CHAPTER 8

State Space Analysis

8.1 Preview

The development of control system analysis and design can be divided into three eras. In the first era, we have classical control theory, which deals with techniques developed before 1950. Classical control embodies such methods as root locus, Bode, Nyquist, and Routh–Hurwitz. These methods have in common the use of transfer functions in the complex frequency (s) domain, emphasis on the use of graphical techniques, the use of feedback, and the use of simplifying assumptions to approximate the time response. Since computers were not available at that time, a great deal of emphasis was placed on developing methods that were amenable to manual computation and graphics. A major limitation of classical control methods was the use of single-input, single-output (SISO) methods. Multivariable (i.e., multiple-input, multiple-output, or MIMO) systems were analyzed and designed one loop at a time. Also, the use of transfer functions and the frequency domain limited one to linear time-invariant systems.

In the second era, we have modern control (which is not so modern any longer), which refers to state-space-based methods developed in the late 1950s and early 1960s. In modern control, system models are directly written in the time domain. Analysis and design are also done in the time domain. It should be noted that before Laplace transforms and transfer functions became popular in the 1920s, engineers were studying systems in the time domain. Therefore, the resurgence of time domain analysis was not unusual, but it was triggered by the development of computers and advances in numerical analysis. Because computers were available, it was no longer necessary to develop analysis and design methods that were strictly manual. An engineer could use computers to numerically solve or simulate large systems that were nonlinear and time-varying. State space methods removed the previously mentioned limitations of classical control. The period of the 1960s was the heyday of modern control.

That period did not last very long, however. For one thing, classical control was already well entrenched, tested, and established. Modern control methods initially enjoyed a great deal of success in academic circles, but they did not perform very well in real applications. Modern control provided a lot of insight into system structure and properties, but it masked other important feedback properties that could be studied and manipulated using classical control. For example, a basic problem in control theory is to design control systems that will work properly when the plant model is uncertain. This issue is tackled in classical control using gain and phase margins; most modern control design methods, however, inherently require a precise model of the plant. During the third era of the 1970s and 1980s, a body of methods finally emerged that tried to provide answers to the plant uncertainty problem. These techniques, commonly known as robust control, are a combination of modern state-space and classical frequency domain techniques.

For a thorough understanding of these new methods, we need to have a basic knowledge of state space analysis. Other advanced techniques in control, such as optimal and adaptive control, are also formulated in state space. Therefore, this chapter presents a brief introduction to state space. Since most of the mathematics associated with the modern approach relies heavily on matrix algebra, Appendix A provides a brief review of matrix algebra. To appreciate the material that follows, matrix algebra must be well understood.

8.2 State Space Representation

Up to this point in this textbook, all control systems have been represented by using transfer functions as functions of the complex frequency variable s. That approach is often called *classical* compared with the "modern" approach in which time domain (differential) equations are used. A fundamental apparatus needed to describe a control system in the time domain is the use of state variables. In general, a system that can be described by an n-order linear differential equation can be defined by creating n state variables. For example, a system whose transfer function has a second-order denominator would require two state variables, because if

$$\frac{Y(s)}{U(s)} = \frac{3(s+1)}{s^2+2s+4}$$

that system can be described by

$$\frac{d^2y}{dt^2} + 2\frac{dy}{dt} + 4y = 3\frac{du}{dt} + 3u$$

which is a second-order linear differential equation in y. By obtaining the transfer function, the system order is determined from the denominator. That order identifies the number of state variables that are needed. The problem, then, is to determine those state variables so the n choices are independent of each other.

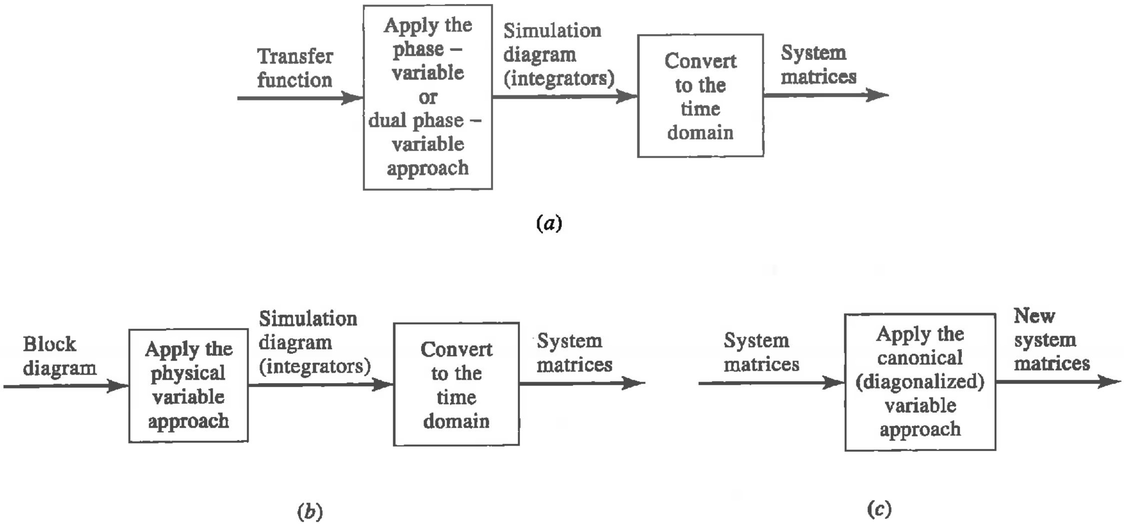


Figure 8.1 Creating a state-variable representation. (a) Phase and dual phase variables. (b) Physical variables. (c) Canonical (diagonalized) variables.

Figure 8.1 shows the three common methods for selecting state variables. If the system transfer function (but not the actual structure of the system) is available, then we see from Figure 8.1(a) that phase or dual phase variables can be used to create a linkage, called a *simulation diagram*, between the classical and the modern. Each integrator output (as we will soon see) is defined as a state variable. If the actual system structure is known in block diagram form (e.g., when a closed-loop system includes compensators), the resulting simulation diagram will follow the system structure, providing integrator outputs that have physical significance. These physical variables might include voltage, current, velocity, and position [e.g., Figure 8.1(b)].

In either case mentioned, the ultimate result is a set of matrices. Finally, if the matrices are available [Figure 8.1(c)], certain operations can be performed on them to create new system matrices of a particularly simple form (providing what are called a *set of canonical state variables*).

Physical variables are most closely related to real-world recognizable quantities, while canonical variables are least related to real-world quantities and are the most theoretical. The phase and dual-phase variables lie between the extremes of practical and theoretical.

8.2.1 Phase-Variable Form

An important problem in control system design is the synthesis of specific transfer functions through the interconnection of simple components, as is needed for many of the controllers (or compensators) of the preceding chapters. Synthesis is important also in the simulation of systems, where system behavior is predicted from a model governed by equivalent equations. Above all, the viewpoint of synthesis leads to fundamental techniques for system description, analysis, and design. These methods are systematic, compact, and suitable for computer analysis. They are also extendable to nonlinear and time-varying systems.

A basic component for synthesis is the integrator, a block or branch having transmittance $1/s$. A block diagram or signal flow graph composed only of constant transmittances and integrators is termed a *simulation diagram*. The order of such a system is simply the number of integrators present. Signal flow graphs are especially convenient for representing simulation diagrams because in many cases, system transfer functions are evident by inspection, making use of Mason's gain rule.

A transfer function that is the ratio of two polynomials in s is termed rational. If the numerator degree is less than or equal to the denominator degree, the transfer function is said to be *proper*. Any proper rational transfer function may be realized with a simulation diagram—that is, using only integration, multiplication by a constant, and summation operations. One very useful realization known as the *phase-variable* form, is described. The development, which is in terms of a specific numerical example for clarity, is applicable to any proper rational transfer function.

For the transfer function

Writing transfer function in Mason's form.

$$T(s) = \frac{-5s^2 + 4s - 12}{s^3 + 6s^2 + s + 3} = \frac{-5/s + 4/s^2 - 12/s^3}{1 + 6/s + 1/s^2 + 3/s^3} = \frac{P_1 + P_2 + P_3}{1 - L_1 - L_2 - L_3}$$

dividing the numerator and denominator by the highest power-of-s term in the denominator places a 1 in the denominator and results in other numerator and denominator terms that are inverse powers of s, representing multiple integrations. In this form the transfer function may be interpreted as a Mason's gain rule expression. The numerator terms

$$\frac{-5}{s} + \frac{4}{s^2} + \frac{-12}{s^3}$$

are each taken to be paths through integrators, and the paths are intermingled as in Figure 8.2(a) to require a minimum number of integrators—in this case, three. The denominator terms

$$\frac{6}{s} + \frac{1}{s^2} + \frac{3}{s^3}$$

are taken to be the negative of the loop gains. By placing each of these loops through the node to which $U(s)$ couples, all loops touch one another, so no product of loop gain terms is involved. All the loops touch each of the paths, so each path cofactor is unity.

In Figure 8.2(b), each integrator output signal has been labeled. These signals are termed the *state variables* of the system. This realization of the example transfer function is then described by the following Laplace-transformed equations:

$$X_1(s) = \frac{1}{s} X_2(s)$$

$$X_2(s) = \frac{1}{s} X_3(s)$$

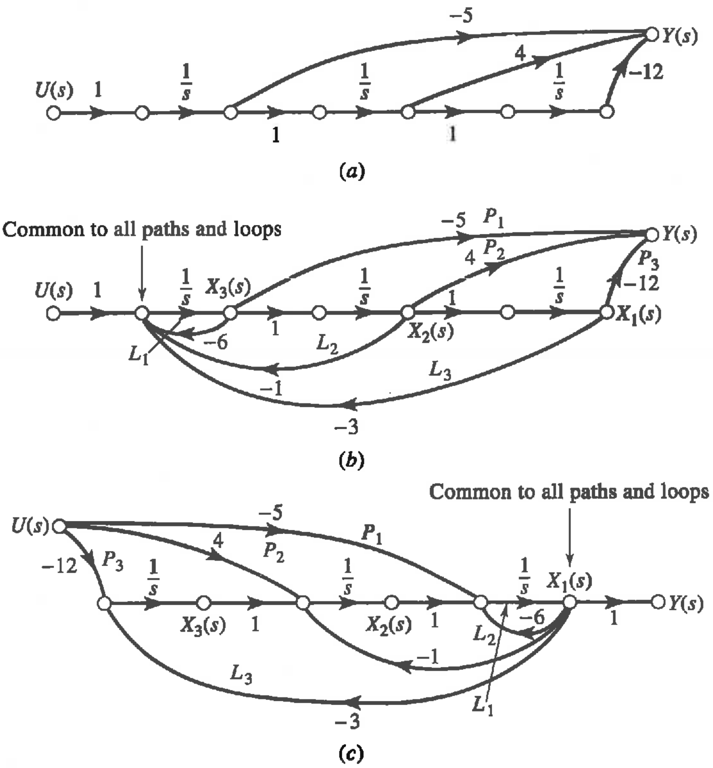


Figure 8.2 Phase-variable realization of a single-input, single-output system. (a) Paths in the simulation diagram. (b) Complete simulation diagram. (c) Realizing a transfer function in the dual phase-variable form.

$$
\begin{aligned}
X_3(s) &= \frac{1}{s}[-3X_1(s) - X_2(s) - 6X_3(s) + U(s)] \\
Y(s) &= -12X_1(s) + 4X_2(s) - 5X_3(s)
\end{aligned}
$$

or

$$
\begin{aligned}
sX_1(s) &= X_2(s) \\
sX_2(s) &= X_3(s) \\
sX_3(s) &= -3X_1(s) - X_2(s) - 6X_3(s) + U(s) \\
Y(s) &= -12X_1 + 4X_2(s) - 5X_3(s)
\end{aligned}
$$

As functions of time, the signals satisfy

$$
\begin{aligned}
\frac{dx_1}{dt} &= x_2(t) \\
\frac{dx_2}{dt} &= x_3(t)
\end{aligned}
$$

$$\frac{dx_3}{dt} = -3x_1(t) - x_2(t) - 6x_3(t) + u(t)$$

$$y(t) = -12x_1(t) + 4x_2(t) - 5x_3(t)$$

which is a set of coupled first-order differential equations.

Denoting time derivatives by $\dot{x}$, and defining the vectors x and $\dot{x}$ by

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \dot{x} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix}$$

allows us to rewrite the foregoing equations in matrix–vector format as follows:

State equations in vector-matrix form.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -3 & -1 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u$$

$$y = [-12 \quad 4 \quad -5] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + 0 \cdot u$$

or more compactly as follows:

$$\dot{x} = Ax + Bu$$

$$y = Cx + Du$$

where the set of four matrices $\{A,\ B,\ C,\ D\}$ are called a *quadruple*. In general, any proper transfer function can be converted to the general form just shown. In fact, any linear differential equation (possibly with variable coefficients) can be converted to this form (in this case, the coefficient matrices may be functions of time rather than constants).

8.2.2 Dual Phase-Variable Form

Another especially convenient way to realize a transfer function with integrators is to arrange the signal flow graph so that all the paths and all the loops touch an output node. For the previous transfer function

$$T(s) = \frac{-5s^2 + 4s - 12}{s^3 + 6s^2 + s + 3}$$

$$= \frac{-5/s + 4/s^2 - 12/s^3}{1 + 6/s + 1/s^2 + 3/s^3}$$

$$= \frac{P_1 + P_2 + P_3}{1 - L_1 - L_2 - L_3}$$

for example, the diagram of Figure 8.2(c) shows this *dual phase-variable* arrangement. The output signal is derived from a single node, while the input signal is coupled to each integrator.

The Laplace transform relations describing this system are, in terms of the indicated state variables,

$$\begin{aligned}
sX_1(s) &= -6X_1(s) + X_2(s) - 5U(s) \\
sX_2(s) &= -X_1(s) + X_3(s) + 4U(s) \\
sX_3(s) &= -3X_1 - 12U(s) \\
Y(s) &= X_1(s)
\end{aligned}$$

or in the time domain

$$\dot{x} = \begin{bmatrix} -6 & 1 & 0 \\ -1 & 0 & 1 \\ -3 & 0 & 0 \end{bmatrix} x + \begin{bmatrix} -5 \\ 4 \\ -12 \end{bmatrix} u$$

$$y = [1 \quad 0 \quad 0]x + 0 \cdot u$$

Phase-variable and dual phase-variable forms are called *canonical forms.* Our constructions always lead to a special form for the $\{A, \ B, \ C, \ D\}$ quadruplets. For example, in the phase-variable form, assuming that the state variables are defined from right to left, we can observe the following patterns. B is a column vector of zeros except the last element which is 1; C is a row vector that contains the coefficients of the transfer function numerator in ascending powers of s. The D term is a scalar, and is always 0 if the transfer function is *strictly proper* (i.e., the degree of the numerator is strictly less than the degree of the denominator). The A matrix has a special form, which can be partitioned as follows:

$$A = \left[\begin{array}{c:cc} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hdashline -3 & -1 & -6 \end{array}\right]$$

Note that the last-row elements are the negative of the coefficients of the transfer function denominator (in ascending powers of s, the highest-degree term is always assumed to have a coefficient of 1). The first column is all zeros (except the last element). The remaining submatrix is an identity matrix.

If you look closely at the matrices in the dual phase-variable form, you will see almost the same pattern. If fact, if you make the following substitutions in the preceding two paragraphs, you will get the dual phase-variable form matrices: replace B and C, row with column, first with last, ascending with descending. A substitution that allows us to go from one form to another is called a *dual.* In fact, you may recall such dualities from basic circuit theory. This explains the name, *dual phase-variable form.*

Because of the above-mentioned patterns, we can obtain these forms directly from the transfer functions. If a system transfer function is given by

$$G(s) = \frac{b_1 s^{n-1} + \cdots + b_{n-1}s + b_n}{s^n + a_1 s^{n-1} + \cdots + a_n}$$

The phase-variable form matrices are given by

Phase-variable form.

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdot & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ & \cdot & & & & \cdot \\ \vdots & \vdots & & \ddots & & 0 \\ & & & & & 1 \\ -a_n & -a_{n-1} & \cdot & \cdots & & -a_1 \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ 0 \\ 1 \end{bmatrix}$$

$$C = [b_n \; b_{n-1} \cdots b_1] \qquad D = [0]$$

The dual phase-variable form matrices are given by

Dual phase-variable form.

$$A = \begin{bmatrix} -a_1 & 1 & 0 & \cdots & 0 \\ -a_2 & 0 & 1 & 0 \cdot\cdot & 0 \\ & & 0 & & \cdot \\ \vdots & \vdots & \vdots & \ddots & 0 \\ & & & & 1 \\ -a_n & 0 & 0 & \cdots & 0 \end{bmatrix} \qquad B = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_n \end{bmatrix}$$

$$C = [1 \quad 0 \quad \cdots \quad 0] \quad D = [0]$$

Matrices that have the special structure observed in the A matrices are called *companion* matrices in matrix algebra. An important property of companion matrices is that their characteristic equation can be obtained by inspection. In particular, the characteristic equation for the above A matrices is as follows:

$$\text{Characteristic equation} = s^n + a_1 s^{n-1} + \cdots + a_n$$

8.2.3 Multiple Inputs and Outputs

Additional system outputs may be easily derived from the phase-variable arrangement. For example, the single-input, two-output system of Figure 8.3(a) has the following transfer functions:

$$T_{11}(s) = \left.\frac{Y_1(s)}{U(s)}\right|_{\substack{\text{initial} \\ \text{conditions} = 0}} = \frac{-5/s + 4/s^2 + (-12/s^3)}{1 + 6/s + 1/s^2 + 3/s^3}$$

$$= \frac{-5s^2 + 4s - 12}{s^3 + 6s^2 + s + 3}$$

$$T_{21}(s) = \left.\frac{Y_2(s)}{U(s)}\right|_{\substack{\text{initial} \\ \text{conditions}=0}} = \frac{3/s + 1/s^2 + (-6/s^3)}{1 + 6/s + 1/s^2 + 3/s^3}$$

$$= \frac{3s^2 + s - 6}{s^3 + 6s^2 + s + 3}$$

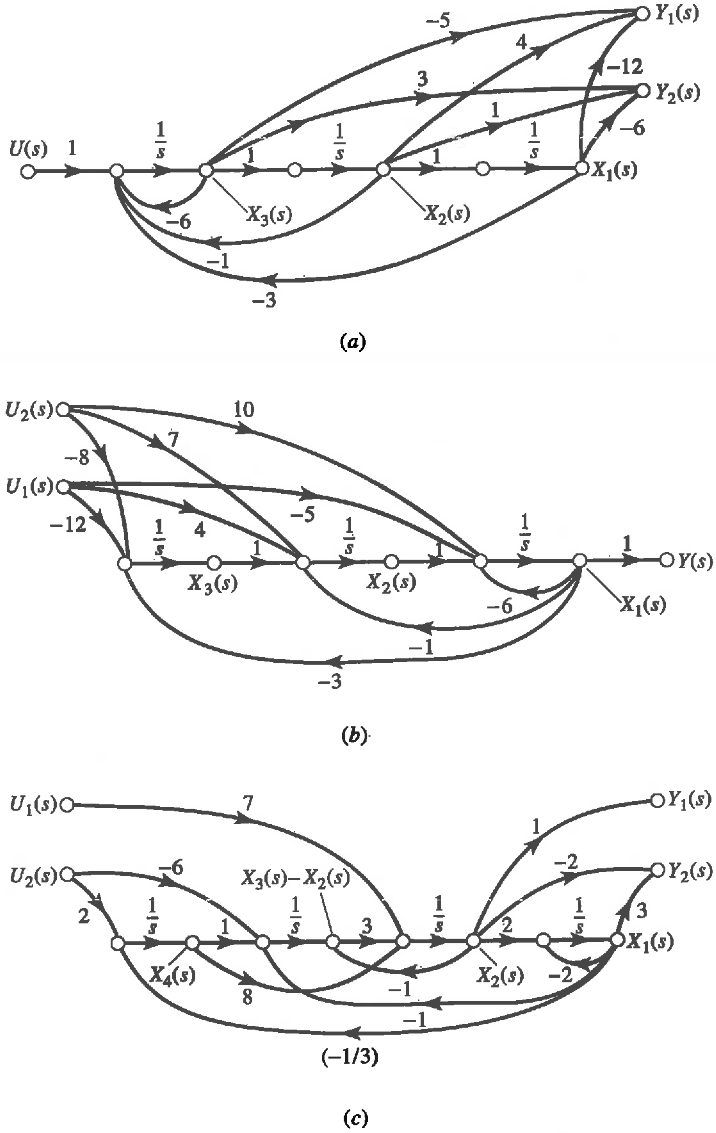


Figure 8.3 Achieving multiple outputs and multiple inputs. (a) Multiple outputs from the phase-variable arrangement. (b) Multiple inputs with the dual phase-variable arrangement. (c) System with both multiple inputs and multiple outputs.

The system transfer function can be written as a vector

$$T(s) = \begin{bmatrix} T_{11}(s) \\ T_{21}(s) \end{bmatrix} = \frac{1}{s^3 + 6s^2 + s + 3} \begin{bmatrix} -5s^2 + 4s - 12 \\ 3s^2 + s - 6 \end{bmatrix}$$

The state space realization for the system in phase-variable form can be obtained from the signal flow graph. Note that both transfer functions share the same denominator and input, hence A and B are the same as the single-input case. The realization is given by

A two-output system.

$$\dot{x} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -3 & -1 & -6 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} -12 & 4 & -5 \\ -6 & 1 & 3 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \end{bmatrix} u$$

Because there are two outputs, y is a two-dimensional vector; hence C and D must have two rows to make the dimensions match. Also note that if the transfer function is available, the phase-variable realization could have been written by inspection.

Additional inputs are easily added to the dual phase-variable arrangement. The example two-input, single-output system of Figure 8.3(b) has the following transfer functions:

$$T_{11}(s) = \left.\frac{Y(s)}{U_1(s)}\right|_{\substack{\text{initial}\\ \text{conditions}\\ \text{and } R_2=0}} = \frac{-5/s + 4/s^2 - 12/s^3}{1 + 6/s + 1/s^2 + 3/s^3}$$

$$= \frac{-5s^2 + 4s - 12}{s^3 + 6s^2 + s + 3}$$

$$T_{12}(s) = \left.\frac{Y(s)}{U_2(s)}\right|_{\substack{\text{initial}\\ \text{conditions}\\ \text{and } R_1=0}} = \frac{10/s + 7/s^2 - 8/s^3}{1 + 6/s + 1/s^2 + 3/s^3}$$

$$= \frac{10s^2 + 7s - 8}{s^3 + 6s^2 + s + 3}$$

You can use the same kind of reasoning as in the preceding case to verify that the dual phase-variable realization for the two-input case is given by

A two-input system.

$$\dot{x} = \begin{bmatrix} -6 & 1 & 0 \\ -1 & 0 & 1 \\ -3 & 0 & 0 \end{bmatrix} x + \begin{bmatrix} -5 & 10 \\ 4 & 7 \\ -12 & -8 \end{bmatrix} u$$

$$y = [1 \quad 0 \quad 0]x + [0 \quad 0]u$$

The system described by the simulation diagram of Figure 8.3(c) is in neither phase-variable nor dual phase-variable form. Its two inputs and two outputs are governed by the following Laplace-transformed equations:

$$sX_1(s) = -2X_1(s) + 2X_2(s)$$

$$sX_2(s) = -3X_2(s) + 3X_3(s) + 8X_4(s) + 7U_1(s)$$

$$sX_3(s) = -X_1(s) + X_4(s) - 6U_2(s)$$
$$sX_4(s) = -\tfrac{1}{3}X_1(s) + 2U_2(s)$$
$$Y_1(s) = X_2(s)$$
$$Y_2(s) = 3X_1(s) - 2X_2(s)$$

We can write the state space realization directly from these equations:

$$\dot{x} = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & -3 & 3 & 8 \\ -1 & 0 & 0 & 1 \\ -\frac{1}{3} & 0 & 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 & 0 \\ 7 & 0 \\ 0 & -6 \\ 0 & 2 \end{bmatrix} u$$

$$y = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3 & -2 & 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} u$$

A two-input two-output system.

Observe that matrices A, B, C, and D are not in any particular form.

❑ DRILL PROBLEM

D8.1 Draw simulation diagrams in either the phase-variable or the dual phase-variable form for systems with the following transfer functions:

(a) $T(s) = \dfrac{-4s+3}{s^2+6s+2}$

(b) $T(s) = \dfrac{-s^2+5s+9}{3s^3+2s^2+4s+1}$

(c) $T_{11}(s) = \dfrac{0.4s^2+1.4s+0.8}{s^3+0.3s^2+1.7s+0.2}$

$T_{12}(s) = \dfrac{-0.5s^2+0.7s-1.9}{s^3+0.3s^2+1.7s+0.2}$

(d) $T_{11}(s) = \dfrac{4s^2-1}{s^3+6s^2+2s+5}$

$T_{21}(s) = \dfrac{3s+6}{s^3+6s^2+2s+5}$

Phase-variable form is particularly convenient for the synthesis of single- and multiple-output systems, while in dual phase-variable form, single- and multiple-input systems are easily arranged. There are a whole spectrum of other ways of connecting integrators to achieve systems with desired transfer functions, including systems with both multiple inputs and multiple outputs. Moreover, the representation of systems in terms of integrators is useful not only for transfer function synthesis, but for the description of systems of all kinds, particularly those that are very complicated, for which a standard, compact notation is especially helpful.

A general state variable description of an nth-order system involves n integrators, the outputs of which are the state variables. The input of each of the integrators are driven with a linear combination of the state signals and the inputs:

$$\begin{aligned}
sX_1(s) &= a_{11}X_1(s) + a_{12}X_2(s) + \cdots + a_{1n}X_n(s) + b_{11}u_1(s) + \cdots + b_{1i}u_i(s) \\
sX_2(s) &= a_{21}X_1(s) + a_{22}X_2(s) + \cdots + a_{2n}X_n(s) + b_{21}u_1(s) + \cdots + b_{2i}u_i(s) \\
&\vdots \\
sX_n(s) &= a_{n1}X_1(s) + a_{n2}X_2(s) + \cdots + a_{nn}X_n(s) + b_{n1}u_1(s) + \cdots + b_{ni}u_i(s)
\end{aligned} \tag{8.1}$$

In the time domain, these are a set of n first-order differential equations in the n state variables and the inputs:

$$\begin{aligned}
\frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + b_{11}u_1 + \cdots + b_{1i}u_i \\
\frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + b_{21}u_1 + \cdots + b_{2i}u_i \\
&\vdots \\
\frac{dx_n}{dt} &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n + b_{n1}u_1 + \cdots + b_{ni}u_i
\end{aligned}$$

These state equations are compactly written in matrix notation as follows:

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{1i} \\ b_{21} & \cdots & b_{2i} \\ \vdots & & \\ b_{n1} & \cdots & b_{ni} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_i \end{bmatrix}$$

or

$$\frac{dx}{dt} = \dot{x} = Ax + Bu$$

The column matrix of state variables

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

is called the *state vector.* The inputs are arranged to form the *input vector*,

$$u = \begin{bmatrix} u_1 \\ \vdots \\ u_i \end{bmatrix}$$

The system outputs are similarly arranged in an *output vector*,

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$

related linearly to the state variables through the output equations:

$$\begin{cases} y_1 = c_{11}x_1 + c_{12}x_2 + \cdots + c_{1n}x_n \\ \vdots \\ y_m = c_{m1}x_1 + c_{m2}x_2 + \cdots + c_{mn}x_n \end{cases}$$

or

$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ \vdots & & & \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

or

$$y = Cx$$

The state equations describe how the system state vector evolves in time. One may imagine the tip of the vector tracing a curve, the *state trajectory*, in an n-dimensional space. The output equations describe how the output signals are related to the state.

For systems described by linear constant-coefficient integrodifferential equations, the state-variable arrangement is simply a standard form for the equations describing a system. Instead of dealing with a mixed collection of simultaneous system equations—some of first order, some of second order, some involving running integrals, and so on—additional manipulation of the original equations is done to place them in the standard form. The advantages of a standard form are that systematic methods may be easily brought to bear upon very involved problems and that a degree of unification results.

8.2.4 Physical State Variables

State space equations are sometimes written directly from first principles. For example, consider a standard series *RLC* circuit driven by a voltage source. From Kirchhoff's voltage law (KVL), we have

$$V_s = Ri(t) + L\frac{di}{dt} + \frac{1}{C}\int i(\tau)d\tau$$

There are many ways to convert this equation to state space form. It is possible, however, to obtain the state equations directly if we know how to choose the states. Voltage across capacitors and current through inductors are frequently chosen as (physical) state variables in circuit analysis. We therefore let

$$V_C = x_1 = \frac{1}{C}\int i(\tau)d\tau \qquad \text{and} \qquad I_L = x_2$$

Because the elements are in series, the loop current is equal to the inductor current. Therefore, from the foregoing definitions we have

$$\dot{x}_1 = \frac{1}{C}x_2$$

and using KVL, we also have

$$V_s = Rx_2 + L\dot{x}_2 + x_1$$

Rearranging, we get

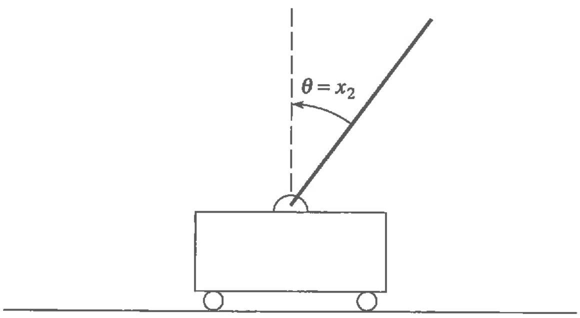
$$\dot{x}_2 = -\frac{1}{L}x_1 - \frac{R}{L}x_2 + \frac{V_s}{L}$$

The output equation depends on what we desire to control, or what we can measure, or both. In fact, in the most general setting, the controlled variables and the measured variables might be different. Suppose we can measure only the loop current (measured output, y), but we want to control the capacitor voltage (controlled output, z); the complete equations in vector–matrix form are (u stands for V_s)

$$\dot{x} = \begin{bmatrix} 0 & \frac{1}{C} \\ \frac{-1}{L} & \frac{-R}{L} \end{bmatrix} x + \begin{bmatrix} 0 \\ \frac{1}{L} \end{bmatrix} u$$

$$y = [0 \quad 1]x$$

$$z = [1 \quad 0]x$$



As another example, consider the well-known problem of balancing an inverted pendulum on a moving cart shown in the figure. The linearized equations of motion for this mechanical system are given by

$$(J + ml^2)\ddot{\theta} + ml\ddot{x} - mgl\theta = 0$$

$$(M + m)\ddot{x} + ml\ddot{\theta} = u$$

where $\theta =$ pendulum angle and $x =$ cart position. The system parameters are as follows: cart mass $M = 26/3$, pendulum mass $m = 4/3$, $l = 3/4 =$ (pendulum length)/2, gravity constant $g = 9.8$ m/s, and J, the pendulum moment of inertia about its center of gravity (assumed to be in the middle of the pendulum= $ml^2/3$. From these hypothetical values, we get

$$\ddot{\theta} + \ddot{x} - g\theta = 0 \quad [8.2]$$

$$10\ddot{x} + \ddot{\theta} = u \quad [8.3]$$

These equations are a set of two coupled second-order differential equations. How we define the state variables for this system depends on the control objectives. That is, the choice of the state variables and the control task are interrelated. We consider the following three situations.

Case I: The objective is to balance the pendulum (i.e., we do not care about the cart position or velocity). In this case, the pendulum angle and its angular velocity are chosen as the state variables. Solving for $\ddot{x}$ from Equation (8.2) and substituting in Equation (8.3), we get

$$\ddot{\theta} = \frac{10}{9}g\theta - \frac{1}{9}u$$

Letting $x_1 = \theta$ and $x_2 = \dot{\theta}$, and assuming we can measure the pendulum angle, we get

$$A = \begin{bmatrix} 0 & 1 \\ \frac{10g}{9} & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \frac{-1}{9} \end{bmatrix} \quad \text{and} \quad C = [1 \quad 0]$$

Case II: The objective is to balance the pendulum and stop the cart. The position of the cart is not important here as long as its velocity is brought down to zero. In this case, we add another state variable for the cart velocity. Writing and solving Equation (8.3) in terms of cart velocity v, we get

$$\dot{v} = -\frac{g}{9}\theta + \frac{1}{9}u$$

Using v as the third state variable, we get the following third-order system equation

$$A = \begin{bmatrix} 0 & 1 & 0 \\ \frac{10g}{9} & 0 & 0 \\ \frac{-g}{9} & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \frac{-1}{9} \\ \frac{1}{9} \end{bmatrix} \quad \text{and} \quad C = [1 \quad 0 \quad 0]$$

Case III: The objective is to balance the pendulum and position the cart (e.g., return it to its original position). In this case, we add the cart position as the fourth state variable.

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{10g}{9} & 0 & 0 & 0 \\ \frac{-g}{9} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ -\frac{1}{9} \\ \frac{1}{9} \\ 0 \end{bmatrix} \quad \text{and} \quad C = [1 \quad 0 \quad 0 \quad 0]$$

Note that as the task complexity increases, so will the order of the model used. Whether it is possible to accomplish the preceding tasks by measuring only the pendulum angle is another question. This is related to the notions of controllability and observability discussed later in the chapter.

❑ DRILL PROBLEMS

D8.2 Draw simulation diagrams for the given state space equations.

(a)
$$\begin{bmatrix} \dot{x} \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -2 & 5 \\ -4 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} -7 \\ 6 \end{bmatrix} u$$
$$y = [1 \quad 0] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

(b)
$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ -3 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$
$$y = [1 \quad 1] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

(c)
$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ -6 \\ 10 \end{bmatrix} u$$
$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

D8.3 Draw simulation diagrams to represent the following systems:

(a)
$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -2 & -3 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 5 \\ -6 \end{bmatrix} u$$
$$y = [7 \quad -8] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

(b)
$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 10 & 3 \\ 5 & 8 & 0 \\ -2 & -7 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 2 & 0 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$y = [0 \quad 4 \quad -3] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

(c)
$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -2 & 0 & -6 \\ 3 & 5 & 0 \\ -4 & 0 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 8 \\ -2 \\ 0 \end{bmatrix} u$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

8.2.5 Transfer Functions

The transfer functions of a system represented in state-variable form may be found by Laplace-transforming the state equations with zero initial conditions. In general, these are Equations (8.1). Collecting the terms involving $X(s)$, there results

$$\begin{bmatrix} (s - a_{11}) & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & (s - a_{22}) & \cdots & -a_{2n} \\ \vdots & & & \\ -a_{n1} & -a_{n2} & \cdots & (s - a_{nn}) \end{bmatrix} \begin{bmatrix} X_1(s) \\ X_2(s) \\ \vdots \\ X_n(s) \end{bmatrix}$$

$$= \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1i} \\ b_{21} & b_{22} & \cdots & b_{2i} \\ \vdots & & & \\ b_{n1} & b_{n2} & \cdots & b_{ni} \end{bmatrix} \begin{bmatrix} U_1(s) \\ U_2(s) \\ \vdots \\ U_i(s) \end{bmatrix}$$

or

$$[sI \quad -A]X(s) = BU(s)$$

where I is the $n \times n$ identity matrix

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & & & & \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

Solving for the Laplace transform of the state vector, we have

$$X(s) = [sI - A]^{-1} BU(s)$$

The output and state vectors are related by

$$\begin{bmatrix} Y_1(s) \\ Y_2(s) \\ \vdots \\ Y_m(s) \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & & & \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{bmatrix} \begin{bmatrix} X_1(s) \\ X_2(s) \\ \vdots \\ X_n(s) \end{bmatrix}$$

or

$$Y(s) = CX(s) = \left\{C[sI - A]^{-1}B\right\} U(s)$$

The $m \times i$ matrix in braces { } consists of the input–output transfer functions of the system, arranged as a matrix:

$$C[sI - A]^{-1}B = \begin{bmatrix} T_{11}(s) & T_{12}(s) & \cdots & T_{1i}(s) \\ T_{21}(s) & T_{22}(s) & \cdots & T_{2i}(s) \\ \vdots & & & \\ T_{m1}(s) & T_{m2}(s) & \cdots & T_{mi}(s) \end{bmatrix}$$

For example, a single-input, single-output system with state equations

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -3 & 1 \\ -2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 4 \\ -5 \end{bmatrix} u$$

$$y = [1 \quad -1] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

has transfer function given by

$$T(s) = [1 \quad -1] \begin{bmatrix} s+3 & -1 \\ 2 & s \end{bmatrix}^{-1} \begin{bmatrix} 4 \\ -5 \end{bmatrix}$$

$$= [1 \quad -1] \frac{\begin{bmatrix} s & 1 \\ -2 & s+3 \end{bmatrix} \begin{bmatrix} 4 \\ -5 \end{bmatrix}}{s^2 + 3s + 2}$$

$$= \frac{[1 \quad -1] \begin{bmatrix} (4s-5) \\ (-5s-23) \end{bmatrix}}{s^2 + 3s + 2}$$

$$= \frac{9s + 18}{s^2 + 3s + 2}$$

The two-input, two-output system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -3 & 1 \\ -2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 4 & 6 \\ -5 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 8 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

is described by the transfer function matrix given by

$$T(s) = \begin{bmatrix} 1 & -1 \\ 8 & 1 \end{bmatrix} \begin{bmatrix} s+3 & -1 \\ 2 & s \end{bmatrix}^{-1} \begin{bmatrix} 4 & 6 \\ -5 & 0 \end{bmatrix}$$

$$= \frac{\begin{bmatrix} 1 & -1 \\ 8 & 1 \end{bmatrix} \begin{bmatrix} s & 1 \\ -2 & s+3 \end{bmatrix} \begin{bmatrix} 4 & 6 \\ -5 & 0 \end{bmatrix}}{s^2+3s+2}$$

$$= \frac{\begin{bmatrix} 1 & -1 \\ 8 & 1 \end{bmatrix} \begin{bmatrix} (4s-5) & 6s \\ (-5s-23) & -12 \end{bmatrix}}{s^2+3s+2}$$

$$= \begin{bmatrix} \dfrac{9s+18}{s^2+3s+2} & \dfrac{6s+12}{s^2+3s+2} \\ \dfrac{27s-63}{s^2+3s+2} & \dfrac{48s-12}{s^2+3s+2} \end{bmatrix} = \begin{bmatrix} T_{11}(s) & T_{12}(s) \\ T_{21}(s) & T_{22}(s) \end{bmatrix}$$

Transfer function matrix.

where

$$T_{11}(s) = \frac{9s+18}{s^2+3s+2} = \left.\frac{Y_1(s)}{U_1(s)}\right|_{\substack{\text{initial} \\ \text{conditions} \\ \text{and } U_2=0}}$$

$$T_{12}(s) = \frac{6s+12}{s^2+3s+2} = \left.\frac{Y_1(s)}{U_2(s)}\right|_{\substack{\text{initial} \\ \text{conditions} \\ \text{and } U_1=0}}$$

$$T_{21}(s) = \frac{27s-63}{s^2+3s+2} = \left.\frac{Y_2(s)}{U_1(s)}\right|_{\substack{\text{initial} \\ \text{conditions} \\ \text{and } U_2=0}}$$

$$T_{22}(s) = \frac{48s-12}{s^2+3s+2} = \left.\frac{Y_2(s)}{U_2(s)}\right|_{\substack{\text{initial} \\ \text{conditions} \\ \text{and } U_1=0}}$$

All the transfer functions of a system share the denominator polynomial

$$|sI - A|$$

where A is the state coupling matrix for the system, since

$$[sI - A]^{-1} = \frac{\text{adjoint}[sI - A]}{|sI - A|}$$

The nth-degree polynomial

$$|sI - A| = 0$$

is termed the characteristic polynomial of an $n \times n$ matrix A, and the n roots of that polynomial are the eigenvalues, or characteristic roots, of the matrix. A system is stable if and only if the eigenvalues of the state coupling matrix are all in the left half of the complex plane.

❑ **DRILL PROBLEM**

D8.4 Find the transfer function matrices of the following systems:

(a)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -2 & 3 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 4 & 0 \\ -5 & 6 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$y = [7 \quad 8] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Ans. $\begin{bmatrix} \dfrac{-12s - 189}{s^2 + 3s + 5} & \dfrac{48s + 222}{s^2 + 3s + 5} \end{bmatrix}$

(b)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -3 & 4 \\ -2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} u$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} -4 & 6 \\ 5 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Ans. $\begin{bmatrix} \dfrac{-2s - 22}{s^2 + 3s + 8} \\ \dfrac{9s + 21}{s^2 + 3s + 8} \end{bmatrix}$

(c)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -4 & 0 & 2 \\ -1 & -1 & 0 \\ 3 & 0 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} -3 & -2 \\ 4 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Ans. $\begin{bmatrix} \dfrac{s^2 + 10s + 15}{(s + 1)(s^2 + 7s + 6)} & \dfrac{-s(s + 5)}{(s + 1)(s^2 + 7s + 6)} \\ \dfrac{3s}{s^2 + 7s + 6} & \dfrac{2s}{s^2 + 7s + 6} \end{bmatrix}$

8.3 State Transformations and Diagonalization

You have already observed that for a given system there is more than one state space representation. We have introduced two canonical forms, namely, the phase-variable and dual phase-variable forms. Hence, state space representation is not unique. In fact, if you change the way you label the states in a simulation diagram (number the states from left to right instead), you will obtain other forms. In general, there are infinitely many representations. These are generally called (state space) *realizations* of a system. Because these realizations correspond to the same transfer function, we want to determine how we can go from one state space realization to another.

The answer is that all state space realizations of the same system are related to each other via a linear transformation.

State transformation.

$$x(t) = Pz(t) \quad P = \text{nonsingular matrix}$$

where $x(t)$ represents the old state, and $z(t)$, the new state vector. How do we obtain the matrices corresponding to the new realization? We do this by differentiating both sides of the equation

$$\dot{x} = P\dot{z} = Ax + Bu = APz + Bu$$

Multiplying the equation on the left by the inverse of P, we get

The transformed equation.

$$\dot{z} = P^{-1}APz + P^{-1}Bu$$

The output equation becomes

$$y = Cx + Du = CPz + Du$$

Summarizing, we have that if the original realization is $\{A, B, C, D\}$, and the new realization is $\{\bar{A}, \bar{B}, \bar{C}, \bar{D}\}$, then the relation between the realizations is given by

$$\bar{A} = P^{-1}AP \qquad \bar{B} = P^{-1}B \qquad \bar{C} = CP \qquad \bar{D} = D$$

Let us demonstrate the procedure by an example. Consider the system shown in Figure 8.4:

$$\dot{x}_1 = x_1 + x_2$$

$$\dot{x}_2 = x_1 + x_2$$

$$y = x_1 + x_2$$

Define new state variables as

$$z_1 = \frac{x_1 - x_2}{2} \qquad \text{and} \qquad z_2 = \frac{x_1 + x_2}{2}$$

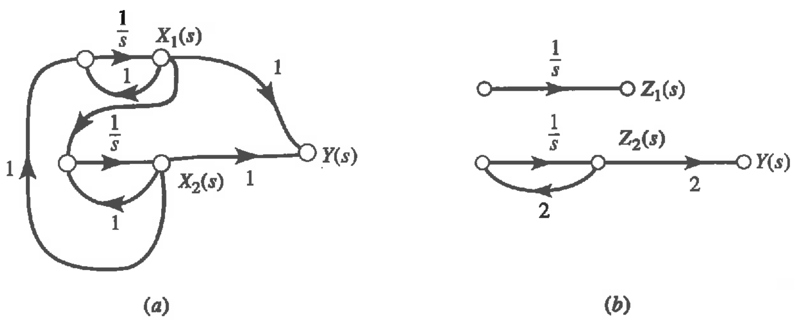


Figure 8.4 (a) Signal flow graph of a system. (b) Signal flow graph of the transformed system.

This can be written in matrix form as

$$z = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} x \quad \text{or} \quad x = Pz = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} z$$

Therefore the state space matrices will be transformed accordingly

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \to \bar{A} = P^{-1}AP$$

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

$$C = [1 \quad 1] \to \bar{C} = CP$$

$$= [1 \quad 1]\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = [0 \quad 2]$$

B and D are zero since the system has no inputs. Writing the new state equations, we get

$$\dot{z}_1 = 0$$

$$\dot{z}_2 = 2z_2$$

$$y = 2z_2$$

Observe that the system in this form is much easier to work with. In the original realization, the second-order system can be viewed as two coupled first-order systems. In the new realization, the same system appears as two uncoupled first-order systems. Every state space realization represents the same system, but each one allows us to look at the system from a different perspective. In modern control theory, state space transformations are used quite frequently for numerical purposes. Some realizations have superior numerical properties to others.

In the case of our present example, we note that even though we have not yet discussed how to solve state space equations, the solution in the new decoupled realization is almost trivial. The solution is

$$z_1(t) = z_1(0)$$

$$z_2(t) = e^{2t} z_2(0)$$

To obtain the solution to our original equation, we note that

$$x(t) = Pz(t) = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} z_1(0) \\ e^{2t} z_2(0) \end{bmatrix}$$

Also note that the initial conditions have to be transformed as

$$z(0) = P^{-1}x(0)$$

Suppose the initial conditions are given by $x(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$; then $z(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and the solution is

$$x(t) = \begin{bmatrix} e^{2t} \\ e^{2t} \end{bmatrix}$$

The input–output relations of a system are unchanged by a nonsingular change of state variables; it is only the internal description, in terms of its state, that is changed. This can be proved by the following line of argument

$$\bar{T}(s) = \bar{C}(sI - \bar{A})^{-1}\bar{B}$$

Substituting for $\bar{A}$, $\bar{B}$, and $\bar{C}$, we get

$$\bar{T}(s) = CP(sI - P^{-1}AP)^{-1}P^{-1}B$$

Let $I = P^{-1}P$, then

$$\bar{T}(s) = CP(sP^{-1}P - P^{-1}AP)^{-1}P^{-1}B = CP[P^{-1}(sI - A)P]^{-1}P^{-1}B$$

Recall from matrix algebra the identity: $(XYZ)^{-1} = Z^{-1}Y^{-1}X^{-1}$

$$\bar{T}(s) = CPP^{-1}(sI - A)^{-1}PP^{-1}B = C(sI - A)^{-1}B = T(s)$$

State transformations leave the transfer function unchanged.

❑ DRILL PROBLEM

D8.5 Make the indicated change of state variables, finding the new set of state and output equations in terms of z.

(a)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ -3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 4 \\ 5 \end{bmatrix} u$$

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

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

Ans.

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ -6 & -5 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 3.5 \\ -3 \end{bmatrix} u$$

$$y = [2 \quad 1] \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

(b)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 2 & -6 \\ 12 & 16 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$y = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

Ans.

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} -74 & -54 \\ 128 & 92 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} -2 & -3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$y = \begin{bmatrix} 2 & 6 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

(c)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -2 & 1 & 1 \\ -3 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} u$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$$

Ans.

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \end{bmatrix} = \begin{bmatrix} -2 & 1 & 5 \\ -3 & 0 & -3 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} + \begin{bmatrix} 3/4 \\ 1/4 \\ 1/4 \end{bmatrix} u$$

$$y = \begin{bmatrix} 2 & -2 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$$

8.3.1 Diagonal Forms

As shown in Figure 8.1(a), a simulation diagram results in a set of system matrices when a transfer function is properly decomposed. Phase variables or dual phase variables result in a recognizable form for the A matrix. In this section, a properly chosen change of variables can create a set of A, B, and C matrices in which the A

matrix has a very simple (diagonal) form. The state variables [Figure 8.1(c)] are often called *canonical* when the A matrix becomes diagonal.

When a nonsingular change of state variables in a system representation is made,

$$z = P^{-1}x, \qquad x = Pz$$

the new state coupling matrix $\bar{A}$ is related to the original one A by

$$\bar{A} = P^{-1}AP$$

Such an operation on a matrix is termed a *similarity transformation*. One of the most important results of matrix algebra is that, provided a square matrix A has no repeated eigenvalues, a similarity transformation P may be found for which

$$\bar{A} = P^{-1}AP$$

is diagonal, with the eigenvalues as the diagonal elements.

A similarity transformation that diagonalizes A can be created using a set of eigenvectors, one for each eigenvalue. The German word "eigen" means "characteristic." These eigenvectors are not unique. Each eigenvector can be multiplied by a constant that works just as well. The following procedure can be used to get P.

1. Find the eigenvalues s_i where

$$|sI - A| = 0$$

2. Find an eigenvector x_i for each s_i

$$[s_iI - A]x_i = 0$$

3. Let P be a matrix consisting of the eigenvectors

$$P = [x_1 : x_2 : \cdots : x_n]$$

$$P^{-1}AP = \begin{bmatrix} s_1 & 0 & 0 & . \\ 0 & s_2 & 0 & . \\ \text{etc.} & & & \end{bmatrix}$$

In the example just analyzed, the P matrix was actually computed by means of the same procedure. You can verify that the eigenvalues of the system are 0 and 2, and the eigenvectors are the columns of P.

As another example, consider

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -1 & -2 & 0 \\ 1 & 2 & 0 \\ -2 & -1 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u = Ax + bu$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = Cx$$

The characteristic polynomial is

$$\begin{vmatrix} s+1 & 2 & 0 \\ -1 & s-2 & 0 \\ 2 & 1 & s+3 \end{vmatrix} = s(s-1)(s+3)$$

The eigenvalues can be selected in any order. If

$$s_1 = 0 \quad s_2 = 1 \quad s_3 = -3$$

Then the first eigenvector is found by solving

$$\begin{bmatrix} s_1+1 & 2 & 0 \\ -1 & s_1-2 & 0 \\ 2 & 1 & s_1+3 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$x_1 = \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \end{bmatrix}$$

With the first eigenvalue being zero

$$\begin{aligned} x_{11} + 2x_{21} \qquad &= 0 \\ -x_{11} - 2x_{21} \qquad &= 0 \\ 2x_{11} + x_{21} + 3x_{31} &= 0 \end{aligned}$$

The first two equations are equivalent, so an infinite number of solutions exist. It is possible to select x_{21} arbitrarily (say -1). Then x_{11} is 2 and x_{31} is -1. By proceeding in a similar way, three eigenvectors result:

$$x_1 = \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix} \quad x_2 = \begin{bmatrix} 4 \\ -4 \\ -1 \end{bmatrix} \quad x_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Any nonzero multiple of any eigenvector also works. Collecting these eigenvectors into P, the transformation

$$P = \begin{bmatrix} 2 & 4 & 0 \\ -1 & -4 & 0 \\ -1 & -1 & 1 \end{bmatrix} \qquad P^{-1} = \begin{bmatrix} 1 & 1 & 0 \\ -\frac{1}{4} & -\frac{1}{2} & 0 \\ \frac{3}{4} & \frac{1}{2} & 1 \end{bmatrix}$$

gives a state-variable representation for which the state coupling matrix is diagonal:

$$\bar{A} = P^{-1}AP = \begin{bmatrix} 1 & 1 & 0 \\ -\frac{1}{4} & -\frac{1}{2} & 0 \\ \frac{3}{4} & \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} -1 & -2 & 0 \\ 1 & 2 & 0 \\ -2 & -1 & -3 \end{bmatrix} \begin{bmatrix} 2 & 4 & 0 \\ -1 & -4 & 0 \\ -1 & -1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 & 0 \\ -\frac{1}{4} & -\frac{1}{2} & 0 \\ \frac{3}{4} & \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 0 & 4 & 0 \\ 0 & -4 & 0 \\ 0 & -1 & -3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

$$\bar{B} = P^{-1}B = \begin{bmatrix} 1 & 1 & 0 \\ -\frac{1}{4} & -\frac{1}{2} & 0 \\ \frac{3}{4} & \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -\frac{1}{4} \\ \frac{3}{4} \end{bmatrix}$$

$$\bar{C} = CP = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 4 & 0 \\ -1 & -4 & 0 \\ -1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 1 \\ 3 & 8 & 0 \end{bmatrix}$$

The system described by the new state variables,

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} + \begin{bmatrix} 1 \\ -\frac{1}{4} \\ \frac{3}{4} \end{bmatrix} u$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 1 \\ 3 & 8 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$$

has the same relation between u and y. Because the state coupling matrix is diagonal, however, the state equations are decoupled from one another. The system is represented in the form of three separate first-order systems, as in the simulation diagram of Figure 8.5.

Finding a transformation (eigenvalues) matrix that diagonalizes a square matrix A with distinct characteristic roots is a fundamental technique of linear algebra. It is termed the *characteristic value problem* and is discussed in detail in most texts on linear algebra, including those cited in the references at the end of this chapter.

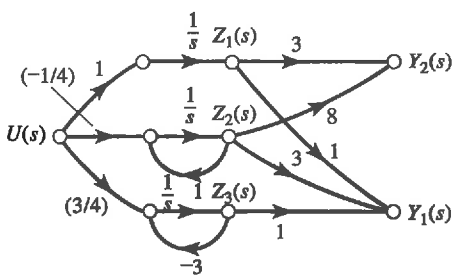


Figure 8.5 Simulation diagram for the diagonalized example system.

8.3.2 Diagonalization Using Partial Fraction Expansion

Another method of determining a diagonal form for a system involves partial fraction expansion. For a single-input, single-output system such as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -1 & -2 & 0 \\ 1 & 2 & 0 \\ -2 & -1 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u$$

$$y = [1 \quad 0 \quad 1] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

the transfer function is

$$\begin{aligned} T(s) &= C(sI - A)^{-1}B \\ &= [1 \quad 0 \quad 1] \begin{bmatrix} s+1 & 2 & 0 \\ -1 & s-2 & 0 \\ 2 & 1 & s+3 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \\ &= [1 \quad 0 \quad 1] \frac{\begin{bmatrix} s^2+s-6 & -2s-6 & 0 \\ s+3 & s^2+4s+3 & 0 \\ -2s+3 & -s+3 & s^2-s \end{bmatrix}}{s^3+2s^2-3s} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \\ &= \frac{[1 \quad 0 \quad 1]}{s^3+2s^2-3s} \begin{bmatrix} (s^2+s-6) \\ (s+3) \\ (-2s+3) \end{bmatrix} = \frac{s^2-s-3}{s^3+2s^2-3s} \end{aligned}$$

Upon expanding this transfer function in partial fractions, there results

$$T(s) = \frac{s^2-s-3}{s(s-1)(s+3)} = \frac{1}{s} + \frac{-\frac{3}{4}}{s-1} + \frac{\frac{3}{4}}{s+3}$$

which may be considered as the tandem (or parallel) connection of first-order systems shown in Figure 8.6(a). Each of these first-order subsystems is drawn in state-variable form in Figure 8.6(b), where the three integrator output signals are labeled as state-variables. The state-variable equations for this alternate system representation, which has the same transfer function as the original system, are

$$\dot{z}_1 = u$$

$$\dot{z}_2 = z_2 + u$$

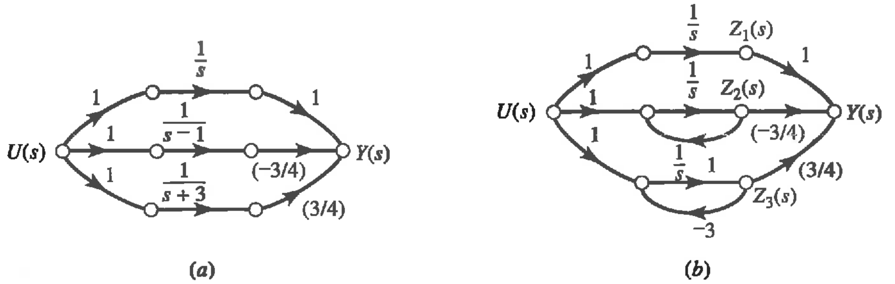


Figure 8.6 Diagonalizing a single-input, single-output system. (a) Tandem first-order sub systems from the partial fraction expansion of the transfer function. (b) Subsystems in simulation diagram form.

$$\dot{z}_3 = -3z_3 + u$$
$$y = z_1 - \tfrac{3}{4}z_2 + \tfrac{3}{4}z_3$$

or

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & -\tfrac{3}{4} & \tfrac{3}{4} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$$

which is diagonal.

❑ DRILL PROBLEMS

D8.6 Use the partial fraction method to find diagonal state equations for single–input, single–output systems with the following transfer functions:

(a)

$$T(s) = \frac{-5s + 7}{s^2 + 7s + 12}$$

Ans.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -3 & 0 \\ 0 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 22 \\ -27 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

(b)

$$T(s) = \frac{3s^2 - 2}{(s + 1)(s + 4)(s + 10)}$$

Ans.
$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} \frac{1}{27} \\ -\frac{23}{9} \\ \frac{149}{27} \end{bmatrix} u$$

$$y = [1 \quad 1 \quad 1] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

(c)
$$T(s) = \frac{4}{s^3 + 3s^2 + 2s}$$

Ans.
$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} u$$

$$y = [2 \quad -4 \quad 2] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

D8.7 Draw a simulation diagram for each of the following state equations.

(a)
$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} -3 & 0 \\ 0 & -4 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 8 \\ 4 \end{bmatrix} u$$

$$y = [1 \quad 1] \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

(b)
$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & -4 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u$$

$$y = [6 \quad -9] \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

8.3.3 Complex Conjugate Characteristic Roots

In general, diagonalized state equations for systems with complex characteristic roots involve state equations with complex coefficients. For example, the single-input, single-output system with transfer function

$$T(s) = \frac{6s^2 + 26s + 8}{(s+2)(s^2 + 2s + 10)} = \frac{-2}{s+2} + \frac{4+j}{s+1+j3} + \frac{4-j}{s+1-j3}$$

may be represented in terms of state variables as in the simulation diagram of Figure 8.7(a). The gains associated with the complex characteristic roots are generally complex numbers. The state equations, in terms of the indicated state variables, are given by

$$\begin{aligned} sX_1(s) &= -2X_1(s) + U(s) \\ sX_2(s) &= (-1 - j3)X_2(s) + U(s) \\ sX_3(s) &= (-1 + j3)X_3(s) + U(s) \\ Y(s) &= -2X_1(s) + (4 + j)X_2(s) + (4 - j)X_3(s) \end{aligned}$$

or

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -1 - j3 & 0 \\ 0 & 0 & -1 + j3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} u$$

$$y = [-2 \quad 4 + j \quad 4 - j] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Although the individual physical components of this representation involve complex numbers, hence cannot be assembled, the mathematical relationships are valid. To build such a system, or to represent it in a convenient form that does not involve complex numbers, the two complex conjugate component parts may be combined just as one commonly combines the corresponding conjugate partial fraction terms:

$$\frac{4 + j}{s + 1 + j3} + \frac{4 - j}{s + 1 - j3} = \frac{8s + 14}{s^2 + 2s + 10}$$

This portion of the system may be represented in phase-variable form, giving the real-number simulation diagram of Figure 8.7(b). The state equations for this alternative arrangement are given by

$$\begin{aligned} sZ_1(s) &= -2Z_1(s) + U(s) \\ sZ_2(s) &= Z_3(s) \\ sZ_3(s) &= -10Z_2(s) - 2Z_3(s) + U(s) \\ Y(s) &= -2Z_1(s) + 14Z_2(s) + 8Z_3(s) \end{aligned}$$

or

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -10 & -2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} u$$

$$y = [-2 \quad 14 \quad 8] \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$$

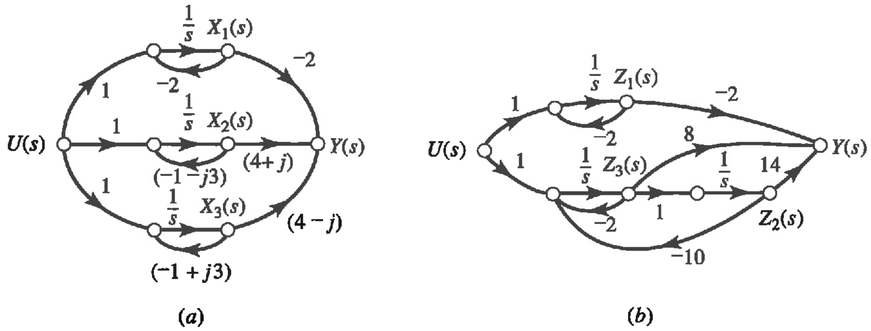


Figure 8.7 A system with complex characteristic roots. (a) Diagonalized system. (b) Alternative form for the diagonalized system, where the complex conjugate root terms have been combined and placed in phase-variable form.

It is thus possible to represent systems with one or more pairs of complex conjugate characteristic roots with diagonalized state equations involving complex numbers or in block diagonal form involving real numbers. For example, the following state equations

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \\ \dot{x}_6 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -17 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -10 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} 6 & -8 & 1 & -5 & 0 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix}$$

are in block diagonal form and represent a system with transfer function

$$T(s) = \frac{6}{s-3} + \frac{-8}{s+4} + \frac{-5s+1}{s^2+2s+17} + \frac{7s}{s^2-3s+10}$$

❑ DRILL PROBLEMS

D8.8 The following transfer functions for single-input, single-output systems involve complex characteristic roots. Find diagonal state equations for these systems. Then find an alternative block diagonal representation that does not involve complex numbers.

(a)

$$T(s) = \frac{10}{s^3 + 2s^2 + 5s}$$

Ans.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -5 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & -4 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

(b)

$$T(s) = \frac{3s^2 - 1}{(s^2 + 4)(s^2 + 4s + 5)}$$

Ans.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -5 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} -\frac{1}{5} & \frac{4}{5} & 0 & -\frac{4}{5} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

(c)

$$T(s) = \frac{s^2 - 4s + 10}{(s + 2)(s^2 + 6s + 13)}$$

Ans.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -13 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} \frac{22}{5} & \frac{118}{5} & -\frac{17}{5} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

8.3.4 Repeated Characteristic Roots

The state equations for a system with repeated characteristic roots may not necessarily be diagonalized. A block diagonal form, termed a *Jordan canonical form*, is commonly used when a simple representation is desired. For example, the single-input, single-output system with transfer function

$$T(s) = \frac{10s^2 + 51s + 56}{(s+4)(s+2)^2} = \frac{3}{s+4} + \frac{-6}{s+2} + \frac{7}{(s+2)^2}$$

may be represented as in Figure 8.8(a). A simplification results when the $1/(s+2)$ transmittance is used in common by two paths, as shown in Figure 8.8(b). A corresponding state-variable representation is given in Figure 8.8(c).

The state equations are given by

$$\begin{aligned} sX_1(s) &= -4X_1(s) + U(s) \\ sX_2(s) &= -2X_2(s) + X_3(s) \\ sX_3(s) &= -2X_3(s) + U(s) \\ Y(s) &= 3X_1(s) + 7X_2(s) - 6X_3(s) \end{aligned}$$

or

Jordan form.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -4 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} u$$

$$y = [3 \quad 7 \quad -6] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

For three repetitions of a characteristic root, the corresponding transfer function partial fraction terms are

$$\frac{k_1}{s+a} + \frac{k_2}{(s+a)^2} + \frac{k_3}{(s+a)^3}$$

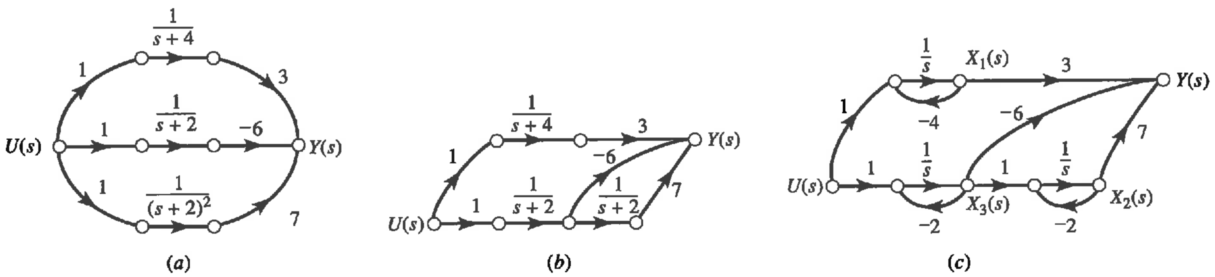


Figure 8.8 State equations for a system with repeated characteristic roots. (a) Diagram showing each partial fraction term. (b) Diagram with common signal path through a repeated transmittance. (c) Diagram showing state variables.

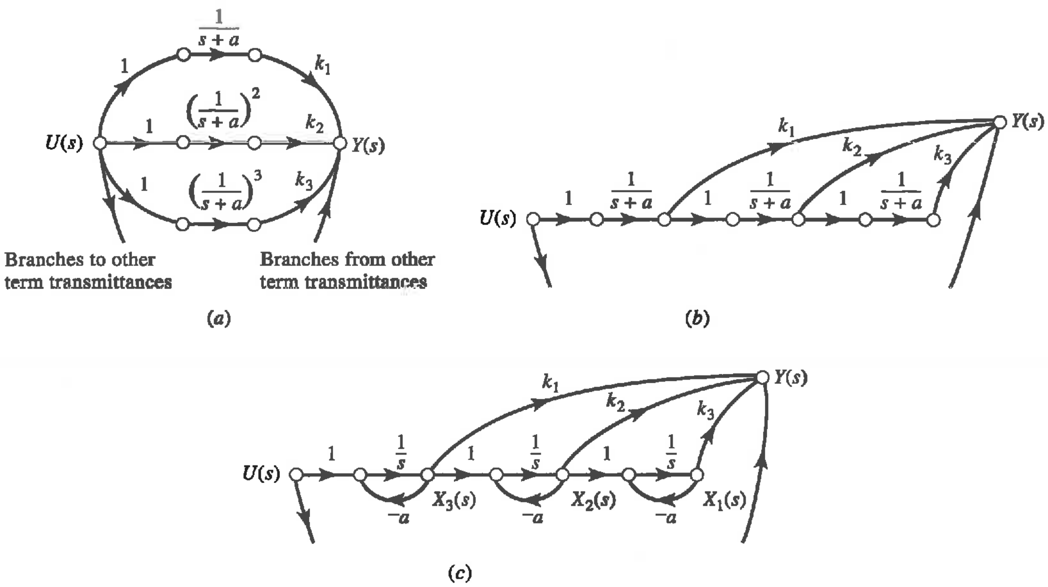


Figure 8.9 State variables for repeated roots. (a) Diagram showing each partial fraction term. (b) Diagram using common signal paths. (c) Diagram showing state variables.

and the state variables may be defined as in Figure 8.9. The resulting Jordan block has the following structure:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} -a & 1 & 0 & 0 & 0 & \cdots \\ 0 & -a & 1 & 0 & 0 & \cdots \\ 0 & 0 & -a & 0 & 0 & \cdots \\ 0 & 0 & 0 & & & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \end{bmatrix} u,$$

$$y = [k_3 \quad k_2 \quad k_1 \quad \cdots] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \end{bmatrix}$$

The state-variable equations

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \\ \dot{x}_6 \end{bmatrix} = \begin{bmatrix} -2 & 1 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & -3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} u$$

$$y = [4 \quad -5 \quad 6 \quad 7 \quad -8 \quad 9]\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix}$$

for example, are in Jordan canonical form. They represent a system with transfer function

$$T(s) = \frac{-5}{s+2} + \frac{4}{(s+2)^2} + \frac{6}{s+3} + \frac{9}{s-4} + \frac{-8}{(s-4)^2} + \frac{7}{(s-4)^3}$$

❑ DRILL PROBLEM

D8.9 The following systems have repeated characteristic roots. Find an alternate set of state equations in Jordan canonical form.

(a)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 2 & 9 \\ -1 & -4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 4 \\ -3 \end{bmatrix} u$$

$$y = [2 \quad -6]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Ans.

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$y = [-60 \quad 26]\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

(b)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -6 & 1 & 0 \\ -9 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} u$$

$$y = [3 \quad -2 \quad 1]\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Ans.
$$\begin{bmatrix}\dot{z}_1\\ \dot{z}_2\\ \dot{z}_3\end{bmatrix}=\begin{bmatrix}-3 & 1 & 0\\ 0 & -3 & 1\\ 0 & 0 & -3\end{bmatrix}\begin{bmatrix}z_1\\ z_2\\ z_3\end{bmatrix}+\begin{bmatrix}0\\ 0\\ 1\end{bmatrix}u$$

$$y=\begin{bmatrix}2 & -2 & 0\end{bmatrix}\begin{bmatrix}z_1\\ z_2\\ z_3\end{bmatrix}$$

❑ Computer-Aided Learning

To create a state space model, we use the "ss" command with the following syntax:

```
SYS=SS(A,B,C,D)
```

For example, to define the following system:

$$\dot{x}=\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}x+\begin{pmatrix}-1\\ 6\end{pmatrix}u$$

$$y=(0 \quad -14)x+8u$$

we enter the following commands

```
>> a=[1, 2; 3, 4]; b=[-1; 6]; c=[0, -14]; d=8;
>> g=ss(a,b,c,d)
```

and MATLAB responds

```
a=
            x1          x2
     x1   1.00000     2.00000
     x2   3.00000     4.00000
b=
            u1
     x1  -1.00000
     x2   6.00000
c=
            x1          x2
     y1     0        -14.00000
d=
            u1
     y1   8.00000
```

Continuous-Time System

We can convert from state space to transfer function form using the "tf" command.

```
>> g_s=tf(g)
```

Transfer function:

```
  8 s^2-124 s+110
-----------------
    s^2-5 s-2
```

We can then transfer back to state space form using the "ss" command

```
g_ss=ss(g_s)
a=
                x1          x2
      x1    5.00000     0.50000
      x2    4.00000           0
b=
                u1
      x1    8.00000
      x2          0
c=
                x1          x2
      y1  -10.50000     3.93750
d=
                u1
      y1    8.00000
```

Continuous-Time System

Note that the A, B, C, D matrices that MATLAB returns are different from the ones we used to define the original system we called g. The reason as explained in the text is that the state space representation is not unique.

Moreover the "tf" and "ss" commands create system objects that MATLAB commands can interpret. If we want direct access to the object features such as numerator, denominator, and the $\{A, B, C, D\}$ matrices, we need to extract them. To extract the $\{A, B, C, D\}$ matrices, we use the "ssdata" command.

```
>>[a,b,c,d]=ssdata(g_ss)
   a=
     5.0000     0.5000
     4.0000          0
   b=
     8
     0
   c=
     -10.5000   3.9375
   d=  8
```

We can now use these matrices to determine stability, controllability and other system properties.

The system can be transformed to other forms using the "canon" and "ss2ss" commands. The "canon" command has the following syntax:

```
sc=canon(sys,'type')
```

where type is either 'companion' or 'modal.' The latter diagonalizes the system (also known as the *modal realization*), and the former converts it to companion form (a variation of the phase or dual-phase variable forms)

```
>>sd=canon(g_ss,'modal')
  a=
              x1           x2
       x1   5.37228         0
       x2      0        -0.37228
  b=
              u1
       x1   9.32759
       x2  -5.59456
  c=
              x1           x2
       y1  -6.07043     4.89359
  d=
              u1
       y1   8.00000
>>[sd]=canon(g_ss,'com')
  a=
              x1           x2
       x1      0        2.00000
       x2   1.00000     5.00000
  b=
              u1
       x1   1.00000
       x2      0
  c=
              x1           x2
       y1  -84.00000  -294.00000
  d=
              u1
       y1   8.00000
```

The "ss2ss" command has the syntax

`sys2=ss2ss(sys1,P)` or `[a2,b2,c2,d2]=ss2ss(a,b,c,d,P)`

where P is the nonsingular state transformation matrix. Note that our definition of P is different from that of MATLAB, which defines P by

`z(t)=P x(t)` or `new state=P.old state`

We, on the other hand, use

`x(t)=P z(t)` or `old state=P.new state`

Hence our P is the inverse of what MATLAB uses. To get the answers in the book using the "ss2ss" command use inv(p) instead of p.

Here is an example:

$$x = \begin{pmatrix} -2 & 1 & 1 \\ -3 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} x + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} u$$

$$y = \begin{pmatrix} 2 & -2 & 1 \\ 0 & -1 & 1 \end{pmatrix} x$$

Use the transformation matrix $P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 4 \end{pmatrix}$

```
>>p=[1 0 1;0 1 3;0 0 4];
>>a=[-2 1 1;-3 0 0;0 0 0];
>>b=[1;1;1];
>>c=[2 -2 1;0 -1 1]; d=[0; 0];
>>g=ss(a,b,c,d);
>>tf(g) % the transfer function of the 2-output system
```

Transfer function from input to output

```
       s^2+8s+9
 #1:  ----------
      s^3+2s^2+3s
         3s+6
 #2:  ----------
      s^3+2s^2+3s
```

To transform using the given P matrix we use "inv (p)" to get the matrix P^{-1}. We then extract the new $\{A, B, C, D\}$ matrices to verify our answers:

```
[a2,b2,c2,d2]=ss2ss(a,b,c,d,inv(p))

a2=
   -2   1   5
   -3   0  -3
    0   0   0
b2=
    0.7500
    0.2500
    0.2500
c2=
   2   -2   0
   0   -1   1
d2=
   0
   0
```

Compare the foregoing answers with the ones from the formulas: $\bar{A} = P^{-1}AP$, $\bar{B} = P^{-1}B$, $\bar{C} = CP$

```
>>p\a*p
ans=
   -2    1   5
   -3    0  -3
    0    0   0
>>p\b
ans=
        0.7500
        0.2500
        0.2500
>>c*p
ans=
    2   -2   0
    0   -1   1
```

C8.1

(a) Find the transfer function of the systems defined in Drill Problem D8.4.

(b) Redo Drill Problem D8.5.

(c) Convert the systems defined in Drill Problem D8.6 to state space form, and then diagonalize using the "canon" command to verify the answers.

8.4 Time Response from State Equations

8.4.1 Laplace Transform Solution

One method of calculating the state of a system as a function of time is to Laplace-transform the equations, solve for the transform of the signals of interest, then invert the transforms. The system outputs, being linear combinations of the state signals, are easily found from the state.

For example, consider the system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -6 & 1 \\ -5 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$y = [1 \quad -2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

with initial state

$$\begin{bmatrix} x_1(0^-) \\ x_2(0^-) \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$$

and input

$$u(t) = 7$$

where 7 represents a step function starting at $t = 0$.

The Laplace-transformed state equations are as follows:

$$\begin{cases} sX_1(s) + 3 = -6X_1(s) + X_2(s) \\ sX_2(s) - 1 = -5X_1(s) + \dfrac{7}{s} \end{cases}$$

$$\begin{cases} (s+6)X_1(s) - X_2(s) = -3 \\ 5X_1(s) + sX_2(s) = 1 + \dfrac{7}{s} = \dfrac{s+7}{s} \end{cases}$$

$$X_1(s) = \frac{\begin{vmatrix} -3 & -1 \\ (s+7)/s & s \end{vmatrix}}{\begin{vmatrix} s+6 & -1 \\ 5 & s \end{vmatrix}} = \frac{-3s + (s+7)/s}{s^2 + 6s + 5}$$

$$= \frac{-3s^2 + s + 7}{s(s+1)(s+5)} = \frac{\frac{7}{5}}{s} + \frac{-\frac{3}{4}}{s+1} + \frac{-\frac{73}{20}}{s+5}$$

$$x_1(t) = \tfrac{7}{5} - \tfrac{3}{4}e^{-t} - \tfrac{73}{20}e^{-5t} \qquad t \geqslant 0$$

8.4.2 Time Domain Response of First-Order Systems

In many situations, it is advantageous to have an expression for the solution of a set of state equations as functions of time rather than in terms of Laplace transforms. For a first-order state-variable system, we write

$$\frac{dx}{dt} = ax + bu$$

$$sX(s) - x(0^-) = aX(s) + bu(s)$$

$$X(s) = \frac{x(0^-)}{s-a} + bu(s)\frac{1}{s-a}$$

$$x(t) = \mathcal{L}^{-1}\left\{\frac{x(0^-)}{s-a} + bu(s)\frac{1}{s-a}\right\}$$

$$= e^{at}x(0^-) + \text{convolution } [bu(t), e^{at}]$$

$$= e^{at}x(0^-) + \int_{0^-}^{t} e^{a(t-\tau)}bu(\tau)d\tau$$

the inverse transform of a product of Laplace transforms being the convolution of the corresponding time functions.

As a numerical example, consider the first-order system

$$\dot{x} = -2x + 3$$

$$y = 4x$$

The general solution for $x(t)$ is

$$x(t) = e^{-2t}x(0^-) + \int_{0^-}^{t} 3e^{-2(t-\tau)}u(\tau)d\tau$$

If

$$x(0^-) = 10 \quad and\ u = 5$$

then

$$\begin{aligned} x(t) &= 10e^{-2t} + \int_{0^-}^{t} 15e^{-2(t-\tau)}d\tau \\ &= 10e^{-2t} + 15e^{-2t}\left.\frac{e^{2\tau}}{2}\right|_{0^-}^{t} \\ &= 10e^{-2t} + 15e^{-2t}\frac{e^{-2t}-1}{2} \\ &= \tfrac{5}{2}e^{-2t} + \tfrac{15}{2} \quad t \geqslant 0 \end{aligned}$$

and the system output is

$$y(t) = 4x(t) = 10e^{-2t} + 30 \quad t \geqslant 0$$

8.4.3 Time Domain Response of Higher-Order Systems

In general, a state-variable system

$$\dot{x} = Ax + Bu$$

has state response given by

$$\begin{aligned} &sX(s) - x(0^-) = AX(s) + BU(s) \\ &[sI - A]X(s) = x(0^-) + BU(s) \\ &X(s) = [sI - A]^{-1}x(0^-) + [sI = -A]^{-1}BU(s) \end{aligned}$$

Denoting the *resolvent matrix* by $\Phi(s) = (sI - A)^{-1}$ and its inverse Laplace transform, the *state transition matrix*, by

$$\Phi(t) = \mathcal{L}^{-1}\{[sI - A]^{-1}\}$$

then

Solution of state equations.

$$x(t) = \Phi(t)x(0^-) + \text{convolution } [Bu(t), \Phi(t)]$$

$$\boxed{= \Phi(t)x(0^-) + \int_0^t \Phi(t-\tau)Bu(\tau)d\tau}$$

For example, for the system

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -3 & 1 \\ -2 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 2 \\ -1 \end{bmatrix} u$$

the state transition matrix is given by

$$\begin{aligned} \Phi(t) &= \mathcal{L}^{-1}\{[sI - A]^{-1}\} \\ &= \mathcal{L}^{-1}\left\{\begin{bmatrix} s+3 & -1 \\ 2 & s \end{bmatrix}^{-1}\right\} \end{aligned}$$

$$= \mathcal{L}^{-1}\begin{bmatrix} \dfrac{s}{s^2+3s+2} & \dfrac{1}{s^2+3s+2} \\ \dfrac{-2}{s^2+3s+2} & \dfrac{s+3}{s^2+3s+2} \end{bmatrix}$$

$$= \mathcal{L}^{-1}\begin{bmatrix} \dfrac{-1}{s+1}+\dfrac{2}{s+2} & \dfrac{1}{s+1}+\dfrac{-1}{s+2} \\ \dfrac{-2}{s+1}+\dfrac{2}{s+2} & \dfrac{2}{s+1}+\dfrac{-1}{s+2} \end{bmatrix}$$

$$= \begin{bmatrix} -e^{-t}+2e^{-2t} & e^{-t}-e^{-2t} \\ -2e^{-t}+2e^{-2t} & 2e^{-t}-e^{-2t} \end{bmatrix}$$

The system state is, in terms of initial conditions and the inputs,

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} -e^{-t}+2e^{-2t} & e^{-t}-e^{-2t} \\ -2e^{-t}+2e^{-2t} & 2e^{-t}-e^{-2t} \end{bmatrix}\begin{bmatrix} x_1(0^-) \\ x_2(0^-) \end{bmatrix}$$

$$+ \int_{0^-}^{t} \begin{bmatrix} -e^{-(t-\tau)}+2e^{-2(t-\tau)} & e^{-(t-\tau)}-e^{-2(t-\tau)} \\ -2e^{-(t-\tau)}+2e^{-2(t-\tau)} & 2e^{-(t-\tau)}-e^{-2(t-\tau)} \end{bmatrix}$$

$$\times \begin{bmatrix} 2 \\ -1 \end{bmatrix} u(\tau)d\tau$$

❑ DRILL PROBLEMS

D8.10 Use Laplace transform methods to find the outputs of the following systems for $t \geqslant 0$ with the given inputs and initial conditions:

(a)

$$\dot{x} = -2x + u(t)$$

$$y = 10x$$

$$x(0^-) = 3$$

$$u(t) = 4e^{5t}$$

Ans. $\frac{170}{7}e^{-2t} + \frac{40}{7}e^{5t}$

(b)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -12 & -7 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u$$

$$y = [1 \quad -1]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{bmatrix} x_1(0^-) \\ x_2(0^-) \end{bmatrix} = \begin{bmatrix} 10 \\ 0 \end{bmatrix}$$

where the input $u(t)$ is the unit step function.

Ans. $Y(s) = (10s^2 + 190s + 20)/s(s+3)(s+4)$

$y(t) = \frac{5}{3} + \frac{460}{3}e^{-3t} - 145e^{-4t} \quad t \geqslant 0$

(c)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -5 & 1 & 0 \\ -6 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u$$

$$y = [1 \quad 0 \quad 0] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$x(0) = 0$

$u(t) = \delta(t)$, where $\delta(t)$ is the unit impulse.

Ans. $\frac{1}{6} + \frac{1}{3}e^{-3t} - \frac{1}{2}e^{-2t} \quad t \geqslant 0$

D8.11 Calculate state transition matrices for system with the following state coupling matrices A, using

$\Phi(t) = \mathcal{L}^{-1}\left\{[sI - A]^{-1}\right\}$:

(a)

$$\begin{bmatrix} -9 & 1 \\ -14 & 0 \end{bmatrix}$$

Ans. $\begin{bmatrix} -\frac{2}{5}e^{-2t} + \frac{7}{5}e^{-7t} & \frac{1}{5}e^{-2t} - \frac{1}{5}e^{-7t} \\ -\frac{14}{5}e^{-2t} + \frac{14}{5}e^{-7t} & \frac{7}{5}e^{-2t} - \frac{2}{5}e^{-7t} \end{bmatrix}$

(b)

$$\begin{bmatrix} 1 & -1 \\ 2 & -4 \end{bmatrix}$$

Ans. $\begin{bmatrix} 1.11e^{0.56t} - 0.11e^{-3.56t} & -0.24e^{0.56t} + 0.24e^{-3.56t} \\ 0.48e^{0.56t} - 0.48e^{-3.56t} & -0.11e^{0.56t} + 1.11e^{-3.56t} \end{bmatrix}$

8.4.4 System Response Computation

One advantage of placing system equations in a state-variable form is that it is well suited to digital computer calculations. Computers are not particularly efficient at equation manipulation, Laplace transformation, and the like, but they excel at such repetitive tasks as matrix addition and multiplication. The capability of simulating a system, that is, investigating and testing its performance by modeling, is important to the designer, particularly for the common situation in which the plant is expensive and the design must be correct when it is first installed.

The state transition matrix can be approximated by an $(m+1)$-term Taylor series

$$\Phi(\Delta t) = I + A(\Delta t) + \frac{1}{2}A^2(\Delta t)^2 + \cdots + \left(\frac{1}{m!}\right)A^m(\Delta t)^m$$

The convolution integral depends on the state transition matrix and on the input, both of which can be functions of time. However, if Δt is a very short time, then $u(t)$ can be removed from the integral so that

$$\text{Convolution integral} = D(\Delta t)Bu(t)$$

$$D(\Delta t) = I\Delta t + \frac{1}{2}A(\Delta t)^2 + \left(\frac{1}{3!}\right)A^2(\Delta t)^3 + \cdots + \left(\frac{1}{(m+1)!}\right)A^m(\Delta t)^{m+1}$$

For sufficiently small time increments Δt, one can start with the initial state $x(0)$ and calculate $x(\Delta t)$ as follows:

$$x(\Delta t) \cong (I + A\Delta t)\,x(0) + (B\Delta t)\,u(0)$$

then $x(2\Delta t)$ may be calculated from $x(\Delta t)$,

$$x(2\Delta t) \cong (I + A\Delta t)\,x(\Delta t) + (B\Delta t)\,u(\Delta t)$$

and so on, obtaining approximate solutions for the state,

$$x\{(k+1)\,\Delta t\} \cong (I + A\Delta t)\,x\,(k\Delta t) + (B\Delta t)\,u\,(k\Delta t)$$

For example, the response of the first-order system

$$\dot{x} = -2x + u$$
$$y = x$$

with

$$x(0^-) = 10$$
$$u(t) = 3\sin t$$

is approximated by

$$x\{(k+1)\,\Delta t\} \cong (1 - 2\Delta t)\,x\,(k\Delta t) + 3\Delta t\ \sin(k\Delta t)$$

with

$$x(0 \cdot \Delta t) = 10$$

Representative computer-generated plots of $x(t)$ are given in Figure 8.10 for various choices of Δt. For a sufficiently small time increment Δt, the approximate response is very nearly the actual system response.

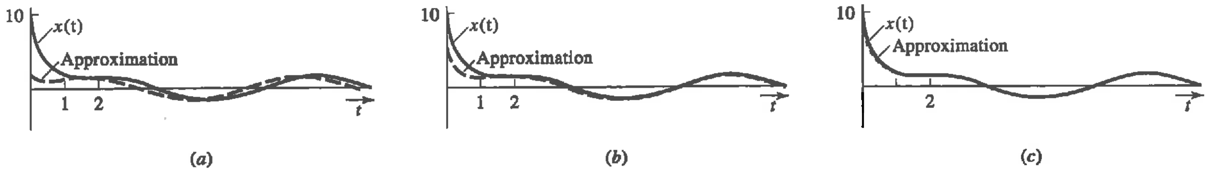


Figure 8.10 Computer-generated response plots for a first-order system. (a) Step size $\Delta t = 0.4$. (b) Step size $\Delta t = 0.2$. (c) Step size $\Delta t = 0.05$.

Another example system is the following:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ -3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 2 \\ -1 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

with

$$\begin{bmatrix} x_1(0^-) \\ x_2(0^-) \end{bmatrix} = \begin{bmatrix} -4 \\ 5 \end{bmatrix}$$

$$u(t) = \cos\ 0.25t$$

It is approximated by

$$\begin{bmatrix} x_1\{(k+1)\,\Delta t\} \\ x_2\{(k+1)\,\Delta t\} \end{bmatrix} = \begin{bmatrix} 1-2\Delta t & \Delta t \\ -3\Delta t & 1 \end{bmatrix} \begin{bmatrix} x_1\,(k\,\Delta t) \\ x_2\,(k\,\Delta t) \end{bmatrix} + \begin{bmatrix} 2 \\ -1 \end{bmatrix} \Delta t\ \cos(0.25k\ \Delta t)$$

$$y\{(k+1)\,\Delta t\} = \begin{bmatrix} 1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} x_1\{(k+1)\,\Delta t\} \\ x_2\{(k+1)\,\Delta t\} \end{bmatrix}$$

or

$$\begin{cases} x_1\{(k+1)\,\Delta t\} = (1-2\Delta t)\,x_1\,(k\,\Delta t) + \Delta t x_2\,(k\,\Delta t) + 2\Delta t\ \cos(0.25k\ \Delta t) \\ x_2\{(k+1)\,\Delta t\} = -3\Delta t x_1\,(k\,\Delta t) + x_2\,(k\,\Delta t) - \Delta t\ \cos\,(0.25k\ \Delta t)) \\ y\{(k+1)\,\Delta t\} = x_1\{(k+1)\,\Delta t\} - \frac{1}{2}x_2\{(k+1)\,\Delta t\} \end{cases}$$

with

$$\begin{cases} x_1(0\cdot\Delta t) = -4 \\ x_2(0\cdot\Delta t) = 5 \end{cases}$$

Computer-generated response plots for this system are given in Figure 8.11, where $\Delta t = 0.05$.

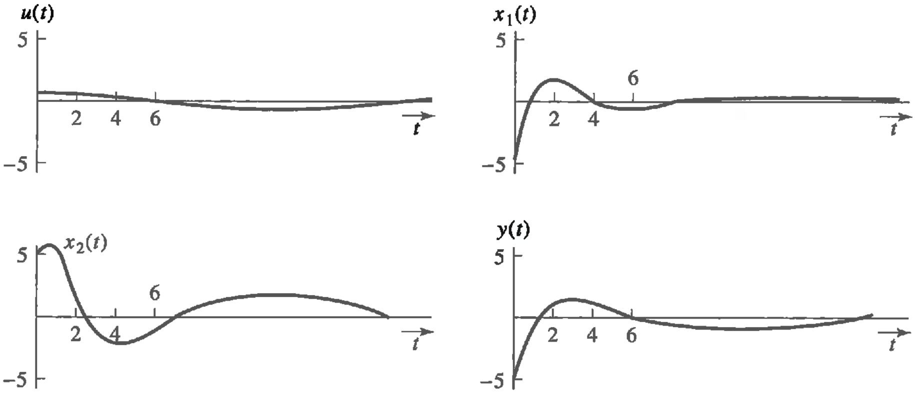


Figure 8.11 Computer-generated response plots for a second-order system.

Improved accuracy and reduced computation time may result from using more involved approximations—for example, matrix power series, predictor correctors, or Runge–Kutta methods.

❑ DRILL PROBLEMS

D8.12 For the following systems, develop discrete-time approximation equations using the indicated time steps Δt.

(a)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 4 \end{bmatrix} u$$

$$y = [-3 \ \ -1] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\Delta t = 0.2$$

Ans.

$$x[(k+1)\ \Delta t] = \begin{bmatrix} 1.6 & 0.2 \\ 0.4 & 0.8 \end{bmatrix} x(k\ \Delta t) + \begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix} u(k\ \Delta t)$$

$$y = [-3 \ \ -1]\, x(k\ \Delta t)$$

(b)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 7 & -2 & -3 \\ 6 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 & -2 \\ -1 & 3 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$y = [5 \ \ -2 \ \ 1] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$\Delta t = 0.01$$

Ans.

$$x[(k+1)\,\Delta t] = \begin{bmatrix} 1.01 & 0.02 & 0.03 \\ 0.07 & 0.98 & -0.03 \\ 0.06 & 0 & 1.04 \end{bmatrix} x(k\,\Delta t) + \begin{bmatrix} 0.01 & -0.02 \\ -0.01 & 0.03 \\ 0 & 0.04 \end{bmatrix} u(k\,\Delta t)$$

$$y = [5 \quad -2 \quad 1]\,x(k\,\Delta t)$$

D8.13 For the set of state equations

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ -3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 4 \end{bmatrix} u$$

$$y = [1 \quad -1] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

a discrete-time approximation is

$$\begin{bmatrix} x_1\{(k+1)\,\Delta t\} \\ x_2\{(k+1)\,\Delta t\} \end{bmatrix} = \begin{bmatrix} (1-2\Delta t) & \Delta t \\ -3\Delta t & 1 \end{bmatrix} \begin{bmatrix} x_1(k\,\Delta t) \\ x_2(k\,\Delta t) \end{bmatrix} + \begin{bmatrix} \Delta t \\ 4\Delta t \end{bmatrix} u(k\,\Delta t)$$

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

If

$$\begin{bmatrix} x_1(0^-) \\ x_2(0^-) \end{bmatrix} = \begin{bmatrix} 10 \\ 0 \end{bmatrix} \quad \text{and} \quad u(t) = 2 \quad t \geqslant 0$$

calculate approximate values for $x(\Delta t)$, $x(2\Delta t)$, and $x(3\Delta t)$ for the following:

(a)

$\Delta t = 0.2$

Ans. $\begin{bmatrix} 6.4 \\ -4.4 \end{bmatrix}, \begin{bmatrix} 3.36 \\ -6.64 \end{bmatrix}, \begin{bmatrix} 1.09 \\ -7.06 \end{bmatrix}$

(b)

$\Delta t = 0.1$

Ans. $\begin{bmatrix} 8.2 \\ -2.2 \end{bmatrix}, \begin{bmatrix} 6.54 \\ -3.86 \end{bmatrix}, \begin{bmatrix} 6.37 \\ -5.02 \end{bmatrix}$

(c)

$$\Delta t = 0.02$$

Ans. $\begin{bmatrix} 9.84 \\ -0.44 \end{bmatrix}, \begin{bmatrix} 9.67 \\ -0.87 \end{bmatrix}, \begin{bmatrix} 9.503 \\ -1.29 \end{bmatrix}$

8.5 Stability

A system is stable if its eigenvalues (or characteristic values) are in the left half-plane (LHP). We want to expand upon this issue and formally define various notions of stability.

8.5.1 Asymptotic Stability

Consider a system represented in state space:

$$\dot{x} = Ax \qquad x(0) = x_0$$

The system is said to be *asymptotically stable* if all the states approach zero with time—that is,

$$x(t) \to 0 \quad \text{as} \quad t \to \infty$$

Now, let us diagonalize the system. This simplifies the task as it will allow us to look at the components of the system one at a time.

$$x(t) = Pz(t) \to \dot{z}(t) = \Lambda z(t)$$

where Λ is the diagonal matrix of eigenvalues of A—that is,

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdot & 0 \\ 0 & \lambda_2 & \cdot & \cdot \\ \cdot & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0 \\ 0 & 0 & \cdot & \lambda_n \end{bmatrix}$$

The individual subsystems are

$$\dot{z}_i(t) = \lambda_i z_i(t)$$

For asymptotic stability eigenvalues must be in the LHP.

The solution to this first-order system is

$$z_i(t) = e^{\lambda_i t} z_i(0)$$

Clearly, if the eigenvalues of the system have negative real parts (i.e., they are in the LHP), the individual states go to zero asymptotically with time: $z_i(t) \to 0$ implies that $z(t) \to 0$. Because $x(t) = Pz(t)$, we conclude that $x(t) \to 0$. Hence, it has been proved that the system is asymptotically stable if its eigenvalues are in the LHP.

8.5.2 BIBO Stability

Another notion of stability that has been used throughout the book, starting from Chapter 2, is input–output stability. This concept is also referred to as *bounded-input, bounded-output* (or *BIBO*) stability. BIBO stability means that the system output is bounded for all bounded inputs. This is,

$$|u(t)| \leqslant N < \infty \rightarrow |y(t)| \leqslant M < \infty \quad \text{for all bounded inputs}$$

where M and N are some finite bounds for u and y. Examples of bounded functions are negative exponentials and sinusoids. For example,

$$\left|e^{-at}\right| \leqslant 1 \quad \text{or} \quad |\sin(\mathrm{t})| \leqslant 1 \rightarrow M = 1 \quad \text{or} \quad N = 1$$

The condition that guarantees BIBO stability is the familiar condition that all the transfer function poles be in the LHP.

What is the difference between the two definitions? Does one imply the other? The answer is obtained from looking at the expression for the transfer function in the terms of state space matrices:

$$T(s) = \frac{N(s)}{D(s)} = C(sI - A)^{-1}B = \frac{C \text{ adjoint}\,[sI - A]\,B}{|sI - A|}$$

Recall that eigenvalues of the system are found from the equation

$$|sI - A| = 0$$

Because this is also the polynomial that appears in the denominator of the transfer function, we can draw the following the conclusion:

In the absence of pole–zero cancellations, transfer function poles are identical to the system eigenvalues, hence BIBO stability and asymptotic stability are equivalent.

The pole–zero cancellation condition is important for the equivalence. This is because, poles are formally to be identified from the simplified transfer function (i.e., after all the numerator and denominator common terms have been canceled out). For instance, consider the following situation,

$$T(s) = \frac{s-1}{(s-1)(s+2)}$$

in which $T(s)$ has one pole at $s = -2$. The term $(s - 1)$ must be canceled out before the poles and zeros are identified. To prove that the transfer function does not have a pole at 1, take the limit of $T(s)$ as s approaches 1.

$$\lim_{s\rightarrow 1} T(s) = \lim_{s\rightarrow 1} \frac{s-1}{(s-1)(s+2)} = \lim_{s\rightarrow 1} \frac{1}{2s+1} = \frac{1}{3} \neq \infty$$

This system does not have a pole at $s = 1$.

where we have used L'Hôpital's rule. It is concluded that the system is BIBO stable because it has no poles in the RHP or on the imaginary axis. To determine asymptotic stability, we need to obtain a state space realization. One realization is given by

$$\dot{x} = \begin{bmatrix} -1 & 2 \\ 1 & 0 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u$$

$$y = [1 \quad -1]x$$

The system eigenvalues are at 1 and -2. Because of the eigenvalue of 1, we conclude that the system is unstable in the asymptotic sense. Such contradictory answers to system stability occur only when the system transfer function has pole–zero cancellations in the RHP or on the imaginary axis. What do we make of the system stability in such cases?

We can make some progress toward the answer by looking at the diagonalized system given by

$$\dot{x} = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} x + \frac{1}{3}\begin{bmatrix} 1 \\ -1 \end{bmatrix} u$$

$$y = [0 \quad -3]x$$

Observing the signal flow graph of the system shown in Figure 8.12, we note that the unstable mode at 1 is not connected to the output. Since the transfer function describes the input–output properties of the system, it is not surprising that it does not detect the offending mode. Imagine an experiment where you apply an input to the system, and connect your measurement instrument to the output. Even though the first state is practically "blowing up," you will not be aware of it. This is what is happening here. Fortunately, such pathological cases do not happen very often in practice. After all, because of modeling uncertainties and component tolerances, it is practically impossible for a system to have its poles and zeros exactly at the same place. We will see in the next section that this issue is related to system controllability and observability properties, and that it can be detected and avoided.

State space description of systems are called *internal representation* because they allow us to observe the internal structure of systems, whereas transfer function description is an external representation. In the case of our example, the state space representation allowed a more accurate description of the system.

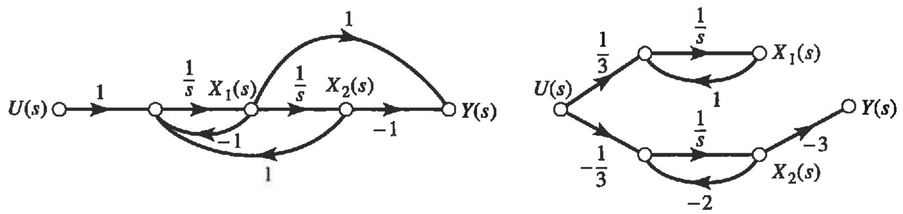


Figure 8.12 (a) Signal flow graph of a system with pole–zero cancellation. (b) Diagonal representation of the system.

Going back to the question of stability, we conclude that the preceding system is practically unstable, because any nonzero initial conditions on the first state will grow indefinitely.

8.5.3 Internal Stability

Asymptotic stability is a notion that is based on the state space description of systems, whereas BIBO stability is based on transfer function description. There is another notion of stability, called *internal stability*, that is based on transfer function description and is stronger than BIBO stability. To define it, we need to consider the general block diagram in Figure 8.13. This figure is a more realistic diagram of feedback control systems. The blocks represent the plant, the controller, and the sensor dynamics. Besides the usual reference (or command) input, the ever-present noise and disturbance inputs are explicitly included.

Internal stability requires that all signals within the feedback system remain bounded for all bounded inputs. This is equivalent to the requirement that all possible transfer functions between all inputs and outputs be stable. It can be shown that only nine transfer functions between the three inputs (R, D, N) and the three outputs taken at the output of the summing junctions (U, V, W) are sufficient. In practice, determining the nine transfer functions and checking them for stability is still a major task. The following result, a necessary and sufficient condition for internal stability, will be used as a test for internal stability. The feedback system is internally stable if and only if the transfer function $1 + KGH$ has no zeros in the RHP (including the imaginary axis), and the product of KGH has no pole–zero cancellations in the RHP (including the imaginary axis).

As an example, consider the case of

$$G(s) = \frac{1}{s-1} \qquad H(s) = \frac{s-1}{s+1} \qquad \text{and} \qquad K = 1$$

Because there is an RHP pole–zero cancellation in KGH, the system is not internally stable. To show that at least one signal will blow up, note the following transfer functions:

$$\frac{Y(s)}{R(s)} = \frac{1}{s+2} \quad \frac{Y(s)}{D(s)} = \frac{s+1}{(s-1)(s+2)}$$

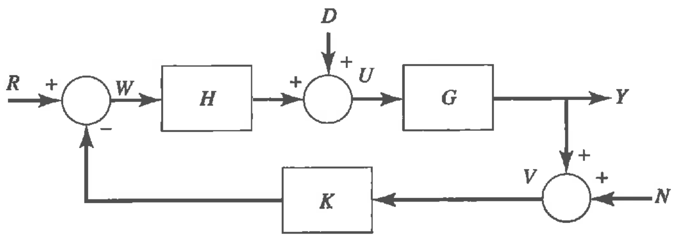


Figure 8.13 Block diagram of a feedback system showing disturbance and noise inputs.

Note that in traditional BIBO stability, it is the transfer function between R and Y that is examined. This transfer function is clearly stable. However, the transfer function between the disturbance D and output Y is unstable. Hence, the slightest disturbances in the system will grow unbounded. For all practical purposes, the system is unstable.

The summary of our discussion of stability is that internal stability is the true stability requirement that must be imposed on feedback control systems. A design lesson that can be drawn from our discussion is the following. We must never cancel the unstable (RHP) plant poles by unstable (RHP) compensator zeros, for this will render the closed-loop system internally unstable. This caution is warranted because canceling poles in undesirable locations by compensator zeros and replacing them by poles in more desirable locations is commonly used by control system designers. This is an acceptable and effective technique, but only for poles that are in the LHP.

❑ DRILL PROBLEMS

D8.14 Determine stability of the following systems. Check for BIBO, asymptotic, and internal stability.

(a)
$$\dot{x} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$
$$y = [1 \quad 1]x$$

(b)
$$\dot{x} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} u$$
$$y = [1 \quad 1 \quad 1]x$$

(c)
$$\dot{x} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x + \begin{bmatrix} -1 \\ 1 \end{bmatrix} u$$
$$y = [1 \quad 1]x + u$$

Ans. (a) $T(s) = 1/(s+1)$, eigenvalues $= 0, -1$; BIBO stable but not asymptotically stable; (b) $T(s) = (3s^2 + 5s + 1)/[s(s+1)]^2$; eigenvalues $= 0, \ -1, \ -1$; neither BIBO nor asymptotically stable; (c) $T(s) = 1$; eigenvalues $= 1, \ 1$; BIBO stable but not asymptotically stable

D8.15 Consider the feedback control system of Figure 8.15.

Let $G(s) = 1/(s-1)$. Determine if the system is internally stable in each case.

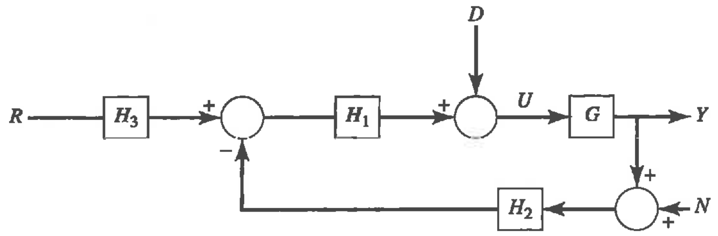


Figure D8.15

(a) $H_1 = \dfrac{s-1}{s+1},\ H_3 = 1,\ H_2 = 1$

(b) $H_1 = 1,\ H_2 = \dfrac{s-1}{s+1},\ H_3 = 1$

(c) $H_1 = 1,\ H_2 = 1,\ H_3 = \dfrac{s-1}{s+1}$

Ans. (a) no; (b) no; (c) no

8.6 Controllability and Observability

In the preceding chapters, a control system was operated upon to provide acceptable phase margin, rise time, or other figures of merit. Perhaps the system is constructed in a way that conspires to thwart efforts at improving performance. By using state-variable methods it is possible to answer fundamental questions about the ability of the control-system designer to effect meaningful improvement in performance and to generate needed sensor measurements. The terms *controllability* and *observability*, respectively, address those needs.

A system is completely controllable if the system state $x(t_f)$ at time t_f can be forced to take on any desired value by applying a control input $u(t)$ over a period of time from t_0 until t_f.

The definition does not restrict the choice of $u(t)$. The idea is that it is possible to move the system state to any desired destination. Perhaps the system is (or is not) constructed in a way that allows control to take place. A test for controllability can easily be constructed.

A system is completely observable if any initial state vector $x(t_0)$ can be reconstructed by examining the system output $y(t)$ over some period of time from t_0 until t_f.

There are no restrictions placed on the output. The definition indicates that any earlier value of the state vector is determinable by watching the output $y(t)$. An automobile would be considered completely observable if, by monitoring speedometer (for speed), odometer (for distance), and steering wheel position (for turning), it is possible to determine where the car was parked before being driven.

For systems of certain kinds (with diagonal A matrices), the tests for controllability and observability are easy to apply. For a nondiagonal system, a test can also be constructed. For a system that is completely controllable, methods will be developed by which an appropriate control can be derived. Similarly, for a system that is completely observable, an observer will be designed to carry out that task of state reconstruction.

Figure 8.14 shows that controllability is tested assuming a zero-state response and that observability is tested assuming a zero-input response. The tests provide a worst-case scenario, where the initial condition does not necessarily aid in control and an input does not necessarily aid in reconstruction of an earlier state. A system that passes the controllability test is usually applied in an environment that has a nonzero initial condition. Similarly, a system that passes the observability test (observers will be considered shortly) is usually applied in an environment that includes an input and control.

Consider the following system:

$$\dot{x}_1 = x_1$$
$$\dot{x}_2 = 2x_2 + u$$

The objective is to force the system states to go to zero. This is another way of stating that we want to make the system asymptotically stable, a common objective. According to definition of controllability, this is an achievable objective if the system is controllable. The solution of the system is

$$x_1(t) = e^t x_1(0)$$
$$x_2(t) = e^{2t} x_2(0) + \int_0^t e^{2(t-\tau)} u(\tau) d\tau$$

Observe that by appropriate choice of the control signal u, the second state can be driven to zero. The first state, however, is uncontrollable. It will always blow up, unless its initial condition is zero. Upon examining the system signal flow graph, shown in Figure 8.15, it is clear that the control signal is not even connected to the first state, so it cannot affect it in any way. Because the system is in decoupled form, its eigenvalues can be obtained by inspection; they are 1 and 2. These are also called system *modes*. When the system is in decoupled form, we can be more specific: The first mode is uncontrollable, whereas the second mode is controllable. Because the objective was to drive all states to zero, the system is declared to be uncontrollable.

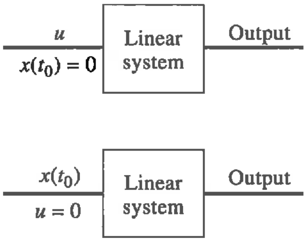


Figure 8.14 Significance of controllability and observability tests. (a) Test for controllability. (b) Test for observability.

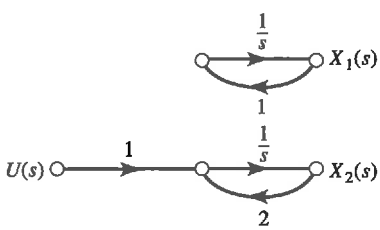


Figure 8.15 Signal flow graph of an uncontrollable system.

How to determine the control signal u to drive the states to arbitrary values. will be discussed later. Driving the states to zero (i.e., stabilization), however, is easy. For instance, in the preceding example, if we choose u as

$$u = -kx_2$$

we get

$$\dot{x}_2 = 2x_2 - kx_2 = (2 - k)x_2$$

Clearly if the gain k is greater than 2, this state will indeed go to zero. The idea behind this choice is to feed the state back to the input using an appropriate gain. This is called *state feedback*, and more will be said about it in the next chapter. That this scheme works is not surprising—after all, feedback has been used for stabilization throughout the book.

Determining controllability and observability is easy when the system is in diagonalized form. To see this, consider the following general example for SISO systems.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} u$$

$$y = [1 \quad 0 \quad 1 \quad 0]x$$

Observing the input connections in Figure 8.16, we conclude that modes 3 and 4 are uncontrollable because they are not connected to the control input. Also, modes 2 and 4 are unobservable because they are not connected to the output. In general we can always categorize the system modes into four categories: controllable and observable

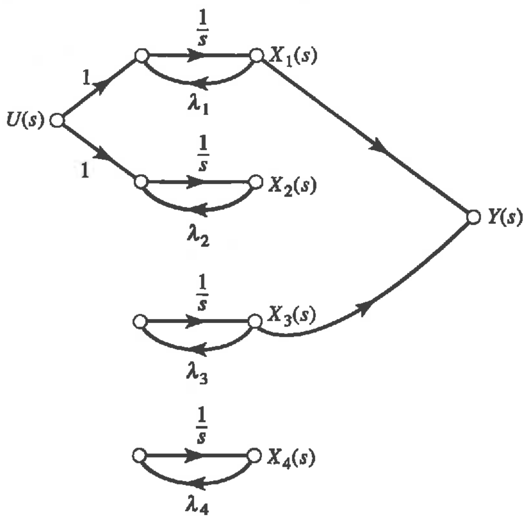


Figure 8.16 Signal flow graph of a system showing controllable, observable, uncontrollable, and unobservable modes.

(as in mode 1), controllable but unobservable (as in mode 2), uncontrollable but observable (as in mode 3), and uncontrollable and unobservable (as in mode 4).

This information is also available from examining the rows and columns of B and C matrices. Uncontrollable modes correspond to zero rows of B: unobservable modes correspond to zero columns of C. The latter applies to multiple-input, multiple-output systems with distinct (i.e., nonrepeated) modes in diagonalized form.

For systems in general (nondiagonalized) form, these properties cannot be determined from the signal flow graph (or block diagram). Similarly, zero rows in B or zero columns in C do not imply anything in general.

8.6.1 The Controllability Matrix

Fortunately, there is a much simpler method of determining system controllability than diagonalization. It can be shown that an nth-order system, with or without repeated modes (eigenvalues),

$$\dot{x} = Ax + Bu$$

is completely controllable if and only if its controllability matrix

$$M_c = \left[B \;\vdots\; AB \;\vdots\; \cdots \;\vdots\; A^{n-1} \;\; B \right]$$

is of full rank. The controllability matrix consists of the columns of B followed by the columns of AB, and so on.

For the system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -2 & 1 & 2 \\ 4 & 0 & 3 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 & 4 \\ -5 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

A is 3×3, so

$$M_c = \left[B \;\vdots\; AB \;\vdots\; A^2 B \right]$$

Upon using

$$AB = \begin{bmatrix} -2 & 1 & 2 \\ 4 & 0 & 3 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 4 \\ -5 & 0 \\ 5 & 0 \end{bmatrix} = \begin{bmatrix} -5 & -8 \\ 0 & 16 \\ 5 & 4 \end{bmatrix}$$

$$A^2 B = A(AB) = \begin{bmatrix} -2 & 1 & 2 \\ 4 & 0 & 3 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} -5 & -8 \\ 0 & 16 \\ 5 & 4 \end{bmatrix} = \begin{bmatrix} 20 & 40 \\ -5 & -20 \\ -5 & -24 \end{bmatrix}$$

we can write

$$M_c = \begin{bmatrix} 0 & 4 & -5 & -8 & 20 & 40 \\ -5 & 0 & 0 & 16 & -5 & -20 \\ 0 & 0 & 5 & 4 & -5 & -24 \end{bmatrix}$$

To be of full rank, the controllability matrix must have three linearly independent columns, which it does, since

$$\begin{vmatrix} 0 & 4 & -5 \\ -5 & 0 & 0 \\ 0 & 0 & 5 \end{vmatrix} \neq 0$$

The system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 6 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ -2 \end{bmatrix} u$$

has controllability matrix

$$M_c = [B \;\vdots\; AB]$$

where

$$AB = \begin{bmatrix} 2 & 3 \\ 6 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} -4 \\ 8 \end{bmatrix}$$

Thus we have

$$M_c = \begin{bmatrix} 1 & -4 \\ -2 & 8 \end{bmatrix}$$

which is not of rank 2, since

$$\begin{vmatrix} 1 & -4 \\ -2 & 8 \end{vmatrix} = 0$$

This system is not completely controllable.

Note that for SISO systems the controllability matrix is square. From matrix algebra, we recall that a square matrix has full rank if and only if its determinant is not zero. Therefore, in the SISO case, the system is controllable if and only if the determinant of M_c is nonzero. The controllability matrix, although easy to apply, provides only a "yes or no" answer on system controllability. To get specific information on individual modes, you have to diagonalize the system.

8.6.2 The Observability Matrix

To determine whether a nondiagonalized nth-order system is completely observable, its observability matrix

$$M_o = \begin{bmatrix} C \\ \text{-----} \\ CA \\ \text{-----} \\ \vdots \\ \text{-----} \\ CA^{n-1} \end{bmatrix}$$

may be formed. The system is completely observable if and only if the observability matrix is of full rank, that is, if M_o has n linearly independent rows.

For example, the system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ -3 & 0 & 1 \\ 4 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} u$$

$$y = [0 \quad 0 \quad 1] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

is completely observable:

$$CA = [0 \quad 0 \quad 1] \begin{bmatrix} 2 & 1 & 0 \\ -3 & 0 & 1 \\ 4 & 0 & 0 \end{bmatrix} = [4 \quad 0 \quad 0]$$

$$CA^2 = (CA)A = [4 \quad 0 \quad 0] \begin{bmatrix} 2 & 1 & 0 \\ -3 & 0 & 1 \\ 4 & 0 & 0 \end{bmatrix}$$

$$= [8 \quad 4 \quad 0]$$

$$M_o = \begin{bmatrix} 0 & 0 & 1 \\ 4 & 0 & 0 \\ 8 & 4 & 0 \end{bmatrix}$$

As another example, consider the system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

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

The observability matrix has two linearly independent rows (1 and 3).

8.6.3 Controllability, Observability, and Pole–Zero Cancellation

It can be shown, in general, that uncontrollable or unobservable (SISO) systems will have pole–zero cancellations in their transfer functions. We will not prove this fact, but it will be investigated by an example. Consider the second-order system

$$\dot{x} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} x + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} u$$

$$y = \begin{bmatrix} c_1 & c_2 \end{bmatrix} x$$

The transfer function of the system is given by

Lack of controllability or observability leads to pole–zero cancellation in the transfer function.

$$T(s) = \frac{(b_1c_1 + b_2c_2)s - (2b_1c_1 + b_2c_2)}{(s-1)(s-2)}$$

Now, if $b_1 = 0$ (or $c_1 = 0$), the mode at 1 becomes uncontrollable (or unobservable) and the pole term $(s-1)$ gets canceled in the transfer function. Similarly, when $b_2 = 0$ (or $c_2 = 0$), the mode at 2 becomes uncontrollable (or unobservable) and the pole term $(s-2)$ will get canceled out.

The example demonstrates that lack of either controllability or observability will lead to pole–zero cancellation in the transfer function. Conversely, pole–zero cancellation in a transfer function implies either uncontrollability or unobservability. As another example, consider the system

$$\dot{x} = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} x + \begin{bmatrix} 1 \\ b_2 \end{bmatrix} u$$

$$y = \begin{bmatrix} c_1 & 1 \end{bmatrix} x$$

$$T(s) = \frac{(b_2 + c_1)s + (-2c_1 - b_2 + b_2c_1)}{(s-1)(s-2)}$$

When $b_2 = 0$, the term $(s-2)$ is canceled. Hence, the corresponding mode is either uncontrollable or unobservable. Let us see if the system is observable.

$$M_o = \begin{bmatrix} c_1 & 1 \\ c_1 & 2 + c_1 \end{bmatrix}$$

Because the determinant of the observability matrix is nonzero, the system is observable. Consequently, the term$(s-2)$ corresponds to an uncontrollable mode. Also, observe that the transfer function becomes 0 when $b_2 = c_1 = 0$. This strange case occurs because there is no path from the input to the output when both parameters are zero.

Never cancel RHP poles or zeros.

Our earlier caution against unstable pole–zero cancellation is worth repeating here. When canceled by a zero, an unstable pole does not really disappear, it simply becomes either uncontrollable or unobservable. In the first case, you will observe the state blowing up, but you cannot do anything about it. In the second case, you will not even be aware that something is wrong because the unstable state does not appear at the output. In either case, the results are disastrous.

8.6.4 Causes of Uncontrollability

What are some of the causes of uncontrollability or unobservability? One cause, as has been indicated, is pole–zero cancellation. Another source of problem is symmetry in the system. For instance, consider the second-order system

$$\dot{x} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u$$

$$y = [1 \quad 1]x$$

The system is neither controllable, nor observable. In fact, the diagonal realization and transfer function of the system indicate that the pole at the origin is canceled out.

$$\dot{x}_1 = 0$$

$$\dot{x}_2 = 2x_2 + u \qquad T(s) = \frac{2s}{s(s-2)} = \frac{2}{s-2}$$

$$y = 2x_2$$

In physical systems, such symmetry is rare. The preceding two cases either can be avoided or are unlikely to occur in practice. Another common cause is redundant state variables. During the process of modeling complex systems, one may introduce unnecessary or redundant state variables. In this case, lack of controllability/observability indicates modeling errors and can be corrected by proper system modeling. The following example demonstrates the case of modeling error.

Consider the simple RL circuit shown in Figure 8.17. The circuit input is a voltage source, and the output is the current flowing through the inductor. From basic circuit theory, we know that the correct equation is given by

$$u(t) = i(t) + \frac{di(t)}{dt}$$

If we let $x = i(t)$, then we have

$$\dot{x} = -x + u$$

If the output is the current, we have

$$y = x = i$$

so

$$T(s) = \frac{Y(s)}{U(s)} = \frac{1}{s+1}$$

Now, it is also known that current is the rate of flow of charge. Suppose electrical charge is selected as a state variable—that is,

An incorrect model can result in an uncontrollable/unobservable realization.

$$x_1 = q$$
$$x_2 = \dot{q} = i$$
$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = -x_2 + u$$
$$y = x_2$$

so that

$$A = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad C = [0 \quad 1]$$

The transfer function between the current and the input is given by

$$\frac{Y(s)}{U(s)} = C(sI - A)^{-1}B = \frac{s}{s(s+1)} = \frac{1}{s+1}$$

The observability matrix of this second-order model is

$$M_o = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} \rightarrow |M_o| = 0$$

which is singular (i.e., the model is not observable). This does not mean that the circuit is unobservable. It simply indicates that the model is not good. In fact, the extra state variable defined for charge is redundant for our purposes.

Another common cause is inappropriate or insufficient control actuators or sensors. The latter cause is an important system design issue. For a given control objective, we need an appropriate model and a sufficient number of control actuators and sensors that are appropriately positioned. To illustrate this issue, consider the classic problem of stabilizing an inverted pendulum on a moving cart shown in Figure 8.18.

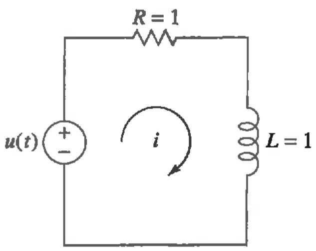


Figure 8.17 *RL* circuit: $R = 1, L = 1$.

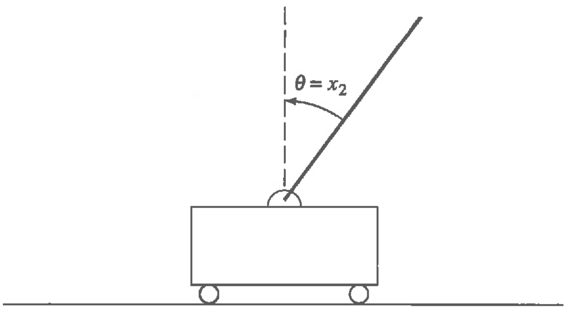


Figure 8.18 Inverted pendulum on a moving cart.

Inverted pendulum on a cart example.

Suppose the objective is to balance the pendulum and stop the cart. A linearized model is given by

$$\dot{x} = \begin{bmatrix} 0 & -a & 0 \\ 0 & 0 & 1 \\ 0 & a & 0 \end{bmatrix} x + \begin{bmatrix} b_1 \\ 0 \\ -b_2 \end{bmatrix} u$$

The state vector components correspond to cart velocity, pendulum angle, and pendulum angular velocity, respectively. To meet the objective, one of the state variables is measured and fed back. That signal is used to drive a motor that will move the cart to stabilize the system. Note that asymptotic stability implies that all states will approach zero, which means that the pendulum will be balanced and the cart will stop moving.

We can show that this system is controllable. Now the question is where to place the sensor—that is, which state should be measured. If we measure the pendulum angle and use that as the feedback signal, we get

$$y = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} x$$

The system is not observable in this case because

$$|M_n| = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & a & 0 \end{vmatrix} = 0$$

Intuitively, we can imagine that it is possible to balance the pendulum while the cart is still moving. Hence, it is obvious that using the pendulum angle as the feedback signal would not be a good choice. Similar arguments can be made against measuring the pendulum angular velocity. Finally, by using the cart velocity as the measured signal—that is,

$$y = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} x$$

we can verify that the system is observable because

$$|M_o| = \begin{vmatrix} 1 & 0 & 0 \\ 0 & -a & 0 \\ 0 & 0 & -a \end{vmatrix} = a^2 \neq 0$$

Therefore, it is theoretically possible to meet the design objectives by measuring the cart velocity.

The preceding example demonstrated one common cause of uncontrollability/unobservability. As system complexity increases, we may no longer have the benefit of our intuition, and we must resort to system concepts.

In general, if system design issues are well thought out, and adequate models are obtained, we need not worry about controllability/observability issues in practice. These issues will appear frequently as theoretical conditions for state space and optimal design, however.

❑ Computer-Aided Learning

System stability can be obtained by looking at the poles or eigenvalues of systems. MATLAB has two commands to find poles and zeros. They are "pole" and "tzero." The "pole" command returns the poles and "tzero" returns the so-called transmission zeros of systems. If the A matrix is available, we can determine asymptotic stability by finding the eigenvalues of A using the "eig" command. For the system defined by

$$\dot{x} = \begin{pmatrix} -2 & 1 & 1 \\ -3 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} x + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} u$$

$$y = \begin{pmatrix} 2 & -2 & 1 \end{pmatrix} x$$

we get

```
>>pole(g)
ans=
-1.0000+1.4142i
-1.0000-1.4112i
 0
>>eig(a)  % same as the poles
ans=
-1.0000+1.4142i
-1.0000-1.4142i
 0
>>tzero(g)
ans=
-6.6458
-1.3542
```

The system controllability and observability can be determined by examining the corresponding matrices. MATLAB has ctrb and obsv to obtain them. For the above-defined system, we get

```
>>CO=ctrb(a,b)
CO=
  1   0  -3
  1  -3   0
  1   0   0
>>OB=obsv(a,c)
OB=
  2  -2  1
  2   2  2
-10   2  2
```

We can then use det and rank to find the determinant and rank of these matrices. Sometimes systems have pole–zero cancellations (owing lack of controllability and/or observability). To obtain the so-called minimal realization of the system, MATLAB has the "mineral" command. As an example, consider the third-order system where all matrices are all 1s (similar to the example in the book).

```
>>a=ones(3,3);b=ones(3,1);c=ones(1,3);d=0;
>>g=ss(a,b,c,d);
>>rank(ctrb(a,b))  % system not controllable
ans=
  1
>>rank(obsv(a,c))  % system not observable
ans=
  1
>>tf(g)  % transfer function has 2 poles and zeros at the origin

Transfer function:
          3s^2
-------------------------
s^3-3s^2-3.077e-015s+1.972e-031

>>eig(a)  % system not asymptotically stable
              (double eigenvalues at the origin)
ans=
  0.0000
  0.0000
  3.0000
>>gmin=mineral(g);  % after pole–zero cancellation
2 state(s) removed
>>tf(gmin)

Transfer function:
3
------
s - 3
```

C8.2

(a) Determine the stability of the systems defined in Drill Problem 8.14.

(b) Determine the stability, controllability, and observability of

```
a=[-1-2 0;1 2 0;-2-1-3]; b=[1;0;0]; c=[1 0 1]; d=0;
```

(c) Use MATLAB to do drill Problem 8.17.

Solution of state equation can be obtained using the "lsim" and "initial" commands.

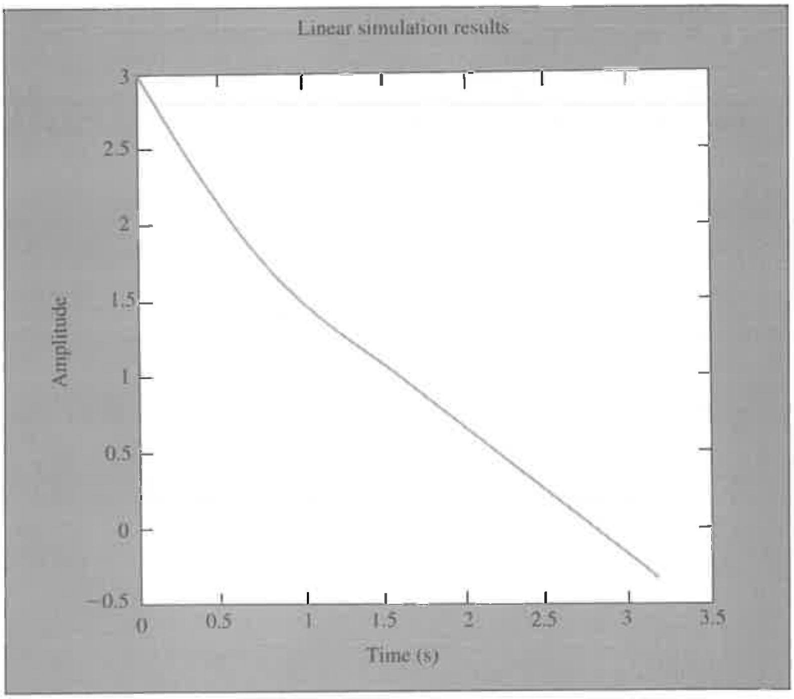
```
LSIM(SYS, U, T, X0)
```

plots the response of a state space system with input $U(T)$ and initial state X_0. The time sequence T must be defined as a vector, and the input must be of the same size as T. Here is an example:

```
g=tf(1, [1 1.4 1])
```

Transfer function:

```
         1
  ---------------
  s^2 + 1.4 s + 1
  gs=ss(g);
  t=linspace(0,pi,100); u=sin(2*t); x0=[-2 3];
  lsim(gs,u,t,x0)
```



When invoked with left-hand arguments,

```
[Y,T,X]=LSIM(SYS,U,T,X0)
```

returns the output Y, the state vector X, and time vector T used for simulation, and no plot is drawn on the screen.

```
  [y,t,x]=lsim(gs,u,t,x0);
  subplot(211), plot(t,x(:,1)), subplot(212), plot(t,x(:,2))
```

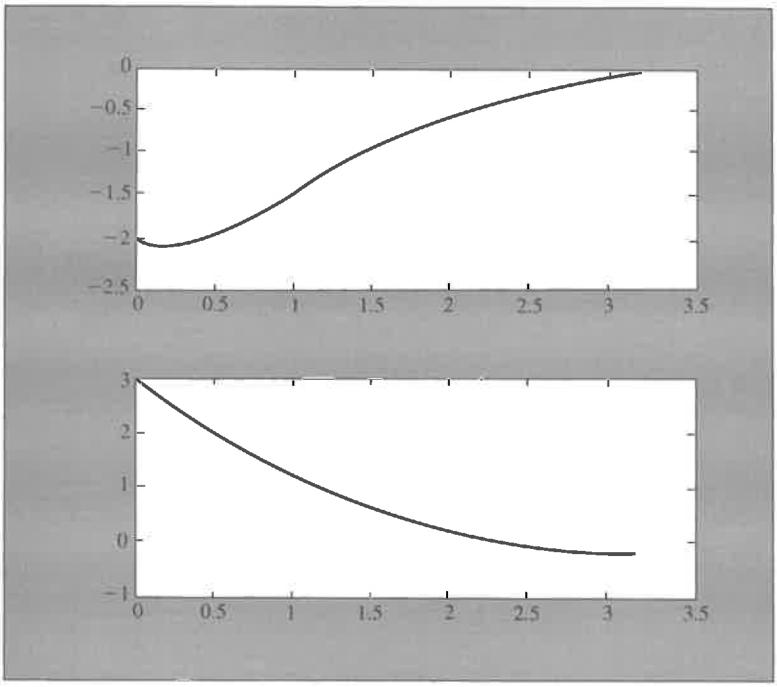
The "initial" command is used to solve for the zero input response (ZIR) of state space systems: $\dot{x} = Ax$ subject to initial conditions:

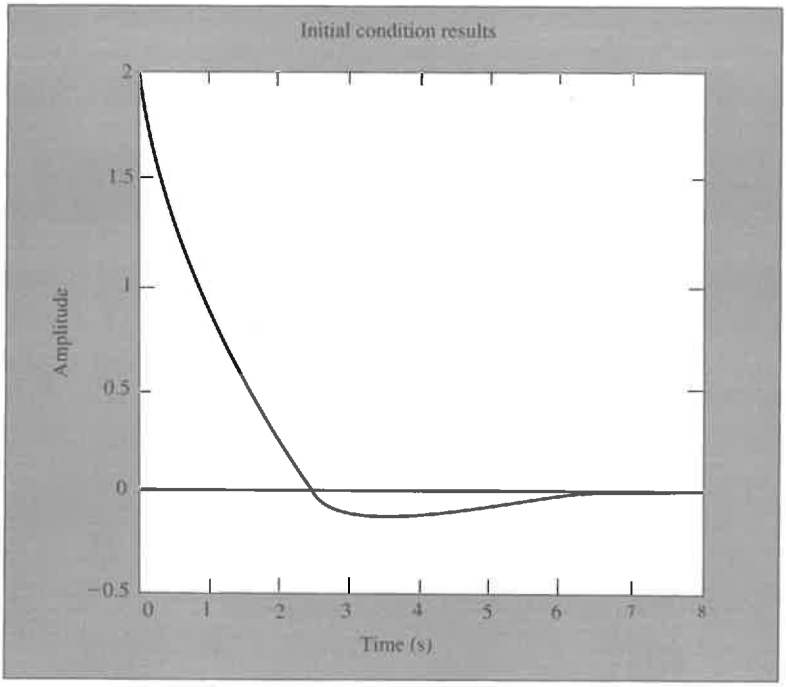
INITIAL(SYS,XO,TF) simulates and plots the time response from $t = 0$ to the final time $t = TF$

INITIAL(SYS,X0,T) specifies a time vector T to be used for simulation

[Y,T,X]=INITIAL(SYS,X0,$\cdots$) returns the left hand side vectors but does not plot.

```
initial(gs, [-1 2])
```





C8.3 Repeat Drill Problem D8.10 using the "lsim" command.

8.7 Inverted Pendulum Problems

One of the most celebrated and well-publicized problems in control is the inverted pendulum (broom balancer) problem. This is an unstable system that may model a rocket before launch. Almost all known and novel control techniques have been tested on the inverted pendulum (IP) problem. In this section we discuss models of a variety of IP-type problems. The IP problem is also highly nonlinear, but it can easily be controlled by using linear controllers in an almost vertical position.

Some of the varieties of IP are single pendulum, single-rotary pendulum, double side-by-side pendulum, double-pendulum, double-rotary, and triple-pendulum problems.

We start with a derivation for a single IP problem (see Figure 8.19).

The horizontal and vertical coordinates of center of gravity of the mass are given by

$$y_1 = x + l \sin\theta$$
$$y_2 = l\cos\theta$$

Newton's law in the horizontal direction gives us

$$u = M\ddot{x} + m\ddot{y}_1 = M\ddot{x} + m\ddot{x} + ml(-\sin\theta\cdot\dot{\theta}^2 + \cos\theta\cdot\ddot{\theta})$$

or

$$\boxed{(M+m)\ddot{x} - ml\sin\theta\cdot(\dot{\theta}^2) + \cos\theta\cdot\ddot{\theta} = u}$$

Newton's law for the rotational motion about the pivot gives

$$m\ddot{y}_1 l\cos\theta - m\ddot{y}_2 l\sin\theta = mgl\sin\theta$$

or

$$(\cos\theta)\cdot ml\ddot{x} + ml^2(-\sin\theta\cdot\dot{\theta}^2 + \cos\theta\cdot\ddot{\theta})\cos\theta$$
$$+ ml^2(\ddot{\theta}\sin\theta + \dot{\theta}^2\cos\theta)\sin\theta = mgl\sin\theta$$

which simplifies to

$$\boxed{m\ddot{x}\cdot\cos\theta + ml\ddot{\theta} = mg\sin\theta}$$

These are the nonlinear IP models. If we assume that θ is small (i.e., we want to control the IP near its vertical equilibrium position), we can linearize these equations. Recall that for small θ, $\sin\theta \approx \theta$ and $\cos\theta \approx 1$, we get

$$\ddot{x} + l\ddot{\theta} = g\theta$$
$$(M+m)\ddot{x} - ml\dot{\theta}^2 + \ddot{\theta} = u$$

Assuming that for small θ, $\dot{\theta}^2$ is negligible, we get our final IP equation.

Linear IP equations.

$$\boxed{\begin{aligned}(M+m)\ddot{x} + \ddot{\theta} &= u\\ \ddot{x} + l\ddot{\theta} &= g\theta\end{aligned}}$$

Obtaining transfer functions, we get

$$\begin{bmatrix} (M+m)s^2 & s^2 \\ s^2 & ls^2-g \end{bmatrix}\begin{bmatrix} X(s) \\ \theta(s) \end{bmatrix} = \begin{bmatrix} u(s) \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} X(s) \\ \theta(s) \end{bmatrix} = \begin{bmatrix} ls^2-g & -s^2 \\ -s^2 & (M+m)s^2 \end{bmatrix}\begin{bmatrix} u(s) \\ 0 \end{bmatrix}\frac{1}{\Delta(s)}$$

$$\Delta(s) = (ls^2-g)(M+m)s^2 - s^4$$
$$= s^2[(M+m)(\ ls^2-g) - 1]$$

$$\frac{X(s)}{U(s)} = \frac{ls^2-g}{\Delta(s)}, \qquad \frac{\theta(s)}{U(s)} = \frac{-s^2}{\Delta(s)}$$

The poles are

$$[(M+m)\ \ g+1] = (M+m)\ \ ls^2 \Rightarrow \mathrm{s} = \pm\sqrt{\frac{(M+m)\ g+1}{l(M+m)}}, 0, 0$$

If $M \gg m$, the poles are at $s = \pm\sqrt{g/l}, 0, 0$ (Figure 8.20). Also if l is small (short pendulum), it is more unstable (RHP pole farther into the plane), Simple experiment confirms that longer pendulums are easier to control, too.

Also notice that the transfer function $\theta(s)/U(s)$ has an unstable pole–zero cancellation (double pole at the origin, s^2 term, cancels out). This indicates that the system cannot be controlled by measuring θ alone. The transfer function $X(s)/U(s)$ does not have this problem, but it has RHP zeros (this makes it more difficult to control the system). Hence, stabilization is possible by measuring cart position.

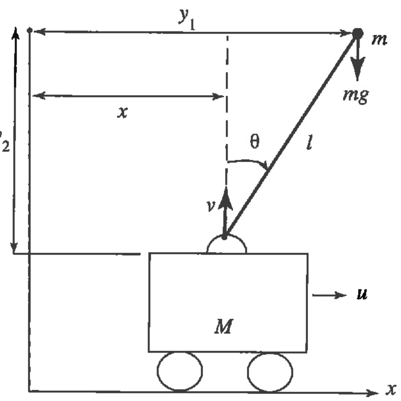


Figure 8.19 The inverted pendulum.

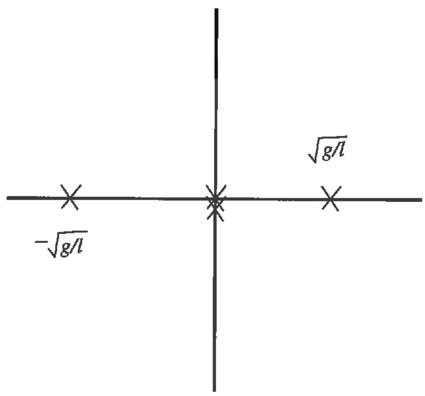


Figure 8.20 Poles of the inverted pendulums

Let us now obtain the state space representation of the IP, rewriting the differential equations:

$$(M+m)\ddot{x} + \left(\frac{g\theta - \ddot{x}}{l}\right) = u \Rightarrow l(M+m)\ddot{x} + g\theta - \ddot{x} = lu$$

$$\left(\frac{u-\ddot{\theta}}{M+m}\right) + l\ddot{\theta} = g\theta \Rightarrow u - \ddot{\theta} + l(M+m)\ddot{\theta} = g(M+m)\theta$$

Simplifying, we have

$$\ddot{x}\,[lc(M+m)-1] + g\theta = lu$$

$$\ddot{\theta}\,[l(M+m)-1] - g(M+m)\theta = -u$$

and

$$\boxed{\begin{aligned}\ddot{x} + \alpha\theta &= \beta lu \\ \ddot{\theta} - \gamma\theta &= -\beta u\end{aligned}}$$

where

$$\alpha = \frac{g}{l(M+m)-1}, \qquad \beta = \frac{1}{l(M+m)-1}, \qquad \gamma = (M+m)\alpha$$

Defining $x_1 = x, \quad x_2 = \alpha, \quad x_3 = \dot{x}, \quad x_4 = \dot{\alpha}$, we write

$$\dot{x}_1 = x_3$$

$$\dot{x}_2 = x_4$$

$$\dot{x}_3 = -\alpha x_2 + \beta lu \Rightarrow \begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\\ \dot{x}_4\end{bmatrix} = \begin{bmatrix}0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ 0 & -\alpha & 0 & 0\\ 0 & \gamma & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\\ x_4\end{bmatrix} + \begin{bmatrix}0\\ 0\\ \beta l\\ -\beta\end{bmatrix}u$$

$$\dot{x}_4 = \gamma x_2 - \beta u$$

The output equation depends on what we measure. If we measure the cart position and the pendulum angle (single input, multi output) problem, we get

$$\begin{bmatrix}y_1\\ y_2\end{bmatrix} = \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\\ x_4\end{bmatrix}$$

Measuring only the pendulum angle (single input, single output), we have

$$y = [\,0 \quad 1 \quad 0 \quad 0]\begin{bmatrix}x_1\\ x_2\\ x_3\\ x_4\end{bmatrix}$$

and sensing only the cart position gives

$$y = [\,1 \quad 0 \quad 0 \quad 0]\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

We can verify system controllability by noting that $\det \mathcal{C} \neq 0$

$$\mathcal{C} = \begin{bmatrix} b & A \cdot b & A^2 \cdot b & A^3 \cdot b \end{bmatrix} = \begin{bmatrix} 0 & \beta l & 0 & -\alpha\beta l \\ 0 & -\beta & 0 & \gamma\beta l \\ \beta l & 0 & -\alpha\beta l & 0 \\ -\beta & 0 & \gamma\beta l & 0 \end{bmatrix}, \det \mathcal{C} \neq 0$$

Checking observability: is a two-step procedure.

1. Sensing pendulum angle: $C = [0 \quad 1 \quad 0 \quad 0]$

$$\mathcal{O} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -\alpha & 0 & 0 \\ 0 & 0 & 0 & -\alpha \end{bmatrix}, \det \mathcal{O} = 0$$

   This is reflected in pole–zero cancellation in the $\theta(s)/u(s)$ transfer function. We cannot achieve stabilization (all states going to zero asymptotically) by measuring only the pendulum angle.

2. Sensing cart position: $C = [1 \quad 0 \quad 0 \quad 0]$

$$\mathcal{O} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -\alpha & 0 & 0 \\ 0 & 0 & 0 & -\alpha \end{bmatrix}, \det \mathcal{O} \neq 0$$

In this case, the system is both controllable and observable [no pole–zero cancellation in the $X(s)/U(s)$ transfer function], and we can stabilize the system.

Note that it makes a big difference which state variable we measure. This is called the "sensor location" problem. In practice, sensitive potentiometers are used to measure both θ and x.

We now consider a slight variation of the IP known as the rotary inverted pendulum (RIP) problem (Figure 8.21). The pendulum, standing on a short arm, can rotate in one plane about a hinge, with a possible potentiometer to measure its angle ϕ. The arm itself can also rotate through an angle of θ.

This version of IP is easier to build very compactly. Regular IP is usually built on a moving car platform or on a track. The track version is large and bulky because the track must be long enough to allow the pendulum to move a distance sufficient to stabilize. The RIP replaces the linear track with a rotating arm, and the arm can rotate

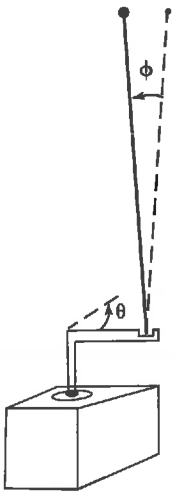


Figure 8.21 The rotary inverted pendulum.

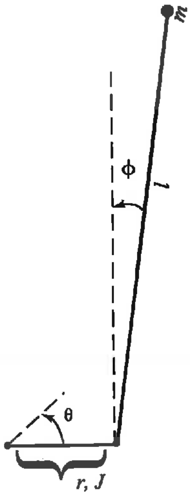


Figure 8.22 Parameter definition for RIP.

as many degrees as desired to stabilize the system. In state space form the linearized equations for the RIP (see Figure 8.22) are given by

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -\alpha & 0 & 0 \\ 0 & \gamma & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ (r/l)\beta \\ -\beta \end{bmatrix} u$$

where the states are $x_1 = \theta$, $x_2 = \phi$, $x_3 = \dot{\theta}$, $x_4 = \dot{\phi}$, and $\beta = 1/J$, $\alpha = mrg/J$, and $\gamma = \left(J + mr^2/Jl\right) g$, in which

J = moment of inertia of the arm
r = length of the arm
m = pendulum mass
l = length of pendulum
g = gravity constant

Note that the equations are dynamically similar to the rectilinear IP system, and similar system properties are expected. It is possible to control the system theoretically by using measurements of θ (i.e., x_1). In practice, it is relatively easy to control the system by measuring both θ and ϕ.

In the double side-by-side inverted pendulum problem (Figure 8.23), we have two pendulums of equal mass m with lengths l_1 and l_2 mounted on a moving cart. The hinge lines of the pendulums are parallel.
The equations of motion are given by

$$\begin{cases} M\ddot{x} = -mg\theta_1 - mg\theta_2 + u \\ m\ddot{x} = mg\theta_i - ml_i\ddot{\theta}_i \quad i = 1, 2 \end{cases}$$

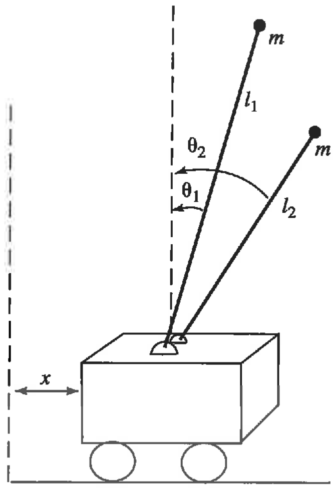


Figure 8.23 Double side-by-side inverted pendulum.

Substituting for $\ddot{x}$ from the first equation into the second set of equations (to eliminate $\ddot{x}$ in the second set) gives

$$\begin{cases} \ddot{\theta}_1 = \dfrac{\alpha}{l_1}\theta_2 + \dfrac{\gamma}{l_1}\theta_1 - \dfrac{1}{l_1}\beta u & \alpha = \dfrac{mg}{M}, \quad \gamma = \alpha + g \\ \ddot{\theta}_2 = \dfrac{\alpha}{l_2}\theta_1 + \dfrac{\gamma}{l_2}\theta_2 - \dfrac{1}{l_2}\beta u & \beta = \dfrac{1}{M} \\ \ddot{x} = -\alpha\theta_1 - \alpha\theta_2 + \beta u \end{cases}$$

Upon defining state variables as $x_1 = x,\ x_2 = \theta_1,\ x_3 = \theta_2,\ x_4 = \dot{x},\ x_5 = \dot{\theta}_1$, and $x_6 = \dot{\theta}_2$, we get the following state equations:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \\ \dot{x}_6 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & -\alpha & -\alpha & 0 & 0 & 0 \\ 0 & \dfrac{\gamma}{l_1} & \dfrac{\alpha}{l_1} & 0 & 0 & 0 \\ 0 & \dfrac{\alpha}{l_2} & \dfrac{\gamma}{l_2} & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \beta \\ -\dfrac{\beta}{l_1} \\ -\dfrac{\beta}{l_2} \end{bmatrix} u$$

Note that the A, B matrices can be partitioned as follows:

$$A = \left[\begin{array}{c|c} 0 & I \\ \hline \bar{A} & 0 \end{array}\right], \quad B = \left[\begin{array}{c} 0 \\ \hline \bar{B} \end{array}\right]$$

where I and 0 stand for identity and zero matrices of appropriate size. This makes it easier to compute the controllability matrix

$$\mathfrak{C} = \begin{bmatrix} 0 & \bar{B} & 0 & \bar{A}\bar{B} & 0 & \bar{A}^2\bar{B} \\ \bar{B} & 0 & \bar{A}\bar{B} & 0 & \bar{A}^2\bar{B} & 0 \end{bmatrix}$$

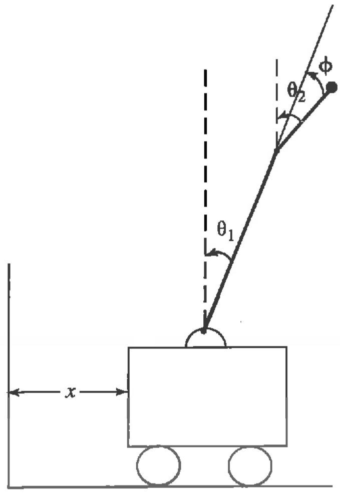


Figure 8.24 The double inverted pendulum.

It is left as an exercise to show that the system is controllable only when $l_1 \neq l_2$. The reader can also verify conditions for observability under various sensor assumptions (sensing x, x and θ_1; θ_1, θ_1 and θ_2; x and θ_1 and θ_2).

We now consider two pendulums, one on top of the other, the double inverted pendulum problem (DIP) (Figure 8.24). Possible sensor variables are the cart position, the angles of the pendulums with respect to the vertical plane (θ_1, θ_2), and also measuring the angle difference between the pendulums (angle ϕ). Researchers have indicated successful control measuring x, θ_1, and ϕ, and even measuring only x and θ_1 (which means stabilizing both pendulums without measuring the apparently crucial angle θ_2 or ϕ). The linearized equations of motion are given by:

$$\begin{cases} r_1\ddot{x} + r_2\ddot{\theta}_1 + r_3\ddot{\theta}_2 = u \\ r_4\ddot{\theta}_1 + r_5\ddot{\theta}_2 + r_2\ddot{x} = r_7\theta_1 \\ r_5\ddot{\theta} + r_6\ddot{\theta}_2 + r_3\ddot{x} = r_8\theta_2 \end{cases}$$

where

$$\begin{aligned} r_1 &= M + m_1 + m_2 \\ r_2 &= m_1 l_1 + m_2 l \\ r_3 &= m_2 l_2 \\ r_4 &= J_1 + m_1 l_1^2 + m_2 l^2 \\ r_5 &= m_2 l_2 l \\ r_6 &= J_2 + m_2 l_2^2 \\ r_7 &= m_1 l_1 g + m_2 l g, \\ r_8 &= m_2 l_2 g \end{aligned}$$

(l_i = distance to center of mass for pendulum i)
l = length of pendulum 1

These implicit differential equations can be solved explicitly for θ_1, θ_2, and u:

$$\begin{bmatrix} r_1 & r_2 & r_3 \\ r_2 & r_4 & r_5 \\ r_3 & r_5 & r_6 \end{bmatrix} \begin{bmatrix} \ddot{x} \\ \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{bmatrix} = \begin{bmatrix} u \\ r_7\theta_1 \\ r_8\theta_2 \end{bmatrix}$$

Solving for $\ddot{x}, \ddot{\theta}_1, \ddot{\theta}_2$, we get

$$\ddot{x} = \alpha_{11}\theta_1 + \alpha_{12}\theta_2 + \beta_1 u$$

$$\ddot{\theta}_1 = \alpha_{21}\theta_1 + \alpha_{22}\theta_2 + \beta_2 u$$

$$\ddot{\theta}_2 = \alpha_{31}\theta_1 + \alpha_{32}\theta_2 + \beta_3 u$$

Letting the state variables be

$$x_1 = x, \quad x_2 = \theta_1, \quad x_3 = \theta_2$$

$$x_4 = \dot{x}, \quad x_5 = \dot{\theta}_1, \quad x_6 = \dot{\theta}_2$$

we write

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \\ \dot{x}_6 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & \alpha_{11} & \alpha_{12} & 0 & 0 & 0 \\ 0 & \alpha_{21} & \alpha_{22} & 0 & 0 & 0 \\ 0 & \alpha_{31} & \alpha_{32} & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} u$$

The A, B matrices can be partitioned as follows.

$$A = \left[\begin{array}{c|c} 0 & I \\ \hline \bar{A} & 0 \end{array}\right], \quad B = \left[\begin{array}{c} 0 \\ \hline \bar{B} \end{array}\right]$$

which is similar in structure to the double side-by-side problem.

The controllability of the system is left as an exercise. The reader can also determine observability under a variety of sensor decisions (measuring x, θ_1, θ_2, x and θ_1, x and θ_2, θ_1 and θ_2, x and θ_1 and θ_2). It is important to appreciate the use of symbolic math programs (such as Symbolic Math Toolbox of MATLAB, Mathematica, Maple, and Macsyma, to name a few) in the various IP problems because they create rather large symbolic matrices that soon test the patience of humans.

8.8 SUMMARY

At least three distinct procedures exist for converting a system describable by an nth-order linear differential equation into a system having n state variables: phase/dual phase variables, canonical (diagonalized) variables, and physical (block diagram) variables. The first of these translates a system transfer function into system matrices, the second provides diagonalized system matrices, while the third preserves actual system quantities (e.g., velocity, current, temperature).

The phase-variable form was shown to be especially convenient for single-and multiple-output system transfer function synthesis, and the dual phase-variable form is convenient for multiple-input system transfer functions. Simulation diagrams are not only useful in transfer function synthesis, they also give a standard, systematic and compact description of a system.

The relationships between signals in a simulation diagram were shown to be a set of coupled first-order differential state equations and linear algebraic output equations relating the system outputs to the state variables. These state-variable equations are compactly expressed using matrix notation:

$$\dot{x} = Ax + Bu$$

$$y = Cx$$

System transfer functions were calculated systematically from the state-variable equations by using matrix algebra:

$$T(s) = C[sI - A]^{-1} B$$

It was seen that all transfer functions of a system share a common characteristic polynomial,

$$|sI - A|$$

A nonsingular change of state variables gives a new representation for a system but leaves the system's input–output relations, its transfer functions, unchanged. Hence, a system characterized by a set of transfer functions may be represented in countless different ways, each differing in the choice of state variables.

A very special set of state variables for a system are those for which the state equations, each of first order, are decoupled from one another. A system so represented is said to be in *normal or diagonal* form:

$$A = \begin{bmatrix} s_1 & 0 & 0 & \cdots & 0 \\ 0 & s_2 & 0 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & & s_n \end{bmatrix}$$

Determining the change of variables that places a system in diagonal form is the *characteristic value problem* of matrix algebra. An alternative transformation method involves expansion of a system transfer function into partial fractions. Systems with repeated characteristic roots cannot be diagonalized: however, they may be placed in a related *Jordan* form, where the repeated root terms involve a distinctive nonzero "block" along the diagonal of the state coupling matrix.

A fundamental system property is stability. It can be studied from transfer function or state space points of view. A system with a given transfer function is BIBO stable if for all bounded inputs, the outputs are bounded. The condition for BIBO stability is that the poles be in the LHP. A system with a given state space realization is asymptotically stable if the states approach 0 as time approaches infinity. The condition for asymptotic stability is that the eigenvalues of the A matrix be in the LHP. These two notions of stability are equivalent when the transfer function of the system contains no common poles and zeros (i.e., there are no pole–zero cancellations). Since most control systems are of the feedback type and are always subject to external disturbances and noise, a more appropriate condition is that all transfer functions between all inputs and outputs be stable. This is called *internal stability*, and it ensures that all signals within the system remain bounded.

Two fundamental properties of a system are controllability and observability. A system is completely controllable if the system state $x(t_f)$ at time t_f can be forced to take on any desired value by applying a control input $u(t)$ over a period of time from t_0 until t_f. A system is completely observable if any initial state vector $x(t_0)$ can be reconstructed by examining the system output $y(t)$ over some period of time from t_0 until t_f.

For a system represented in diagonal form, controllability and observability are apparent from inspection of the input and output coupling matrices. For systems in other than diagonal form, simple rank tests of the controllability and observability matrices,

$$M_c = \left[B \;\vdots\; AB \;\vdots\; \cdots \;\vdots\; A^{n-1}B \right]$$

$$M_o = \begin{bmatrix} C \\ \text{-------} \\ CA \\ \text{-------} \\ \vdots \\ \text{-------} \\ CA^{n-1} \end{bmatrix}$$

may be used.

Whenever the system's transfer function has pole–zero cancellations, its state space realizations will be either uncontrollable or unobservable or both. The uncontrollable or unobservable modes correspond to the canceled poles. If a canceled pole is in the RHP, contradictory answers between BIBO and asymptotic stability result. Most physical systems that are properly modeled satisfy controllability and observability conditions, hence checking the poles for stability is sufficient in most cases.

For an nth-order system

$$x(t) = \Phi(t)\, x\left(0^-\right) + \int_{0-}^{t} \Phi(t-\tau)\, Br(\tau) d\tau$$

where $\Phi(t)$ is the $n \times n$ state transition matrix, we have

$$\Phi(t) = \mathcal{L}^{-1}\left\{(sI - A)^{-1}\right\}$$

The time response can be approximated by using a Taylor series to compute the state transition matrix and the convolution integral. The classic inverted pendulum problem and its variations are discussed at the end of this chapter to demonstrate state space models and system concepts such as stability, controllability, and observability.

REFERENCES

Simulation Diagrams

Jackson, A. S., *Analog computation*. New York: McGraw-Hill, 1960.

Korn, G. A., and Korn, T. M., *Electronic Analog Computers*. New York: McGraw-Hill, 1952.

State Variables

Brockett, R. W., "Poles, Zeros and Feedback: State Space Interpretation." *IEEE Trans. Autom. Control* (April 1965).

Brogan, W. L., *Modern Control Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1991.

DeCarlo, R. A., *Linear Systems: A State Variable Approach with Numerical Implementation*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

De Russo, P. M., Roy, R. J., and Close, C. M., *State Variables for Engineers*. New York: Wiley, 1965.

Friedland, B., *Control Systems Design: An Introduction to State-Space Methods*. New York: McGraw-Hill, 1986.

Gupta, S. C., *Transform and State Variable Methods in Linear Systems*. New York: Wiley, 1966.

Horowitz, I. C., and Shaked, U., "Superiority of Transfer Function Over State Variable Methods in Linear Time-Invariant Feedback System Design." *IEEE Trans. Autom. Control* (February1975).

Kailath, T., *Linear Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1980.

Kalman, R. E., "Mathematical Description of Linear Dynamical Systems." *SIAM J. Control ser.* A, (1): (1963).

Luenberger, D. G., *Introduction to Dynamic Systems: Theory, Models, & Applications*. New York: Wiley, 1979.

Ogata, K., *State Space Analysis of Control Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1967.

Timothy, L. K., and Bona, B. E., *State Space Analysis: An Introduction*. New York: McGraw-Hill, 1968.

Zadeh, L. A., and Desoer, C. A., *Linear System Theory: The State Space Approach*. New York: McGraw-Hill, 1963.

Matrix Algebra and the Characteristic Value Problem

Bellman, R., *Introduction to Matrix Analysis*. New York: McGraw-Hill, 1960.

Noble, B., *Applied Linear Algebra*. Englewood Cliffs, NJ: Prentice-Hall, 1969.

Strang, G., *Linear Algebra and Its Applications*, 2nd ed. New York: Academic Press, 1980.

Strang, G., *Introduction to Applied Mathematics*.: Wellesley-Cambridge Press, 1986.

Controllability and Observability

Gillbert, E. G., "Controllability and Observability in Multivariable Control Systems." *J. SIAM*, ser. A (1963): 128–151.

Kalman, R. E., "Canonical Structure of Linear Dynamical Systems." *Proc. Natl. Acad. Sci. USA* (April 1962): 596–600.

Stubberud, A. R., "A Controllability Criterion for a Class of Linear Systems." *IEEE Trans. Appl. Ind.* 68 (1964): 411–413.

Computational Methods

Faddeeva, D. K., and Faddeeva, V. N., *Computational Methods of Linear Algebra.* San Fransisco: Freeman, 1963.

Shahian, B., and Hassul, M., *Control System Design Using MATRIX$_X$.* Englewood Cliffs, NJ: Prentice-Hall, 1992.

Shahian, B., and Hassul, M., *Control System Design Using MATLAB.* Englewood Cliffs, NJ: Prentice-Hall, 1993.

PROBLEMS

1. Draw phase-variable form simulation diagrams for systems with the following transfer functions. Then write the state-variable equations in matrix form.

(a) $T(s) = \dfrac{-2s+8}{s^2+8}$

(b) $T(s) = \dfrac{10s}{s^3+12s^2+7s+2}$

Ans.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -7 & -12 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} 0 & 10 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

(c) $T(s) = \dfrac{7s^3-2s^2+s}{s^4+3s^3+9s^2+s+1}$

(d) Two outputs:

$$T_{11}(s) = \frac{-s^2+9}{s^3+3s^2+s+4}$$

$$T_{21}(s) = \frac{s^2+s+10}{s^3+3s^2+s+4}$$

Ans.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -4 & -1 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 9 & 0 & -1 \\ 10 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

2. Draw dual phase-variable form simulation diagrams for systems with the following transfer functions. Then write the state-variable equations in matrix form.

(a) $T(s) = \dfrac{-2s+8}{s^2+8}$

(b) $T(s) = \dfrac{2s+8}{3s^3+7s^2+8s+2}$

Ans.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -\frac{7}{3} & 1 & 0 \\ -\frac{8}{3} & 0 & 1 \\ -\frac{2}{3} & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{2}{3} \\ \frac{8}{3} \end{bmatrix} u$$

$$y = [1 \quad 0 \quad 0] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

(c) $T(s) = \dfrac{-s^3+4s^2-9s+4}{s^4+8s^3+2s^2+s+9}$

(d) Two inputs:

$$T_{11}(s) = \frac{3s^2+9}{s^3+3s^2+s+9}$$

$$T_{12}(s) = \frac{s-4}{s^3+3s^2+s+9}$$

Ans.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -3 & 1 & 0 \\ -1 & 0 & 1 \\ -9 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 3 & 0 \\ 0 & 1 \\ 9 & -4 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$y = [1 \quad 0 \quad 0] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

3. Draw simulation diagrams to represent the following systems:

(a)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 1 & 8 & -1 \\ 2 & 0 & 4 \\ -2 & 1 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u$$

$$y = [1 \quad -4 \quad 0] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

(b)
$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 4 & 0 \\ -1 & -1 & 4 \\ 8 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 8 \\ -5 \end{bmatrix} u$$

$$y = [0 \quad -3 \quad 6] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

(c)
$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ -8 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 8 \\ 2 & 4 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

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

(d)
$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ -1 & 2 & 4 \\ -8 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 4 & 0 \\ 0 & 0 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 4 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

4. Draw a simulation diagram to represent the following state equation.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -1 & 3 & 0 \\ 0 & -4 & 8 \\ 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 & 8 \\ -3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$y = [9 \quad 2 \quad 0] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

5. For the following systems, find the transfer function matrices:

(a)
$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ -8 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 2 \\ -4 \end{bmatrix} u$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

(b)
$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix}=\begin{bmatrix}0 & 1\\ -3 & 8\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}+\begin{bmatrix}0 & 1\\ 4 & -7\end{bmatrix}\begin{bmatrix}u_1\\ u_2\end{bmatrix}$$
$$y=[-1 \quad 4]\begin{bmatrix}x_1\\ x_2\end{bmatrix}$$

Ans. $\begin{bmatrix}\dfrac{16s-4}{s^2-8s+3} & \dfrac{-29s+3}{s^2-8s+3}\end{bmatrix}$

(c)
$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\end{bmatrix}=\begin{bmatrix}0 & 4 & 1\\ 0 & 0 & 4\\ -2 & 8 & -4\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}+\begin{bmatrix}4 & 2\\ 0 & 1\\ 3 & -8\end{bmatrix}\begin{bmatrix}u_1\\ u_2\end{bmatrix}$$
$$\begin{bmatrix}y_1\\ y_2\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0 & 4 & -8\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}$$

6. Find the characteristic equations of the following systems. Then determine whether each is stable.

(a)
$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix}=\begin{bmatrix}-1 & 3\\ -3 & 8\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}+\begin{bmatrix}4 & 1\\ -3 & 3\end{bmatrix}\begin{bmatrix}u_1\\ u_2\end{bmatrix}$$
$$\begin{bmatrix}y_1\\ y_2\end{bmatrix}=\begin{bmatrix}1 & 2\\ 0 & 7\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}$$

(b)
$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\end{bmatrix}=\begin{bmatrix}0 & -2 & 3\\ 0 & -3 & -1\\ 0 & 1 & -8\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}+\begin{bmatrix}2\\ -8\\ 4\end{bmatrix}u$$
$$y=[4 \quad 1 \quad 6]\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}$$

Ans. s^3+11s^2+25s, marginally stable

(c)
$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\end{bmatrix}=\begin{bmatrix}-1 & -2 & 2\\ 2 & 0 & 6\\ -1 & 2 & -4\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}+\begin{bmatrix}2 & -1\\ 0 & 0\\ 2 & 8\end{bmatrix}\begin{bmatrix}u_1\\ u_2\end{bmatrix}$$
$$y=[4 \quad 0 \quad -1]\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}$$

7. Although the following transfer functions do not share a common denominator polynomial, they may be made to have a common denominator by multiplying their numerators and denominators by appropriate factors. Find a simulation

diagram and a matrix state-variable diagram and a matrix state-variable representation for a single-input, two-output system with the following two transfer functions:

$$T_{11}(s) = \frac{4s + 1}{(s + 2)(s + 4)}$$

$$T_{21}(s) = \frac{10s}{(s + 1)(s + 4)}$$

The best solutions will involve only three integrators.

8. Use the partial fraction method to find diagonal state equations for single-input, single-output systems with the following transfer functions:

(a) $T(s) = \dfrac{-7s + 4}{s^2 + 8s + 12}$

(b) $T(s) = \dfrac{2s^2 + 3s - 7}{(s + 2)(s + 8)(s + 5)}$

Ans.
$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -8 & 0 \\ 0 & 0 & -5 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} -\frac{5}{18} & \frac{97}{18} & -\frac{28}{9} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$$

(c) $T(s) = \dfrac{10}{s^3 + 8s^2 + 15s}$

9. The following transfer functions for single-input, single-output systems involve complex characteristic roots. Find diagonal state equations for these systems. Then find an alternative block-diagonal representation that does not involve complex numbers.

(a) $T(s) = \dfrac{4s}{s^2 + 2s + 7}$

(b) $T(s) = \dfrac{s^2 + 3s - 8}{(s + 8)(s + 3 + j)(s + 3 - j)}$

(c) $T(s) = \dfrac{4}{(s + 2)\left(s^2 + 2s + 17\right)}$

10. The following transfer functions for single-input, single-output systems involve repeated characteristic roots. Find block diagonal Jordan canonical form state equations for these systems.

(a) $T(s) = \dfrac{3s - 1}{s^2 + 4s + 4}$

(b) $T(s) = \dfrac{s^3 - 4s^2 + s - 2}{(s + 2)(s + 3)^3}$

Ans.

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \\ \dot{z}_4 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & -3 & 1 & 0 \\ 0 & 0 & -3 & 1 \\ 0 & 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} u$$

$$y = [-28 \quad 68 \quad 16 \quad 29] \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix}$$

(c) $T(s) = \dfrac{7s^3}{(s+2)^2(s+6)^2}$

11. The following systems have real characteristic roots. Find alternative diagonal state equations.

(a)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -9 & 1 \\ -20 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 4 \end{bmatrix} u$$

$$y = [2 \quad -3] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

(b)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -6 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

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

Ans.

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & -3 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 3 & 1 \\ -2 & -1 \end{bmatrix} u$$

$$y = [-2 \quad -3] \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

(c)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -5 & 1 \\ -4 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

(d)
$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -3 & 1 & 0 \\ -2 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} u$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

12. The following system has a set of complex conjugate characteristic roots. Find an alternative diagonal set of state equations. Then find another alternative set of state equations where the complex root terms are placed in real number block diagonal form.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -17 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

13. The following system has a repeated characteristics root. Find an alternative set of state equations in Jordan form:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -9 & -15 & -7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

14. Find diagonal state equations for systems with the following transfer function matrices:

(a) $$T(s) = \begin{bmatrix} \dfrac{-6s}{s^2 + 4s + 3} & \dfrac{4}{s^2 + 4s + 3} \end{bmatrix}$$

(b) $$T(s) = \begin{bmatrix} \dfrac{s^2 - 4}{s^3 + 3s^2 + 2s} & \dfrac{4s - 8}{s^3 + 3s^2 + 2s} & \dfrac{s^2 + 3s - 4}{s^3 + 3s^2 + 2s} \end{bmatrix}$$

Ans. One possibility is the following:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} -2 & -4 & -2 \\ 3 & 12 & 6 \\ 0 & -8 & -3 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

$$y = [1 \quad 1 \quad 1]\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

(c)

$$T(s) = \begin{bmatrix} \dfrac{3s-1}{s^2+4} \\ \dfrac{-s+8}{s^2+4} \end{bmatrix}$$

(d)

$$T(s) = \frac{\begin{bmatrix} s \\ -3s^2-4 \\ 8 \end{bmatrix}}{s^3+3s^2+2s}$$

Ans. One possibility is the following:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} u$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & -1 \\ -2 & 7 & -8 \\ 4 & -8 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

15. Find a simulation diagram and a matrix state-variable representation for a two-input, two-output system with the following transfer function matrix:

$$T(s) = \begin{bmatrix} \dfrac{4s}{s^2+3s+2} & \dfrac{s-3}{s^2+3s+2} \\ \dfrac{-6}{s^2+3s+2} & \dfrac{s+4}{s^2+3s+2} \end{bmatrix}$$

16. A transfer function with equal numerator and denominator polynomial degrees may be expanded as a constant plus a proper remainder, as in the following example:

$$T(s) = \frac{3s^2+2s-4}{s^2+3s+2} = 3 + \frac{-7s-10}{s^2+3s+2}$$

It may be realized by adding to the system output a term that is proportional to the system input. The resulting state-variable equations have the form

$$\dot{x} = Ax + Bu$$
$$y = Cx + Du$$

Find matrices A, B, C, and D for a system with the transfer function above. For such a second-order single-input, single-output system, the state-variable equations will be of the form

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} u$$

$$y = [c_1 \quad c_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + du$$

17. Use controllability and observability matrices to determine whether the following systems are completely controllable and whether these systems are completely observable. In addition, determine BIBO and asymptotic stability in each case.

(a)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 2 & -4 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u$$

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

(b)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 3 & 0 & -5 \\ -2 & 1 & 5 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 4 & 1 & -3 \\ 3 & 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Ans. completely controllable but not completely observable

(c)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -2 \\ 3 & -3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 1 & -1 \\ 2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 & 4 & 1 \\ 0 & -2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

18. Write state equations for systems—each with modes e^{2t}, e^{-3t}, $e^{(-4+j)t}$, and $e^{(-4-j)t}$—that have the following properties:

(a) The mode e^{2t} is uncontrollable.

(b) The mode e^{-3t} is unobservable.

(c) The mode e^{2t} is both uncontrollable and unobservable.

(d) The modes $e^{(-4+j)t}$ and $e^{(-4-j)t}$ are uncontrollable.

(e) The mode e^{-3t} is uncontrollable and the mode e^{2t} is unobservable.

19. The system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

is unstable. Can the instability be detected from input–output measurements? Determine whether the system is completely observable. Then calculate the system transfer function. A common factor in the numerator and the denominator should cancel.

Repeat if instead the output equation is

$$y = \begin{bmatrix} -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

20. Find a third-order system, if possible, in phase-variable form, that is not completely controllable.

21. Show that an nth-order system with n outputs is completely observable if its $(n \times n)$-output coupling matrix is nonsingular.

22. Use the time domain method involving convolution to solve

$$\dot{x} = -2x + u(t)$$

with

$$x(0^-) = 7$$
$$u(t) = 3e^{3t}$$

23. Use Laplace transform methods to find the state response of the following systems for $t \geqslant 0$ with the given inputs and initial conditions. Also find the system output.

(a)
$$\dot{x} = -3x + 2r(t)$$
$$y = 4x$$
$$x(0^-) = 7$$
$$r(t) = 5u(t), \text{ where } u(t) \text{ is the unit step function}$$

(b)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -3 & 1 \\ 7 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u$$

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

$$\begin{bmatrix} x_1(0^-) \\ x_2(0^-) \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$

$u(t) = \delta(t)$, the unit impulse function

Ans. $\begin{bmatrix} 1.01e^{1.54t} & 2.99e^{-4.54t} \\ 4.61e^{1.54t} & -4.61e^{-4.54t} \end{bmatrix}$; $y(t) = 1.01e^{1.54t} + 2.99e^{-4.54t}$

(c)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -3 & -8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 & 4 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

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

$$\begin{bmatrix} x_1(0^-) \\ x_2(0^-) \end{bmatrix} = \begin{bmatrix} 7 \\ -3 \end{bmatrix}$$

$$\begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} = \begin{bmatrix} e^t \\ 5 \end{bmatrix}$$

(d)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -3 & 1 & 0 \\ -2 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} u$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$\begin{bmatrix} x_1(0^-) \\ x_2(0^-) \\ x_3(0^-) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$u(t) = 2$

24. The state transition matrix for a certain system is

$$\Phi(t) = \begin{bmatrix} \frac{1}{2}e^{-t} + \frac{1}{2}^{-2t} & 2(e^{-t} - e^{-2t}) \\ 3e^{-t} - 3e^{-2t} & -e^{-t} + 2e^{-2t} \end{bmatrix}$$

Find the state $x(t)$ for $t \geqslant 0$ if all system inputs are zero and

$$x(0^-) = \begin{bmatrix} x_1(0^-) \\ x_2(0^-) \end{bmatrix} = \begin{bmatrix} -7 \\ 2 \end{bmatrix}$$

25. Calculate state transition matrices for systems with the following state coupling matrices A, using

$$\Phi(t) = \mathcal{L}^{-1}\left\{[sI - A]^{-1}\right\}$$

(a) $\begin{bmatrix} 0 & 1 \\ -6 & -4 \end{bmatrix}$

(b) $\begin{bmatrix} -1 & 4 \\ 8 & -2 \end{bmatrix}$

Ans. $\begin{bmatrix} 0.4e^{-7.1t} + 0.6e^{4.1t} & 0.3e^{-7.1t} - 0.3e^{4.1t} \\ 0.6e^{-7.1t} - 0.6e^{4.1t} & 0.54e^{-7.1t} + 0.46e^{4.1t} \end{bmatrix}$

(c) $\begin{bmatrix} -5 & 1 & 0 \\ -4 & 0 & 5 \\ 0 & 0 & 0 \end{bmatrix}$

26. Show that the state transition matrix for a diagonalized system is diagonal, with the system modes along the diagonal.

27. Find state-variable equations for each of the systems of Fig. P8.27. Then find the transfer function(s) from the original drawing and compare with the transfer function(s) of the state-variable model.

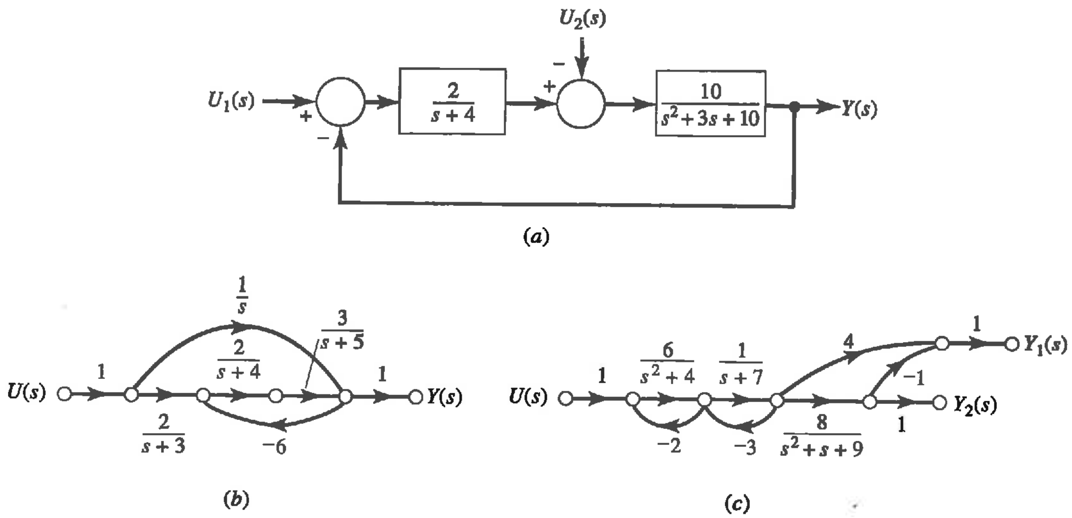


Figure P8.27

CHAPTER

9

State Space Design

9.1 Preview

Feedback along with many of its properties has been the underlying theme in control engineering. In classical design, the plant output is fed back and processed by standard compensators (lead–lag, PID) to modify system dynamics with a view to satisfying stability and performance requirements. Typically, a design engineer tries to reshape the system root locus, or Bode plot, to meet the requirements. These methods allow limited control of the closed loop poles.

In state space design, the basic idea of feedback is maintained. Rather than feeding back one or two outputs, we feedback the complete system state vector to modify system dynamics. We will see that feeding back the complete state vector gives the designer total control over the closed-loop poles. Such complete control is possible because, ideally, the system's state contains all the available information about the system. In a sense, we are using *full information* about the system to control its behavior.

9.2 State Feedback and Pole Placement

Consider the simple first-order system described by

$$\dot{x} = x + u \qquad x(0) = x_0$$

$$y = x$$

where u is the input to the plant. The pole of this open loop system is 1, indicating that the open-loop system is unstable. To stabilize the system, we may feed the state back using some gain k where

$$u = -kx$$

Therefore, the compensated system becomes

$$\dot{x} = x - kx = (1-k)x$$

The closed-loop pole is $1-k$, which results in an asymptotically stable system for $k > 1$. In fact, by a suitable choice of k, we can place the closed-loop pole anywhere on the real axis.

As another example of placing the system poles at desired locations with state feedback, consider the following single-input, single-output system, which is described in phase-variable form:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -5 & -7 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u$$

$$y = [-2 \quad 4 \quad 3] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

This system is shown in the simulation diagram of Figure 9.1(a) and, from Mason's gain rule, we find that it has transfer function

$$T(s) = \frac{3/s + 4/s^2 + -2/s^3}{1 + 3/s + 7/s^2 + 5/s^3} = \frac{3s^2 + 4s - 2}{s^3 + 3s^2 + 7s + 5}$$

Since the characteristic equation factors as follows:

$$s^3 + 3s^2 + 7s + 5 = (s + 1 + j2)(s + 1 - j2)(s + 1)$$

its poles are at $s = -1 - j2, -1 + j2$, and -1.

With state feedback,

$$u(t) = -k_1x_1 - k_2x_2 - k_3x_3 + r(t)$$

the state equations are of the form

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -k_1 - 5 & -k_2 - 7 & -k_3 - 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} r(t)$$

$$y = [-2 \quad 4 \quad 3] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

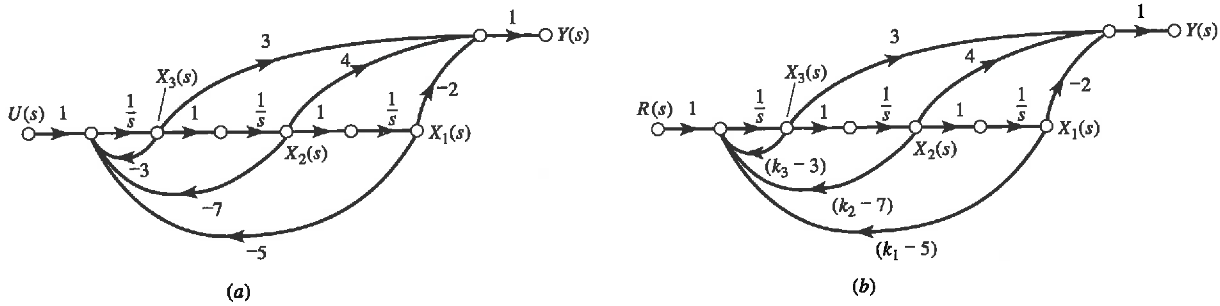


Figure 9.1 State feedback example. (a) Open-loop system. (b) System with state feedback.

as diagrammed in Figure 9.1(b). The feedback system has transfer function, in terms of the feedback gain constants k,

$$T(s) = \frac{3/s + 4/s^2 - 2/s^3}{1 + (3 + k_3)/s + (7 + k_2)/s^2 + (5 + k_1)/s^3}$$

$$= \frac{3s^2 + 4s - 2}{s^3 + (3 + k_3)s^2 + (7 + k_2)s + (5 + k_1)}$$

The coefficients of the characteristic equation may be chosen at will, by appropriately selecting k_1, k_2, and k_3. If, for instance, it is desired that the system poles be located at $s = -4$, -4, and -5, the characteristic polynomial should be

$$(s + 4)(s + 4)(s + 5) = s^3 + 13s^2 + 56s + 80$$
$$= s^3 + (3 + k_3)s^2 + (7 + k_2)s + (5 + k_1)$$

which will be the case for

$$k_1 = +75$$
$$k_2 = +49$$
$$k_3 = +10$$

In general, given a system in state space form

$$\dot{x} = Ax + Bu$$

Using state feedback, we get

$$u = -Kx \rightarrow \dot{x} = Ax - BKx = (A - BK)x$$

The closed-loop system matrix has been modified from A to $A - BK$. The closed-loop characteristic polynomial is given by

$$\Delta_c(s) = |sI - (A - BK)|$$

The closed-loop poles (or eigenvalues) are the roots of the foregoing polynomial. If the original system is represented in phase-variable form, the A matrix will be in companion form. An important property of companion form matrices is that their characteristic polynomial can be written by inspection. In fact the coefficients of the characteristic polynomial can be read off the last row of the A matrix. For instance, in the preceding example, the last row of the $A - BK$ matrix is

$$\text{Last row} = [-k_1 - 5 \quad -k_2 - 7 \quad -k_3 - 3]$$

and the characteristic polynomial is

$$\text{Characteristic polynomial} = s^3 + (3 + k_3)s^2 + (7 + k_2)s + (5 + k_1)$$

It is clear from this equation that, in this case, any desired polynomial can be achieved by selecting the feedback gains.

Can all systems be stabilized by using state feedback? To answer this question, consider the following example.

$$\dot{x} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} x + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} u$$

Using state feedback, the closed-loop characteristic equation becomes

$$\Delta_c(s) = \left| \begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} [k_1 \quad k_2] \right|$$

$$= s^2 + (-3 + b_2k_2 + b_1k_1)s + (2 - b_2k_2 - 2b_1k_1)$$

To see whether we can place the closed-loop poles anywhere in the complex plane, consider an arbitrary second-order polynomial

$$\text{Desired characteristic polynomial} = \Delta_d(s) = s^2 + \alpha s + \beta$$

To find the state feedback gains, these polynomials must be identical. Thus

$$b_1k_1 + b_2k_2 - 3 = \alpha$$

$$-b_2k_2 - 2b_1k_1 + 2 = \beta$$

Writing the equations in matrix form, we get

$$\begin{bmatrix} b_1 & b_2 \\ -2b_1 & -b_2 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} 3 + \alpha \\ -2 + \beta \end{bmatrix}$$

We can solve for the unknown gains for any right-hand side (i.e., for any desired polynomial) if and only if the determinant of the coefficient matrix is nonzero.

$$\begin{vmatrix} b_1 & b_2 \\ -2b_1 & -b_2 \end{vmatrix} = -b_1b_2 + 2b_1b_2 = b_1b_2$$

Hence, both b_1 and b_2 must be nonzero. But note that this is the same condition as complete controllability of the system (refer to Section 8.6). We conclude that to place the closed-loop poles of this system arbitrarily, the system must be controllable. Let us see what happens if this condition is violated. Suppose b_1 is zero. In this case, the equations will be consistent when

$$3 + \alpha = 2 - \beta \qquad \text{or} \qquad \beta = -1 - \alpha$$

Imposing this condition on the desired polynomial

$$\text{Achievable desired polynomial} = s^2 + \alpha s - 1 - \alpha$$

But this polynomial is unstable. Hence, we can never stabilize this system if it is uncontrollable.

The results of the previous example can be generalized as follows:

The closed-loop poles of a system can be arbitrarily placed anywhere in the complex plane if and only if the system is completely controllable.

Controllability was defined earlier as the ability to move the system states from any initial state to any final state. We observe here that this is equivalent to the ability of placing (or shifting) the system poles anywhere in the complex plane. The foregoing result is commonly called the *pole placement* (or *pole-shifting*) theorem. If a system is not controllable, we may still be able to move some of the poles but not all of them. In general, controllable modes can be shifted, whereas uncontrollable modes are fixed. For instance, in the preceding example where b_1 was 0, the mode at $s = 1$ was uncontrollable (could not be moved), but the mode at $s = 2$ could be placed anywhere (we can show that to move the mode from 2 to -2, $k_2 = 4/b_2$).

9.2.1 Stabilizability

Careful study of the pole placement theorem reveals that controllability is too strong a condition. It allows arbitrary pole placement, even in the RHP. But we normally are not interested in placing system poles in the RHP. It turns out in practice that a weaker notion than controllability is sufficient for most purposes. This notion is called *stabilizability*. It refers to the ability to move only the unstable modes of the system. Therefore, we say a system is *stabilizable* if the unstable modes are controllable or, equivalently, if the uncontrollable modes are stable. The easiest way to check this is to convert the system to modal form, then check each mode and the corresponding row in the B matrix. The next example illustrates the notion.

Defining stabilizabilty.

$$\dot{x} = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u, \quad \dot{z} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} z + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u$$

Neither of these systems is controllable. In the first system, the stable mode at -1 is not controllable, whereas the unstable mode at 2 is controllable. Hence, the system is stabilizable. For instance, by using state feedback control—$u = -kx_1$, with $k > 2$—the system is stabilized. In the second system, observe that the unstable mode at 1 is not controllable; therefore, the system is not stabilizable. Note that either stability or controllability implies that the system is stabilizable. For control system design, stabilizability is the minimum condition the system must satisfy for any problem. A system model that does not satisfy this condition is a poor model. Either the system must be remodeled or its structure must be modified to render it stabilizable.

Several formulas exist for computation of the state-feedback gain. *Ackermann's formula* is an example of one (for SISO controllable systems). Given the desired characteristic equation

$$\Delta_d(s) = s^n + \alpha_1 s^{n-1} + \cdots + \alpha_n$$

The state feedback gain vector is given by

$$k = [0, 0, \ldots, 1]M_c^{-1}\Delta_d(A)$$

Ackermann's formula.

where $\Delta_d(A) = A^n + \alpha_1 A^{n-1} + \cdots + \alpha_{n-1}A + \alpha_n I$ and M_c is the controllability matrix.

As an application of this formula, consider the following double-integrator plant.

$$G(s) = \frac{1}{s^2}$$

or in state space form

State feedback control of the double-integrator system.

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} x$$

Let us place the system poles at $-1 \pm j$. That is, $\Delta_d(s) = s^2 + 2s + 2$. Then, k is given by

$$k = [0 \quad 1] \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{-1} \left\{ \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}^2 + 2 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + 2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\} = [2 \quad 2]$$

Let the plant input be

$$u = -kx + r$$

Then the compensated system block diagram will be as shown in Figure 9.2(a).

Rearranging the block diagram, as in Figure 9.2(b), shows that the state feedback controller is equivalent to a feedback PD compensator of the form

$$H(s) = 2(1 + s)$$

The open-loop transfer function is given by

$$G(s)H(s) = \frac{2(s + 1)}{s^2}$$

This can be used to perform classical root locus and Bode analysis on the system. We can also obtain stability margins for the system. The open-loop and closed-loop transfer functions can also be obtained directly in terms of the state space matrices as shown next.

The plant is represented by

$$\dot{x} = Ax + Bu$$

Laplace-transforming, we have

$$sX(s) = AX(s) + BU(s) \qquad X(s) = (sI - A)^{-1}BU(s)$$

Let $\Phi(s) = (sI - A)^{-1}$; then $X(s) = \Phi(s)BU(s)$, and

$$U(s) = -KX(s) = -K\Phi(s)BU(s)$$

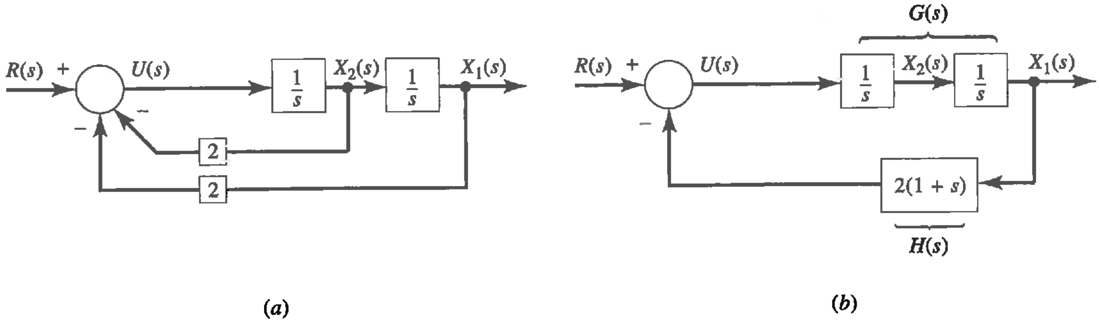


Figure 9.2 (a) Compensated system block diagram. (b) Rearranged block diagram.

State feedback open-loop transfer function.

Figure 9.3(c) shows that if the plant input is u, the signal fed back to the summing junction is $-K\Phi B$. Hence, the open-loop transfer function is $K\Phi B$. If this is compared to a classical feedback configuration, shown in Figure 9.3(d), we get

$$G(s)H(s) = K\Phi(s)B$$

The closed-loop transfer function of the state feedback system, shown in Figure 9.4, is given by

$$u = -Kx + r$$

Applying the input to the system (i.e., closing the loop), Laplace-transforming, and solving for the state vector, we get

$$\dot{x} = (A - BK)x + Br \rightarrow X(s) = (sI - A + BK)^{-1}BR(s)$$

Closed-loop transfer function under state feedback.

Substituting the state into the output equation gives the closed-loop transfer function

$$Y(s) = CX(s) = C(sI - A + BK)^{-1}BR(s) \rightarrow$$

Applying these to the present example, we get

$$k\Phi(s)B = [2 \quad 2]\begin{bmatrix} s & -1 \\ 0 & s \end{bmatrix}^{-1}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = 2\frac{s+1}{s^2}$$

$$T(s) = [1 \quad 0]\left\{s\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ -2 & -2 \end{bmatrix}\right\}^{-1}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{s^2 + 2s + 2}$$

9.2.2 Choosing Pole Locations

State feedback gives the designer the option of relocating all system closed-loop poles. This is in contrast with classical design, where the designer can only hope to achieve a pair of complex conjugate poles that are dominant. Because all other poles and zeros may fall anywhere, meeting the design specifications becomes a matter of trial and error. With the freedom of choice rendered by state feedback comes the responsibility of selecting these poles judiciously.

Although there is no magic choice, there are some guidelines we can follow. Moving the poles around is costly. Suppose a first-order system has a pole at -1.

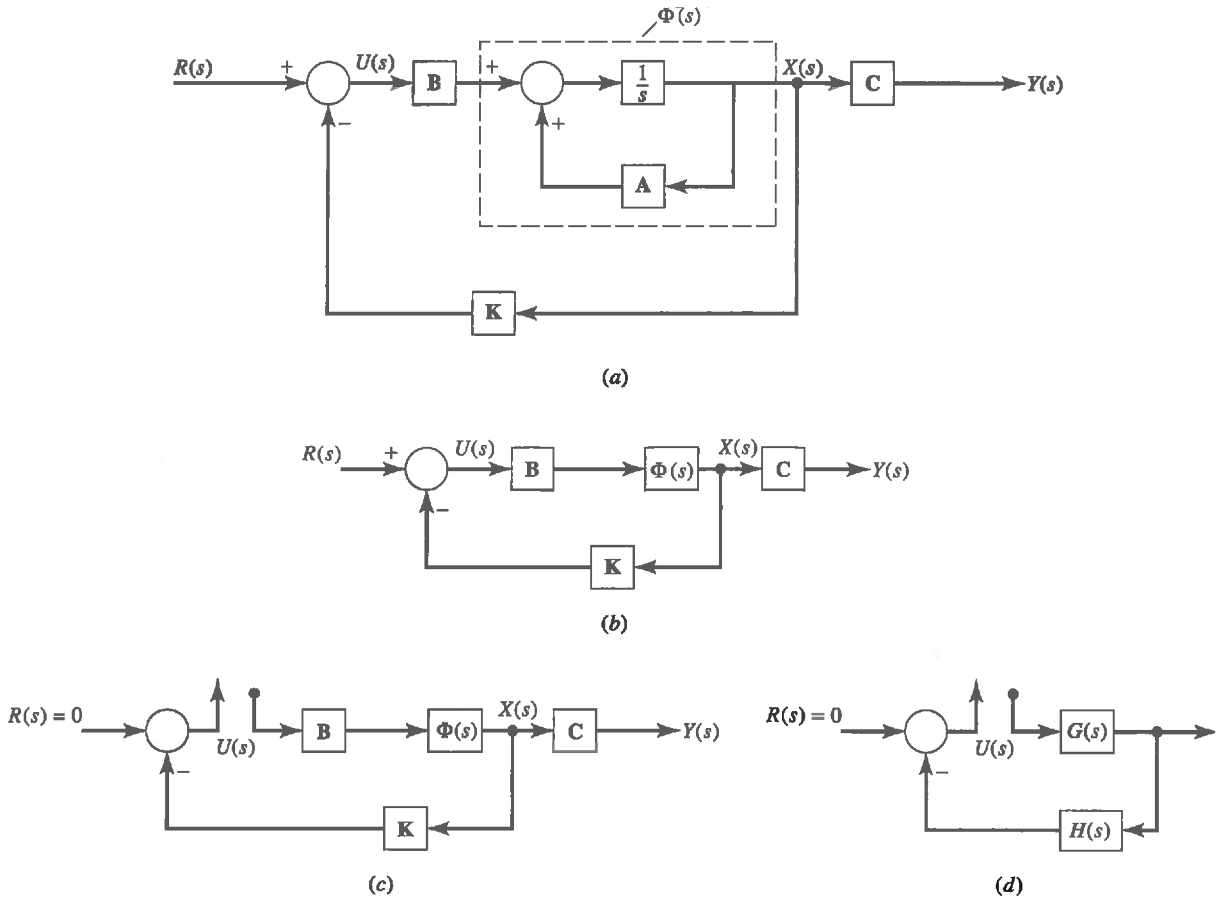


Figure 9.3 (a) Block diagram of a state feedback compensated system. (b) Simplified diagram (c) Diagram for computing the open-loop transfer function with the loop broken at the input. (d) Classical feedback configuration.

If this pole is moved to -10, the time constant is reduced, resulting in a faster system. The system output may be voltage, pressure, position, velocity, temperature, and so on. Sensors and an actuator are needed to do the job. A faster system may require a more accurate sensor and a larger or stronger actuator (such as a motor) to perform the task. These are some of the obvious costs that may be associated with pole shifting. Therefore, one guideline is that if an LHP pole has an acceptable location, leave it alone. Poles in the RHP, or poles on or close to the imaginary axis must be moved. A rule, suggested by optimal control, is that RHP poles must be reflected about the imaginary axis to minimize control energy (i.e., a pole at 2 must be shifted to -2). A pair of complex conjugate poles can be placed to meet transient response requirements. One should also be cautious about the temptation to push the poles too far into the LHP. The consequence of this is that the system bandwidth increases, and the system becomes sensitive to noise.

Guidelines from optimal control.

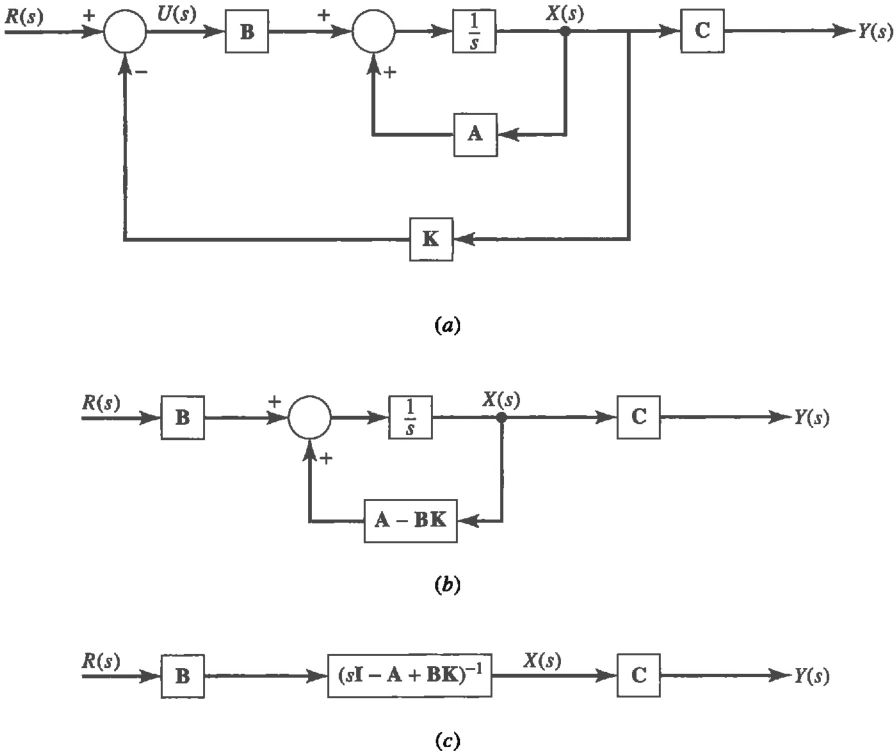


Figure 9.4 (a) State feedback compensated system. (b) The system after closing the feedback loop. (c) Diagram for the closed-loop transfer function.

Suppose an unstable system has poles at $\{2, -10, -0.1 - j10, -0.1 + j10\}$. It is specified that overshoot be less than 10%, and the settling time be less than 5. The pole at 2 can be shifted to -2. The pole at -10 can remain. The complex poles have a very small damping ratio. They could be moved to $-1 - j$ and $-1 + j$ to meet the transient response specifications. Hence, the desired pole locations are $\{-2, -10, -1 - j, -1 + j\}$. Figure 9.5 shows the resulting step response; clearly, the specifications are satisfied.

State feedback does not affect plant zeros.

Note that the closed-loop zero locations were not specified. The reason is that state feedback does not affect the system zeros. Therefore, if they are at undesirable locations, nothing can be done about them by using state feedback. Because steady state tracking properties depend on poles and zeros, this means that tracking properties cannot be helped by state feedback alone. For instance, the plant in the preceding example was a double integrator, hence it is a type 2 system with zero steady state error to unit step and ramp inputs. After state feedback, it became a type 0 system. For example, we put

$$T_E(s) = 1 - T(s) = \frac{s^2 + 2s + 1}{s^2 + 2s + 2}$$

and for a unit step input, $T_E(0) = \frac{1}{2} = 0.5$; hence, the steady state error is 0.5.

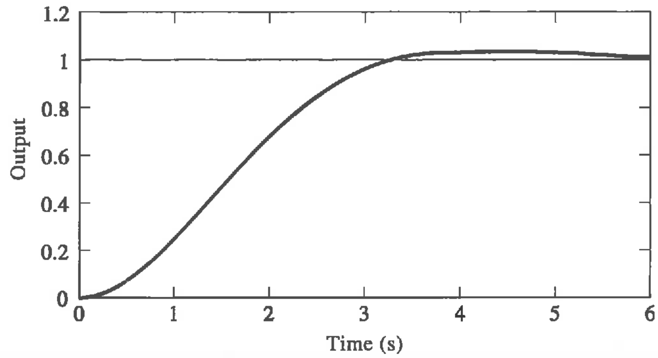


Figure 9.5 Step response of a fourth-order system with specified poles.

9.2.3 Limitations of State Feedback

State feedback is not usually practical.

The major limitation of state feedback is that it is not usually practical. It is not practical for two reasons. One is that state feedback leads to PD-type compensators, which have infinite bandwidth, whereas real components and compensators always have finite bandwidth. Another reason is that it is simply not possible or practical to sense all the states and feed them back. In reality, only certain states or combinations of them are measurable as outputs. Consequently, any practical compensator must rely on system outputs and inputs only for compensation; that is called *output feedback* and is discussed in the following sections.

❑ DRILL PROBLEMS

D9.1 For the state-feedback systems described by the following equations, choose the feedback gain constants k_i to place the closed-loop system poles at the indicated locations:

(a)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -3 & -6 & -7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u$$

$$u = \begin{bmatrix} -k_1 & -k_2 & -k_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + r$$

$$y = \begin{bmatrix} 2 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Closed-loop poles at $s = -3, -4,$ and -5

Ans. $k_1 = +57, k_2 = +41, k_3 = +5$

(b)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -2 & 1 & 0 \\ 4 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u$$

$$u = \begin{bmatrix} -k_1 & -k_2 & -k_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + r$$

$$y = \begin{bmatrix} 2 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Closed-loop poles at $s = -3 \pm j3, -3$

Ans. $k_1 = +30, k_2 = +26, k_3 = +7$

D9.2 Design a state-feedback controller for the following systems. Determine the controller gains, open-loop transfer functions, and closed-loop transfer functions. Use the open-loop transfer functions to obtain root locus, Bode plots, and gain and phase margins.

(a)

$$\dot{x} = \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & -1 \end{bmatrix} x$$

$$u = -kx + r$$

Closed-loop poles at $s = -1 \pm j$

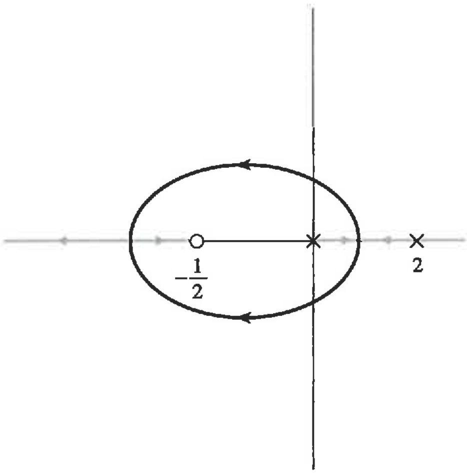


Figure D9.2a

Ans. $k = [4 \quad 2]$, $G(s)H(s) = (4s+2)/s(s-2)$, gain margin $= -6$ dB, phase margin $= 52°$, $T(s) = (s-1)/(s^2+2s+2)$

(b)

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u$$

$$y = [0 \quad 1]x$$

$$u = -kx + r$$

Closed-loop poles at $s = -1 \pm j$

Ans. $k = [2 \quad 3]$, $G(s)H(s) = (2s+3)/(s^2-1)$, gain margin $= -9.5$ dB, phase margin $= 53°$, $T(s) = 1/(s^2+2s+2)$

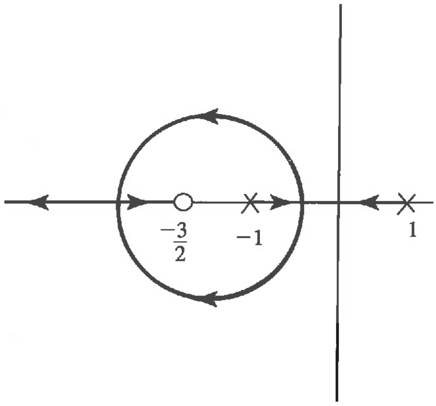


Figure D9.2b

9.3 Tracking Problems

The state feedback compensator design discussed so far has been a regulator problem. The command input has been ignored or set to zero. The objective has been to design a system that is stable and rejects disturbances. The issue of steady state error was not dealt with. In fact, if you check the step response of the double-integrator example, you will discover some steady state error. This is true even though the original plant is a type 2 system. The compensator has changed the system type. How do we incorporate the command input into the system, and design for a given steady state accuracy?

For a step input, this is easy. We simply add a summing junction with a gain, such as

$$u = -Kx + \bar{N}r$$

where r is the reference (or command) input. The constant, $\bar{N}$, can be computed to ensure zero steady state error to step inputs. Let us derive this for the case of state feedback.

$$\dot{x} = Ax + Bu = Ax + B(-Kx + \bar{N}r) = (A - BK)x + B\bar{N}r$$

By definition of steady state, the states and output must reach a constant value—that is,

$$\dot{x}_{ss} = 0 \qquad \text{which implies that} \qquad 0 = (A - BK)x_{ss} + B\bar{N}r$$

Solving for the steady state output

$$x_{ss} = -(A - BK)^{-1}B\bar{N}r \rightarrow y_{ss} = -C(A - BK)^{-1}B\bar{N}r$$

This inverse exists because $(A - BK)$ is a stable matrix. The steady state error to a constant input is the difference between the input and the output. It is therefore given by

$$e_{ss} = r - y_{ss} = r + C(A - BK)^{-1}B\bar{N}r = [1 + C(A - BK)^{-1}B\bar{N}]r$$

For zero step tracking error, the steady state output must be equal to the command input, therefore

$$\bar{N} = \frac{-1}{C(A - BK)^{-1}B}$$

9.3.1 *Integral Control*

The preceding technique places a gain outside of the feedback loop. As you know, when elements of a control system are not within a feedback loop, the overall system will be quite sensitive to elements outside the loop. An alternate method that allows us to achieve zero steady state error to step inputs is integral control. The idea is a classical one; we place an integrator in the forward path in series with the system, thereby increasing its system type. The block diagram of a state feedback controller using integral control is shown in Figure 9.6

Because the integral term increases the order of the system by 1, we need to augment the plant model with an added state variable to account for this. Define the new state variable as

$$x_i = \int e\,dt = \int (r - y)dt = \int (r - Cx)dt$$

Therefore, $\dot{x}_i = r - Cx$, and the augmented plant equation becomes

$$\begin{bmatrix} \dot{x} \\ \dot{x}_i \end{bmatrix} = \begin{bmatrix} A & 0 \\ -C & 0 \end{bmatrix}\begin{bmatrix} x \\ x_i \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u + \begin{bmatrix} 0 \\ 1 \end{bmatrix} r$$

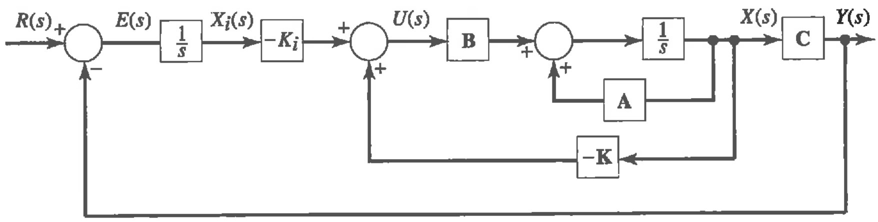


Figure 9.6 Block diagram of a state feedback compensated system with integral control.

where the zero matrices have compatible dimensions. The controller is modified to

$$u = -Kx - K_i x_i = -[K \quad K_i]\begin{bmatrix} x \\ x_i \end{bmatrix}$$

Using this controller, the compensated system becomes

$$\begin{bmatrix} \dot{x} \\ \dot{x}_i \end{bmatrix} = \begin{bmatrix} A - BK & -BK_i \\ -C & 0 \end{bmatrix}\begin{bmatrix} x \\ x_i \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} r$$

The characteristic polynomial of the compensated system is then equated with the desired one to solve for the controller gains.

Double-integrator example with integral control.

Let us design a state feedback controller for the double-integrator system incorporating integral control action. First we need to augment the plant

$$\begin{bmatrix} \dot{x} \\ \dot{x}_i \end{bmatrix} = \left[\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \\ \hline -1 & 0 & 0 \end{array}\right]\begin{bmatrix} x \\ \hline x_i \end{bmatrix} + \left[\begin{array}{c} 0 \\ 1 \\ \hline 0 \end{array}\right] u$$

$$y = [1 \quad 0 \mid 0]\begin{bmatrix} x \\ \hline x_i \end{bmatrix}$$

The poles of the system will be shifted to $\{-1 \pm j, -5\}$. Note that an extra pole needs to be selected because of the extra state. Solving for K, we get

$$K = [12 \quad 7 \quad -10]$$

The steady state output due to unit step input is

$$y_{ss} = -C(A - BK)^{-1}B = -\left[1 \quad 0 \mid 0\right]\left[\begin{array}{cc|c} 0 & 1 & 0 \\ -12 & -7 & 10 \\ \hline -1 & 0 & 0 \end{array}\right]^{-1}\left[\begin{array}{c} 0 \\ 0 \\ \hline 1 \end{array}\right] = 1$$

❑ DRILL PROBLEMS

D9.3 For the plants described in Drill Problem D9.2, use state feedback to design an integral controller. The desired pole locations are as indicated.

(a) Closed-loop poles at $s = -1 \pm j, -5$
Ans. $k = [9 \quad 22], K_i = 10$

(b) Closed-loop poles at $s = -1 \pm j, -5$
Ans. $k = [7 \quad 13], K_i = -10$

D9.4 For the plant described by $G(s) = 1/(s-1)$, use state feedback to design an integral controller that will place closed-loop poles at $s = -1 \pm j$. Also, draw a block diagram or signal flow graph of the system.
Ans. $k = 3, K_i = 2$

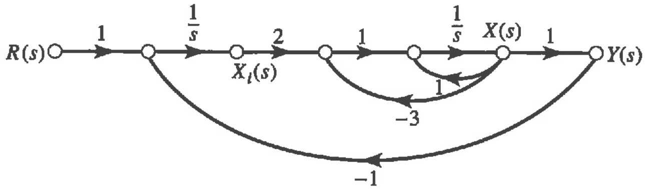


Figure D9.4

9.4 Observer Design

An observer estimates the states.

To fully implement the advantages of state feedback, all states should be fed back. Typically, some of the states are measured by sensors and the rest must be estimated by another device. An observer, shown in Figure 9.7, is a device that uses the inputs and outputs of a system to produce estimates of its states. Observers either are built using electronic components (hardware) or are equivalent computer (or microprocessor) programs (software) in digital control implementations. The word "device" implies either implementation. The idea of an observer is that if we have all system parameters, we can always simulate the model on an analog or digital computer. Even though we do not have access to system states, we have full access to the states of our simulation. Therefore, an observer is a device which simulates the original system. Letting $\hat{x}$ denote the state estimates in Figure 9.8, we have

$$\dot{x} = Ax + Bu \qquad x(0) = x_0$$
$$y = Cx$$

The observer dynamics are

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - C\hat{x}) \quad \hat{x}(0) = \hat{x}_0$$

The observer design proceeds by defining the error between the states and their estimates. Let

$$\tilde{x} = x - \hat{x}$$

To see how the error evolves with time, a differential equation for the error must be obtained.

$$\dot{\tilde{x}} = \dot{x} - \dot{\hat{x}} = Ax + Bu - A\hat{x} - Bu - L(y - C\hat{x}) = (A - LC)\tilde{x}$$
$$\tilde{x}(0) = \tilde{x}_0$$

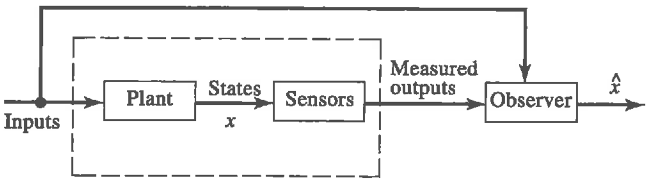


Figure 9.7 Block diagram of system and its observer.

The observer error will go to zero asymptotically if and only if the matrix $(A - LC)$ is a stable matrix (i.e., its eigenvalues are in the LHP). But this matrix contains the matrix L, which is to be determined. It turns out that if the system is completely observable, L can be chosen such that the eigenvalues of $(A - LC)$ are arbitrary. The L matrix, which is a column vector for single-output systems, is called the *observer gain*. Also note that the observer equation can be written as

The observer works if the system is observable.

$$\dot{\hat{x}} = (A - LC)\hat{x} + Bu + Ly \quad \hat{x}(0) = \hat{x}_0$$

It can be seen that the eigenvalues of $(A - LC)$ are the observer poles. Hence, controllability allows arbitrary plant pole placement, whereas observability allows arbitrary observer pole placement.

Recall from Chapter 8 that there was a duality between phase-variable and dual phase-variable forms, in that one form could be converted into the other form via an algorithm. There is a similar duality between the notions of controllability and observability. For instance, if we transpose the controllability matrix M_c (rank remains the same), and replace A' with A and B' with C, we get M_0. Because stabilizability is a weaker version of controllability, we can also define its dual. The dual notion of *stabilizability* is called *detectability*. We say a system is *detectable* if the unstable modes are observable, or equivalently, the unobservable modes are stable.

Detectability.

Consider the following example.

$$\dot{x} = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u \qquad \dot{z} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} z + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u$$

$$y = [1 \quad 0]x \qquad w = [1 \quad 0]z$$

In the first system, the mode at -1 is unobservable but detectable. In the second system, the mode at 1 is unobservable and undetectable. Let us design an observer for the double-integrator plant $G(s) = 1/s^2$.

An observer for the double-integrator system.

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$y = [1 \quad 0]x$$

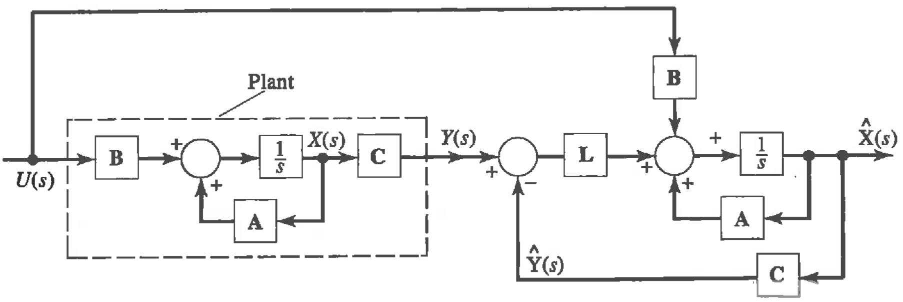


Figure 9.8 Block diagram of system with observer.

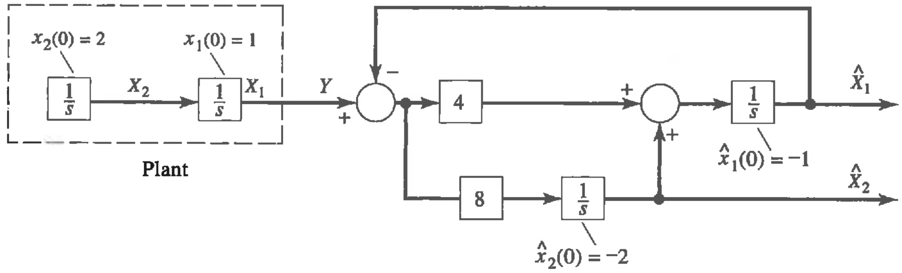


Figure 9.9 Structure of the observer for the double-integrator plant. Plant initial conditions are (1, 2), observer initial conditions are (−1, −2).

You can verify that the system is observable. We will design an observer with poles at $\{-2 \pm j2\}$. This choice is arbitrary, and more will be said about it later. The design starts with setting the observer characteristic polynomial equal to our desired polynomial, and solving for L.

$$|sI - (A - LC)| = \left| s \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} l_1 \\ l_2 \end{bmatrix} [1 \quad 0] \right|$$
$$= s^2 + l_1 s + l_2$$

$$s^2 + l_1 s + l_2 = s^2 + 4s + 8 \rightarrow l_1 = 4, l_2 = 8, \text{ i.e., } L = \begin{bmatrix} 4 \\ 8 \end{bmatrix}$$

The observer equations are (with the input u set to zero).

$$\dot{\hat{x}}_1 = \hat{x}_2 + 4(y - \hat{x}_1)$$
$$\dot{\hat{x}}_2 = 8(y - \hat{x}_1)$$

The structure of this observer is shown in Figure 9.9.

This observer was simulated to verify its convergence. The zero-input response plots are shown in Figure 9.10. The plant initial conditions are (1, 2) and the observer initial conditions are (−1, −2). Because there are no inputs, the second state stays at a constant value of 2 and the first state is a ramp starting at 1. The observer estimates have an initial error, but they converge to true state values in about 2 s. The convergence time is the settling time of the observer, which is controlled by the real part of its poles. Because the real parts of observer poles are −2, the observer is expected to converge in about 4 time constants (i.e., 2 s).

❑ DRILL PROBLEM

D9.5 For each of the following systems, x_1 is measured while x_2 must be estimated by an observer. Select the observer gain L so both eigenvalues are as required.

Write the observer equations and create a block diagram or signal flow graph for the interconnected system and observer. Also use computer software to simulate the system and the observer, verifying the convergence of the observer.

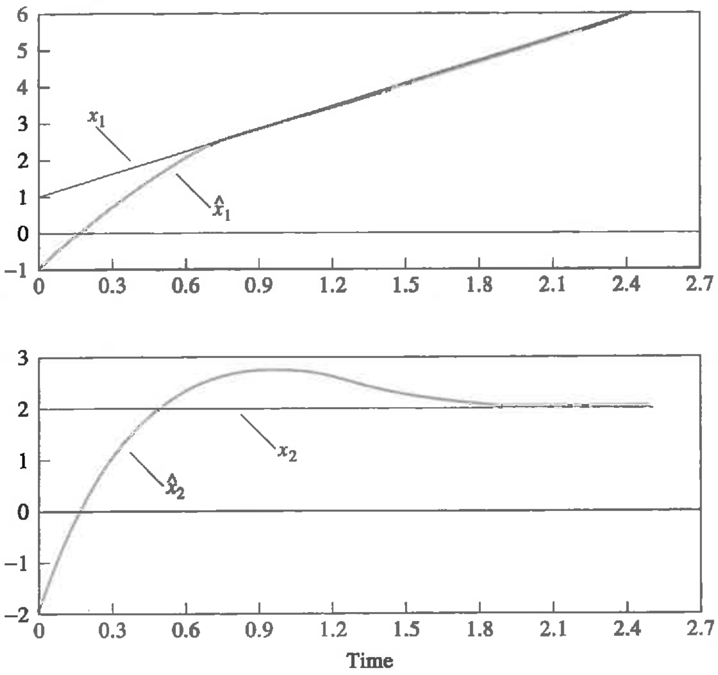


Figure 9.10 Simulation of the double-integrator plant and its observer. (a) Plot of the first state and its estimate. (b) Plot of the second state and its estimate.

(a)

$$A = \begin{bmatrix} -2 & -4 \\ 1 & -4 \end{bmatrix} \quad B = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \quad C = [1 \quad 0]$$

Observer eigenvalues should be $\{-50, -50\}$

Ans.

$$L = \begin{bmatrix} 94 \\ -528 \end{bmatrix}$$

$$\dot{\hat{x}}_1 = -2\hat{x}_1 - 4\hat{x}_2 + 2u + 94(x_1 - \hat{x}_1)$$

$$\dot{\hat{x}}_2 = \hat{x}_1 - 4x_2 - 528(x_1 - \hat{x}_1)$$

(b)

$$A = \begin{bmatrix} -4 & -4 \\ 1 & -2 \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \quad C = [1 \quad 0]$$

Observer eigenvalues should be $\{-10, -10\}$.

Ans.

$$L = \begin{bmatrix} 14 \\ -15 \end{bmatrix}$$

$$\dot{\hat{x}}_1 = -4\hat{x}_1 - 4\hat{x}_2 + 14(x_1 - \hat{x}_1)$$

$$\dot{\hat{x}}_2 = \hat{x}_1 - 2x_2 + 2u - 15(x_1 - \hat{x}_1)$$

(c)

$$A = \begin{bmatrix} 0 & 1 \\ -10 & -3 \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 10 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

Observer eigenvalues should be $\{-20, -20\}$.

Ans.

$$L = \begin{bmatrix} 37 \\ 279 \end{bmatrix}$$

$$\dot{\hat{x}} = \hat{x}_2 + 37(x_1 - \hat{x}_1)$$

$$\dot{\hat{x}}_2 = -\ 10\hat{x}_1 - 3\hat{x}_2 + 10u + 279(x_1 - \hat{x}_1)$$

9.4.1 Control Using Observers

It was pointed out earlier that state feedback requires that all states be available for feedback. What happens when state estimates obtained using observers replace the actual states? Will we still be able to stabilize the system, or even to place its poles arbitrarily? The answer is yes. To see this, we need to obtain the equations of the closed-loop system, and examine its eigenvalues.

If the system is of order n, the observer will also be of the same order. When the observer estimates are fed back into the system, the order of the closed-loop system will be $2n$. The states of the composite system are the original plant states, x, and their estimates, $\hat{x}$. The composite state equations are obtained as follows.

$$\dot{x} = Ax + Bu \quad x(0) = x_0$$

$$y = Cx$$

Estimated state feedback:

$$u = -K\hat{x}$$

Observer:

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - C\hat{x}), \quad \hat{x}(0) = \hat{x}_0$$

Closing the loop (i.e., feeding the control back into the system and the observer), we get

$$\dot{x} = Ax - BK\hat{x}$$

$$\dot{\hat{x}} = (A - LC)\hat{x} - BK\hat{x} + Ly = (A - LC - BK)\hat{x} + LCx$$

Combining these two equations gives the closed-loop composite system

Closed-loop system.

$$\begin{bmatrix} \dot{x} \\ \dot{\hat{x}} \end{bmatrix} = \begin{bmatrix} A & -BK \\ LC & A - BK - LC \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix}$$

The system will be closed-loop stable if the eigenvalues of this block matrix are in the LHP. Although, this is not obvious from the matrix, we will soon show that it is indeed the case. For now, let us return to our example and see whether the double-integrator system is stabilized when the observer-based control is used. We recall that the control and observer gains were

$$k = [2 \quad 2] \qquad L = \begin{bmatrix} 4 \\ 8 \end{bmatrix}$$

The closed-loop system becomes

$$\begin{bmatrix} \dot{x} \\ \dot{\hat{x}} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & -2 \\ 4 & 0 & -4 & 1 \\ 8 & 0 & -10 & -2 \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix} \qquad \begin{bmatrix} x(0) \\ \hat{x}(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 1.5 \\ -1 \\ -2 \end{bmatrix}$$

The eigenvalues of this matrix are $\{-1 \pm j, -2 \pm j2\}$. These are the plant and observer pole locations selected earlier. The zero-input response of the system due to the specified initial conditions is shown in Figure 9.11.

The system is clearly asymptotically stable. Therefore, the system has been stabilized by using state estimates instead of actual states. Figure 9.12, compares the

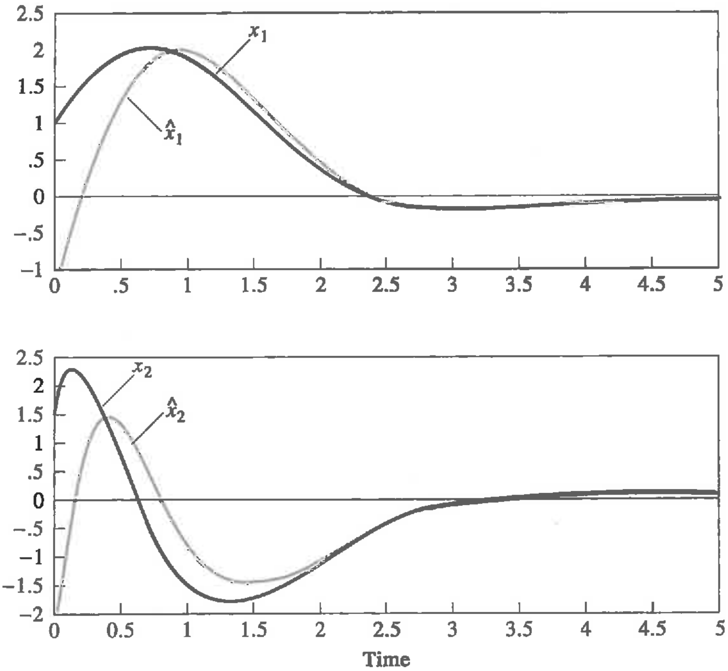


Figure 9.11 States and their estimates of the observer-based compensated double-integrator system. (a) First state and its estimate. (b) Second state and its estimate.

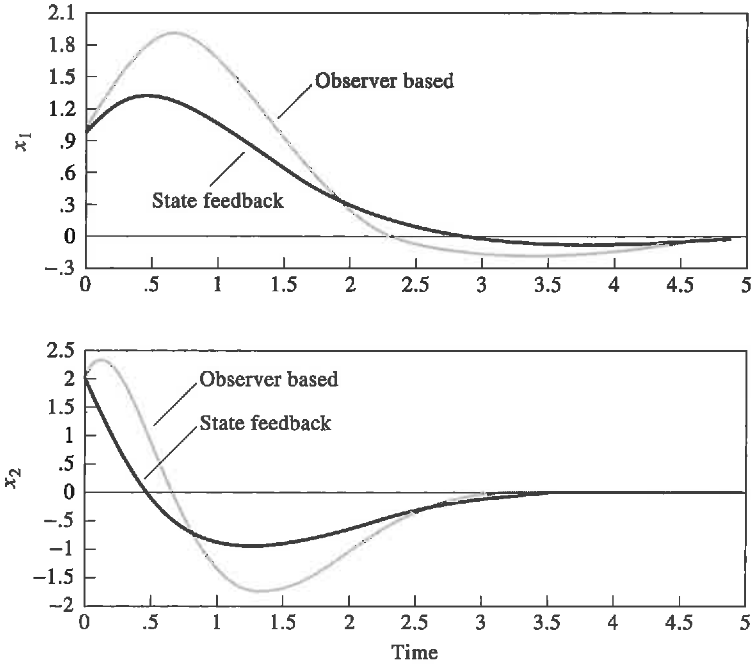


Figure 9.12 Comparison of state feedback and observer-based compensated systems. (a) First state. (b) Second state.

two designs: the zero-input responses of the system using full state feedback and using observer-based control. The figure indicates some performance degradation due to estimated states. The observer-based response has a longer settling time and a larger overshoot.

9.4.2 Separation Property

We want to show that the observer-based control results in a stable system. Recall from Section 8.3 that state space transformations allow us to look at the system in a different, and possibly more informative, way. Let us introduce the following state transformation.

$$\begin{bmatrix} x \\ \hat{x} \end{bmatrix} = P \begin{bmatrix} z \\ w \end{bmatrix} \qquad \text{where } P = \begin{bmatrix} I_n & 0_n \\ I_n & -I_n \end{bmatrix}$$

It turns out that the P matrix is its own inverse; hence the new states are

$$\begin{bmatrix} z \\ w \end{bmatrix} = \begin{bmatrix} I_n & 0_n \\ I_n & -I_n \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix} = \begin{bmatrix} x \\ x - \hat{x} \end{bmatrix} = \begin{bmatrix} x \\ \tilde{x} \end{bmatrix}$$

This transformation allows us to look at the states and observer errors. Recall that under state transformation, the "A" matrix becomes "$P^{-1}AP$". Therefore the transformed system becomes

$$\begin{bmatrix} \dot{x} \\ \dot{\tilde{x}} \end{bmatrix} = \begin{bmatrix} A - BK & BK \\ 0 & A - LC \end{bmatrix} \begin{bmatrix} x \\ \tilde{x} \end{bmatrix}$$

Note that, in this new realization, the matrix is block triangular (a block matrix is a matrix whose elements are matrices, as in our case). Now, we use a fact from matrix algebra.

FACT: Eigenvalues of a block triangular matrix are equal to the eigenvalues of the matrices along the diagonal blocks.

Using this fact, and the knowledge that system eigenvalues remain invariant under state transformations, we conclude that the closed-loop poles (or eigenvalues) of an observer-based control system are the union of the observer poles, and the poles of the system selected under state feedback (also known as *controller poles*). Because controllability allows us to place the eigenvalues of $A - BK$ arbitrarily, and observability does the same for the eigenvalues of $A - LC$, we see that under these two conditions we have complete freedom in controller and observer pole selections. Moreover, these selections can be made independently of each other because of the block-triangular structure of the closed-loop system matrix. This property is known as the *separation property*.

Closed-loop poles = controller poles plus observer poles.

The separation between the control and the observer problem implies that we can find the controller gain, assuming the states are available, design an observer to estimate the states, and then use the estimates in place of the actual states.

9.4.3 Observer Transfer Function

It is instructive to obtain the transfer function of the observer-based compensator and compare it with classical designs. For instance, we note that state feedback resulted in a feedback PD-type compensator. The compensator output is the plant input u, and the compensator input is the plant output y. The observer also has a feedback from u, but this can be eliminated by substituting u into the observer equation. The derivation follows.

$$u = -K\hat{x}$$

$$\begin{aligned} \dot{\hat{x}} &= A\hat{x} + Bu + L(y - C\hat{x}) = A\hat{x} - BK\hat{x} + Ly - LC\hat{x} \\ &= (A - BK - LC)\hat{x} + Ly \end{aligned}$$

Taking the Laplace transform of this equation and solving it for $\hat{X}(s)$, we get

$$\hat{X}(s) = (sI - A + BK + LC)^{-1}LY(s)$$

Substituting this into the control equation gives the transfer function

$$U(s) = -K(sI - A + BK + LC)^{-1}LY(s)$$

$$U(s) = -H(s)Y(s) \rightarrow H(s) = K(sI - A + BK + LC)^{-1}L$$

Transfer function for observer-based compensator.

The observer-based compensator for the double-integrator system is

$$H(s) = \frac{24s + 16}{s^2 + 6s + 18}$$

The open-loop transfer function is given by

$$G(s)H(s) = \frac{24s + 16}{s^2(s^2 + 6s + 18)}$$

The root locus and Bode plots are shown in Figures 9.13 and 9.14.

Stable but no guaranteed stability margins.

From the Bode plots, we read the gain and phase margins. They are approximately 10 dB and 36°, respectively. Observe that the gain margin is fairly small (i.e., raising the gain by a factor of about 3.16 would be destabilizing). One limitation of observer-based design is that we have no direct control over the stability margins. Thus designs that are perfect on paper might not work in a real situations. This is because we usually have imperfect models of our systems, and stability margins provide some protection against model uncertainties. System designed with low margins are inherently sensitive to model errors and may become unstable in actual operation.

No control over compensation poles and zeros.

Observing the resulting compensator, we note that it has a pair of complex conjugate poles and one zero, so it has no classical counterpart. In fact, the compensator poles and zeros could end up anywhere in the complex plane, including the RHP, and we have no control over this issue.

In the preceding example, the observer initial conditions were chosen rather arbitrarily. How do we choose them in general? The best choice is to use the plant initial conditions. This we do not have, however. If we did, given the plant model, we could have numerically solved for the states. From the convergence point of

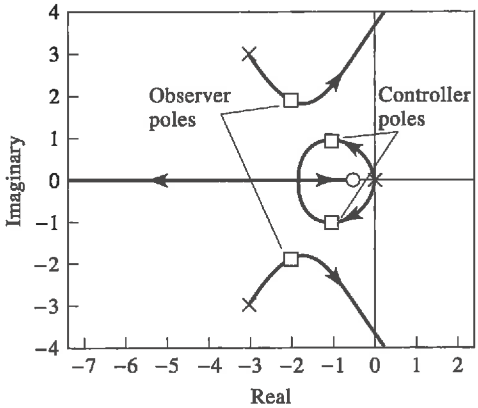


Figure 9.13 Root locus of the observer-based compensated double-integrator system.

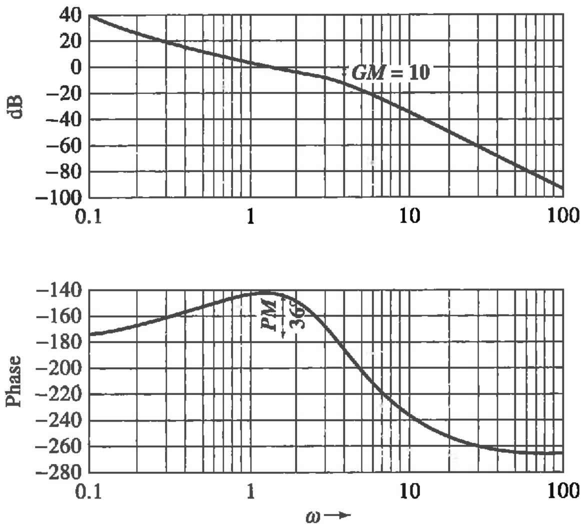


Figure 9.14 Bode plots of an observer-based compensated double-integrator system.

view, this is not an issue because the matrix $(A - LC)$ is a stable matrix, and the observer errors will go to zero independently of the initial conditions. It can be shown that because the initial output is known, an optimal choice (*in a mathematical sense, not necessarily optimal in a physical sense*) for observer initial conditions is

Choosing observer initial conditions.

$$\hat{x}(0) = C'(CC')^{-1}y(0)$$

What about the location of observer poles? There are several guidelines available. One guideline is to choose the observer poles that are faster than the controller poles. This choice ensures that the observer converges faster than the system, so the controller will be using more accurate estimates, thereby reducing the degradation caused by the observer. Again, one must caution against pushing the observer poles too far into the LHP, because this increases the system bandwidth and makes it more susceptible to noise. An alternate guidelines suggested by results from robust control is to choose the observer poles at the plant zeros (if the system has RHP zeros, use their LHP images).

Choosing observer poles.

❑ DRILL PROBLEMS

D9.6 For the system of Drill Problem D9.1, choose control gain k to place the closed-loop system poles at the indicated locations. In addition, design observers with the indicated poles. Use MATLAB to verify that the closed-loop system is stable using estimated states.

(a) Controller poles at $\{-6, -1 \pm j\}$

observer poles at $\left\{-5 \pm j5, -\sqrt{2}\right\}$.

$$\textbf{Ans.}\quad k = [9 \quad 8 \quad 1] \quad L = \begin{bmatrix} 2.336 \\ 2.3037 \\ 0.2579 \end{bmatrix}$$

(b) Controller poles at $\{-3, -1 \pm j\}$

observer poles at $\{-3, -5 \pm j5\}$

$$\textbf{Ans.}\quad k = [6 \quad 6 \quad 3], \quad L = \begin{bmatrix} -2.7619 & 5.1429 \\ 1.1905 & 1.7143 \\ -9.6667 & 14.0000 \end{bmatrix}$$

D9.7 For the systems of Drill Problem D9.2, design observers with indicated poles. Using the same control gains simulate the closed-loop systems to verify stability. Obtain compensator transfer functions, and root locus and Bode plots. Use these to determine stability margins.

(a) observer poles $= \{-1, -2\}$

$$\textbf{Ans.}\quad L = \begin{bmatrix} 12 \\ 7 \end{bmatrix}, \; H(s) = \frac{62s - 4}{s^2 + 7s - 38}$$

(b) observer poles = $\{-1, -2\}$

$$\textbf{Ans.} \quad L = \begin{bmatrix} 3 \\ 3 \end{bmatrix}, H(s) = \frac{15(s+1)}{s^2+5s+11}$$

9.5 Reduced-Order Observer Design

There is no need to estimate states that are directly measured.

The observer introduced in Section 9.4 has the same order as the system and is referred to as the *identity* or *full-order observer*. If the system has n states and m measurements are available, it is possible to build an observer that estimates the states that are not measured, hence reducing the order of the observer. It seems reasonable that an observer of order $(n - m)$ should be sufficient. This was first introduced by D. Luenberger and is referred to as *reduced-order* (or *Luenberger*) *observer.* The reduction in order leads to simpler and more economical compensators. If the number of measurements m is large, the benefits could be substantial.

Before we drive this observer, we make an assumption on the structure of the measurement matrix C. We will assume that C has the form

$$C = [I \quad 0]$$

where $I = m \times m$ and $0 = m \times (n - m)$

The consequence of this is that it allows us to divide the states into two categories: measured, unmeasured. That is,

$$y = [I \quad 0]\begin{bmatrix} x_m \\ x_u \end{bmatrix} = x_m$$

We can then partition the system accordingly as

$$\begin{bmatrix} \dot{x}_m \\ \dot{x}_u \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} x_m \\ x_u \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u$$

$$y = [I \quad 0]\begin{bmatrix} x_m \\ x_u \end{bmatrix}$$

The unmeasured portion of the system is

$$\dot{x}_u = A_{22}x_u + (A_{21}x_m + B_2u)$$

The terms within the parentheses are known quantities, so they are collected together. Because there are m measured states, the number of unmeasured states is $n - m$, so we will build an observer of order $n - m$ to estimate these states. The observer structure is given by the following procedure (this is the same procedure used for the full-order observer): copy the system equation, replace unknown quantities by their estimates, and add a correction term multiplied by the observer gain. The correction term is the difference between the plant output and the observer output.

$$\dot{\hat{x}}_u = A_{22}\hat{x}_u + (A_{21}x_m + B_2u) + L(\text{correction term})$$

The correction term in the full-order observer case was $(y - C\hat{x})$. In the present case, it is

$$y - [I \quad 0]\begin{bmatrix} x_m \\ \hat{x}_u \end{bmatrix} = y - x_m = 0$$

Therefore, using the output will not provide any correction. However, we note that if the outputs are available, we can assume that their derivatives are also available. Now, observe that the derivative of the plant output is equal to the measured portion of the system, that is,

$$\dot{y} = \dot{x}_m = A_{12}x_u + (A_{11}x_m + B_1u) \rightarrow \dot{y} - A_{11}x_m - B_1u = A_{12}x_u$$

where we have collected the known, or measured, quantities on the left-hand side. We can use the known quantities on the left as a substitute for plant output (we are basically using all the information available from the system), and the right-hand side as the observer output. Substituting this in the observer equation, we get

$$\dot{\hat{x}}_u = A_{22}\hat{x}_u + (A_{21}x_m + B_2u) + L(\dot{y} - A_{11}x_m - B_1u - A_{12}\hat{x}_u)$$

To verify that this scheme works, we need to show that the error dies out. Define the error as

$$\tilde{x}_u = x_u - \hat{x}_u$$

and derive a differential equation for the error.

Reduced-order observer error.

$$\dot{\tilde{x}}_u = (A_{22} - LA_{12})\,\tilde{x}_u$$

This is similar to the full-order observer error equation

Full-order observer error.

$$\dot{\tilde{z}} = (A - LC)\tilde{z}$$

Recall that if the pair (C, A) is observable (i.e., the system is observable), L can be chosen to place observer poles anywhere in the complex plane. Comparing the two error equations, by analogy we conclude that the same would be true for the reduced-order case under the condition that the pair of matrices (A_{12}, A_{22}) are observable. Luenberger showed that this condition is equivalent to the observability of the original system—that is, (C, A).

We are almost done. The last step is to eliminate the output derivative in the observer. Differentiation is to be avoided in system design because it is a noise-enhancing operation. Looking at the observer equation suggests that a simple change of variable will eliminate the derivative term. This change of variable is given by

$$z = \hat{x}_u - Ly \qquad \text{or} \qquad \hat{x}_u = z + Ly$$

A bit of algebra results in the final form of the reduced-order observer

Final form of the reduced-order observer.

$$\boxed{\begin{aligned} \dot{z} &= Dz + Fy + Gu \\ \hat{x}_u &= z - Ly \end{aligned}}$$

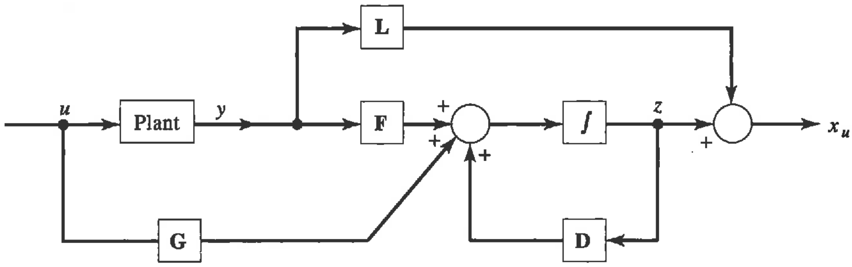


Figure 9.15 Block diagram of the reduced-order observer.

where

$$\boxed{\begin{aligned} D &= A_{22} - LA_{12} \\ F &= DL + A_{21} - LA_{11} \\ G &= B_2 - LB_1 \end{aligned}}$$

The block diagram of the observer is shown in Figure 9.15.

Let us design a reduced-order observer for the double-integrator system. Because there are two states and one measurement, we require a first-order observer. The plant equations are repeated:

Designing a reduced-order observer for the double-integrator system.

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$y = [1 \quad 0]x$$

Although, we can simply plug into the formulas, we will repeat the observer derivation for this simple example from scratch. The unmeasured portion of the system is given by

$$\dot{x}_2 = u \rightarrow A_{21} = A_{22} = 0, \ B_2 = 1$$

The measured portion of the system, to be used for the correction term, is given by

$$\dot{x}_1 = \dot{y} = x_2 \rightarrow (\dot{y} - \hat{x}_2) \quad \text{(correction term)}$$

Now, we copy the unmeasured portion, and add the correction term to get the observer.

$$\dot{\hat{x}}_2 = u + L(\dot{y} - \hat{x}_2)$$

The observer error equation becomes

$$\dot{\tilde{x}}_2 = \dot{x}_2 - \dot{\hat{x}}_2 = u - u - L(\dot{y} - \hat{x}_2) = -L(x_2 - \hat{x}_2) = -L\tilde{x}_2$$

Choosing the observer pole at -2 yields $L = 2$. To eliminate the derivative term, let

$$z = \hat{x}_2 - Ly = \hat{x}_2 - 2y$$

First-order observer to estimate x_2.

The final form of the observer becomes

$$\begin{aligned} \dot{z} &= -2z - 4y + u \\ \hat{x}_2 &= z + 2y \end{aligned}$$

Figure 9.16 is a diagram of this reduced-order observer.

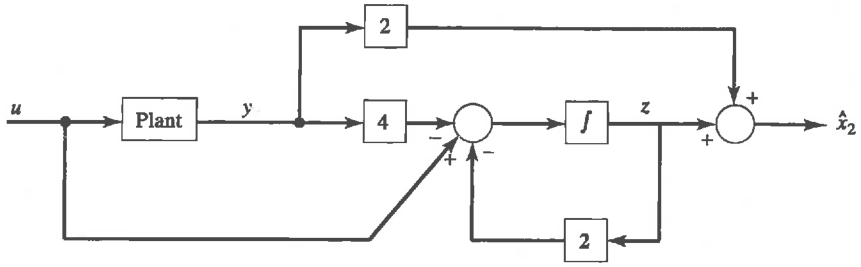


Figure 9.16 Block diagram of the reduced-order observer for the double-integrator example.

Note that in this example, the system was second-order, and we had one measurement. Solving for L involved solving a first-order algebraic equation, so we obtained a unique solution. In general, when there are n states and m measurements, there are $(m-1)$ degrees of freedom in solving for L, and the solution is not unique.

❑ DRILL PROBLEM

D9.8 For the systems of Drill Problem D9.5, design a reduced-order, first-order observer to estimate only x_2. Select the observer eigenvalue as indicated. Write the observer equation and use computer software to verify its operation.

(a) The observer eigenvalue should be −50.

Ans. $D = -50, F = 553, G = 23, L = -11.5$

$$\begin{cases} \dot{z} = -50z + 23u + 553y \\ \hat{x}_2 = z - 11.5y \end{cases}$$

(b) The observer eigenvalue should be −10.

Ans. $D = -10, F = 13, G = 2, L = -2$

$$\begin{cases} \dot{z} = -10z + 2u + 13y \\ \hat{x}_2 = z - 2y \end{cases}$$

9.5.1 Separation Property

The separation property also holds in the reduced-order case. It can be derived by combining the closed-loop system and error system equations

$$\begin{bmatrix} \dot{x} \\ \dot{\tilde{x}}_u \end{bmatrix} = \begin{bmatrix} A - BK & BK_2 \\ 0 & A_{22} - LA_{12} \end{bmatrix} \begin{bmatrix} x \\ \tilde{x}_u \end{bmatrix}$$

Therefore, the closed-loop eigenvalues are the union of the controller and observer eigenvalues. In the above, the control gain vector has been partitioned as

$$u = -\begin{bmatrix} K_1 & K_2 \end{bmatrix} \begin{bmatrix} x_m \\ \hat{x}_u \end{bmatrix} = -K_1 x_m - K_2 \hat{x}_u$$

9.5.2 Reduced-Order Observer Transfer Function

The compensator transfer function is derived by substituting for u in the observer equation

$$H(s) = K_2 [sI - (D - GK_2)]^{-1} (F - GK_1 - GK_2 L) + (K_1 + K_2 L)$$

The same techniques used for the full-order case can be used to handle tracking problems and integral control.

A first-order compensator will be obtained for the double-integrator problem. The control gain, observer gain, and observer parameters, obtained earlier, are as follows:

$$k = [k_1 \quad k_2] = [2 \quad 2]\,, \; L = 2, \; D = -2, \; F = -4, \; G = 1$$

First-order compensator for the double-integrator system.

By using the preceding parameters, we get the compensator transfer function

$$H(s) = \frac{6s + 4}{s + 4}$$

The compensator is recognized as a classical lead compensator. Root locus and Bode plots of the compensated system are shown in Figures 9.17 and 9.18, respectively.

The compensated system has 45° phase margin and infinite gain margin. Comparing this with the full-order observer-based compensator, we note that the reduced-order case has resulted in a simpler (first-order) compensator with better stability margins. The zero-input response of the compensated system is shown in Figure 9.19. The zero-input responses of all three designs (i.e., state feedback, observer-based, and reduced-order observer-based compensators) are shown in Figure 9.20 for comparison.

At the beginning of Section 9.5 we assumed that the C matrix is of the form

$$C = [I \quad 0]$$

The case when C is not of the form $[I \quad 0]$

This is not a restrictive assumption, because through a linear transformation, we can always convert C to this form. The transformation is given as follows. Choose any

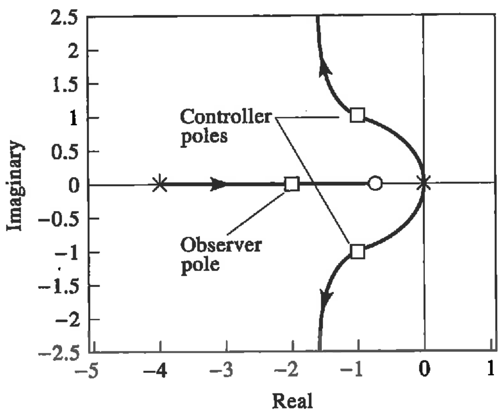


Figure 9.17 Root locus of the reduced-order, observer-based, compensated double-integrator system.

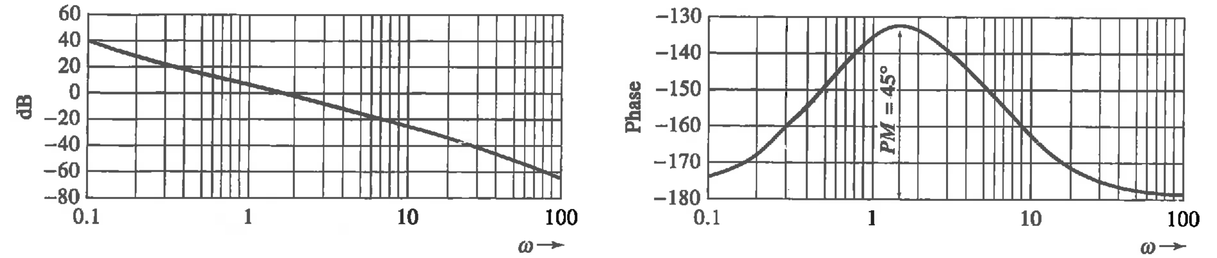


Figure 9.18 Bode plots of the reduced-order, observer-based, compensated double-integrator system.

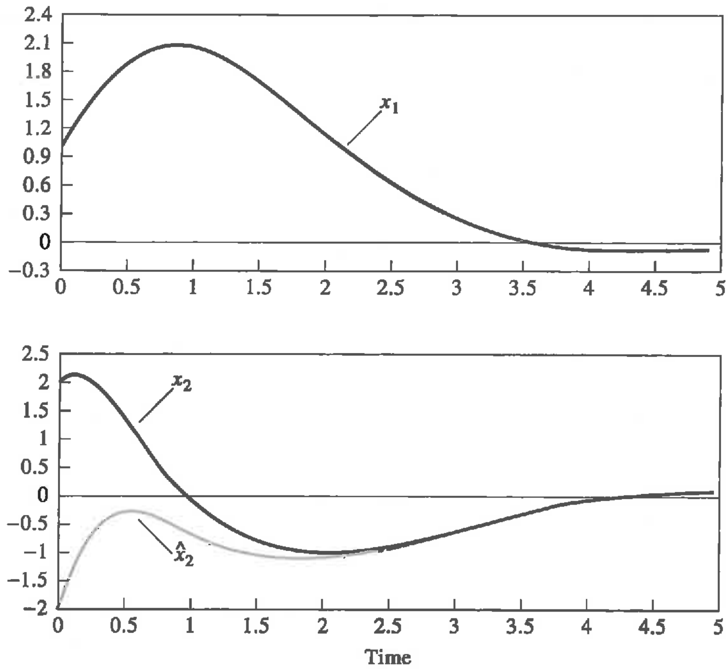


Figure 9.19 Using a reduced-order observer to obtain zero-input response of the double-integrator system.

arbitrary matrix, T, such that when it is stacked on top of C, the result is a nonsingular matrix—that is,

$$P = \begin{bmatrix} C \\ T \end{bmatrix} \text{ is nonsingular}$$

The inverse of this matrix, called Q, is our sought-after transformation. This is because P and Q are inverses of each other, and therefore satisfy

$$PQ = I \rightarrow \begin{bmatrix} C \\ T \end{bmatrix} [Q_1 \quad Q_2] = \begin{bmatrix} CQ_1 & CQ_2 \\ TQ_1 & TQ_2 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

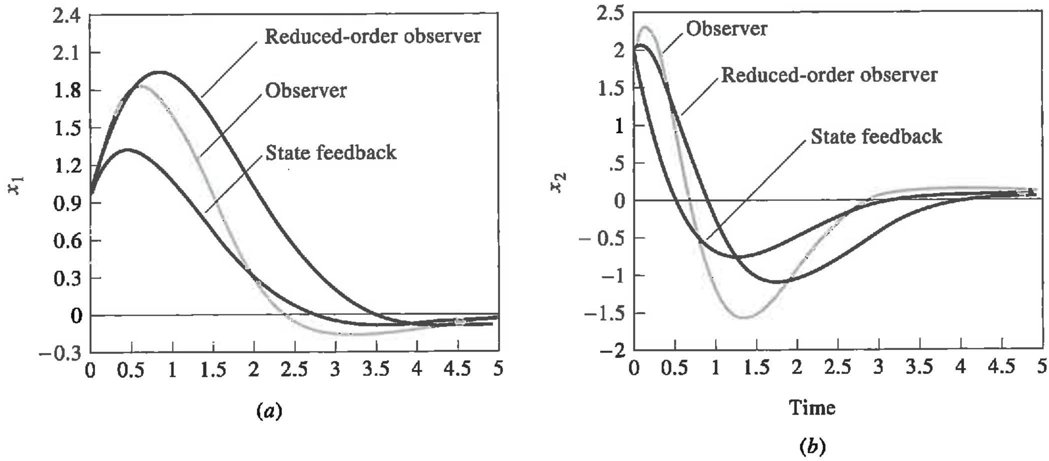


Figure 9.20 Comparison of the zero-input responses of the double-integrator system under state feedback, full-order observer, and reduced-order observer.

where the identity and zero matrices have compatible dimensions. Recall from Chapter 8 that under a linear transformation, the new C matrix becomes

$$\bar{C} = CQ = C[Q_1 \quad Q_2] = [CQ_1 \quad CQ_2] = [I \quad 0]$$

As an example, consider the following equivalent model for the double-integrator system (obtained by labeling the states from left to right):

$$\dot{x} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$y = [0 \quad 1]\, x$$

Choosing T as shown below results in

$$T = [1 \quad 0] \rightarrow P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \rightarrow Q = P^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \rightarrow \bar{C} = CQ = [1 \quad 0]$$

$\bar{A}$ and $\bar{B}$ will turn out to be the same matrices as in the original example (this is just a coincidence in this simple example, usually the new matrices will be quite different, but the C matrix will have the required form).

❑ DRILL PROBLEM

D9.9 Consider the third-order system given by

$$G(s) = \frac{3}{s(s^2 + 4s + 5)}$$

or, in state-space form,

$$\dot{x} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -5 & -4 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} u$$

Suppose the states x_1 and x_2 are measured but the state x_3 is to be reconstructed using a first-order observer.

(a) Find the observer gain and parameters. The observer pole is to be at -20. Also write down the observer equation.

(b) Find the controller gain to place poles at -10 and $-2 \pm j2$.

(c) Find the transfer function of the compensator (note that the compensator is a two-input, one-output system).

Ans. (a) $D = -20,\ G = 3,\ L = [0 \quad 16],\ F = [0 \quad -325]$

$$\dot{z} = -20z + [0 \quad -325]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + 3u = -20z - 325x_2 + 3u$$

$$\hat{x}_3 = z + [0 \quad 16]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = z + 16x_2$$

(b) $k = [80 \quad 43 \quad 10]\frac{1}{3}$

(c) $H(s) = \dfrac{1}{s+30}[-26.66s - 533.33,\ 30s + 2660]$

9.6 A Magnetic Levitation System

Beginning in 1969, West Germany sought to develop a high-speed electric train system to span central Europe. State space analysis and aircraft technology were used to design, build, and test such a train for operation at speeds as high as 400 km/h (248 mi/h).The train is suspended in midair by magnetic fields. This type of suspension is called magnetic levitation or MAGLEV.

Figure 9.21, shows the cross section of a MAGLEV vehicle. The track is a T-shaped concrete guideway. Once under way, the train does not touch the guideway, resulting in greatly reduced friction and reduced guideway construction costs. Electromagnets are distributed along the guideway and along the length of the train in matched pairs. The magnetic attraction of the vertically paired magnets balances the force of gravity and levitates the vehicle above the guideway. The horizontally paired magnets stabilize the vehicle against sideways forces. Forward propulsion is produced by linear induction motor action between train and guideway. Only the vertical motion and control of the suspended vehicle will be considered here.

The equations characterizing the train's vertical motion are now developed. It is desired to control the gap distance d within a close tolerance in normal operation of the train. The gap distance d between the track and the train magnets is

$$d = z - h$$

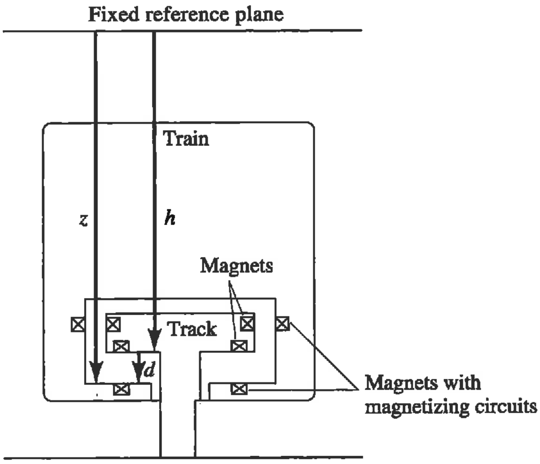


Figure 9.21 Cross section of a MAGLEV train.

Then

$$\dot{d} = \dot{z} - \dot{h}$$
$$\ddot{d} = \ddot{z} - \ddot{h}$$

where the dots denote time derivatives. The magnet produces a force that is dependent upon residual magnetism and upon the current passing through the magnetizing circuit. For small changes in the magnetizing current i and the gap distance d, that force is approximately

$$f_1 = -Gi + Hd$$

where G and H are positive constants. That force acts to accelerate the mass M of the train in a vertical direction, so

$$f_1 = M\ddot{z} = -Gi + Hd$$

For increased current, the distance z diminishes, reducing d as the vehicle is attracted to the guideway.

A network model for the magnetizing circuit is given in Figure 9.22. This circuit represents a generator driving a coil wrapped around the magnet on the vehicle. The voltage induced in the coil by the vehicle motion is represented by the term $(LH/G)\dot{d}$, for which it is assumed that the magnetic flux loss is negligible. For that circuit

$$Ri + L\dot{i} - \frac{LH}{G}\dot{d} = v$$

The three state variables

$$x_1 = d$$
$$x_2 = \dot{d}$$
$$x_3 = i$$

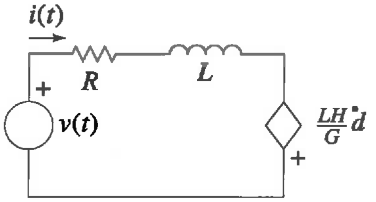


Figure 9.22 Magnetizing circuit model.

are convenient, and in terms of them the vertical motion state equations are

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ \frac{H}{M} & 0 & -\frac{G}{M} \\ 0 & \frac{H}{G} & -\frac{R}{L} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & -1 \\ \frac{1}{L} & 0 \end{bmatrix} \begin{bmatrix} v \\ f \end{bmatrix}$$

where

$$f = \ddot{h}$$

If the gap distance d is considered to be the system output, then the state variable output equation is

$$d = x_1$$

The voltage v is considered to be a control input, while guideway irregularities $f = \ddot{h}$ constitute a disturbance. Figure 9.23 shows block diagram and signal flow graph representations of the state equations.

The characteristic equation for the system, the roots of which are the transfer function poles, is given by

$$|sI - A| = \begin{vmatrix} s & -1 & 0 \\ -\frac{H}{M} & s & \frac{G}{M} \\ 0 & -\frac{H}{G} & s + \frac{R}{L} \end{vmatrix} = 0$$

$$= s \begin{vmatrix} s & \frac{G}{M} \\ -\frac{H}{G} & s + \frac{R}{L} \end{vmatrix} + \begin{vmatrix} -\frac{H}{M} & \frac{G}{M} \\ 0 & s + \frac{R}{L} \end{vmatrix}$$

$$= s\left(s^2 + \frac{R}{L}s + \frac{H}{M}\right) - \frac{H}{M}\left(s + \frac{R}{L}\right)$$

$$= s^3 + \frac{R}{L}s^2 - \frac{HR}{ML} = 0$$

The system is thus unstable, since its characteristic polynomial has coefficients with differing algebraic signs. Also, the coefficient of s in the characteristic equation is zero. The system instability is quite understandable when one considers the action

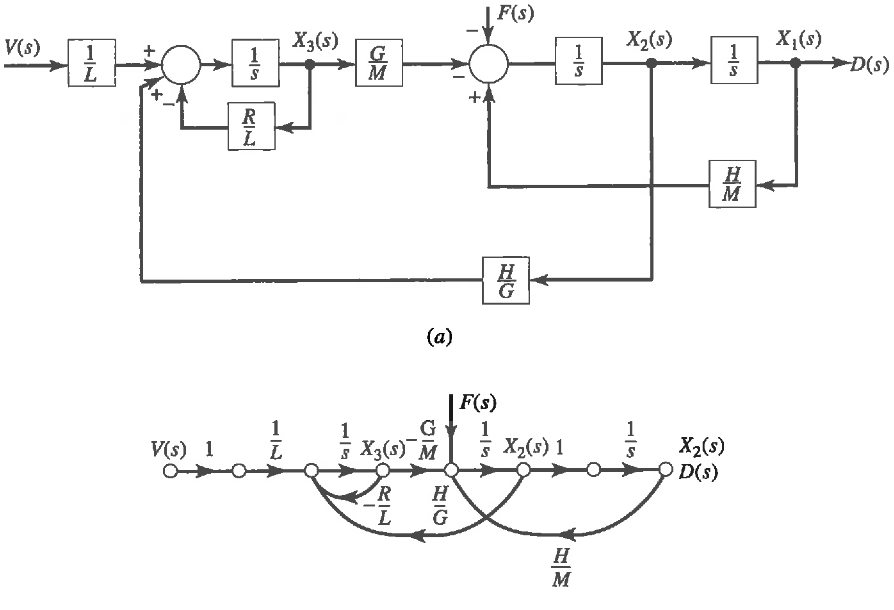


Figure 9.23 Diagrams of the state equations. (a) Block diagram. (b) Signal flow graph.

of the magnets. If the gap distance d should increase slightly, the magnetic attraction will decrease, tending to further increase the gap, and so on.

To control the system, the magnetizing circuit voltage is chosen to be a linear combination of the state signals plus a tracking input $u_1(t)$:

$$v = k_1x_1 + k_2x_2 + k_3x_3 + u_1(t)$$

The feedback signals are produced from sensors that monitor the state variables, namely, gap distance d, gap velocity $\dot{d}$, and magnetizing current i. The resulting feedback system is described by

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ \dfrac{H}{M} & 0 & -\dfrac{G}{M} \\ \dfrac{k_1}{L} & \dfrac{H}{G} + \dfrac{k_2}{L} & -\dfrac{R}{L} + \dfrac{k_3}{L} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & -1 \\ \dfrac{1}{L} & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ f \end{bmatrix}$$

$$d = x_1$$

Appropriate choice of the feedback gain constants k_1, k_2, and k_3, that is, the feedback gain matrix,

$$K = [k_1 \quad k_2 \quad k_3]$$

will place the system poles at any desired locations.

To proceed with state-variable design methods, the parameters M, G, L, and R must be estimated. The following values do not necessarily represent those of any specific existing system, but the methods and values are representative of the design process in general.

Suppose an engineer finds that each train car weighs 8000 kg. Each car is supported by four magnets, each of which must therefore support 2000 kg. Each subsystem can be analyzed using $M = 2000$ kg.

A static test is performed without control. The air gap is clamped shut, causing d to be zero. A -120 V source is applied to the magnetizing circuit. With a time constant of $\frac{1}{30}$ s, -8 A eventually flows at steady state. A resultant force of 4000 N is measured (in addition to that of gravity).

The static test is concluded and the voltage is carefully varied until, at equilibrium, the car levitates with $d = 10$ mm under the influence of 8 A of current.

If the magnetizing circuit is at steady state, the static test can be used to get R and L, since

$$R = \frac{v}{i} = \frac{-120}{-8} = 15\ \Omega$$

and, from the time constant during the static test

$$T = \frac{L}{R} \qquad \text{so} \qquad L = RT = \frac{15}{30} = 0.5H$$

The data from when the air gap was clamped shut ($d = 0$) permit G to be computed:

$$f = -Gi + H \times 0$$

$$G = \frac{-f}{i} = \frac{-4000}{-8} = 500\ \text{N/A}$$

The data from when the car was levitated to equilibrium provides H:

$$0 = -500 \times 8 + H \times 10$$

$$H = 400\ \text{N/mm}$$

The parameter values are, therefore,

$$M = 2000 \qquad H = 400$$

$$G = 500 \qquad L = 0.5$$

$$R = 15$$

For these, the feedback system equations are

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0.2 & 0 & -0.25 \\ 2k_1 & 0.8 + 2k_2 & -30 + 2k_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & -1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ f \end{bmatrix}$$

$$d = x_1$$

The characteristic equation for the feedback system is given by

$$\begin{vmatrix} s & -1 & 0 \\ -0.2 & s & 0.25 \\ -2k_1 & -0.8 - 2k_2 & s + 30 - 2k_3 \end{vmatrix}$$

$$= s\begin{vmatrix} s & 0.25 \\ -0.8-2k_2 & s+30-2k_3 \end{vmatrix} + \begin{vmatrix} -0.2 & 0.25 \\ -2k_1 & s+30-2k_3 \end{vmatrix}$$

$$= s^3 + (30-2k_3)s^2 + (0.5k_2)s + 0.4k_3 + 0.5k_1 - 6$$

The feedback gains k_1, k_2, and k_3 may be chosen to give any desired coefficients of the characteristic equation of the feedback system. For example, if it is desired to have the system poles at $s = -1 + j2,\ -1 - j2,$ and -3, the characteristic polynomial would be

$$(s+1-j2)(s+1+j2)(s+3) = s^3 + 5s^2 + 11s + 15$$
$$= s^3 + c_2 s^2 + c_1 s + c_0$$

which is achieved with

$$k_3 = 0.5(30 - c_2) = 12.5$$
$$k_2 = 2c_1 = 22$$
$$k_1 = 2(c_0 + 0.2c_2) = 32$$

For this choice of feedback gains, the feedback system model is

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0.2 & 0 & -0.25 \\ 64 & 44.8 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & -1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ f \end{bmatrix}$$

$$d = [1 \quad 0 \quad 0] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

The steady state output d due to a unit step disturbance input f is given by

$$0 = [A + BK]x + Bu$$
$$d = x_1 \qquad u_1 = 0 \qquad f = 1$$

So that

$$[A + BK]x = -Bu = -\begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Cramer's rule can be used to obtain $d = x_1$. It is instructive to write the gain values in terms of the desired characteristic polynomial coefficients.

$$d = x_1 = \frac{\begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & -0.25 \\ 0 & 4c_1 + 0.8 & -c_2 \end{vmatrix}}{\begin{vmatrix} 0 & 1 & 0 \\ 0.2 & 0 & -0.25 \\ 4(c_0 + 0.2c_2) & 4c_1 + 0.8 & -c_2 \end{vmatrix}} = \frac{c_2}{-c_0} = -\frac{1}{3}$$

which depends only on the desired performance. Further study would be needed to determine whether this amount of disturbance rejection from track irregularities is sufficient. The negative algebraic sign simply means that a positive step in $f = \ddot{h}$ results in a steady state decrease in the gap distance. Types of disturbance other than constant ones should also be considered in the design.

The reference input u_1 would normally be a constant that sets the nominal gap distance. The steady state gap distance d due to a constant reference input u_1 where f is zero (level track) is given by

$$0 = [A + BK]x + Bu \qquad u = \begin{bmatrix} u_1 \\ 0 \end{bmatrix}$$

$$[A + BK]x = -Bu = \begin{bmatrix} 0 \\ 0 \\ -2u_1 \end{bmatrix}$$

$$d = x_1$$

Again, Cramer's rule can be used to obtain $d = x_1$ where it is instructive to write the gain values in terms of the desired characteristic polynomial coefficients

$$d = x_1 = \frac{\begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & -0.25 \\ -2u_1 & 4c_1 + 0.8 & -c_2 \end{vmatrix}}{-c_0}$$

$$= -\frac{0.5u_1}{c_0}$$

For a nominal gap of 10 mm, the reference input should be

$$u_1 = -(20)c_0 = -300$$

which depends only on the nominal gap and on a coefficient of the desired characteristic polynomial.

Figure 9.24 shows calculated system response where the train accelerates from a standstill and traverses an irregular guideway with a downgrade followed by an upgrade. In the German system, the nominal air gap distance is 14 mm (about $\frac{1}{2}$ in.). Improvement in disturbance rejection is obtained by modeling the track irregularities by differential equations that are included as part of an observer. Three levels of complexity are used depending on whether the track is level (actually somewhat curved between towers), following a hill, or following a curve.

Space does not permit a complete discussion of the system; however, one feature is of interest. The rate of change of the air gap is also estimated by using an observer. The state vector may be reordered as follows:

$$x = \begin{bmatrix} d \\ i \\ \cdots \\ \dot{d} \end{bmatrix} = \begin{bmatrix} y \\ x_u \end{bmatrix}$$

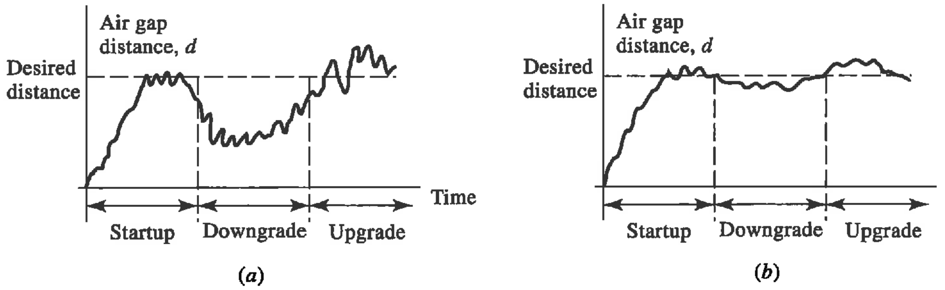


Figure 9.24 MAGLEV system response. (a) Response of the system with state feedback. (b) Improved response with disturbance modeling and feedback.

because d and i can be measured and the rate of change of d must be estimated by a reduced-order, first-order observer. The concepts of control and estimation can be separated (as mentioned earlier in this chapter); therefore, the open-loop system matrix is used to compute the observer dynamics. As a result of reordering, we have

$$A = \left[\begin{array}{cc|c} 0 & 0 & 1 \\ 0 & -30 & 0.8 \\ \hline 0.2 & -0.25 & 0 \end{array}\right] \qquad B = \left[\begin{array}{c} 0 \\ 2 \\ \hline 0 \end{array}\right]$$

The reordering is the result of augmenting the C matrix by T and transforming the state equations.

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{choose } T \text{ as } T = [0 \quad 1 \quad 0] \text{ then}$$

$$P = \begin{bmatrix} C \\ T \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad Q = P^{-1} \quad \text{and} \quad \bar{C} = CQ = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Suppose the observer pole is selected at -10. Following the given design procedures, we get

$$A_{22} - LA_{12} = 0 - \begin{bmatrix} l_1 & l_2 \end{bmatrix} \begin{bmatrix} 1 \\ 0.8 \end{bmatrix} = -l_1 - 0.8 l_2 = -10$$

This results in one equation and two unknowns. We can set l_1 to zero so that $l_2 = 12.5$. Computing the remaining observer parameters, we get

$$F = \begin{bmatrix} 0.2 & -249.25 \end{bmatrix} \qquad G = -25 \qquad D = -10$$

Therefore, the observer equation is given by

$$\dot{z} = -10z + 0.2x_1 + 249.75x_2 - 25u$$

$$\hat{x}_3 = z + 12.5x_2$$

Figure 9.24(b) shows the improved performance when an observer estimates track motion. The closed-loop system also uses the observer estimate of d for feedback control.

❑ Computer-Aided Learning

To find the state feedback gain we use the "place" command with the following syntax:

```
K=place(A,B,DP)
```

where DP stands for the vector of desired pole locations. The only restriction is that desired poles must be distinct (not repeated). The "place" command also works in the multiinput case.

For example, let us place the poles of the following system in $-1 \pm j$.

$$\dot{x} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} x + \begin{pmatrix} -1 \\ 6 \end{pmatrix} u$$

$$y = (0 \quad -14)x + 8u$$

We enter the following commands:

```
a=[1,2;3,4];b=[-1;6];c=[0,-14];d=8;
g=ss(a,b,c,d);
dp=[-1+j, -1-j];
k=place(a,b,dp)
place: ndigits=15
k=
  1.0    1.3333
```

We can verify our work by finding the closed-loop poles by

```
eig(a-b*k)
ans=
  -1.0000+1.0000i
  -1.0000-1.0000i
```

MATLAB also has an implementation of the Ackerman formula under the "acker" command. This command only works for single-input systems but does not require that the poles be distinct. Here is an example:

```
k=acker(a,b,dp)
k=
  1.0000  1.3333
```

We can place both poles at $-1, -1$:

```
dp2=[-1,-1];
k2=acker(a,b,dp2)
k2=
  0.9310  1.3218
```

C9.1 Redo Drill Problems D9.1–9.3 using the "place" and "acker" commands.

Observer Design

Because of the duality that exists between control and observer problems, the same commands can be used for observer design with modified inputs as shown next.

For full-order observer, use

```
L=place(A',C',ODP)'
```

For reduced-order observers, use

```
Lr=place(A22',A12',ODP)'
```

Where ODP are the observer desired poles. Note that in the single-output cases we can use the "acker" command if the observer poles are repeated.

We will design a full-order observer with poles at $-5 \pm j5$ for the system just described.

```
odp=5*[-1+j, -1-j]; l=place(a',c',odp)'
Place: ndigits=15
l=
  -1.5952
  -1.0714
```

C9.2 (a) Use MATLAB to solve Drill Problem 9.5.

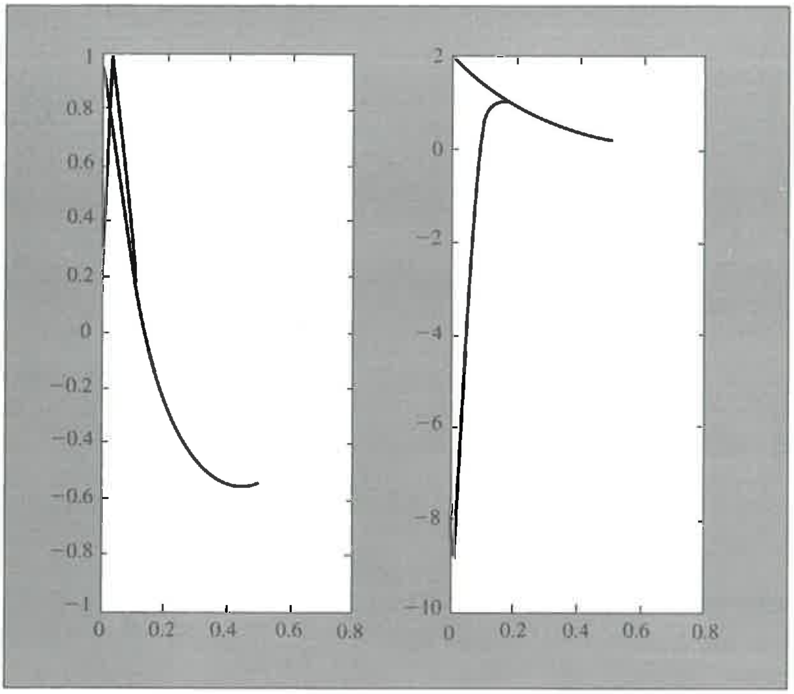


Figure C9.2

Ans. (a)
```
a=[-2,-4;1-4]; b=[2;0]; c=[1 0];
odp=[-50 -50]; l=acker(a',c',odp)'
l=
   94
  -528
ac=[a zeros (2, 2); l*c a-l*c] ;
bc=zeros (4,1); cc=eye (4); d=zeros (4,1);
[y,x,t]=initial (ac, bc, cc, zeros (4,1), [1 2 -1 -2], .5);
subplot (1, 2, 1), plot (t, y(:, 1), t, y (:,3)),
subplot (1, 2, 2), plot (t, y(:, 2), t, y(:, 4))
```

(b) Use MATLAB to solve Drill Problem 9.6, verify closed-loop stability.

(c) Use MATLAB to solve Drill Problem D9.7.

(d) Use MATLAB to solve Drill Problem D9.9.

9.7 SUMMARY

The techniques discussed in this chapter are very powerful and have expanded the range of problems that can be solved. If all the states are available, the advantages of state feedback become apparent. All closed-loop poles can be placed at desired locations in the complex plane as long as the system is controllable. If the system is not controllable, we can still stabilize the system as long as the system is stabilizable (i.e., the unstable modes are controllable). If some of the states are not available for measurement, for technological or economic reasons, observers can be implemented to estimate the states. Full-order or identity observer has a simple structure; its structure is the copy of the system plus a correction term, multiplied by observer gains. The observer gains can be computed to place the observer poles anywhere in the complex plane as long as the system is observable. Reduced-order observers reconstruct only the states that are not measured. For linear systems, there exists a seperation between the control and the estimation problems. This means that the poles of the closed-loop system (the interconnection of the controlled plant and the observer) are the union of the poles of the controlled plant and the observer.

Integral control, which allows us to eliminate steady state errors to constant inputs can be designed by using state space methods. This is done by augmenting the plant model by an integrator. Although not discussed here, it is also possible to track general command inputs using the methods discussed in this chapter. A major limitation of observer-based controllers is the lack of guaranteed stability margins. The methods rely heavily on the plant models. Because we rarely have accurate models of our plants, adequate stability margins are required to protect against these model uncertainties. In some situations we may design a controller that works perfectly under computer simulations but turns out to be unstable in practice. Therefore, any design must be tested thoroughly to prevent disastrous results.

The design of a magnetically levitated train exemplifies state space representation and controller and observer design.

REFERENCES

State Feedback

Anderson, B. D. O., and Moore, J. B., *Optimal Control*. Englewood Cliffs, NJ: Prentice-Hall, 1990.

Davison, E. J., "On Pole Assignment in Multivariable Linear Systems." *IEEE Trans. Autom. Control* AC-13 (December 1968): 747–748.

Wonham, W. M., "On Pole Assignment in Multi-Input Controllable Linear Systems." *IEEE Trans. Autom. Control* AC-12 (December 1967): 660–665.

Observers

Doyle. J. C., and Stein, G., "Robustness with Observers." *IEEE Trans. Autom. Control* (August 1979): 607–611.

Krogh, B., and Cruz. J. B., "Design of Sensitivity-Reducing Compensators Using Observers." *IEEE Trans. Autom. Control* (December 1978): 1058–1062.

Luenberger, D. G., "Observers for Multivariable Systems." *IEEE Trans. Autom. Control* AC-11 (April 1966): 190–197.

———, "An Introduction to Observers." *IEEE Trans. Autom. Control* AC-16 (December 1971): 596–602.

Nuyan, S., and Carroll, R. L., "Minimum Order Arbitrarily Fast Adaptive Observers and Identifiers." *IEEE Trans. Autom. Control* (April 1979): 289–297.

Sage, A. P., and White, C. C.,*Optimum Systems Control*. Englewood Cliffs. NJ: Prentice-Hall, 1977.

Stefani, R. T., "Reducing the Sensitivity to Parameter Variations of a Minimum-Order Reduced-Order Observer" *Int. J. Control* (1982): 983–995.

Stefani, R. T., "Observer Steady State Errors Induced by Errors in Realization." *IEEE Trans. Autom. Control* (April 1976): 280–282.

Magnetic Levitation of Trains

Brock, K. H., Gottzein, E., Pfefferl, J., and Schneider, E., "Control Aspects of a Tracked Magnetic Levitation High Speed Test Vehicle." *Automatica.* vol. 13. no. 3, 1977, pp. 205–233.

Glatzel, K., Khurdok, G., and Rogg, D., "The Development of the Magnetically Suspended Transportation System in the Federal Republic of Germany." *IEEE Trans. Vehic. Technol.* (February 1980): 3–17.

Glatzel, K., and Schulz, H., "Transportation: The Promise of MAGLEV." *IEEE Spectrum* (March 1980): 63–66.

Gottzein, E., Meisinger, R., and Miller, L., " The Magnetic Wheel in the Suspension of High Speed Ground Transportation Vehicles." *IEEE Trans. Vehic. Technol.* (February 1980): 17–22.

Kaplan, G., "Rail Transportation." *IEEE Spectrum* (January 1984): 82–85.

———. " Transportation." *IEEE Spectrum* (January 1985): 81–84.

PROBLEMS

1. For the state feedback systems described by the following equations, choose the feedback gain constants k_i to place the closed-loop system poles at the indicated locations. Then, for the feedback system, find the steady state outputs due to a unit step input.

(a)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -8 & -3 & -7 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} u$$

$$u = -\begin{bmatrix} k_1 & k_2 & k_3 & k_4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + r$$

$$y = \begin{bmatrix} 2 & -1 & 0 & 3 \\ 1 & 0 & 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

Closed-loop poles at $s = -5 \pm j3, \ -4 \pm j4$

(b)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -10 & -5 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 0 & 0 \\ 1 & 7 \end{bmatrix} \begin{bmatrix} u \\ r \end{bmatrix}$$

$$u = -\begin{bmatrix} k_1 & k_2 & k_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$y = [1 \quad 0 \quad 0] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Closed-loop poles at $s = -4$ and $-4 \pm j2$

Ans. $k_1 = +70, \ k_2 = +47, \ k_3 = +10, \ y(\infty) = -0.5625$

2. Design first-order observers of the following plants. Choose the observer eigenvalues to be at $s = -30$.

(a)
$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -2 & 1 & 0 \\ -4 & 0 & 1 \\ -2 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} u$$
$$y = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

(b)
$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -4 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 3 \end{bmatrix} u$$
$$y = x_1$$

3. Design identity observers for the following plants. Choose the observer eigenvalues to be at -20.

(a)
$$A = \begin{bmatrix} 0 & 1 \\ -4 & -4 \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$
$$y = x_1$$

(b)
$$A = \begin{bmatrix} 0 & 1 \\ -3 & -8 \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$
$$y = x_1$$

4. For the systems of Problem 3, design control gains to place the desired closed-loop poles at -5 and -8, assuming the measurements are available. Next, close the loop by using a reduced-order observer to furnish an estimate of x_2. Show that the characteristic polynomial of the closed-loop system including the observer contains the desired closed-loop roots and the observer root.

5. Consider the system given by

$$\dot{x} = \begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix} x + \begin{bmatrix} 1 \\ -1 \end{bmatrix} u$$
$$y = \begin{bmatrix} c_1 & c_2 \end{bmatrix} x$$

(a) Determine whether the system is controllable.

(b) Determine whether the system is stabilizable.

(c) Find the transfer function, $T(s)$, if $c_1 = c_2 = 1$.

(d) Repeat (c) for $c_1 = 1,\ c_2 = -1$.

(e) Can state feedback be used to stabilize the system?

Ans. (a) uncontrollable; (b) not stabilizable; (c) $T(s) = 0$; (d) $T(s) = 2(s-1)/[(s-1)(s-2)] = 2/(s-2)$ (e) no.

6. Consider the system given by

$$\dot{x} = \begin{bmatrix} 0 & 1 & -1 \\ -2 & -3 & 0 \\ p & 1 & 1 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u$$

$$y = [0 \quad 0 \quad 1]x$$

(a) Determine the values of the parameter p for which the system is controllable (observable).

(b) Find the transfer function of the system.

(c) Determine for what values of the parameter p the system is stabilizable (detectable).

7. Consider the system

$$\dot{x}_2 = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ -2 & 0 & 0 \end{bmatrix} x + \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} u$$

$$y = [1 \quad 0 \quad 0]x$$

(a) Determine its controllability and observability.

(b) Diagonalize the system; that is, find $\{\bar{A},\ \bar{B},\ \bar{C}\}$.

(c) Identify the modes that are either uncontrollable or unobservable.

(d) Find the system transfer function.

(e) Determine the stabilizability and detectability of the system.

8. Consider the linear system

$$\dot{x} = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 0 \\ 1 & -4 & 3 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u$$

$$y = [1 \quad -1 \quad 1]x$$

Determine the controllability, observability, stabilizability, and detectability of each mode. Also find the system transfer function and note any pole–zero cancellations.

9. Consider the following plant and answer the following questions.

$$\dot{x} = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} c_1 & c_2 \end{bmatrix} x$$

(a) Is the system controllable?

(b) Is the system observable?

(c) Find a control gain vector, k, to place plant poles at $\{-1, \quad -2\}$, if possible.

(d) Repeat (c) for poles at $\{-2, \quad -2\}$.

(e) Explain any discrepancy in answers to (c) and (d).

(f) Find the transfer function when $c_1 = 1, \ c_2 = 0$.

(g) Repeat (f) for $c_1 = 0, \ c_2 = 1$.

(h) Given the answers in (f) and (g), which state variable should be measured to stabilize the system using observers?

10. Consider the plant, $G(s)$.

$$G(s) = \frac{1}{s(s-2)}$$

Use state feedback to move the closed-loop poles to $s = -1, \ -1$.

(a) Find control gain, k.

(b) Design a reduced-order observer with pole at $s = -2$.

(c) Find the open-loop transfer function of the compensated system and use it to plot the root locus.

11. Design a first-order observer for the following system. Place the observer pole at -1. Also design a controller to place system poles at $-1 \pm j$. Obtain the compensator transfer function and draw the root locus of the open-loop transfer function.

Use computer software to simulate the system. For simulation purposes, the plant initial conditions are 1 and 2. Set the initial condition for $\hat{x}_2$ at -2.

$$\dot{x} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u, \qquad x(0) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$y = [1 \quad 0]x$$

12. It was mentioned in this chapter that an optimal choice for observer initial conditions is given by $\hat{x}(0) = C'(CC')^{-1}y(0)$. Repeat Problem 11 by designing a full-order observer with poles at $-2 \pm 2j$. Simulate the closed-loop system using observer initial conditions of $\hat{x}(0) = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$. Then, to repeat the simulation, use optimal initial conditions [note that $y(0) = 1$]. Compare the zero-input responses in both cases.

13. Consider the following plant: $\dot{x} = Ax + Bu$, where the state space matrices are given by

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

(a) We want to place the closed-loop poles at $\{-10,\ -1+j,\ -1-j\}$. Find the state feedback gain vector.

(b) Obtain the equivalent transfer function of the compensator, root locus, Bode plots, and closed-loop step response. Tabulate the step response and frequency response features such as percent overshoot, rise time, settling time, and phase and gain margins.

(c) Design a full-order observer. Choose observer poles at $\{-40, -4+j4, -4-j4\}$. Repeat (b).

(d) Design a reduced-order observer. Choose observer poles at $\{-4+j4, -4-j4\}$. Repeat (b).

(e) Repeat (c) with the observer poles at $\{-40,\ -1,\ -2\}$.

(f) Repeat (d) with the observer poles at $\{-1,\ -2\}$

Note: In (e) and (f), the observer poles are chosen at the plant zeros. It is known that such a choice increases the robustness of the system. Because phase and gain margins are classical measures of robustness (protection against uncertainty), compare the margins in all cases. Does the choice of observer poles in (e) and (f) really improve the margins?

(g) To verify robustness, let us assume that the A matrix of the true plant model is

$$A_{\text{true}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & 0 & -1 \end{bmatrix}$$

Use the control gain vector of (a) to obtain the step responses and frequency responses for the true system, using the observers designed in (c)–(f). Compare the responses, and determine which observer design is more robust to parameter uncertainty and variation.

14. For the MAGLEV system choose instead values of the feedback gain constants k_1, k_2, and k_3 to place all three of the overall system poles at $s = -5$. For this system, find the steady state response d to a unit step disturbance f and the value of constant reference input u_1 to give a nominal gap distance $d = 15$.

15. For the open-loop MAGLEV system, suppose the vertical track elevation varies sinusoidally with time as the train is in motion, according to

$$\dot{h}(t) = 0.2 \sin \frac{\pi t}{10}$$

Find the second-order differential equation satisfied by $\dot{h}(t)$, then augment the original state equations with two more equations and two more state variables

$$x_4 = \dot{h}(t)$$

$$x_5 = \ddot{h}(t)$$

in place of the disturbance input $f = \ddot{h}$. With additional sensors for the signals x_4 and x_5 and feedback of the form

$$u = +k_1x_1 + k_2x_2 + k_3x_3 + k_4x_4 + k_5x_5 + u_1$$

find the state equations for the feedback system in terms of the k constants.

CHAPTER 10

Advanced State Space Methods

10.1 Preview

In the preceding chapters, compensators were designed to satisfy specified requirements for steady state error, transient response, stability margins, or closed-loop pole locations. Meeting all objectives is usually difficult because of various tradeoffs that must be made and because of the limitations of the design techniques. For example, classical Bode design allows us to satisfy phase margin and steady state error requirements, but the step response characteristics may not be desirable. State space observer-based techniques allow arbitrary pole placement, but the stability margins cannot be controlled directly. Also, none of the techniques discussed so far address practical issues such as plant model uncertainty or actuator signal limits. In addition, none of the techniques result in the best possible performance. This chapter addresses some of these issues. In particular, we present optimization-based techniques that result in an optimal solution.

Optimization refers to the science of maximizing or minimizing objectives. Optimization requires a measure of performance. When mathematically formulated, this measure of performance is called the *objective* (or *cost*) *function*. Optimization of control systems is called *optimal control*. Optimization problems are either constrained or unconstrained. For example, finding the minimum of a parabola is an unconstrained optimization problem. Finding the minimum of a parabola in a given interval of its domain is a constrained optimization problem. You have seen examples of these types in calculus. Calculus-type problems are usually static problems because the constraints are algebraic equations. Optimal control problems are usually constrained dynamic optimization problems because the constraints are the system equations, which are differential equations (i.e., they are dynamic). Simple examples of optimal control appear in Section 3.5, where a gain or damping ratio is found to minimize the tracking error in a control system. In this chapter a more general and systematic treatment of optimal control is presented.

A typical optimal control problem formulation is the following:

$$\min_u J = \int_0^T L(x, u, t)dt \qquad \text{subject to} \qquad \dot{x} = f(x, u, t)$$

Here the plant is a nonlinear system, which is the constraint, and the cost function is the integral of some nonlinear function of the state x, control u, and time. The objective is to find a control function u that will minimize the cost function. Examples of optimal control problems are minimum time, minimum fuel, and minimum energy (more examples appear in Section 3.5). This formulation, in general, leads to controllers that are time-varying and nonlinear. Analog implementation of these nonlinear controllers is not usually practical or worthwhile.

10.2 The Linear Quadratic Regulator Problem

We will now restrict our attention to linear systems (or linearized versions of non-linear systems) and choose a cost function that is a quadratic function of states and controls. In this case, the solution is a linear controller that is easily implemented. Hence, we will consider a special optimal control problem, called the *linear quadratic regulator* (*LQR*) problem. The formulation of the problem follows. Given the linear system

$$\boxed{\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx \end{aligned}}$$

find a control function *u(t)* that will minimize the cost function J given by

$$\boxed{J = \frac{1}{2}\int_o^\infty \left(x'Qx + u'Ru\right) dt}$$

The function inside the integral is a quadratic form and the matrices Q and R are usually symmetric (see Appendix A for a brief review of quadratic forms). It is assumed that R is positive definite (i.e., it is symmetric and has positive eigenvalues) and Q is positive semi definite (i.e., it is symmetric and its eigenvalues are nonnegative). These assumptions imply that the cost is nonnegative, so its minimum value is zero. For the cost function to achieve its minimum value, both x and u must go to zero. This type of control problem is called a *regular problem.* When the state vector is to track nonzero values, J can be redefined to create an optimal *servomechanism* (tracking) problem. Many of the control systems considered in earlier chapters were servomechanisms. Regulator behavior is important for control systems of many types (e.g., attitude control of satellites or spacecraft, where a zero reference should be maintained in spite of disturbances).

A simple interpretation of the cost function is as follows. If the system is scalar (i.e., a first-order system), the cost function becomes

Scalar LQR problem.

$$J = \frac{1}{2}\int_0^\infty \left(qx^2 + ru^2\right) dt$$

Now we see that J represents the weighted sum of energy of the state and control. Small q and r are used, respectively, when x and u are scalars. If r is very large relative to q, which implies that the control energy is penalized heavily, the control effort will diminish at the expense of larger values for the state. This physically translates into smaller motors, actuators, and amplifier gains needed to implement the control law. Likewise if q is much larger than r, which means that the state is penalized heavily, the control effort rises to reduce the state, resulting in a damped system. In the general case, Q and R represent respective weights on different states and control channels. For example, if

$$Q = \begin{bmatrix} 10 & 0 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad R = r \text{ (a scalar)}$$

Then

$$x'Qx + u'Ru = 10x_1^2 + x_2^2 + ru^2$$

By putting a larger weight on the first state, we are putting more emphasis on controlling this state and restricting its fluctuations.

Several procedures are available to solve the LQR problem. Since optimization can easily become the subject for several textbooks, we will present only the main results. The work of mathematicians and engineers such as Hamilton, Euler, Lagrange, Jacobi, Pontryagin, Kalman, and Bellman have resulted in a rather complete understanding of the optimal control problem. Actual implemented control systems that have been designed by these methods were few in number as of the early 1990s, but they are now more popular.

One approach to finding a controller that minimizes the LQR cost function is based on finding the positive-definite solution of the following *algebraic Riccati equation* (ARE).

LQR Solution.

$$\boxed{\begin{array}{l} A'P - PA + Q - PBR^{-1}B'P = 0 \\ u = -Kx \qquad K = R^{-1}B'P \end{array}}$$

It turns out that under the conditions stated shortly, the positive-definite solution of the ARE results in an asymptotically stable closed-loop system. The conditions are the following. The system is controllable, R is positive definite (this ensures that its inverse exists), and Q can be factored as $Q = C_q'C_q$, where C_q is any matrix such that (C_q, A) is observable. These conditions are necessary and sufficient for the existence and uniqueness of the optimal controller that will asymptotically stabilize the system (these assumptions can be relaxed to stabilizability and detectability). Note that the assumption on Q allows us to define another output vector z as

$$z = C_q x$$

therefore

$$x'Qx = x'C_q'C_qx = z'z$$

the vector z is called the *controlled* or *regulated output*, and may differ from the *measured output* y.

Manually solving the Riccati equation is tedious and almost impossible for third- or higher-order systems; the second-order example, however, can be solved. Let us consider the double-integrator example considered in Chapter 9. The system matrices are

Solving the double integrator problem by LQR.

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad C_q = [1 \quad 0]$$

Let us assume Q and R are given by $Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $R = 1$.

First, we will check to see if the conditions are satisfied. The system is controllable because the matrix $[B \quad AB]$ has rank 2. Q can be factored as

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} [1 \quad 0] = C_q'C_q$$

The observability condition is also satisfied because the matrix $\begin{bmatrix} C_q \\ C_qA \end{bmatrix}$ has rank 2. Therefore, the ARE will have a stabilizing solution. Now, the ARE becomes

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} p_1 & p_2 \\ p_2 & p_3 \end{bmatrix} + \begin{bmatrix} p_1 & p_2 \\ p_2 & p_3 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$
$$- \begin{bmatrix} p_1 & p_2 \\ p_2 & p_3 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} p_1 & p_2 \\ p_2 & p_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Multiplying and adding the matrices, and setting the sides of the equation equal to each other, element by element, we get three coupled algebraic-quadratic equations. In this case, the equations are very simple (usually, they are quite horrendous) because of the number of zero elements in the matrices. The equations are

$$p_2^2 = 1$$
$$p_1 = p_2p_3$$
$$2p_2 - p_3^2 = 0$$

Therefore,

$$P = \begin{bmatrix} \sqrt{2} & 1 \\ 1 & \sqrt{2} \end{bmatrix} \qquad \text{and} \qquad K = R^{-1}B'P = [1 \quad \sqrt{2}]$$

The closed-loop system matrix becomes

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

The closed-loop characteristic equation and its roots (closed-loop eigen values) are

$$\lambda^2 + \sqrt{2}\lambda + 1 = 0$$

$$\lambda = \frac{\sqrt{2}}{2}(-1 \pm j)$$

Therefore, the system has been stabilized and it has a damping ratio of 0.707. Observe that the optimal controller is of the state feedback form (i.e., we are assuming all states are available for feedback). We discuss an observer design in the next section. Let us obtain the open-loop transfer function and use it to obtain Bode plots and determine the stability margins. The open-loop transfer function L(s) is given by (refer to Section 9.2 and Figure 9.3)

$$L(s) = K\Phi(s)B = K(sI - A)^{-1}B = \frac{\sqrt{2}[s + (\sqrt{2}/2)]}{s^2}$$

The Bode plots in Figure 10.1 show that the system has 65° of phase margin and infinite gain margin.

❑ DRILL PROBLEMS

D10.1 For each of the following systems described by A and B matrices and LQR performance criteria measured by Q and R, solve the associated algebraic Riccati equation and find the optimal control gains.

(a) $A = -2, \; B = 4, \; Q = 4, \; R = 1$

(b) $A = 2, \; B = 4, \; Q = 4, \; R = 1$

(c) $A = \begin{bmatrix} 0 & 1 \\ -10 & -2 \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \; Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \; R = 1$

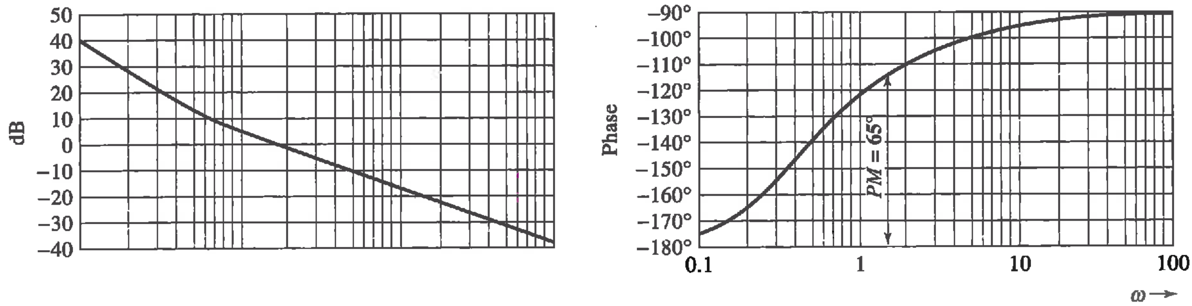
(d) $A = \begin{bmatrix} 0 & 1 \\ -10 & -2 \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \; Q = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \; R = 1$

Ans. (a) $P = 0.3904, k = +1.462;$

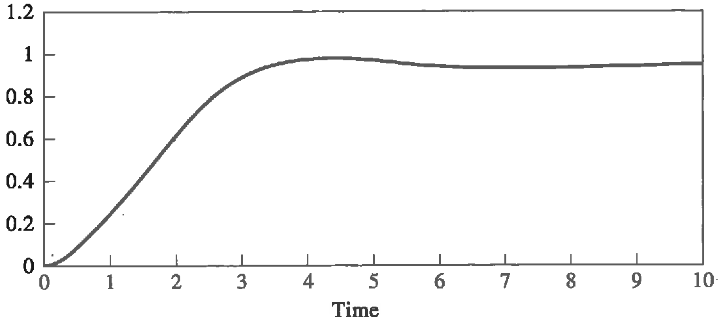
(b) $P = 0.6404, k = +2.562;$

(c) $P = \begin{bmatrix} 0.3455 & 0.0495 \\ 0.0495 & 0.0242 \end{bmatrix}, k = [0.1 \quad 0.048];$

(d) $p = \begin{bmatrix} 2.07 & 0 \\ 0 & 2.07 \end{bmatrix}, k = [0 \quad -0.414]$



(a)



(b)

Figure 10.1 (a) Bode plots for the LQR design of the double-integrator plant. (b) The step response for LQR design.

10.2.1 Properties of the LQR Design

LQR has many desirable properties. Among them are good stability margins and sensitivity properties. We will also discuss the effects of weights in the LQR setting. Most of these properties can be derived using the *return-difference inequality* first derived by Kalman.

10.2.2 Return Difference Inequality

The algebraic Riccati equation can be manipulated to arrive at the following relation:

$$|1 + L(j\omega)|^2 = 1 + \frac{1}{\rho}\left|G_q(j\omega)\right|^2$$

Where $L(s)$ is the loop gain (open-loop transfer function) given by

$$L(s) = K\Phi(s)B$$

where $\Phi(s) = (sI - A)^{-1}$, relation assumes that $Q = C_q' C_q$ and $G_q(s) = C_q \Phi(s)B$.

Because the right-hand side of the return-difference equality (RDE) is always greater than 1, the following inequality holds:

$$|1 + L(j\omega)| \geqslant 1$$

The preceding return-difference inequality (RDI) implies that for all frequencies, the Nyquist plot of the open-loop transfer function of an LQR-based design always stays outside a unit circle centered at $(-1, 0)$. A typical Nyquist plot is shown in Figure 10.2. The term *return difference*, introduced by Bode, means the following. Suppose a feedback loop is broken at a given point; inject a 1-volt signal at the entrance of the point and measure the signal returned at the exit of that point; the difference between the injected signal and the returned signal is called the *return difference*. If the gain around the loop is $-GH$, the return difference is $1 - (-GH)$ or $1 + GH$. The return difference is a measure of the amount of feedback in a feedback loop. It is an important quantity and appears in many expressions, such as the denominator of the closed-loop transfer function and the sensitivity function defined in earlier chapters.

The return-difference inequality, along with simple geometric arguments, can be used to show that the LQR solution, in the SISO case, has at least 60° phase margin, infinite gain margin, and a gain reduction tolerance of -6 dB. The latter means that the gain can be reduced by a factor of $\frac{1}{2}$ before instability occurs. Therefore, an LQR design behaves quite well from a classical control point of view. It not only always results in an asymptotically stable system but also provides guaranteed stability margins. This is to be compared with the state feedback pole placement technique discussed in Chapter 9, where stability margins are not known or guaranteed ahead of time. Finally, observe that the LQR margins are a bit excessive in that lower gain and phase margins are generally acceptable in most designs.

Stability margins of LQR.

Another frequency domain property of the LQR solution is its high-frequency roll-off rate. Recall that the closed-loop transfer function of state feedback design is

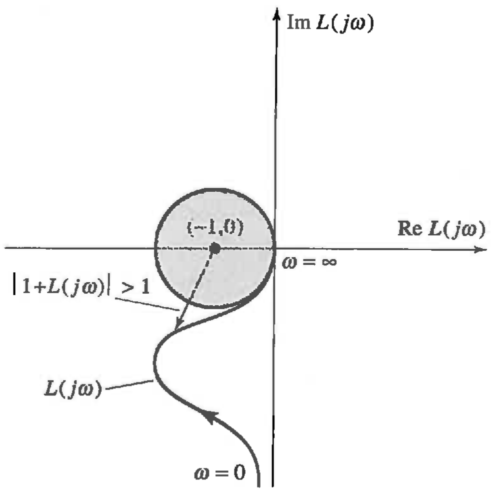


Figure 10.2 Nyquist plot of an LQR-based design. The plot always stays outside the unit circle centered at $(-1, 0)$.

given by

$$T(j\omega) = -K(j\omega I - A + BK)^{-1} B$$

it can be shown that

$$\lim_{\omega\to\infty} T(j\omega) = \frac{-1}{j\omega} KB = \frac{-1}{j\omega} R^{-1} B' P B < 0$$

LQR Bode plot rolls off at −20 dB at high frequencies.

The preceding implies that $|T(j\omega)|$ drop as$1/(j\omega)$ in the SISO case, indicating a roll-off rate (i.e., a slope) of −20 dB per decade at high frequencies [see Figure 10.1(a)]. This, of course, affects the noise suppression properties of the optimal system and as such is not very good. It can be argued that this defect is the result of the excessive stability margins of the LQR solution.

10.2.3 Optimal Root Locus

We will see that a special choice of Q and R allows us to investigate the effects of weights on the location of closed-loop poles. Let us assume that Q and R are given by

$$Q = C_q' C_q \quad \text{and} \quad R = \rho I$$

where ρ is a positive scalar. Then

$$x' Q x = z' z$$

where $z = C_q x$
and the cost function becomes

$$J = \frac{1}{2} \int_0^{\alpha} \left(z'z + \rho u'u \right) dt$$

This means that we are minimizing the system output and control energy, and by increasing ρ, we can put more emphasis on minimizing control energy. Definite the following matrix, called the *Hamiltonian* matrix

The Hamiltonian matrix for LQR.

$$\mathcal{H} = \begin{bmatrix} A & -\frac{1}{\rho} BB' \\ -C_q' C_q & -A' \end{bmatrix}$$

Eigenvalues of $\mathcal{H}$ are symmetric with respect to the imaginary axis.

Because of the special structure of the Hamiltonian matrix, its characteristic polynomial is an even polynomial (i.e., if s is a root, so is $-s$). Therefore, it can be factored as a polynomial with only LHP roots and a polynomial with only RHP roots (the Hamiltonian has no roots on the imaginary axis). The Hamiltonian matrix is used in formal proofs of the LQR problem, and the eigenvalues and eigenvectors of the Hamiltonian matrix are used to solve the ARE (Potter's method).

The optimal closed-loop poles will be the stable (i.e., LHP) eigenvalues of the Hamiltonian matrix. If we denote the characteristics polynomial of the Hamiltonian matrix by $\Delta_c = |sI - \mathcal{H}|$, after a series of matrix manipulations, we arrive at the following equation (n = number of poles, m = number of zeros, with $m < n$, $r = n - m$)

$$\Delta_c(s) = (-1)^n \Delta(s) \Delta(-s) \left[1 + \frac{1}{\rho} G_q(s) G_q(-s) \right]$$

where $\Delta(s) = |sI - A|$ and preceding $G_q(s) = n_q(s)/d(s)$ is the transfer function from u to z (the regulated output). The preceding equation simplifies to

$$(-1)^n \Delta_c(s) = \Delta(s)\,\Delta(-s)\left[1 + \frac{1}{\rho} G_q(s)\,G_q(-s)\right]$$
$$= d(s)\,d(-s) + \frac{1}{\rho} n_q(s)\,n_q(-s)$$

This has the standard root locus form. It implies that the optimal closed-loop poles can be obtained from the root locus of $G_q(s)\,G_q(-s)$. Such root loci are generally called a *symmetric root locus* or *root-square locus.* We will discuss the effects of limiting values of ρ.

Root locus of $G_q(s)\,G_q(-s)$ gives the optimal pole locations.

Minimum Energy Control (or Expensive Control) Case

$$As \quad \rho \to \infty, \quad (-1)^n \Delta_c(s) \to d(s)\,d(-s)$$

Because the optimal closed-loop poles are always in the LHP, we conclude that as the control weighting is increased, the stable open-loop poles will remain where they are and the unstable ones will be reflected about the imaginary axis. This property can be used as a guideline for pole placement.

Cheap Control Case

$$\text{As } \rho \to 0, \quad (-1)^n \Delta_c(s) \to n_q(s)\,n_q(-s) \quad \text{for finite } s$$

Hence, the closed-loop poles approach the plant finite zeros or their stable images. For values of s approaching , the closed-loop poles will approach zeros at infinity in the so-called *Butterworth pattern.*

$$\text{for } |s| \to \infty \quad s = \left(\frac{\alpha_m^2}{\rho}\right)^{1/2r} \exp\left[j\frac{\pi k(r+1)}{2r}\right] \qquad k = \text{odd integer}$$

where α_m is the coefficient of the highest-order term in $n_q(s)$.

As an example of root-square locus (RSL) consider the following system

RSL example.

$$G_q(s) = \frac{5}{s^2 + s + 5}$$

Then the optimal characteristic equation is given by

$$1 + \frac{1}{\rho}\frac{25}{\left(s^2 + s + 5\right)\left(s^2 - s + 5\right)} = 0$$

The optimal closed-loop poles are along the LHP branches of the RSL shown in Figure 10.3. Note that the RSL is symmetric with respect to both the imaginary and the real axes. The RSL shows what happens to the poles as the control cost weight (ρ) increases from 0 to infinity (note that the actual root locus gain $1/\rho$). When ρ approaches infinity (i.e., the root locus gain goes to 0), the closed-loop poles approach the plant open-loop poles; therefore, when control is expensive and the

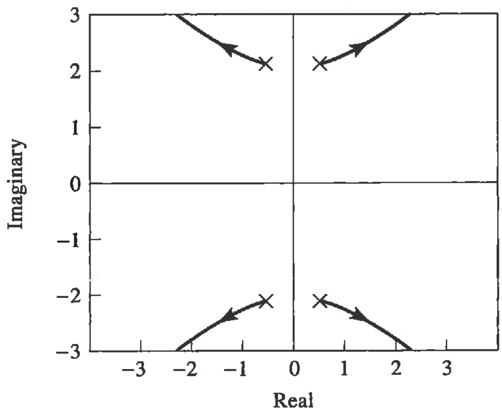


Figure 10.3 Root-Square locus for $G_q(s)$.

plant is stable, the best choice is to do nothing and leave the poles where they are. When ρ approaches 0 (i.e., the root locus gain goes to infinity), the closed-loop poles approach the plant open-loop zeros at infinity.

Now suppose that the plant is unstable and is given by

Another RSL example.

$$G_q(s) = \frac{5}{s^2 - s + 5}$$

We obtain the same RSL as shown in Figure 10.3. The interpretation is slightly different here. In the minimum-energy control case (expensive control), the best choice is to reflect the unstable poles about the imaginary axis. In either of these examples, the optimal characteristic equation for the minimum-energy control case is

$$s^2 + s + 5$$

To find the optimal control gain K using RSL, we first determine the optimal pole location from the RSL, form the characteristic polynomial, and set this equal to the characteristic polynomial of $A - BK$ and solve for K. For example, in the preceding case, we get

$$\begin{aligned} |sI - (A - BK)| &= s^2 + (-1 + 5k_2)\,s + 5 + 5k_1 \\ &= s^2 + s + 5 \end{aligned}$$

Therefore,

$$k_1 = 0 \quad \text{and} \quad k_2 = 0.4$$

Let us consider another example.

RSL example.

$$G_q(s) = \frac{s}{s^2 + 2s + 10} \quad \text{and} \quad 1 + \frac{1}{\rho}\frac{-s^2}{\left(s^2 + 2s + 10\right)\left(s^2 - 2s + 10\right)} = 0$$

The RSL for positive ρ is shown in Figure 10.4(a). Note that for positive ρ the locus has imaginary poles, and this cannot correspond to an optimal system. Therefore, we have to use negative-gain sketching rules for root locus. The optimal locus is shown in Figure 10.4(b). In this example, this was clear because of the negative sign in the numerator. In general, if r (the number of poles minus the number of zeros in the plant

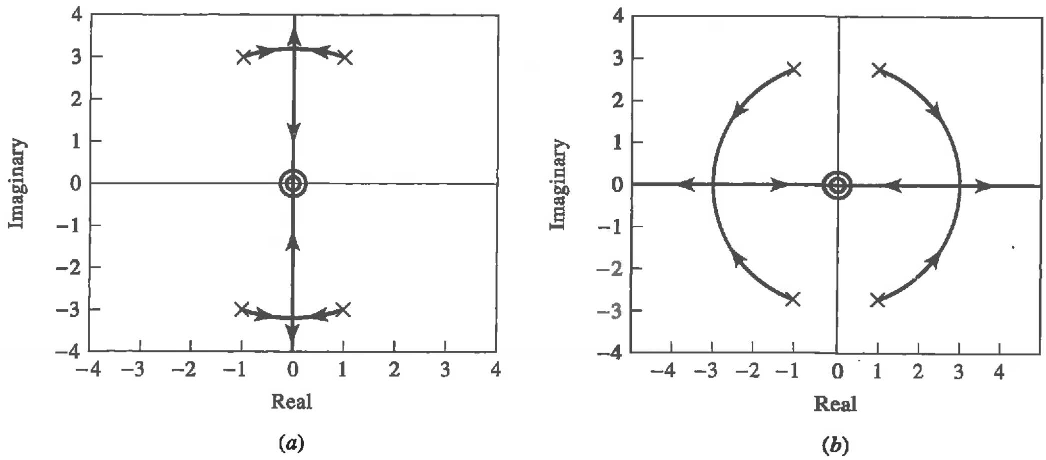


Figure 10.4 (a) Root-square locus using positive gain. (b) Root-square locus using negative gain.

transfer function) is odd, we have to use negative-gain sketching rules; otherwise we use positive-gain sketching rules for plotting the root locus.

❑ DRILL PROBLEMS

D10.2 Consider the systems presented in Drill Problem 10.1. In each case, factor $Q = C_q' C_q$ and let $R = \rho$. Then find $G_q(s)\, G_q(-s)$ and plot the appropriate root-square locus.

Ans. (a) $Q = 4 = 2\,(2)\,,\, C_q = 2,\, G_q(s)\, G_q(-s) = [8/(s+2)]\,[8/(-s+2)]$

(b) $Q = (2)(2),\, C_q = 2,\, G_q(s) = [8/(s-2)]\,[8/(-s-2)]$

(c) $Q = \begin{bmatrix} 1 \\ 0 \end{bmatrix} [1 \quad 0],\, C_q = [1 \quad 0],\, G_q(s) = \left[2/(s^2+2s+10)\right]\left[2/(s^2-2s+10)\right]$

(d) $Q = \begin{bmatrix} 0 \\ 1 \end{bmatrix} [0 \quad 1],\, C_q = [0 \quad 1],$
$G_q(s) = \left[2s/(s^2+2s+10)\right]\left[-2s/(s^2-2s+10)\right]$

See Figure 10.4(b) for the RSL

10.3 Optimal Observers—the Kalman Filter

The LQR solution is basically a state-feedback type of controller—i.e., it requires that all states be available for feedback. It was argued in the previous chapter that this is usually an unreasonable assumption and some form of state estimations necessary.

In addition, the concept of the observer was introduced and it was shown that the combination of the state feedback and observer will always result in a stable closed-loop system. In chapter 9, however, the designer was left with the responsibility of choosing the controller and observer poles. We have seen in the previous section that the controller performance can be optimized according to some quadratic cost function, resulting in optimal controller pole locations. The next obvious question is whether the observer design can also be done in an optimal manner. The answer is affirmative, provided the problem is formulated in a probabilistic (or stochastic) sense. The formulation of the state estimation problem is as follows

$$\dot{x} = Ax + Bu + \omega$$

$$y = Cx + v$$

where ω represents random noise disturbance input (process noise) and ν represents random measurement (sensor) noise; we also have to assume some statistical knowledge of the noise processes. For instance, in the case of an aircraft, the plant is subject to random wind disturbances (or process noise), and the measurement instrumentations (sensors) are not always accurate and may include random errors (sensor noise). Ships and other marine vessels are subjected to random wave motions (which may also have strong periodic components), and in general, most systems are subject to both kinds of random inputs.

The state-space solution to this problem was first provided by R.E. Kalman and R.S. Bucy. The optimal observer (commonly known as the *Kalman filter*) is given by

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - C\hat{x})$$

where $\hat{x}$ is the estimate of x. The observer gain is computed from

$$L = \Sigma C' R^{-1}$$

and Σ is found as the positive semi-definite solution of

$$A\Sigma + \Sigma A' + Q_o - \Sigma C' R_o^{-1} C \Sigma = 0$$

Note that the equation for the filter gain and Σ are very similar to the equations for the LQR solution. In particular, the equation for Σ is an algebraic Riccati equation. There are two matrices that appear in the filter equation that require some explanation. They are Q_o and R_o. These matrices represent the intensity of the process and sensor noise inputs and are the only parameters that are to be provided by the user. In the mathematical subject of random processes, these matrices are known as *co-variance matrices*. Their size (usually measured by their trace, the sum of the diagonal elements) is a measure of how strong the noise is—the larger the size, the more random or intense the noise—hence we refer to it as *noise intensity*.The Kalman filter attempts to minimize the size of the estimation error intensity (the intensity of the estimation error is given by Σ). Finally, we note that the mathematical conditions that are needed for the solution of the Kalman filter problem to exist are the following: $Q_o$and R_o must be positive semidefinite and positive definite respectively, and the system must be observable.

Estimation theory and, in particular, Kalman filter theory are vast and important areas that are common to control and communications. There are many reported successful applications of Kalman filters in a wide range of areas (many more than LQR implementations). Because our interest lies in control system design rather than pure state estimation, we will return to the control problem without pursuing this subject any further. Hence, we view the Q_o and R_o matrices as design parameters, not necessarily related to physical noise processes.

10.4 The Linear Quadratic Gaussian (LQG) Problem

Given an optimal filter to estimate the states, the next question is whether the closed-loop system remains stable and optimal when we combine the LQR controller of Section 10.2 and the Kalman filter of Section 10.3. This problem is known as the *linear quadratic Gaussian* (or *LQG*) problem. The term *Gaussian* refers to the statistical distribution of the noise processes. The plant equations and the problem solution are repeated:

$$\dot{x} = Ax + Bu + W$$
$$y = Cx + v$$

The controller portion is given by

$$u = -K\hat{x}(t)$$
$$K = R^{-1}B'P$$
$$A'P + PA - PBR^{-1}B'P + Q = 0$$

The observer (or filter) portion is given by

$$\dot{\hat{x}} = A\hat{x} + Bu + L\left(y - C\hat{x}\right)$$
$$L = \Sigma C'R^{-1}$$
$$A\Sigma + \Sigma A' + Q_o - \Sigma C' R_o^{-1} C\Sigma = 0$$

It can be shown that the LQG solution results in an asymptotically stable closed-loop system. In addition, the controller minimizes the average of the LQR cost function (i.e., the weighted variance of the state and input), resulting in an optimal solution. Because the structure of the controller and the Kalman filter are similar to the observer-based compensator discussed in Chapter 9 (the major difference is how the control and filter gains are computed), the LQG compensator will also exhibit the separation property (the mathematical proof of this fact is actually quite involved). Hence, the closed-loop poles will be the union of the controller poles and the filter poles, and the controller and the filter can be designed independently of each other (this means that the filter equation do not contain K or P, and the control equations do not depend on L or Σ).

The transfer function of the LQG compensator is similar to the observer-based compensator, and is given by

$$H(s) = K\left(sI - A + BK + LC\right)^{-1} L$$

LQG compensator.

Let us obtain an LQG compensator for the double-integrator plant. The controller portion has already been found, so we will design the Kalman filter. We will assume the noise intensities are

$$Q_o = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad R_o = 1$$

The filter Riccati equation results in three coupled algebraic nonlinear equations:

$$a^2 = 2b + 1$$
$$ab = c$$
$$b^2 = 1$$

where $\Sigma = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$

Therefore,

$$\Sigma = \begin{bmatrix} \sqrt{3} & 1 \\ 1 & \sqrt{3} \end{bmatrix} \quad \text{and} \quad L = \Sigma C' R_o^{-1} = \begin{bmatrix} \sqrt{3} \\ 1 \end{bmatrix}$$

LQG design for the double-integrator system.

The transfer function of the compensator is given by

$$H(s) = \frac{3.14\,(s + 0.31)}{(s + 1.57 + j1.4)\,(s + 1.57 - j1.4)}$$

Further computation shows that the closed-loop poles are at the locations of the controller and filter eigenvalues, respectively:

$$\text{Closed-loop poles} = \frac{\sqrt{2}}{2}(-1 \pm j)\,,\ \frac{-\sqrt{3} \pm j}{2}$$

The root locus, open-loop magnitude and phase plots, closed-loop magnitude plot, and closed-loop step response of the system are shown in Figure 10.5. The Bode plots indicate a gain margin of 10.7 dB and phase margin of 34.5°. Let us now compare the LQR and LQG designs.

LQR and LQG comparison for the double-integrator.

1. LQR has much higher stability margins.
2. The low-frequency gain in LQR is 40 dB and in LQG is 27 dB. Hence, LQR will have better steady state tracking properties (recall that error coefficients are obtained by letting s approach 0, so as low-frequency gain in the open-loop magnitude plot determines steady state error properties).
3. The gain-crossover frequency is higher in LQR. This means that LQR has a higher bandwidth, so it passes more noise into the system. Also, since gain crossover frequency is inversely related to the speed of response, this indicates a faster response in LQR, as can be seen in the step response.
4. The high-frequency roll-off rates, approximated by the slope of Bode magnitude plot in Figure 10.5(b), are −60 dB and −20 dB in LQG and LQR, respectively. This means that LQG has better noise suppression properties.

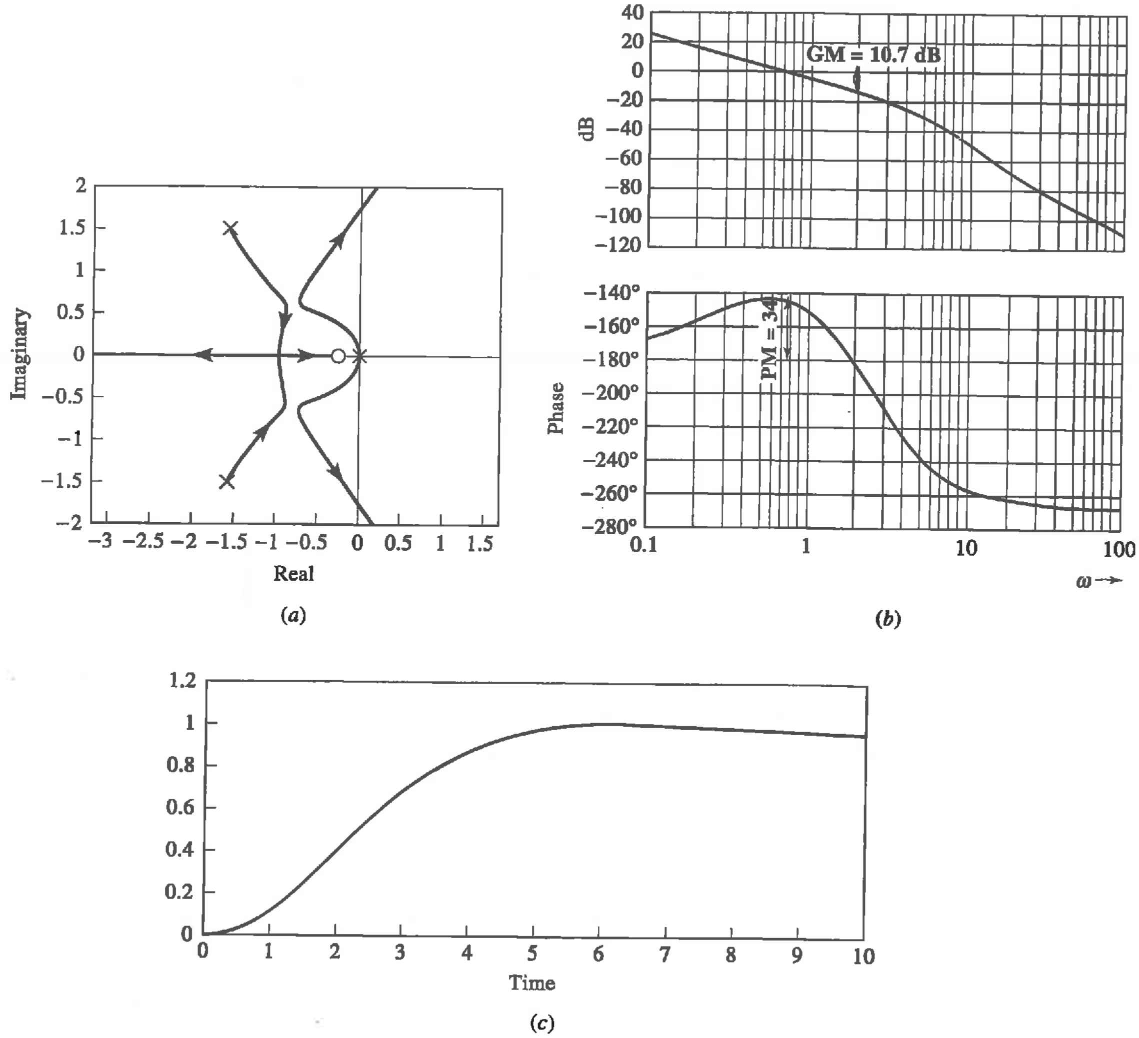


Figure 10.5 Classical root locus for LQG design. (b) Open-loop Bode plots for LQG design. (c) Closed-loop step response for LQG design.

In this example, we see the trade-offs involved in control system design. Stability margin is traded off with high-frequency roll-off rate. Gain crossover frequency (or bandwidth) is traded off with speed of response.

We will now examine the Nyquist plots for both cases shown in Figure 10.6. As predicted by the return-difference inequality, the LQR plot avoids the unit circle centered at $(-1, 0)$, whereas the LQG plot enters it. This shows that the LQG open-loop transfer function does not satisfy the return-difference inequality. This has very important implications because the RDI is the basis for the guaranteed stability margins of LQR. In fact, it has been shown by counterexamples that LQG has no guaranteed stability margins and its margins can be dangerously low.

LQG has no guaranteed stability margins.

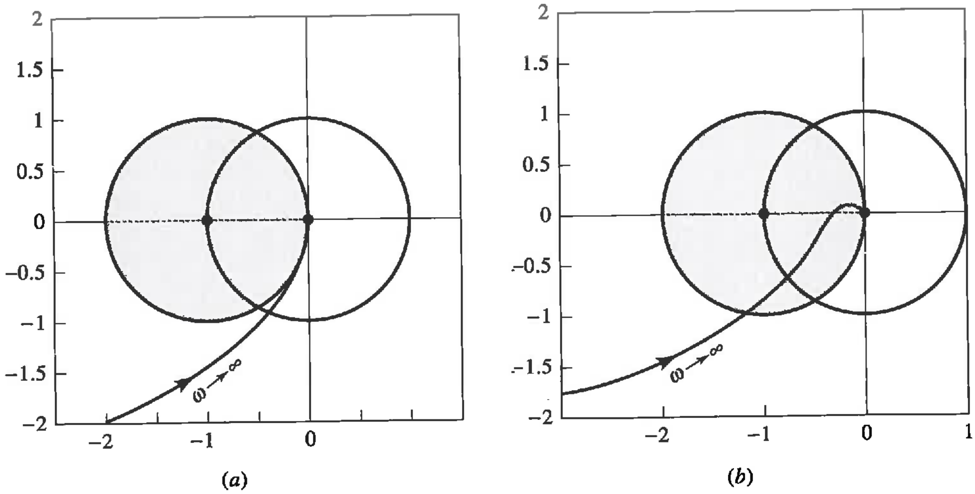


Figure 10.6 (a) Nyquist plot for LQR design. (b) Nyquist plot for LQG design. The unit circles centered at (0, 0) and (-1, 0) are shown.

You can experiment by changing the design parameters Q and R and the noise intensities, and you will observe that some parameters can have drastic effects on the system properties. But how does one choose these parameters? We will try to investigate this question in the subsequent sections.

10.4.1 Critique of LQG

Early pioneers of control, particularly H. W. Bode and I. M. Horowitz, studied and delineated most of the properties of feedback. In the early 1960s, with the birth of modern control, optimality and the design of optimal control systems became the dominant concern. The solution of the LQG problem was probably the highlight of this era. However, the LQG paradigm failed to meet the main objectives of control system designers. That is LQG control failed to work in real environments. The major problem with the LQG solution was lack of robustness. In a series of papers, researchers showed that LQG-based designs can become unstable in practice as more realism it added to the plant model. The same kinds of failure were also observed in industrial experiments with LQG. It became apparent that the main culprit was too much emphasis on optimality and not enough attention to the model uncertainty issue. During the 1980s, much of the attention was shifted back to feedback properties and frequency domain techniques (which were the main features of classical control), and their generalization to multivariable systems.

Section 10.6 discusses the *loop transfer recovery* (LQG/LTR) technique. This method maintains the LQG machinery but modifies the design procedure to address some of the shortcomings of the original LQG approach.

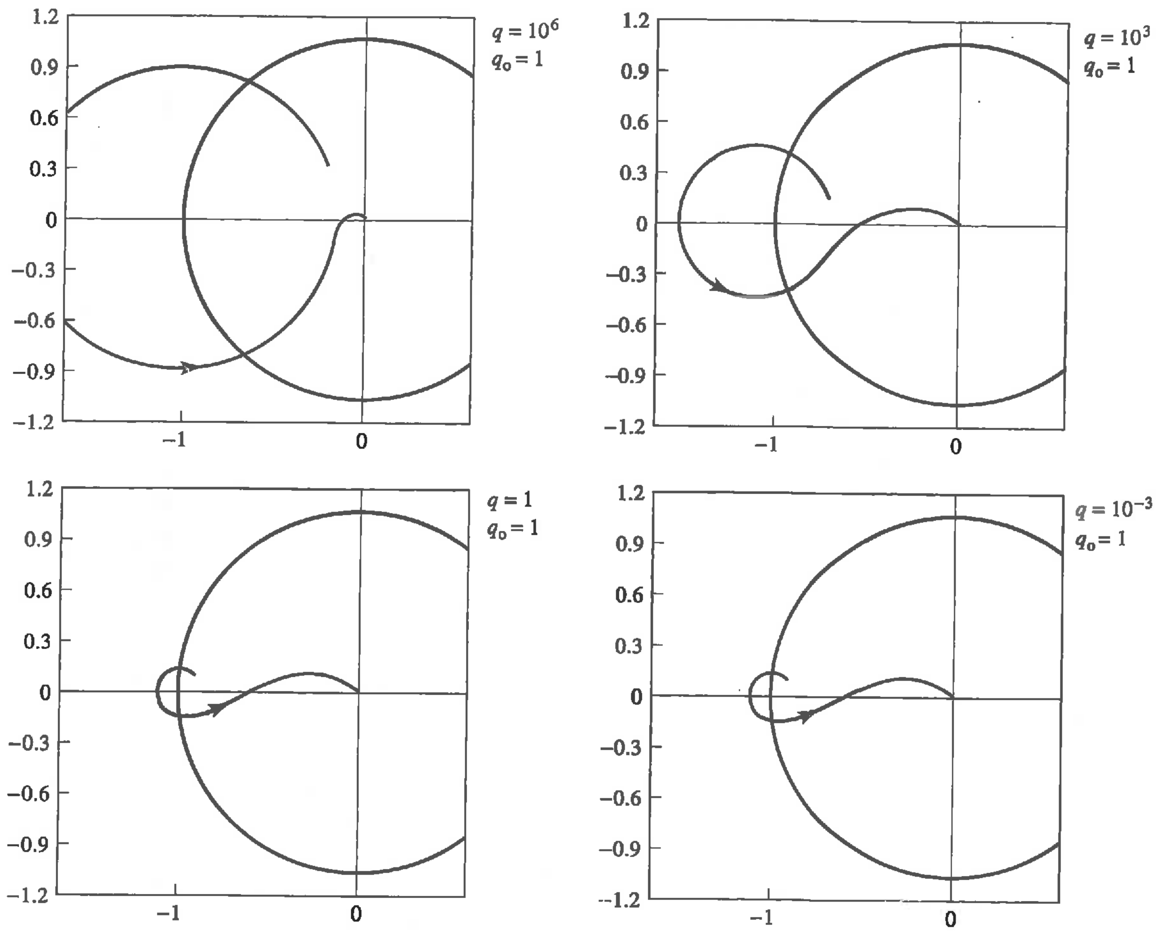
❑ DRILL PROBLEM

D10.3 Consider the following system (Doyle, J. and Stein, G., *IEET Trans. (Auto— Control,* August 1979).

$$\dot{x} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u + \begin{bmatrix} 1 \\ 1 \end{bmatrix} w$$

$$y = [1 \quad 0] x + v$$

Let the LQR parameters be $Q = qC'C$, $R = 1$, and let the filter parameters be $Q_o = q_o \begin{bmatrix} 1 \\ 1 \end{bmatrix} [1 \quad 1]$, $R_o = 1$. For each case given, compute gain and phase margins and draw the Nyquist plot overlaid on the unit circle centered at the origin.



(*a*)

Figure D10.3

(a) $q = 10^6, 10^3, 1, 10^{-3}$ and $q_o = 1$

(b) $q_o = 10^6, 10^3, 1, 10^{-3}$ and $q = 1$

Ans. (a) $GM = \begin{bmatrix} -5.7 & -3.4 & -0.8 & -0.7 \\ 21.3 & 9.8 & 5.7 & 5.6 \end{bmatrix}$

$PM = \begin{bmatrix} -51 & -25.6 & -4.5 & -3.5 \\ 51 & 26.5 & 9.2 & 8.5 \end{bmatrix}$

(b) $GM = \begin{bmatrix} -1.2 & -1.2 & -0.8 & -0.8 \\ 49.7 & 20.7 & 5.7 & 5.3 \end{bmatrix}$

$PM = \begin{bmatrix} 19 & 17.7 & -4.5 & -4.6 \\ - & - & 9.2 & 8.6 \end{bmatrix}$

10.5 Robustness

The ultimate goal of a control system designer is to build a system that will work in real environment. Since the real environment may change with time (as components age or their parameters vary with temperature or other environmental conditions) or operating conditions may vary (load changes, disturbances), the control system must be able to withstand these variations. Assuming that the environment does not change, the second fact of life is the issue of model uncertainty. A mathematical representation of a system often involves simplifying and sometimes wishful assumptions. Nonlinearities are either unknown, hence unmodeled, or modeled and later ignored to simplify analysis. Different components of systems (actuators, sensors, amplifiers, motors, gears, belts, etc.) are sometimes modeled by constant gains, even though they may have dynamics or nonlinearities. Dynamic structures (e.g., aircrafts, satellites, missiles) have complicated dynamics in high frequencies, and these may initially be ignored. Since control systems are typically designed using much simplified models of systems, they may not work on the real plant in real environments.

The particular property a control system must possess to operate properly in realistic situations is called *robustness*. Mathematically this means that the controller must perform satisfactorily not just for one plant, but for a family (or set) of plants. Let us be more specific. Suppose the following plant is to be stabilized.

$$G(s) = \frac{1}{s - a}$$

It is suspected that the value of the parameter a is equal to 1, but this value could be off by 50%. If we design a controller that will stabilize the system for all values of $0.5 \leqslant a \leqslant 1.5$,we say the system has *robust stability.* If in addition the system is to satisfy performance specifications such as steady state tracking, disturbance rejection, and speed of response requirements, and the controller satisfies all requirements for all values of a in the specified range, we say the system possesses *robust performance.* The problem of designing controllers that satisfy robust stability and performance requirements is called *robust control*. This problem was investigated intensely during the 1980s and is still under investigation by many researchers following a variety of approaches. We will present a brief introduction to the robust control problem in the ensuring sections.

The underlying concept within control theory that has made it into a field of science is feedback. The study of feedback and its properties is responsible for the rapid growth of this field. What are these properties and why do we use feedback? The answer is that feedback has many properties that are discussed, either implicitly or explicitly, in this book. But there are two properties that a feedback system possesses that an open-loop system cannot have: sensitivity and disturbance rejection. By *sensitivity* it is meant that feedback reduces the sensitivity of the closed-loop system with respect to uncertainties or variations in elements located in the forward path of the system. *Disturbance rejection* refers to the fact that feedback can eliminate or reduce the effects of unwanted disturbances occuring within the feedback loop. It is mainly for these reasons that feedback is used. An open-loop system (i.e., a system with no feedback) does not have these properties. Of course, an open-loop system can also eliminate certain disturbances, but it requires full knowledge of the disturbance, which is not always available. Feedback is also used to stabilize unstable systems, but feedback itself is frequently the cause of instability. The stabilizing effects of feedback are emphasized so much in most texts that its other important properties are forgotten by beginning (or even experienced) students of control.

Two very important properties of feedback.

10.5.1 Feedback Properties

A feedback control system must satisfy certain performance specifications, and it must tolerate model uncertainties. We will study these issues from a frequency domain perspective. Consider the feedback system in Figure 10.7. The system has the following inputs:

$R(s)$ = command (or reference) input. This is the input that the system must be able to follow or track.

$D(s)$ = disturbance input. Disturbances are known or unknown inputs that the system must be able to reject. Disturbances may represent actual physical disturbances acting on the system such as wind gusts disturbing aircraft, disturbances owing to actuators such as motors, or uncertainties resulting from model errors in plant or actuator. Model uncertainties include neglected nonlinearities in plant or actuator and neglected or unknown modes in the system.

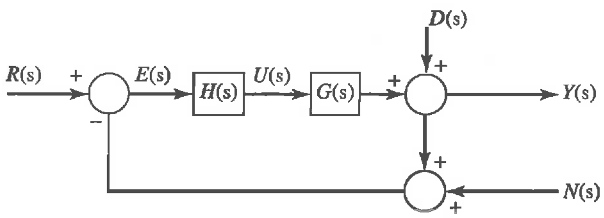


Figure 10.7 Block diagram of a feedback control system including disturbance and measured noise inputs.

$N(s)$ = sensor or measurement noise. This is introduced into the system via sensors, which are usually random high-frequency signals.

A properly designed control system must track reference inputs with small error and reject disturbance and noise inputs. The contribution of general disturbances to the output must be small. The total output of the closed-loop system is

$$Y(s) = \frac{G(s)H(s)}{1+G(s)H(s)}R(s) + \frac{1}{1+G(s)H(s)}D(s) - \frac{G(s)H(s)}{1+G(s)H(s)}N(s)$$

If we define the tracking error as $e = r - y$, we get

$$E(s) = \frac{1}{1+G(s)H(s)}R(s) - \frac{1}{1+G(s)H(s)}D(s) + \frac{G(s)H(s)}{1+G(s)H(s)}N(s)$$

Finally, the actuator output (i.e., the plant input) is given by

$$U(s) = \frac{H(s)}{1+G(s)H(s)}[R(s) - D(s) - N(s)]$$

Several quantities appear frequently in these relationships, they are

$$J(s) = 1 + G(s)H(s) \qquad \textit{return difference}$$

$$S(s) = \frac{1}{1+G(s)H(s)} = \frac{1}{J(s)} \qquad \text{sensitivity}$$

$$T(s) = \frac{G(s)H(s)}{1+G(s)H(s)} \qquad \textit{complementary sensitivity}$$

It can be seen that, for all frequencies, the following equality holds:

$$S(s) + T(s) = 1$$

Using the earlier definitions, we can write

$$Y(s) = S(s)D(s) + T(s)[R(s) - N(s)]$$
$$E(s) = S(s)[R(s) - D(s)] + T(s)N(s)$$
$$U(s) = H(s)S(s)[R(s) - D(s) - N(s)]$$

We are now ready to draw the following conclusions from these relations:

1. **Disturbance rejection:** $S(s)$ must be kept small to minimize the effects of disturbances. From the definition of S, this can be met if the loop gain (i.e., GH) is large.
2. **Tracking:** $S(s)$ must be kept small to keep tracking errors small.
3. **Noise suppression:** $T(s)$ must be kept small to reduce the effects of measurement noise on the output and errors. From the definition of T, this is met if the loop gain is small.
4. **Actuator limits:** $H(s)S(s)$ must be bounded to ensure that the actuating signal driving the plant does not exceed plant tolerances. Another reason for taking this relation into consideration is to reduce the control energy so that we can use smaller actuators (such as motors).

Tracking and disturbance rejection require small sensitivity. Noise suppression requires small complementary sensitivity. Because these two transfer functions add up to unity, we cannot reduce both transfer functions to zero simultaneously. We can, however, avoid this conflict by noticing that, in practice, command inputs and disturbances are low-frequency signals (i.e., they vary slowly with time), whereas measurement noise is a high-frequency signal. Therefore, we can meet both objectives by keeping S small in the low-frequency range and T small in high frequencies. The control energy constraint requires keeping HS small. Note that

$$H(s)S(s) = \frac{H(s)}{1 + G(s)H(s)} = \frac{T(s)}{G(s)}$$

Hence, by keeping T small we can reduce control energy. Putting together these effects, we arrive at a general desired shape for the open-loop transfer function (or loop gain) of a properly designed feedback system. This is shown in Figure 10.8. The general feature of this loop gain is that it has high gain at low frequencies (for good tracking and disturbance rejection) and low gain at high frequencies (for noise suppression). The intermediate frequencies typically control the gain and phase margins. Bode has shown that for a stable system, the slope of the magnitude plot should not exceed −40 dB/decade; that is, the transition from low- to high-frequency range must be smooth (e.g., −20 dB/decade). Desirable shapes for sensitivity and complementary sensitivity transfer functions are shown in Figure 10.9. Note that S must be small at low frequencies and roll off to 1 (0 dB) at high frequencies, whereas T must be at 1 (0 dB) at low frequencies and diminish at high frequencies. These properties are summarized in Table 10.1.

10.5.2 Uncertainty Modeling

The preceding performance specifications apply to a stable feedback system. As we discussed earlier, a stable system is not our final objective; rather, the stability must

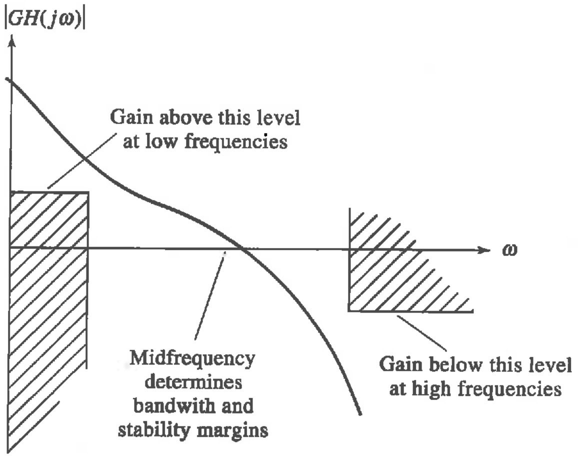


Figure 10.8 Desirable shape for the open-loop transfer function of a feedback system.

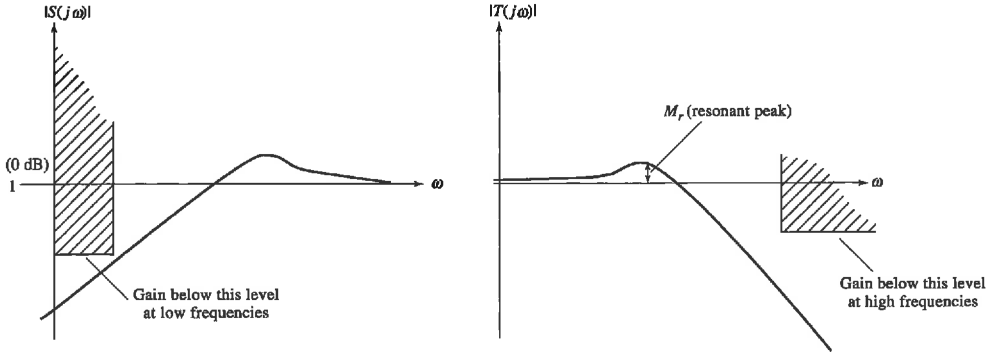


Figure 10.9 Desirable shape for the sensitivity and complementary sensitivity of a feedback system.

Table 10.1 *Loop Transfer Function Properties*

	Low Frequency	Mid Frequency	High Frequency
Performance (R)	High gain	Smooth transition	
Disturbance rejection (D)	High gain		
Noise suppression (N)			Low gain

be maintained despite model uncertainties. Model uncertainty is generally divided into two categories: *structured uncertainty* and *unstructured uncertainty*. Structured uncertainty assumes that the uncertainty is modeled and we have ranges and bounds for uncertain parameters in the system. For example, we may have a valid transfer function model of a system but have some uncertainty about the exact location of the poles, zeros, or gain of the system. In the case of an *RLC* circuit we know that it can be adequately modeled by a second-order transfer function (in a given frequency range), but the components may have up to 20–30% tolerance. These kinds of uncertainties are structured. Unstructured uncertainties assume less knowledge of the system. We only assume that the frequency response of the system lies between two bounds. Both kinds of uncertainties are usually present in most applications. We will discuss only unstructured uncertainty. (Dealing with structured uncertainty is still under investigation; owing to the complexity of the problem and space limitations, we will not discuss this case.)

Unstructured uncertainty can be modeled in different ways. We will discuss *additive* and *multiplicative* uncertainty. Suppose we model a system by $G(s)$, where the actual system is given by $\tilde{G}(s)$—That is,

$$\tilde{G}(s) = G(s) + \Delta_a(s)$$

Therefore, the model error, or the additive uncertainty, is given by

$$\Delta_a(s) = \tilde{G}(s) - G(s)$$

In the multiplicative uncertainty case, we assume the true model, $\tilde{G}(s)$ is given by

$$\tilde{G}(s) = [1 + \Delta_m(s)]G(s)$$

The uncertainty, or the model error, is given by

$$\Delta_m(s) = \frac{\tilde{G}(s) - G(s)}{G(s)}$$

Block diagram representations of these uncertainty models are shown in Figure 10.10. Because multiplicative uncertainty represents the relative error in the model, whereas the additive model represents absolute error, the multiplicative model is used more often.

As an example, consider the flexible spacecraft example in Section 4.10. The nominal plant model consists of the rigid mode, and is given by $G(s)$. The true plant model, $\tilde{G}(s)$, must also include the flexible mode.

$$G(s) = \frac{2}{s^2} \quad \tilde{G}(s) = \frac{s^2 + 2s + 2}{s^2(s^2 + s + 1)}$$

Modeling the flexible mode as additive uncertainty, we get

$$\Delta_a(s) = \tilde{G}(s) - G(s) = \frac{-1}{s^2 + s + 1}$$

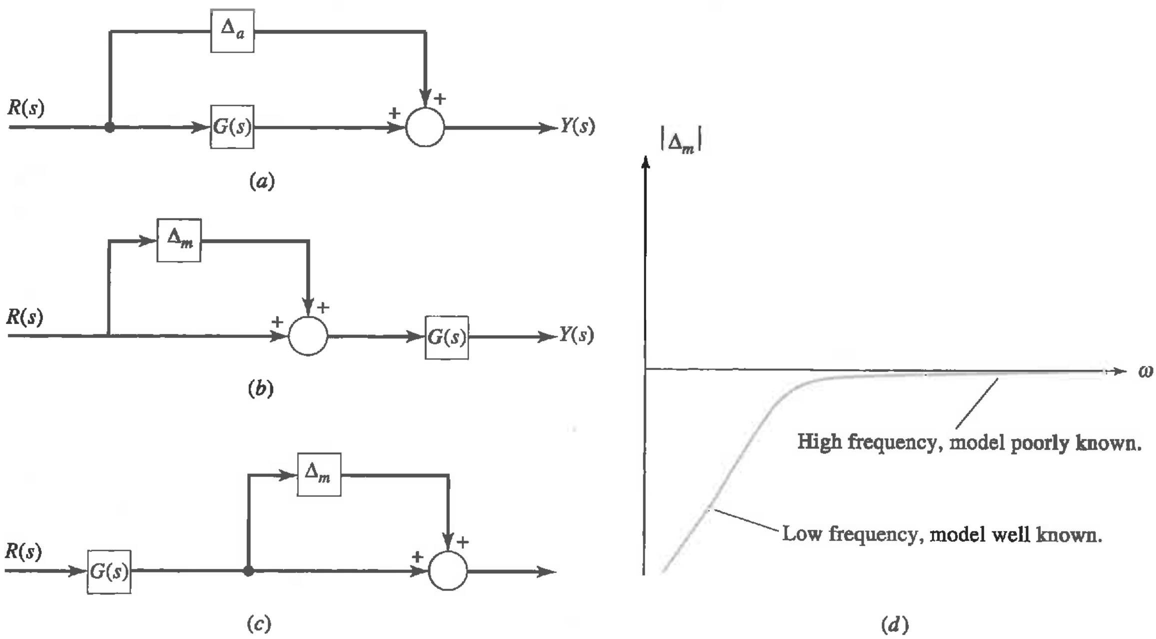


Figure 10.10 (a) Additive uncertainty. (b) Multiplicative uncertainty at the plant input. (c) Multiplicative uncertainty at the plant output. (d) Typical shape for multiplicative uncertainty.

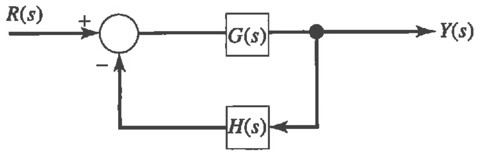


Figure 10.11 Block diagram of a feedback control system.

Using the multiplicative model, we have

$$\Delta_m(s) = \frac{\tilde{G}(s) - G(s)}{G(s)} = \frac{-s^2}{2(s^2 + s + 1)}$$

10.5.3 Robust Stability

Consider a feedback system containing a plant and a compensator. Suppose the compensator stabilizes the nominal plant model $G(s)$. We say that the compensator *robustly stabilizes* the system if the closed-loop system remains stable for the true plant $\tilde{G}(s)$. Most of the results and conditions for robust stability can be derived from variations of the Nyquist stability criterion or the following very powerful result, called the *small-gain theorem.*

Small-Gain Theorem

Consider the feedback system in Figure 10.11. Assume the plant and the compensator are stable. Then the closed-loop system will remain stable if

$$|G(s)H(s)| < 1$$

Also, because of the following inequality

$$|G(s)H(s)| \leqslant |G(s)|\,|H(s)|$$

We can also guarantee closed-loop stability if

$$|G(s)|\,|H(s)| < 1$$

In essence, the small-gain theorem states that for closed-loop stability, the loop gain must be small. The Nyquist stability criterion can be used to justify the validity of this theorem. Because we are requiring the open-loop transfer function to be inside the unit circle, there can be no encirclements of the $(-1,\ 0)$ point. In addition, we are assuming that the system is open-loop stable; it follows from the Nyquist stability criterion that the system has no closed-loop RHP poles and is therefore closed-loop stable. We should also add that the small-gain theorem guarantees internal stability, so all possible closed-loop transfer functions are stable and all internal signals will remain bounded for bounded inputs.

We can use the small-gain theorem to answer two kinds of question about robust stability. First, given that the uncertainty is stable and bounded, will the closed-loop system be stable for the given uncertainty? Second, for a given system, what is the smallest uncertainty that will destabilize the system? To use the small-gain theorem, it is helpful to convert our system block diagram to a two-block structure, shown in

Figure 10.11. Let us now derive the condition for robust stability under multiplicative uncertainty. Consider the feedback system shown in Figure 10.12(a). To obtain the two-block structure in Figure 10.11, we need to find the transfer function seen by the uncertainty block. The input and output of this transfer function are shown at the indicated points in Figure 10.12(b). It is given by [see Figure 10.12(c)]

$$M(s) = \frac{-G(s)H(s)}{1 + G(s)H(s)}$$

By the small-gain theorem, if the transfer function and the uncertainty transfer function are stable, the closed-loop system will be robustly stable if

$$|\Delta_m| < \frac{1}{\left|GH(1+GH)^{-1}\right|}$$

Condition for robust stability.

Observe that the denominator of the right-hand side of the inequality is the complementary sensitivity, T, so the robust stability condition becomes

$$|\Delta_m| < \frac{1}{|T|}$$

We can use this result to answer the two questions posed earlier. Suppose the stable uncertainty is bounded by

$$|\Delta_m| < \gamma$$

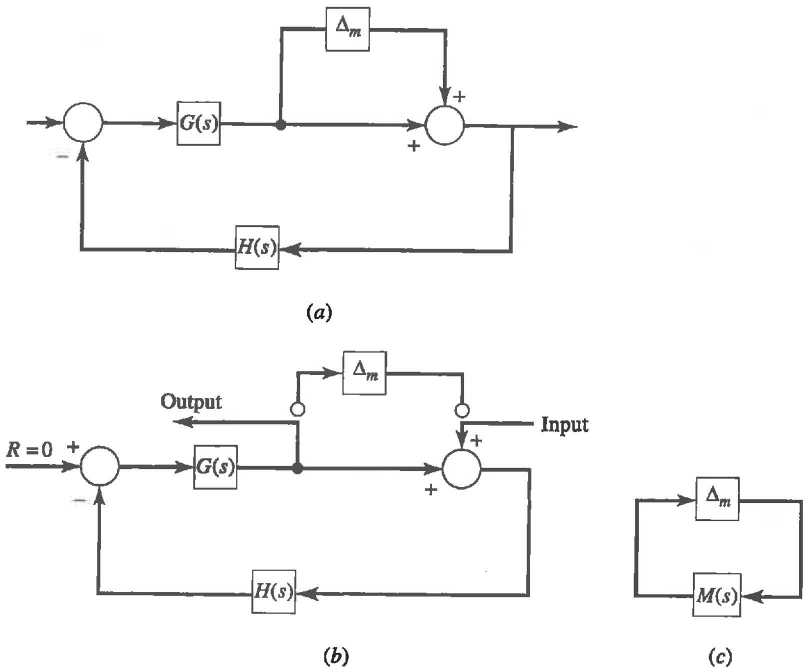


Figure 10.12 (a) Feedback system with multiplicative uncertainty. (b) Obtaining the transfer function seen by the uncertainty. (c) The system as seen by the uncertainty.

Then the closed-loop system will be stable if

$$|T| < \frac{1}{\gamma} \quad \text{or} \quad |\gamma T| < 1$$

To answer the second question, we are interested in finding the size of the smallest stable uncertainty that will destabilize the system. Because the uncertainty must be smaller than $1/T$, it must be smaller than the minimum of $1/T$. To minimize the right-hand side of the inequality, we must maximize T. The maximum of T over all frequencies is its peak value (also called the *resonant peak* in second-order systems: see Figure 10.9). Hence, the smallest destabilizing uncertainty (we call this the *multiplicative stability margin* or MSM) is given by

$$\text{MSM} = \frac{1}{M_r}$$

where

$$M_r = \sup_{\omega} |T(j\omega)|$$

and the symbol "sup" stands for the *supremum* of the function. The supremum (or least upper bound) of a function is its maximum value, even if it is not attained. This is needed for mathematical reasons. We frequently encounter transfer functions that have no maximum. For instance, the following transfer function (a lead network)

$G(s)$ has no maximum, but it has a supremum.

$$G(s) = \frac{s+1}{s+5}$$

has no maximum (if you take the derivative of its magnitude and set it equal to zero, you will get the minimum value of 0.2). However, a glance at its frequency response shows that it approaches the value of 1 as the frequency approaches infinity. But because we never reach the infinite frequency, we never reach the maximum value (although we get very close to it). That is why it does not have a maximum. In these situations, we use the notion of the supremum (or sup for short). We have

$$\sup_{\omega} \frac{|j\omega + 1|}{|j\omega + 5|} = 1$$

The condition for robust stability under additive uncertainty modeling can be derived using the same approach. The transfer function seen by the uncertainty, in this case, is given by (see Figure 10.13)

$$M(s) = \frac{-H(s)}{1 + G(s)H(s)}$$

Hence, the closed-loop system will be robustly stable if

Condition for robust stability (additive case).

$$|\Delta_a| < \frac{1}{\left|H(1 - GH)^{-1}\right|} \quad \text{or} \quad |\Delta_a| < \frac{1}{|HS|}$$

If the uncertainty is stable and bounded by

$$|\Delta_a| < \gamma$$

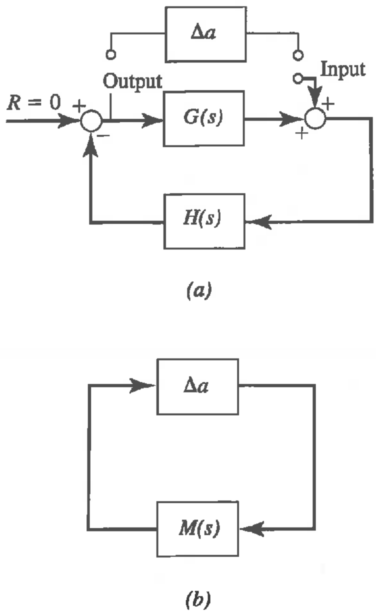


Figure 10.13 (a) Obtaining transfer function seen by the additive uncertainty. (b) The system as seen by the uncertainty.

then we can guarantee closed-loop stability if

$$|HS| < \frac{1}{\gamma} \qquad \text{or} \qquad |\gamma HS| < 1$$

We can also define the *additive stability margin* (ASM) by

$$\text{ASM} = \frac{1}{\sup\limits_{\omega} |H(j\omega)S(j\omega)|}$$

Note that for increased protection against destabilizing multiplicative uncertainties, MSM must be large, implying that the complementary sensitivity must be small. This is compatible with good noise suppression but conflicts with tracking and disturbance rejection. Therefore, small loop gain at high frequencies will protect against multiplicative uncertainties in the high-frequency range. Similarly, observe that the appropriate transfer function for ASM is the same transfer function that determines control energy (actuator limits). Therefore, these requirements are compatible.

Let us apply these results to an example. Consider the following plant and compensator (the compensator is in the feedback path).

$$G(s) = \frac{(5-s)}{(s+5)(s^2+0.2s+1)} \qquad \text{and} \qquad H(s) = \frac{5(s+0.1)}{s}\frac{s+0.2}{s+5}$$

The open-loop Bode plot of the system is shown in Figure 10.14. The system has a phase margin of 38° and a gain margin of 9 dB. This means that a phase lag of 38° or a gain increase factor of 2.8 (9 dB) will destabilize the system. Let us compute the ASM and MSM for the system. For the MSM, we need to obtain the complementary sensitivity and find its peak value. The plot is shown in Figure 10.15. The peak value is 1.52, resulting in an MSM of 0.65. The interpretation of MSM is the following: the

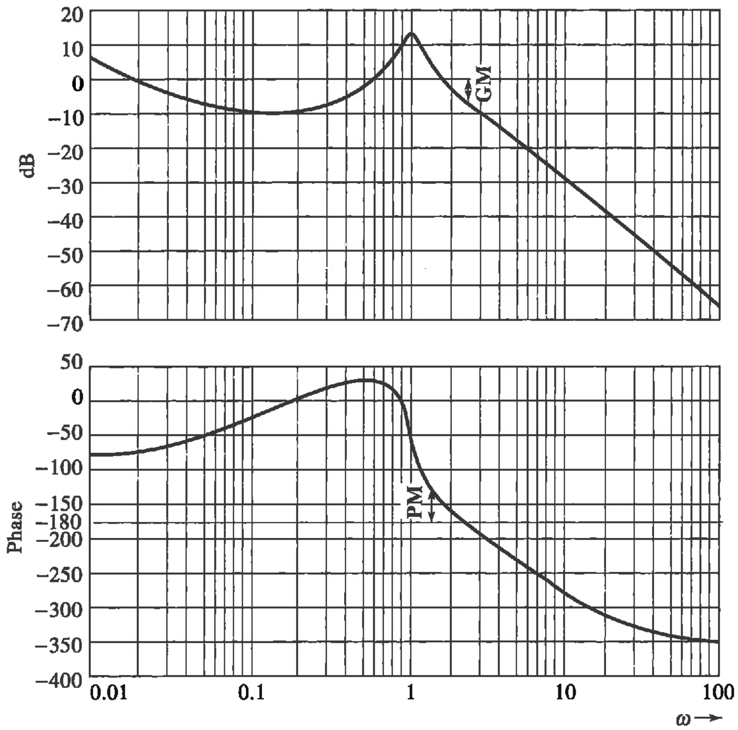


Figure 10.14 Bode plots of $G(s)H(s)$.

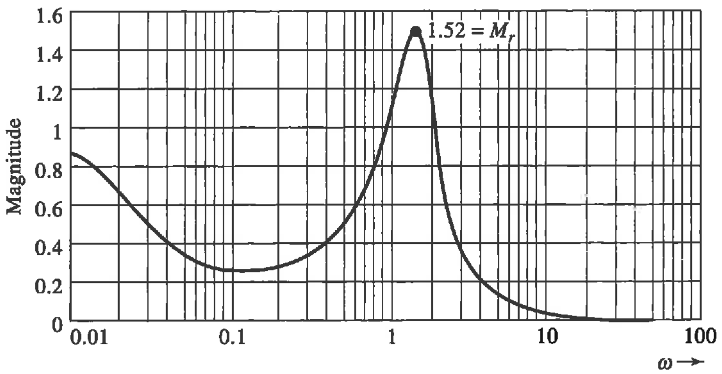


Figure 10.15 Frequency response of the complementary sensitivity for determining the MSM.

system will be robustly stable against unmodeled multiplicative uncertainties with transfer function magnitudes below 0.65. Two points need to be emphasized:

1. The uncertainty can be *any* stable transfer function, provided its magnitude is below our bound.

2. The small-gain theorem is only a sufficient condition (i.e., even if it is violated, the system can still be stable). These stability margins (ASM, MSM) are sometimes very conservative. Hence, the system may be able to tolerate uncertainties that violate the bounds.

To check the conservativeness of the MSM, we modeled the uncertainty by a first-order transfer function and varied its gains. With

$$\Delta_m(s) = \text{MSM}\frac{1 + 0.2k}{s + 1}$$

for $k = 0$, the gain of the uncertainty is the upper limit guaranteed by our theory. We varied k from 0 to 10 in steps of 2, and it was discovered that the system becomes unstable for $k = 8$ or

$$\Delta_m(s) = \frac{1.7}{s + 1} \quad \text{and} \quad 1 + \Delta_m(s) = \frac{s + 2.7}{s + 1}$$

The step responses for these values of k (along with the nominal system (i.e., with no uncertainty) are shown in Figure 10.16. The figure shows that the uncertainty causes oscillations that will eventually lead to instability. Note that the transfer function that is actually in series with the plant is $[1 + \Delta_m(s)]$, which has a maximum destabilizing gain of 2.7. Now, you may wonder why the system is unstable for a gain of 2.7 even though it has gain margin of 2.8. To answer this question, we must be specific about the meaning of ***gain, phase,*** and ***multiplicative stability margins.*** The gain margin is the factor by which the gain can be increased before instability occurs. This assumes no phase change, which implies that the gain margin is a measure of tolerance of pure gain uncertainty. Likewise, the definition of phase margin assumes that the gain is fixed, so phase margin is a measure of tolerance of pure phase uncertainty. MSM, however, allows simultaneous gain *and* phase changes. For example, the gain and phase of $[1 + \Delta_m(s)]$ near the gain phase crossover frequencies are:

At the gain crossover frequency: $\omega = 1.57$, gain $= 1.67$, phase $= -27$
At the phase crossover frequency: $\omega = 2.8$, gain $= 1.30$, phase $= -24$

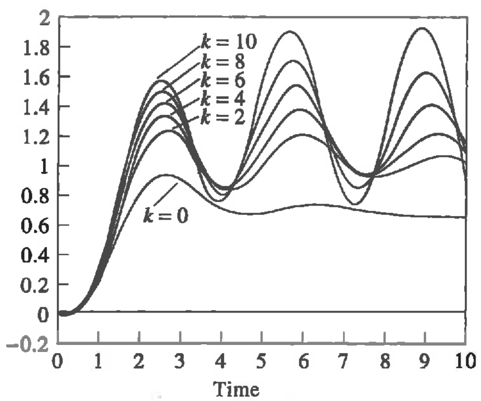


Figure 10.16 Step responses for the uncertain (or perturbed) system.

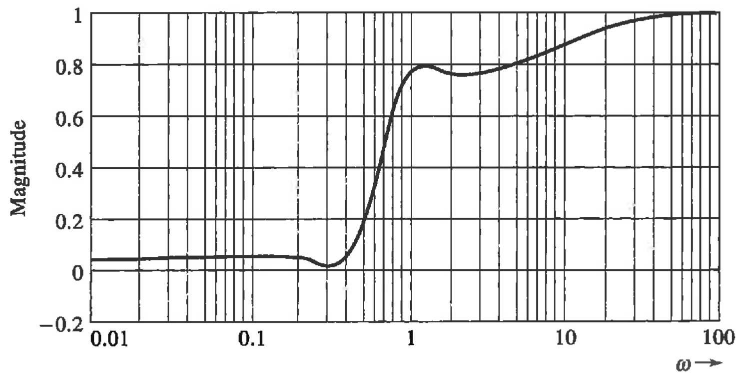


Figure 10.17 Frequency response of $0.2\,H(s)S(s)$ for determining additive robust stability.

Therefore, at the gain crossover frequency, the uncertainty introduces a phase lag of 27° (the phase margin is 38) in addition to a gain increase of 1.67. Also, at the phase crossover frequency, it multiplies the gain by 1.3, and adds 24° of phase lag. That is why it is destabilizing: it is adding both gain and phase lag near the critical frequencies of the system. In a sense, MSM is more general than gain and phase margins. For this reason, MSM is sometimes called *gain-phase margin*.

We can also study the system tolerance to additive uncertainty. We ask whether the system can withstand additive uncertainty transfer functions with magnitude less then $0.2 (\gamma = 0.2)$. To answer this question, we obtain the frequency response of *HS*, and robust stability is guaranteed if

$$|0.2H(j\omega)S(j\omega)| < 1 \qquad \text{for} \qquad |\Delta_a(s)| < 0.2$$

Figure 10.17 shows the response. Because, its peak is less than 1, we conclude that the system is stable in the robust sense, and ASM = 1.

❑ DRILL PROBLEM

D10.4 Consider the double-integrator plant compensated by a feedback lead compensator.

$$G(s) = \frac{1}{s^2} \qquad H(s) = \frac{20(s+1)}{s+10}$$

Suppose the actual plant model contains an additive uncertainty given by

$$\Delta_a(s) = \frac{-1}{s^2+s+1}$$

(a) Determine if the compensator $H(s)$ stabilizes the plant $G(s)$.

(b) Determine if the compensator $H(s)$ stabilizes the actual plant given by $G_a(s) = G(s) + \Delta_a(s)$.

(c) Find $M(s)$, the transfer function "seen" by the uncertainty.

(d) Draw Bode plots of $\Delta_a(s)$ and $M(s)$ to determine the robust stability of the system.

Ans. (a) characteristic equation $= s^3 + 10s^2 + 20s + 20$, stable;

(b) characteristic equation $= s^5 + 11s^4 + 11s^3 + 50s^2 + 80s + 40$, unstable;

(c) $M(s) = [20s^2(1+s)]/(s^3 + 10s^2 + 20s + 20)$;

(d) system not robustly stable

10.6 Loop Transfer Recovery (LTR)

It was discussed earlier that the LQR solution has excellent stability margins (infinite gain margin and 60° phase margin); we know that LQR is usually, but not always, considered impractical because it requires that all states be available for feedback. Doyle and Stein showed that under certain conditions, the LQG can asymptotically recover the LQR properties. One of the problems with LQG is that it requires statistical information of the noise processes. In most cases, however, this information is not available or is impractical to obtain. Mathematical arguments and simulations had shown that the LQG design parameters (Q, R, Q_o, and R_o) have a strong influence on the performance of the system. It was suggested that because Q_o, and R_o are not usually available, they should be used as tuning parameters to improve system performance.

Consider the block diagrams in Figure 10.18. With the loop broken at the indicated point, the (open) loop transfer function of the LQR is given by

LQR.

$$L(s) = K_{\Phi(s)}B$$

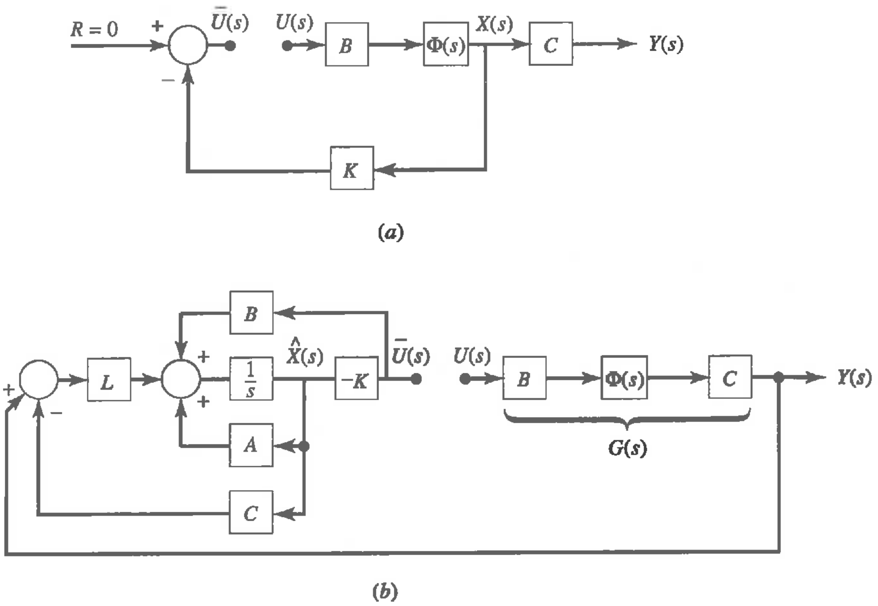


Figure 10.18 (a) Block diagram of an LQR controller. (b) Block diagram of an LQG controller.

where

$$\Phi(s) = (sI - A)^{-1}$$

The loop transfer function for LQG is likewise given by

LQG.

$$\underset{LQG}{L(s)} = K(sI - A + BK + LC)^{-1}LC\Phi(s)B$$

Under the following two conditions

LTR assumptions.

$G(s)$ is minimum-phase (i.e., it has no zeros in the RHP)
$R_o = 1$ and $Q_o = q^2BB'$

it can be shown that

$$\lim_{q\to\infty} \underset{LQG}{L(s)} = L(s)$$

The following procedure for design is suggested by the foregoing conditions. Choose the LQR parameters such that the LQR loop transfer function (also called the *target feedback loop* or *TFL*) has desirable time and/or frequency domain properties. Design an observer with parameters specified in condition 2. Increase the tuning parameter q until the resulting loop transfer function is as close as possible to the TFL.

In many situations, the variable that is measured is different from the variable we want to control. For example, we may desire to control thrust in a jet engine, but we can sense only temperature and turbine speed. Let y denote the measured states, and z denote the controlled states, then

$$y = Cx \quad \text{and} \quad z = C_q x$$

Because the loop transfer function of LQG approaches that of LQR, it will asymptotically recover its properties. A more detailed procedure follows.

Loop Shaping Step

1. Determine the controlled variable and set

$$Q = C'C \quad \text{or} \quad Q = C_q'C_q$$

2. Convert the design specifications into a desired TFL. At this stage, if the system is type 0 and we want a type 1 system, we can add an integrator to the system.
3. Vary the parameter R until the resulting loop transfer function is similar to the TFL. One may use the RSL approach here. Also, check the sensitivity and complementary sensitivity transfer functions (S and T), to make sure they have desirable shapes.

Recovery Step

4. Select a scalar, q, and solve the filter Riccati equation

$$A\Sigma + \Sigma A' + q^2BB' - \Sigma C'C\Sigma = 0 \quad \text{and set } L = \Sigma C'$$

Increase q until the resulting loop transfer function is close to the TFL.

The higher the value of q, the closer the LQG system comes to LQR performance. It should be noted that the value of q should not be increased indefinitely, because this may lead to unreasonably large values for the filter gain L. Also, because LQR has -20 dB slope at high frequencies, large values for q will also recover this slow roll-off rate. Smaller values for q will tend to trade off lower stability margins with higher roll-off rates at high frequencies.

LTR solution of the double integrator system.

We will now use LTR on the double-integrator system to recover the LQR properties. Because LTR requires solving the Riccati equation a number of times, the problem must be solved on the computer. First, we chose the LQR loop transfer function as the TFL. Therefore, our objective is to recover the Bode plots shown in Figure 10.1. We next let the parameter q vary over the range (1, 10, 100, 1,000). The plots for the closed-loop step response and open-loop Bode plots for the LQR case and LTR, for the specified values of q, are shown in Figure 10.19.

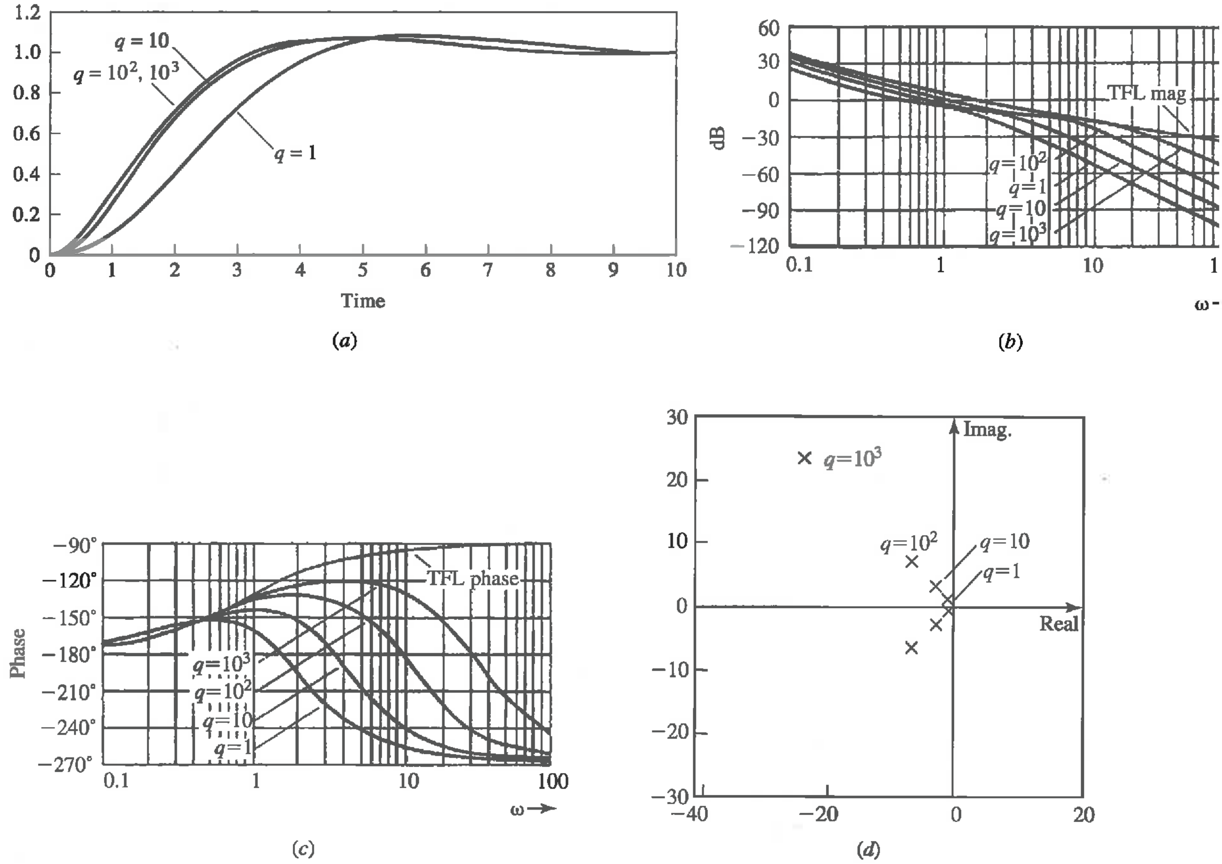


Figure 10.19 Step response, Bode plots, and filter oples for LTR using $q = (1, 10, 100, 1000)$. (a) Closed-loop step response. (b) Open-loop magnitude Bode plot.(c) Open-loop phase Bode plot. (d) Filter poles.

Table 10.2 *Results of LTR Design*

q	1	10	100	1000
PM	32.6	41.9	55.0	61.7
GM	9.5	13.0	21.1	30.4
L	1.4	4.5	14.1	44.7
	1.0	10.0	100.0	1000.0
Filter	$-0.7+0.7j$	$-2.2+2.2j$	$-7.0+7.0j$	$-22.3+22.3j$
poles	$-0.7-0.7j$	$-2.2-2.2j$	$-7.0-7.0j$	$-22.3-22.3j$

Note how the step response approaches the LQR case for increasing values of q. Also, as q increases, the low frequency gain of the system goes from 28 dB to 40 dB while the high frequency gain goes from −110 dB to −40. The values for the filter gain L, its eigenvalues, and the stability margins (*GM* and *PM*) are given in Table 10.2. The data show that the LQR phase margin is recovered. The gain margin increases from 10 dB to 30 dB. This can clearly be increased by increasing q. But note that increasing the margins will cost us in terms of higher values for the filter gain L, higher gain crossover frequency, and lower high-frequency gain. This will make the system more sensitive to noise and uncertainties at high frequencies. It appears that a value of q between 100 to 1000 is a reasonable compromise.

Note that the procedure uses the machinery of LQG (i.e., two Riccati equations) and its guaranteed stability. However, it allows us to work strictly with Bode plots of various transfer functions and to satisfy frequency domain measures (similar to classical control). Therefore, it can be considered a frequency domain design procedure that uses state-space equations for computation. This is the common feature of control system design after LQR/LQG, sometimes called *postmodern control* (i.e., frequency domain techniques that use state space machinery for computation).

❑ DRILL PROBLEM

D10.5 Consider the following system:

$$\dot{x} = Ax + Bu + \omega$$

$$y = cx + v$$

$$A = \begin{bmatrix} 0 & 1 \\ -3 & -4 \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad c = [2 \quad 1]$$

Let $Q = E'E$ where $E = 4\sqrt{5}[\sqrt{35} \quad 1]$

$$R = 1$$

$$Q_o = \begin{bmatrix} 35 \\ -61 \end{bmatrix} [35 \quad -61]$$

$$R_o = 1$$

(a) Find an LQR compensator $H(s)$. Compute K, the closed-loop poles, and the gain and phase margins.

(b) Design an LQG compensator. Compute the gain and phase margins in this case. Compare with (a).

(c) Design an LTR compensator by solving the following Riccati equation:

$$A\Sigma + \Sigma A' + (Q_o + q^2 BB') - \Sigma C'C\Sigma = 0$$

Increase the value of q to 10^6 and observe the effects. Compute the gain and phase margins for $q = 100$.

Ans. (a) $K = [50 \quad 10], -7 \pm j2, \ GM = \infty, \ PM = 85°$;

(b) $PM < 15°, \ GM = 6.7$ dB;

(c) $PM = 74°, \ GM = 37$ dB

10.7 H_∞ Control

10.7.1 A Brief History

One of the major challenges in control has been the analysis and design of multivariable (multiple-input, multiple-output or MIMO) control systems. This is a difficult problem, because the transfer function of a MIMO system is a matrix. Even very basic concepts such as system order, poles, and zeros run into difficulty in this case. For instance, there are at least five to ten different definitions of zeros of a multivariable system! Successful concepts and tools of classical control such as root locus, Bode plots, Nyquist stability criterion, and gain and phase margins ran into difficulty. State space techniques, based in the time domain, avoided the complexities of transfer function matrices and provided tools for analysis and design of MIMO systems. Within the state space framework, the only difference between a SISO system and a MIMO system is the number of columns of the B matrix (number of inputs) and the number of rows in the C matrix (number of outputs). Note that in all the techniques we have discussed, these dimensions play no part. In fact, the most important feature of LQR/LQG is that they are systematic methods for designing MIMO systems.

At the about the same time that most researchers were developing, extending, and refining time domain optimal control methods. Other researchers, mostly in Britain, (A. G. J. MacFarlane and H. H. Rosenbrock), were busy extending classical control tools to the multivariable case. They were largely successful in these endeavors. Classical tools such as root locus (renamed *characteristic locus*), Nyquist techniques (renamed *Nyquist arrays*), and Bode plots (renamed *singular value plots*) were extended to the multivariable case. As the shortcomings of LQG methods became more apparent in the 1970s, more attention was paid to classical control concepts and concerns.

During the 1980s a new paradigm emerged, H_∞ control. This control problem was first formulated by G. Zames. It was essentially a frequency domain optimization method for designing robust control systems. Robustness became the main concern in the control community, and other techniques for designing robust multivariable control systems soon followed. They are H_∞ control, μ synthesis (by J. Doyle, also called k_m synthesis by M. Safonov), QFT (quantitative feedback theory) by I. Horowitz, and methods based on Kharitonov's theorem for structured uncertainty. All these techniques are still being developed and refined today.

Our purpose in this section is to present a brief introduction to H_∞ control. Although this is a powerful technique for the MIMO case, our presentation is limited to the SISO case. The transition to the MIMO case is straightforward in theory but not necessarily in practice.

10.7.2 Some Preliminaries

H_∞ control has developed its own terminology, notation, and paradigm. For example, the classical block diagram has been modified to handle problems of more general types. Also, because the design equations are very lengthy, some shorthand notation is introduced to simplify the presentation. Because these notations have become standard in the literature, and because they could be confusing to the novice, we will introduce and use them in this discussion to ease the transition to more advanced books and the literature for the readers.

We first discuss the name. H_∞ refers to the space of stable and proper transfer functions. We generally desire that the closed-loop transfer functions be proper (i.e., the degree of the denominator $\geqslant$ the degree of the numerator) and stable (poles strictly in the LHP). Instead of repeating these requirements, we say $G(s)$ is in H_∞. The basic object of interest in H_∞ control is a transfer function. In fact, we will be optimizing over the space of transfer functions. Optimization presupposes a cost (or objective) function, because we want to compare different transfer functions and choose the best one in the space. In H_∞ control, we compare transfer functions according to their H_∞ norm (a mathematical term for the concept of size). The H_∞ norm of a transfer function is defined by

Measuring size of transfer functions.

$$\|G\|_\infty = \sup_\omega |G(j\omega)|$$

This is easy to compute graphically: it is simply the peak in the Bode magnitude plot of the transfer function (it is finite when the transfer function is proper with no imaginary poles). We have already seen this quantity before in the section on robust stability. For instance, the multiplicative stability margin (MSM) can be written

$$\text{MSM} = \frac{1}{\|T\|_\infty}$$

As an example

$$\left\|\frac{1}{s+1}\right\|_\infty = 1$$

In H_∞ control, the objective is to minimize the H_∞ norm of some transfer function, so we will try to minimize the peak in the Bode magnitude plots.

A notation that is rapidly becoming popular is the *packed-matrix* notation for representing transfer functions in state space. Recall that the transfer function of a system with state-space matrices $\{A,\ B,\ C,\ D\}$ is given by

$$G(s) = C(sI - A)^{-1}B + D$$

This transfer function in packed-matrix notation is written

$$G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

New notation: packed-matrix.

We emphasize that this is not a matrix in the ordinary sense; it is just a shorthand notation for the foregoing expression for $G(s)$. For example, the transfer function of an LQG compensator, given in Section 10.4, can be expressed as follows:

$$H(s) = \left[\begin{array}{c|c} A - BK - LC & L \\ \hline K & 0 \end{array}\right]$$

The solution to the H_∞ control problem contains very messy Riccati equations, so the following notation is introduced to simplify solution representation. Consider the following Riccati equation:

$$A'X + XA - XRX + Q = 0$$

The stabilizing solution of this equation will be denoted by $X = \text{Ric}\ (H)$, where H is the following *Hamiltonian* matrix:

The Ric notation.

$$H = \begin{bmatrix} A & -R \\ -Q & -A' \end{bmatrix} \quad \text{and} \quad (A - RX) \text{ is stable}$$

Instead of writing the Riccati equation, we will specify its associated Hamiltonian matrix and the reader can create the appropriate Riccati equation.

Finally, we introduce a more general block diagram representation of control systems shown in Figure 10.20. This new diagram is able to represent a variety of problems of interest. The diagram contains two main blocks, the plant and the controller. The plant section has two inputs and two outputs. The plant inputs are classified as control inputs and exogenous inputs. The control input u is the output of the controller, which becomes the input to the actuators driving the plant. The exogenous input w is actually a collection of inputs (a vector). The main distinction between w and u is that the controller cannot manipulate these exogenous inputs. Typical inputs that are lumped into w are external disturbances, noise from the sensors, and tracking or command signals.

The plant outputs are also categorized in two groups.The first group, y, are signals that are measured and fed back. These become the inputs to the controller. The second group, z, are the regulated outputs. These are all the signals we are interested to control or regulate. They could be states, error signals, and control signals. Even if the original

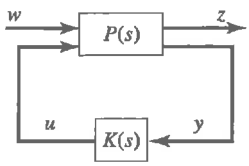


Figure 10.20 Generic block diagram for H_∞ control.

system is SISO (i.e., u and$y$ are scalars), the new formulation is essentially MIMO. Most realistic control system problem formulations are of the MIMO kind.

A transfer function representation of the system is given by

$$\begin{aligned} z &= P_{zw}w + P_{zu}u \\ y &= P_{yw}w + P_{yu}u \\ u &= Ky \end{aligned}$$

(Note the change in notation: although the inputs and outputs are s-domain quantities, we write them in lowercase letters and omit the dependency on s; as a rule we will use lowercase italic letters for scalars and vectors, and capital letters for matrices in our presentation.)

The closed-loop transfer function between the regulated outputs and the exogenous inputs is obtained as follows. First, we substitute for u in the equation for y.

$$y = P_{yw}w + P_{yu}Ky$$

then, we solve for y (note that all capital letters are matrices, so we have to use matrix inversion and watch for the order of multiplication)

$$(I - P_{yu}K)y = P_{yw}w \rightarrow y = (I - P_{yu}K)^{-1}P_{yw}w$$

Therefore, u becomes

$$u = Ky = K(I - P_{yu}K)^{-1}P_{yw}w$$

Substituting this into the equation for z, we get

$$z = P_{zw}w + P_{zu}K(I - P_{yu}K)^{-1}P_{yw}w = [P_{zw} + P_{zu}K(I - P_{yu}K)^{-1}P_{yw}]w$$

Finally

The linear fractional transformation.

$$z = T_{zw} \qquad \text{where } \mathrm{T}_{zw} = P_{zw} + P_{zu}K(I - P_{yu}L)^{-1}P_{yw}$$

This expression for the closed-loop transfer function of P and K is called the *linear fractional transformation.*

The plant can also be represented in state-space form as follows:

$$\begin{aligned} \dot{x} &= Ax + B_1w + B_2u \\ z &= C_1x + D_{11}w + D_{12}u \\ y &= C_2x + D_{21}w + D_{22}u \end{aligned}$$

Using the packed-matrix notation, we get

New plant representation.

$$P(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right]$$

10.7.3 H_∞ Control: Solution

The H_∞ control problem is formulated as follows. Consider the generic block diagram in Figure 10.20 and find an internally stabilizing controller, $K(s)$, for the plant $P(s)$, such that the infinity norm of the closed-loop transfer function, T_{zw}, is below a given level γ (a positive scalar). This problem is called the *standard H_α control problem.* The *optimal H_∞ control problem* is

$$\text{Optimal problem} \quad \underset{K(s)\text{stabilizing}}{\text{Min}} \quad \|T_{zw}\|_\infty$$
$$\text{Standard problem} \quad \underset{K(s)\text{stabilizing}}{\text{Find}} \quad \|T_{zw}\|_\infty \leqslant \gamma$$

The standard problem is more practical. In practice, control system design is more like a balancing act and trade-offs, and a mathematically optimal solution may not be so desirable after all the other real-life constraints have been taken into account. To solve the optimal problem, we start with a value for γ and reduce it until the problem fails to have a solution. As a starting value for γ, we can solve an LQG problem; find the peak in the resulting closed-loop transfer function and use this value. To lower γ, we can use a search algorithm (such as a binary search) to reach the optimal value. This procedure is called *γ-iteration.*

For the problem to have a solution, certain assumptions must be satisfied. They are listed below. The dimensions of various variables are listed first.

Dimensions: $\dim x = n$, $\dim w = m_1$, $\dim u = m_2$, $\dim z = p_1$, $\dim y = p_2$

H_∞ problem assumptions.

1. The pair (A, B_2) is stabilizable and (C_2, A) is detectable. Recall from Chapter 8 that these are weaker versions of controllability and obsevability conditions. This assumption is necessary for a stabilizing controller to exist. It simply guarantees that the controller can reach all unstable states and these states show up on the measurements.
2. Rank $D_{12} = m_2$, rank$D_{21} = p_2$. These conditions are needed to ensure that the controllers are proper. They also imply that the transfer function from w to y is nonzero at high frequencies. Unlike the first assumption, which is usually satisfied, this assumption is frequently violated (e.g., if the original plant is strictly proper: i.e., if it has more poles than zeros, this condition will be violated) unless the problem is formulated in a way that ensures its satisfaction.
3. Rank$\begin{bmatrix} A - j\omega I & B_2 \\ C_1 & D_{12} \end{bmatrix} = n + m_2$ for all frequencies.
4. Rank$\begin{bmatrix} A - j\omega I & B_1 \\ C_2 & D_{21} \end{bmatrix} = n + p_2$ for all frequencies.
5. $D_{11} = 0$ and $D_{22} = 0$. This assumption is not needed, but it will simplify the equations for the solution. It also implies that the transfer functions from w to z and u to y roll off at high frequencies, respectively.

Before we present the solution, it should be pointed out that the solutions of the H_∞ and LQG problems are very similar. Both use a state estimator and feed back the estimated states. The controller and estimator gains are also computed from two

Riccati equations. The differences are in the coefficients of the Ricatti equations, and the H_∞ state estimator contains an extra term. The compensator equations follow.

The controller is given by, K_c corresponds to K, the controller gain, in the LQG case

$$u = -K_c\hat{x}$$

and the state estimator is given by

$$\dot{\hat{x}} = A\hat{x} + B_2u + B_1\hat{w} + Z_\infty K_e(y - \hat{y})$$

where

$$\hat{w} = \gamma^{-2}B_1'X_\infty\hat{x}$$

$$\hat{y} = C_2\hat{x} + \gamma^{-2}D_{21}B_1'X_\infty\hat{x}$$

We can also write this in packed-matrix notation as follows:

And finally the solution.

$$K(s) = \left[\begin{array}{c|c} A - B_2K_c - Z_\infty K_eC_2 + \gamma^{-2}(B_1B_1' - Z_\infty K_eD_{21}B_1')X_\infty & Z_\infty K_e \\ \hline -K_c & 0 \end{array}\right]$$

The extra term, $\hat{w}$, is an estimate of the worst-case input disturbance to the system, and $\hat{y}$ is the output of the estimator. The estimator gain is $Z_\infty K_e(K_e$ corresponds to L in the LQG case). The controller gain K_c and estimator gain K_e are given by

$$K_c = \tilde{D}_{12}(B_2'X_\infty + D_{12}'C_1) \quad \text{where } \tilde{D}_{12} = (D_{12}'D_{12})^{-1}$$

$$K_e = (Y_\infty C_2' + B_1D_{21}')\tilde{D}_{21} \quad \text{where } \tilde{D}_{21} = (D_{21}D_{21}')^{-1}$$

The term Z_∞ is given by

$$Z_\infty = (I - \gamma^{-2}Y_\infty X_\infty)^{-1}$$

The terms X_∞ and Y_∞ are solutions to the controller and estimator Riccati equations—that is,

The Riccati equations.

$$X_\infty = \text{Ric}\begin{bmatrix} A - B_2\tilde{D}_{12}D_{12}'C_1 & \gamma^{-2}B_1B_1' - B_2\tilde{D}_{12}B_2' \\ -\tilde{C}_1'\tilde{C}_1 & -(A - B_2\tilde{D}_{12}D_{12}'C_1') \end{bmatrix}$$

$$Y_\infty = \text{Ric}\begin{bmatrix} (A - B_1D_{21}'\tilde{D}_{21}C_2)' & \gamma^{-2}C_1'C_1 - C_2'\tilde{D}_{21}C_2 \\ -\tilde{B}_1\tilde{B}_1' & -(A - B_1D_{21}'\tilde{D}_{21}C_2) \end{bmatrix}$$

where $\tilde{C}_1 = (I - D_{12}\tilde{D}_{12}D_{12}')C_1$ and $\tilde{B}_1 = B_1(I - D_{21}'\tilde{D}_{21}D_{21})$.

The closed-loop system becomes

$$\begin{bmatrix} \dot{x} \\ \dot{\hat{x}} \end{bmatrix} = \begin{bmatrix} A & -B_2K_c \\ Z_\infty K_eC_2 & \begin{array}{c} A - B_2K_c + \gamma^{-2}B_1B_1'X_\infty \\ -Z_\infty K_e(C_2 + \gamma^{-2}D_{21}B_1'X_\infty) \end{array} \end{bmatrix}\begin{bmatrix} x \\ \hat{x} \end{bmatrix} + \begin{bmatrix} B_1 \\ Z_\infty K_eD_{21} \end{bmatrix} w$$

$$\begin{bmatrix} z \\ y \end{bmatrix} = \begin{bmatrix} C_1 & -D_{12}K_c \\ C_2 & 0 \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix} + \begin{bmatrix} 0 \\ D_{21} \end{bmatrix} w$$

Closed-loop system.

As we had promised, the equations are quite complicated and messy!

Finally, it can be proved that there exists a stabilizing compensator if and only if there exist positive semidefinite solutions to the two Riccati equations and the following condition:

$$\rho(X_\infty Y_\infty) < \gamma^2$$

where $\rho(A)$ = spectral radius of A = largest eigenvalue of $A = \lambda_{\max}(A)$.

The block diagrams of the LQG and H_∞ control systems are shown in Figure 10.21. Compare these diagrams to see the similarities and differences between them.

It should be fairly obvious that H_∞ problems cannot be solved manually. Computer programs such as Program CC, MATRIX_X, and MATLAB have special functions and utilities for solving these problems. For every value of γ, two Riccati equations must be solved. In addition, even if the plant is first-order, we still may need to add weights to the system either to satisfy design requirements or to satisfy the necessary assumptions for a feasible solution. This increases the order of the equations and makes manual solution almost impossible. The steps can be summarized as follows.

Solution procedure.

1. Set up the problem to obtain the state space representation for $P(s)$.
2. Check to see whether the assumptions (the rank conditions) are satisfied. If they are not, reformulate the problem by adding weights or adding (fictitious) inputs or outputs.
3. Select a large positive value of γ.
4. Solve the two Riccati equations. Determine if the solutions are positive semidefinite; also, verify that the spectral radius condition is met.
5. If all the conditions given are satisfied, lower the value of γ. Otherwise, increase it. Repeat steps 4 and 5 until either an optimal or satisfactory solution is obtained.

10.7.4 Weights in H_∞ Control Problems

Practical control problems require weighting the inputs and outputs. There are a few reasons for using weights. Constant weights are used for scaling inputs and outputs, they are also used for unit conversions. Transfer function weights are used to shape the various measures of performance in the frequency domain. In H_∞ control problems, weights are also used to satisfy the rank conditions. These assumptions are frequently violated unless appropriate weights are selected. In fact, the weights are the only parameters that the designer must specify. Proper selection of these weights depends a great deal on the experience of the user and on his or her understanding of the physics of the problem and other practical engineering constraints.

Tracking and disturbance rejection require that the sensitivity transfer function be small in the low-frequency range. This can be formulated as specifying that the

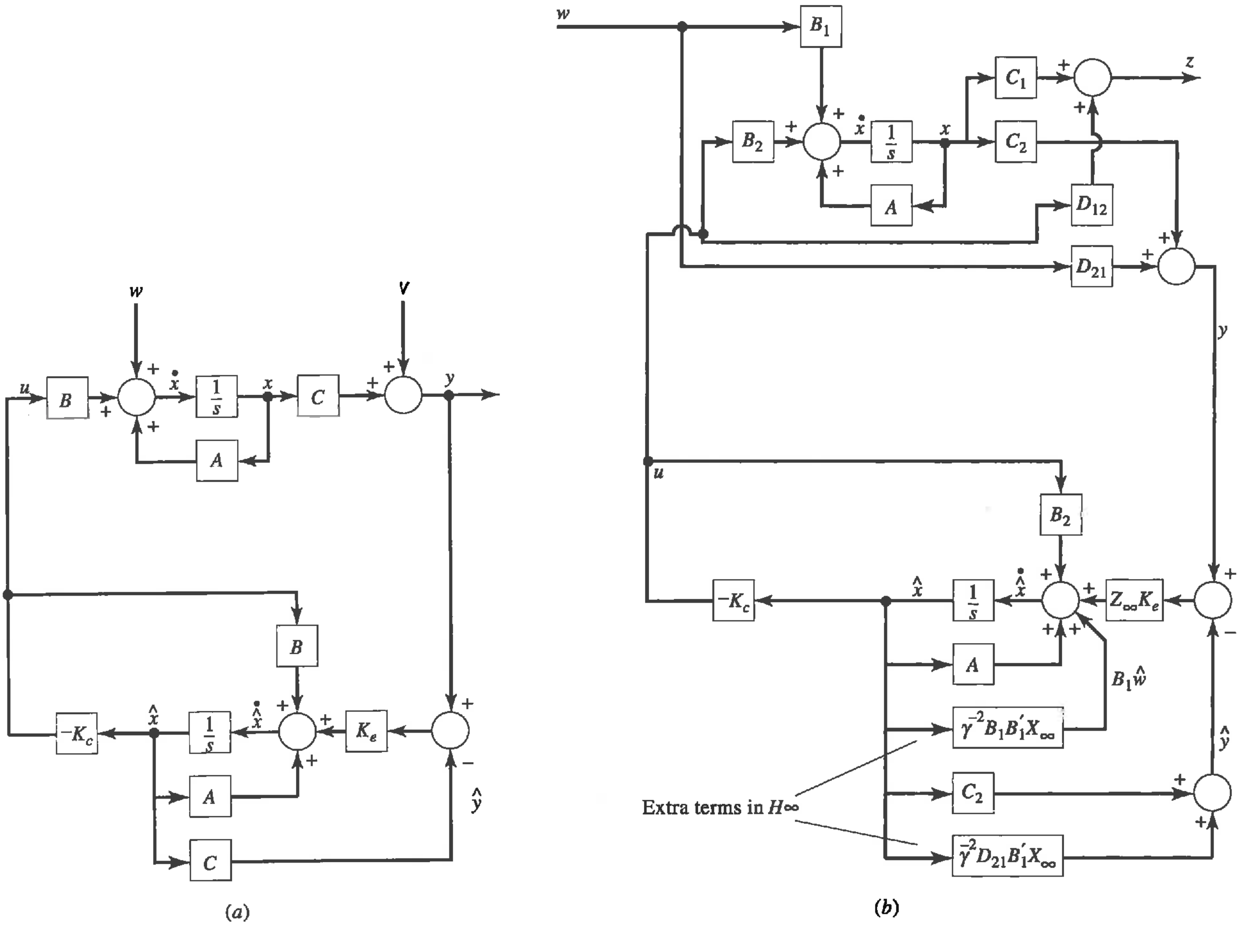


Figure 10.21 (a) The LQG controller block diagram. (b) Block diagram showing the structure of the H_∞ control system.

Weights give us frequency domain control over system behavior.

sensitivity remain below a given frequency-dependent weight—that is,

$$|S| \leqslant W_s^{-1} \quad \text{or} \quad |W_s S| \leqslant 1$$

Similarly we can specify that the complementary sensitivity be kept below a given weight in the high-frequency range—that is,

$$|T| \leqslant W_t^{-1} \quad \text{or} \quad |W_t T| \leqslant 1$$

Finally, both requirements can be satisfied by solving what is called the *mixed-sensitivity problem*. Typical plots for both cases including weights are shown in Figure 10.22.

As an example, we will use the H_∞ approach to design a controller for the double-integrator system. The first step is setting up the problem appropriately. The

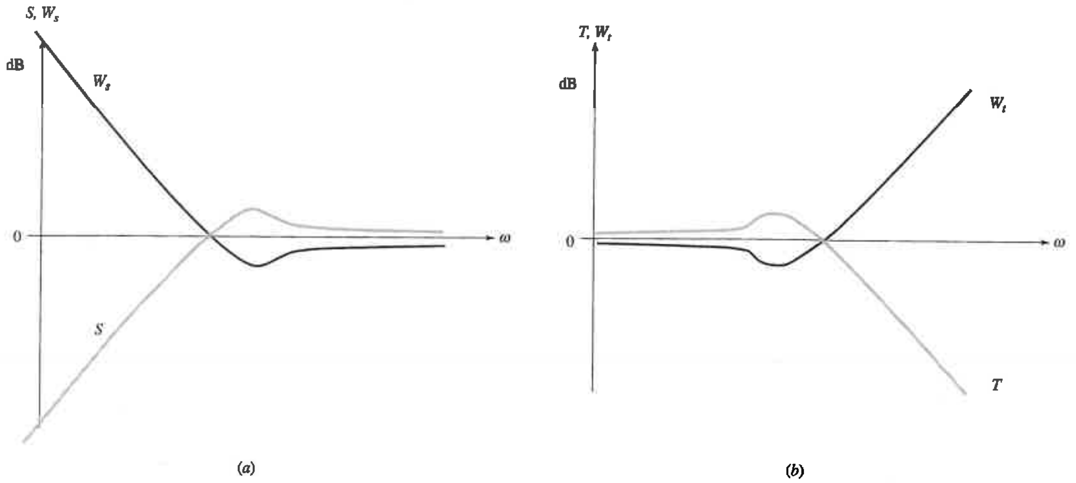


Figure 10.22 (a) Plot of sensitivity function and its weight. (b) Plot of the complementary sensitivity and its corresponding weight.

plant equations are given by

$$\dot{x}_1 = d + u$$
$$\dot{x}_2 = x_1$$

H_∞ solution of the double-integrator.

The new term that we have added is the disturbance term, d; this term corresponds either to an actual disturbance or to unmodeled dynamics in the system. The regulated outputs are given by

$$z = \begin{bmatrix} x_2 \\ u \end{bmatrix}$$

It is important that the control signal be included in the regulated outputs so that we can bound its magnitude. This is also needed to ensure that the rank condition on D_{12} is satisfied. The measurement equation is given by

$$y = x_2 + n$$

The noise term, n, is either actual sensor noise or, perhaps, it represents high-frequency unmodeled dynamics. It is also needed to ensure that the rank condition on D_{21} is met. Collecting these equations, we obtain the system equations in packed-matrix notation as follows:

$$P(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right] = \left[\begin{array}{cc|cc:c} 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ \hdashline 0 & 1 & 0 & 1 & 0 \end{array}\right]$$

The block diagram of the system is shown in Figure 10.23 in the usual form and in the generic H_∞ form.

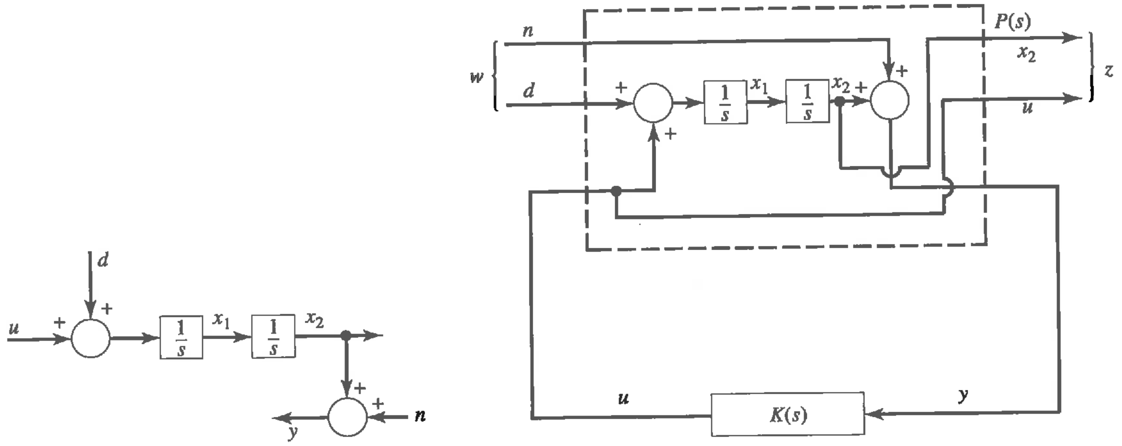


Figure 10.23 (a) The block diagram for the double-integrator system. (b) The generic H_∞ block system.

This problem was solved on the computer, and after several trials we found that the value of γ could not be reduced below 2.62. Hence, we conclude that 2.62 is the optimal value (note that the solution of the optimal H_∞ control problem involves a search over γ and we can get as close to it as possible but not achieve it). The following are the relevant data obtained:

$$\gamma = 2.62 \qquad X_\infty = \begin{bmatrix} 1.59 & 1.08 \\ 1.08 & 1.47 \end{bmatrix} \qquad Y_\infty = \begin{bmatrix} 1.47 & 1.08 \\ 1.08 & 1.59 \end{bmatrix}$$

$$K_c = [\,1.59 \quad 1.08\,] \qquad K_e = \begin{bmatrix} 1.08 \\ 1.59 \end{bmatrix}$$

The transfer function of the compensator and the resulting closed-loop poles are given by

The H_∞ compensator. Looks like a lead plus an extra pole.

$$K(s) = \frac{-578.3(s+0.39)}{(s+2.33)(s+220.72)}$$

Closed-loop poles = $\{-0.71, -0.81 \pm 0.91j, -200.7\}$

The Bode and Nyquist plots of the system are shown in Figure 10.24(a, b). We have obtained a gain margin of 44 dB and phase margin of 45°. The responses of the system to a unit-pulse disturbance and random sensor noise also is shown in Figure 10.24 (c). As expected, both responses approach zero asymptotically.

We will end our brief introduction to H_∞ control at this point. We point out that this subject is still very novel and is rapidly progressing. We have also limited our discussion to the treatment of unstructured uncertainty and have presented only one of the approaches to robust control. The different approaches, however, have one feature in common: They are all frequency domain, computer-assisted tools for

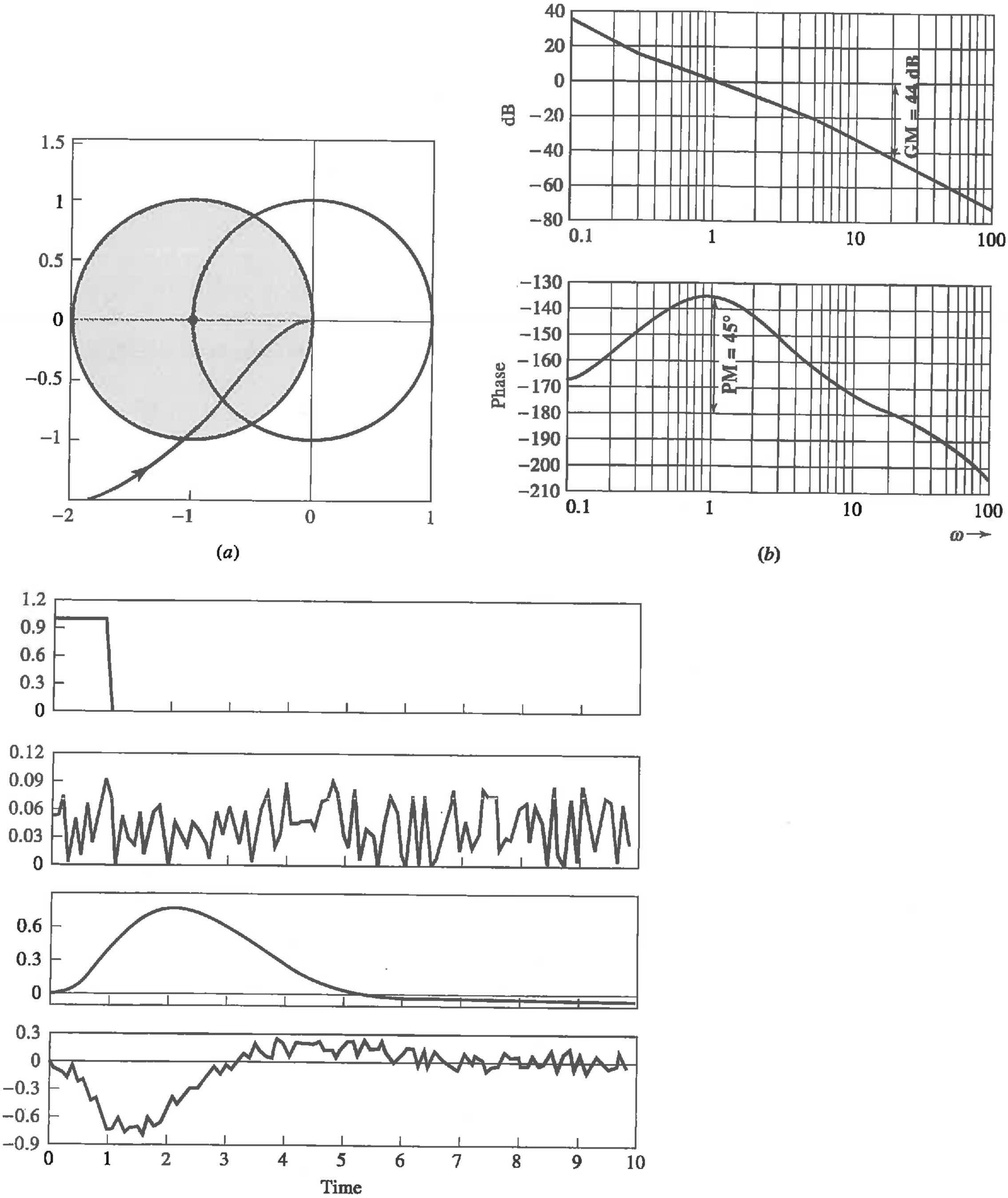


Figure 10.24 (a) The Nyquist plot of the H_∞ compensated system. (b) The Bode plots of the system. (c) Unit pulse, random noise, and plant response to these inputs.

design of MIMO systems satisfying practical constraints. For this reason, H_∞ control is expected to find a permanent place in the control engineer's toolbox.

❑ DRILL PROBLEMS

D10.6 (*Note:* As much as possible, try to solve this problem manually.) Consider the first-order system given by

$$\dot{x} = x + u + d$$
$$y = x + n$$

Our objective is to regulate the state and control signals (x, u) in presence of disturbance and noise inputs (d, n).

Let $w = \begin{bmatrix} d \\ n \end{bmatrix}$ and $z = \begin{bmatrix} x \\ u \end{bmatrix}$

(a) Obtain the plant equation, $P(s)$, in packed-matrix notation.

(b) Draw the generic H_∞ block diagram of the system.

(c) Verify that all the rank conditions are met.

(d) Compute $\tilde{D}_{12}, \tilde{D}_{21}, \tilde{C}_1, \tilde{B}_1$.

(e) Find the controller Hamiltonian matrix and solve for X_∞ in terms of γ.

(f) Repeat (e) for the estimator, and solve for Y_∞.

(g) Compute $(X_\infty Y_\infty)$ in terms of γ. Find out if the spectral radius condition is met for $\gamma = 2$. Repeat for $\gamma = 3$.

For the rest of the drill, let $\gamma = 3$.

(h) Compute X_∞, Y_∞, Z_∞, K_c, K_e.

(i) Find the compensator transfer function and the closed-loop poles.

(j) Draw the Nyquist plot and obtain gain and phase margins.

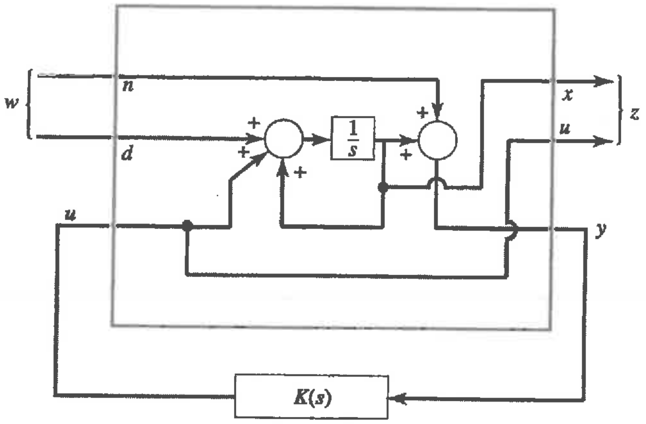
Ans.

(a) $\dot{x} = x + [1 \quad 0]w + u$

$$z = \begin{bmatrix} 1 \\ 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \qquad P(s) = \left[\begin{array}{c|cc|c} 1 & 1 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 1 & 0 \end{array}\right]$$

$y = x + [0 \quad 1]w$

(b)



(c) Rank D_{12} =rank $\begin{bmatrix} 0 \\ 1 \end{bmatrix} = 1$, rank D_{21} = rank $[0 \quad 1] = 1$,

$D_{11} = 0, D_{22} = 0$ $\qquad$ rank $\begin{bmatrix} A - j\omega & B_2 \\ C_1 & D_{12} \end{bmatrix}$

$= \text{rank} \begin{bmatrix} j\omega - 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = 2$ $\qquad$ for all ω $\quad$ rank $\begin{bmatrix} A - j\omega & B_1 \\ C_2 & D_{21} \end{bmatrix}$

$= \text{rank} \begin{bmatrix} j\omega - 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = 2$ $\qquad$ for all ω

(d) $\tilde{D}_{12} = \tilde{D}_{21} = 1,\ \tilde{C}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ \tilde{B}_1 = [1 \quad 0]$

(e) $X_\infty = \text{Ric} \begin{bmatrix} 1 & \gamma^{-2} - 1 \\ -1 & -1 \end{bmatrix} = \dfrac{-\gamma^2 - \gamma\sqrt{2\gamma^2 - 1}}{1 - \gamma^2}$

(f) $Y_\infty = X_\infty$ for this problem

(g) $X_\infty Y_\infty = 1/(1 - \gamma^2)^2 \left[\gamma^4 + \gamma^2(2\gamma^2 - 1) + 2\gamma^3\sqrt{2\gamma^2 - 1}\right]$

for $\gamma = 2$, $X_\infty Y_\infty > 4$, violated

for $\gamma = 3$, $X_\infty Y_\infty < 9$, satisfied

(h) $X_\infty = Y_\infty = 2.67$, $Z_\infty = 4.82$, $K_c = K_e = 2.67$

(i) $K(s) = -34.43 \quad 1/(s + 14.26)$, poles at $s = -1.75, -11.51$

(j) phase margin= 56.5°, infinite gain margin

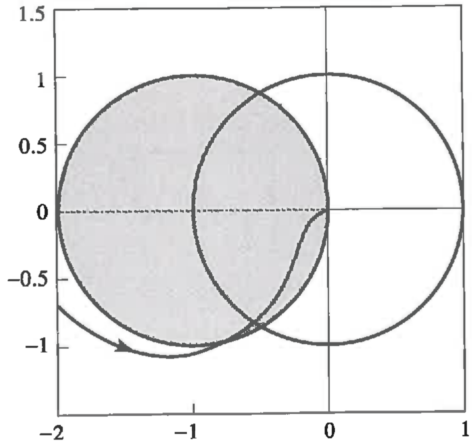


Figure D10.6j

10.8 SUMMARY

Linear quadratic methods for control system design have been discussed. These techniques lead to linear controllers that are easy to implement. Another important feature is guaranteed closed-loop stability. The LQR technique requires that all states be available for measurement. If the system is controllable (or at least stabilizable), this method gives excellent stability margins. The guaranteed margins are 60° phase margin, infinite gain margin, and−6 dB gain reduction margin. The design involves selection of the state and control weights, Q and R matrices, and solution of the Riccati equation. This can be also accomplished by using the root-square locus approach.

If all the states are not available for feedback, a Kalman filter (observer or estimator) can be designed. The combination of the filter and the LQ controller is called the LQG compensator. The design starts with selection of the state control weights Q and R, selection (or determination) of the process and measurement noise intensities, and solution of two Riccati equations. The compensator structure is of the observer-based type seen in Chapter 9, and the solution satisfies the separation principle. Although the closed-loop system is guaranteed to be stable, it will have no guaranteed stability margins. The design requires perfect knowledge of the system model and consequently is not robust.

Feedback properties of systems can be adequately characterized by the sensitivity or complementary sensitivity transfer functions. Performance specifications impose bounds of the open-loop transfer function of the system in various frequency ranges. Stability or performance of a system is robust if it is maintained in spite of model uncertainty. Model uncertainty can be modeled as either structured or unstructured uncertainty. Unstructured uncertainty can be modeled as either additive or multiplicative uncertainty. The small-gain theorem can be used to determine if the system is robustly stable under model uncertainties.

Loop transfer recovery (LTR) is a modification of the LQG technique; it allows recovery of the LQR stability margins. One begins with selection the LQR parameters until a desired open-loop transfer function (the target feedback loop) is obtained. The recovery step involves iterating on the filter design parameters until the desired loop transfer function shape has been obtained. The LTR methodology converts the LQG technique from a rigid time domain method to a flexible frequency domain design technique. The computations are still based on state space techniques, but they can remain hidden from the user.

H_∞ control is the newest tool for control-system design. It is a computer-aided frequency domain method for design of multivariable systems. The exogenous inputs (disturbances, command inputs, sensor noise) are collected into one vector; the regulated outputs (control signals, errors) are collected into another vector. This will result in a dual-input, dual-output block diagram, called the *generic* (or *synthesis*) *block diagram*. The objective is to maintain the peak in the closed-loop frequency response of the system below a specified level γ. The optimal problem can be solved by iteration on γ. The solution involves selecting weights (possibly frequency-dependent weights) and solving two Riccati equations. The compensator structure is similar to LQG with some added terms. It can be shown that if the value of γ is allowed to go to infinity, the solution approaches the LQG solution.

REFERENCES

LQR/LQG/LTR

Anderson, B. D. O., and Moore, J. B., *Linear Optimal Control.* Englewood Cliffs, NJ: Prentice-Hall, 1990.

Bryson, A. E, Jr., and Ho, Y. C., *Applied Optimal Control: Optimization, Estimation and Control.* Washington, DC: Hemisphere Publishing Corporation, 1975.

Chen, C. T., *Control System Design: Conventional, Algebraic and Optimal Methods.* Pond Woods Press, 1987.

Doyle, J. C., and G. Stein. Robustness with Observers. *IEEE Trans. Autom. Control*, vol. 24, 1979, pp. 607–611.

B. Friedland. *Control Systems Design: An Introduction to State Space Methods.* New York: McGraw-Hill, 1986.

IEEE Trans. Autom. Control. Special issue on linear multivariable control systems, (February 1981).

R. E. Kalman. When Is a Linear Control System Optimal? *J. Basic Eng. Trans. ASME D*, 86 (1964): 51–60.

Kirk, D. E., *Optimal Control Theory: An Introduction.* Englewood Cliffs, NJ: Prentice-Hall, 1970.

Kwakernaak, H., and Sivan, R., *Linear Optimal Control Systems.* New York: Wiley-Interscience, 1972.

Lewis, F. L., *Optimal Control.* New York: Wiley, 1986.

H_∞ Control

Boyd, S, P., and Barratt, C. H., *Linear Controller Design.* Englewood Cliffs, NJ: Prentice-Hall, 1991.

Doyle, J. C., Francis, B. A., and Tannenbaum, A. R., *Feedback Control Theory.* New York: Macmillan, 1992.

Hauser, Frank, Lecture Notes for Modern Control Theory for the Practitioner, UCLA Extension course, Los Angeles, 1992.

Levine, W.S., and Reichert, R. T., "An Introduction to H_∞ Control System Design." *Proceedings of the 29th Conference on Decision and Control,* Honolulu, December 1990.

Maciejowski, J. M., *Multivariable Feedback Design.* Reading, MA: Addison-Wesley, 1989.

Safanov, M. G., Lecture Notes for a Course in H_∞ Control offered at the University of Southern California, Los Angeles 1992.

Computational Techniques

Balas, G. J., Doyle, J. C., Glover, K., Packard, A., and Smith, R., *User's Guide to the μ-Analysis and Synthesis Toolbox for MATLAB.* MUSYN Inc. and Math Works, Inc., S. Natick, MA, 1991.

Balas, G. J., Packard, A., Doyle, J. C., Glover, K., and Smith, R., "Development of Advanced Control Design Software for Researchers and Engineers. *Proceedings of the American Control Conference*, Boston, June 26–28, 1991.

Chiang, R. Y., and Safonov, M. G., "A Hierarchical Data Structure and New Capabilities of the Robust-Control Toolbox." *Proceedings of the American Control Conference*, Boston, June 26–28, 1991.

Chiang, R. Y., and Safonov, M. G., *User's Guide to the Robust Control Toolbox for MATLAB*. South Natick, MA: MathWorks, Inc., 1988.

Integrated Systems Inc., *User's Guide to the Robust Control Module for MATRIX$_X$*. Santa Clara, CA:, Integrated Systems Inc., 1990.

Shahian, B., and Hassul, M., *Control System Design Using MATLAB*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

Shahian, B., and Hassul, M., *Control System Design Using MATRIX$_X$*. Englewood Cliffs, N J: Prentice-Hall, 1992.

Thompson, Peter, *Tutorial and User's Guide for program CC, Version 4.* Systems Technology, Hawthorn, CA: 1988.

PROBLEMS

1. For each of the following systems, obtain the optimal control gain using LQR by solving the Riccati equation.

(a) $A = \begin{bmatrix} 0 & 1 \\ -4 & -4 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 3 \end{bmatrix}, Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, R = 1$

(b) $A = \begin{bmatrix} 0 & 1 \\ -4 & -4 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 3 \end{bmatrix}, Q = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, R = 1$

(c) $A = \begin{bmatrix} 0 & 1 \\ -4 & -4 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 3 \end{bmatrix}, Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, R, = 1$

Ans. (a) $K = [0.33 \quad 0.08]$

2. Obtain the root-square locus in each case for Problem 1. Let $R = \rho$

3. Obtain the root-square locus for each of the following systems.

(a) $A = \begin{bmatrix} 0 & 1 \\ -3 & -8 \end{bmatrix}, \ B = \begin{bmatrix} 0 \\ 4 \end{bmatrix}, \ Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \ R = \rho$

(b) same system as (a) but let $Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, R = \rho$

(c) $A = \begin{bmatrix} 0 & 1 \\ -4 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, R = \rho$

4. Use the Riccati equation to solve Problem 3. Find the gain for $\rho = 0.1, \ 1, \ 10$ in each case.

5. Obtain the compensator transfer function, draw Bode plots, and compute gain and phase margins for each system in Problem 1.

   Let $C = [1 \quad 0]$ in each case.

6. Consider the system given by

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad Q = \begin{bmatrix} 1 & 0 \\ 0 & q \end{bmatrix} \quad R = 1$$

   Find the optimal control gain and the closed-loop poles as a function of q. Discuss what happens to these quantities as q increases ($q > 0$).

7. Consider the following system:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -12 & 0 & 12 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 \end{bmatrix}$$

   Let $Q = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad R = 10^{-6}$

(a) Find the control gain vector K and the optimal closed-loop poles

(b) Let $Q_o = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ and $R_o = \begin{bmatrix} 2 \times 10^{-6} & 0 \\ 0 & 10^{-6} \end{bmatrix}$

Design a Kalman filter. Find the filter gain L and filter eigenvalues.

(c) Find the equation for the LQG compensator of (a) and (b).

(d) Find the eigenvalues of the closed-loop system.

(e) Plot the impulse and step response of the system.

8. Consider the system given by

$$A = \begin{bmatrix} -54 & 2 & 10 \\ 2 \times 10^{-4} & -10^{-3} & -5 \times 10^{-3} \\ -10^{-3} & -24 \times 10^{-3} & -0.14 \end{bmatrix}$$

$$B = \begin{bmatrix} -10^4 \\ 0.25 \\ -2 \end{bmatrix} \quad C = [0 \quad 3 \quad 0.05]$$

The open-loop transfer function must meet the following specification:

(i) $|GH| > 20$ dB for $\omega \leqslant 0.1$ rad/s
for good disturbance rejection and command tracking.

(ii) $|1 + GH| \geqslant 25$ dB for $\omega \leqslant 0.05$ rad/s
for insensitivity to parameter variations.

(iii) $|GH| \leqslant -20$ dB for $\omega \geqslant 5$ rads
for good immunity to noise.

(a) Design an LQG compensator and record its performance with respect to the foregoing specifications.

(b) Use LTR to meet the specification as closely as possible.

9. Consider the problem in Drill Problem D10.4.

$$G(s) = \frac{1}{s^2} \qquad H(s) = \frac{20(s+1)}{(s+10)}$$

Suppose the actual plant is given by

$$\tilde{G}(s) = \frac{2(s+1)}{s^2(s^2+s+1)}$$

(a) Find a multiplicative uncertainty model for the system.

(b) Find $M(s)$, the transfer function as seen by the multiplicative uncertainty.

(c) Determine if the closed-loop system is robustly stable under the multiplicative uncertainty computed in (a).

Ans. (a) $\Delta_m(s) = (-s^2 + s + 1)/(s^2 + s + 1)$;

(b) $M(s) = -20(s + 1)/(s^3 + 10s^2 + 20s + 20)$

10. This problem is adapted from a paper by W. S. Levine and R. T. Reichert (*Proceedings of Conference on Decision and Control,* December, 1990).

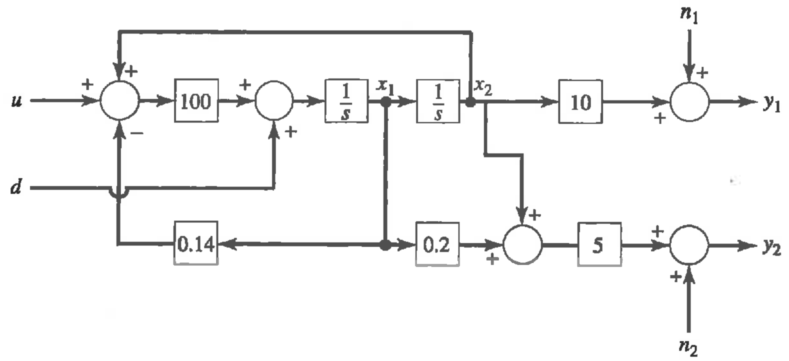


Figure P10.10

Consider the single-input, dual output plant, in Figure P10.10(a). where $u =$ control, $d =$ disturbance, n_1 and $n_2 =$ sensor noise sources. The specifications are to have integral tracking performance for output y_1 with a time constant of 0.6 s. The second output y_2, is also available for feedback.

It has been suggested that for integral tracking, one should regulate the integral of the tracking error. Hence, we will introduce a command input y_c, and an integrator at $y_1 - y_c$(before the entry of the noise source n_1).

In addition, all regulated outputs and disturbance and noise inputs will be weighted. The new block diagram is shown in Figure P10.10 (b).

The following weights are suggested: $W_d = 0.1$, $W_{n1} = W_{n2} = 1$, $W_e = 0.5$, $W_u = 1$.

(By varying these weights, one can manipulate the performance of the system).

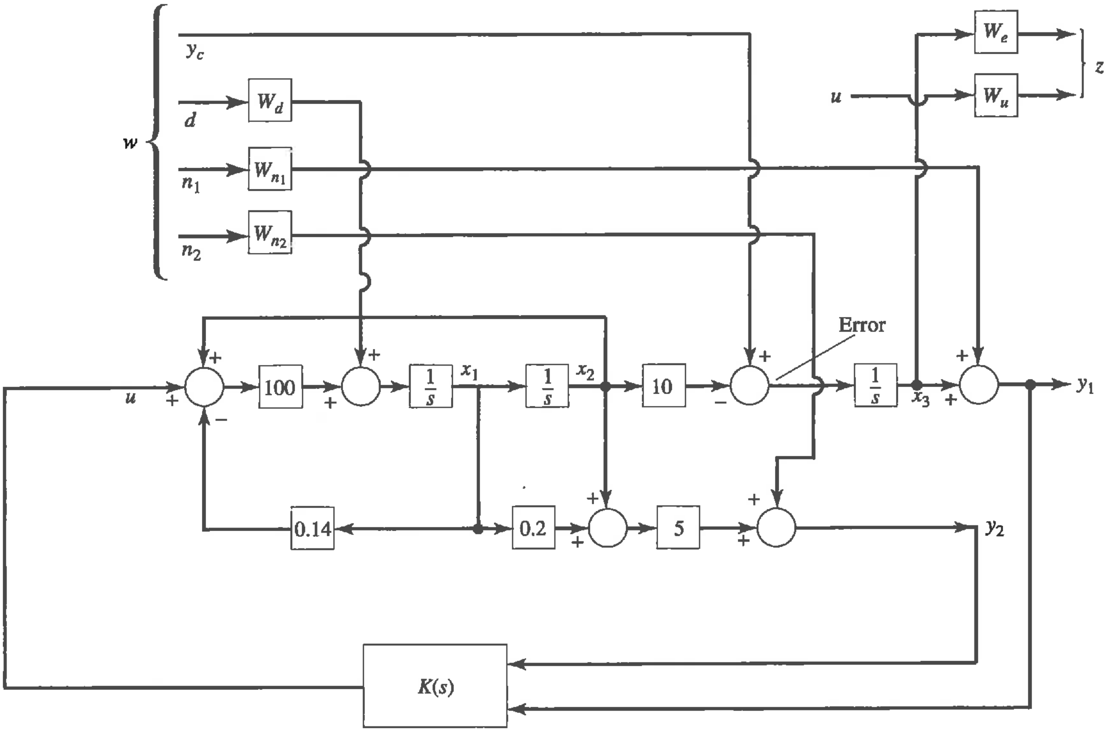


Figure P10.10

(a) Verify that the plant equations are given by

$$\dot{x}_1 = -14x_1 + 100x_2 + 0.1d + 100u$$

$$\dot{x}_2 = x_1$$

$$\dot{x}_3 = -10_{x2} + y_c$$

$$z = \begin{bmatrix} 0.5x_3 \\ u \end{bmatrix}$$

$$y = \begin{bmatrix} x_3 + n_1 \\ x_1 + 5x_2 + n_2 \end{bmatrix}$$

(b) Obtain the plant matrix $P(s)$.

(c) For a value of $\gamma = 0.5922$ (optimal value obtained by computer) find K_c, K_e, and the compensator transfer function.

(d) Draw the Nyquist plot, and compute the stability margins.

(e) Plot the unit step response of the system (set all inputs to zero and let $y_c = 1$, then plot y_1).

(f) Use γ-iteration to verify that the $\gamma = 0.5922$ is optimal.

$$\textbf{Ans.}\quad (b)\quad P(s) = \left[\begin{array}{ccc|cccc|c} -14 & 100 & 0 & 0 & 0.1 & 0 & 0 & 100 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -10 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 5 & 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right]$$

Note: The (1, 3) element of C_2 is incorrect in the original source.

(c) $K_c = [\ 0.14 \quad 3 \quad -0.52]$

$$K_e = \begin{bmatrix} -0.89 & 5.26 \\ -0.17 & 1.01 \\ 2.16 & -1.76 \end{bmatrix}$$

$$K(s) = \frac{1}{(s+4.9)(s-16.6)(s+318.2)}[(s+3.6)(s+19.2)$$

$$(s+0.5)(s+19.2)]$$

(d) $PM = 51, GM = -5.5.$

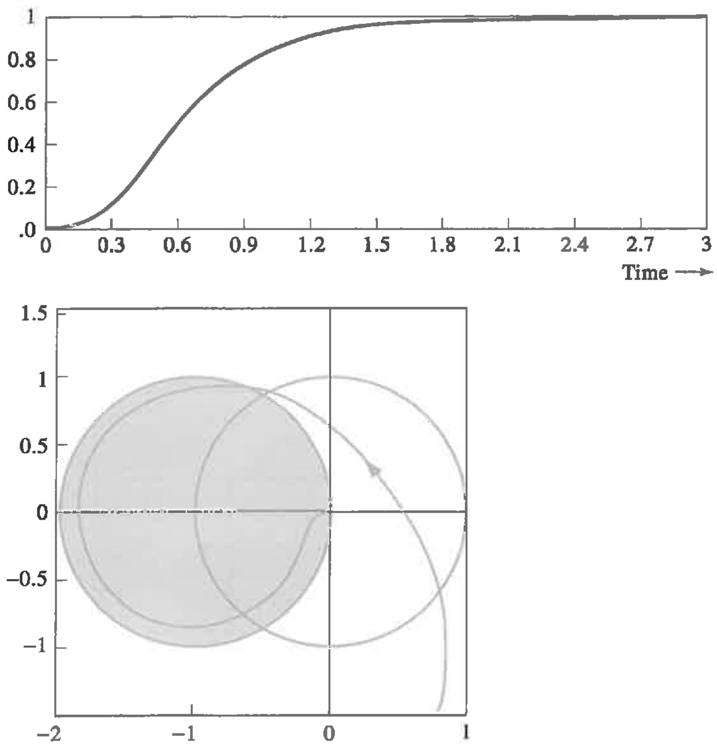
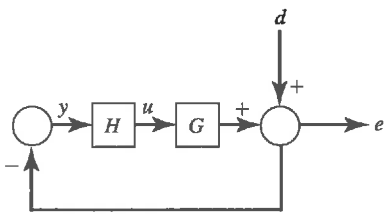


Figure P10.10

11. Consider the plant given by $G(s) = (s-1)/[(s+1)(s^2+s+1)]$. The specification is $|S(j\omega)| \leqslant 0.1$ for $\omega \leqslant 0.01$ rad/s. This is a sensitivity minimization problem. The requirement means that we want to reject low-frequency disturbances (or equivalently, to reduce the system sensitivity to parameters variations or model uncertainties). To set up the problem for H_∞, consider Figure P10.11(a).



(a)

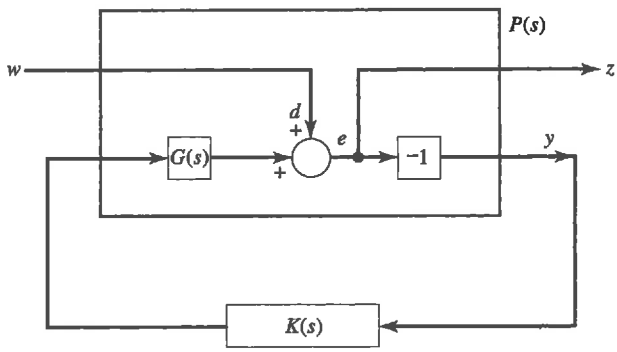
Figure P10.11

The transfer function from disturbance d to error e is given by

$$\frac{E(s)}{D(s)} = \frac{1}{1+GH} = S$$

Now, we redraw the diagram as in Figure P 10.11(b). (We have changed the notation to correspond to Section 10.7.)

Obtain the system matrix, $P(s)$, and use γ-iteration to solve the problem. Plot the sensitivity, $S(j\omega)$, and vary γ until the specification is met. For the final design, display $|S(j\omega)|$, the compensator transfer function $K(s)$, and the step response of the system.



(b)

Figure P10.11

12. Consider the H_∞ set up in Figure P10.12.

$$G(s) = \frac{1}{(s+10^{-3})^2}$$

$$G(s) = \frac{1}{(s+10^{-3})^2}$$

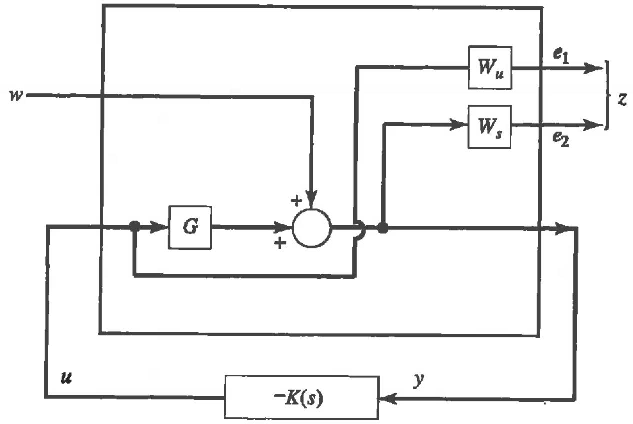


Figure P10.12

The objective is to minimize $\|T_{zw}\|_\infty$. This corresponds to minimizing the effects of the disturbance w. Note that the transfer function from w to e_2 is the sensitivity weighted by W_s, and the transfer function from w to e_1 is related to the ASM (Section 10.5).

$$\frac{e_2}{w} = \frac{1}{1+KG}W_s \qquad \frac{e_1}{w} = \frac{-K}{1+KG}W_u$$

Hence by reducing the $\|e_1/w\|_\infty$, we are also increasing the additive stability margin.

Let

$$W_s = \frac{1}{(s+10^{-3})^2}$$

and

$$W_u = \begin{cases} 1 & \text{Case I} \\ 10^{-4} & \text{Case II} \end{cases}$$

Find the H_∞ optimal compensator in each case. In each case display the Bode magnitude plot for GK, S, and T. Also give the optimal value of γ. Discuss the effects of the weight W_u and compare the two cases.

13. Repeat Problem 12 for $G(s) = (s - 1)/[(s + 1)(s + 10^{-3})^2]$.

$$W_u = 10^{-4} \qquad W_s = \begin{cases} \dfrac{1}{(s + 10^{-3})^2} & \text{Case I} \\ \dfrac{s^2 + 1.4s + 1}{(s + 10^{-3})^2(1 + 10^{-3}s)} & \text{Case I} \\ 1 & \text{Case III} \end{cases}$$

14. It is possible to use the Nyquist plot and some geometry to obtain formulas for gain and phase margins in terms of return difference and sensitivity.

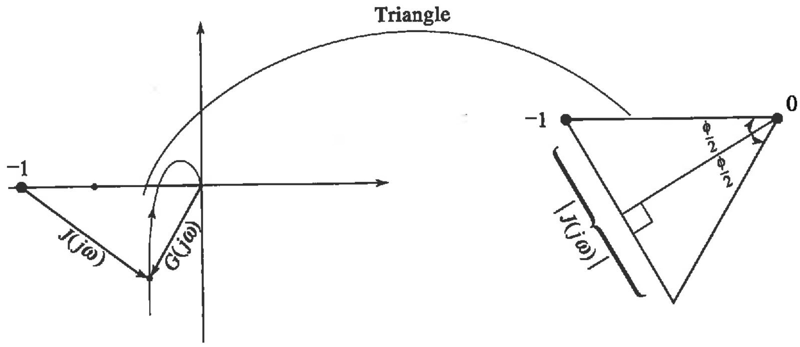


Figure P10.14

In Figure P10.14, if

$J(\omega) = 1 + G(j\omega) =$ return difference

and

$$S(j\omega) = \frac{1}{1 + G(j\omega)} = \text{sensitivity}$$

use the figure as a hint to show that

$$PM = 2\sin^{-1}\left|\frac{J(j\omega_{gc})}{2}\right| \qquad \omega_{gc} = \text{gain crossover frequency}$$

$$GM = -20\ \log\left(1 - \frac{1}{|S(j\omega)|}\right)$$

$$PM = 2\sin^{-1}\ \frac{1}{2\,|S(j\omega)|}$$

Tabulate $|S(j\omega)|$, GM, PM for $|S(j\omega)|$ from 1 to 3 in 0.2 increments. [Note: for $|S| = 1$, we get LQR stability margins.]