5.55 Show that an impulse response matrix $\mathbf{H}(t, \tau)$ is realizable by a finite dimensional dynamical system if and only if there exist two continuous matrices $\mathbf{P}(t)$ and $\mathbf{Q}(\tau)$ such that

$$\mathbf{H}(t, \tau) = \mathbf{P}(t)\mathbf{Q}(\tau) \quad \text{for all } t, \tau$$

Hint: See Section 2.8.

5.56 The impulse response matrix (for a fixed T) of the system

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} + \mathbf{B}(t)\mathbf{v}$$
$$\mathbf{y} = \mathbf{C}(t)\mathbf{x}$$

is given by

$$\mathbf{H}(T, \tau) = \mathbf{C}(T)\boldsymbol{\phi}(T, \tau)\mathbf{B}(\tau) \qquad T \geqslant \tau$$
$$= 0 \qquad T < \tau$$

(*a*) Show that the impulse response matrix of the adjoint system

$$\dot{\boldsymbol{\alpha}} = -\boldsymbol{\alpha}\mathbf{A}(t) + \mathbf{v}\mathbf{C}(t) \qquad (\boldsymbol{\alpha} \text{ and } \mathbf{q} \text{ are row vectors})$$
$$\mathbf{q} = \boldsymbol{\alpha}\mathbf{B}(t)$$

is given by

$$\mathbf{H}^*(t, T) = \mathbf{C}(T)\boldsymbol{\phi}(T, t)\mathbf{B}(t) \qquad t \geqslant T$$
$$= 0 \qquad t < T$$

(*b*) If the change in variable $t_1 = T - t$ is made for the adjoint system, and the change in variable $\tau_1 = T - \tau$ is made for the original system, show that

$$\mathbf{H}^*(t_1, 0) = \mathbf{C}(T)\boldsymbol{\phi}(T, T - t_1)\mathbf{B}(T - t_1) \qquad t_1 \geqslant 0$$
$$= [0] \qquad t_1 < 0$$

and

$$\mathbf{H}(T, T - \tau_1) = \mathbf{C}(T)\boldsymbol{\phi}(T, T - \tau_1)\mathbf{B}(T - \tau_1) \qquad \tau_1 \geqslant 0$$
$$= [0] \qquad \tau_1 < 0$$

(*c*) From the result of part *b*, show that observation of the response of the modified adjoint system over the t_1 axis (the running variable of the simulator) is identical with observing the cross-plot at time T of the response of the original system over the τ_1 axis. In addition, show that an impulse placed on one of the inputs to the adjoint system produces a *row* of the desired impulse response matrix.

(*d*) Show that the result of the preceding parts proves that the complete modified adjoint system is obtained by interchanging the inputs and outputs of the original system and making the change in variable $t = T - t_1$.

6

State Variables and Linear Discrete Systems

6.1 INTRODUCTION

The state variable viewpoint is applied to linear discrete time processes in this chapter. Much of the effort is directed toward sampled-data systems. It is also shown by example that the state variable approach is quite useful in dealing with linear sequential systems, which naturally arise out of the theory of coding. Viewed in this fashion, the state variable approach is a unifying concept, as both continuous and discrete systems fall within its general framework.

The theory of linear discrete time systems follows the theory of linear continuous systems closely. Therefore, much of what is said in this chapter is based on the preceding chapters. The similarity between the theory of linear continuous systems and the theory of linear discrete systems is indicated.

6.2 SIMULATION DIAGRAMS

The basic building blocks required to construct a block diagram of a system described by linear difference equations are the adder, the amplifier, and the unit delay. The adder and the amplifier are the same blocks that were used for continuous systems, and the unit delay for difference equations is somewhat analogous to the integrator for differential equations. It is shown in Fig. 6.2-1. The input to the unit delay is $y(kT + T)$, and the corresponding output is $y(kT)$. Thus the input to the unit delay appears at its output one period later, or delayed

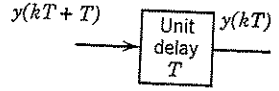


Fig. 6.2-1

by T. It should be noted that unit delays of the order generally required in control systems are quite difficult to obtain in practice, and seldom is a system actually simulated in real time by this method. For sequential circuits, where the variables are either binary in nature or take on discrete values, this type of simulation can be performed in real time.

The approach used to generate a block diagram of a linear difference equation is to assume that the variable $y(nT + kT)$ is available, and then successively pass this variable through unit delays until $y(kT)$ is obtained. The block diagram then is completed by satisfying the requirements of the difference equation, or "closing the loop."

Example 6.2-1. Find the simulation diagram for the system governed by the difference equation $y(kT + 2T) + ay(kT + T) + by(kT) = v(kT)$.

The first step is to solve for $y(kT + 2T)$, as

$$y(kT + 2T) = v(kT) - ay(kT + T) - by(kT)$$

The terms $y(kT + T)$ and $y(kT)$ are obtained as shown in Fig. 6.2-2*a*. Assuming ideal distortionless delays, a signal which appears at terminal 1 appears at terminal 2 one time period later, and at terminal 3 two time periods later. Similarly, a signal at terminal 2 appears at terminal 3 one time period later. The completed block diagram (Fig. 6.2-2*b*) is obtained by satisfying the requirements of the difference equation.

If the initial conditions are given in terms of $y(0)$ and $y(T)$, then $y(0)$ is the initial signal at the output of the first delay, and $y(T)$ is the initial output of the second delay. After one time period, $y(T)$ appears at the output of the first delay unit. After two time periods, the output of the first delay is $y(2T) = v(0) - ay(T) - by(0)$.

If a comparison is made between Figs. 5.2-2*b* and 6.2-2*b*, it is evident that similar rules hold for constructing block diagrams of difference equations and differential equations. The integrator used in simulating differential equations is analogous to the unit delay used in simulating difference equations.

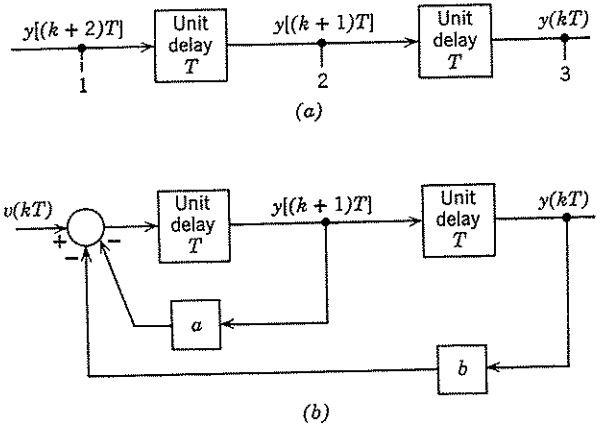


Fig. 6.2-2

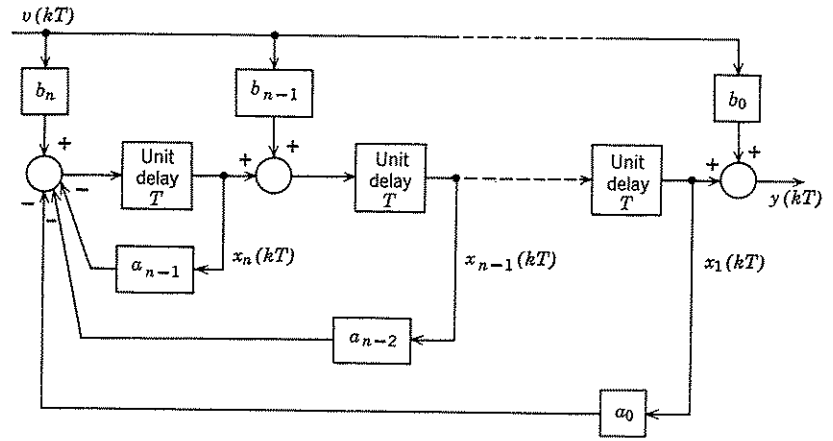


Fig. 6.2-3

Example 6.2-2. Find the simulation diagram for the nth order difference equation

$$y(nT + kT) + \alpha_{n-1}y(nT + kT - T) + \cdots + \alpha_0 y(kT)$$
$$= \beta_n v(nT + kT) + \beta_{n-1}v(nT + kT - T) + \cdots + \beta_0 v(kT)$$

The general simulation diagram for this system is shown in Fig. 6.2-3, by analogy to Fig. 5.5-4. The a's and b's of the block diagram are given by Eqs. 5.5-7 and 5.5-8.

For the specific case of the difference equation

$$y(kT + 3T) + 3y(kT + 2T) + 4y(kT + T) + y(kT) = 2v(kT + 3T) + 3v(kT + 2T) + v(kT + T) + 2v(kT)$$

$$b_0 = \beta_n = 2$$
$$b_1 = \beta_{n-1} - \alpha_{n-1}b_0 = -3$$
$$b_2 = \beta_{n-2} - \alpha_{n-1}b_1 - \alpha_{n-2}b_0 = 2$$
$$b_3 = \beta_{n-3} - \alpha_{n-1}b_2 - \alpha_{n-2}b_1 - \alpha_{n-3}b_0 = 6$$

The simulation diagram for this system is shown in Fig. 6.2-4. A comparison of this

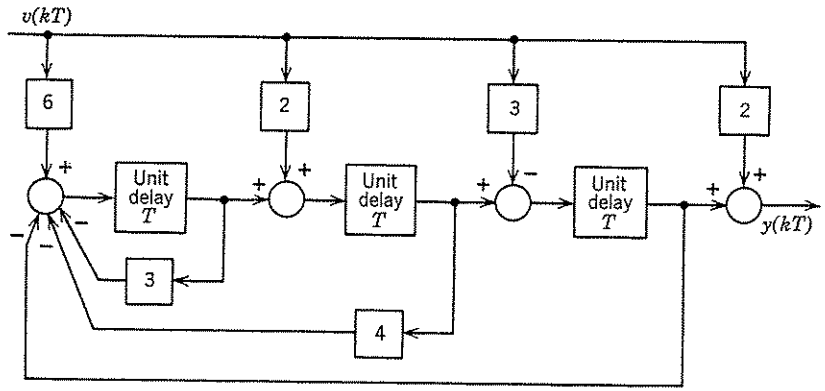


Fig. 6.2-4

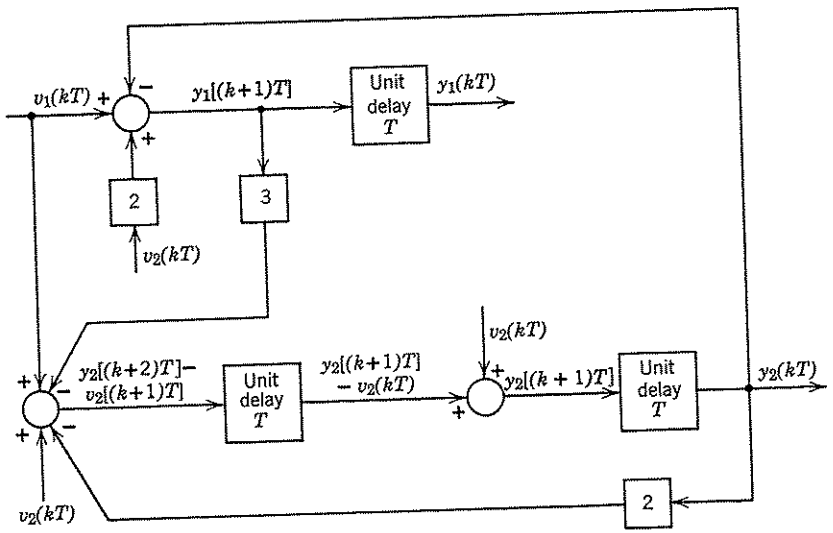


Fig. 6.2-5

diagram with Fig. 5.5-5 shows that the only difference between the two diagrams is that the integrators of Fig. 5.5-5 have been replaced by unit delays.

Example 6.2-3. Find the simulation diagram for the system governed by the difference equations

$$y_1(kT + T) + y_2(kT) = v_1(kT) + 2v_2(kT)$$
$$y_2(kT + 2T) + 3y_1(kT + T) + 2y_2(kT) = v_2(kT + T) + v_2(kT) + v_1(kT)$$

These equations can be rewritten as

$$y_1(kT + T) = v_1(kT) + 2v_2(kT) - y_2(kT)$$
$$y_2(kT + 2T) - v_2(kT + T) = v_2(kT) + v_1(kT) - 3y_1(kT + T) - 2y_2(kT)$$

Using $y_2(kT + 2T) - v_2(kT + T)$ as the input to one delay chain, the block diagram appears as shown in Fig. 6.2-5. The approach is similar to that used for continuous systems.

6.3 TRANSFER FUNCTION MATRICES

The transfer function $H(z)$ for a single input-single output discrete time system is equal to the ratio of the Z transforms of the output and input of the system. For multivariable systems, the transfer function between various input-output terminals is similarly defined. Thus

$$H_{ij}(z) = \frac{Y_i(z)}{V_j(z)}; \quad V_k(z) = 0, \quad k \neq j \tag{6.3-1}$$

where $Y_i(z)$ is the Z transform of the output at terminal i, and $V_j(z)$ is the Z transform of the input at terminal j.

The transfer function matrix is simply the ordered array of these transfer functions, where i denotes the row and j denotes the column in which $H_{ij}(z)$ appears. If the transfer function matrix $\mathbf{H}(z)$ is known, then the output vector transform $\mathbf{Y}(z)$ is given by

$$\mathbf{Y}(z) = \mathbf{H}(z)\mathbf{V}(z) \qquad \text{(assuming zero initial conditions)} \tag{6.3-2}$$

where $\mathbf{V}(z)$ is the column matrix of the Z transform of the input vector $\mathbf{v}(kT)$.

Example 6.3-1. Find the transfer function matrix for the two input-two output system described by the difference equations

$$y_1(kT + 3T) + 6y_1(kT + 2T) + 11y_1(kT + T) + 6y_1(kT) = v_1(kT + T) + v_1(kT) + v_2(kT)$$
$$y_2(kT + 2T) + 5y_2(kT + T) + 6y_2(kT) = v_2(kT + T) + v_2(kT)$$

Taking the Z transform of both sides of these equations, assuming zero initial conditions,

$$(z^3 + 6z^2 + 11z + 6)Y_1(z) = (z + 1)V_1(z) + V_2(z)$$
$$(z^2 + 5z + 6)Y_2(z) = (z + 1)V_2(z)$$

Since $z^3 + 6z^2 + 11z + 6 = (z + 1)(z + 2)(z + 3)$ and $z^2 + 5z + 6 = (z + 2)(z + 3)$,

$$H_{11}(z) = \frac{1}{(z + 2)(z + 3)} \qquad H_{12}(z) = \frac{1}{(z + 1)(z + 2)(z + 3)}$$
$$H_{21}(z) = 0 \qquad H_{22}(z) = \frac{z + 1}{(z + 2)(z + 3)}$$

or

$$\mathbf{H}(z) = \begin{bmatrix} \dfrac{1}{(z + 2)(z + 3)} & \dfrac{1}{(z + 1)(z + 2)(z + 3)} \\ 0 & \dfrac{(z + 1)}{(z + 2)(z + 3)} \end{bmatrix}$$

The transfer function block diagram appears in Fig. 6.3-1.

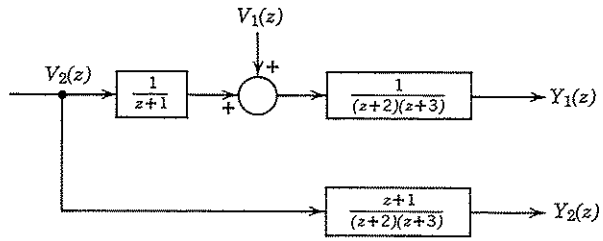


Fig. 6.3-1

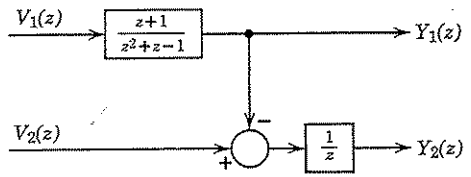


Fig. 6.3-2

Example 6.3-2. Find the transfer function matrix for the system governed by the difference equations

$$y_1(kT+2T)+y_1(kT+T)+y_2(kT+T)=v_1(kT+T)+v_1(kT)+v_2(kT)$$
$$y_2(kT+T)+y_1(kT)=v_2(kT)$$

Transforming both sides of these equations, assuming zero initial conditions,

$$(z^2+z)Y_1(z)+zY_2(z)=(z+1)V_1(z)+V_2(z)$$
$$Y_1(z)+zY_2(z)=V_2(z)$$

Solving for $Y_1(z)$ and $Y_2(z)$,

$$Y_1(z)=\frac{(z+1)}{z^2+z-1}V_1(z) \qquad Y_2(z)=\frac{V_2(z)}{z}-\frac{(z+1)}{z(z^2+z-1)}V_1(z)$$

The transfer function matrix is

$$\mathbf{H}(z)=\begin{bmatrix}\dfrac{z+1}{z^2+z-1} & 0\\ \dfrac{-(z+1)}{z(z^2+z-1)} & \dfrac{1}{z}\end{bmatrix}$$

The transfer function block diagram is shown in Fig. 6.3-2.

The unit delay, represented by the Laplace transform ϵ^{-sT}, corresponds to $1/z$ in the z domain. Thus the integrator, as represented by $1/s$ for continuous systems, has the unit delay as its analog in the diagram of discrete systems. Hence the transfer functions $1/(z+a)$ and $z/(z+a)$ are obtained in a manner similar to that used for continuous systems (see Fig. 6.3-3.).

The same note of caution that was injected into Section 5.3 should be added here. Inspection of only the transfer function matrix may lead to incorrect conclusions about the order of a system, as the transfer function relates only the controllable and observable aspects of the system. It may

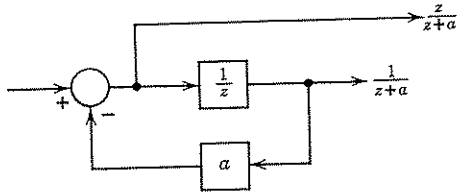


Fig. 6.3-3

mask system properties which would be obtained from the physical system.

6.4 THE CONCEPT OF STATE[1]

The state of a discrete time system can be intuitively defined as the minimum amount of information about the system which is necessary to determine both the output and the future states of the system, if the input function is known. More precisely, a set of variables $\mathbf{x}$ qualifies as a state vector, if two single-valued functions $\mathbf{f}$ and $\mathbf{g}$ can be found such that

$$\mathbf{x}(kT+T)=\mathbf{f}[\mathbf{x}(kT),\mathbf{v}(kT)]$$
$$\mathbf{y}(kT)=\mathbf{g}[\mathbf{x}(kT),\mathbf{v}(kT)] \qquad (6.4\text{-}1)$$

where $\mathbf{y}(kT)$ = the output vector at time kT, and $\mathbf{v}(kT)$ = the input vector at time kT. It is interesting to note that these requirements for the state of a sequential device were independently determined by Huffman and by Moore while working on related but different problems.[2,3]

In most of the subsequent material, the outputs of the delay elements of the simulation diagram of a system are taken as the state of the system. These outputs provide a convenient and sufficient choice for the state vector.

Example 6.4-1. The capacitor step charger of Fig. 6.4-1 is so designed that the voltage on the capacitor increases in steps each time the input pulse appears.[4] Assume that initially there is no charge on capacitor C_2. The first input pulse charges C_1 through diode D_1 to a voltage V_0. After the input pulse disappears, the charge on C_1 distributes itself between C_1 and C_2 according to the inverse ratio of the capacitances. The second pulse again charges C_1 to a voltage V_0, but the subsequent additional charge which is placed on C_2 is less than that from the first pulse, owing to the previous charge left on C_2. The input pulses are continued, and it is expected that the voltage changes appearing on capacitor C_2 diminish asymptotically. Evaluate these changes and verify that the voltage on C_2 is a state variable.

If $x(k)$ is the voltage on capacitor C_2 after the kth input pulse, then the difference equation for $x(k)$ is given by the principle of conservation of charge as $C_1V_0+C_2x(k)=(C_1+C_2)x(k+1)$ or

$$x(k+1)-\left(\frac{C_2}{C_1+C_2}\right)x(k)=\frac{C_1V_0}{C_1+C_2}$$

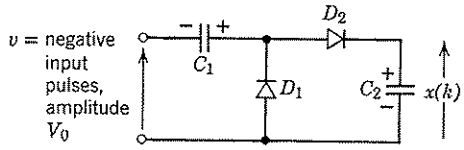


Fig. 6.4-1

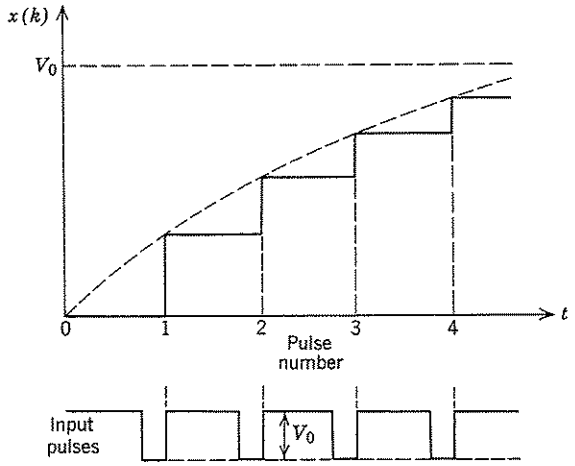


Fig. 6.4-2

Boolean Identities

And	Or
$0 \cdot 0 = 0$	$0 + 0 = 0$
$0 \cdot 1 = 0$	$0 + 1 = 1$
$1 \cdot 0 = 0$	$1 + 0 = 1$
$1 \cdot 1 = 1$	$1 + 1 = 1$
Complement	
$\bar{0} = 1$	$\bar{1} = 0$

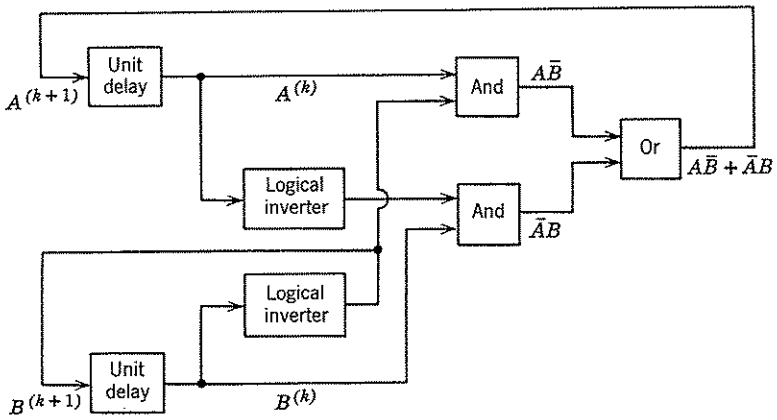


Fig. 6.4-3

Solving this equation in the standard fashion indicated in Section 2.12 yields

$$x(k) = V_0\left[1 - \left(\frac{C_2}{C_1 + C_2}\right)^k\right]$$

Therefore the capacitor voltage increases in discrete steps as shown in Fig. 6.4-2. The step heights are given by

$$x(k+1) - x(k) = V_0\left(\frac{C_1}{C_1 + C_2}\right)\left(\frac{C_2}{C_1 + C_2}\right)^k$$

The state of this circuit is given by the capacitor voltage $x(k)$. Note that this state takes on only discrete values. The next state of the system, $x(k+1)$, is uniquely determined by the present state of the system and the present input. The general state equations for this system are given by Eqs. 6.4-1. In this case, both **f** and **g** are linear, single-valued, scalar functions.

Example 6.4-2.[5] A modulo 4 counter is designed so that it cycles through the counts 00, 01, 10, 11, 00, ... The circuit for this device is shown in Fig. 6.4-3, and the reader familiar with sequential circuits can verify that this circuit does indeed cycle through these counts. The outputs of the delay elements represent the state of the system, as well as the outputs of the system. Write the state equations.

The logical equations for this counter are normally written in the form

$$A^{(k+1)T} = (A\bar{B} + \bar{A}B)^{kT}$$
$$B^{(k+1)} = (\bar{B})^{kT}$$

where the superscripts denote time instants and are not exponents. This is an autonomous sequential circuit, and the general state equations for such a circuit are

$$\mathbf{x}[(k+1)T] = \mathbf{f}[\mathbf{x}(kT)] \qquad \mathbf{x} = \text{state (output of delay elements)}$$
$$\mathbf{y}[(kT)] = \mathbf{g}[\mathbf{x}(kT)] \qquad \mathbf{y} = \text{output (outputs of delay elements)}$$

6.5 MATRIX REPRESENTATION OF LINEAR STATE EQUATIONS

The general form of the state equations for a multivariable discrete time system were given by Eq. 6.4-1 as

$$\begin{aligned} \mathbf{x}[(k+1)T] &= \mathbf{f}[\mathbf{x}(kT), \mathbf{v}(kT)] \\ \mathbf{y}(kT) &= \mathbf{g}[\mathbf{x}(kT), \mathbf{v}(kT)] \end{aligned} \tag{6.5-1}$$

If the system is linear, then Eq. 6.5-1 can be written as the set of linear vector-matrix difference equations

$$\begin{aligned} \mathbf{x}[(k+1)T] &= \mathbf{A}(kT)\mathbf{x}(kT) + \mathbf{B}(kT)\mathbf{v}(kT) \\ \mathbf{y}(kT) &= \mathbf{C}(kT)\mathbf{x}(kT) + \mathbf{D}(kT)\mathbf{v}(kT) \end{aligned} \tag{6.5-2}$$

$\mathbf{A}(kT)$, $\mathbf{B}(kT)$, $\mathbf{C}(kT)$, and $\mathbf{D}(kT)$ have been indicated as time-varying matrices. If the system is nontime-varying, these matrices can be written

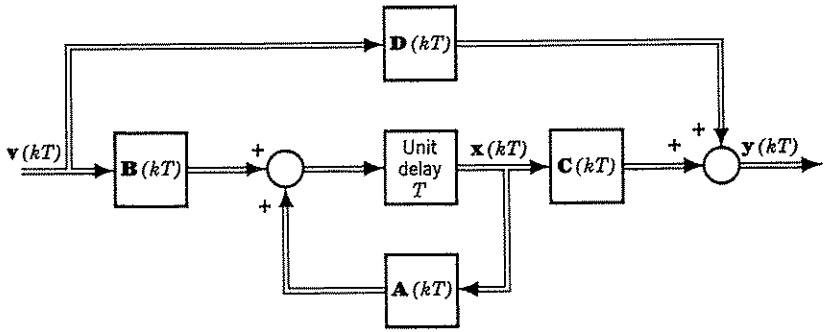


Fig. 6.5-1

as the constant matrices **A**, **B**, **C**, and **D**. The general block diagram, similar to Fig. 5.5-1, is shown in Fig. 6.5-1.

For a system described by a set of nth order difference equations, the form shown in Eq. 6.5-2 can be obtained by writing the given equations as a set of first order difference equations.

Example 6.5-1. Express the second order difference equation

$$y[(k+2)T] + ay[(k+1)T] + by(kT) = v(kT)$$

represented in Fig. 6.2-2*b* as a set of first order difference equations.

Let

$$y(kT) = x_1(kT)$$

$$y[(k+1)T = x_1[(k+1)T] = x_2(kT)$$

Then

$$\begin{bmatrix} x_1(kT+T) \\ x_2(kT+T) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -b & -a \end{bmatrix} \begin{bmatrix} x_1(kT) \\ x_2(kT) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} v(kT)$$

$$y(kT) = [1 \quad 0] \begin{bmatrix} x_1(kT) \\ x_2(kT) \end{bmatrix}$$

The variables $x_1(kT)$ and $x_2(kT)$ are the outputs of the delay elements, and they represent the state of the system. These equations are of the general form of Eq. 6.5-2, where

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -b & -a \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{C} = [1 \quad 0], \quad \mathbf{D} = [0]$$

Example 6.5-2. The general form for an nth order difference equation is given in Example 6.2-2. The simulation diagram appears in Fig. 6.2-3. Find the **A**, **B**, **C**, and **D** matrices.

Analogous to the continuous case, the **A**, **B**, **C**, **D** matrices for this system are given by Eq. 5.5-9.

Linear Binary Sequential Networks†

An interesting application of the state variable approach is the analysis of linear binary sequential networks. A linear binary sequential network consists of pure time delays and modulo 2 summing junctions. A modulo 2 summing junction is an exclusive-OR function having the logical equation $f = (x_1 + x_2)\overline{x_1 x_2} = \bar{x}_1 x_2 + x_1 \bar{x}_2$. The output of such a summer is zero if the two inputs x_1 and x_2 are the same, and one if the inputs are different. Although this discussion is limited to modulo 2 networks, the same type of analysis can be used for modulo p networks.[6–8] These linear modular sequential networks have found a limited application in error-correcting codes, computer circuits, and in certain types of radar systems.[9]

For the purpose of illustrating the use of the state variable approach in the analysis of linear binary sequential networks, consider the network of Fig. 6.5-2. This network consists of four unit delays (a four-element shift register) and a single modulo 2 summing junction (a logical exclusive-OR circuit). The equations of this network are

$$x_1(k+1) = x_2(k)$$
$$x_2(k+1) = x_3(k)$$
$$x_3(k+1) = x_4(k)$$
$$x_4(k+1) = x_4(k) \oplus x_3(k) \oplus x_1(k)$$

In matrix form,

$$\mathbf{x}(k+1) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix} \mathbf{x}(k) = \mathbf{A}\mathbf{x}(k) \bmod 2$$

The sequence of states through which this network will pass is $\mathbf{x}(0)$, $\mathbf{A}\mathbf{x}(0)$, $\mathbf{A}^2\mathbf{x}(0)$, ... If **A** is nonsingular, then each state $\mathbf{x}(k)$ has a unique preceding state $\mathbf{x}(k-1)$. For a mod 2 network, the determinant $|\mathbf{A}|$ is either one or zero. In this particular example $|\mathbf{A}| = 1$, so that **A** is

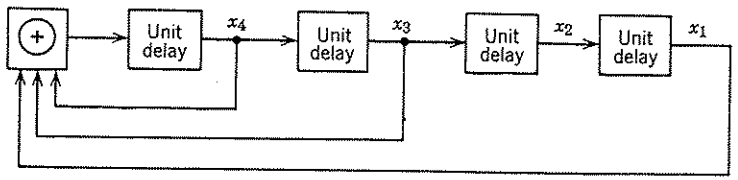


Fig. 6.5-2

† Readers completely unfamiliar with sequential networks should omit this section.

nonsingular. Hence the inverse $\mathbf{A}^{-1}$ does exist.

$$\mathbf{A}^{-1} = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

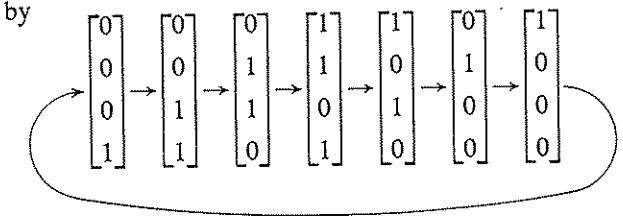
Since there are four state variables, each of which can have the value zero or one, there are $2^4 = 16$ possible states of the system. Therefore the state sequence of an initial state is either periodic with period $P \leq 15$ or goes to an equilibrium state, where the equilibrium state is defined by $\mathbf{x}(k+1) = \mathbf{x}(k)$. The trivial case where all the state variables are zero is called the *null* state. All other equilibrium states are called *finite* equilibrium states.

Consider the case in which the system reaches an equilibrium state. It is assumed that $\mathbf{A}$ is nonsingular, such that $\mathbf{A}^k$ is not equal to the null matrix. This removes the possibility of a nonzero initial state going to the null state. For the finite equilibrium case $\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) = \mathbf{x}(k)$. Since $(\mathbf{A} - \mathbf{I})\mathbf{x}(k) = \mathbf{0}$, the existence of an equilibrium requires that the characteristic equation $|\lambda\mathbf{I} - \mathbf{A}| = 0$ have at least one unity root. For the network of Fig. 6.5-2, the characteristic equation is $\lambda^4 + \lambda^3 + \lambda^2 + 1 = 0 \bmod 2$. This can be factored into $(\lambda + 1)(\lambda^3 + \lambda + 1) = 0 \bmod 2$. Therefore the characteristic equation has one unity root, and there exist initial states for which this network goes to an equilibrium state. One of these initial states is $\mathbf{x}(0) = (1, 1, 1, 1)$.

When the network has a periodic state sequence of period P, then

$$\mathbf{x}(k+P) = \mathbf{A}^P\mathbf{x}(k) = \mathbf{x}(k) \bmod 2$$

For this condition, $|\lambda\mathbf{I} - \mathbf{A}^P| = 0 \bmod 2$, must have at least one unity root. The integer P can be found by determining the smallest integer such that $\mathbf{A}^P = \mathbf{I} \bmod 2$. If the characteristic polynomial $p(\lambda)$ divides $\lambda^P - 1$ without remainder, then $\lambda^P - 1 = p(\lambda)q(\lambda)$, where $q(\lambda)$ is the quotient polynomial. Therefore $\mathbf{A}^P - \mathbf{I} = \mathbf{P}(\mathbf{A})\mathbf{Q}(\mathbf{A}) = [0] \bmod 2$, and $\mathbf{A}^P = \mathbf{I} \bmod 2$. Hence the technique to determine the length of the minimum periodic sequence of a network is to determine the smallest integer P such that $p(\lambda)$ divides $\lambda^P - 1$ without remainder. For this particular problem, the length of the minimum periodic sequence is seven. A typical sequence is given by



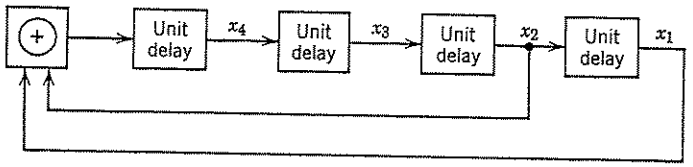


Fig. 6.5-3

The class of networks of order n, whose minimum periodic sequence is equal to $2^n - 1$, are known as *maximal-period* networks. The necessary and sufficient condition for a network to have only maximal-period sequences is that the characteristic polynomial be irreducible and not a divisor for $\lambda^k - 1$, $k < 2^n - 1$. (An irreducible polynomial $f(\lambda)$ is one which cannot be factored into the form $h(\lambda)g(\lambda)$, except for the trivial factorization when $g(\lambda)$ is a constant.) The network of Fig. 6.5-3 has only maximal-period sequences. The equations of this network are

$$\mathbf{x}(k+1) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix} \mathbf{x}(k)$$

$$p(\lambda) = |\lambda\mathbf{I} - \mathbf{A}| \bmod 2$$

$$= \lambda^4 + \lambda^3 + 1 = 0 \bmod 2$$

Since $\lambda^4 \oplus \lambda^3 \oplus 1$ is an irreducible polynomial, and since it is a prime factor of $\lambda^{15} - 1$, the minimum length of the periodic sequence is fifteen. The prime factors of $\lambda^{15} - 1$ are

$$\lambda^{15} - 1 = (\lambda^4 + \lambda^3 + 1)(\lambda^4 + \lambda + 1)(\lambda^4 + \lambda^3 + \lambda^2 + \lambda + 1) \times (\lambda^2 + \lambda + 1)(\lambda - 1)$$

Since the cycle structure of such a network is completely determined by the characteristic polynomial (assuming that the network matrix $\mathbf{A}$ is nonsingular and that the characteristic polynomial is also the minimum polynomial of the network), it is possible to synthesize these linear sequential networks analogously to ordinary lumped element network synthesis. A set of synthesis procedures is given in Reference 6.

State Equations-Partial Fractions Technique

The partial fractions technique for deriving the state equations for a linear, fixed, discrete time process follows the same procedure as that used for continuous systems. For a single input-output system with

transfer function $H(z)$, the transform of the output $Y(z)$ is given by $Y(z) = H(z)V(z)$. If the denominator polynomial of the transfer function $H(z)$ has distinct roots $\lambda_1, \lambda_2, \ldots, \lambda_n$, and if the order of the numerator polynomial of $H(z)$ is less than the order of its denominator polynomial, then $H(z)$ can be written as the partial fraction expansion

$$H(z) = \sum_{i=1}^{n} \frac{c_i}{z - \lambda_i}$$

The output $Y(z)$ is given by

$$Y(z) = \sum_{i=1}^{n} \frac{c_i}{z - \lambda_i} V(z)$$

Therefore $y(kT)$ can be written as a sum of terms of the form

$$y(kT) = \sum_{i=1}^{n} c_i x_i(kT)$$

where the $x_i(kT)$ must satisfy the first order difference equation

$$x_i[(k+1)T] = \lambda_i x_i(kT) + v(kT)$$

The state equations for the system can then be written in the form

$$\begin{bmatrix} x_1[(k+1)T] \\ x_2[(k+1)T] \\ \cdot \\ \cdot \\ \cdot \\ x_n[(k+1)T] \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} x_1(kT) \\ x_2(kT) \\ \cdot \\ \cdot \\ \cdot \\ x_n(kT) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix} v(kT)$$

$$[y(kT)] = [c_1 \quad c_2 \quad \cdots \quad c_n] \begin{bmatrix} x_1(kT) \\ x_2(kT) \\ \cdot \\ \cdot \\ \cdot \\ x_n(kT) \end{bmatrix} \tag{6.5-3}$$

or

$$\begin{aligned} \mathbf{x}[(k+1)T] &= \mathbf{\Lambda x}(kT) + \mathbf{B}v(kT) \\ \mathbf{y}(kT) &= \mathbf{C x}(kT) \end{aligned} \tag{6.5-4}$$

where the matrices $\mathbf{\Lambda}$, $\mathbf{B}$, and $\mathbf{C}$ are defined by the equivalence of Eqs. 6.5-3 and 6.5-4. Note that the symbol $\mathbf{\Lambda}$ is used in place of $\mathbf{A}$, to denote that $\mathbf{A}$ is diagonal. Also note $\mathbf{B} = \mathbf{b} = \mathbf{1}$, a vector.

This form of the state equations is of particular importance when dealing with concepts and proofs, since the diagonal form enables one to

make concise statements about the properties of the system. This form is particularly convenient when dealing with forcing functions, as the corresponding state transition matrix is also of diagonal form. However, there does arise a computational difficulty, because the transfer function $H(z)$ gives no information about the initial conditions of the system. In fact, to compute the initial conditions on the x_i's in this form, one must find $\mathbf{y}(0), \mathbf{y}(1), \ldots, \mathbf{y}(n-1)$ and then solve a set of simultaneous equations to find the relationships between the boundary conditions on the y's and those on the state variables of Eq. 6.5-4.

Example 6.5-3. Find the **A**, **B**, and **C** matrices for the sampled-data system of Fig. 6.5-4*a*.

With respect to the sampled input and output, the transfer function $H(z)$ of this system is given by

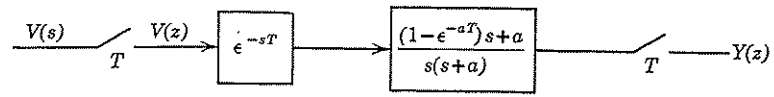
$$\frac{Y(z)}{V(z)} = H(z) = \frac{1}{z-1} - \frac{\epsilon^{-aT}}{z - \epsilon^{-aT}}$$

The output transform $Y(z)$ is then

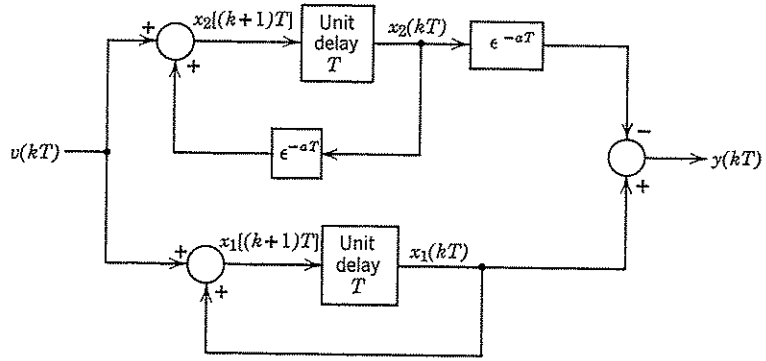
$$Y(z) = H(z)V(z) = \frac{V(z)}{z-1} - \frac{V(z)\epsilon^{-aT}}{z - \epsilon^{-aT}} = X_1(z) + X_2(z)\epsilon^{-aT}$$

where the difference equations for x_1 and x_2 are given by

$$\begin{aligned} x_1[(k+1)T] - x_1(kT) &= v(kT) \\ x_2[(k+1)T] - \epsilon^{-aT}x_2(kT) &= v(kT) \end{aligned}$$



(*a*)



(*b*)

Fig. 6.5-4

The matrix equations for this system are

$$\mathbf{x}(k+1)T = \mathbf{\Lambda x}(kT) + \mathbf{B v}(kT)$$
$$\mathbf{y}(kT) = \mathbf{C x}(kT)$$

where

$$\mathbf{\Lambda} = \begin{bmatrix} 1 & 0 \\ 0 & \epsilon^{-aT} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{C} = [1 \quad -\epsilon^{-aT}]$$

The simulation diagram for this system is shown in Fig. 6.5-4*b*.

For the frequently arising sampled-data case in which the numerator polynomial of $H(z)$ is of the same order as the denominator polynomial of $H(z)$, the **D** matrix is not zero. Since **D** represents a feedthrough term, it can be found by dividing the numerator polynomial by the denominator polynomial, stopping after the first term. The first term represents the feedthrough, or **D** matrix, term. A simple example of this is $H(z) = 1/(1 - z^{-1})$, which represents the z transform corresponding to $1/s$. Using the above criterion for handling this problem,

$$H(z) = \frac{z}{z-1} = 1 + \frac{1}{z-1}$$

This operation is equivalent to

$$\mathbf{D} = D = \lim_{z \to \infty} H(z)$$

The remaining transfer function $H_1(z) = H(z) - D$ can be handled in the same manner as previously described. Actually, once D is found, there is no need to find $H_1(z)$, since $H_1(z)$ has the same poles and the same residues at these poles as does $H(z)$. Therefore, when the order of the numerator polynomial of $H(z)$ is of the same order as the denominator polynomial of $H(z)$ and the poles of $H(z)$ are of first order (distinct roots of the denominator polynomial), the general state equations are

$$\mathbf{x}[(k+1)T] = \mathbf{\Lambda x}(kT) + \mathbf{B v}(kT)$$
$$\mathbf{y}(kT) = \mathbf{C x}(kT) + \mathbf{D v}(kT) \tag{6.5-5}$$

where $\mathbf{\Lambda}$ is the diagonal matrix whose elements are the characteristic roots $\lambda_1, \lambda_2, \ldots, \lambda_n$ of the denominator polynomial of $\mathbf{H}(z)$, and

$\mathbf{B}$ = column matrix whose elements are equal to one

$\mathbf{C}$ = row matrix all of whose elements are equal to $c_i = (z - \lambda_i)H(z)\big|_{z=\lambda_i}$

$\mathbf{D} = d_0$ = single-element matrix $= \lim_{z \to \infty} H(z)$

A general block diagram is shown in Fig. 6.5-5.

For the case where the roots of the denominator polynomial of $H(z)$ are not distinct, the procedure to be followed is the same as that given for the continuous case in Section 5.5. Rather than repeat the procedure here, an illustrative example is given.

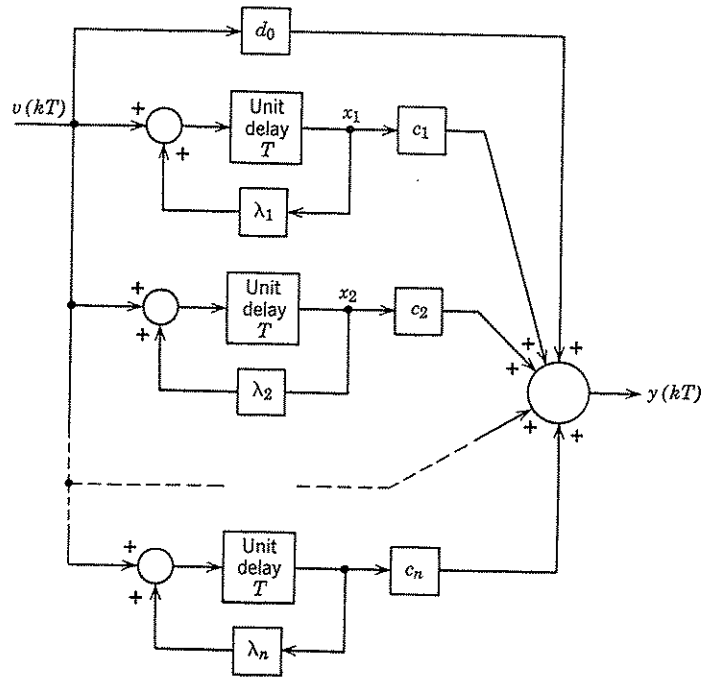


Fig. 6.5-5

Example 6.5-4. Find the **A**, **B**, **C**, and **D** matrices for the system whose transfer function is

$$H(z) = \frac{4z^3 - 12z^2 + 13z - 7}{(z-1)^2(z-2)}$$

This transfer function can be expressed in the partial fraction form

$$H(z) = \frac{c_1}{(z-1)^2} + \frac{c_2}{(z-1)} + \frac{c_3}{(z-2)} + d_0$$

where

$$d_0 = \lim_{z \to \infty} H(z) = 4$$
$$c_1 = (z-1)^2 H(z)\big|_{z=1} = 2$$
$$c_2 = \frac{d}{dz}[(z-1)^2 H(z)]_{z=1} = 1$$
$$c_3 = (z-2)H(z)\big|_{z=2} = 3$$

Therefore the output $Y(z)$ can be written

$$Y(z) = \frac{2V(z)}{(z-1)^2} + \frac{V(z)}{z-1} + \frac{3V(z)}{z-2} + 4V(z) = 2X_1(z) + X_2(z) + 3X_3(z) + 4V(z)$$

Since $X_1(z) = X_2(z)/(z-1)$,

$$x_1[(k+1)T] - x_1(kT) = x_2(kT)$$

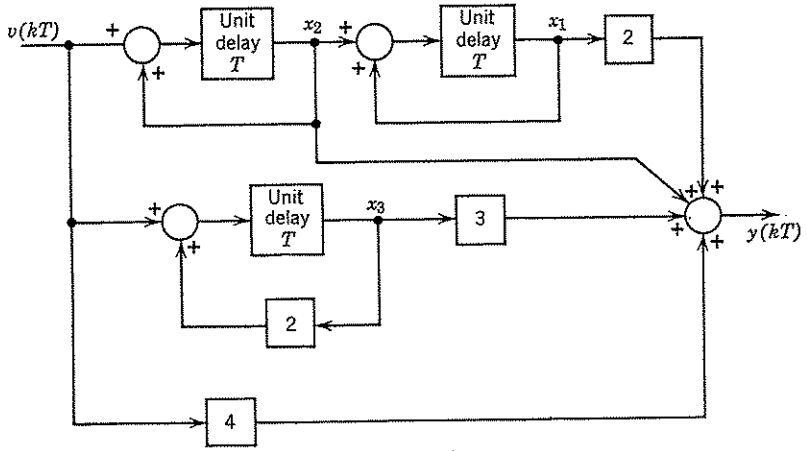


Fig. 6.5-6

The state equations are

$$\mathbf{x}[(k+1)T] = \mathbf{J}\mathbf{x}(kT) + \mathbf{b}v(kT)$$
$$\mathbf{y}(kT) = \mathbf{C}\mathbf{x}(kT) + d_0 v(kT)$$

where

$$\mathbf{J} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{C} = [2 \quad 1 \quad 3], \quad d_0 = 4$$

The simulation diagram for this system is shown in Fig. 6.5-6.

For the multivariable case where there are multiple inputs or outputs, the transfer function matrix $\mathbf{H}(z)$ can be used in a similar manner to find the state equations. However, the approach is not so clear-cut as in the single input-single output case, as there is generally a greater freedom of choice in assigning elements to the **B** and **C** matrices.

Example 6.5-5. Find the **A**, **B**, **C**, and **D** matrices for the system whose transfer function diagram is given in Fig. 6.5-7.

The outputs $Y_1(z)$ and $Y_2(z)$ are given by

$$Y_1(z) = \frac{z}{(z+1)(z+2)} V_1(z) + \frac{1}{z+2} V_2(z)$$

$$Y_2(z) = \frac{z^2}{(z+1)(z+3)} V_1(z) + 4V_2(z)$$

Expanding into partial fractions yields

$$Y_1(z) = -\frac{1}{z+1} V_1(z) + \frac{2}{z+2} V_1(z) + \frac{1}{z+2} V_2(z)$$

$$Y_2(z) = \frac{\frac{1}{2}}{z+1} V_1(z) - \frac{\frac{9}{2}}{z+3} V_2(z) + V_1(z) + 4V_2(z)$$

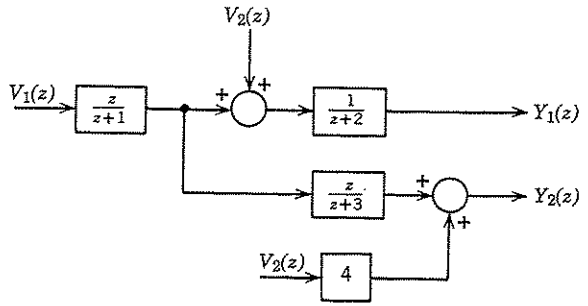


Fig. 6.5-7

Using these equations, Fig. 6.5-8 was drawn. It is readily apparent that the corresponding state equations are

$$\begin{bmatrix} x_1[(k+1)T] \\ x_2[(k+1)T] \\ x_3[(k+1)T] \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} x_1(kT) \\ x_2(kT) \\ x_3(kT) \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ -\frac{9}{2} & 0 \end{bmatrix} \begin{bmatrix} v_1(kT) \\ v_2(kT) \end{bmatrix}$$

$$\begin{bmatrix} y_1(kT) \\ y_2(kT) \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(kT) \\ x_2(kT) \\ x_3(kT) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} v_1(kT) \\ v_2(kT) \end{bmatrix}$$

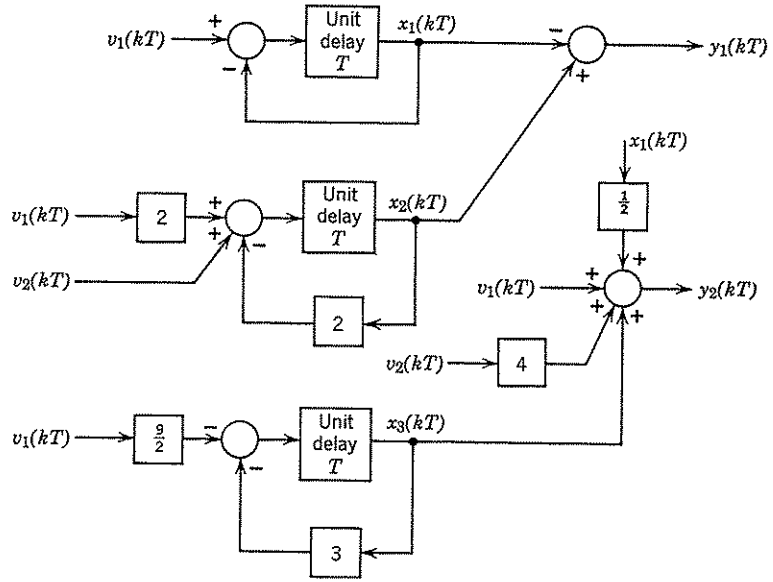


Fig. 6.5-8

Notice that, although the **A** matrix in the preceding example is easily found, the **B** and **C** matrices are by no means unique. Certainly, some of the elements in **B** could be interchanged with some of the elements in **C** and the conditions of the transfer function matrix would still be satisfied. Without any knowledge of the physical properties of the system, either choice is equally valid. As long as all the transfer products $b_{ij}c_{ki}$ remain the same, **B** and **C** can be arranged in any number of ways. The use of the transfer function, or the transfer function matrix, to obtain the state equations of the system is at best a compromise, and it should be used only if the original difference equations are not available.

If the original difference equations are available and **A** can be diagonalized, a diagonal form for **A** can be obtained by use of the modal matrix **M**. Assume then, that the state equations are available in the form of Eq. 6.5-6 below, and that it is desired that a new set of state variables be obtained such that the **A** matrix is a diagonal matrix $\mathbf{\Lambda}$.

$$\begin{aligned}\mathbf{x}[(k+1)T] &= \mathbf{A}\mathbf{x}(kT) + \mathbf{B}\mathbf{v}(kT)\\ \mathbf{y}(kT) &= \mathbf{C}\mathbf{x}(kT) + \mathbf{D}\mathbf{v}(kT)\end{aligned} \tag{6.5-6}$$

Define a new set of state variables **q**, such that

$$\mathbf{x}(kT) = \mathbf{M}\mathbf{q}(kT) \tag{6.5-7}$$

Equation 6.5-6 can be rewritten in terms of these *normal coordinates* as

$$\begin{aligned}\mathbf{M}\mathbf{q}[(k+1)T] &= \mathbf{A}\mathbf{M}\mathbf{q}(kT) + \mathbf{B}\mathbf{v}(kT)\\ \mathbf{y}(kT) &= \mathbf{C}\mathbf{M}\mathbf{q}(kT) + \mathbf{D}\mathbf{v}(kT)\end{aligned}$$

Premultiplying both sides of the first equation by $\mathbf{M}^{-1}$ yields

$$\mathbf{q}[(k+1)T] = \mathbf{M}^{-1}\mathbf{A}\mathbf{M}\mathbf{q}(kT) + \mathbf{M}^{-1}\mathbf{B}\mathbf{v}(kT)$$

However, since $\mathbf{M}^{-1}\mathbf{A}\mathbf{M} = \mathbf{\Lambda}$, vector-matrix state equations in *normal form* are

$$\begin{aligned}\mathbf{q}[(k+1)T] &= \mathbf{\Lambda}\mathbf{q}(kT) + \mathbf{B}_n\mathbf{v}(kT)\\ \mathbf{y}(kT) &= \mathbf{C}_n\mathbf{q}(kT) + \mathbf{D}\mathbf{v}(kT)\end{aligned} \tag{6.5-8}$$

where $\mathbf{B}_n = \mathbf{M}^{-1}\mathbf{B}$, the normal form input matrix, and $\mathbf{C}_n = \mathbf{C}\mathbf{M}$, the normal form output matrix. This is a general procedure, and, since it originates from the difference equations of the system, it is to be preferred over the transfer function matrix approach.

Example 6.5-6. The simulation diagram of Fig. 6.5-9 represents the original system whose transfer function matrix was given in Example 6.5-5. Determine the state equations in normal form.

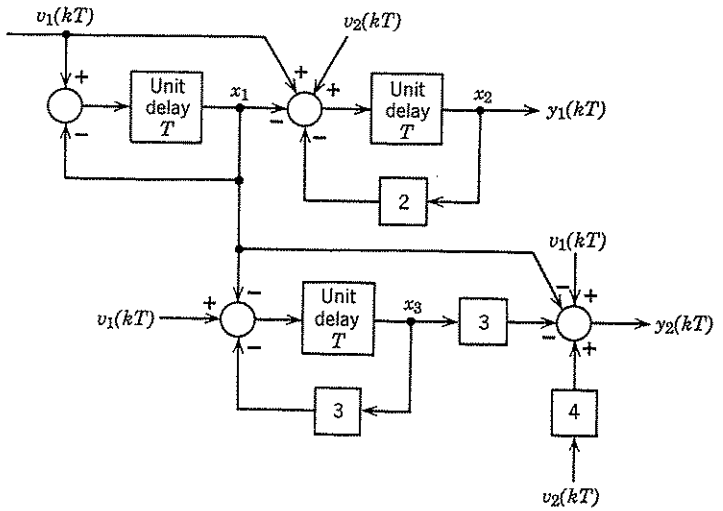


Fig. 6.5-9

The vector-matrix equations for this system are

$$\begin{bmatrix}x_1[(k+1)T]\\ x_2[(k+1)T]\\ x_3[(k+1)T]\end{bmatrix} = \begin{bmatrix}-1 & 0 & 0\\ -1 & -2 & 0\\ -1 & 0 & -3\end{bmatrix}\begin{bmatrix}x_1(kT)\\ x_2(kT)\\ x_3(kT)\end{bmatrix} + \begin{bmatrix}1 & 0\\ 1 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}v_1(kT)\\ v_2(kT)\end{bmatrix}$$

$$\begin{bmatrix}y_1(kT)\\ y_2(kT)\end{bmatrix}\begin{bmatrix}0 & 1 & 0\\ -1 & 0 & -3\end{bmatrix}\begin{bmatrix}x_1(kT)\\ x_2(kT)\\ x_3(kT)\end{bmatrix}\begin{bmatrix}0 & 0\\ 1 & 4\end{bmatrix}\begin{bmatrix}v_1(kT\\ v_2(kT\end{bmatrix}$$

For the A matrix above, the modal matrix M, and its inverse $\mathbf{M}^{-1}$ are given by

$$\mathbf{M} = \begin{bmatrix}2 & 0 & 0\\ -2 & -1 & 0\\ -1 & 0 & 1\end{bmatrix}, \quad \mathbf{M}^{-1} = \begin{bmatrix}\frac{1}{2} & 0 & 0\\ -1 & -1 & 0\\ \frac{1}{2} & 0 & 1\end{bmatrix}$$

Substituting into Eq. 6.5-8, the resulting normal form matrices are

$$\mathbf{\Lambda} = \begin{bmatrix}-1 & 0 & 0\\ 0 & -2 & 0\\ 0 & 0 & -3\end{bmatrix}, \quad \mathbf{B}_n = \begin{bmatrix}\frac{1}{2} & 0\\ -2 & -1\\ \frac{3}{2} & 0\end{bmatrix},$$

$$\mathbf{C}_n = \begin{bmatrix}-2 & -1 & 0\\ 1 & 0 & -3\end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix}0 & 0\\ 1 & 4\end{bmatrix}$$

A check of all the transfer products $b_{ij}c_{ki}$ shows that this set of matrices represents the same system as that of Example 6.5-5.

The new state variables **q** are related to the old state variables **x** by the relationship $\mathbf{q} = \mathbf{M}^{-1}\mathbf{x}$, or

$$q_1 = \tfrac{1}{2}x_1 \qquad q_2 = -(x_1 + x_2) \qquad q_3 = \tfrac{1}{2}x_1 + x_3$$

The use of the relationship $\mathbf{q} = \mathbf{M}^{-1}\mathbf{x}$ removes the previous difficulty of finding the initial conditions on the state variables $\mathbf{q}$ in terms of the known system initial conditions.

Finding the modal matrix may involve no more labor than any of the other methods for finding the state transition matrix. In view of the advantages of having a diagonal $\mathbf{A}$ matrix, the normal form is quite desirable. It is also interesting to note that the mode expansion technique, to be considered next, and the normal form produce the same effect of uncoupling the state equations. In this respect, they are identical approaches but are written in different forms. The form in which the system equations are expressed is frequently one of personal preference and familiarity.

6.6 MODE INTERPRETATION

The concept of expanding the response of a linear fixed system into the sum of responses along the characteristic vectors of the $\mathbf{A}$ matrix can also be applied to discrete systems with distinct characteristic values. The development follows directly from the equations in normal form.

From Eq. 6.5-8,

$$\mathbf{q}(T) = \mathbf{\Lambda}\mathbf{q}(0) + \mathbf{B}_n\mathbf{v}(0) \tag{6.6-1}$$

and

$$\mathbf{q}(2T) = \mathbf{\Lambda}\mathbf{q}(T) + \mathbf{B}_n\mathbf{v}(T) \tag{6.6-2}$$

Substitution of Eq. 6.6-1 into Eq. 6.6-2 yields

$$\mathbf{q}(2T) = \mathbf{\Lambda}^2\mathbf{q}(0) + \mathbf{\Lambda}\mathbf{B}_n\mathbf{v}(0) + \mathbf{B}_n\mathbf{v}(T) \tag{6.6-3}$$

Similarly,

$$\mathbf{q}(3T) = \mathbf{\Lambda}\mathbf{q}(2T) + \mathbf{B}_n\mathbf{v}(2T)$$

and Eq. 6.6-3 give

$$\mathbf{q}(3T) = \mathbf{\Lambda}^3\mathbf{q}(0) + \mathbf{\Lambda}^2\mathbf{B}_n\mathbf{v}(0) + \mathbf{\Lambda}\mathbf{B}_n\mathbf{v}(T) + \mathbf{B}_n\mathbf{v}(2T)$$

Continuation of this procedure leads to

$$\mathbf{q}(kT) = \mathbf{\Lambda}^k\mathbf{q}(0) + \sum_{j=0}^{k-1} \mathbf{\Lambda}^{k-j-1}\mathbf{B}_n\mathbf{v}(jT) \tag{6.6-4}$$

Then, since $\mathbf{q} = \mathbf{M}^{-1}\mathbf{x}$ and $\mathbf{B}_n = \mathbf{M}^{-1}\mathbf{B}$,

$$\mathbf{x}(kT) = \mathbf{M}\mathbf{\Lambda}^k\mathbf{M}^{-1}\mathbf{x}(0) + \sum_{j=0}^{k-1} \mathbf{M}\mathbf{\Lambda}^{k-j-1}\mathbf{M}^{-1}\mathbf{B}\mathbf{v}(jT) \tag{6.6-5}$$

Now, recalling from Section 4.7 that the columns of

$$\mathbf{M} = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_n]$$

form a basis and that the rows of $\mathbf{M}^{-1}$, where

$$\mathbf{M}^{-1} = \begin{bmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{r}_n^T \end{bmatrix} \qquad (\mathbf{r}_i = \text{column vectors})$$

form a reciprocal basis, Eq. 6.6-5 becomes

$$\mathbf{x}(kT) = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_n]\mathbf{\Lambda}^k \begin{bmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{r}_n^T \end{bmatrix} \mathbf{x}(0)$$

$$+ \sum_{j=0}^{k-1} [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_n]\mathbf{\Lambda}^{k-j-1} \begin{bmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{r}_n^T \end{bmatrix} \mathbf{B}\mathbf{v}(jT)$$

This can be rewritten as

$$\mathbf{x}(kT) = \sum_{i=1}^{n} [\langle \mathbf{r}_i, \mathbf{x}(0)\rangle \lambda_i^k + \sum_{j=0}^{k-1} \langle \mathbf{r}_i, \mathbf{B}\mathbf{v}(jT)\rangle \lambda_i^{k-j-1}]\mathbf{u}_i \tag{6.6-6}$$

Equation 6.6-6 is the discrete system analog of Eq. 5.6-10.

The "modes" of the system are given by the terms of Eq. 6.6-6 for $i = 1, 2, \ldots, n$. The mode expansion technique separates or uncouples these modes, so that the response $\mathbf{x}(kT)$ is expressed as a linear weighted sum of the modes. Each mode is directed along the characteristic vector $\mathbf{u}_i$, defined in terms of the state space $x_1, x_2, \ldots, x_n$. For an unforced system, the amount of excitation of each mode is given by the scalar product $\langle \mathbf{r}_i, \mathbf{x}(0)\rangle$. For a forced system, the scalar product $\langle \mathbf{r}_i, \mathbf{B}\mathbf{v}\rangle$ is the amplitude of the forcing function that is coupled to the ith mode.

In effect, the characteristic vectors represent a new coordinate system, such that each mode of the system is directed along one of the coordinate axes. The *normal form* for the state equations performs the same task but expresses it in slightly different form. Thus if the original state equation is

$$\mathbf{x}[(k+1)T] = \mathbf{A}\mathbf{x}(kT) + \mathbf{B}\mathbf{v}(kT)$$

the transformation $\mathbf{q} = \mathbf{M}^{-1}\mathbf{x}$, where $\mathbf{M}$ is the *normalized* modal matrix, results in the new state equation

$$\mathbf{q}[(k+1)T] = \mathbf{\Lambda}\mathbf{q}(kT) + \mathbf{M}^{-1}\mathbf{B}v(kT)$$

In this form the state variables $q_1, q_2, \ldots, q_n$ are uncoupled, since $\mathbf{\Lambda}$ is a diagonal matrix. The state variable q_i is directed along the characteristic vector $\mathbf{u}_i$. The scalar product $\langle \mathbf{r}_i, \mathbf{x}(0)\rangle$ in the mode expansion is simply the ith component of the normal form column vector $\mathbf{q}(0) = \mathbf{M}^{-1}\mathbf{x}(0)$, and the scalar product $\langle \mathbf{r}_i, \mathbf{B}v\rangle$ is the ith component of the column vector $\mathbf{M}^{-1}\mathbf{B}v$.

Example 6.6-1. Analyze the system of Fig. 6.6-1 using both the mode expansion technique and the normal form of the state equations.

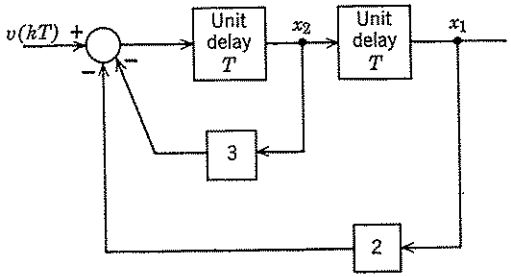


Fig. 6.6-1

The A matrix for this system is

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}$$

The characteristic roots are $\lambda_1 = -2$, and $\lambda_2 = -1$. The normalized characteristic vectors are

$$\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \qquad \mathbf{u}_2 = \begin{bmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}$$

The reciprocal basis is then

$$\mathbf{r}_1 = \begin{bmatrix} 2\sqrt{2} \\ \sqrt{2} \end{bmatrix} \qquad \mathbf{r}_2 = \begin{bmatrix} -\sqrt{5} \\ -\sqrt{5} \end{bmatrix}$$

The normalized modal matrix $\mathbf{M}$ and its inverse $\mathbf{M}^{-1}$ are given by

$$\mathbf{M} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{5} \\ -1/\sqrt{2} & -2/\sqrt{5} \end{bmatrix} \qquad \mathbf{M}^{-1} = \begin{bmatrix} 2\sqrt{2} & \sqrt{2} \\ -\sqrt{5} & -\sqrt{5} \end{bmatrix}$$

Note that the rows of $\mathbf{M}^{-1}$ are the reciprocal basis vectors $\mathbf{r}_i$.

The scalar products $\langle \mathbf{r}_i, \mathbf{x}(0)\rangle$ are

$$\langle \mathbf{r}_1, \mathbf{x}(0)\rangle = 2\sqrt{2}\,x_1(0) + \sqrt{2}\,x_2(0)$$
$$\langle \mathbf{r}_2, \mathbf{x}(0)\rangle = -\sqrt{5}\,x_1(0) - \sqrt{5}\,x_2(0)$$

The initial conditions on the $\mathbf{q}$ vector are

$$\mathbf{q}(0) = \mathbf{M}^{-1}\mathbf{x}(0) = \begin{bmatrix} 2\sqrt{2}\,x_1(0) + \sqrt{2}\,x_2(0) \\ -\sqrt{5}\,x_1(0) - \sqrt{5}\,x_2(0) \end{bmatrix}$$

The forcing function $\mathbf{B}v(kT)$ is the vector $\mathbf{b}$ times the scalar $v(kT)$, where

$$\mathbf{b} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Thus the scalar products $\langle \mathbf{r}_i, \mathbf{B}v(kT)\rangle$ are

$$\langle \mathbf{r}_1, \mathbf{B}v(kT)\rangle = \sqrt{2}\,v(kT)$$
$$\langle \mathbf{r}_2, \mathbf{B}v(kT)\rangle = -\sqrt{5}\,v(kT)$$

The forcing function $\mathbf{B}_n v(kT)$ is

$$B_n v(kT) = \mathbf{M}^{-1}\mathbf{b}v(kT) = \begin{bmatrix} \sqrt{2} \\ -\sqrt{5} \end{bmatrix} v(kT)$$

The general expression for the time response $\mathbf{x}(kT)$ is

$$\mathbf{x}(kT) = [2\sqrt{2}\,x_1(0) + \sqrt{2}\,x_2(0)](-1)^k\mathbf{u}_1 + \sum_{j=0}^{k-1} \sqrt{2}\,(-1)^{(k-j-1)}v(jT)\mathbf{u}_1$$
$$+ [-\sqrt{5}\,x_1(0) - \sqrt{5}\,x_2(0)](-2)^k\mathbf{u}_2 + \sum_{j=0}^{k-1} (-\sqrt{5})(-2)^{(k-j-1)}v(jT)\mathbf{u}_2$$

The general expression for the time response $\mathbf{q}(kT)$ is, from the above and Eq. 6.6-4,

$$\mathbf{q}(kT) = \begin{bmatrix} (-1)^k q_1(0) + \sum_{j=0}^{k-1} \sqrt{2}\,(-1)^{(k-j-1)}v(jT) \\ (-2)^k q_2(0) + \sum_{j=0}^{k-1} (-\sqrt{5})(-2)^{(k-j-1)}v(jT) \end{bmatrix}$$

where $q_1(0) = 2\sqrt{2}\,x_1(0) + \sqrt{2}\,x_2(0)$ and $q_2(0) = -\sqrt{5}\,x_1(0) - \sqrt{5}\,x_2(0)$.

Obviously both methods are equivalent, the $\mathbf{q}$ coordinates of the *normal form* being the normalized characteristic vectors of the *mode expansion*.

6.7 CONTROLLABILITY AND OBSERVABILITY

The controllability and observability concepts presented in Section 5.7 carry over directly to the linear discrete system, so there is little need for further discussion on these points. To restate the principal ideas:

Controllability is a function of the coupling between the inputs to the system and the various modes of the system. If the system equations can

be written with distinct λ's in the normal form

$$\mathbf{q}[(k+1)T] = \mathbf{\Lambda q}(kT) + \mathbf{B}_n\mathbf{v}(kT)$$
$$\mathbf{y}(kT) = \mathbf{C}_n\mathbf{q}(kT) + \mathbf{D}\mathbf{v}(kT)$$

then all the modes are controllable if there are no zero *rows* of $\mathbf{B}_n$. Stated in terms of the mode expansion method, this means that none of the scalar products $\langle \mathbf{r}_i, \mathbf{Bv} \rangle$ vanishes.

Observability is a function of the coupling between the modes of the system and the output of the system. All the modes of the system are observable if there are no zero *columns* of $\mathbf{C}_n$. Alternatively, this requirement could be stated as: The kth mode is not observable if all the scalar products $\langle \mathbf{c}_i, \mathbf{u}_k \rangle = 0$ for all i's, where the vector $\mathbf{c}_i$ constitutes the ith row of the original $\mathbf{C}$ matrix.

For a sampled-data system there is an additional requirement. If the continuous system has a partial fraction expansion which contains the term $\beta/[(s+a)^2+\beta^2]$, and if the sampling interval $T = \pi/\beta$, then the Z transform of this term,

$$\mathscr{Z}\left[\frac{\beta}{(s+a)^2+\beta^2}\right] = \frac{z^{-1}\epsilon^{-aT}\sin\beta T}{1 - 2z^{-1}\epsilon^{-aT}\cos\beta T + z^{-2}\epsilon^{-2aT}} = 0$$

The system may even be unstable, with $a < 0$, but this fact could not be inferred from observations of the output. These are called "hidden oscillations," and they occur when the zeros of the oscillation coincide exactly with the time that the system is sampled.[10-12] In this situation the system is neither completely controllable nor completely observable. Therefore the additional requirement for complete controllability and observability of sampled-data systems is that, if a characteristic root of the continuous system is $-a \pm j\beta$, then $T \neq \pi/\beta$.

6.8 THE STATE TRANSITION MATRIX

The state transition matrix for the linear discrete time system is investigated in this section. Similarly to the continuous case, the state transition matrix is the fundamental matrix of Eq. 6.8-1 below, subject to the condition that $\mathbf{\Phi}[(k_0, k_0)T] = \mathbf{I}$, the unit matrix. Consider, then, the time-varying state difference equation

$$\mathbf{x}[(k+1)T] = \mathbf{A}(kT)\mathbf{x}(kT) \tag{6.8-1}$$

If the initial conditions $\mathbf{x}(k_0T)$ are known, then

$$\mathbf{x}[(k_0+1)T] = \mathbf{A}(k_0T)\mathbf{x}(k_0T)$$

Similarly,

$$\mathbf{x}[(k_0+2)T] = \mathbf{A}[(k_0+1)T]\mathbf{A}(k_0T)\mathbf{x}(k_0T)$$

By a process of iteration, the continued product[13]

$$\mathbf{x}(kT) = \prod_{n=k_0}^{k-1} \mathbf{A}(nT)\mathbf{x}(k_0T) \qquad (k > k_0) \tag{6.8-2}$$

is obtained.

Since the state transition matrix $\mathbf{\Phi}[(k, k_0)T]$ is defined by the relationship

$$\mathbf{x}(kT) = \mathbf{\Phi}[(k, k_0)T]\mathbf{x}(k_0T) \tag{6.8-3}$$

then

$$\mathbf{\Phi}[(k, k_0)T] = \prod_{n=k_0}^{k-1} \mathbf{A}(nT) \qquad (k > k_0) \tag{6.8-4}$$
$$= \mathbf{I} \qquad (k = k_0)$$

This process of obtaining the state transition matrix by iteration is similar to the iterative procedure for computing the matrizant of the analogous continuous system.

For the case where $\mathbf{A}(kT)$ is a constant matrix $\mathbf{A}_0$, the state transition matrix $\mathbf{\Phi}_0[(k, k_0)T]$ is

$$\mathbf{\Phi}_0[(k, k_0)T] = \mathbf{A}_0^{(k-k_0)} \qquad \mathbf{A}(kT) = \mathbf{A}_0, \text{ a constant matrix} \tag{6.8-5}$$

This is analogous to the continuous case, where the solution for a fixed system depends only upon time differences; whereas for a time-variable case the solution depends upon both the time of application of cause and the time of observation of effect.

For the time-varying case where $\mathbf{A}(kT)$ can be written as the sum of two matrices $\mathbf{A}_0$ and $\mathbf{A}_1(kT)$, a perturbation technique can be used to obtain the state transition matrix. This procedure is useful if the time-varying matrix $\mathbf{A}_1(kT)$ represents a small perturbation upon the constant matrix $\mathbf{A}_0$. For this case

$$\mathbf{x}[(k+1)T] = [\mathbf{A}_0 + \mathbf{A}_1(kT)]\mathbf{x}(kT) \tag{6.8-6}$$

This equation can be viewed as a constant system $\mathbf{A}_0$ with a forcing function $\mathbf{v}(kT)$ applied, where $\mathbf{v}(kT) = \mathbf{A}_1(kT)\mathbf{x}(kT)$. Thus

$$\mathbf{x}[(k+1)T] = \mathbf{A}_0\mathbf{x}(kT) + \mathbf{v}(kT), \qquad \mathbf{v}(kT) = \mathbf{A}_1(kT)\mathbf{x}(kT) \tag{6.8-7}$$

The solution $\mathbf{x}(kT)$ for the system of Eq. 6.8-7 can be determined by the same process of iteration used to give Eqs. 6.6-4 and 6.8-2. Thus

$$\mathbf{x}(kT) = \mathbf{A}_0^{(k-k_0)}\mathbf{x}(k_0T) + \sum_{n=k_0}^{k-1} \mathbf{A}_0^{(k-n-1)}\mathbf{v}(nT) \tag{6.8-8}$$

Substituting $\mathbf{v}(nT) = \mathbf{A}_1(nT)\mathbf{x}(nT)$ into Eq. 6.8-8 yields

$$\mathbf{x}(kT) = \mathbf{A}_0^{(k-k_0)}\mathbf{x}(k_0T) + \sum_{n=k_0}^{k-1} \mathbf{A}_0^{(k-n-1)}\mathbf{A}_1(nT)\mathbf{x}(nT) \tag{6.8-9}$$

This is a summation equation, and it can be solved by the usual methods of iteration.

The first iteration is

$$\mathbf{x}(kT) = \mathbf{A}_0^{(k-k_0)}\mathbf{x}(k_0T) + \sum_{n=k_0}^{k-1} \mathbf{A}_0^{(k-n-1)}\mathbf{A}_1(nT) \times \left[\mathbf{A}_0^{(n-k_0)}\mathbf{x}(k_0T) + \sum_{m=k_0}^{n-1} \mathbf{A}_0^{(n-m-1)}\mathbf{A}_1(mT)\mathbf{x}(mT)\right]$$

Further iterations yield

$$\mathbf{x}(kT) = [\mathbf{I} + \mathbf{S}(\boldsymbol{\phi}_0\mathbf{A}_1) + \mathbf{S}(\boldsymbol{\phi}_0\mathbf{A}_1\mathbf{S}(\boldsymbol{\phi}_0\mathbf{A}_1) + \mathbf{S}(\boldsymbol{\phi}_0\mathbf{A}_1\mathbf{S}(\boldsymbol{\phi}_0\mathbf{A}_1\mathbf{S}(\boldsymbol{\phi}_0\mathbf{A}_1))) + \cdots]\boldsymbol{\phi}_0\mathbf{x}(k_0T) \tag{6.8-10}$$

where the $\mathbf{S}$ indicates a summation of all terms to the right and $\boldsymbol{\phi}_0$ is given by Eq. 6.8-5. The state transition matrix $\boldsymbol{\phi}(kT, k_0T)$ can then be written as

$$\boldsymbol{\phi}[(k, k_0)T] = [\mathbf{I} + \mathbf{S}(\boldsymbol{\phi}_0\mathbf{A}_1) + \mathbf{S}(\boldsymbol{\phi}_0\mathbf{A}_1\mathbf{S}(\boldsymbol{\phi}_0\mathbf{A}_1)) + \mathbf{S}(\boldsymbol{\phi}_0\mathbf{A}_1\mathbf{S}(\boldsymbol{\phi}_0\mathbf{A}_1\mathbf{S}(\boldsymbol{\phi}_0\mathbf{A}_1))) + \cdots]\boldsymbol{\phi}_0 \tag{6.8-11}$$

which properly reduces to $\boldsymbol{\phi}_0$ for the case when the system is fixed, i.e., $\mathbf{A}_1(kT) = [0]$.

Equation 6.8-11 for discrete systems is analogous to Eq. 5.9-21 for continuous systems. If $\mathbf{A}_1$ represents a small perturbation upon the $\mathbf{A}_0$, then this series is rapidly convergent, and only a few terms are required to find $\boldsymbol{\phi}[(k, k_0)T]$. The advantage of using this form for the state transition matrix is that the general time-varying $\boldsymbol{\phi}$ is expressed in terms of successive corrections upon a constant $\boldsymbol{\phi}$.

Properties of the State Transition Matrix

The state transition matrix for discrete systems has a set of properties which are directly analogous to the properties listed for a continuous system in Section 5.9. Namely,

$$\boldsymbol{\phi}[(k_0, k_0)T] = \mathbf{I} \tag{6.8-12}$$

$$\boldsymbol{\phi}[(k_2, k_1)T]\boldsymbol{\phi}[(k_1, k_0)T] = \boldsymbol{\phi}[(k_2, k_0)T] \tag{6.8-13}$$

$$\boldsymbol{\phi}[(k_1, k_2)T] = \boldsymbol{\phi}^{-1}[(k_2, k_1)T] \tag{6.8-14}$$

For fixed systems,

$$\boldsymbol{\phi}[(k + n)] = \boldsymbol{\phi}(k)\boldsymbol{\phi}(n) \tag{6.8-15}$$

$$\boldsymbol{\phi}(k) = \boldsymbol{\phi}^{-1}(-k) \tag{6.8-16}$$

Computation of $\boldsymbol{\phi}$

In general, the computation of the state transition matrix for the time-varying case is a formidable task. Clearly, for any large value of n, Eq. 6.8-4 becomes most unwieldy. In certain cases, where the difference equations of the system can be handled, an analytical solution can be obtained (see Section 2.13). However, this occurs rarely. Use of a computer is generally the best method to obtain a solution.

For the fixed system, an analytical solution can generally be obtained. Equation 4.10-20 provides one method of computing the state transition matrix. Some others follow.

1. *Cayley-Hamilton Method.* For the discrete time case, the Cayley-Hamilton procedure can be used for computing $\mathbf{A}^k$. Here the $f(\lambda_i)$ to be used is λ_i^k rather than the $\epsilon^{\lambda_i t}$ used for the continuous case.

Example 6.8-1. Compute $\boldsymbol{\phi}(k)$ for the difference equation

$$y[(k + 2)T] + 5y[(k + 1)T] + 6y(kT) = 0$$

The $\mathbf{A}$ matrix for this system is

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -6 & -5 \end{bmatrix}$$

assuming $x_1(kT) = y(kT)$, $x_2(kT) = y[(k + 1)T]$. The characteristic equation $|\lambda\mathbf{I} - \mathbf{A}| = 0$ has two characteristic roots $\lambda_1 = -2$, $\lambda_2 = -3$. Therefore

$$F(\lambda_1) = \lambda_1^k = (-2)^k = \alpha_0 + \alpha_1\lambda_1 = \alpha_0 - 2\alpha_1$$

$$F(\lambda_2) = \lambda_2^k = (-3)^k = \alpha_0 + \alpha_1\lambda_2 = \alpha_0 - 3\alpha_1$$

From these two equations, $\alpha_0 = 3(-2)^k - 2(-3)^k$ and $\alpha_1 = (-2)^k - (-3)^k$. Hence

$$\mathbf{F}(\mathbf{A}) = \mathbf{A}^k = \alpha_0\mathbf{I} + \alpha_1\mathbf{A}$$

or

$$\boldsymbol{\phi}(k) = \begin{bmatrix} 3(-2)^k - 2(-3)^k & (-2)^k - (-3)^k \\ -6[(-2)^k - (-3)^k] & -2(-2)^k + 3(-3)^k \end{bmatrix}$$

Example 6.8-2. Compute $\boldsymbol{\phi}(k)$ for the system whose A matrix is given by

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix}$$

The characteristic equation $|\lambda\mathbf{I} - \mathbf{A}| = 0$ has two characteristic roots located at -1. For this case of repeated roots, the conditions which must be used to obtain the α's are

$$\left.\frac{d^s F(\lambda)}{d\lambda^s}\right|_{\lambda=\lambda_i} = \frac{d^s}{d\lambda^s}\left[\sum_{r=0}^{n-1}\alpha_r\lambda^r\right]_{\lambda=\lambda_i} = \frac{d^s}{d\lambda^s}[\alpha_0 + \alpha_1\lambda]_{\lambda=\lambda_i} \qquad s = 0, 1, 2, \ldots, p_i - 1$$

where p_i is the order of the root. Hence $(-1)^k = \alpha_0 + \alpha_1\lambda_1 = \alpha_0 - \alpha_1$ and $-k(-1)^k = \alpha_1$, or $\alpha_1 = -k(-1)^k$ and $\alpha_0 = (-1)^k(1 - k)$. Therefore

$$\boldsymbol{\phi}(k) = \alpha_0\mathbf{I} + \alpha_1\mathbf{A} = \begin{bmatrix} \alpha_0 & \alpha_1 \\ -\alpha_1 & \alpha_0 - 2\alpha_1 \end{bmatrix} = (-1)^k\begin{bmatrix} 1 - k & -k \\ k & 1 + k \end{bmatrix}$$

Example 6.8-3. It is informative to take up the case in which the A matrix may have complex roots. For this reason, determine $\Phi(k)$ corresponding to

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -2 & 2 \end{bmatrix}$$

The characteristic values of this matrix are $1 \pm j$. For purposes of evaluating $\mathbf{A}^k$, the polar form $\sqrt{2}\,\epsilon^{\pm j\pi/4}$ is most useful. The computation of $\mathbf{A}^k$ is then

$$F(\lambda_1) = (2)^{k/2}\epsilon^{jk\pi/4} = \alpha_0 + \alpha_1 + j\alpha_1$$

$$F(\lambda_2) = (2)^{k/2}\epsilon^{-jk\pi/4} = \alpha_0 + \alpha_1 - j\alpha_1$$

Adding and subtracting these equations yields $(2)^{k/2}\cos(k\pi/4) = \alpha_0 + \alpha_1$ and $\alpha_1 = (2)^{k/2}\sin(k\pi/4)$, or $\alpha_0 = (2)^{k/2}[\cos(k\pi/4) - \sin(k\pi/4)]$. Since

$$\Phi(k) = \alpha_0\mathbf{I} + \alpha_1\mathbf{A} = \begin{bmatrix} \alpha_0 & \alpha_1 \\ -2\alpha_1 & \alpha_0 + 2\alpha_1 \end{bmatrix}$$

then

$$\Phi(k) = (2)^{k/2}\begin{bmatrix} \left(\cos\dfrac{k\pi}{4} - \sin\dfrac{k\pi}{4}\right) & \sin\dfrac{k\pi}{4} \\ -2\sin\dfrac{k\pi}{4} & \left(\cos\dfrac{k\pi}{4} + \sin\dfrac{k\pi}{4}\right) \end{bmatrix}$$

2. *Frequency-Domain Method.* The Z transform, analogously to the Laplace transform, can be used to find the state transition matrix of the equation

$$\mathbf{x}[(k+1)T] = \mathbf{A}\mathbf{x}(kT) \tag{6.8-17}$$

Transforming both sides of Eq. 6.8-17, $z\mathbf{X}(z) - z\mathbf{x}(0) = \mathbf{A}\mathbf{X}(z)$, where use has been made of Eq. 3.12-3. Thus

$$\mathbf{X}(z) = (z\mathbf{I} - \mathbf{A})^{-1}z\mathbf{x}(0) \tag{6.8-18}$$

or $\mathbf{X}(z) = \mathbf{\Phi}(z)\mathbf{x}(0)$, where

$$\mathbf{\Phi}(z) = (z\mathbf{I} - \mathbf{A})^{-1}z \tag{6.8-19}$$

This form is slightly different from the analogous form $\mathbf{\Phi}(s) = (s\mathbf{I} - \mathbf{A})^{-1}$ for continuous systems. The state transition matrix is given by the inverse z transform of $\mathbf{\Phi}(z)$, or

$$\begin{aligned} \Phi(k) &= \mathscr{Z}^{-1}[(z\mathbf{I} - \mathbf{A})^{-1}z] \\ &= \sum \text{residues of } (z\mathbf{I} - \mathbf{A})^{-1}z^k \end{aligned} \tag{6.8-20}$$

Example 6.8-4. Using the A matrix of Example 6.8-1, determine $\Phi(k)$ by the frequency-domain method.

For this case,

$$(z\mathbf{I} - \mathbf{A}) = \begin{bmatrix} z & -1 \\ 6 & z+5 \end{bmatrix}$$

so that

$$(z\mathbf{I} - \mathbf{A})^{-1} = \frac{\begin{bmatrix} z+5 & 1 \\ -6 & z \end{bmatrix}}{(z+2)(z+3)}$$

Determining the sum of the residues of $(z\mathbf{I} - \mathbf{A})^{-1}z^k$ gives

$$\Phi(k) = \begin{bmatrix} 3(-2)^k - 2(-3)^k & (-2)^k - (-3)^k \\ -6[(-2)^k - (-3)^k] & -2(-2)^k + 3(-3)^k \end{bmatrix}$$

Example 6.8-5. Using the A matrix of Example 6.8-2, determine $\Phi(k)$ by the frequency-domain method.

For this case,

$$(z\mathbf{I} - \mathbf{A}) = \begin{bmatrix} z & -1 \\ 1 & z+2 \end{bmatrix}$$

so that

$$(z\mathbf{I} - \mathbf{A})^{-1} = \frac{\begin{bmatrix} z+2 & 1 \\ -1 & z \end{bmatrix}}{(z+1)^2}$$

Determination of the sum of the residues of $(z\mathbf{I} - \mathbf{A})^{-1}z^k$ yields

$$\Phi(k) = (-1)^k\begin{bmatrix} 1-k & -k \\ k & 1+k \end{bmatrix}$$

3. *Transfer Function Method.* As for the continuous case, the simulation diagram can be used to obtain the terms $\Phi_{ij}(z)$. The basis for this method is that the solution to Eq. 6.8-17 is

$$\mathbf{x}(kT) = \mathbf{\Phi}(k)\mathbf{x}(0) \tag{6.8-21}$$

Therefore $x_i(kT)$ is given by

$$x_i(kT) = \sum_{j=1}^{n} \phi_{ij}(k)x_j(0) \tag{6.8-22}$$

If all the state variables except the jth are set equal to zero, and a unit initial condition is placed on x_j, then the response at the ith state variable $x_i(kT)$ represents the term $\phi_{ij}(k)$. Therefore the transfer function from the *output* of the jth delay to the *output* of the ith delay represents the term $\Phi_{ij}(z)$. This is slightly different from the continuous case, where the transfer function was calculated between the *input* to the jth integrator and the *output* of the ith integrator. The difference is due to the fact that the transfer function in the continuous case is expressed as the transform of the

impulse response (in the continuous case, a unit impulse input establishes a unit initial condition *immediately*), while in the discrete case it is the transform of the *unit initial condition response*. If an analogy is desired, then the transfer function from the *input* of the jth delay to the *output* of the ith delay must be multiplied by z to obtain $\Phi_{ij}(z)$. Since what is generally desired is $(z\mathbf{I} - \mathbf{A})^{-1} = z^{-1}\mathbf{\Phi}(z)$, the transfer function from the *input* of the jth delay to the *output* of the ith delay is perfectly suitable.

Example 6.8-6. The simulation diagram of Fig. 6.8-1*a* represents the system of Example 6.8-1. Determine $\boldsymbol{\phi}(k)$.

Figure 6.8-1*b* is the same diagram redrawn for the convenience of computing transfer functions. Since 1/(1—loop transfer function) is equal to $z(z + 5)/(z^2 + 5z + 6)$, the

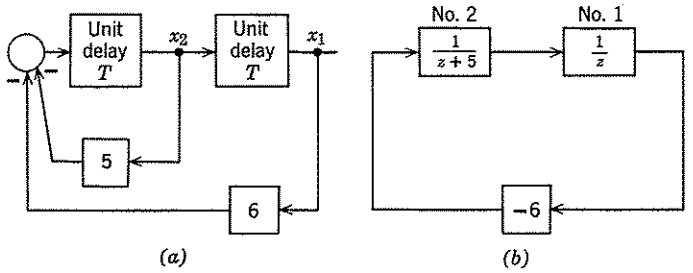


Fig. 6.8-1

various transfer functions $z^{-1}\Phi_{ij}(z)$ can be obtained by multiplying the forward transfer function from j to i by $z(z + 5)/(z^2 + 5z + 6)$. By performing this operation, the matrix $(z\mathbf{I} - \mathbf{A})^{-1} = z^{-1}\mathbf{\Phi}(z)$ is obtained. Thus

$$[z\mathbf{I} - \mathbf{A}]^{-1} = z^{-1}\mathbf{\Phi}(z) = \frac{\begin{bmatrix} z + 5 & 1 \\ -6 & z \end{bmatrix}}{z^2 + 5z + 6}$$

The inverse transform of $\mathbf{\Phi}(z)$ gives $\boldsymbol{\phi}(k)$ as in Example 6.8-4.

Adjoint System—The State Transition Matrix

Similar to the adjoint of an unforced linear continuous system, there exists an adjoint corresponding to an unforced discrete linear system. The adjoint operator $\mathbf{L}^*$ is defined in terms of the system operator $\mathbf{L}$ by

$$\langle \boldsymbol{\alpha}, \mathbf{Lx} \rangle = \langle \mathbf{L}^*\boldsymbol{\alpha}, \mathbf{x} \rangle \tag{6.8-23}$$

where $\boldsymbol{\alpha}$ is the adjoint vector and the inner product denotes†

$$\langle \mathbf{a}, \mathbf{b} \rangle = \sum_{i=k_0}^{k-1} \mathbf{a}^T[(i + 1)T]\mathbf{b}(iT)$$

† In the case of complex elements, the transpose is replaced by the conjugate transpose.

Considering $\langle \boldsymbol{\alpha}, \mathbf{Lx} \rangle$,

$$\begin{aligned}\langle \boldsymbol{\alpha}, \mathbf{Lx} \rangle &= \sum_{i=k_0}^{k-1} \boldsymbol{\alpha}^T[(i + 1)T]\{\mathbf{x}[(i + 1)T] - \mathbf{A}(iT)\mathbf{x}(iT)\} \\ &= \boldsymbol{\alpha}^T(kT)\mathbf{x}(kT) + \sum_{i=k_0-1}^{k-2} \boldsymbol{\alpha}^T[(i + 1)T]\mathbf{x}(i + 1)T] \\ &\qquad - \boldsymbol{\alpha}^T(k_0T)\mathbf{x}(k_0T) - \sum_{i=k_0}^{k-1} \boldsymbol{\alpha}^T[(i + 1)T]\mathbf{A}(iT)\mathbf{x}(iT) \\ &= \sum_{i=k_0}^{k-1} \{\boldsymbol{\alpha}^T(iT) - \boldsymbol{\alpha}^T[(i + 1)]\mathbf{A}(iT)\}\mathbf{x}(iT) \\ &\qquad + \boldsymbol{\alpha}^T(kT)\mathbf{x}(kT) - \boldsymbol{\alpha}^T(k_0T)\mathbf{x}(k_0T)\end{aligned} \tag{6.8-24}$$

Then by identifying

$$\mathbf{L}^*\boldsymbol{\alpha} = \boldsymbol{\alpha}^T(iT) - \boldsymbol{\alpha}^T[(i + 1)T]\mathbf{A}(iT) \tag{6.8-25}$$

so that the adjoint equation is

$$\boldsymbol{\alpha}^T(iT) = \boldsymbol{\alpha}^T[(i + 1)T]\mathbf{A}(iT) \tag{6.8-26}$$

the unforced system equation

$$\mathbf{x}[(i + 1)T] = \mathbf{A}(iT)\mathbf{x}(iT)$$

can be multiplied by $\boldsymbol{\alpha}^T[(i + 1)T]$ and combined with Eq. 6.8-26 to yield

$$\boldsymbol{\alpha}^T[(i + 1)T]\mathbf{x}[(i + 1)T] = \boldsymbol{\alpha}^T(iT)\mathbf{x}(iT)$$

Then by iteration

$$\boldsymbol{\alpha}^T(kT)\mathbf{x}(kT) = \boldsymbol{\alpha}^T(k_0T)\mathbf{x}(k_0T) \tag{6.8-27}$$

Substitution of Eqs. 6.8-25 and 6.8-27 into Eq. 6.8-24 yields Eq. 6.8-23. Thus Eq. 6.8-26 rewritten as

$$\boldsymbol{\alpha}[(i + 1)T] = [\mathbf{A}^{-1}(iT)]^T\boldsymbol{\alpha}(iT) \tag{6.8-28}$$

is the equation for the adjoint system.

Since the system state transition matrix $\boldsymbol{\phi}[(k, k_0)T]$ must satisfy Eq. 6.8-1,

$$\boldsymbol{\phi}[(k + 1, k_0)T] = \mathbf{A}(kT)\boldsymbol{\phi}[(k, k_0)T]$$

Taking the inverse and transposing,†

$$\{\boldsymbol{\phi}^{-1}[(k + 1, k_0)T]\}^T = [\mathbf{A}^{-1}(kT)]^T\{\boldsymbol{\phi}^{-1}[(k, k_0)T]\}^T \tag{6.8-29}$$

Comparison of Eqs. 6.8-28 and 6.8-29 indicates that the state transition matrix of the adjoint system is

$$\{\boldsymbol{\phi}^{-1}[(k, k_0)T]\}^T = \boldsymbol{\phi}^T[(k_0, k)T] \tag{6.8-30}$$

† If $\mathbf{A}(kT)$ contains complex elements, the conjugate transpose must be taken.

Note that the presentation here is the reverse of that of Chapter 5. There the continuous analogs of Eqs. 6.8-28 and 6.8-30 were defined, and the continuous analog of Eq. 6.8-23 resulted. Here Eqs. 6.8-28 and 6.8-30 are obtained from the definition of the adjoint operator, Eq. 6.8-23.

For the discrete time adjoint system, the state transition matrix

$$\{\boldsymbol{\Phi}^{-1}[(k, k_0)T]\}^T$$

can be found by iteration of Eq. 6.8-28. The result is

$$\{\boldsymbol{\Phi}^{-1}[(k, k_0)T]\}^T = \prod_{n=k_0}^{k-1} [\mathbf{A}^{-1}(nT)]^T \qquad k > k_0 \tag{6.8-31}$$

The state transition matrix for the original system, for $k_0 < k$, can be found by reverse iteration of Eq. 6.8-1, i.e., Eq. 6.8-1 can be written as $\mathbf{x}(nT) = \mathbf{A}^{-1}(nT)\mathbf{x}[(n+1)T]$, and this expression can then be iterated from $n = k$ down to $n = k_0$. The result is

$$\mathbf{x}(k_0T) = \left[\prod_{n=k_0}^{k-1} \mathbf{A}^{-1}(nT)\right]\mathbf{x}(kT)$$

so that

$$\boldsymbol{\Phi}[(k_0, k)T] = \prod_{n=k_0}^{k-1} \mathbf{A}^{-1}(nT) \qquad k > k_0 \tag{6.8-32}$$

These two equations show the validity of Eq. 6.8-14, namely,

$$\boldsymbol{\Phi}^{-1}[(k, k_0)T] = \boldsymbol{\Phi}[(k_0, k)T] \qquad k > k_0$$

However, Eq. 6.8-31 was obtained by a forward iteration or running the adjoint system forward in time from k_0T to kT, while Eq. 6.8-32 was obtained by a reverse iteration, or running the original system backward in time from kT to k_0T. This is shown in Fig. 6.8-2.

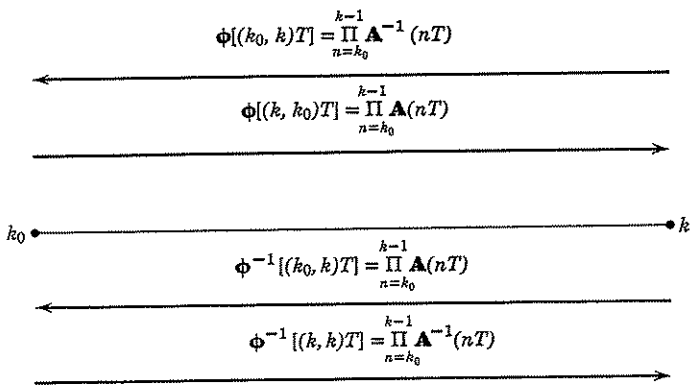


Fig. 6.8-2

Note that the form of the state transition matrix for increasing time (Eq. 6.8-4) is different from the form for the state transition matrix for decreasing time (Eq. 6.8-32). The reason for this difference is that reversing the time direction for a discrete system entails an inverse $\mathbf{A}$ matrix $\{\mathbf{x}(kT) = \mathbf{A}^{-1}(kT)\mathbf{x}[(k+1)T]\}$. Reversing the time direction for a continuous system simply entails reversing the sign of the $\mathbf{A}$ matrix $[d/d(-t) = -d/dt]$. A set of alternative forms of the state transition matrix is shown in Table 6.8-1.

Table 6.8-1

Original System		Adjoint System (Transposed)	
Forward Direction		*Forward Direction*	
$\boldsymbol{\Phi}[(k, k_0)T] = \prod_{n=k_0}^{k-1} \mathbf{A}(nT)$	$k > k_0$	$\boldsymbol{\Phi}^{-1}[(k, k_0)T] = \prod_{n=k_0}^{k-1} \mathbf{A}^{-1}(nT)$	$k > k_0$
$\boldsymbol{\Phi}[(k_0, k)T] = \prod_{n=k}^{k_0-1} \mathbf{A}(nT)$	$k_0 > k$	$\boldsymbol{\Phi}^{-1}[(k_0, k)T] = \prod_{n=k}^{k_0-1} \mathbf{A}^{-1}(nT)$	$k_0 > k$
Reverse Direction		*Reverse Direction*	
$\boldsymbol{\Phi}[(k_0, k)T] = \prod_{n=k_0}^{k-1} \mathbf{A}^{-1}(nT)$	$k > k_0$	$\boldsymbol{\Phi}^{-1}[(k_0, k)T] = \prod_{n=k_0}^{k-1} \mathbf{A}(nT)$	$k > k_0$
$\boldsymbol{\Phi}[(k, k_0)T] = \prod_{n=k}^{k_0-1} \mathbf{A}^{-1}(nT)$	$k_0 > k$	$\boldsymbol{\Phi}^{-1}[(k, k_0)T] = \prod_{n=k}^{k_0-1} \mathbf{A}(nT)$	$k_0 > k$

6.9 THE COMPLETE SOLUTION

The complete solution to the set of state equations

$$\begin{aligned} \mathbf{x}[(k+1)T] &= \mathbf{A}(kT)\mathbf{x}(kT) + \mathbf{B}(kT)\mathbf{v}(kT) \\ \mathbf{y}(kT) &= \mathbf{C}(kT)\mathbf{x}(kT) + \mathbf{D}(kT)\mathbf{v}(kT) \end{aligned} \tag{6.9-1}$$

can be found by a process of iteration and induction, similar to the method used to obtain the state transition matrix. However, to illustrate the analogy between discrete and continuous systems, the adjoint system is used in a method similar to the integrating factor method.

The state transition matrix of the adjoint system $\{\boldsymbol{\Phi}^{-1}[(k, k_0)T]\}^T$ must satisfy Eq. 6.8-29. Taking the transpose and postmultiplying by $\mathbf{A}(kT)$ yields

$$\boldsymbol{\Phi}^{-1}[(k, k_0)T] = \boldsymbol{\Phi}^{-1}[(k+1, k_0)T]\mathbf{A}(kT) \tag{6.9-2}$$

If Eq. 6.9-1 for $\mathbf{x}$ is premultiplied by $\boldsymbol{\phi}^{-1}[(k+1, k_0)T]$ and Eq. 6.9-2 is postmultiplied by $\mathbf{x}(kT)$ and the difference between the two equations taken, the result is

$$\boldsymbol{\phi}^{-1}[(k+1, k_0)T]\mathbf{x}[(k+1)T] - \boldsymbol{\phi}^{-1}[(k, k_0)T]\mathbf{x}(kT) = \boldsymbol{\phi}^{-1}[(k+1, k_0)T]\mathbf{B}(kT)\mathbf{v}(kT) \quad (6.9\text{-}3)$$

If k is replaced by m in Eq. 6.9-3 and both sides are then summed from k_0 to $k-1$, the result is

$$\boldsymbol{\phi}^{-1}[(k, k_0)T]\mathbf{x}(kT) - \boldsymbol{\phi}^{-1}[(k_0, k_0)T]\mathbf{x}(k_0T) = \sum_{m=k_0}^{k-1} \boldsymbol{\phi}^{-1}[(m+1, k_0)T]\mathbf{B}(mT)\mathbf{v}(mT) \quad (6.9\text{-}4)$$

Since $\boldsymbol{\phi}^{-1}[(k_0, k_0)T] = \mathbf{I}$, premultiplication by $\boldsymbol{\phi}[(k, k_0)T]$ gives

$$\mathbf{x}(kT) = \boldsymbol{\phi}[(k, k_0)T]\mathbf{x}(k_0T) + \sum_{m=k_0}^{k-1} \boldsymbol{\phi}[(k, m+1)T]\mathbf{B}(mT)\mathbf{v}(mT) \quad (6.9\text{-}5)$$

The first term on the right side of Eq. 6.9-5 represents the initial condition response of the system, while the second term represents a *superposition* summation of the effects of the forcing function. This equation is analogous to Eq. 5.10-6 for the continuous case.

When the system under investigation is fixed, Eq. 6.9-5 can be written as the sum of an initial condition response and a *convolution* summation:

$$\mathbf{x}(kT) = \boldsymbol{\phi}(k - k_0)\mathbf{x}(k_0T) + \sum_{m=k_0}^{k-1} \boldsymbol{\phi}(k - m - 1)\mathbf{B}\mathbf{v}(mT) \quad (6.9\text{-}6)$$

The corresponding output $\mathbf{y}(kT)$ is obtained by substitution of Eq. 6.9-5 or 6.9-6 into Eq. 6.9-1. Thus

$$\mathbf{y}(kT) = \mathbf{C}(kT)\boldsymbol{\phi}[(k, k_0)T]\mathbf{x}(k_0T) + \sum_{m=k_0}^{k-1} \mathbf{C}(kT)\boldsymbol{\phi}[(k, m+1)T]\mathbf{B}(mT)\mathbf{v}(mT) + \mathbf{D}(kT)\mathbf{v}(kT) \quad (6.9\text{-}7)$$

for a time-varying system, or

$$\mathbf{y}(kT) = \mathbf{C}\boldsymbol{\phi}(k - k_0)\mathbf{x}(k_0T) + \sum_{m=k_0}^{k-1} \mathbf{C}\boldsymbol{\phi}(k - m - 1)\mathbf{B}\mathbf{v}(mT) + \mathbf{D}\mathbf{v}(kT) \quad (6.9\text{-}8)$$

for a fixed system.

For the case in which the system is fixed, it is frequently convenient to use the frequency-domain approach. For this case, $\mathbf{Y}(z)$ can be found by transforming Eq. 6.9-1 directly. The result of this operation is

$$\mathbf{X}(z) = (z\mathbf{I} - \mathbf{A})^{-1}z\mathbf{x}(0) + (z\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}\mathbf{V}(z) \quad (6.9\text{-}9)$$

$$\mathbf{Y}(z) = \mathbf{C}(z\mathbf{I} - \mathbf{A})^{-1}z\mathbf{x}(0) + [\mathbf{C}(z\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}]\mathbf{V}(z) \quad (6.9\text{-}10)$$

An example is now given which illustrates the various approaches that can be taken.

Example 6.9-1. Determine the solution to the difference equation

$$y(k+2) + 5y(k+1) + 6y(k) = 1 \qquad (T = 1)$$

by each of the methods indicated below.

1. *Classical Solution—Time Domain.* Assume $y_H(k) = \beta^k$. Then $(\beta^2 + 5\beta + 6)\beta^k = 0$, or $\beta_1 = -2$ and $\beta_2 = -3$. The homogeneous solution is then $y_H(k) = C_1(-2)^k + C_2(-3)^k$. Assume that $y_P(k) = C_3$. Then $C_3 + 5C_3 + 6C_3 = 1$, or $C_3 = \frac{1}{12}$. The total solution is

$$y(k) = C_1(-2)^k + C_2(-3)^k + \tfrac{1}{12}$$

Using the initial conditions $y(0)$ and $y(1)$,

$$y(0) = C_1 + C_2 + \tfrac{1}{12} \quad \text{and} \quad y(1) = -2C_1 - 3C_2 + \tfrac{1}{12}$$

The constants C_1 and C_2 are

$$C_1 = 3y(0) + y(1) - \tfrac{1}{3}, \quad C_2 = \tfrac{1}{4} - 2y(0) - y(1)$$

Substituting these constants into the total solution gives the complete solution

$$y(k) = [3(-2)^k - 2(-3)^k]y(0) + [(-2)^k - (-3)^k]y(1) + [\tfrac{1}{4}(-3)^k - \tfrac{1}{3}(-2)^k + \tfrac{1}{12}]$$

2. *State Variables Technique—Time Domain.* From Example 6.8-1, the state transition matrix for this system is given by

$$\boldsymbol{\phi}(k) = \begin{bmatrix} 3(-2)^k - 2(-3)^k & (-2)^k - (-3)^k \\ -6[(-2)^k - (-3)^k] & -2(-2)^k + 3(-3)^k \end{bmatrix}$$

The $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ matrices are

$$\mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{C} = [1 \quad 0], \quad \mathbf{D} = [0]$$

The output $y(k)$ is given by Eq. 6.9-8 as

$$\mathbf{y}(k) = \mathbf{C}\boldsymbol{\phi}(k)\mathbf{x}(0) + \sum_{m=0}^{k-1} \mathbf{C}\boldsymbol{\phi}[(k - m - 1)]\mathbf{B}\mathbf{v}(mT) + \mathbf{D}\mathbf{v}(kT)$$

which, for this case, reduces to

$$y(k) = \phi_{11}(k)y(0) + \phi_{12}y(1) + \sum_{m=0}^{k-1} \phi_{12}(k - m - 1)$$

The computation of the summation for a forced input may involve some skill in finding a closed-form expression for the resulting series. The summation formula

$$\sum_{m=0}^{k-1} a^{k-m-1} = \frac{1 - a^k}{1 - a} \qquad \text{(sum of a geometric series)}$$

is of particular use in expressions of this type.[13] Since $\phi_{12}(k - m - 1) = (-2)^{k-m-1} - (-3)^{k-m-1}$, the summation in this case is equal to

$$\frac{1 - (-2)^k}{1 - (-2)} - \frac{1 - (-3)^k}{1 - (-3)} = \frac{1}{12} + \frac{1}{4}(-3)^k - \frac{1}{3}(-2)^k$$

The complete solution is then

$$y(k) = [3(-2)^k - 2(-3)^k]y(0) + [(-2)^k - (-3)^k]y(1) + [\tfrac{1}{4}(-3)^k - \tfrac{1}{3}(-2)^k + \tfrac{1}{12}]$$

The expression above for the sum of a geometric series is particularly helpful if the matrix summation

$$\sum_{m=0}^{k-1} \mathbf{A}^{k-m-1}$$

is to be found. The Cayley-Hamilton method can be applied where $F(\lambda_i) = (1 - \lambda_i^k)/(1 - \lambda_i)$.

A useful relation in finding the closed-form expression, if one exists, for a summation is the formula for *summation by parts*, analogous to the formula for *integration by parts*. This formula is given by (see Problem 2.12)

$$\sum_{M}^{N} u(k)\,\Delta v(k) = [u(k)v(k)]_M^{N+1} - \sum_{M}^{N} v(k+1)\,\Delta u(k)$$

As an example, the summation

$$\sum_{0}^{N} kr^k \qquad r \neq 1$$

can be found by setting

$$u(k) = k,\ \Delta v(k) = r^k.\ \text{Then } \Delta u(k) = 1,$$

$$v(k) = \sum_{n=0}^{k-1} r^n + C_1 = \frac{r^k - 1}{r - 1} + C_1 = \frac{r^k}{r-1} + C_2$$

For convenience, let $C_2 = 0$. Then

$$\sum_{0}^{N} kr^k = \left[\frac{kr^k}{r-1}\right]_0^{N+1} - \frac{1}{r-1}\sum_{0}^{N} r^{k+1}$$

$$= \frac{1}{(r-1)^2}[Nr^{N+2} - (N+1)r^{N+1} + r] \qquad r \neq 1$$

3. *Standard Z Transform Technique.* Taking the Z transform of both sides of the given difference equation, there results

$$z^2[Y(z) - y(0) - z^{-1}y(1)] + 5z[Y(z) - y(0)] + 6Y(z) = \frac{z}{z-1}$$

or

$$Y(z) = \frac{z/(z-1) + z[zy(0) + y(1) + 5y(0)]}{(z+2)(z+3)}$$

The poles of this function are at $z = 1, -2, -3$. Since

$$y(k) = \sum \text{residues of } Y(z)z^{k-1} \text{ at poles of } Y(z)$$

there are three residues to compute.

$$R_1 = (z-1)\,Y(z)z^{k-1}\big|_{z=1} = \tfrac{1}{12}$$

$$R_2 = (z+2)\,Y(z)z^{k-1}\big|_{z=-2} = [-\tfrac{1}{3} + 3y(0) + y(1)](-2)^k$$

$$R_3 = (z+3)\,Y(z)z^{k-1}\big|_{z=-3} = [\tfrac{1}{4} - 2y(0) - y(1)](-3)^k$$

The complete solution is then

$$y(k) = [3(-2)^k - 2(-3)^k]y(0) + [(-2)^k - (-3)^k]y(1) + [\tfrac{1}{4}(-3)^k - \tfrac{1}{3}(-2)^k + \tfrac{1}{12}]$$

4. *State Space Technique—Frequency Domain.* From Eq. 6.9-10, the Z transform of the output $\mathbf{y}(k)$ is

$$\mathbf{Y}(z) = \mathbf{C}(z\mathbf{I} - \mathbf{A})^{-1}z\mathbf{x}(0) + [\mathbf{C}(z\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}]\mathbf{V}(z)$$

From Example 6.8-4, the matrix $(z\mathbf{I} - \mathbf{A})^{-1}$ is given by

$$(z\mathbf{I} - \mathbf{A})^{-1} = \frac{\begin{bmatrix} z+5 & 1 \\ -6 & z \end{bmatrix}}{(z+2)(z+3)}$$

The $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ matrices were given in part 2 of this example as

$$\mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{C} = [1 \quad 0], \quad \mathbf{D} = [0]$$

The transform of the output, $Y(z)$, is then

$$Y(z) = \frac{z(z+5)}{(z+2)(z+3)}y(0) + \frac{z}{(z+2)(z+3)}y(1) + \frac{z}{(z-1)(z+2)(z+3)}$$

This z transform corresponds to the $Y(z)$ found in part 3 of this example. Therefore the answer is identical with that given in the three preceding parts of this example.

5. *State Variables Technique—Normal Form.* The modal matrix $\mathbf{M}$ is found by successive substitution of the characteristic values of $\mathbf{A}$ into the matrix $\mathbf{Adj}\,[\lambda\mathbf{I} - \mathbf{A}]$. The characteristic values of $\mathbf{A}$ are $\lambda_1 = -2$, $\lambda_2 = -3$, and $\mathbf{Adj}\,[\lambda\mathbf{I} - \mathbf{A}]$ is given by

$$\mathbf{Adj}\,[\lambda\mathbf{I} - \mathbf{A}] = \begin{bmatrix} \lambda+5 & 1 \\ -6 & \lambda \end{bmatrix}$$

Therefore $\mathbf{M}$ is given by

$$\mathbf{M} = \begin{bmatrix} 1 & 1 \\ -2 & -3 \end{bmatrix} \quad \text{and} \quad \mathbf{M}^{-1} = \begin{bmatrix} 3 & 1 \\ -2 & 1 \end{bmatrix}$$

Since

$$\mathbf{M}^{-1}\mathbf{A}\mathbf{M} = \mathbf{\Lambda} = \begin{bmatrix} -2 & 0 \\ 0 & -3 \end{bmatrix}$$

the transformation $\mathbf{q} = \mathbf{M}^{-1}\mathbf{x}$ leads into the form

$$\mathbf{q}(k+1) = \mathbf{\Lambda}\mathbf{q}(k) + \mathbf{B}_n\mathbf{v}(k) \qquad (T = 1)$$

$$\mathbf{y}(k) = \mathbf{C}_n\mathbf{q}(k) + \mathbf{D}\mathbf{v}(k)$$

where $\mathbf{B}_n = \mathbf{M}^{-1}\mathbf{B}$ and $\mathbf{C}_n = \mathbf{C}\mathbf{M}$. Therefore $q_1 = 3x_1 + x_2$, $q_2 = -2x_1 - x_2$, and

$$\mathbf{B}_n = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{and} \quad \mathbf{C}_n = [1 \quad 1]$$

The output $\mathbf{y}(k)$ is then

$$\mathbf{y}(k) = \mathbf{C}_n\mathbf{\Lambda}^k\mathbf{q}(0) + \sum_{m=0}^{k-1} \mathbf{C}_n\mathbf{\Lambda}^{k-m-1}\mathbf{B}_n\mathbf{v}(m)$$

from Eq. 6.6-4 and the above. Since $\mathbf{\Lambda}$ is a diagonal matrix, finding $\mathbf{\Lambda}^k$ means simply raising the elements of $\mathbf{\Lambda}$ to the kth power. This is one of the advantages of using this method. It follows that

$$y(k) = (-2)^k q_1(0) + (-3)^k q_2(0) + \sum_{m=0}^{k-1} \{(-2)^{k-m-1} - (-3)^{k-m-1}\}$$

or

$$y(k) = (-2)^k q_1(0) + (-3)^k q_2(0) + [\tfrac{1}{4}(-3)^k - \tfrac{1}{3}(-2)^k + \tfrac{1}{12}]$$

If it is desired that $y(k)$ be expressed in terms of the original initial conditions, then $q_1(0) = 3x_1(0) + x_2(0) = 3y(0) + y(1)$ and $q_2(0) = -2x_1(0) - x_2(0) = -2y(0) - y(1)$ yield

$$y(k) = [3y(0) + y(1)](-2)^k - [2y(0) + y(1)](-3)^k + [\tfrac{1}{4}(-3)^k - \tfrac{1}{3}(-2)^k + \tfrac{1}{12}]$$

6. *Mode Expansion Method.* Since

$$\mathbf{M} = \begin{bmatrix} 1 & 1 \\ -2 & -3 \end{bmatrix} \quad \text{and} \quad \mathbf{M}^{-1} = \begin{bmatrix} 3 & 1 \\ -2 & -1 \end{bmatrix}$$

from part 5, it follows that

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \quad \text{and} \quad \mathbf{u}_2 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$$

form a basis, and

$$\mathbf{r}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{r}_2 = \begin{bmatrix} -2 \\ -1 \end{bmatrix}$$

form a reciprocal basis. Thus, from Eq. 6.6-6,

$$y(k) = x_1(k) = [3x_1(0) + x_2(0)](-2)^k + [-2x_1(0) - x_2(0)](-3)^k + \sum_{m=0}^{k-1} [(-2)^{k-m-1} - (-3)^{k-m-1}]$$

This is the same expressions as for $y(k)$ in terms of the ϕ's in part 2. Thus performance of the indicated summations and substitution of $y(0) = x_1(0)$ and $y(1) = x_2(0)$ yields the desired result.

7. *State Variables Technique—Partial Fraction Expansion.* Since

$$H(z) = \frac{1}{z^2 + 5z + 6} = \frac{1}{z+2} - \frac{1}{z+3}$$

then

$$Y(z) = \frac{V(z)}{z+2} - \frac{V(z)}{z+3} = X_1(z) + X_2(z)$$

or $y(k) = x_1(k) + x_2(k)$, where x_1 and x_2 satisfy the first order difference equations

$$x_1(k+1) + 2x_1(k) = v(k) \quad \text{and} \quad x_2(k+1) + 3x_2(k) = -v(k)$$

Therefore the **A**, **B**, **C**, and **D** matrices are

$$\mathbf{A} = \begin{bmatrix} -2 & 0 \\ 0 & -3 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathbf{C} = [1 \quad 1], \quad \mathbf{D} = [0]$$

Since these are exactly the same matrices which were derived in part 5, $y(k)$ is then

$$y(k) = (-2)^k x_1(0) + (-3)^k x_2(0) + [\tfrac{1}{4}(-3)^k - \tfrac{1}{3}(-2)^k + \tfrac{1}{12}]$$

However, since the transfer function gives no clue to the relationship between **y** and **x**, this relationship must be obtained from examination of $y(k)$. This is necessary because the known initial conditions are in terms of $y(0)$ and $y(1)$. The desired relationships can be found by substituting $k = 0$ and $k = 1$ into $y(k) = x_1(k) + x_2(k)$. This gives $y(0) = x_1(0) + x_2(0)$ and $y(1) = x_1(1) + x_2(1)$. Now, using the equations for $x_1(k+1)$ and $x_2(k+1)$ at $k = 0$, these expressions become

$$y(0) = x_1(0) + x_2(0) \quad \text{and} \quad y(1) = -2x_1(0) - 3x_2(0)$$

Solving these two simultaneous equations yields

$$x_1(0) = 3y(0) + y(1) \quad \text{and} \quad x_2(0) = -2y(0) - y(1)$$

which are the same relationships found in part 5 for the q's. Although this procedure is not too difficult for a second order system, it may prove to be laborious for higher order systems with many inputs and outputs.

In looking through these various approaches, there are advantages and disadvantages to each. For single input-output fixed systems, the standard Z transform approach is certainly the easiest to use. For multiple input-output systems, a matrix technique is advisable. Which one of the matrix techniques to use is a different question. Since time-varying systems are almost impossible to solve analytically, the time-domain formulation of a time-varying system is best for computer purposes. For the fixed case, the Z transform of the state equations is quite useful, since numerous Z transform tables are available. The use of the Z transform bypasses some of the difficulties in evaluating the summation forms that are obtained in a time-domain computation. It appears that the normal form is perhaps the best form conceptually, for mathematical proofs, and for time-domain analysis of the unforced system. When the system is subject to a forcing function, the Z transform of the normal form is most convenient.

Stability of Fixed Discrete Systems

For a discrete fixed system, the state transition matrix approaches zero as k approaches infinity, if the characteristic values of **A** are located inside the unit circle. If a characteristic value lies on the unit circle and is of order one, then $\mathbf{\Phi}(k)$ is bounded as k approaches infinity. For any characteristic values which lie outside the unit circle or for multiple characteristic values which lie on the unit circle, $\mathbf{\Phi}(k)$ becomes infinite as k approaches infinity. These statements can be proved from the Cayley-Hamilton method of obtaining $\mathbf{A}^k$, which depends upon obtaining certain elements α_m such that

$$\mathbf{A}^k = \sum_{m=0}^{n-1} \alpha_m \mathbf{A}^m$$

where n is the order of the $\mathbf{A}$ matrix. The α_m are obtained from the equations

$$F(\lambda)\Big|_{\lambda=\lambda_i} = \sum_{m=0}^{n-1} \alpha_m \lambda_i{}^m = \lambda_i{}^k$$

or

$$\frac{d^p}{d\lambda^p} F(\lambda)\Big|_{\lambda=\lambda_i} = \frac{d^p}{d\lambda^p}\left(\sum_{m=0}^{n-1} \alpha_m \lambda^m\right)\Big|_{\lambda=\lambda_i} = \frac{d^p}{d\lambda^p}(\lambda^k)\Big|_{\lambda=\lambda_i}$$

where $p = 0, 1, 2, \ldots, r-1$ for the case where the characteristic value is of order r. When the characteristic values are distinct, the elements of $\mathbf{A}^k$ contain linear combinations of elements such as $(\lambda_i)^k$. These elements vanish as k approaches infinity if $|\lambda_i| < 1$, become bounded if $|\lambda_i| = 1$, and become unbounded if $|\lambda_i| > 1$. When there are multiple characteristic values, the elements of $\mathbf{A}^k$ contain linear combinations of elements such as $k\lambda_i^{k-1}$. Clearly, these elements are unbounded for $|\lambda_i| \geq 1$ and approach zero as k approaches infinity for $|\lambda_i| < 1$.

6.10 THE UNIT FUNCTION RESPONSE MATRIX

The output $\mathbf{y}(kT)$ of a linear time-varying discrete system can be written in terms of the $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ matrices as

$$(kT) = \sum_{m=-\infty}^{k-1} \mathbf{C}(kT)\boldsymbol{\phi}[(k, m+1)T]\mathbf{B}(mT)\mathbf{v}(mT) + \mathbf{D}(kT)\mathbf{v}(kT) \quad (6.10\text{-}1)$$

by setting k_0 equal to $-\infty$ in Eq. 6.9-7. In terms of the unit function response matrix $\mathbf{H}[(k, m)T]$, the output $\mathbf{y}(kT)$ is given by the superposition summation

$$\mathbf{y}(kT) = \sum_{m=-\infty}^{k} \mathbf{H}[(k, m)T]\mathbf{v}(mT) \quad (6.10\text{-}2)$$

A comparison of Eqs. 6.10-1 and 6.10-2 shows that the unit function response matrix is

$$\mathbf{H}[(k, k_0)T] = \begin{cases} \mathbf{C}(kT)\boldsymbol{\phi}[(k, k_0+1)T]\mathbf{B}(k_0T) & k \geqslant k_0 + 1 \\ \mathbf{D}(kT) & k = k_0 \\ = [0] & k < k_0 \end{cases} \quad (6.10\text{-}3)$$

For a fixed system the unit function response matrix is

$$\mathbf{H}[(k-k_0)T] = \begin{cases} \mathbf{C}\boldsymbol{\phi}[(k-k_0-1)]\mathbf{B} & k \geqslant k_0 + 1 \\ \mathbf{D} & k = k_0 \\ = [0] & k < k_0 \end{cases} \quad (6.10\text{-}4)$$

In the frequency domain, from Eq. 6.9-10, the unit function response matrix $\mathbf{H}(z)$ for a fixed system is given by

$$\mathbf{H}(z) = \mathbf{C}(z\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D} \quad (6.10\text{-}5)$$

Example 6.10-1. The system of Fig. 6.10-1 represents a simple sampled-data system. The block $(1 - \epsilon^{-sT})/s$ is commonly called a zero order hold, since it takes the input sample at time kT and provides this value as its output until time $(k+1)T$. Determine the unit function response matrix $\mathbf{H}(kT)$ assuming $T = 1$.

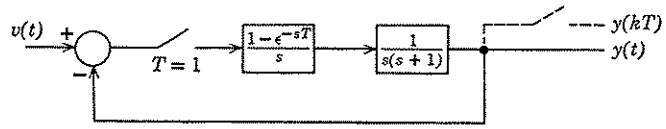


Fig. 6.10-1

The transfer function $G(z)$ of the forward transmission is

$$G(z) = (1 - z^{-1})\mathscr{Z}\left[\frac{1}{s^2(s+1)}\right] = \frac{\epsilon^{-1} - 1}{z - \epsilon^{-1}} + \frac{1}{z - 1}$$

Using this transfer function, the sampled-data system of Fig. 6.10-1 can be redrawn as the discrete time system of Fig. 6.10-2. In this figure, the forward path has been broken

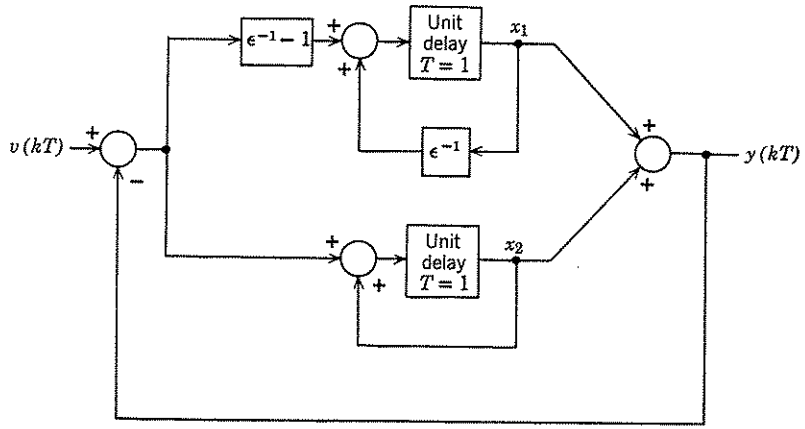


Fig. 6.10-2

up into its partial fraction expansion. The closed-loop system state equations can then be written down by inspection. They are

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = \begin{bmatrix} 1 & 1-\epsilon^{-1} \\ -1 & 0 \end{bmatrix}\begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + \begin{bmatrix} \epsilon^{-1} - 1 \\ 1 \end{bmatrix} v(k)$$

$$y(k) = [1 \quad 1]\begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix}$$

From Eq. 6.10-4, the unit function matrix $\mathbf{H}(k)$ is

$$\mathbf{H}(k) = [1 \quad 1]\begin{bmatrix}\phi_{11}(k-1) & \phi_{12}(k-1)\\ \phi_{21}(k-1) & \phi_{22}(k-1)\end{bmatrix}\begin{bmatrix}\epsilon^{-1}-1\\ 1\end{bmatrix}$$

Since

$$\boldsymbol{\Phi}(k-1) = \mathbf{A}^{(k-1)} = \alpha_0\mathbf{I} + \alpha_1\mathbf{A} = \begin{bmatrix}\alpha_0+\alpha_1 & \alpha_1-\alpha_1\epsilon^{-1}\\ -\alpha_1 & \alpha_0\end{bmatrix}$$

then $H(k) = \alpha_0\epsilon^{-1} + \alpha_1(1-\epsilon^{-1})$, where α_0 and α_1 are to be determined. The roots of the characteristic equation

$$|\lambda\mathbf{I} - \mathbf{A}| = \lambda^2 - \lambda + (1-\epsilon^{-1}) = 0$$

are $\lambda_{1,2} = N\epsilon^{\pm j\theta}$ where $N = 0.795$, $\theta = 0.680$ radian. Applying the Cayley-Hamilton method, where $F(\lambda_i) = \lambda^{k-1}$, the equations

$$N^{(k-1)}\epsilon^{j(k-1)\theta} = \alpha_0 + \alpha_1 N\epsilon^{j\theta}$$

$$N^{(k-1)}\epsilon^{-j(k-1)\theta} = \alpha_0 + \alpha_1 N\epsilon^{-j\theta}$$

are obtained. Solving for α_0 and α_1,

$$\alpha_0 = N^{(k-1)}[\cos(k-1)\theta - \cot\theta\sin(k-1)\theta]$$

$$\alpha_1 = \frac{N^{(k-2)}\sin[(k-1)\theta]}{\sin\theta}$$

The unit function response is then

$$H(k) = \epsilon^{-1}N^{(k-1)}[\cos(k-1)\theta - \cot\theta\sin(k-1)\theta] + \frac{(1-\epsilon^{-1})N^{(k-2)}}{\sin\theta}\sin(k-1)\theta$$

or

$$H(k) = (0.795)^{(k-1)}\{0.368\cos[0.680(k-1)] + 0.724\sin[0.680(k-1)]\}$$

This result can be checked by using conventional feedback methods.

$$H(z) = \frac{G(z)}{1+G(z)} = \frac{0.368(z+0.72)}{(z-N\epsilon^{j\theta})(z-N\epsilon^{-j\theta})} \qquad \begin{cases}N = 0.795\\ \theta = 0.680 \text{ radian}\end{cases}$$

Taking the inverse Z transform,

$$H(k) = 0.368N^{(k-1)}\left[\frac{(N\epsilon^{j\theta}+0.72)\epsilon^{j(k-1)\theta} - (N\epsilon^{-j\theta}+0.72)\epsilon^{-j(k-1)\theta}}{N(\epsilon^{j\theta}-\epsilon^{-j\theta})}\right]$$

After some manipulation, this can be written as

$$H(k) = 0.368N^{(k-1)}\left[\cos(k-1)\theta + \left(\cot\theta + \frac{0.72}{N\sin\theta}\right)\sin(k-1)\theta\right]$$

$$= (0.795)^{(k-1)}\{0.368\cos[0.680(k-1)] + 0.724\sin[0.680(k-1)]\}$$

For a time-varying system, the elements of the unit function response can be obtained by simulation in a manner analogous to that used for continuous systems. For an m input-p output system, the ith output $y_i(kT)$ is

given by

$$y_i(kT) = \sum_{n=-\infty}^{k}\sum_{j=1}^{m} h_{ij}[(k,n)T]v_j(nT) \tag{6.10-6}$$

The jth *column* of $\mathbf{H}[(k,n)T]$ can be obtained by setting all the inputs except the jth equal to zero, and placing a unit function on the jth input at time nT. The outputs of the system are $h_{ij}[(k,n)T]$ for $i = 1, 2, \ldots, p$ and for fixed j. In order to obtain the complete response $h_{ij}[(k,n)T]$ as a function of both k and n, a set of runs must be performed, each run starting at a different time $n_1T, n_2T, \ldots$. The results of these runs must then be cross-plotted to obtain the variation with respect to nT, the point in time of application. Proceeding to different values of j, these tests must be repeated until all m columns of the unit function response matrix are obtained. This is the same problem that was presented in the last chapter where the impulse response matrix $\mathbf{H}(t,\tau)$ was obtained by a similar cross-plotting procedure. In this discrete case, however, the difference equations can be solved on a digital computer and the necessary cross-plotting also done by the computer. Thus the discrete modified adjoint system is not discussed.

Example 6.10-2. In Section 5.11, the differential equation $\ddot{y} + t\dot{y} + y = 0$ was analyzed, and the impulse response $h(T,\tau)$ was obtained by performing a set of simulation runs for different values of τ and cross-plotting the results for fixed $t = T$. The results of the simulation runs are shown in Fig. 5.11-2*a*, and the results of the cross-plotting are shown in Fig. 5.11-2*b*. Perform the same task, but for the discrete version of the same differential equation. This method is often used to solve a time-varying differential equation numerically by either a desk calculator or a simple digital computer routine.

A discrete version of the equation can be obtained in the following manner. Since

$$\frac{d^2y}{dt^2} = \lim_{h\to 0}\left[\frac{y(t+h) - 2y(t) + y(t-h)}{h^2}\right]$$

$$\frac{dy}{dt} = \lim_{h\to 0}\left[\frac{y(t+h) - y(t)}{h}\right]$$

an approximate solution can be obtained by letting $h = T$ and $t = kT$, and using a "small" value for T. Thus

$$\frac{y(kT+T) - 2y(kT) + y(kT-T)}{T^2} + \frac{kT[y(kT+T) - y(kT)]}{T} + y(kT) = 0$$

or

$$(1+kT^2)[y(kT+T)] + (T^2 - kT^2 - 2)[y(kT)] + [y(kT-T)] = 0$$

The initial conditions $y(\tau) = 0$, $\dot{y}(\tau) = 1$ are replaced by

$$y(k_0T) = 0$$

$$\frac{y(k_0T+T) - y(k_0T)}{T} = 1$$

or $y(k_0T) = 0$ and $y(k_0T+T) = T$.

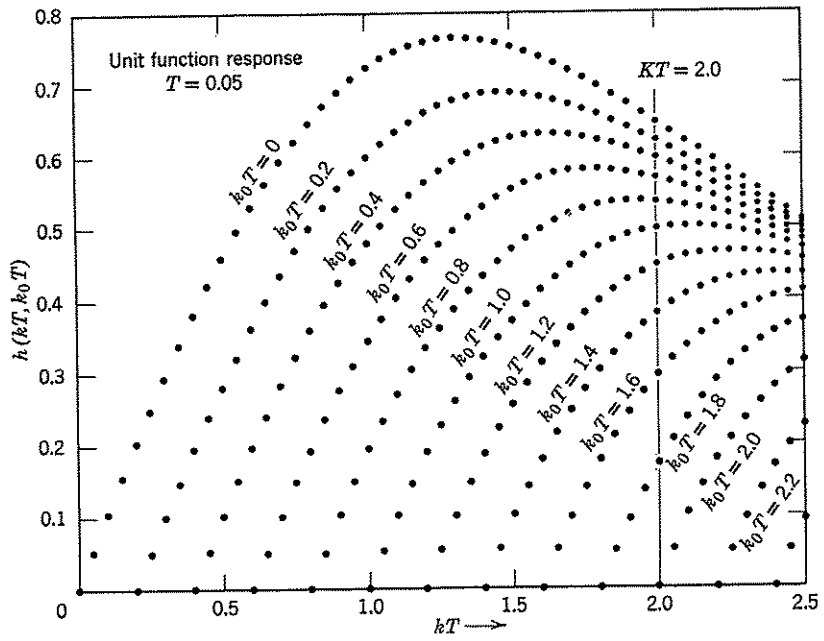


Fig. 6.10-3*a*

Using a value of $T = 0.05$ and $k_0 = 0, 4, 8, 12, \ldots, 48$, the points shown on Fig. 6.10-3*a* are obtained. Obtaining these points is a relatively simple task for a digital computer, and the cross-plotting can also be performed by the computer. A desk calculator can also be used. However, care must be taken to use sufficient accuracy in computing each point, as the round-off errors can build up rapidly. The interested reader can consult any of the many texts written on numerical solution of differential equations.[14,15]

As a comparison of the accuracy that can be obtained by simple numerical methods the results of the continuous system simulation run and the results of the approximate numerical solution for $\tau = k_0 = 0$ are listed in Table 6.10-1. The numerical solution is given to three places, while the simulation run is given to two places, this being the accuracy of reading from the original recording. For this comparison, the discrete simulation reproduces the results of the continuous simulation within the accuracy of the recording of the continuous information.

A crossplot of the points at $KT = 2.0$ is shown in Fig. 6.10-3*b*. The points of Fig. 6.10-3*b* represent $h(2.0, k_0T)$, the unit response of the system as $kT = 2.0$ as a function of the time of application of the unit input. A comparison of the cross-plot obtained by discrete simulation and the cross-plot obtained by the continuous simulation (Fig. 5.11-2*b*) is shown in Table 6.10-2. This comparison shows that the discrete simulation is fairly good, but not within the accuracy of the continuous system. The differences between the continuous and discrete systems are due to the accumulation of round-off error and the basic approximation involved. Because these effects are more noticeable farther out along a simulation run, a cross-plot at $kT = KT$ shows these effects more than a cross-plot at $kT < KT$.

Table 6.10-1

$T = 0.05$					
k	$h(k, 0)$	$h(t, 0)$	k	$h(k, 0)$	$h(t, 0)$
0	0	0	22	0.751	0.75
2	0.100	0.10	24	0.766	0.76
4	0.198	0.20	26	0.771	0.77
6	0.291	0.29	28	0.769	0.77
8	0.380	0.38	30	0.760	0.76
10	0.461	0.46	32	0.746	0.75
12	0.534	0.53	34	0.727	0.73
14	0.597	0.60	36	0.705	0.71
16	0.651	0.65	38	0.679	0.68
18	0.695	0.69	40	0.653	0.65
20	0.728	0.72			

Figure 6.10-3*b* can also be obtained by a discrete simulation of the modified adjoint differential equation. The original differential equation is $\ddot{y} + t\dot{y} + y = 0$, and the adjoint differential equation is $\ddot{\alpha} - (d/dt)(t\alpha) + \alpha = 0$ or $\ddot{\alpha} - t\dot{\alpha} = 0$. Making the change in variable $t = T_0 - t_1$, the modified adjoint differential equation is then

$$\frac{d^2\alpha}{dt_1^2} + (T_0 - t_1)\frac{d\alpha}{dt_1} = 0$$

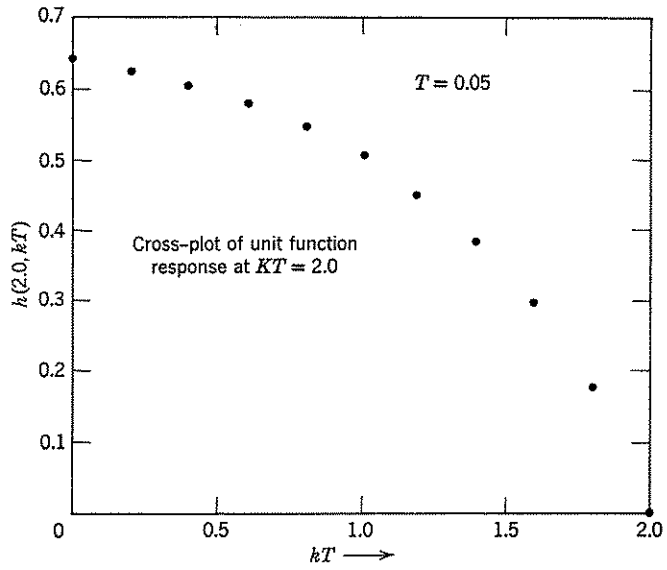


Fig. 6.10-3*b*

Table 6.10-2

$T = 0.05$		
k	$h(2.0, k_0T)$	$h(2.0, \tau)$
0	0.653	0.65
4	0.628	0.63
8	0.603	0.60
12	0.574	0.56
16	0.541	0.53
20	0.501	0.48
24	0.450	0.43
28	0.385	0.37
32	0.297	0.28
36	0.175	0.17
40	0	0

The discrete simulation of this equation is found by the procedure used to find the discrete simulation of the original differential equation. The resulting difference equation is

$$[1 + (K - k')T^2]y(k'T + T) + [-2 - (K - k')T^2]y(k'T) + y(k'T - T) = 0$$

where $KT = T_0$ is the fixed end time. A comparison of the points obtained by this simulation and those obtained from the continuous system simulation is shown in Table 6.10-3.

Table 6.10-3

$KT = 1.0, \quad T = 0.1$		
k	$h_{k'}{}^*(k', 0)$	$h_{t_1}{}^*(t_1, 0)$
0	0	0
1	0.100	—
2	0.191	0.18
3	0.276	—
4	0.350	0.33
5	0.420	—
6	0.487	0.47
7	0.550	—
8	0.612	0.60
9	0.673	—
10	0.733	0.72

Transmission Matrices

For single input-output systems, the *transmission matrix* is sometimes used to describe the unit response of the system. The transmission matrix is simply an ordered array of the elements $h(iT, jT)$, where i indicates the row and j indicates the column where the element is located. The general form for the transmission matrix is then

$$\mathbf{H}_T(kT, k_0T) = \begin{bmatrix} h(0,0) & 0 & 0 & \cdots & 0 \\ h(T,0) & h(T,T) & 0 & \cdots & 0 \\ h(2T,0) & h(2T,T) & h(2T,2T) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ h(mT,0) & h(mT,T) & h(mT,2T) & \cdots & h(mT,mT) \end{bmatrix} \tag{6.10-7}$$

All elements to the right of the element $h(iT, iT)$ are zero, since the system is assumed to be physically realizable, or nonanticipative. If the input $v(kT)$ is ordered into a column vector whose components are $v(0), v(T), v(2T), \ldots, v(mT)$ then the output of the system can be written as

$$\mathbf{y}(kT) = \mathbf{H}_T(kT, k_0T)\mathbf{v}(k_0T) \tag{6.10-8}$$

It is understood that the output $y(kT)$ is ordered into a column vector whose components are $y(0), y(T), y(2T), \ldots, y(mT)$. The ith component of the column vector $\mathbf{y}(kT)$ is then

$$y(iT) = \sum_{k_0=0}^{i} h[(i, k_0)T]v(k_0T) \tag{6.10-9}$$

For a fixed system, the elements of the transmission matrix are $h(iT - jT)$, the argument being the difference between the time of observation and the time of application. Thus

$$\mathbf{H}_T(kT) = \begin{bmatrix} h(0) & 0 & 0 & \cdots & 0 \\ h(T) & h(0) & 0 & \cdots & 0 \\ h(2T) & h(T) & h(0) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ h(mT) & h(mT - T) & h(mT - 2T) & \cdots & h(0) \end{bmatrix} \tag{6.10-10}$$

These matrices are not very useful when dealing with systems with several inputs and outputs, but they have some application when dealing with single input-single output systems or systems comprised of interconnected single input-single output systems.

Example 6.10-3. Repeat Example 1.9-1 using the transmission matrix description. $y(k)$, $k = 0, 1, 2, \ldots$, can be written as

$$\mathbf{y}(k) = \mathbf{D}_T(k)\mathbf{v}(k)$$

where

$$\mathbf{y}(k) = \begin{bmatrix} y(0) \\ y(1) \\ y(2) \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}$$

and from Example 1.9-1,

$$\mathbf{D}_T(k) = \begin{bmatrix} 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 2 & 0 & 0 & \cdot & \cdot & \cdot & \cdot \\ 4 & 2 & 0 & 0 & \cdot & \cdot & \cdot \\ 8 & 4 & 2 & 0 & 0 & \cdots & \\ \cdot & \cdot & \cdot & \cdot & \cdot & & \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \\ \cdot & \cdot & \cdot & \cdot & & & \cdot \end{bmatrix} \quad \text{and} \quad \mathbf{v}(k) = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}$$

Note that by defining

$$\mathbf{V}_T(k) = \begin{bmatrix} v(0) & 0 & \cdot & \cdots & \cdots \\ v(1) & v(0) & 0 & \cdots & \cdots \\ v(2) & v(1) & v(0) & 0 & \cdots \\ \cdot & \cdot & & \cdot & \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \end{bmatrix}$$

and

$$\mathbf{d}(k) = \begin{bmatrix} d(0) \\ d(1) \\ d(2) \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}$$

$\mathbf{y}(k)$ can also be written as $\mathbf{y}(k) = \mathbf{V}_T(k)\mathbf{d}(k)$.

Example 6.10-4. Write an expression for $y(kT)$, $k = 0, 1, 2, \ldots$, for the system of Example 3.13-6.

Defining the indicated transmission matrices and vectors as above,

$$\mathbf{q}(kT) = \mathbf{v}(kT) - \mathbf{G}\mathbf{H}_T(kT)\mathbf{q}(kT)$$

or

$$\mathbf{q}(kT) = [\mathbf{I} + \mathbf{G}\mathbf{H}_T(kT)]^{-1}\mathbf{v}(kT)$$

Then

$$\mathbf{y}(kT) = \mathbf{G}_T(kT)[\mathbf{I} + \mathbf{G}\mathbf{H}_T(kT)]^{-1}\mathbf{v}(kT)$$

6.11 THE METHOD OF LEAST SQUARES

At this point in the presentation of the preceding chapter, the evaluation of the mean square outputs of continuous systems was developed from the viewpoint of the modified adjoint system. As indicated in the previous section, the discrete modified adjoint system has limited value, because a digital computer can be programmed to overcome the analogous difficulty encountered with continuous systems. Thus this discussion of linear discrete systems departs from paralleling the preceding chapter to consider the general topic of least squares. This topic has considerable importance in the areas of communications, control, numerical analysis, prediction, and others, and, upon completion of this section, the reader is encouraged to investigate the continuous analogy to this section.[16,17]

Suppose that the output $y(kT)$ of a linear, stationary, discrete system is given by a weighted sum of the present and a finite number of past values of the input $v[(k - m)T]$, $m = 0, 1, 2, \ldots, M - 1$. This can be written in terms of the weighting factors $h(mT)$ as

$$y(kT) = v(kT)h(0) + v[(k-1)T]h(T) + \cdots + v[(k-M-1)T]h[(M-1)T] = \sum_{m=0}^{M-1} v[(k-m)T]h(mT)$$

In order to determine the system-weighting factors, $J \geqslant M$ sets of measurements of $v[(k - m)T]$, $m = 0, 1, \ldots, M - 1$, and the corresponding $y(kT)$ could be made. This would yield J equations of the form

$$\begin{aligned} v_{11}h_1 + v_{12}h_2 + \cdots + v_{1M}h_M &= y_1 \\ v_{21}h_1 + v_{22}h_2 + \cdots + v_{2M}h_M &= y_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ v_{J1}h_1 + v_{J2}h_2 + \cdots + v_{JM}h_M &= v_J \end{aligned} \tag{6.11-1}$$

where the subscript on the y's and the first subscript on the v's denote a particular set of measurements. The subscript on the h's and the second subscript on the v's denote the argument $(m + 1)T$. If only M sets of measurements are made, i.e., $J = M$, unique solutions exist for the h_i, but noise and measurement errors generally cause Eq. 6.11-1 to yield incorrect values. Thus, in such an experimental situation, many measurements are usually taken, and more equations than the number of unknowns are found. Hence $J > M$. Now, however, values for the h_i cannot be found which satisfy all the equations. For example, substitution of the values h_1^0, $h_2^0, \ldots, h_m^0$ for the unknown $h_1, h_2, \ldots, h_m$ in the left side of Eq. 6.11-1 might yield $y_1^0, y_2^0, \ldots, y_J^0$ which differ from $y_1, y_2, \ldots, y_J$ by $e_i = y_i^0 - y_i$, $i = 1, 2, \ldots, J$. Faced with this dilemma, one might attempt

to determine the h_i in such a way that each of Eqs. 6.11-1 is at least approximately valid, and such that some measure of the total approximation error is as small as possible. For example, the h_i^0 can be determined such that the e_i have the smallest possible mean square deviation, i.e.,

$$I = \sum_{i=1}^{J} e_i^2 = \sum_{i=1}^{J} (y_i^0 - y_i)^2$$

is a minimum. This is known as the *method of least squares*. It is, in essence, an attempt to find the "best" values for the h_i.

The preceding problem can be viewed geometrically. Let the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_M$ denote the columns of

$$\mathbf{V} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1M} \\ v_{21} & v_{22} & \cdots & v_{2M} \\ \cdots & \cdots & \cdots & \cdots \\ v_{J1} & v_{J2} & \cdots & v_{JM} \end{bmatrix}$$

Then the vector

$$\mathbf{y}^0 = \begin{bmatrix} y_1^0 \\ y_2^0 \\ \cdot \\ \cdot \\ \cdot \\ y_J^0 \end{bmatrix}$$

is given by $\mathbf{y}^0 = h_1^0\mathbf{v}_1 + h_2^0\mathbf{v}_2 + \cdots + h_M^0\mathbf{v}_M$. The problem becomes one of determining $h_1^0, h_2^0, \ldots, h_M^0$, such that $\|e\|^2 = \|y^0 - y\|^2$ is a minimum, where

$$\mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ \cdot \\ e_J \end{bmatrix} \quad \text{and} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_J \end{bmatrix}$$

Thus the problem is to determine

$$\mathbf{h}^0 = \begin{bmatrix} h_1^0 \\ h_2^0 \\ \cdot \\ \cdot \\ \cdot \\ h_M^0 \end{bmatrix}$$

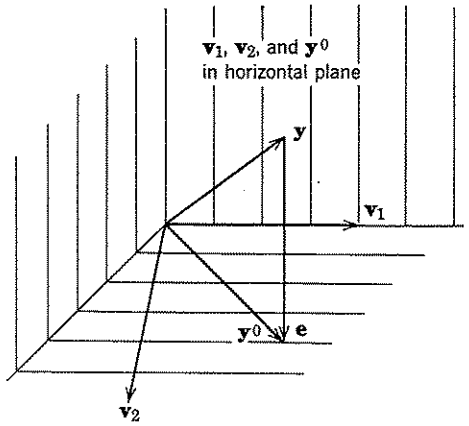


Fig. 6.11-1

such that $\mathbf{y}^0$ has the smallest possible deviation in norm from $\mathbf{y}$. This is represented for the case $M = 2$ in Fig. 6.11-1. In the general case, the set of all linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_M$ forms a space R^0, and the *orthogonal projection* of $\mathbf{y}$ on R^0 is the vector in R^0 which is the closest to $\mathbf{y}$. This is a simple generalization of what is geometrically apparent for $M = 2$ in Fig. 6.11-1. Thus $\mathbf{h}^0$ is to be chosen so that the linear combination $\mathbf{y}^0 = h_1^0\mathbf{v}_1 + h_2^0\mathbf{v}_2 + \cdots + h_M^0\mathbf{v}_M$ is the orthogonal projection of $\mathbf{y}$ on R^0.

Given a space R^0 and a vector $\mathbf{y}$, which is in general not contained in R^0, $\mathbf{y}$ can always be represented in the form

$$\mathbf{y} = \mathbf{y}^0 - \mathbf{e}$$

where the vector $\mathbf{y}^0$ belongs to R^0 and $\mathbf{e}$ is orthogonal to R^0. This is the geometrical idea behind the Gram-Schmidt orthogonalization procedure of Section 4.5. Taking $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_M$ as a basis in R^0, $\mathbf{y}^0 = h_1^0\mathbf{v}_1 + h_2^0\mathbf{v}^2 + \cdots + h_M^0\mathbf{v}_M$, where $h_1^0, h_2^0, \ldots, h_M^0$ are to be determined. The vector $\mathbf{e} = \mathbf{y}^0 - \mathbf{y}$ must be orthogonal to R^0, for which is is necessary and sufficient that

$$\langle \mathbf{e}, \mathbf{v}_i \rangle = \langle \mathbf{y}^0 - \mathbf{y}, \mathbf{v}_i \rangle = 0, \qquad i = 1, 2, \ldots, M$$

Substitution for $\mathbf{y}^0$ yields

$$\langle h_1^0\mathbf{v}_1 + h_2^0\mathbf{v}_2 + \cdots + h_M^0\mathbf{v}_M - \mathbf{y}, \mathbf{v}_i \rangle = 0, \quad i = 1, 2, \ldots, M$$

or

$$\begin{aligned} h_1{}^0\langle \mathbf{v}_1, \mathbf{v}_1\rangle + h_2{}^0\langle \mathbf{v}_2, \mathbf{v}_1\rangle + \cdots + h_M{}^0\langle \mathbf{v}_M, \mathbf{v}_1\rangle &= \langle \mathbf{y}, \mathbf{v}_1\rangle \\ h_1{}^0\langle \mathbf{v}_1, \mathbf{v}_2\rangle + h_2{}^0\langle \mathbf{v}_2, \mathbf{v}_2\rangle + \cdots + h_M{}^0\langle \mathbf{v}_M, \mathbf{v}_2\rangle &= \langle \mathbf{y}, \mathbf{v}_2\rangle \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ h_1{}^0\langle \mathbf{v}_1, \mathbf{v}_M\rangle + h_2{}^0\langle \mathbf{v}_2, \mathbf{v}_M\rangle + \cdots + h_M{}^0\langle \mathbf{v}_M, \mathbf{v}_M\rangle &= \langle \mathbf{y}, \mathbf{v}_M\rangle \end{aligned} \qquad (6.11\text{-}2)$$

The determinant of this set of equations is the Gram determinant G defined by Eq. 4.5-11, and it is nonzero as long as $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_M$ are linearly independent. This is, in essence, an observability condition. Assuming this, the value of $\mathbf{h}^0$, as determined by least squares, is given by

$$h_i{}^0 = \frac{1}{G}\begin{vmatrix} \langle \mathbf{v}_1, \mathbf{v}_1\rangle & \langle \mathbf{v}_2, \mathbf{v}_1\rangle & \cdots & \langle \mathbf{v}_{i-1}, \mathbf{v}_1\rangle & \langle \mathbf{y}, \mathbf{v}_1\rangle & \langle \mathbf{v}_{i+1}, \mathbf{v}_1\rangle & \cdots & \langle \mathbf{v}_M, \mathbf{v}_1\rangle \\ \langle \mathbf{v}_1, \mathbf{v}_2\rangle & \langle \mathbf{v}_2, \mathbf{v}_2\rangle & \cdots & \langle \mathbf{v}_{i-1}, \mathbf{v}_2\rangle & \langle \mathbf{y}, \mathbf{v}_2\rangle & \langle \mathbf{v}_{i+1}, \mathbf{v}_2\rangle & \cdots & \langle \mathbf{v}_M, \mathbf{v}_2\rangle \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \langle \mathbf{v}_1, \mathbf{v}_M\rangle & \langle \mathbf{v}_2, \mathbf{v}_M\rangle & \cdots & \langle \mathbf{v}_{i-1}, \mathbf{v}_M\rangle & \langle \mathbf{y}, \mathbf{v}_M\rangle & \langle \mathbf{v}_{i+1}, \mathbf{v}_M\rangle & \cdots & \langle \mathbf{v}_M \mathbf{v}_M\rangle \end{vmatrix} \qquad (6.11\text{-}3)$$

for $i = 1, 2, \ldots, M$.

The corresponding minimum value of the mean square deviation, $I = \|\mathbf{e}\|^2$, can also be determined from geometrical considerations. For $M = 2$, it is the square of the magnitude of the altitude of the parallelepiped determined by $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{y}$. In general, it is the square of the magnitude of the altitude of the hyperparallelepiped determined by $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_M, \mathbf{y}$. If V_y is used to denote the volume of this hyperparallelepiped, then $V_y = V\,\|\mathbf{e}\|$, or

$$\|\mathbf{e}\|^2 = \frac{V_y^2}{V^2}$$

where V is the volume determined by $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_M$. But $V_y{}^2$ is G_y, the Gramian of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_M, \mathbf{y}$, and V is G, the Gramian of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_M$. This is readily apparent for the case in which the vectors determining the hyperparallelepiped are orthogonal. For the nonorthogonal case, the Gram-Schmidt orthogonalization procedure can be used to arrive at the same conclusion. Thus

$$I_{\min} = \|\mathbf{e}\|^2_{\min} = \frac{G_y}{G}$$

is the minimum value of the mean square deviation.

It is useful to rewrite Eq. 6.11-2 as $\mathbf{V}^T\mathbf{V}\mathbf{h}^0 = \mathbf{V}^T\mathbf{y}$. Then Eq. 6.11-3 is

$$\mathbf{h}^0 = (\mathbf{V}^T\mathbf{V})^{-1}\mathbf{V}^T\mathbf{y} \qquad (6.11\text{-}4)$$

where the observability condition is that $\mathbf{V}^T\mathbf{V}$ be nonsingular. It is interesting now to assume that one additional measurement y_{J+1} is made,

which is supposed to equal $v_{J+1,1}h_1{}^0 + v_{J+1,2}h_2{}^0 + \cdots + v_{J+1,M}h_M{}^0$, and determine a new least squares estimate of $\mathbf{h}^0$ based on these $J + 1$ measurements. Let

$$\mathbf{v} = \begin{bmatrix} v_{J+1,1} \\ v_{J+1,2} \\ \cdot \\ \cdot \\ \cdot \\ v_{J+1,M} \end{bmatrix}$$

Then Eq. 6.11-4 becomes

$$\mathbf{h}^0_{J+1} = \left([\mathbf{V}^T \mid \mathbf{v}] \begin{bmatrix} \mathbf{V} \\ \hline \mathbf{v}^T \end{bmatrix} \right)^{-1} [\mathbf{V}^T \mid \mathbf{v}] \begin{bmatrix} \mathbf{y} \\ \hline y_{J+1} \end{bmatrix} \qquad (6.11\text{-}5)$$

where the subscript on $\mathbf{h}^0_{J+1}$ indicates that $J + 1$ measurements are used. Equation 6.11-5 provides a means of updating the least squares estimate of $\mathbf{h}^0$. However, it is not satisfactory from a computational viewpoint if the updating is to be continued, because of the matrix inversion required for each new estimate.

In an attempt to avoid repeated matrix inversions, let $\mathbf{P}_J = (\mathbf{V}^T\mathbf{V})^{-1}$. Then define

$$\mathbf{P}_{J+1} = \left([\mathbf{V}^T \mid \mathbf{v}] \begin{bmatrix} \mathbf{V} \\ \hline \mathbf{v}^T \end{bmatrix} \right)^{-1} = [\mathbf{V}^T\mathbf{V} + \mathbf{v}\mathbf{v}^T]^{-1}$$

so that

$$\mathbf{P}_{J+1} = [\mathbf{P}_J{}^{-1} + \mathbf{v}\mathbf{v}^T]^{-1} \qquad (6.11\text{-}6)$$

Direct substitution of

$$\mathbf{P}_{J+1} = \mathbf{P}_J - \mathbf{P}_J\mathbf{v}(\mathbf{v}^T\mathbf{P}_J\mathbf{v} + 1)^{-1}\mathbf{v}^T\mathbf{P}_J \qquad (6.11\text{-}7)$$

into Eq. 6.11-6 indicates that Eq. 6.11-7 is a valid expression for $\mathbf{P}_{J+1}$. Then, denoting the $\mathbf{h}^0$ based on the first J measurements by $\mathbf{h}_J{}^0$, Eq. 6.11-5 can be written, after simplification, as

$$\mathbf{h}^0_{J+1} = \mathbf{h}_J{}^0 + \mathbf{P}_J\mathbf{v}(\mathbf{v}^T\mathbf{P}_J\mathbf{v} + 1)^{-1}(y_{J+1} - \mathbf{v}^T\mathbf{h}_J{}^0) \qquad (6.11\text{-}8)$$

Since $(\mathbf{v}^T\mathbf{P}_J\mathbf{v} + 1)$ is a scalar, repeated matrix inversions are not required to update the estimation of $\mathbf{h}$. The updated $\mathbf{h}^0$ is the former $\mathbf{h}^0$ plus a weighting of the difference between the new value of y and the estimate of y based on J measurements.

REFERENCES

1. J. E. Bertram, "The Concept of State in the Analysis of Discrete Time Control Systems" 1962 Joint Automatic Control Conference, New York University, June 27–29, 1962, Paper No. 11-1.
2. D. A. Huffman, "The Synthesis of Sequential Switching Circuits," *J. Franklin Inst.*, March–April 1954, pp. 161–190, 275–303.