CHAPTER 2

SYSTEMS

In Chapter 1, we defined a signal as a "physical quantity, or quality, which conveys information." For this book, a key mathematical representation of a signal is as a function of a single variable we call time. A *system* takes one or more signals as input, performs operations on the signals, and produces one or more signals as output. So, the input is the stimulus or excitation applied to a system from an external source, usually in order to produce a specified response. The output is the actual response obtained from a system. In an algebraic framework, we can represent a system as an operator that maps input signals onto output signals [8].

We have already talked about sampling and quantizing continuous-time signals into digital signals in Chapter 1. In practice, this conversion is accomplished by an analog-to-digital (A/D) converter. An ideal representation of an A/D conversion system is as a cascade of a lowpass filter, a sampling device, and a quantizer. The lowpass filter will ideally pass components of continuous-time signals that are below a specific frequency and reject those above the specific frequency. Lowpass filtering is necessary for guaranteeing that we can choose a sampling period T (or equivalently a sampling rate $f_0 = \frac{1}{T}$) that satisfies the Sampling Theorem.

An A/D converter is an example of a dictionary definition of a system as "a group of related parts working together, or an ordered set of ideas, methods, or ways of working" [9]. From an implementation point-of-view, a system is an arrangement of physical components connected or related in such a manner as to form and/or act as an entire unit. From a signal processing perspective, a system can be viewed as any process that results in the transformation of signals, in which systems act on signals in prescribed ways [10, 11].

In this book, we define a *system* as a mapping of N input signals onto M output signals. The mapping carries out a transformation on the input signals according to a set of rules. The rest of this chapter presents basic definitions and background mathematics that will be used in the remainder of this book. Since many readers will already be familiar with this material, we shall aim to be logically consistent rather than mathematically rigorous.

2.1 BASIC DEFINITIONS

A system is said to be a *single-variable system* if it has only one input and only one output. A system is said to be a *multivariable system* if it has more than one input or more than one output. An equation that describes the relation between the input and the output of a system is called the *input–output relationship*, also known as the external description or the input–output description, of the system. In developing this relationship, we assume that the knowledge of the internal structure of a system is unavailable to us. Instead, the only access to the system is by means of the input ports and the output ports. Under this assumption, a system may be considered as a "*black box*." We apply inputs to a black box, measure their corresponding outputs, and then try to abstract key properties of the system from these input–output pairs [12].

We represent the input–output relation of a system using the notation

$$x \to y$$

or

$$x \xrightarrow{\mathcal{R}} y$$

where $\mathcal{R}$ is an operator—a *rule*, formula, or set of formulas for transforming a signal x into a signal y. Some authors put braces around the signal designation to indicate that there are, in general, several input signals and several output signals

$$\{x\} \xrightarrow{\mathcal{R}} \{y\}$$

or explicitly specify the input and the output

$$(x_1, x_2, \ldots, x_N) \xrightarrow{\mathcal{R}} (y_1, y_2, \ldots, y_M)$$

This notation assumes that the output is excited solely and uniquely by the input. It is also legitimate to write $y = \mathcal{R}(x)$, where $\mathcal{R}$ is some operator or function that specifies uniquely the output y in terms of the input x of the system. Other authors prefer to define a system as an interconnected collection of abstract objects, defined by relations of input–output pairs rather than by operators or functions [13].

Although systems of prime concern are those required for processing a signal that is a function of the single independent variable *time*, t, it should be appreciated that the independent variable in many applications need not be time. When more than one independent variable is involved, the system is multidimensional. For example, in a video processing systems, we have two spatial variables for each frame of video and a discrete-time variable to index the frames. For now, we focus on one-dimensional systems.

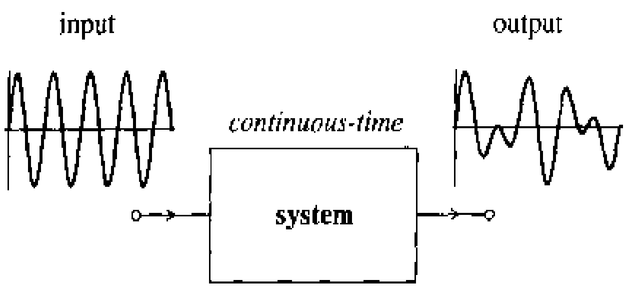


Figure 2.1 Continuous-time system.

The *time response* of a system is the output signal(s) as a function of time, following the application of a set of prescribed input signals, under specified operating conditions. In a *continuous-time system* (Fig. 2.1), the input and output signals are continuous-time. A *discrete-time system* (Fig. 2.2) has discrete-time input and output signals. A *hybrid-time system* (Fig. 2.3) is one in which both continuous-time and discrete-time signals appear at the inputs and outputs. An A/D converter is an example of a hybrid system: The input is continuous-time, but the output is discrete-time. Therefore, a continuous-time system is one in which continuous-time input signals are transformed into continuous-time output signals, and a discrete-time system is one that transforms discrete-time input signals into discrete-time output signals. A continuous-time system is symbolically represented as $x(t) \to y(t)$, and a discrete-time system is represented as $x(k) \to y(k)$ or $x_k \to y_k$, where k stands for the integer sample number.

A discrete-time system is *digital* if it operates on discrete-time signals whose amplitudes are quantized. Quantization maps each continuous amplitude level into a number. The digital system employs digital hardware either (a) explicitly in the form of the usual collection of logic circuits or (b) implicitly when the operations on the signals are executed by writing a computer program [14].

A *lumped system* is one that can be decomposed into a finite number of components, each with a finite number of inputs and outputs, and such that the values of the outputs at every time are functions of the inputs, their derivatives, and their integrals, at the same instant of time. Since the velocity of light and velocity of sound are finite, all physical objects with inputs and outputs at different places are not lumped. Nevertheless, the lumped system can provide a good approximation for a physical system.

The quantity that is responsible for activating the system to produce the output is called the *control action*. An *open-loop system* is one in which the control action is independent of the output. A *closed-loop* system is one in which the control action

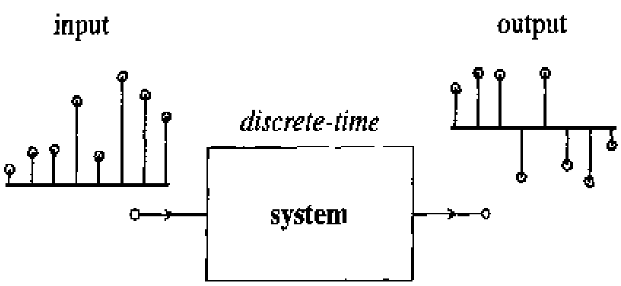


Figure 2.2 Discrete-time system.

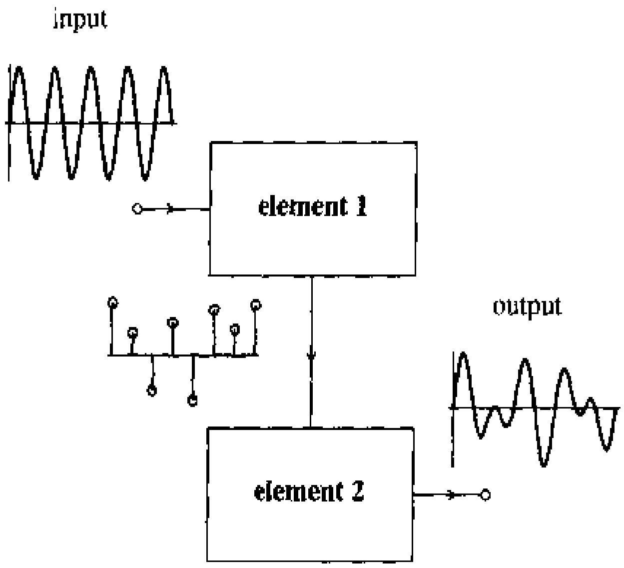


Figure 2.3 Hybrid-time system.

is dependent on the output. Closed-loop systems are commonly called feedback systems. *Feedback* is that property of a closed-loop system which permits the output to be compared with the input to the system so that the appropriate control action may be formed as some function of the output and input. A *control system* is one that commands, directs, or regulates itself or another system.

In many cases, we are presented with a specific system and are interested in characterizing it in detail to understand how it will respond to various inputs. In another context, our interest may be focused on the problem of designing systems to process signals in particular way. *Analysis* of a system is investigation of the properties and the behavior (response) of an existing system. *Design* of a system is the choice and arrangement of systems components to perform a specific task. *Design by analysis* is accomplished by modifying the characteristics of an existing system. *Design by synthesis* means that we define the form of the system directly from its specifications.

In order to analyze, design, and evaluate a system, the description of its components and their interconnections must be put into a suitable form. A mathematical or graphical representation of a system is called the *model*. The model resembles the system in its salient features but is easier to study. Some authors refer to models of physical systems as systems. Therefore, a physical system is a device or a collection of devices existing in the real world, and a system is a model of a physical system [12].

A *mathematical model* is a set of mathematical relations representing the system. These relations may assume many forms, including linear equations, nonlinear equations, integral equations, differential equations, and difference equations. The solution of these equations represents the system's behavior. Often, this solution is difficult, if not impossible, to find, and certain simplifying assumptions must be made in the mathematical description. Frequently, these approximations and simplifications lead to systems describable by linear ordinary differential equations or difference equations.

2.2 BLOCK DIAGRAMS

A *block diagram* is a shorthand, pictorial representation of the cause and effect relationship between the input and output of a system. It provides a convenient and useful method for characterizing the functional relationships among the various components of a system. Block diagrams are representations of either (a) the schematic diagram of a physical system or (b) the set of mathematical equations characterizing its parts. The simplest form of the block diagram is the single *block*, with one input and one output (Fig. 2.4). The interior of the rectangle representing the block usually contains (a) the name of the component, (b) a description of the component, or (c) the symbol for the mathematical operation to be performed on the input to yield the output. The *arrows* represent the direction of unilateral information or signal flow. The standard symbols used to represent various types of blocks are shown in Fig. 2.5.

The operations of addition and subtraction are represented by a circle, called a *summing point*, with the appropriate plus or minus sign associated with the arrows entering the circle (Fig. 2.6). The output is the algebraic sum of the inputs. Any number of inputs may enter a summing point. Some authors put the plus sign "+" or the Greek letter "Σ" in the circle.

In order to employ the same signal or variable as an input to more than one block or summing point, a *takeoff point* is used, as shown in Fig. 2.7. This permits the signal to proceed unaltered along several different paths to several destinations. A block with N inputs and M outputs is represented by a rectangle as shown in Fig. 2.8. The blocks representing the various components of a system are connected in a fashion which characterizes their functional relationship within the system. The arrows connecting one block with another represent the direction of flow of signals or information. In general, a block diagram consists of a specific configuration of four types of elements: blocks, summing points, takeoff points, and arrows representing unidirectional signal flow.

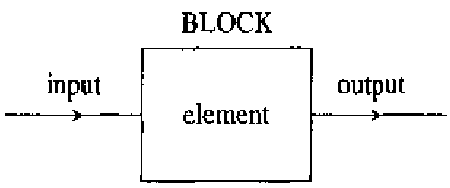


Figure 2.4 Single block with one input and one output.

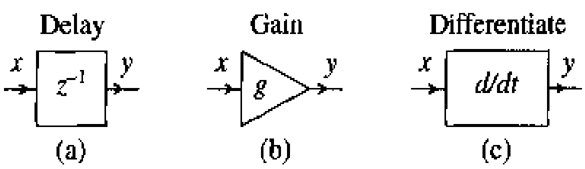


Figure 2.5 Various blocks: (a) Delay, $y(t) = x(t - T)$, some authors put the amount of delay T in the box, (b) Gain, multiplier by constant, or amplifier, $y = gx$, (c) Differentiate with respect to time, $y = dx/dt$.

Basically, two blocks may be connected in cascade (Fig. 2.9), in parallel (Fig. 2.10), or in feedback (Fig. 2.11). We combine cascade, parallel, and feedback interconnections to obtain more complicated interconnections. Next, we can use these interconnections to construct new systems or models out of existing ones.

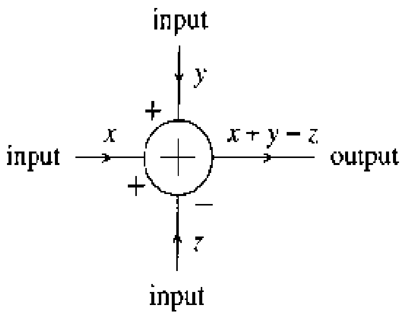


Figure 2.6 Summing point—representation of addition and subtraction.

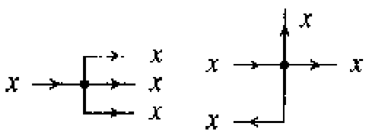


Figure 2.7 Takeoff point in a block diagram.

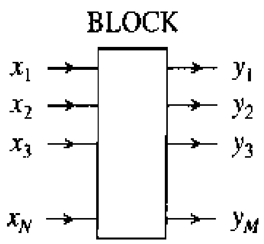


Figure 2.8 An N-input, M-output block.

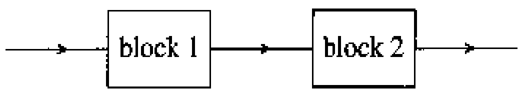


Figure 2.9 Blocks connected in cascade.

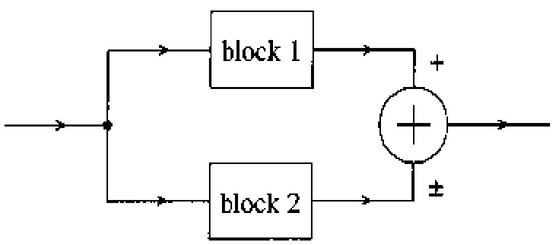


Figure 2.10 Blocks connected in parallel.

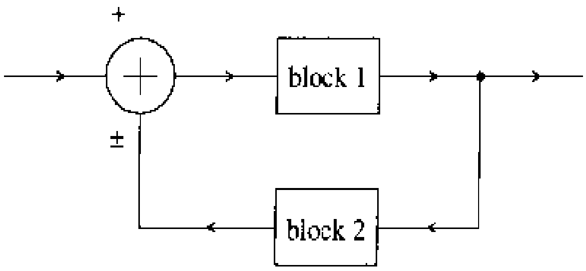


Figure 2.11 Blocks connected in feedback. The plus sign refers to a *positive* feedback, and the minus sign refers to a *negative* feedback.

2.3 SYSTEM PROPERTIES

2.3.1 State and Relaxed Systems

For most systems of interest, the output at time t_0 depends not only on the input applied at t_0, but also on the input applied before and after t_0. If an input $x(t)$, $t \geq t_0$, is applied to a system, unless we know the input applied before t_0, the output $y(t)$, $t \geq t_0$, is generally not uniquely determinable. For different inputs applied before t_0, we will obtain different output $y(t)$, $t \geq t_0$, although the same input $x(t)$, $t \geq t_0$, is applied. Hence, in developing the input–output description, before an input is applied, the system must be assumed to be *relaxed* or *at rest*, and that **the output is excited solely and uniquely by the input applied**. This definition is general and it is valid for an arbitrary system, continuous-time, discrete-time, and so on. If the concept of energy is applicable to a system, the system is said to be relaxed at time t_0 if no energy is stored in the system at that instant. A system is said to be relaxed at time t_0 if the output $y(t)$, $t \geq t_0$, is solely and uniquely excited by the input $x(t)$ defined for $t \geq t_0$.

We shall assume that every system is relaxed at time $-\infty$. We shall call a system that is initially relaxed at $-\infty$ an *initially relaxed system*, or a *relaxed system*, for short. In this section, whenever we talk about the input–output pairs of a system, we mean only those input–output pairs derived or formulated under relaxed system assumptions.

The input–output description of a system is applicable only when the system is initially relaxed. If a system is not initially relaxed, say at time t_0, then it depends also on the set of initial conditions at t_0, which we call the *state*. The state is the information that, together with the input, determines **uniquely** the output. The state of a system at time t_0 is the minimum amount of information needed about the system at time t_0, in addition to the system diagram (or system equations) and the system inputs between t_0 and $t_1 > t_0$, so that all outputs at any time t_1 can be determined. The set of equations that describes the unique relations between the input, output, and state is called a *dynamical equation*.

If we can estimate the state of a system from its output, in a finite time interval, the system is said to be *observable*. To be more precise, a system is observable at t_0, if for any state at time t_0, there exists a finite $t_1 > t_0$ such that the knowledge of the input and the output over the time interval $t_0 \leq t \leq t_1$ suffices to determine the state. If we can steer the state of a system from the input, in a finite time interval, the system is referred to as *controllable*. We can find inputs capable of moving any state of a controllable system to any other state in a finite time. Furthermore, there is no constraint imposed on the input—its magnitude can be as large as desired.

2.3.2 Causality and Realizable Systems

A system in which time is the independent variable is called *causal* if the output depends only on the present and past values of the input. If $y(t)$ is the output, then $y(t)$ depends only on the input $x(\tau)$ for values of $\tau \le t$. Such a system is often referred to as being *nonanticipative*, because the system output does not anticipate future values of the input.

Causal systems are important, but they are not the only systems that are of practical significance. Causality is not of fundamental importance in

- applications, such as image processing, in which the independent variable is not time;
- processing recorded data for which time is the independent variable, as often happens with speech; and
- many applications, such as in stock market analysis and demographic studies, when we want to determine a slowly varying trend in data, and when we average data over an interval to smooth out the fluctuations and keep only the trend.

Finally, noncausal systems may arise in the course of analysis, when decomposing or recombining causal systems. Causal systems are sometimes called *physically realizable* systems.

2.3.3 Stability

The stability of a system is determined by its response to inputs or disturbances. Intuitively, a stable system is one that will remain at rest unless excited by an external source and will return to rest if all excitations are removed. The definition of a stable system can be based upon the response of the system to *bounded inputs*. Bounded inputs have magnitudes that are less than some finite value for all time. A relaxed system is said to be bounded-input bounded-output (BIBO) stable if every bounded input produces a bounded output. The stability that is defined in terms of the input–output description is applicable only to relaxed systems. This stability is referred to as the *input–output stability*. A zero-input system is said to be *asymptotically stable* if the response approaches zero asymptotically as time t approaches infinity. If the response remains bounded for $t > t_0$, we speak about the *bounded stability*. There are several methods (criteria) for determining system stability: Routh, Hurwitz, Liénard–Chipart, Nyquist, Lyapunov, and so on.

Consideration of the stability of a system provides valuable information about its behavior and is an important issue in system design. Often, it is desirable to determine a range of values of a particular system parameter for which the system is stable. The concept of stability is extremely important, because almost every workable system is designed to be stable. If a system is not stable, it is usually of no use in practice.

2.3.4 Time Invariance

A relaxed system is *time-invariant*, also known as fixed or stationary, if a time shift in the input signal causes a time shift in the output signal. Therefore, if an input $x(t)$ produces an output $y(t)$, then an input $x(t-t_0)$ produces an output $y(t-t_0)$. In other words, no matter at what time an input is applied to a time-invariant relaxed system, the waveform of the output is the same. We can also say that a (relaxed) system is time-invariant if delaying the input by t_0 seconds merely delays the response by t_0 seconds. A relaxed discrete-time system is said to be time-invariant if a shifted input $x(k-k_0)$ produces a shifted output $y(k-k_0)$. In the case of discrete-time digital systems, we often use the term *shift-invariant* instead of time-invariant. The characteristics and parameters of a time-invariant system do not change with time. A relaxed system that is not time-invariant is said to be *time-varying*.

2.3.5 Linearity and Superposition

Consider a relaxed system in which there is one independent variable t. A *linear system* is a system which has the property that if:

(a) an input $x_1(t)$ produces an output $y_1(t)$ and

(b) an input $x_2(t)$ produces an output $y_2(t)$, then

(c) an input $c_1x_1(t) + c_2x_2(t)$ produces an output $c_1y_1(t) + c_2y_2(t)$ for all pairs of inputs $x_1(t)$ and $x_2(t)$ and all pairs of constants c_1 and c_2.

Otherwise the (relaxed) system is *nonlinear*. This is the concept of zero-state linearity. The concept of linearity can be represented by the principle of superposition.

Theorem 2.1 Principle of Superposition: The response $y(t)$ of a linear system due to several inputs $x_1(t), x_2(t), \ldots, x_N(t)$ acting simultaneously is equal to the sum of the responses of each input acting alone, that is, if $y_i(t)$ is the response due to the input $x_i(t)$, then $y(t) = \sum_{i=1}^{N} y_i(t)$.

The principle of superposition follows directly from the definition of linearity. Any system which satisfies the principle of superposition is linear.

In the engineering literature, linearity is often defined in terms of the following two properties:

1. Property of *additivity*. If an input $x_1(t)$ produces an output $y_1(t)$, and an input $x_2(t)$ produces an output $y_2(t)$, then an input $x_1(t) + x_2(t)$ produces an output $y_1(t) + y_2(t)$ for all pairs of inputs $x_1(t)$ and $x_2(t)$.
2. Property of *homogeneity*. If an input $x(t)$ produces an output $y(t)$, then an input $a\,x(t)$ produces an output $a\,y(t)$ for any input $x(t)$ and any constant a.

The property of homogeneity does not imply the property of additivity and vice versa.

Linear systems can often be represented by linear differential equations or difference equations. Also, any continuous-time system is linear if its input–output relationship can be described by the ordinary linear differential equation

$$\sum_{i=0}^{n} a_i(t)\frac{\mathrm{d}^i y}{\mathrm{d}t^i} = \sum_{k=0}^{m} b_k(t)\frac{\mathrm{d}^k x}{\mathrm{d}t^k} \tag{2.1}$$

where $y = y(t)$ is the system output and $x = x(t)$ is the system input. The coefficients $a_i(t)$ and $b_k(t)$ depend only upon the independent variable t (e.g., time).

Initially relaxed linear systems possess the following property: Zero input yields zero output. An *incrementally linear system* is one that responds linearly to changes in the input; that is, the difference in the response to any two inputs is a linear function of the difference between the two inputs. Many of the characteristics of such systems can be analyzed using the techniques developed for linear systems. In reality, no physical system can be described exactly by a linear differential or difference equation. Many systems, however, can be represented over a limited operating range, or approximated by such equations.

2.3.6 Memoryless Systems

A system is said to be *memoryless*, also known as zero-memory or instantaneous, if its output for each value of the independent variable is dependent only on the input at that same time. The simplest memoryless system is the *identity system*, whose output is identical to its input. The input–output relationship for the continuous-time identity system is $y(t) = x(t)$, and the corresponding relationship in discrete time is $y(k) = x(k)$. All memoryless systems are causal.

2.3.7 Sensitivity

A first step in the analysis or design of a system is the generation of models for the various elements in the system. The system characteristics are fixed when a finite number of constant parameters have been chosen. The values given to these parameters are called the *nominal values*, and the corresponding characteristics are called the *nominal characteristics*. The accuracy of the model depends on how closely these nominal parameter values approximate the actual parameter values, and how much these parameters deviate from the nominal values during the course of system operation.

The *sensitivity* of a system can be defined as a measure of the amount by which a system characteristic differs from its nominal value when one of its parameters differs from the number chosen as its nominal value. Consider a characteristic C that depends on a parameter p; that is, the mathematical model of the characteristic is $C = C(p)$.

Usually, p is a real or complex quantity representing some identifiable parameter of the system. The sensitivity of $C(p)$ with respect to the parameter p is defined by

$$S_p^{C(p)} = \frac{\mathrm{d}C(p)/C(p)}{\mathrm{d}p/p} = \frac{\mathrm{d}C(p)}{\mathrm{d}p}\frac{p}{C(p)} = \frac{\mathrm{d}\ln C(p)}{\mathrm{d}\ln p} \tag{2.2}$$

where p is regarded as a variable; $S_p^{C(p)}$ is called the *single-parameter relative sensitivity*. The relative change of C due to the variation of N parameters, $p_1, p_2, \ldots, p_N$, is given by

$$\frac{\Delta C}{C} = \sum_{i=1}^{N} S_{p_i}^{C}\frac{\Delta p_i}{p_i} \tag{2.3}$$

If we define the *variation* of a function C due to a relative change in a parameter p by

$$V_p^C = S_p^C\frac{\mathrm{d}p}{p} \tag{2.4}$$

then the *worst-case* variation follows as

$$W^C = \sum_{i=1}^{N}|V_{p_i}^C| \tag{2.5}$$

Another useful figure-of-merit is the so-called *Schoeffler criterion*, or the sum-of-squares sensitivity:

$$S^C = \sum_{i=1}^{N}|V_{p_i}^C|^2 \tag{2.6}$$

In some cases it is useful to define the *semirelative sensitivity*

$$H_p^C = C(p)\,S_p^C \tag{2.7}$$

In general, the sensitivity is a complex number [15].

2.3.8 Optimal and Adaptive Systems

The basic goal of system design is meeting performance specifications. *Performance specifications* are the constraints put on mathematical functions describing system characteristics. They may be stated in any number of ways. The desired (or prescribed) system characteristics specify important properties of the system: speed of response, stability, system accuracy, allowable error, sensitivity, and so on. Satisfactory performance is determined by the application and the characteristics of the particular system.

The system measure of performance is called *performance index*. In more general designs, it is not specified. Instead, the system components, their interconnections, and the parameter values are chosen so that the performance index is maximized or minimized. The systems designed in this way are called *optimal systems*.

In many problems, the performance index is a measure or function of the error between the actual and ideal (or desired) responses. It is formulated in terms of the design parameters chosen, subject to existing physical constraints, to optimize the performance index. The measures of system performance are essentially *criteria of optimality*, and systems are designed taking them into account.

In some systems, certain parameters are either not constant or vary in an unknown manner. It may be desirable to design for the capability of continuously measuring them and changing the system so that the system performance criteria are always satisfied. Systems designed with these objectives are called *adaptive systems*.

Adaptive systems are an important issue in channel equalization, which is the recovery of a signal distorted in transmission through a communication channel with a nonflat magnitude or a nonlinear phase response. When the channel response is unknown, the process of signal recovery is called *blind equalization*.

2.3.9 Invertibility

A system is said to be *invertible* if distinct inputs lead to distinct outputs. By observing its output we can determine its input. An *inverse system* of a given system is one which, when cascaded with the given system, yields an output equal to the input: that is, if $x(t)$ is an input to the cascaded interconnection, its output will be $y(t) = x(t)$.

2.4 LINEAR TIME-INVARIANT SYSTEMS

When a system is both linear and time-invariant, it is called a *linear time-invariant* (LTI) system, and it is amenable to analysis using many techniques. A lumped continuous-time system is linear time-invariant if its input–output relationship can be described by the ordinary linear constant coefficient differential equation

$$\sum_{i=0}^{n} a_i \frac{d^i y}{dt^i} = \sum_{k=0}^{m} b_k \frac{d^k x}{dt^k} \tag{2.8}$$

where $y = y(t)$ is the system output, and $x = x(t)$ is the system input. The coefficients a_i and b_k do not depend on the independent variable t (time). If the equation describes a physical system, then generally $m \le n$, and n is called the *order* of the system.

A lumped discrete-time system is linear time-invariant if its input–output relationship can be described by means of a linear constant coefficients difference equation

$$\sum_{i=0}^{n} a_i\, y(\nu - i) = \sum_{k=0}^{m} b_k\, x(\nu - k) \tag{2.9}$$

where the notation $y(\nu)$ means that the value of the sequence y at the νth sample. Some authors prefer to write (2.9) in the form

$$y_{(\nu)} + \sum_{i=1}^{n} a_i\, y_{(\nu-i)} = \sum_{k=0}^{m} b_k\, x_{(\nu-k)} \tag{2.10}$$

The parameter n is called the order of the difference equation. Discrete-time systems described by (2.10) are known as *recursive systems*, because the output depends on the previous values of the output as well as on the input. Another class of discrete systems is described by

$$y_{(\nu)} = \sum_{k=0}^{m} b_k \, x_{(\nu-k)} \tag{2.11}$$

and is known as *nonrecursive*, because the previous values of the output do not appear.

Differential and difference equations have broad application in the description of physical laws and are useful for relating rates of change of variables and other parameters of a system or its components. In actuality, LTI systems do not exist. All physical systems are nonlinear to some extent. Fortunately, a large percentage of systems can be represented by LTI models over a limited operating range. Many systems always operate within this linear range. Other systems exceed the limits of linear operation, but may be approximated by LTI systems.

2.4.1 The Differential Operator

Differential equations can be presented more compactly by introducing a *differential operator*

$$\mathrm{D} = \frac{\mathrm{d}}{\mathrm{d}t}$$

an ith-*order differential operator*

$$\mathrm{D}^i = \frac{\mathrm{d}^i}{\mathrm{d}t^i}$$

and more generally a *polynomial differential operator* (PDO)

$$A(\mathrm{D}) = \sum_{i=0}^{n} a_i \mathrm{D}^i$$

The differential equation (2.8) can be written as

$$A(\mathrm{D})\, y(t) = B(\mathrm{D})\, x(t)$$

where $A(\mathrm{D})$ maps the function $y(t)$ into the expression $\sum_{i=0}^{n} a_i \mathrm{D}^i y(t)$, and $B(\mathrm{D})\, x(t)$ represents the expression $\sum_{k=0}^{m} b_k \mathrm{D}^k x(t)$.

The polynomial in complex variable s

$$A(s) = \sum_{i=0}^{n} a_i s^i = a_n s^n + a_{n-1} s^{n-1} + \cdots + a_1 s + a_0$$

is called the *characteristic polynomial*. The equation $A(s) = 0$ is called the *characteristic equation*. From the definition of polynomial differential operators it follows:

$$A(\mathrm{D})\,(K\, y(t)) = (K\, A(\mathrm{D}))\, y(t)$$

$$A(\mathrm{D})\,(y_1(t) \pm y_2(t)) = A(\mathrm{D})\, y_1(t) \pm A(\mathrm{D})\, y_2(t)$$

$$A_1(\mathrm{D})\,(A_2(\mathrm{D})\, y(t)) = (A_1(\mathrm{D})\, A_2(\mathrm{D}))\, y(t) = A_2(\mathrm{D})\,(A_1(\mathrm{D})\, y(t))$$

where K is a constant, $y(t)$, $y_1(t)$, and $y_2(t)$ are functions of t, and $A(\mathrm{D})$, $A_1(\mathrm{D})$, and $A_2(\mathrm{D})$ are polynomial differential operators. Addition, $A_1(\mathrm{D}) + A_2(\mathrm{D})$, subtraction, $A_1(\mathrm{D}) - A_2(\mathrm{D})$, and multiplication, $A_1(\mathrm{D})\,A_2(\mathrm{D})$, are performed following the rules for the corresponding operations defined for algebraic polynomials in complex variables. Division $A_1(\mathrm{D})/A_2(\mathrm{D})$ is not defined.

2.4.2 Response of Continuous-Time LTI Systems

Polynomial differential operators are important in analysis of continuous-time linear systems. If a system is specified by its block diagram, we can use polynomial differential operators to simplify derivation of the system input–output description. To specify the problem completely so that the unique solution $y(t)$ can be obtained for a known time function $x(t)$ and the coefficients a_i and b_k, we must specify (1) the interval of time over which a solution is desired and (2) a set of n initial conditions for $y(t)$ and its first $n-1$ derivatives. In many cases the time interval is defined by $0 \le t < +\infty$ and the set of conditions is

$$y(0),\ \left.\frac{\mathrm{d}y}{\mathrm{d}t}\right|_{t=0},\ \ldots,\ \left.\frac{\mathrm{d}^{n-1}y}{\mathrm{d}t^{n-1}}\right|_{t=0}$$

The response of a continuous-time LTI system—that is, the solution $y(t)$—can be divided into two parts, namely, a free response and a forced response. The sum of two responses constitutes the total response.

The *free response* of an LTI system, $y_0(t)$, is the solution of the differential equation when the input $x(t)$ is identically zero. The free response depends only on the n initial conditions; some authors call it the *zero-input response*. The *forced response* of an LTI system, $y_x(t)$, is the solution of the differential equation when all initial conditions are identically zero; some authors call it the *zero-state response*. The forced response depends only on the input $x(t)$. The *total response* of an LTI system is the sum of the free response and the forced response, $y(t) = y_0(t) + y_x(t)$. The total response can be viewed, also, as the sum of the steady-state response and transient response. These terms are often used for specifying system performance. The *steady-state response* of an LTI system is that part of the total response which does not approach zero as time approaches infinity. The *transient response* is that part of the total response which approaches zero as time goes to infinity.

The general procedure for analyzing a system is as follows:

1. Determine the equations for each system component.
2. Choose a model for representing the system (e.g., block diagram).
3. Formulate the system model by appropriately connected the components.
4. Determine the system characteristics.

The direct solution of the system differential equation may be employed to find the total response or its parts, as well as the steady-state response and the transient response.

In the study of systems and differential equations which describe them, a particular family of functions (signals), the *singularity functions*, is used. These functions represent idealizations of physical phenomena and allow us to obtain simple representations of signals and systems. Assuming that signals described by singularity functions excite a system, we define the following special responses:

- *Unit step response* is the output of an LTI system to a unit step input when all initial conditions are zero.
- *Unit ramp response* is the output of an LTI system to a unit ramp input when all initial conditions are zero.
- *Unit impulse response* is the output of an LTI system to a unit impulse input when all initial conditions are zero.

If we know the unit impulse response, $y(t) = y_\delta(t)$, of an initially relaxed causal CTLTI system, then the forced response, $y_x(t)$, of the system, when the input is a known time function $x(t)$, can be written in terms of the *convolution integral*:

$$y_x(t) = \int_{-\infty}^{t} y_\delta(t - \tau)\, x(\tau)\, d\tau$$

This formula directly follows from the principle of superposition and the sifting property of the unit impulse function.

2.4.3 Transient System Specifications

In many cases the desired, or prescribed, system characteristics are specified in terms of time response. Time-domain specifications are customarily defined in terms of unit step function. Typical transient performance specifications are:

1. *Overshoot*—the maximum difference between the transient and steady-state response for a unit step function input (Fig. 2.12).
2. *Delay time*—the time required for the response to a unit step function input to reach 50% of its final value (Fig. 2.13).
3. *Rise time*—the time required for the response to a unit step function input to rise from 10% to 90% of its final value (Fig. 2.13).
4. *Settling time*—the time required for the response to a unit step function input to reach and remain within a specified percentage (frequently 2% or 5%) of its final value (Fig. 2.12).

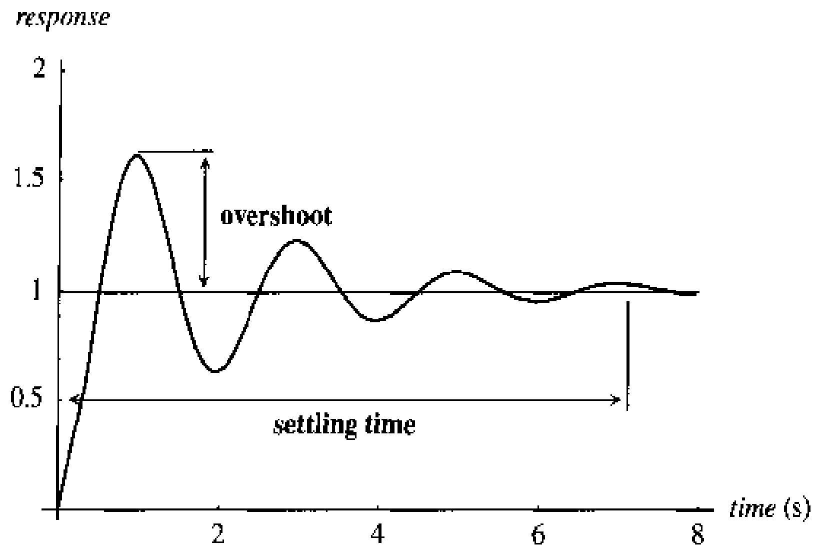


Figure 2.12 Overshoot and settling time.

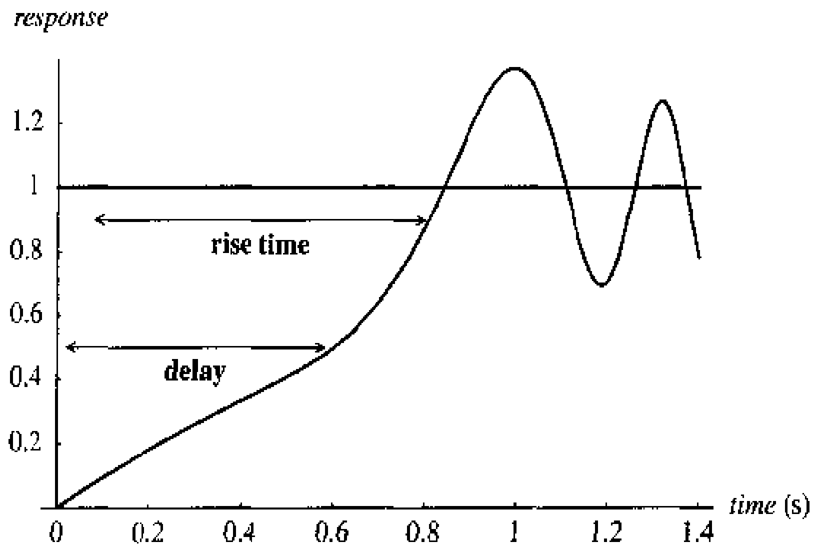


Figure 2.13 Rise time and delay time.

2.4.4 The Shifting Operator for Difference Equations

Difference equations can be presented more compactly by introducing a *shifting operator*

$$Qx(\nu) = x(\nu + 1)$$

an *ith-order shifting operator*

$$Q^i x(\nu) = x(\nu + i)$$

and more generally a polynomial shifting operator

$$A(Q) = \sum_{i=0}^{n} a_i Q^i$$

The difference equation (2.9) can be written as

$$A(\mathrm{Q})\, y(\nu) = B(\mathrm{Q})\, x(\nu)$$

where $A(\mathrm{Q})$ maps the sequence $y(\nu)$ into the expression $\sum_{i=0}^{n} a_i \mathrm{Q}^{-i}\, y(\nu)$, and $B(\mathrm{Q})x(\nu)$ represents the expression $\sum_{k=0}^{m} b_k \mathrm{Q}^{-k} x(\nu)$.

The polynomial in complex quantity z^{-1}

$$A(z) = \sum_{i=0}^{n} a_i z^{-i} = a_0 + a_1 z^{-1} + \cdots + a_{n-1} z^{-n+1} + a_n z^{-n}$$

is called the *characteristic polynomial.* The equation $A(z) = 0$ is called the *characteristic equation.*

From the definition of polynomial shifting operators, it follows that

$$\begin{aligned} A(\mathrm{Q})\ (K\, y(\nu)) &= (K\, A(\mathrm{Q}))\ y(\nu) \\ A(\mathrm{Q})\ (y_1(\nu) \pm y_2(\nu)) &= A(\mathrm{Q})\, y_1(\nu) \pm A(\mathrm{Q})\, y_2(\nu) \\ A_1(\mathrm{Q})\ (A_2(\mathrm{Q})\, y(\nu)) &= (A_1(\mathrm{Q})\, A_2(\mathrm{Q}))\ y(\nu) = A_2(\mathrm{Q})\ (A_1(\mathrm{Q})\, y(\nu)) \end{aligned}$$

where K is a constant, $y(\nu)$, $y_1(\nu)$, and $y_2(\nu)$ are sequences of ν, and $A(\mathrm{Q})$, $A_1(\mathrm{Q})$, and $A_2(\mathrm{Q})$ are polynomial shifting operators. Addition, $A_1(\mathrm{Q}) + A_2(\mathrm{Q})$, subtraction, $A_1(\mathrm{Q}) - A_2(\mathrm{Q})$, and multiplication, $A_1(\mathrm{Q})A_2(\mathrm{Q})$, are performed following the rules for the corresponding operations defined for algebraic polynomials in complex variables. Division, $A_1(\mathrm{Q})/A_2(\mathrm{Q})$, is not defined.

2.4.5 Discrete-Time LTI Systems

Polynomial shifting operators are useful in analysis of discrete-time linear systems. If a system is specified by its block diagram, we can use polynomial shifting operators to derive the system input–output relationship in a simpler way. To specify the problem completely so that a unique solution $y(\nu)$ can be obtained, for a known sequence $x(\nu)$ and the coefficients a_i and b_k, we must specify (1) the range over which a solution is desired and (2) a set of n initial conditions for $y(\nu)$. In many cases the range is defined by $0 \le \nu < +\infty$ and the set of conditions is

$$y(-1), y(-2), \ldots, y(-n)$$

The response of a discrete-time LTI system—that is, the solution $y(\nu)$—can be divided into two parts, namely, a free response and a forced response. The sum of two responses constitutes the total response. The *free response* (or *initial condition response*), $y_0(\nu)$, is the solution of the difference equation when the input $x(\nu)$ is identically zero. The free response depends only on the n initial conditions. Some authors call it the *zero-input response.* The *forced response*, $y_x(\nu)$, is the solution of the difference equation when all initial conditions are identically zero. Some authors call it the *zero-state response*. The forced response depends only on the input $x(\nu)$. The *total response* is the sum of the free response and the forced response, $y(\nu) = y_0(\nu) + y_x(\nu)$. The total response can be viewed, also, as the sum of the steady-state response and transient response. These terms are often used for specifying system performance. The *steady-state response* is that part of the total response which does not approach zero as ν approaches infinity. The *transient response* is that part of the total response which approaches zero as ν goes to infinity.

In the study of discrete-time systems and difference equations which describe them, two special sequences (signals) are used that allow us to obtain simple representations of signals and systems. Assuming that signals described by these sequences excite a system, we define the following special responses:

- *Unit step response* of a discrete-time LTI system is the output of the system to a unit step input with all initial conditions set to zero.
- *Unit impulse response* of a discrete-time LTI system is the output of the system to a unit impulse input with all initial conditions set to zero.

If we know the unit impulse response, $y(\nu) = y_\delta(\nu)$, of an initially relaxed causal discrete-time LTI system, the forced response, $y_x(\nu)$, of the system, then when the input is a known sequence $x(\nu)$, can be written in terms of the *convolution sum*:

$$y_x(\nu) = \sum_{i=-\infty}^{\nu} y_\delta(\nu - i)\, x(i)$$

This formula directly follows from the principle of superposition and the sifting property of the unit impulse sequence.

2.4.6 Properties of LTI Systems

In the preceding sections we showed that the response of LTI systems can be expressed in terms of the unit impulse responses. It implies that the characteristics of an LTI system are completely determined by its impulse response. In the following paragraphs we revisit several important system properties and define them in terms of the impulse responses of LTI systems:

- Memoryless: A continuous-time LTI system is memoryless if $y_\delta(t) = K\delta(t)$, and a discrete-time LTI system is memoryless if $y_\delta(\nu) = K\delta(\nu)$, where K represents a constant. If $K = 1$, the systems become **identity** systems.
- Causal: A continuous-time LTI system is causal if $y_\delta(t) = 0|_{t<0}$, and a discrete-time LTI system is causal if $y_\delta(\nu) = 0|_{\nu<0}$.

- Stability: A sufficient condition for BIBO stability of continuous-time LTI systems is

$$\int_{-\infty}^{\infty} |y_\delta(t)| \mathrm{d}t < \infty$$

which means that the impulse response is absolutely integrable. The corresponding condition for discrete-time LTI systems is

$$\sum_{\nu=-\infty}^{\infty} |y_\delta(\nu)| < \infty$$

which means that the impulse response is absolutely summable.

The **unit step response**, $s(t)$ or $s(\nu)$, can be computed from the unit impulse response by evaluating $s(t) = \int_{-\infty}^{t} y_\delta(\tau)\,\mathrm{d}\tau$, or $s(\nu) = \sum_{i=-\infty}^{\nu} y_\delta(i)$. The above-listed properties can be examined and verified by using the convolution integral and convolution sum as a model of an LTI system.

PROBLEMS

2.1 Construct block diagrams for each of the following equations:

(a) $y = Ax + B$

(b) $y = K_0 + \sum_{i=1}^{n} K_i x_i$

where $A = 100$, $B = 1$, $K_0 = 0.1$, $n = 3$, $K_1 = -2$, $K_2 = \pi$, $K_3 = \sqrt{2}$, x and x_i denote the inputs, and y is the output. Assume that the following basic blocks are available: gain (multiplication by a constant), adder (summing/subtracting point), takeoff point, generator (excitation).

2.2 Consider the following equations in which x_1, x_2, x, represent the inputs and y represents the output of a system:

(a) $a_2 \dfrac{\mathrm{d}^2}{\mathrm{d}t^2} y + a_1 \dfrac{\mathrm{d}}{\mathrm{d}t} y + a_0 y = B \dfrac{\mathrm{d}}{\mathrm{d}t} x_1 + C x_2$

(b) $A \dfrac{\mathrm{d}}{\mathrm{d}t} y + B \int y\,\mathrm{d}t + Cy = Kx$

(c) $y = A \dfrac{\mathrm{d}}{\mathrm{d}t} x + B \int_0^t x\,\mathrm{d}t + Cx$

where $a_2 = 1$, $a_1 = 1$, $a_0 = 4$, $A = -1$, $B = 3$, $C = \dfrac{1}{2}$, $K = \pi^2$, and t designates time. Draw a block diagram for each equation. Assume that the following basic blocks are available: gain (multiplication by a constant), adder (summing/subtracting point), takeoff point, generator (excitation), differentiator (with respect to time t), integrator (with respect to time t).

2.3 Classify the following systems, described by the input-output relation, according to whether they are time-variable or time-invariant:

(a) $y = 100x_1 - 10^{-6} x_2$

(b) $y = \dfrac{1}{1200} \dfrac{\mathrm{d}}{\mathrm{d}t} x$

(c) $12t \int y \, dt = \pi x$

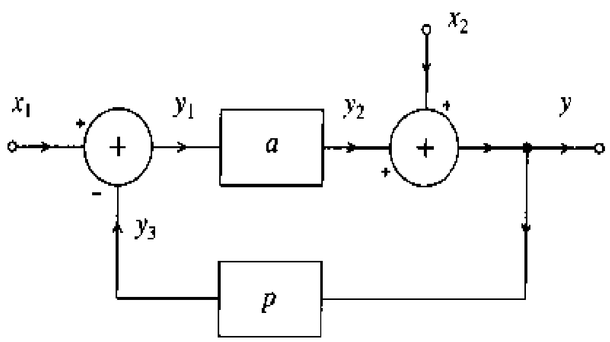
(d) $\frac{d^2}{dt^2} y + (\sin(2\pi t))y = 0$

2.4 A system is described by the convolution integral

$$y(t) = \int_{-\infty}^{t} h(t - \tau)x(\tau)d\tau$$

where x is the input and y is the output. Show that the system is linear.

2.5 Use the *Principle of Superposition* to find the output y of the two-input single-output system



A two-input single-output system.

2.6 Are either of the following systems, described by the input–output relation, causal?

(a) $y(t) = 120x(t - \frac{1}{125})$

(b) $y(t) = 10^{-3}t^2x(t + 0.01)$

2.7 A system is characterized by the following equation

$$y(t) = \sqrt{x(t)^2}$$

where $x(t)$ is the system input and $y(t)$ is the system output. Is this system a linear system?

2.8 The input x and the output y of a system are related by

$$y = A\exp(-kx)$$

where $A = -1000$ and $k = 0.01$. Is this system invertible? If yes, find the inverse system to the given system?

2.9 Consider systems described by the following equations

(1) $y = 2x - \sqrt{3}$

(2) $\frac{d}{dt}y = x$

(3) $y = x(t - 0.1)$

(4) $y(\nu) = 5x(\nu - 1) - 7x(\nu - 2)$, ν is integer.

Classify the systems according to whether they are lumped, linear

(a) continuous-time, time-invariant systems

(b) discrete-time, shift-invariant systems.

2.10 Find the output y of a system described by the differential equation

$$a_1 \frac{\mathrm{d}y}{\mathrm{d}t} + a_0 y = b_1 \frac{\mathrm{d}x}{\mathrm{d}t} + b_0 x, \quad x = u(t)$$

with an initial condition $y(0-) = K$. Assume that $a_1 = 1, a_0 = 2, b_0 = -1, K = 5$, and

(a) $b_1 = 0$

(b) $b_1 = 4$.

2.11 Identify the transient response and the steady-state response if the output of a system is

(a) $y(t) = (2 - 2\exp(-0.12t))\,u(t)$

(b) $y(n) = (2^{-n} + 0.1\sin(n))\,u(n)$.

2.12 Find the transient and steady-state response of a system described by the differential equation

$$a_1 \frac{\mathrm{d}y}{\mathrm{d}t} + a_0 y = b_1 \frac{\mathrm{d}x}{\mathrm{d}t} + b_0 x$$

with an initial condition $y(0-) = K$. Assume that $a_1 = 2$, $a_0 = \frac{1}{2}$, $b_1 = 1$, $b_0 = -1$, $K = 3$, and

(a) $x = 2\sin(8\pi t + \frac{\pi}{4})\,u(t)$

(b) $x = \sqrt{2}\cos(\pi t - \frac{\pi}{3})\,u(t)$

2.13 Show that the unit step response $y_u(t)$, of a causal CTLTI system, is related to the unit impulse response $y_\delta(t)$ by the equation

$$y_u(t) = \int_0^t y_\delta(\tau)\,\mathrm{d}\tau$$

2.14 Show that the unit ramp response $y_r(t)$, of a causal CTLTI system, is related to the unit step response $y_u(t)$ by the equation

$$y_r(t) = \int_0^t y_u(\tau)\,\mathrm{d}\tau$$

2.15 The unit impulse response of a continuous-time system is given by

$$y_\delta(t) = (1 - \exp(-t))\,u(t)$$

Find the system output, $y_x(t)$, if the system is excited by $x(t) = \sin(t)\,u(t)$. Assume that the system is initially relaxed.

2.16 We know that the unit impulse response of an initially relaxed discrete-time system is

$$y_\delta(\nu) = \delta(\nu) - \delta(\nu - 1)$$

Determine the system output, $y_x(\nu)$, for the input $x(\nu) = u(\nu)$.

2.17 Two CTLTI systems are connected in cascade. The unit impulse responses of the systems are $h_1(t) = 2(1 - \exp(-0.2t))\,u(t)$ and $h_2(t) = \cos(\pi t)\,u(t)$. Find the overall unit impulse response of the cascaded systems.

2.18 Two DTLTI systems are connected in parallel. The unit impulse responses of the systems are $h_1(\nu) = 4^{-\nu}u(\nu)$ and $h_2(\nu) = 0.4\cos(2\pi\nu/3)$. Find the overall unit impulse response of the systems.

2.19 Show that the response y of a relaxed causal DTLTI system, excited by a causal input sequence x, is

$$y(n) = \sum_{\nu=0}^{n} x(\nu)h(n-\nu) = \sum_{\nu=0}^{n} h(\nu)x(n-\nu)$$

where h is the unit impulse response of the system.

2.20 The response of a relaxed LTI system, $y_s(t)$, to the unit step input, $x = u(t)$, is

$$y_s(t) = \exp(-0.1t)\,(0.3\sin(2t) - \cos(2t))\,u(t) + u(t)$$

Determine

(a) the peak overshoot,

(b) the delay time,

(c) the rise time, and

(d) the settling time.

2.21 Unit impulse responses of two discrete-time LTI systems are given by

$$h_1(\nu) = \begin{cases} 0, \ \nu < -1 \\ 2^{\nu}, -1 \le \nu \le 1 \\ 0, \ \nu > 1 \end{cases}$$

$$h_2(\nu) = h_1(\nu - 2)$$

Which one of the two systems is causal?

2.22 Find the response of the system described in Problem 2.20 if the system is excited by $x(t) = t\,u(t)$.

2.23 The impulse responses of several systems are given below. For each case determine if the impulse response represents a stable or an unstable system.

(a) $y_\delta(t) = (-2t + 2)\exp(0.7t)u(t)$

(b) $y_\delta(t) = (10\cos(t) + \sqrt{5}\sin(t))\exp(t)u(t)$

(c) $y_\delta(t) = (t + 2)\sin(\pi t)\exp(-0.6t)u(t)$

(d) $y_\delta(t) = \cos(100\pi t)\sin(\pi t)u(t)$.

2.24 The unit step signal is applied to the input of a system and the output is of the form

$$y_u(t) = 2t\sin(4\pi t)\,u(t)$$

Is the system stable or unstable?

2.25 The gain of the system from Problem 2.5 is defined as

$$G = \frac{y}{x_1}, \quad x_2 = 0$$

Find the sensitivity of the gain with respect to the parameter p at the nominal value $p = \frac{1}{2}$, $a = 10$. Compute the gain variation if p changes $\pm 10\%$ about the nominal value.

2.26 The unit impulse response of a DTLTI system is $h(n) = 0.7^n\,u(n)$. Find the response $y(n)$ if the signal $x(n) = u(n) - u(n-4)$ is applied to the input of the system. Assume that the system is initially relaxed.

2.27 A DTLTI system is described by the equation

$$y(n) = x(n) - ax(n-1)$$

Find the response $y(n)$ to the input $x(n) = (\sin(\pi n/3) + \sin(\pi n/21))\, u(n)$, for

(a) $a = 0.1$

(b) $a = 0.9$

(c) $a = 10$.

Assume that the system is initially relaxed.

MATLAB EXERCISES

2.1 Write a MATLAB script that computes and plots the output $y(t)$, $0 < t < 5$, from Problem 2.4 for $h(t) = \exp(-t)\, u(t)$, and sinusoidal input $x(t) = \sin(\omega t)\, u(t)$, if the angular frequency takes a value (a) $\omega = 0.5$, (b) $\omega = 1$, (c) $\omega = 5$, (d) $\omega = 10^3$, (e) $\omega = 10^6$.

2.2 Use MATLAB to find the response from Problem 2.10 and Problem 2.12.

2.3 Use MATLAB to compute the response from Problem 2.15 for $0 \le t \le 5$.

2.4 Write a MATLAB script that computes (a) the peak overshoot, (b) the delay time, (c) the rise time, and (d) the settling time, from Problem 2.20.

2.5 Write a MATLAB script that computes and plots the response from Problem 2.26 for $0 \le n \le 100$.

2.6 A DTLTI system is described by the equation

$$y(n) = \frac{1}{M+1} \sum_{\nu=0}^{M} x(n-\nu)$$

Write a MATLAB script that computes and plots the response $y(n)$, $0 \le n \le 10$, to the input $x(n) = 0.6^{-n}\, u(n)$, for

(a) $M = 3$

(b) $M = 7$.

Assume that the system is initially relaxed.

2.7 Automate the computation of Problem 2.27 in MATLAB. Plot the response and the input on the same diagram.

MATHEMATICA EXERCISES

2.1 Analyze the system from Problem 2.5 symbolically in *Mathematica*.

(a) Write a notebook that sets the analysis equations and derives the output.

(b) Find the gain and the sensitivity from Problem 2.25.

(c) Determine the unit impulse response for the input x_1.

(d) Determine the unit step response for the input x_2.

2.2 Use *Mathematica* to find the response of each system from Problem 2.23 if the following signals are applied at the system input:

(a) $x(t) = -4u(t)$

(b) $x(t) = 0.6\exp(0.25t)u(t)$

(c) $x(t) = \sqrt{2}\cos(\pi t)u(t)$

(d) $x(t) = 2\sin(2\pi(t-5))u(t-5)$.

Assume that the systems are initially relaxed.

2.3 Automate response computation from Problem 2.10 and Problem 2.12 in *Mathematica*.

2.4 Write a *Mathematica* code that computes (a) the peak overshoot, (b) the delay time, (c) the rise time, and (d) the settling time, from Problem 2.20.

2.5 Can you find the inverse system from Problem 2.8 in *Mathematica*?

2.6 Use *Mathematica* to compute the response from Problem 2.15.

2.7 Write a *Mathematica* code that computes and plots the response from Problem 2.26.