(\tau)d\tau, \quad (2.85)$$

where $H_P : R\times R \to R^{p\times m}, u \in U$, and $y \in Y$ and where H_P is assumed to be integrable. This means that each element of H_P, $h_{P_{ij}} : R\times R \to R$ is integrable (in the Riemann or Lebesgue sense).

Now let y_1 and y_2 denote the response of system (2.85) corresponding to the input u_1 and u_2, respectively, let α_1 and α_2 be real scalars, and let y denote the response of system (2.85) corresponding to the input $\alpha_1 u_1 + \alpha_2 u_2 = u$. Then

$$\begin{aligned} y = P(u) = P(\alpha_1 u_1 + \alpha_2 u_2) &= \int_{-\infty}^{\infty} H_P(t,\tau)[\alpha_1 u_1(\tau) + \alpha_2 u_2(\tau)]d\tau \\ &= \alpha_1 \int_{-\infty}^{\infty} H_P(t,\tau)u_1(\tau)d\tau + \alpha_2 \int_{-\infty}^{\infty} H_P(t,\tau)u_2(\tau)d\tau \\ &= \alpha_1 P(u_1) + \alpha_2 P(u_2) = \alpha_1 y_1 + \alpha_2 y_2, \end{aligned} \quad (2.86)$$

which shows that system (2.85) is indeed a *linear* system in the sense defined in (2.50).

Next, we let all components of $u(\tau)$ in (2.85) be zero, except for the jth component. Then the ith component of $y(t)$ in (2.85) assumes the form

$$y_i(t) = \int_{-\infty}^{\infty} h_{P_{ij}}(t,\tau)u_j(\tau)d\tau. \quad (2.87)$$

According to the results of the previous subsection [see (2.84)], $h_{P_{ij}}(t,\tau)$ denotes the response of the ith component of the output of system (2.85), measured at time t, due to an impulse applied to the jth component of the input of system (2.85), applied at time τ, whereas all of the remaining components of the input are zero. Therefore, we call $H_P(t,\tau) = [h_{P_{ij}}(t,\tau)]$ the *impulse response matrix* of system (2.85).

Now suppose that it is known that system (2.85) is *causal*. Then its output will be identically zero before an input is applied. It follows that system (2.85) is causal if and only if

$$H_P(t,\tau) = 0 \quad \text{for all } t < \tau.$$

Therefore, when system (2.85) is causal, we have in fact that

$$y(t) = \int_{-\infty}^{t} H_P(t,\tau)u(\tau)d\tau. \tag{2.88}$$

We can rewrite (2.88) as

$$\begin{aligned} y(t) &= \int_{-\infty}^{t_0} H_P(t,\tau)u(\tau)d\tau + \int_{t_0}^{t} H_P(t,\tau)u(\tau)d\tau \\ &\triangleq y(t_0) + \int_{t_0}^{t} H_P(t,\tau)u(\tau)d\tau. \end{aligned} \tag{2.89}$$

We say that the continuous-time system (2.85) is *at rest at* $t = t_0$ if $u(t) = 0$ for $t \geq t_0$ implies that $y(t) = 0$ for $t \geq t_0$. Note that our problem formulation mandates that the system be at rest at $t_0 = -\infty$. Also, note that if a system (2.85) is known to be causal and to be at rest at $t = t_0$, then according to (2.89) we have

$$y(t) = \int_{t_0}^{t} H_P(t,\tau)u(\tau)d\tau. \tag{2.90}$$

Next, suppose that it is known that the system (2.85) is *time-invariant*. This means that if in (2.87) $h_{P_{ij}}(t,\tau)$ is the response y_i at time t due to an impulse applied at time τ at the jth component of the input [i.e., $u_j(\tau) = \delta(t)$], with all other input components set to zero, then a $-\tau$ time shift in the input [i.e., $u_j(t-\tau) = \delta(t-\tau)$] will result in a corresponding $-\tau$ time shift in the response, which results in $h_{P_{ij}}(t-\tau, 0)$. Since this argument holds for all $t, \tau \in R$ and for all $i = 1, \ldots, p$, and $j = 1, \ldots, m$, we have $H_P(t,\tau) = H_P(t-\tau, 0)$. If we define (using a slight abuse of notation) $H_P(t-\tau,0) = H_P(t-\tau)$, then (2.85) assumes the form

$$y(t) = \int_{-\infty}^{\infty} H_P(t-\tau)u(\tau)d\tau. \tag{2.91}$$

Note that (2.91) is consistent with the definition of the integral representation of a linear time-invariant operator introduced in the previous subsection.

The right-hand side of (2.91) is the familiar *convolution integral* of H_P and u and is written more compactly as

$$y(t) = (H_P * u)(t). \tag{2.92}$$

We note that since $H_P(t-\tau)$ represents responses at time t due to impulse inputs applied at time τ, then $H_P(t)$ represents responses at time t due to impulse function inputs applied at $\tau = 0$. Therefore, a linear time-invariant system (2.91) is causal if and only if $H_P(t) = 0$ for all $t < 0$.

If it is known that the linear time-invariant system (2.91) is causal and is at rest at t_0, then we have

$$y(t) = \int_{t_0}^{t} H_P(t-\tau)u(\tau)d\tau = \int_{t_0}^{t} H_P(\tau)u(t-\tau)d\tau. \tag{2.93}$$

In this case it is customary to choose, without loss of generality, $t_0 = 0$. We thus have

$$y(t) = \int_{0}^{t} H_P(t-\tau)u(\tau)d\tau, \quad t \geq 0. \tag{2.94}$$

If we take the Laplace transform of both sides of (2.94), provided it exists, we obtain

$$\hat{y}(s) = \widehat{H}_P(s)\hat{u}(s), \tag{2.95}$$

where $\hat{y}(s) = [\hat{y}_1(s), \dots, \hat{y}_p(s)]^T$, $\widehat{H}_P(s) = [\hat{h}_{P_{ij}}(s)]$, $\hat{u}(s) = [\hat{u}_1(s), \dots, \hat{u}_m(s)]^T$ where the $\hat{y}_i(s)$, $\hat{u}_j(s)$, and $\hat{h}_{P_{ij}}(s)$ denote the Laplace transforms of $y_i(t)$, $u_j(t)$, and $h_{Pij}(t)$, respectively [see Chapter 3 for more details concerning Laplace transforms]. Consistent with (2.78), we note that $\widehat{H}_P(s)$ represents the Laplace transform of the impulse response matrix $H_P(t)$. We call $\widehat{H}_P(s)$ a *transfer function matrix*.

Now suppose that the input–output relation of a system is specified by the state and output equations (2.3), repeated here as

$$\dot{x} = A(t)x + B(t)u, \tag{2.96a}$$

$$y = C(t)x + D(t)u. \tag{2.96b}$$

If we assume that $x(t_0) = 0$ so that the system is at rest at $t_0 = 0$, we obtain for the response of this system,

$$y(t) = \int_{t_0}^{t} C(t)\Phi(t,\tau)B(\tau)u(\tau)d\tau + D(t)u(t) \tag{2.97}$$

$$= \int_{t_0}^{t} [C(t)\Phi(t,\tau)B(\tau) + D(t)\delta(t-\tau)]u(\tau)d\tau, \tag{2.98}$$

where in (2.98) we have made use of the interpretation of δ given in Subsection 2.4.3. Comparing (2.98) with (2.90), we conclude that the impulse response matrix for system (2.96) is given by

$$H_P(t,\tau) = \begin{cases} C(t)\Phi(t,\tau)B(\tau) + D(t)\delta(t-\tau), & t \geq \tau, \\ 0, & t < \tau. \end{cases} \tag{2.99}$$

Finally, for time-invariant systems described by the state and output equations (2.7), repeated here as

$$\dot{x} = Ax + Bu, \tag{2.100a}$$

$$y = Cx + Du, \tag{2.100b}$$

we obtain for the impulse response matrix the expression

$$H_P(t-\tau) = \begin{cases} Ce^{A(t-\tau)}B + D\delta(t-\tau), & t \geq \tau, \\ 0, & t < \tau, \end{cases} \tag{2.101}$$

or, as is more commonly written,

$$H_P(t) = \begin{cases} Ce^{At}B + D\delta(t), & t \geq 0, \\ 0, & t < 0. \end{cases} \tag{2.102}$$

We will pursue the topics of this section further in Chapter 3.

2.5 Summary and Highlights

Internal Descriptions

- The *response of the time-varying continuous-time system*

$$\dot{x} = A(t)x + B(t)u, \quad y = C(t)x + D(t)u, \tag{2.3}$$

with $x(t_0) = x_0$ is given by

$$y(t) = C(t)\Phi(t, t_0)x_0 + C(t)\int_{t_0}^{t} \Phi(t, s)B(s)u(s)ds + D(t)u(t). \tag{2.6}$$

- The *response of the time-invariant continuous-time system*

$$\dot{x} = Ax + Bu, \quad y = Cx + Du, \tag{2.7}$$

is given by

$$y(t) = Ce^{A(t-t_0)}x_0 + C\int_{t_0}^{t} e^{A(t-s)}Bu(s)ds + Du(t). \tag{2.9}$$

- The *response of the discrete-time system*

$$x(k+1) = A(k)x(k) + B(k)u(k), \quad y(k) = C(k)x(k) + D(k)u(k), \tag{2.14}$$

with $x(k_0) = x_{k_0}$ is given by

$$y(k) = C(k)\Phi(k, k_0)x_{k_0} + C(k)\sum_{j=k_0}^{k-1}\Phi(k, j+1)B(j)u(j) + D(k)u(k), \quad k > k_0 \tag{2.42}$$

and

$$y(k_0) = C(k_0)x_{k_0} + D(k_0)u(k_0), \tag{2.43}$$

where the state transition matrix

$$\Phi(k, k_0) = \prod_{j=k_0}^{k-1} A(j), \quad k > k_0, \tag{2.31}$$

$$\Phi(k_0, k_0) = I. \tag{2.32}$$

In the time-invariant case

$$x(k+1) = Ax(k) + Bu(k), \quad y(k) = Cx(k) + Du(k), \tag{2.15}$$

with $x(k) = x_{k_0}$, the system response is given by

$$y(k) = CA^{k-k_0}x_{k_0} + C\sum_{j=k_0}^{k-1} A^{k-(j+1)}B(j)u(j) + Du(k), \quad k > k_0, \tag{2.44}$$

and

$$y(k_0) = Cx_{k_0} + Du(k_0). \tag{2.45}$$

External Descriptions

- Properties: *Linearity* (2.50); with *memory; causality; time-invariance*
- *The input–output description* of a *linear, discrete-time, causal, time-invariant* system that is at rest at $k = k_0$ is given by

$$y(n) = \sum_{k=k_0}^{n} h(n-k)u(k) = \sum_{k=k_0}^{n} h(k)u(n-k). \tag{2.59}$$

  $h(n-k)(= h(n-k, 0))$ is the discrete-time unit impulse response of the system.
- For the *discrete-time, time-invariant system*

$$x(k+1) = Ax(k) + Bu(k), \quad y(k) = Cx(k) + Du(k),$$

  the discrete-time unit impulse response (for the MIMO case) is

$$H(n-k) = \begin{cases} CA^{n-(k+1)}B, & n > k, \\ D, & n = k, \\ 0, & n < k. \end{cases} \tag{2.68}$$

- *The unit impulse* (*Dirac delta distribution*) $\delta(t)$ satisfies

$$\int_a^b f(\tau)\delta(t-\tau)d\tau = f(t),$$

where $a < t < b$ [see (2.75)].
- *The input–output description of a linear, continuous-time, causal, time-invariant system* that is at rest at $t = t_0$ is given by

$$y(t) = \int_{t_0}^t H_P(t-\tau)u(\tau)d\tau = \int_{t_0}^t H_P(\tau)u(t-\tau)d\tau. \tag{2.93}$$

$H_P(t-\tau)(= H_P(t-\tau, 0))$ is the continuous-time unit impulse response of the system.
- For the *time-invariant system*

$$\dot{x} = Ax + Bu \quad y = Cx + Du, \tag{2.100}$$

the continuous-time unit impulse response is

$$H_P(t-\tau) = \begin{cases} Ce^{A(t-\tau)}B + D\delta(t-\tau), & t \geq \tau, \\ 0, & t < \tau. \end{cases} \tag{2.101}$$

2.6 Notes

An original standard reference on linear systems is by Zadeh and Desoer [7]. Of the many excellent texts on this subject, the reader may want to refer to Brockett [2], Kailath [5], and Chen [3]. For more recent texts on linear systems, consult, e.g., Rugh [6] and DeCarlo [4]. The presentation in this book relies mostly on the recent text by Antsaklis and Michel [1].

References

1. P.J. Antsaklis and A.N. Michel, *Linear Systems*, Birkhäuser, Boston, MA, 2006.
2. R.W. Brockett, *Finite Dimensional Linear Systems*, Wiley, New York, NY, 1970.
3. C.T. Chen, *Linear System Theory and Design*, Holt, Rinehart and Winston, New York, NY, 1984.
4. R.A. DeCarlo, *Linear Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1989.
5. T. Kailath, *Linear Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1980.
6. W.J. Rugh, *Linear System Theory, Second Edition*, Prentice-Hall, Englewood Cliffs, NJ, 1996.
7. L.A. Zadeh and C.A. Desoer, *Linear System Theory - The State Space Approach*, McGraw-Hill, New York, NY, 1963.

Exercises

2.1. (a) For the mechanical system given in Exercise 1.2a, we view f_1 and f_2 as making up the system input vector, and y_1 and y_2 the system output vector. Determine a state-space description for this system.
(b) For the same mechanical system, we view $f_1 + 5f_2$ as the (scalar-valued) system input and we view $8y_1 + 10y_2$ as the (scalar-valued) system output. Determine a state-space description for this system.
(c) For part (a), determine the input–output description of the system.
(d) For part (b), determine the input–output description of the system.

2.2. In Example 1.3, we view e_a and θ as the system input and output, respectively.

(a) Detemine a state-space representation for this system.
(b) Determine the input–output description of this system.

2.3. For the second-order section digital filter in direct form, given in Figure 2.1, determine the input–output description, where $x_1(k)$ and $u(k)$ denote the output and input, respectively.

2.4. In the circuit of Figure 2.5, $v_i(t)$ and $v_0(t)$ are voltages (at time t) and R_1 and R_2 are resistors. There is also an ideal diode that acts as a short circuit when v_i is positive and as an open circuit when v_i is negative. We view v_i and v_0 as the system input and output, respectively.

(a) Determine an input–output description of this system.
(b) Is this system linear? Is it time-varying or time-invariant? Is it causal? Explain your answers.

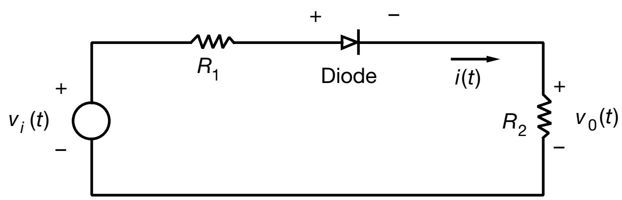

Figure 2.5. Diode circuit

2.5. We consider the *truncation operator* given by

$$y(t) = T_\tau(u(t))$$

as a system, where $\tau \in R$ is fixed, u and y denote system input and output, respectively, t denotes time, and $T_\tau(\cdot)$ is specified by

$$T_\tau(u(t)) = \begin{cases} u(t) & t \le \tau, \\ 0 & t > \tau. \end{cases}$$

Is this system causal? Is it linear? Is it time-invariant? What is its impulse response?

2.6. We consider the *shift operator* given by

$$y(t) = Q_\tau(u(t)) = u(t - \tau)$$

as a system, where $\tau \in R$ is fixed, u and y denote system input and system output, respectively, and t denotes time. Is this system causal? Is it linear? Is it time-invariant? What is its impulse response?

2.7. Consider the system whose input–output description is given by

$$y(t) = \min\{u_1(t), u_2(t)\},$$

where $u(t) = [u_1(t), u_2(t)]^T$ denotes the system input and $y(t)$ is the system output. Is this system linear?

2.8. Suppose it is known that a linear system has impulse response given by $h(t, \tau) = \exp(-|t - \tau|)$. Is this system causal? Is it time-invariant?

2.9. Consider a system with input–output description given by

$$y(k) = 3u(k + 1) + 1, \quad k \in Z,$$

where y and u denote the output and input, respectively (recall that Z denotes the integers). Is this system causal? Is it linear?

2.10. Use expression (2.54),

$$x(n) = \sum_{k=-\infty}^{\infty} x(k)\delta(n - k),$$

and $\delta(n) = p(n) - p(n - 1)$ to express the system response $y(n)$ due to any input $u(k)$, as a function of the unit step response of the system [i.e., due to $u(k) = p(k)$].