

## Chapter 3

# Control Design Techniques

### 3.1 Aims and Objectives

- To introduce linear control system concepts of transfer function, impulse response, step response, and frequency response, as well as the measures of control system performance and robustness
- To present an overview of design techniques for single-variable and multivariable linear control systems
- To highlight useful control theory concepts

### 3.2 Transfer Function and Singularity Inputs

A linear time-invariant (LTI), *single-input, single-output* (SISO) system's response to a specified input is most efficiently analyzed and designed with the concept of *transfer function*. Consider an LTI, SISO system, with input,  $u(t)$ , and output,  $y(t)$ , initially at rest and not experiencing the input initially, i.e.,  $u(t) = 0, y(t) = 0$ , ( $t < 0$ ). All the time derivatives of the input and output are also zeros for  $t < 0$ . In the following discussion, the system must have the property of *causality*, i.e., it produces a nonzero output only when acted upon by a nonzero input for  $t \geq 0$ . The transfer function,  $G(s)$ , of such a system is then defined as the ratio of the output's Laplace transform,  $Y(s)$ , to that of the input,  $U(s)$ , subject to the zero initial condition,

$$G(s) \doteq \frac{Y(s)}{U(s)}. \quad (3.1)$$

The zero initial condition is commonly denoted by  $u(0_-) = 0, y(0_-) = 0$ , where  $t = 0_-$  indicates the time immediately before the application of the input. By the definition of Laplace transform [11] and noting the causality of the system, we write

$$U(s) = \mathcal{L}\{u(t)\} = \int_0^\infty e^{-st} u(t) dt; \quad Y(s) = \mathcal{L}\{y(t)\} = \int_0^\infty e^{-st} y(t) dt. \quad (3.2)$$

**Table 3.1** Basic Laplace transforms and properties

$f(t)$	$F(s) = \mathcal{L}\{f(t)\}$
$\delta(t) = \begin{cases} \infty & (t = 0) \\ 0 & (t \neq 0) \end{cases}$	1
$u_s(t) = \begin{cases} 1 & (t \geq 0) \\ 0 & (t < 0) \end{cases}$	$\frac{1}{s}$
$tu_s(t)$	$\frac{1}{s^2}$
$t^n u_s(t); (n = 1, 2, \dots)$	$\frac{n!}{s^{n+1}}$
$e^{at} u_s(t)$	$\frac{1}{s - a}$
$\sin(\omega t) u_s(t)$	$\frac{\omega}{s^2 + \omega^2}$
$\cos(\omega t) u_s(t)$	$\frac{s}{s^2 + \omega^2}$
$e^{at} h(t)$	$H(s - a); H(s) = \mathcal{L}\{h(t)\}$
$h(t - a) u_s(t - a)$	$e^{-as} H(s); H(s) = \mathcal{L}\{h(t)\}$
$-th(t)$	$\frac{dH(s)}{ds}; H(s) = \mathcal{L}\{h(t)\}$
$\frac{h(t)}{t}$	$\int_s^\infty H(p) dp; H(s) = \mathcal{L}\{h(t)\}$
$\frac{dh(t)}{dt}$	$sH(s) - h(0_-); H(s) = \mathcal{L}\{h(t)\}$
$\int_0^t h(\tau) d\tau$	$\frac{H(s)}{s}; H(s) = \mathcal{L}\{h(t)\}$

The existence of the Laplace transforms requires that the integrals in (3.2) should converge to finite values for a given complex variable,  $s$ . If  $U(s)$  and  $Y(s)$  exist then they are unique. It can be shown [11] that a Laplace integral,  $\mathcal{L}\{f(t)\}$ , converges if and only if the function,  $f(t)$ , is piecewise continuous (i.e., any time interval, however large, can be broken up into a finite number of subintervals over each of which  $f(t)$  is continuous, and at the ends of each subinterval,  $f(t)$  is finite) and bounded by an exponential (i.e., there exists a positive, real constant,  $a$ , such that  $|e^{-at} f(t)|$  is finite at all times). Practical Laplace transforms of functions defined for  $t \geq 0$  (along with some important properties) are listed in Table 3.1. Most of the commonly applied inputs are Laplace transformable, and the definition of the transfer function by (3.1) requires that the output of the LTI system to such an input is also Laplace transformable (which is true for all physical systems).

*Example 3.1.* Find the transfer function of a causal system described by the following governing differential equation:

$$\frac{d^3 y}{dt^3} - 4 \frac{d^2 y}{dt^2} + \frac{dy}{dt} - 2y = 0.1 \frac{du}{dt} - 0.5u,$$

where  $u(t)$  is the input that begins acting at  $t = 0$  and  $y(t)$  the resulting output.

We apply the property of Laplace transform of a time derivative (Table 3.1) with zero initial conditions on both input and output, and write

$$\mathcal{L}\left\{\frac{dy}{dt}\right\} = sY(s); \quad \mathcal{L}\left\{\frac{d^2y}{dt^2}\right\} = s^2Y(s); \quad \mathcal{L}\left\{\frac{d^3y}{dt^3}\right\} = s^3Y(s),$$

and

$$\mathcal{L}\left\{\frac{du}{dt}\right\} = sU(s),$$

where  $U(s) = \mathcal{L}\{u(t)\}$  and  $Y(s) = \mathcal{L}\{y(t)\}$ . Thus, taking the Laplace transform of both sides of the governing differential equation yields:

$$(s^3 - 4s^2 + s - 2)Y(s) = (0.1s - 0.5)U(s),$$

or

$$G(s) = \frac{Y(s)}{U(s)} = \frac{0.1s - 0.5}{s^3 - 4s^2 + s - 2}.$$

This example indicates that an LTI, SISO system's transfer function is a ratio of polynomials, called a *rational function*, in the Laplace variable,  $s$ . Note that here the degree of the numerator polynomial (i.e., the highest power of  $s$ ) is less than that of the denominator polynomial.

For a general LTI, SISO system with the governing differential equation,

$$\frac{d^n y}{dt^n} + a_n \frac{d^{n-1} y}{dt^{n-1}} + \cdots + a_2 \frac{dy}{dt} + a_1 y = b_{m+1} \frac{d^m u}{dt^m} + b_m \frac{d^{m-1} u}{dt^{m-1}} + \cdots + b_2 \frac{du}{dt} + b_1 u, \quad (3.3)$$

the transfer function is the following:

$$G(s) = \frac{b_{m+1}s^m + b_m s^{m-1} + \cdots + b_2 s + b_1}{s^n + a_n s^{n-1} + \cdots + a_2 s + a_1}. \quad (3.4)$$

Here,  $n$  is the order of the system. If the degree of the numerator polynomial of  $G(s)$  is either less than or equal to that of the denominator polynomial ( $m \leq n$ ) then the LTI system is said to be *proper*. If  $m < n$  then we have no direct connection between input and output, and the system is called *strictly proper*.

We can express the LTI, SISO system given by (3.3) in the following state-space form (Chap. 1):

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t) \\ y(t) &= \mathbf{C}\mathbf{x}(t) + Du(t), \end{aligned} \quad (3.5)$$

where  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  are constant coefficient matrices,  $D$  a constant scalar, and  $\mathbf{x}(t)$  is the state vector. For a strictly proper system,  $D = 0$ . By taking Laplace transform of both sides of the state and output equations subject to zero initial condition,  $\mathbf{x}(0_-) = \mathbf{0}$ , we have

$$\begin{aligned} s\mathcal{X}(s) &= \mathbf{A}\mathcal{X}(s) + \mathbf{B}U(s) \\ Y(s) &= \mathbf{C}\mathcal{X}(s) + DU(s), \end{aligned} \quad (3.6)$$

or

$$Y(s) = [\mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1} + D] U(s). \quad (3.7)$$

Comparing (3.1) and (3.7) we have the following relationship between the transfer function and the state-space coefficient matrices:

$$G(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1} + D. \quad (3.8)$$

As

$$(s\mathbf{I} - \mathbf{A})^{-1} = \frac{\text{adj}(s\mathbf{I} - \mathbf{A})}{|s\mathbf{I} - \mathbf{A}|} \quad (3.9)$$

we see that the denominator polynomial of  $G(s)$  is the same as the characteristic polynomial,  $|s\mathbf{I} - \mathbf{A}|$ . Thus, the roots of the denominator polynomial, called the *poles* of the transfer function,

$$|s\mathbf{I} - \mathbf{A}| = s^n + a_n s^{n-1} + \cdots + a_2 s + a_1 = 0, \quad (3.10)$$

are the same as the eigenvalues of  $\mathbf{A}$ , and determine the LTI system's characteristics. The roots of the numerator polynomial of the transfer function,  $b_{m+1}s^m + b_m s^{m-1} + \cdots + b_2 s + b_1 = 0$  are called the *zeros* of the system. Clearly, the constant coefficients of the denominator polynomial,  $a_1, \dots, a_n$ , have an important role to play in stability and performance of an LTI system, while the coefficients  $b_1, \dots, b_{m+1}$  also have an influence on the system's response to applied inputs.

*Example 3.2.* Find the poles and zeros of the system given in Example 3.1 and determine its stability.

The poles are the roots of the denominator polynomial of  $G(s)$ ,  $s^3 - 4s^2 + s - 2 = 0$ , which are determined numerically by the use of MATLAB operator *roots.m* as follows:

```
>> a=[1 -4 1 -2];
>> roots(a)

ans =
    3.8751
    0.0624 + 0.7157i
    0.0624 - 0.7157i
```

Thus, the poles of the third-order system are  $s_1 = 3.8751$  and  $s_{2,3} = 0.0624 \pm 0.7157i$ . By the stability criteria of Chap. 1, the system is unstable (presence of positive real parts). The system has only one zero, i.e., the root of the numerator polynomial of  $G(s)$ ,  $0.1s - 0.5 = 0$ , or  $s = 5$ .

### 3.2.1 Impulse Response

The transfer function of an LTI, SISO system is also the Laplace transform of its output,  $y(t)$ , when the initial condition is zero and the input is a *unit impulse* (or Dirac delta) function applied at  $t = 0$ ,  $\delta(t)$ , and given by:

$$\delta(t) = \begin{cases} \infty & (t = 0) \\ 0 & (t \neq 0) \end{cases}, \quad (3.11)$$

such that

$$\int_{-\infty}^{\infty} \delta(t) dt = 1. \quad (3.12)$$

The unit impulse function,  $\delta(t)$ , can be visualized as a rectangular pulse of width,  $\epsilon$ , and height,  $1/\epsilon$ , centered at  $t = 0$ , by taking the limit,  $\epsilon \rightarrow 0$ . The output of an LTI system with zero initial condition to a unit impulse input is called its *impulse response*,  $g(t)$ , and is given by:

$$g(t) = \mathcal{L}^{-1}\{G(s)\} \quad (t \geq 0). \quad (3.13)$$

A useful method of calculating the impulse response is by the *partial fraction expansion* of the transfer function as follows:

$$G(s) = \frac{r_1}{s - p_1} + \frac{r_2}{s - p_2} + \cdots + \frac{r_n}{s - p_n} \\ + \frac{r_{m1}}{s - p_m} + \frac{r_{m2}}{(s - p_m)^2} + \cdots + \frac{r_{m(k-1)}}{(s - p_m)^{k-1}} + \frac{r_{mk}}{(s - p_m)^k}, \quad (3.14)$$

where  $p_1, p_2, \dots, p_n$  are  $n$  distinct poles, and  $p_m$  a pole of multiplicity  $k$ , of  $G(s)$ . The numerator coefficients of the series,  $r_1, r_2, \dots, r_n$ , and  $r_{m1}, \dots, r_{mk}$ , are called the *residues* of  $G(s)$  and are determined by the following scheme:

$$\begin{aligned} r_1 &= (s - p_1)G(s)|_{s=p_1} \\ r_2 &= (s - p_2)G(s)|_{s=p_2} \\ r_n &= (s - p_n)G(s)|_{s=p_n} \\ r_{mk} &= (s - p_m)^k G(s)|_{s=p_m} \end{aligned} \quad (3.15)$$

$$\begin{aligned}
 r_{m(k-1)} &= \left. \frac{d(s - p_m)^k G(s)}{ds} \right|_{s=p_m} \\
 r_{m2} &= \left. \frac{d^{k-2}(s - p_m)^k G(s)}{ds^{k-2}} \right|_{s=p_m} \\
 r_{m1} &= \left. \frac{d^{k-1}(s - p_m)^k G(s)}{ds^{k-1}} \right|_{s=p_m}
 \end{aligned}$$

*Example 3.3.* Calculate the impulse response of a system with the following transfer function:

$$G(s) = \frac{s + 3}{(s + 1)^2(s + 2)}.$$

The poles of the system are all real – a single pole  $p_1 = -2$ , and a double pole,  $p_1 = -1$ . Thus, we write

$$G(s) = \frac{r_1}{s + 2} + \frac{r_2}{s + 1} + \frac{r_3}{(s + 1)^2},$$

where the residues are calculated as follows:

$$\begin{aligned}
 r_1 &= (s + 2)G(s)|_{s=-2} = \frac{-2 + 3}{(-2 + 1)^2} = 1 \\
 r_3 &= (s + 1)^2 G(s)|_{s=-1} = \frac{-1 + 3}{(-1 + 2)} = 2
 \end{aligned} \tag{3.16}$$

$$r_2 = \left. \frac{d(s + 1)^2 G(s)}{ds} \right|_{s=-1} = \left. \frac{s + 2 - (s + 3)}{(s + 2)^2} \right|_{s=-1} = -1. \tag{3.17}$$

Thus, we have

$$G(s) = \frac{1}{s + 2} - \frac{1}{s + 1} + \frac{2}{(s + 1)^2},$$

or, by Table 3.1,

$$g(t) = \mathcal{L}^{-1}\{G(s)\} = (e^{-2t} - e^{-t} + 2te^{-t})u_s(t).$$

In case of complex poles (which always occur in conjugate pairs), the residues are also complex conjugates; hence, the partial fraction expansion involving a complex conjugate pair of poles can be combined to produce a quadratic (or second-order) subsystem.

*Example 3.4.* Calculate the impulse response of the system given in Example 3.2.

$$G(s) = \frac{0.1s - 0.5}{s^3 - 4s^2 + s - 2}.$$

The third-order system with distinct poles is expanded as follows:

$$G(s) = \frac{r_1}{s - p_1} + \frac{r_2}{s - p_2} + \frac{r_3}{s - p_3}.$$

Example 3.2 showed that the system had a real pole,  $p_1 = 3.8751$ , and a complex conjugate pair,  $p_{2,3} = 0.0624 \pm 0.7157i$ . Thus, we expect the residue,  $r_1$ , to be real and the residues,  $r_{2,3}$ , to be complex conjugates. The complex terms are combined by writing  $r_{2,3} = a \pm ib$  and  $p_{2,3} = c \pm id$ , resulting in

$$\begin{aligned} \frac{r_2}{s - p_2} + \frac{r_3}{s - p_3} &= \frac{a + ib}{s - (c + id)} + \frac{a - ib}{s - (c - id)} \\ &= \frac{(a + ib)(s - c + id) + (a - ib)(s - c - id)}{(s - c)^2 + d^2} \\ &= \frac{2(as - bd - ac)}{(s - c)^2 + d^2}. \end{aligned}$$

Hence, we have

$$G(s) = \frac{r_1}{s - p_1} + \frac{2(as - bd - ac)}{(s - c)^2 + d^2},$$

and we first calculate  $r_1$  by

$$r_1 = (s - p_1)G(s)|_{s=p_1} = \frac{0.1s - 0.5}{(s - c)^2 + d^2} \Big|_{s=p_1} = -0.00747479.$$

Next,  $r_1$  is substituted into expanded  $G(s)$ , and the complex residue parts,  $a, b$ , calculated by comparing the coefficients of the numerator polynomial:

$$a = -\frac{r_1}{2} = 0.003737395,$$

and

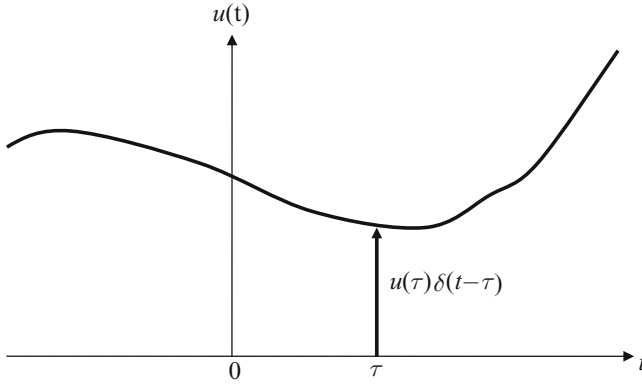
$$b = -\frac{r_1 c + ac + ap_1 + 0.05}{d} = -0.089772745$$

Finally, we take the inverse Laplace transform of the expanded transfer function according to Table 3.1, yielding

$$g(t) = \mathcal{L}^{-1}\{G(s)\} = \left[ r_1 e^{p_1 t} + 2e^{ct} \left( \frac{1}{d} \cos(dt) - \left\{ b + \frac{ac}{d} \right\} \sin(dt) \right) \right] u_s(t),$$

which, by substituting the values computed earlier, becomes

$$\begin{aligned} g(t) = & \left[ -0.0075e^{3.8751t} + 2e^{0.0624t} (267.566 \cos(0.7157t) \right. \\ & \left. + 0.0894467 \sin(0.7157t)) \right] u_s(t). \end{aligned}$$



**Fig. 3.1** The functional value,  $u(\tau)$ , represented by an impulse at  $t = \tau$

The impulse response of a linear system (thus the transfer function) is a useful mathematical construct, as it helps us in deriving the system's response to an arbitrary input by linear superposition (Chap. 1). Consider a piecewise continuous, arbitrary input,  $u(t)$ , represented by series of impulse inputs,  $\delta(t - \tau)$ , applied at various times,  $-\infty < \tau \leq t$ , and scaled by the current input magnitude,  $u(\tau)$ , as shown in Fig. 3.1. By the *mean value theorem* of integral calculus [11], (3.12) yields the following expression for input magnitude at time  $t$ , called the *sampling property* of the Dirac delta function:

$$u(t) = \int_{-\infty}^{\infty} u(\tau)\delta(t - \tau)d\tau. \quad (3.18)$$

From the definition of the unit impulse function, it is clear that the limits of integration in (3.18) need not be infinite, but should only bracket the time instant,  $\tau$ , at which the impulse is applied.

If the system is initially at rest (i.e., the output and all its time derivatives are zero at  $t = 0_-$ ) then the system's response at a subsequent time,  $t$ , is simply given by the summation of the individual impulse responses, scaled by the current input magnitude,  $u(\tau)$ :

$$y(t) = \int_{-\infty}^{\infty} u(\tau)g(t - \tau)d\tau. \quad (3.19)$$

The integral on the right-hand side of (3.19) is called the *convolution integral*, which is expressed as  $(u * g)(t)$ , and is symmetric in  $u(\cdot)$  and  $g(\cdot)$  such that we can write

$$y(t) = (u * g)(t) = \int_{-\infty}^{\infty} u(t - \tau)g(\tau)d\tau. \quad (3.20)$$

In the common control application, the input,  $u(t)$ , begins acting at  $t = 0$ , and since  $g(t - \tau) = 0$  for  $t < \tau$ , we have

$$y(t) = (u * g)(t) = \int_0^t u(\tau)g(t - \tau)d\tau = \int_0^t u(t - \tau)g(\tau)d\tau. \quad (3.21)$$

Note that (3.21) can be alternatively derived by taking the inverse Laplace transform of  $Y(s) = G(s)U(s)$  and using (3.13). Numerical evaluation of (3.21) is easily carried out by quadrature [11] for an arbitrary, piecewise continuous input,  $u(t)$ . (The concept of convolution can be extended beyond impulse response. Any two functions that satisfy the convolution property given by (3.21) are said to *convolve* with each other, and the Laplace transform of the convolution integral is a product of the Laplace transforms of the two functions.)

*Example 3.5.* Consider the single-axis rotation of a rigid spacecraft with the following equation of motion:

$$J\ddot{\theta} = u, \quad (3.22)$$

where  $\theta(t)$  is the angle of rotation,  $J$  the constant moment of inertia, and  $u(t)$  is the applied torque input. The spacecraft is equipped with a pair of rocket thrusters that can apply a torque impulse of magnitude,  $\pm u_0$ , whenever required.<sup>1</sup> It is desired to rotate the spacecraft by angle,  $\theta_f$ , beginning from rest at  $t = 0$ . Find the final time,  $t_f$ , for achieving the desired rotation by application of only two impulses.

Let us apply the first impulse at  $t = 0$  to initiate the rotation in the desired direction, and the second impulse in the opposite direction at  $t = t_f$  in order to stop the rotation when the desired angle is reached. Thus, the input profile is given by:

$$u(t) = u_0 [\delta(t) - \delta(t - t_f)]. \quad (3.23)$$

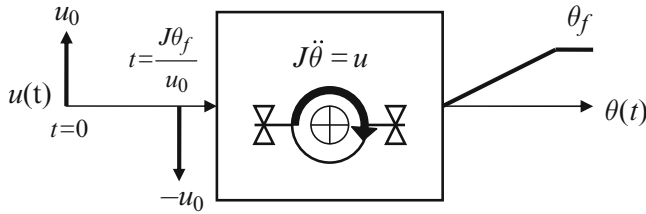
Substitution of (3.23) into (3.22) and integration in time yields

$$\dot{\theta}(t) = \frac{1}{J} \int_0^t u(\tau)d\tau = \begin{cases} \frac{u_0}{J} & (0 \leq t < t_f) \\ 0 & (t \geq t_f) \end{cases} \quad (3.24)$$

$$\theta(t) = \begin{cases} \frac{u_0}{J}t & (0 \leq t < t_f) \\ \frac{u_0}{J}t_f & (t \geq t_f) \end{cases}. \quad (3.25)$$

By equating  $\theta_f$  with  $u_0 t_f / J$ , we have  $t_f = J\theta_f / u_0$ . The larger the magnitude of the impulse,  $u_0$ , the smaller would be the required time for achieving the desired angle. Such an open-loop control by the application of equal and opposite impulses separated by a given interval is called *bang–bang* control, as depicted in Fig. 3.2.

<sup>1</sup>Note that a torque impulse of magnitude  $u_0$  translates into an input profile (torque vs. time) of area  $u_0$  N m s applied instantaneously, i.e., in zero time.



**Fig. 3.2** Bang–bang rotation of a rigid spacecraft

If the impulses are replaced by pulses of the same area,  $u_0$ , the required time of rotation would be larger than  $t_f = J\theta_f/u_0$ , as seen later. In fact, of all the possible input profiles of a given amplitude, the smallest time of rotation is achieved by a bang–bang profile. This forms the basis of *time-optimal control*.

### 3.2.2 Step Response

Related to the transfer function are responses to other singularity inputs, such as the *unit step* and *unit ramp* functions. It is clear from the integral property of Laplace transform (last row of Table 3.1) that the unit step function,  $u_s$ , is the time integral of the unit impulse function,  $\delta(t)$ ,

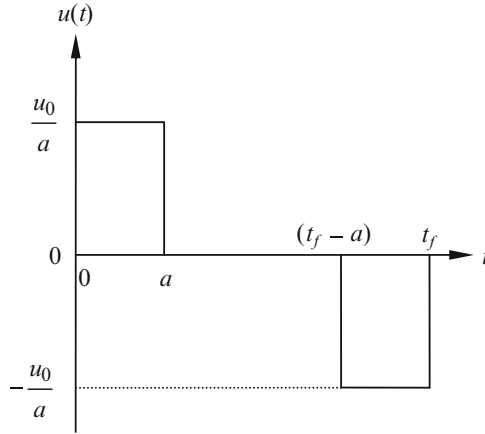
$$u_s(t) = \mathcal{L}^{-1} \left\{ \frac{1}{s} \right\} = \int_0^t \delta(\tau) d\tau \quad (t > 0). \quad (3.26)$$

Similarly, the *step response*,  $y_s(t)$ , of an LTI system,  $G(s)$ , defined as the output when the input is a unit step function applied at  $t = 0$  with a zero initial condition, is the time integral of the impulse response:

$$y_s(t) = \mathcal{L}^{-1} \left\{ \frac{G(s)}{s} \right\} = \int_0^t g(\tau) d\tau \quad (t > 0). \quad (3.27)$$

The step response (also called indicial response) is valuable in studying a stable LTI system's *performance* when a sudden change is desired in the output, and can be derived in a closed-form by partial fraction expansion. The time integral of the step response is the *ramp response*, which is useful in such applications as tracking an object moving with a constant velocity.

**Example 3.6.** For the single-axis rotation of a rigid spacecraft (Example 3.5), consider the application of two equal and opposite torque pulses of area  $u_0$  N m s (width,  $a$ , and height,  $u_0/a$ ) as shown in Fig. 3.3. The torque profile is represented as follows:



**Fig. 3.3** Torque pulses of width  $a$  and height  $u_0/a$  for rigid spacecraft rotation

$$u(t) = \frac{u_0}{a} [u_s(t) - u_s(t-a) - u_s(t-a-t_f) + u_s(t-t_f)]. \quad (3.28)$$

Substitution of (3.28) into (3.22) and integration in time yields

$$\dot{\theta}(t) = \frac{1}{J} \int_0^t u(\tau) d\tau = \begin{cases} \frac{u_0 t}{aJ} & (0 \leq t \leq a) \\ \frac{u_0}{J} & (a < t \leq t_f - a) \\ \frac{u_0}{J} (t_f - t) & (t_f - a < t \leq t_f) \\ 0 & (t > t_f) \end{cases} \quad (3.29)$$

$$\theta(t) = \begin{cases} \frac{u_0 t^2}{2Ja} & (0 \leq t \leq a) \\ \frac{u_0}{J} \left( t - \frac{3}{2}a \right) & (a < t \leq t_f - a) \\ \frac{u_0}{Ja} \left[ a(t_f - 2a) - \frac{1}{2}(t_f - t)^2 \right] & (t_f - a < t \leq t_f) \\ \frac{u_0}{J} (t_f - 2a) & (t > t_f) \end{cases}. \quad (3.30)$$

Therefore, the final rotation produced by the torque pulses is  $\theta_f = u_0(t_f - 2a)/J$ . Thus, for a given rotation angle,  $\theta_f$ , the required time is

$$t_f = 2a + \frac{J}{u_0} \theta_f,$$

which is larger than the time required by a bang–bang impulse profile,  $J\theta_f/u_0$ , in Example 3.5.

*Example 3.7.* Many SISO systems are of second order, governed by the following differential equation:

$$m\ddot{y}(t) + c\dot{y}(t) + ky(t) = u(t), \quad (3.31)$$

where the constants  $m, c, k$  are referred to as *inertia*, *damping*, and *stiffness* parameters, respectively. The transfer function of the second-order system is

$$\frac{Y(s)}{U(s)} = \frac{1}{ms^2 + cs + k}, \quad (3.32)$$

or in the traditional form,

$$\frac{Y(s)}{U(s)} = \frac{1}{m(s^2 + 2\zeta\omega_n s + \omega_n^2)}, \quad (3.33)$$

where  $\omega_n \doteq \sqrt{\frac{k}{m}}$  is called the *natural frequency* and  $\zeta \doteq \frac{c}{2m\omega_n}$  is the *damping-ratio*. It is clear that the stability of the second-order system depends upon the roots of the characteristic equation  $s^2 + 2\zeta\omega_n s + \omega_n^2 = 0$ , which are written as

$$s_{1,2} = -\zeta\omega_n \pm \omega_n \sqrt{\zeta^2 - 1}. \quad (3.34)$$

Applying the linear stability criteria of Chapter 1, it is evident that the system is stable only for  $\zeta \geq 0$ . Furthermore, if  $0 \leq \zeta < 1$ , the eigenvalues are complex conjugates,

$$s_{1,2} = -\zeta\omega_n \pm i\omega_n \sqrt{1 - \zeta^2}, \quad (3.35)$$

and the unforced system's initial response displays an oscillatory behavior with time, while  $\zeta \geq 1$  is the case of real, negative eigenvalues, representing a purely exponential behavior (similar to a first-order LTI system). Most second-order control systems are designed with  $0 \leq \zeta < 1$ , because the overdamped case of  $\zeta \geq 1$  generally requires large control input magnitudes.

The performance of a stable second-order control system is analyzed by its step response, derived as follows.

$0 \leq \zeta < 1$ :

$$Y(s) = \frac{(1/m)U(s)}{s^2 + 2\zeta\omega_n s + \omega_n^2} = \frac{1/m}{s(s^2 + 2\zeta\omega_n s + \omega_n^2)}, \quad (3.36)$$

which is expanded by partial fractions into real and complex conjugate poles,

$$Y(s) = \frac{r_1}{s} + \frac{as + b}{s^2 + 2\zeta\omega_n s + \omega_n^2}, \quad (3.37)$$

where

$$r_1 = sY(s)|_{s=0} = \frac{1}{m\omega_n^2} = \frac{1}{k}.$$

By comparing the numerator polynomials of (3.36) and (3.37), we have

$$a = -\frac{1}{k}; \quad b = -\frac{2\zeta\omega_n}{k}.$$

Thus, we write

$$Y(s) = \frac{1}{k} \left[ \frac{1}{s} - \frac{s + 2\zeta\omega_n}{(s + \zeta\omega_n)^2 + \omega_n^2(1 - \zeta^2)} \right], \quad (3.38)$$

or, taking the inverse Laplace transform (Table 3.1):

$$y_s(t) = \frac{1}{k} \left[ 1 - e^{-\zeta\omega_n t} \left\{ \cos(\omega_d t) + \frac{\zeta}{\sqrt{1 - \zeta^2}} \sin(\omega_d t) \right\} \right] u_s(t), \quad (3.39)$$

where  $\omega_d \doteq \omega_n \sqrt{1 - \zeta^2}$  is the *damped natural frequency* of the system.

$\zeta = 1$ :

$$Y(s) = \frac{1}{ks} - \frac{1}{m\omega_n} \left[ \frac{1}{\omega_n(s + \omega_n)} + \frac{1}{(s + \omega_n)^2} \right], \quad (3.40)$$

or, taking the inverse Laplace transform (Table 3.1):

$$y_s(t) = \left[ \frac{1}{k} (1 - e^{-\omega_n t}) - \frac{t}{\sqrt{km}} e^{-\omega_n t} \right] u_s(t). \quad (3.41)$$

$\zeta > 1$ :

$$Y(s) = \frac{1}{ks} + \frac{r_1}{s - p_1} + \frac{r_2}{s - p_2}, \quad (3.42)$$

which, by taking the inverse Laplace transform (Table 3.1), becomes the step response,

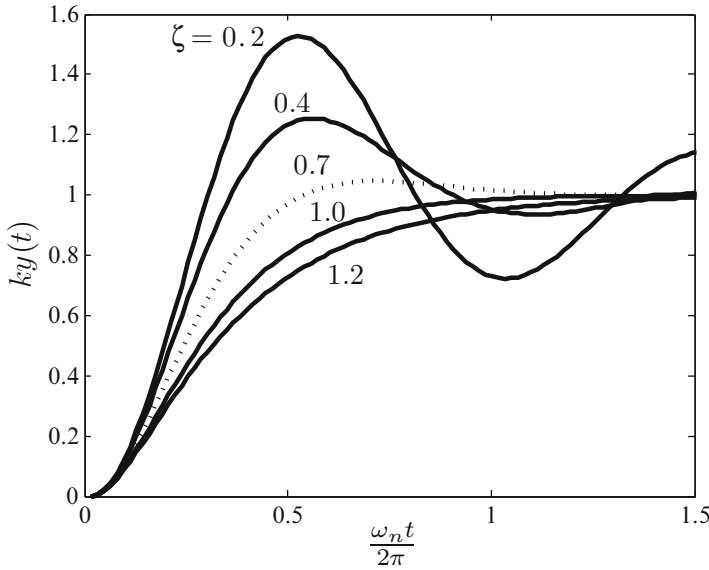
$$y_s(t) = \frac{1}{k} (1 + r_1 e^{p_1 t} + r_2 e^{p_2 t}) u_s(t), \quad (3.43)$$

where

$$p_{1,2} = -\zeta\omega_n \pm \omega_n \sqrt{\zeta^2 - 1},$$

and

$$r_{1,2} = \frac{1}{2\sqrt{\zeta^2 - 1} (\mp\zeta + \sqrt{\zeta^2 - 1})}.$$



**Fig. 3.4** Step response of a second-order, SISO, LTI system

Figure 3.4 shows a plot of  $k y_s(t)$  against the nondimensional time,  $\omega_n t / (2\pi)$ , for different values of  $\zeta$  in the stable range. As  $\zeta$  is increased, the system takes a longer time to reach the desired value of unity for the first time. This time is called the *rise time* and indicates the speed (or alacrity) of the system's response to a step input. On the other hand, the time taken by the system to achieve an equilibrium (or steady) state after being disturbed by a step input is indicated by the *settling-time*, which is defined as the time taken by the amplitude of the step response to settle within  $\pm 2\%$  of the steady-state value. The settling-time and rise time are nearly the same in the nonoscillatory response of system with  $\zeta \geq 1$ , which is a behavior identical to a first-order system. On the other hand, for  $\zeta < 1$ , the system does not stop at the steady state value when it reaches it, but instead crosses it. Such a crossing of the steady state is called an *overshoot*. Repeated crossings of the steady-state value results in an oscillatory response, which decays in amplitude for  $\zeta > 0$ . The maximum overshoot becomes larger and larger as the damping ratio,  $\zeta$ , is reduced as seen in Fig. 3.4, causing an increase in the settling-time.

The following important approximation for the settling-time,  $t_s$ , of step response of a stable second-order system can be obtained by noting that the system essentially behaves like a first-order system<sup>2</sup> in the neighborhood of the steady state (Fig. 3.3):

<sup>2</sup>An asymptotically stable first-order system with transfer function,

$$\frac{Y(s)}{U(s)} = \frac{y_0}{s + k} ; \quad (k > 0),$$

has settling-time  $t_s = \frac{4}{k}$ .

$$e^{-\zeta\omega_n t_s} \simeq 0.02, \quad (3.44)$$

or,

$$t_s \simeq -\frac{\ln(0.02)}{\zeta\omega_n} \simeq \frac{4}{\zeta\omega_n}. \quad (3.45)$$

The rise time, maximum overshoot, and settling-time are important performance parameters of a system. It is evident from Fig. 3.4 that the increase of speed (rise time) and decrease of maximum overshoot are contradictory design requirements. A trade-off between the two can be obtained by selecting a value of  $\zeta$  which is neither too large (causing a sluggish response) nor too small (leading to a large overshoot). In most design applications,  $\zeta = \frac{1}{\sqrt{2}} = 0.707$  (shown as dotted line in Fig. 3.4) is considered ideal in striking a compromise between the speed and the maximum overshoot.

The concept of first- and second-order system analysis can be extended to higher-order systems, which have one or two poles (eigenvalues) nearest the imaginary axis and the other poles further away from it. Because the system's transient response is primarily governed by the poles that have the smallest real part magnitude, such poles are called the *dominant poles* of the system. Most higher-order physical systems can be approximated by a dominant second-order system (i.e., a pair of complex conjugate poles), called the primary (or predominant) *mode*. The other poles occurring either singly or in pairs are treated as less dominant (or secondary) modes that have a smaller contribution to the response, in proportion to the magnitudes of their respective real parts. Such an approximation is commonly applied to linear dynamical systems and is called *modal analysis*. For this reason, the second-order system is a valuable analysis tool, and terms such as damping-ratio and natural frequency can be used to denote either primary or secondary modes represented by each pair of complex conjugate poles.

### 3.2.3 Frequency Response

Apart from the singularity functions, some smooth test functions can be applied as inputs for analyzing an unknown physical system. The response of an LTI system to a simple harmonic input at a particular frequency, subject to zero initial condition, is called *frequency response*, and reveals important control system properties such as the natural frequencies, and the behavior in the presence of high-frequency, unmodeled disturbances, called *noise*. Consider the applied harmonic input,

$$u(t) = u_0 \cos(\omega t) u_s(t), \quad (3.46)$$

where  $\omega$  is called the frequency of excitation. As the LTI system is initially at rest, and the input begins acting at  $t = 0$ , the system's output at a large time,  $t \geq T > 0$ , is also harmonic and is given by the frequency response,

$$y_f(t) = y_0 \cos(\omega\{t - T\}) u_s(t - T), \quad (3.47)$$

provided the system is asymptotically stable, i.e., the transient response due to the step input,  $u_0 u_s(t)$ , decays to zero in the limit  $t \rightarrow \infty$ , or practically vanishes at  $t = T$ . This implies shifting the zero initial condition from  $t = 0$  to  $t = T$ , and taking the Laplace transform of the frequency response with the time-shift property (ninth row, Table 3.1):

$$Y_f(s) = \mathcal{L}\{y_f(t)\} = y_0 e^{Ts} \mathcal{L}\{\cos(\omega t) u_s(t)\}. \quad (3.48)$$

As the Laplace transform of the harmonic input is given by  $U(s) = u_0 \mathcal{L}\{\cos(\omega t) u_s(t)\}$ , we write

$$G(s) = \frac{Y_f(s)}{U(s)} = \frac{y_0 e^{Ts}}{u_0}, \quad (3.49)$$

or, in the harmonic limit,  $s = i\omega$ , we have

$$G(i\omega) = \frac{Y_f(i\omega)}{U(i\omega)} = \frac{y_0 e^{i\omega T}}{u_0}. \quad (3.50)$$

Here, we introduce the *gain* of the frequency response,  $y_0/u_0$ , as the magnitude of  $G(i\omega)$ ,

$$\frac{y_0}{u_0} = |G(i\omega)|, \quad (3.51)$$

and the *phase angle*,  $\phi(\omega) = \omega T$ , as the phase of  $G(i\omega)$ :

$$\phi = \omega T = \tan^{-1} \frac{\Im\{G(i\omega)\}}{\Re\{G(i\omega)\}}. \quad (3.52)$$

Thus, the frequency response of an LTI, SISO system is given by the transfer function on the imaginary axis of the Laplace domain,  $s = i\omega$ . In phasor representation, the gain and phase of the frequency response are related to the harmonic transfer function by

$$G(i\omega) = |G(i\omega)| e^{i\phi(\omega)}. \quad (3.53)$$

*Example 3.8.* Consider the response of the spring, mass, and damper system of Example 3.7 with  $0 \leq \zeta < 1$  to the simple harmonic input, (3.46). From Table 3.1, for the zero initial condition, we have

$$Y(s) = \frac{(1/m)U(s)}{(s + \zeta\omega_n)^2 + \omega_d^2} = \frac{(u_0/m)s}{[(s + \zeta\omega_n)^2 + \omega_d^2](s^2 + \omega^2)}, \quad (3.54)$$

where,  $\omega_d = \omega_n \sqrt{1 - \zeta^2}$ . Expansion of the response into two quadratic denominator factors yields the following expression:

$$Y(s) = \frac{u_0}{m} \left[ \frac{a/s + b}{(s + \zeta\omega_n)^2 + \omega_d^2} + \frac{d/s + e}{s^2 + \omega^2} \right]. \quad (3.55)$$

Comparing the coefficients of the numerator polynomial, we have

$$\begin{aligned} a &= -d\alpha^2 \\ b &= -e = \frac{2\omega_n\zeta}{\omega^2(\alpha^2 - 1)}d \\ d &= \frac{1 - \alpha^2}{(1 - \alpha^2)^2 + 4\alpha^2\zeta^2}, \end{aligned}$$

where

$$\alpha = \omega_n/\omega.$$

By taking the inverse Laplace transform of (3.55) we obtain the following expression for the output in the time domain:

$$\begin{aligned} y(t) &= \frac{u_0}{m} \left[ \frac{a}{\omega_n^2} - e^{-\zeta\omega_n t} \left\{ \frac{a}{\omega_n^2} \cos(\omega_d t) + \left( \frac{a\zeta - b\omega_n}{\omega_n\omega_d} \right) \sin(\omega_d t) \right\} \right. \\ &\quad \left. + \frac{c}{\omega^2} \{1 - \cos(\omega t)\} - \frac{b}{\omega} \sin(\omega t) \right] u_s(t). \end{aligned} \quad (3.56)$$

Note that the term in the first row on the right-hand side of (3.56) decays to  $au_0/k$  in the limit  $t \rightarrow \infty$  (or practically for  $t > T$ ), thereby leaving the term in the second row on the right-hand side as the only function of time. Furthermore, the zero initial condition is easily seen to be satisfied,

$$y(0) = \frac{u_0}{m} \left( \frac{a}{\omega_n^2} + \frac{c}{\omega^2} \right) = 0. \quad (3.57)$$

Hence, the frequency response of the system,  $y_f(t)$ , expressed in the time domain is the following:

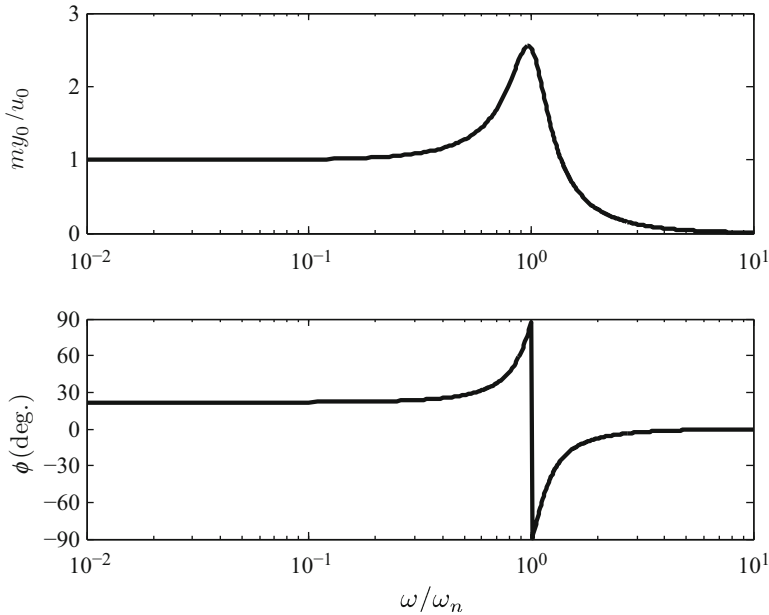
$$y_f(t) = -\frac{u_0}{m} \left[ \frac{c}{\omega^2} \cos(\omega t) + \frac{b}{\omega} \sin(\omega t) \right] u_s(t - T). \quad (3.58)$$

Comparison with (3.47),

$$y_f(t) = y_0[\cos(\omega t) \cos(\omega T) + \sin(\omega t) \sin(\omega T)]u_s(t - T), \quad (3.59)$$

reveals the phase,

$$\tan \phi = \tan(\omega T) = \frac{b\omega}{d} = \frac{2\zeta\omega_n}{\omega_n^2 - \omega^2}, \quad (3.60)$$



**Fig. 3.5** Frequency response of a spring, mass, and damper system for  $\zeta = 0.2$

and the gain,

$$\frac{y_0}{u_0} = \frac{1}{m\omega^2} \sqrt{d^2 + \left( \frac{2d\zeta\omega\omega_n}{\omega_n^2 - \omega^2} \right)^2}. \quad (3.61)$$

Clearly, the gain,  $|G(i\omega)|$ , attains a maximum value at  $\omega = \omega_n$ , for a given damping ratio,  $0 \leq \zeta < 1$ . This is the well known condition of *resonance* where the excitation at the natural frequency builds up the response to a maximum amplitude by continually pumping-in energy. Furthermore, it is evident from (3.60) that the phase angle,  $\phi(\omega)$ , changes by  $\pm 180^\circ$  at  $\omega = \omega_n$ . While resonance is a second-order behavior, it can be used to isolate the modes of a higher-order system (provided they are not clustered together). Thus, by exciting a stable, higher-order system at various frequencies and studying its gain and phase, we can identify its distinct natural frequencies. Such an approach of system identification is called *spectrum analysis*. Figure 3.5 shows a plot of  $m y_0 / u_0$  and the phase of the frequency response for  $\zeta = 0.2$  in the frequency range,  $0.01 \leq (\omega/\omega_n) \leq 10$ . Note the peak corresponding to the natural frequency, the rapid decay of gain at higher frequencies, and the change of phase by  $-180^\circ$  at the natural frequency.

Typically, the gain of the frequency response is expressed on a logarithmic *decibel* scale:

$$\text{Gain} = 20 \log_{10} |G(i\omega)| \text{ (dB)}. \quad (3.62)$$

A plot of gain (dB) and phase ( $^{\circ}$ ) against a logarithmic frequency scale is termed a *Bode plot*. Bode plots are commonly used for studying stability, performance, and robustness properties of control systems. Each second-order subsystem corresponding to a quadratic factor in the characteristic polynomial generally appears as a peak in the gain plot and a phase change by  $\pm 180^{\circ}$  (unless it is affected by the numerator polynomial). Thus, the associated natural frequencies of the respective modes can be easily identified in the Bode plot. The gain at zero frequency,  $|G(0)|$ , is called *DC gain* of the system, and is useful in calculating the steady state response to a step input. The range of frequencies ( $0 \leq \omega \leq \omega_b$ ) over which the gain stays above  $0.707 |G(0)|$  is called the system's *bandwidth* and is denoted by the frequency,  $\omega_b$ . The bandwidth indicates the highest frequency signal to which the system has an appreciable response. For higher frequencies ( $\omega > \omega_b$ ), the gain of a strictly proper system generally declines rapidly with increasing frequency, called *roll-off*. As noise is typically a high-frequency signal, a system with a large roll-off has a reduced sensitivity to noise – a desirable property for a robust system.

The roll-off and phase of a linear system in the high-frequency limit,  $\omega \rightarrow \infty$ , tell us about the difference in the degree of numerator and denominator polynomials of the transfer function. For example, the roll-off is approximately  $-20(n - m)$  dB/decade, where  $n$  is the degree of the denominator (characteristic) polynomial and  $m$  is that of the numerator polynomial of the transfer function,  $G(s)$ .

*Example 3.9.* Consider the longitudinal dynamics of an aircraft described by the following transfer function:

$$G(s) = -\frac{(s + 0.01)(s + 0.5)}{(s^2 + 0.001s + 0.005)(s^2 + 1.5s + 1)}. \quad (3.63)$$

The computation of the numerator and denominator polynomials, followed by the poles and their corresponding natural frequencies and damping-ratios, is carried out using basic MATLAB commands:

```
>> n1=[-1 -.01];n2=[1 .5];num=conv(n1,n2) % numerator polynomial
num =    -1.0000    -0.5100    -0.0050

>> d1=[1 .001 .005];d2=[1 1.5 1];den=conv(d1,d2) % denominator polynomial
den =    1.0000    1.5010    1.0065    0.0085    0.0050

>> pole=roots(den) % poles of the system
pole =
   -0.7500 + 0.6614i
   -0.7500 - 0.6614i
   -0.0005 + 0.0707i
   -0.0005 - 0.0707i

>> wn=abs(pole), zeta=-real(pole)./wn % natural freq., damping-ratio
wn =
  1.0000e+000
  1.0000e+000
  7.0711e-002
  7.0711e-002
```

```

zeta =
    7.5000e-001
    7.5000e-001
    7.0711e-003
    7.0711e-003

```

Hence, the system has two modes, each corresponding to a pair of complex conjugate poles. The lower frequency mode at  $\omega_n = 0.707$  rad/s is lightly damped ( $\zeta = 0.007$ ), while the higher frequency mode at  $\omega_n = 1$  rad/s is well damped ( $\zeta = 0.75$ ). The Bode plot of the plant in the frequency range  $10^{-4} \leq \omega \leq 100$  rad/s is generated using the following basic MATLAB commands:

```

>> w=logspace(-4,2);
>> N=num(:,1)*(i*w).^2+num(:,2)*(i*w)+num(:,3)*ones(size(w,1));
>> D=den(:,1)*(i*w).^4+den(:,2)*(i*w).^3+den(:,3)*(i*w).^2 ...
    +den(:,4)*(i*w)+den(:,5)*ones(size(w,1));
>> G=N./D;
>> gain=20*log10(abs(G));
>> phase=angle(G)*180/pi+180;
>> subplot(212),semilogx(w,phase),grid,xlabel('Frequency (rad/s)'),
    ylabel('Phase (deg.)'),
    subplot(211),semilogx(w,gain),grid,ylabel('Gain (dB)')

```

Alternatively, the Bode plot can be created more easily using the following MATLAB-Control Systems Toolbox (CST) commands:

```

>> sys=tf(num,den)

Transfer function:
          -s^2 - 0.51 s - 0.005
-----
s^4 + 1.501 s^3 + 1.007 s^2 + 0.0085 s + 0.005
>> bode(sys)

```

The resulting Bode plot is shown in Fig. 3.6. Note the  $180^\circ$  phase change corresponding to the mode of natural frequency 0.707 rad/s, which we will call the first mode. A sharp peak in the gain plot is also evident at  $\omega = 0.707$  rad/s signifying the dominance of the first mode. The second mode at  $\omega = 1$  rad/s is well damped and the corresponding pole is close to the system's zeros, hence the corresponding peak and  $180^\circ$  phase change are absent. This indicates that the second mode does not contribute significantly to the system's frequency response for the concerned output variable. The DC gain is 0 dB and the bandwidth is seen to be approximately 1.05 rad/s. A roll-off of  $-40$  dB per decade is evident in Fig. 3.6, indicating low noise sensitivity (high robustness) at high frequencies. As expected, the phase tends to  $-180^\circ$  in the limit  $\omega \rightarrow \infty$ .

The step response of the system,

```

>> t=0:0.1:100;
>> step(sys,t)

```

plotted in Fig. 3.7, confirms the dominance of the first mode in the system's response with a characteristic time period of approximately  $2\pi/(.707) = 89$  s and the light damping causing an amplitude decay of about 5% per cycle. This indicates the fact that the given transfer function can be approximated by only the first mode.

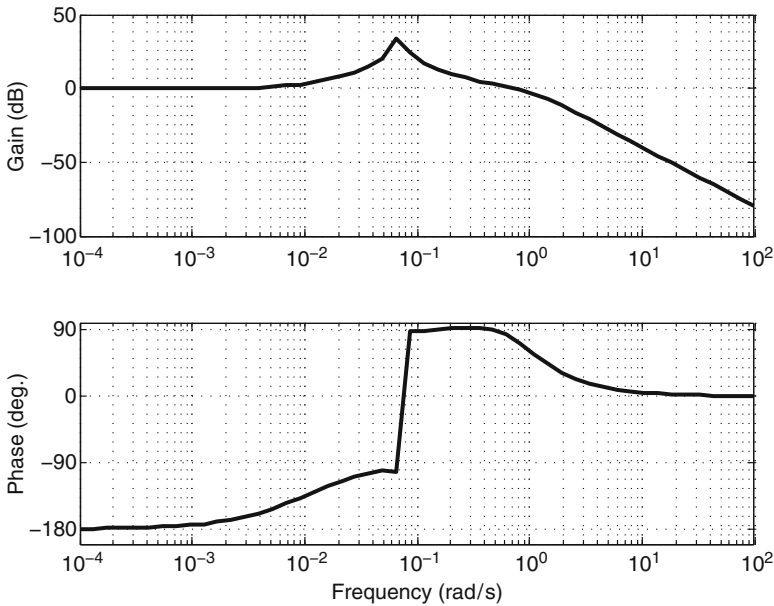


Fig. 3.6 Frequency response of aircraft longitudinal dynamics

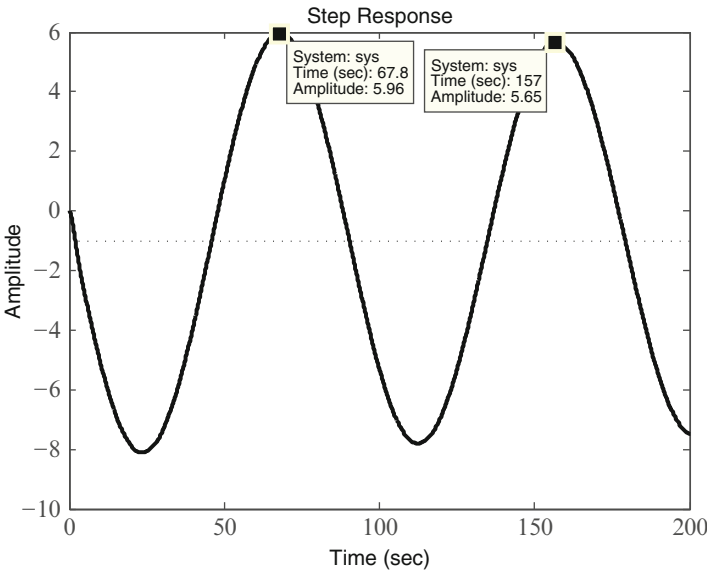


Fig. 3.7 Step response of aircraft longitudinal dynamics

### 3.3 Single Variable Design

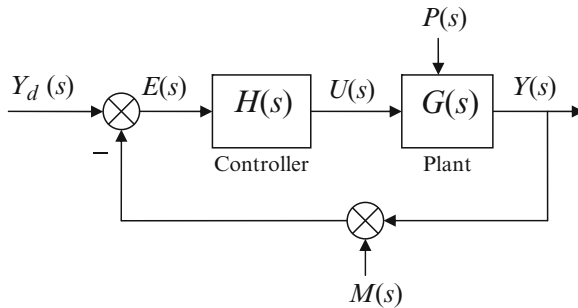
The basic closed-loop system consists of a plant in a feedback loop with a controller, as shown in Fig. 3.8. As we are interested in controlling single-input, single-output, linear time-invariant plants, the same characteristics are usually present in the controller, and both plant and controller are modeled by their respective transfer functions,  $G(s)$  and  $H(s)$ . However, the plant model is never exact and actually contains some unmodeled dynamics that we regard as random process noise. Furthermore, the plant output,  $y(t)$ , is neither sensed exactly nor fed back to the controller in a pristine form, but generally contains an additive random error called the measurement noise.

The design objective of the control system is to maintain the plant output,  $y(t)$ , close to the desired (or commanded) output,  $y_d(t)$ , by driving the error,  $e(t) = y_d(t) - y(t)$ , through the application of control input,  $u(t)$ , in the presence of unmodeled disturbance inputs,  $p(t)$  (process noise), and  $m(t)$  (measurement noise). All the variables are assumed to be Laplace transformable [20]. The Laplace transforms are indicated by upper-case symbols of the respective time-domain variables. In the basic controller design, we ignore the presence of the disturbance inputs; thus from Fig. 3.8, we have

$$\begin{aligned} E(s) &= Y_d(s) - Y(s) \\ Y(s) &= G(s)U(s) \\ U(s) &= H(s)E(s), \end{aligned} \tag{3.64}$$

which yield the following transfer function of the closed-loop system without disturbances:

$$\frac{Y(s)}{Y_d(s)} = \frac{G(s)H(s)}{1 + G(s)H(s)} \tag{3.65}$$



**Fig. 3.8** Closed-loop, SISO, LTI control system

The closed-loop poles are the roots of the following characteristic equation:

$$1 + G(s)H(s) = 0, \quad (3.66)$$

If the command signal,  $y_d(t)$ , is zero the control system is called a regulator; otherwise, the control system is referred to as a tracking system. As a tracking system is expected to respond to a variety of command signals (including step inputs), it must have asymptotic stability as its primary feature. Therefore, the closed-loop poles must all lie in the left-half  $s$ -plane. Other closed-loop design features can be derived from the performance and robustness requirements.

### 3.3.1 Steady-State Error

In order to assess the closed-loop performance, we must know what are the likely desired outputs (command signals) the system is expected to respond to. For a given desired output,  $y_d(t)$ , the *steady-state error*,  $e_{ss}$ , can be obtained by applying the *final-value theorem* of the Laplace transform [20]:

$$e_{ss} = \lim_{s \rightarrow 0} sE(s) = \lim_{s \rightarrow 0} \frac{sY_d(s)}{1 + G(s)H(s)}. \quad (3.67)$$

The steady-state error crucially depends upon the *type* of the control system shown in Fig. 3.8. The system's type is defined as the number of pure integrators (or poles at  $s = 0$ ) of the open-loop transfer function,  $G(s)H(s)$ . Thus, a control system having one pole of  $G(s)H(s)$  at origin is called Type I system, the one with two poles at origin is said to be of Type II, and so on. As the most common control application involves tracking a step command,  $Y_d(s) = y_0/s$ , the steady-state error in such a case is given by:

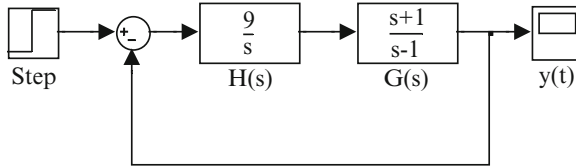
$$e_{ss} = \lim_{s \rightarrow 0} sE(s) = \lim_{s \rightarrow 0} \frac{sy_0Y_d(s)/s}{1 + G(s)H(s)} = \frac{y_0}{1 + G(0)H(0)}. \quad (3.68)$$

Clearly, if the system is of Type I (or higher), its response to a step command will result in a zero steady-state error. For a ramp command,  $Y_d(s) = y_0/s^2$ , a control system of Type II (or higher) will produce a zero steady-state error. If the open-loop plant,  $G(s)$ , has an insufficient number of poles at origin then its type can be increased by putting poles at  $s = 0$  in the controller transfer function,  $H(s)$ .

*Example 3.10.* A plant has the following transfer function:

$$G(s) = \frac{s + 1}{s - 1}. \quad (3.69)$$

Devise a controller in the feedback configuration of Fig. 3.8 such that the closed-loop response to a unit step input has zero steady-state error with a settling-time of 3 s.



**Fig. 3.9** Simulink block-diagram of the control system of Example 3.10

The plant is unstable due to the pole,  $s = 1$ . Furthermore, it is of Type 0 because there are no poles at the origin. Hence, we select a controller transfer function for both stabilizing and increasing the type of the control system as follows:

$$H(s) = \frac{k}{s}, \quad (3.70)$$

where the pole  $s = 0$  adds integral action to the forward path. The closed-loop transfer function is then given by:

$$\frac{Y(s)}{Y_d(s)} = \frac{G(s)H(s)}{1 + G(s)H(s)} = \frac{k(s+1)}{s^2 + (k-1)s + k}. \quad (3.71)$$

Thus, the closed-loop characteristic equation is the following:

$$1 + G(s)H(s) = s^2 + (k-1)s + k = 0,$$

which implies a natural frequency,  $\omega_n = \sqrt{k}$ , and damping ratio,

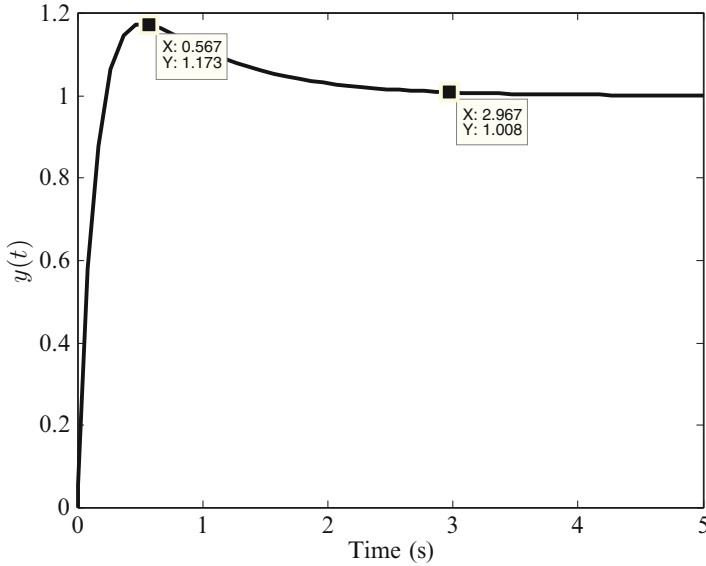
$$\zeta = \frac{k-1}{2\sqrt{k}}.$$

For meeting the settling-time requirement, we choose  $k = 9$ , which yields,  $\omega_n = 3$  rad/s,  $\zeta = 4/3$ . The system's step response is simulated with the Simulink model of Fig. 3.9, and the result plotted in Fig. 3.10. Note the maximum overshoot of 17.3%. A reduction in the value of  $k$  increases both the maximum overshoot and the settling-time. For having a better control over the closed-loop performance, one must be able to change both the natural frequency and damping ratio in an independent manner, which requires an additional controller constant.

### 3.3.2 Proportional-Integral-Derivative Control

For the feedback control system of Fig. 3.8, the simplest control law is *proportional control*,

$$u = ke = k(y_d - y), \quad (3.72)$$



**Fig. 3.10** Closed-loop step response of the control system of Example 3.10

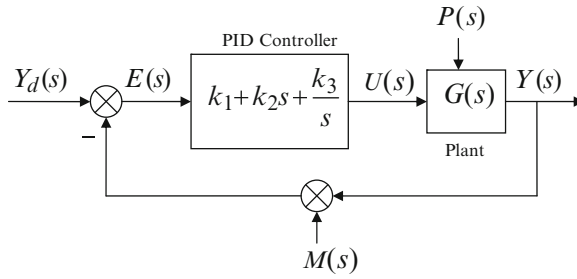
where  $k$  is a positive constant, called the *feedback gain*. With a suitable choice of the gain, an acceptable closed-loop performance may be obtained. However, this is not always guaranteed, as the proportional feedback can amplify even small errors (perhaps due to the measurement noise), leading to a high sensitivity to disturbances. Thus, the proportional controller often acts analogous to a spring. In order to damp out the oscillations caused by the proportional gain, either the plant must have inherent damping or an artificial damping mechanism must be provided by the controller. Therefore, a *proportional-derivative* (PD) control,

$$u = k_1 e + k_2 \dot{e}, \quad (3.73)$$

is often applied instead of proportional control. A properly designed PD controller produces a satisfactory transient behavior, but can cause a large steady-state error,  $e_{ss} \doteq \lim_{t \rightarrow \infty} e(t)$ , in plants which do not have a sufficient number of poles at the origin ( $s = 0$ ). In such cases, it becomes imperative to add a controller pole at the origin, which translates into an integral action in the time domain and increased type of the system. The resulting control law is termed *proportional-integral-derivative* (PID), and is given by:

$$u = k_1 e + k_2 \dot{e} + k_3 \int e dt. \quad (3.74)$$

The controller gains,  $k_1, k_2, k_3$ , are suitably chosen by a design process called *PID tuning* in order to achieve an acceptable response, as well as a low sensitivity



**Fig. 3.11** Block diagram of a proportional-integral-derivative (PID) control system

to unmodeled disturbances. Due to their excellent closed-loop performance and ease of design, PID controllers are commonly used in SISO control systems. They are especially useful in controlling a second-order plant as the controller gains,  $k_1, k_2, k_3$ , provide the necessary parameters for suitably placing the poles of the resulting third-order control system. A block diagram of PID controller is shown in Fig. 3.11.

*Example 3.11.* As an application of the PID control system, consider the aircraft long-period (phugoid) longitudinal dynamics (Chap. 4) involving a slow oscillation in airspeed,  $y(t)$  m/s, caused by elevator deflection,  $u(t)^\circ$  with the following transfer function:

$$\frac{Y(s)}{U(s)} = G(s) = \frac{-s^2 + 99.5s + 50}{2,000s^2 + 2s + 10}. \quad (3.75)$$

It is desired to put in place a controller that enables a unit step change in airspeed by elevator deflection in about 30 s.

Since the plant does not have a pole at origin, we require integral action by the controller in order to make the steady-state error vanish for a step command. Furthermore, the plant has very large settling-time (8,000 s) due to low damping. Hence, we choose a proportional-integral (PI) controller with the transfer function

$$H(s) = k_1 + \frac{k_2}{s} \quad (3.76)$$

in order to both reduce the settling-time (with some help from the numerator polynomial of  $G(s)$ ) and the steady-state error. The resulting closed-loop system is of third-order with transfer function given by:

$$\begin{aligned} \frac{Y(s)}{Y_d(s)} &= \frac{G(s)H(s)}{1 + G(s)H(s)} \\ &= \frac{(-s^2 + 99.5s + 50)(k_1s + k_2)}{(2,000 - k_1)s^3 + (99.5k_1 - k_2 + 2)s^2 + (50k_1 + 99.5k_2 + 10)s + 50k_2}. \end{aligned} \quad (3.77)$$

Clearly, we cannot place all three poles at desired locations with only two controller constants. Happily, we don't need to do so as we have only the settling-time requirement, which can be easily met by properly selecting only the *real parts* of the closed-loop poles. To this end, we write the desired closed-loop characteristic polynomial as follows:

$$1 + G(s)H(s) = (s + a)(s^2 + 2\zeta\omega_n s + \omega_n^2), \quad (3.78)$$

and note that the real parts of the closed-loop poles are  $a$  and  $-\zeta\omega_n$ . Since we want the dominant closed-loop dynamics to be of second-order (as the plant), the real pole,  $s = -a$ , must be placed further into the left-half s-plane than the complex conjugate poles,  $s_{2,3} = -\zeta\omega_n \pm i\omega_n \sqrt{1 - \zeta^2}$ . Thus, we specify  $\zeta\omega_n = 4/30$  ( $\approx 4/t_s$ ) and  $a = 1/3$ . Upon comparing the desired characteristic polynomial with the numerator polynomial of the closed-loop transfer function (3.77) we have

$$\begin{aligned} k_1 &= 12.05425 \\ k_2 &= 8.633789 \\ \omega_n &= \sqrt{\frac{150k_2}{2000 - k_1}} = 0.8071 \text{ rad/s} \\ \zeta &= 0.1652. \end{aligned} \quad (3.79)$$

In order to verify our design, we compute the closed-loop step response by the following MATLAB-CST commands:

```
>> num=[-1 99.5 50];den=[2000 2 10];sys=tf(num,den) % Plant
>> k2=8.633789; k1=12.05425; % Controller constants
>> numb=conv(num,[k1 k2]);denb=[2000-k1 99.5*k1-k2+2 50*k1+99.5*k2+10 50*k2];
    sysb=tf(numb,denb)
% Closed-loop system

Transfer function: -12.05 s^3 + 1,191 s^2 + 1,462 s + 431.7
-----
1,988 s^3 + 1,193 s^2 + 1,472 s + 431.7

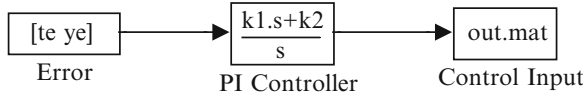
>> [y,t]=step(sysb);

>> syserr=tf(conv([1 0],den),denb)

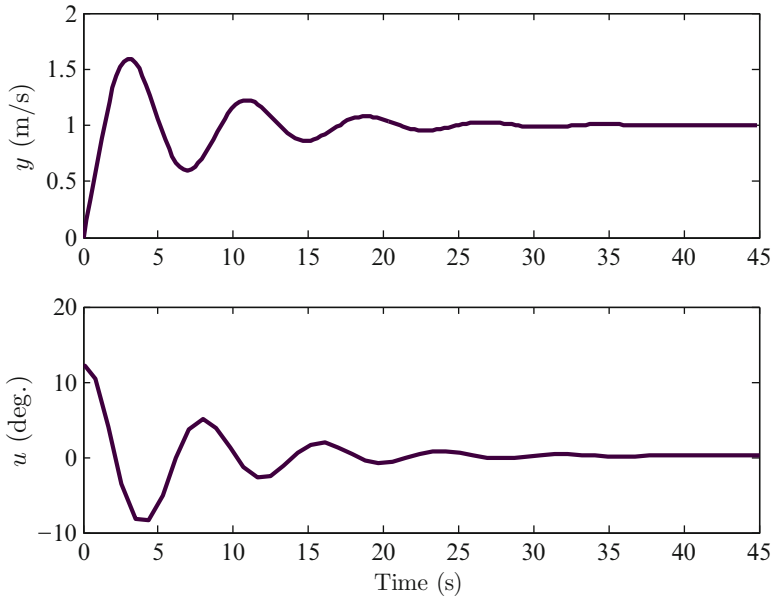
Transfer function:
      2,000 s^3 + 2 s^2 + 10 s
-----
1988 s^3 + 1,193 s^2 + 1,472 s + 431.7

>> [ye,te]=step(syserr); % System error
```

From the system error calculated earlier, we can estimate the required control input using the Simulink block diagram of Fig. 3.12. The closed-loop step response and control input are plotted in Fig. 3.13. The settling-time is quite close to the desired value of 30 s, while the maximum elevator deflection is slightly greater than  $10^\circ$  – an acceptable magnitude for a linear control system.



**Fig. 3.12** Simulink block diagram for PI control input simulation



**Fig. 3.13** Closed-loop step response for aircraft with PI airspeed tracking system

### 3.3.2.1 Roll Autopilot for Aircraft and Banking Missiles

Rolling motion of high-speed aircraft and bank-to-turn missiles requires automatic control that is often designed using single-input techniques. In such applications, the wings/fins are of much smaller span than a conventional aircraft, hence the aerodynamic coupling among roll, yaw, and sideslip can be considered negligible. Even when there is a significant roll-yaw-sideslip coupling due to larger span wings, the pure rolling motion is well damped and has a much smaller time-scale than the yaw-sideslip dynamics. Thus, a predominantly rolling mode generally exists – called the *pure rolling mode* – and a roll autopilot exercises control over it for quickly achieving a desired bank angle. Consider the pure roll dynamics of the vehicle represented by:

$$\begin{aligned}
 J_{xx}\ddot{\sigma} &= -L_p p + L_A \delta_A \\
 p &= \dot{\sigma}
 \end{aligned}
 \tag{3.80}$$

resulting in the following transfer function:

$$\frac{\Sigma(s)}{\Delta_A(s)} = \frac{L_A}{J_{xx}s^2 - L_P s}. \quad (3.81)$$

As the roll dynamics plant lacks asymptotic stability, a PD feedback control is required to ensure a well damped response to a step banking command,  $\sigma_d(t)$ , resulting in the following closed-loop transfer function:

$$\frac{\Sigma(s)}{\Sigma_d(s)} = \frac{\frac{L_A}{J_{xx}}(K_P + K_D s)}{s^2 + \left(K_D \frac{L_A}{J_{xx}} - \frac{L_P}{J_{xx}}\right)s + \frac{L_A}{J_{xx}} K_P}. \quad (3.82)$$

A suitable choice of the controller gains,  $K_P, K_D$ , results in the desired settling-time and maximum overshoot. Due to the open-loop transfer function having a pole at the origin, the steady-state error to a step command is zero. The bank angle,  $\sigma(t)$ , is sensed and fed back through a vertical gyro whose gain is included in the PD roll controller.

In addition to the roll dynamics, the aileron's rotation about its hinge is modeled by the following second-order dynamics:

$$J_A \ddot{\delta}_A = H_{\delta_A} \dot{\delta}_A + H_{\delta_A} \delta_A + K_A u, \quad (3.83)$$

where  $u(t)$  is the applied input voltage to the DC motor whose dynamics is modeled simply by the gain,  $K_A$ . As the aileron hinge-moment damping derivative,  $H_{\delta_A}$ , is seldom sufficient to produce acceptable transient characteristics, a closed-loop control system – called *aileron servo* – is invariably required. The servo must also provide integral action for ensuring zero steady-state error to a step command,  $\delta_{Ad}(t)$ . Hence, a PID control with gains  $k_1, k_2, k_3$  is selected for the aileron servo, which has the following closed-loop transfer function:

$$\frac{\Delta_A(s)}{\Delta_{Ad}(s)} = \frac{\frac{K_A}{J_A}(k_3 + k_1 s + k_2 s^2)}{s^3 + \left(k_2 \frac{K_A}{J_A} - \frac{H_{\delta_A}}{J_A}\right)s^2 + \left(\frac{K_A}{J_A} k_1 - \frac{H_{\delta_A}}{J_A}\right)s + \frac{K_A}{J_A} k_3}. \quad (3.84)$$

The servo must have a faster transient response than that of the roll control loop. A likely choice for the gains,  $k_1, k_2, k_3$ , is such that the servo dynamics is ten times faster than the closed-loop roll dynamics (3.82), resulting in the following servo poles:

$$\begin{aligned} s_1 &= -\zeta\omega \\ s_{2,3} &= -\zeta\omega \pm \omega\sqrt{1 - \zeta^2}, \end{aligned} \quad (3.85)$$

where  $\zeta = 1/\sqrt{2}$  and

$$\omega = 10 \sqrt{\frac{L_A}{J_{xx}}} K_P. \quad (3.86)$$

There is usually a limit imposed on the motor input voltage,  $|u(t)| \leq u_{\max}$ . Such a limit – called *saturation* – renders the control system nonlinear, even though its design is based upon an LTI system. Alternatively, saturation limits may be individually imposed on the aileron deflection,  $\delta_A(t)$ , and its rate,  $\dot{\delta}_A(t)$ . Input saturation as well as process and measurement noise cause a departure of the autopilot from its nominal design performance that should be carefully analyzed before implementation.

### 3.3.3 Feedforward/Feedback Tracking

An alternative to PID control is feedforward/feedback (FF/FB) tracking of command signals wherein the integral action of the PID control is replaced by a suitable feedforward path in order to reduce the steady-state error. Consider the block diagram of a SISO FF/FB control system shown in Fig. 3.14, with output,  $y(t)$ , and error,  $e(t) = y_d(t) - y(t)$ . The plant transfer function,  $G(s)$ , does not have an adequate number of poles at the origin in order to make  $e_{ss} \rightarrow 0$  for a given command signal (desired output),  $y_d(t)$ . Furthermore, the feedback controller,  $H(s)$ , lacks integral action, to compensate for which a feedforward path,  $F(s)$  is added. The closed-loop transfer function is thus the following:

$$\frac{Y(s)}{Y_d(s)} = \frac{G(s)[F(s) + H(s)]}{1 + G(s)H(s)}, \quad (3.87)$$

and the error is given by

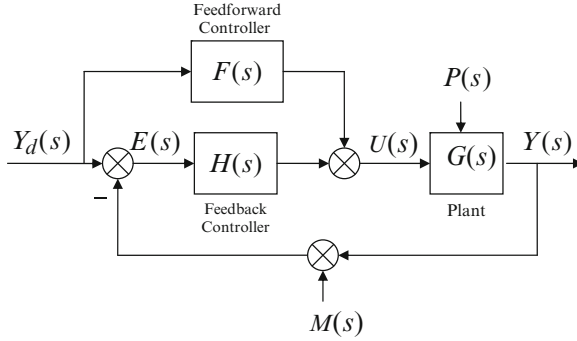
$$\frac{E(s)}{Y_d(s)} = \frac{1 - G(s)F(s)}{1 + G(s)H(s)}. \quad (3.88)$$

Clearly, the closed-loop characteristic polynomial,  $1 + G(s)H(s)$ , is unaffected by the feedforward path. However, the steady-state error,

$$e_{ss} = \lim_{s \rightarrow 0} sE(s) = \lim_{s \rightarrow 0} \left\{ sY_d(s) \frac{1 - G(s)F(s)}{1 + G(s)H(s)} \right\}, \quad (3.89)$$

is greatly influenced by  $F(s)$  since  $G(s)H(s)$  does not have an adequate number of poles at  $s = 0$  for  $e_{ss}$  to vanish identically. Therefore, the primary function of the feedforward path is to reduce the steady-state error, by providing a similar effect as that of a pole at origin in  $G(s)H(s)$ .

The feedforward path requires that the command signal,  $y_d(t)$ , be measured separately from the output,  $y(t)$ . This requirement is usually more stringent than the



**Fig. 3.14** Block diagram of a feedforward/feedback (FF/FB) tracking system

error measurement required in a pure feedback control. Typically, a command signal is either operator specified or generated as an output of another control system; hence, it can often be measured quite easily.

*Example 3.12.* An air-to-air missile autopilot is tasked to track a step normal acceleration signal,  $Y_d(s) = 0.1/s$  g, commanded by an outer guidance loop. The transfer function of the missile's rigid body, short-period longitudinal dynamics with control-surface deflection input is given by:

$$G(s) = -\frac{0.05(s^2 - 50)}{s^2 + s + 10} \text{ g/deg.} \quad (3.90)$$

A proportional feedback controller,  $H(s) = 19.75$ , is selected in order to have a closed-loop settling-time of 0.1 s, with the attendant damping ratio,  $\zeta = 0.58$ . However, as neither  $G(s)$  nor  $H(s)$  has a pole at origin, the steady-state error for a step command is nonzero. Hence, a proportional feedforward controller,  $F(s) = K$ , is added such that  $e_{ss} = 0$ , which requires

$$e_{ss} = \lim_{s \rightarrow 0} sE(s) = \frac{1 - KG(0)}{1 + 19.75G(0)} = 0, \quad (3.91)$$

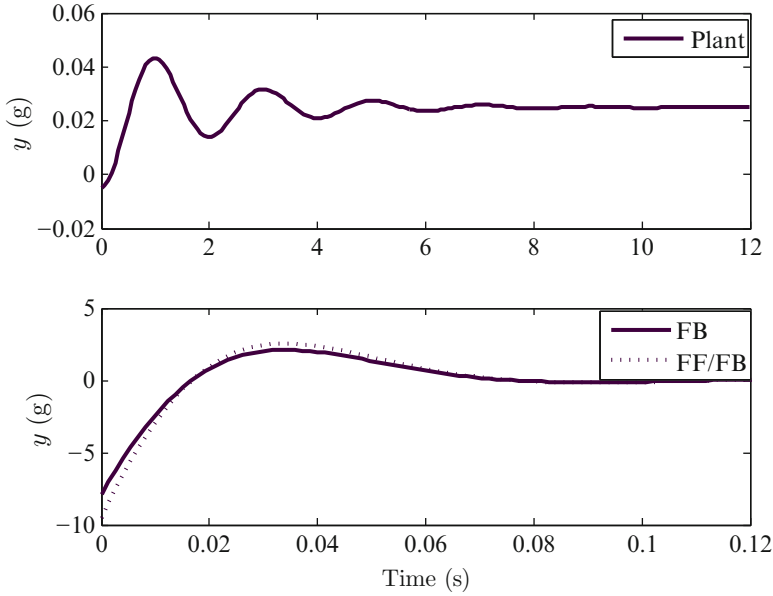
or

$$K = \frac{1}{G(0)} = 4. \quad (3.92)$$

The responses of the plant, the pure feedback control system, and the FF/FB system are compared in Fig. 3.15 generated using the following MATLAB-CST commands:

```
>> num=-.05*[1 0 -50];den=[1 1 10];sys=tf(num,den) % Plant

Transfer function:
-0.05 s^2 + 2.5
-----
s^2 + s + 10
```



**Fig. 3.15** Step response of an air-to-air missile longitudinal dynamics plant compared with a pure feedback (FB) controller and a feedforward/feedback (FF/FB) tracking system

```
>> numfb=19.75*num;denfb=[(1-0.05*19.75) 1 (10+2.5*19.75)];
sysfb=tf(numfb,denfb)% FB system

Transfer function:
      -0.9875 s^2 + 49.38
      -----
      0.0125 s^2 + s + 59.38

>> numfffb=(4+19.75)*num;sysfffb=tf(numfffb,denfb) % FF/FB system

Transfer function:
      -1.188 s^2 + 59.38
      -----
      0.0125 s^2 + s + 59.38

>> [y,t]=step(sys);
>> [yb,tb]=step(sysfb);
>> [yf,tf]=step(sysfffb);
>> subplot(211),plot(t,y/10),subplot(212),plot(tb,yb/10,tf,yf/10)
```

The steady state outputs of the plant, the FB and the FF/FB systems to the applied command  $Y_d(s) = 0.1/s$  g are the following:

$$y_{ss} = \lim_{s \rightarrow 0} s Y_d(s) G(s) = 0.025 \text{ g} \quad (\text{Plant})$$

$$y_{ss} = \lim_{s \rightarrow 0} s Y_d(s) \frac{G(s)H(s)}{1 + G(s)H(s)} = (0.1) \frac{49.38}{59.38} = 0.0832 \text{ g} \quad (\text{FB})$$

$$y_{ss} = \lim_{s \rightarrow 0} s Y_d(s) \frac{G(s)[F(s) + H(s)]}{1 + G(s)H(s)} = (0.1) \frac{59.38}{59.38} = 0.1 \text{ g} \quad (\text{FF/FB})$$

Thus, only the FF/FB system has a zero steady-state error, as expected.

### 3.3.4 Robustness Analysis from Frequency Response

*Robustness* of a control system is related to its sensitivity to unmodeled dynamics, i.e., part of the behavior of an actual control system, which is not included in the mathematical model of the control system. As we have to deal with actual control systems in which it is difficult to mathematically model all relevant physical processes, we are always left with the question: will the control system based upon a mathematical model really work? We have already discussed the presence of random disturbances in the form of process and measurement noise. If a control system meets its performance and stability objectives in the presence of such disturbances then the control system is said to be robust. Hence, robustness is a desirable property that dictates whether a control system is immune to uncertainties in its mathematical model. More specifically, robustness can be subdivided into *stability robustness* and *performance robustness*, depending upon whether we are interested in range of controller parameters for which either the stability, or the performance of the control system is guaranteed.

Taking the case of a feedback control system with, we can write the closed-loop characteristic equation as follows:

$$1 + G(s)H(s) = 0, \quad (3.93)$$

which becomes the following on the imaginary axis of the Laplace domain (frequency domain),  $s = i\omega$ :

$$1 + G(i\omega)H(i\omega) = 0. \quad (3.94)$$

Note that (3.94) satisfies the closed-loop system's characteristic equation *if and only if* any of the roots of (3.93) falls on the imaginary axis,  $s = i\omega$ . As the closed-loop system would become unstable if any of its poles crosses over into the right-half  $s$ -plane (Chap. 1), the margin of stability (i.e., stability robustness) is measured by how far the poles are from crossing the imaginary axis, or in other words, how shy is the locus of the open-loop frequency response,  $G(i\omega)H(i\omega)$ , from the point  $-1$ . In the Laplace domain, the point  $-1$  is denoted by the gain of unity and a phase of  $\pm 180^\circ$ . Thus, the stability robustness is determined by the proximity of the open-loop frequency response from a gain of unity and a phase of  $\pm 180^\circ$ . The gain of the frequency response when its phase is  $\pm 180^\circ$  is thus called *gain margin* and indicates the factor by which the controller gain,  $|H(i\omega)|$ , can be increased before the closed-loop system becomes unstable. Similarly, the phase difference of

$G(i\omega)H(i\omega)$  from  $\pm 180^\circ$  when its gain is unity is called the *phase margin*. Note that a negative gain margin indicates an unstable closed-loop system. The frequency at which the phase of  $G(i\omega)H(i\omega)$  crosses the  $\pm 180^\circ$  line is called the *phase crossover frequency*, while the frequency at which  $|G(i\omega)H(i\omega)|$  becomes unity is called the *gain crossover frequency*.

*Example 3.13.* Let us find the gain and phase margins of the aircraft longitudinal dynamics plant of Example 3.9. From the frequency response shown in Fig. 3.6, we note that the phase of  $\pm 180^\circ$  is never crossed, which implies an infinite gain margin. The gain of unity (0 dB) is crossed at the gain crossover frequency of 0.74 rad/s, at which the phase is approximately  $77^\circ$ . Therefore, the phase margin is  $180 - 77 = 103^\circ$ .

Another important measure of a control system's robustness is its ability to reject high-frequency process and measurement noise. A system that has a rapid decline of the gain,  $|G(i\omega)|$ , with frequency (roll-off) beyond the bandwidth of the system, is little affected by random, high-frequency noise inputs. An adequate roll-off is considered to be about  $-20$  dB per decade of frequency in rad/s.

### 3.3.4.1 Lead and Lag Compensation for Robust Design

Often the controller transfer function must be compensated for the unmodeled dynamics of a control system, such as higher (neglected) modes of the plant and actuator nonlinearities, which are invariably present in the actual system. Such a compensation may involve placing a simple, linear system (called a *filter*) in series at the controller output. Since the open-loop DC gain (that has already been selected for a small steady-state error) must be unaffected by a filter, its transfer function should be such that only a phase change occurs in a desired frequency range. In order to increase or decrease the speed of the closed-loop step response, the filter must introduce either a phase-lead or phase-lag, respectively. Hence, such a filter is appropriately called either a lead or a lag compensator and can be used to increase the gain and phase margins or the roll-off of the linear control system. The compensating filters are generally chosen to be of first-order to avoid resonant peaks associated with second- and higher-order filters.

A first-order filter has the following transfer function:

$$F(s) = \frac{s + b}{as + b}, \quad (3.95)$$

where  $a, b$  are real coefficients. For a lag compensator, we have  $a > 1$ , whereas  $a < 1$  for the lead compensator. A lag compensator is useful in reducing the steady-state error of Type 0 plants (i.e., plants without any pure integrators), and decreasing the gain of the closed-loop system at high frequencies (which is desirable for reducing the sensitivity to the measurement noise). However, lag compensation slows down the closed-loop transient response. Lag compensation is relatively

simple to use, because the passive circuit through which it can be implemented is quite inexpensive. On the other hand, a lead compensator is useful in increasing the speed of the closed-loop response, and increasing the phase margin of the closed-loop system, which also results in smaller overshoots of the transient response. Lead compensation usually requires amplification of error signals through a relatively expensive electrical circuit. Also, lead compensation increases the gain at high frequencies, which is undesirable due to increased sensitivity to measurement noise.

In order to combine the useful features of both lead and lag compensators, one can have a lead–lag compensator with the following transfer function:

$$F(s) = \left( \frac{s + b_1}{s + b_1/\alpha} \right) \left( \frac{s + b_2}{s/\alpha + b_2} \right), \quad (3.96)$$

where  $\alpha > 1$ . Note that the lead–lag compensator applies a phase reduction in the approximate frequency range,  $b_1/\alpha \leq \omega \leq b_1$ , and also increases the DC gain of the system, while the increasing the phase in the approximate frequency range  $b_2 \leq \omega \leq \alpha b_2$ . In the limit  $\alpha \rightarrow \infty$ , the lead–lag compensator becomes a PID controller. Hence, one can approximate PID controller by a passive lead–lag electrical network with a large value of the positive constant,  $\alpha$  [20].

### 3.4 Multivariable Control Design

The concept of transfer function ceases to be useful for multi-input, multi-output (MIMO) plants, for which the state-space methods are more appropriate. The design methodology for MIMO control systems is thus carried out directly in the time domain, rather than in the Laplace (or frequency) domain. Consider the plant dynamics expressed in a linear, state space form as follows:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}, \quad (3.97)$$

where  $\mathbf{x}(t)$  is the *state vector*,  $\mathbf{u}(t)$  the control input vector, and  $\mathbf{A}(t), \mathbf{B}(t)$  are the state space coefficient matrices that could be time varying. However, for the time being, we will confine our attention to a linear plant with constant coefficient matrices,  $\mathbf{A}, \mathbf{B}$ , i.e., linear time-invariant (LTI) system.

#### 3.4.1 Regulator Design by Eigenstructure Assignment

Design of a linear, state feedback regulator for the LTI plant of (3.97) with the control law,

$$\mathbf{u} = -\mathbf{K}\mathbf{x}, \quad (3.98)$$

is possible by assigning a structure for the eigenvalues and eigenvectors of the closed-loop dynamics matrix,  $\mathbf{A} - \mathbf{BK}$ . In case of single-input plants, this merely involves selecting the locations for the closed-loop poles (*pole-placement*) by the following *Ackermann's formula* that yields the desired closed-loop characteristics [20]:

$$\mathbf{K} = (\mathbf{a}_d - \mathbf{a})(\mathbf{PP}')^{-1}, \quad (3.99)$$

where  $\mathbf{a}$  is the row vector formed by the coefficients,  $a_i$ , of the plant's characteristic polynomial in *descending order* [ $\mathbf{a} = (a_n, a_{n-1}, \dots, a_2, a_1)$ ]:

$$|s\mathbf{I} - \mathbf{A}| = s^n + a_n s^{n-1} + a_{n-1} s^{n-2} + \dots + a_2 s + a_1, \quad (3.100)$$

$\mathbf{a}_d$  is the row vector formed by the characteristic coefficients of the closed-loop system in descending order [ $\mathbf{a}_d = (a_{dn}, a_{d(n-1)}, \dots, a_{d2}, a_{d1})$ ];

$$|s\mathbf{I} - \mathbf{A} + \mathbf{BK}| = s^n + a_{dn} s^{n-1} + a_{d(n-1)} s^{n-2} + \dots + a_{d2} s + a_{d1}, \quad (3.101)$$

$\mathbf{P}$  is the controllability test matrix of the plant, and  $\mathbf{P}'$  is the following upper triangular matrix:

$$\mathbf{P}' = \begin{pmatrix} 1 & a_{n-1} & a_{n-2} & \cdots & a_2 & a_1 \\ 0 & 1 & a_{n-1} & \cdots & a_3 & a_2 \\ 0 & 0 & 1 & \cdots & a_4 & a_3 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \cdots & 1 & a_{n-1} \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}. \quad (3.102)$$

Of course, this requires that the plant must be controllable,  $|\mathbf{P}| \neq 0$ . A popular choice of the closed-loop poles is the *Butterworth pattern* [20] wherein all the poles are equidistant from the origin,  $s = 0$ . Such a pattern generally requires the least control effort for a given bandwidth, and is thus considered to be optimal placement of poles.

*Example 3.14.* Find the regulator gain matrix that can move the poles of the following single-input plant to  $s_{1,2} = -1 \pm i$ ,  $s_{3,4} = -2 \pm 2i$ , and  $s_{5,6} = -3 \pm 3i$ :

$$\mathbf{A} = \begin{pmatrix} 0 & -1 & 3 & 0 & 1 & -2 \\ -1 & 0.5 & 0 & -1 & 0 & 1 \\ 0 & 1 & -1 & 2 & -1 & 1 \\ 0 & -2 & -1 & 0 & 0.5 & 1 \\ -1 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1 \end{pmatrix}; \quad \mathbf{B} = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 2 \\ 1 \\ -0.5 \end{pmatrix}.$$

We begin by checking the controllability of the plant as follows:

```
>> A=[0 -1 3 0 1 -2; -1 0.5 0 -1 0 1; 0 1 -1 2 -1 1;
      0 -2 -1 0 0.5 1; -1 0 0 0 1 0; 0 -1 0 0 0 1];
>> B=[-1 0 1 2 1 -0.5]';

>> P=ctrb(A,B) % controllability test matrix

P =
    -1.0000    5.0000    9.0000   -22.2500   18.8750   58.9375
         0    -1.5000   -5.2500   -12.6250    4.6875   -12.9063
    1.0000    1.5000   -7.5000   10.2500   30.3750   13.4375
    2.0000   -1.0000    2.0000   17.5000   15.2500   -15.7500
    1.0000    2.0000   -3.0000   -12.0000   10.2500   -8.6250
   -0.5000   -0.5000    1.0000    6.2500   18.8750   14.1875

>> rank(P) % rank of controllability test matrix

ans =
     6
```

Thus,  $P$  is of full-rank and one can design a regulator by pole-placement.

```
% characteristic coefficients of the plant (in descending order):
>> a=poly(A);a=a(2:7)

%      a_6      a_5      a_4      a_3      a_2      a_1
a =   -1.5000    0.5000   -1.0000   -4.5000   15.0000   -6.0000

>> ad=poly([-1-i,-1+i,-2-2i,-2+2i,-3-3i,-3+3i]);ad=ad(2:7)

% desired characteristic coefficients (in descending order):
%      a_d6 a_d5 a_d4 a_d3 a_d2 a_d1
ad =   288   528   484   240    72    12

% Upper triangular matrix follows:
>> Pprime=[1 a(1) a(2) a(3) a(4) a(5); 0 1 a(1) a(2) a(3) a(4);
           0 0 1 a(1) a(2) a(3); 0 0 0 1 a(1) a(2);
           0 0 0 0 1 a(1); 0 0 0 0 0 1]

Pprime =
    1.0000   -1.5000    0.5000   -1.0000   -4.5000   15.0000
         0    1.0000   -1.5000    0.5000   -1.0000   -4.5000
         0         0    1.0000   -1.5000    0.5000   -1.0000
         0         0         0    1.0000   -1.5000    0.5000
         0         0         0         0    1.0000   -1.5000
         0         0         0         0         0    1.0000

>> K=(ad-a)*inv(P*Pprime) % regulator gain matrix

K =   53.7609   16.3270   52.2065   49.8244  -96.4548  -23.7209
```

It is prudent to check whether the closed-loop poles are indeed at the desired locations, which we do as follows:

```
>> eig(A-B*K) % eigenvalues of the closed-loop system

ans =
   -3.0000 + 3.0000i
   -3.0000 - 3.0000i
   -2.0000 + 2.0000i
   -2.0000 - 2.0000i
   -1.0000 + 1.0000i
   -1.0000 - 1.0000i
```

Alternatively, one can employ the MATLAB-Control Systems Toolbox (CST) command *acker* for computing the gains by a single statement:

```
>> K=acker(A,B,[-1-i,-1+i,-2-2i,-2+2i,-3-3i,-3+3i])
K =    53.7609    16.3270    52.2065    49.8244   -96.4548   -23.7209
```

The pole-placement method is inapplicable to multi-input plants, which have many more controller gains to be found than the number of equations available from the pole locations. In such a case, we need additional equations that can be derived from the shape of the eigenvectors using the method of *eigenstructure assignment*. A popular method in this regard is the *robust pole assignment method* of [10] wherein the eigenvectors,  $\mathbf{v}_i$ ,  $i = 1, 2, \dots, n$ , corresponding to the eigenvalues,  $\lambda_i$ , respectively, and satisfying the eigenvalue problem,

$$(\mathbf{A} - \mathbf{BK})\mathbf{v}_i = \lambda_i \mathbf{v}_i, \quad (3.103)$$

are chosen such that the modal matrix,

$$\mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n), \quad (3.104)$$

is as well-conditioned<sup>3</sup> as possible. An eigenstructure assignment algorithm of [10] is coded in the MATLAB-CST command called *place*.

*Example 3.15.* Repeat the regulator design of Example 3.14 if the matrix  $\mathbf{B}$  is replaced by the following:

$$\mathbf{B} = \begin{pmatrix} -1 & 0 \\ 0 & -0.5 \\ 1 & 0 \\ 0 & -1 \\ 2 & 0 \\ 0 & 0.5 \end{pmatrix}.$$

The eigenstructure assignment is carried out by the MATLAB-CST command, *place*, as follows:

```
>> B=[-1 0;0 -0.5;1 0;0 -1;2 0;0 0.5];
>> [K,prec]=place(A,B,[-1-i,-1+i,-2-2i,-2+2i,-3-3i,-3+3i])

K =
    11.0966   -30.3800     3.6203    18.3561     7.6079    -0.5582
    -8.7649    20.7055    -4.3927   -13.6522    -4.1298     4.9219

prec = 13 % indicates precision of pole-placement (13th decimal place)
```

---

<sup>3</sup>Conditioning of a square matrix refers to how close the matrix is from being singular (i.e., has a zero determinant). A scalar measure called the *condition number* is assigned to the matrix that reflects its conditioning. A large condition number implies the matrix is close to being singular.

```
>> eig(A-B*K) % eigenvalues of the closed-loop system

ans =
-3.0000 + 3.0000i
-3.0000 - 3.0000i
-2.0000 + 2.0000i
-2.0000 - 2.0000i
-1.0000 + 1.0000i
-1.0000 - 1.0000i
```

### 3.4.2 Linear, Quadratic Regulator

An alternative design strategy to eigenstructure assignment for MIMO control systems is the method of linear optimal control. An optimal state feedback regulator with gain matrix,  $\mathbf{K}$ , is designed with the control law,

$$\mathbf{u} = -\mathbf{K}\mathbf{x}, \quad (3.105)$$

for minimizing the following quadratic, infinite time objective function with symmetric cost parameter matrices,  $(\mathbf{Q}, \mathbf{R})$ ,

$$\mathcal{J} = \frac{1}{2} \int_0^{\infty} [\mathbf{x}^T(\tau)\mathbf{Q}\mathbf{x}(\tau) + \mathbf{u}^T(\tau)\mathbf{R}\mathbf{u}(\tau)] d\tau, \quad (3.106)$$

subject to the state equation constraint, (3.97). The stabilizing steady state solution for the feedback gain matrix is given by (Appendix A)

$$\mathbf{K} = \mathbf{R}^{-1}\mathbf{B}^T\mathbf{S}, \quad (3.107)$$

where  $\mathbf{S}$  is the constant, symmetric, positive semidefinite solution of the following *algebraic Riccati equation*:

$$\mathbf{S}\mathbf{A} + \mathbf{A}^T\mathbf{S} - \mathbf{S}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T\mathbf{S} + \mathbf{Q} = 0. \quad (3.108)$$

A sufficient condition for the existence of a unique, positive semidefinite solution to (3.108) is that the plant  $(\mathbf{A}, \mathbf{B})$  is controllable,  $\mathbf{R}$  is positive definite, and  $\mathbf{Q}$  is positive semidefinite (Appendix A). A regulator designed in such a way is termed linear, quadratic regulator (LQR).

**Example 3.16.** Re-design the regulator for the multi-input plant of Example 3.15 by LQR method with the cost matrices,  $\mathbf{Q} = \mathbf{I}_{6 \times 6}$ ,  $\mathbf{R} = \mathbf{I}_{2 \times 2}$ .

The design is carried out by the MATLAB-CST command *lqr* as follows:

```
>> [K,S,E]=lqr(A,B,eye(6),eye(2)) % LQR design

K =
    6.0852   -17.1092    2.2125    10.2854    5.2204   -2.0313
   -6.3399    14.1600   -3.1633   -10.0378   -2.9616    4.5359
```

```
% Positive semi-definite solution to algebraic Riccati equation:

S =
    18.5603   -46.3295     8.2108    29.7079     8.2174     0.4066
   -46.3295    122.9885   -18.5601   -77.0397   -22.4393   -2.7710
     8.2108   -18.5601     4.8149    12.8668     2.8041     0.8470
    29.7079   -77.0397    12.8668    49.9388    13.5633     2.7622
     8.2174   -22.4393     2.8041    13.5633     5.3168    -1.2358
     0.4066    -2.7710     0.8470     2.7622    -1.2358    11.8252

% Eigenvalues of the closed-loop system:

E =
   -2.3539 + 0.9632i
   -2.3539 - 0.9632i
   -1.2325 + 1.9804i
   -1.2325 - 1.9804i
   -1.5605 + 0.5196i
   -1.5605 - 0.5196i
```

Thus, the new regulator design results in closed-loop eigenvalues,  $s_{1,2} = -2.3539 \pm 0.9632i$ ,  $s_{3,4} = -1.2325 \pm 1.9804i$ , and  $s_{5,6} = -1.5605 \pm 0.5196i$ .

A variation of the state-weighted LQR design is the output weighted LQR (or LQRY) problem with the following objective function:

$$\mathcal{J} = \frac{1}{2} \int_0^\infty \left[ \mathbf{y}^T(\tau) \bar{\mathbf{Q}} \mathbf{y}(\tau) + \mathbf{u}^T(\tau) \bar{\mathbf{R}} \mathbf{u}(\tau) \right] d\tau, \quad (3.109)$$

where  $\mathbf{y}(t)$  satisfies the following output equation:

$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u}. \quad (3.110)$$

The LQRY problem is more general than a corresponding LQR problem, because it results in a cross-weighting term between state and control variables in its objective function (Appendix A):

$$\mathbf{y}^T \bar{\mathbf{Q}} \mathbf{y} + \mathbf{u}^T \bar{\mathbf{R}} \mathbf{u} = \mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{x}^T \mathbf{L} \mathbf{u} + \mathbf{u}^T \mathbf{L}^T \mathbf{x} + \mathbf{u}^T \mathbf{R} \mathbf{u}. \quad (3.111)$$

However, the LQRY problem being more general requires a greater care in selecting the cost parameters for a positive semidefinite Riccati solution, compared to those of the LQR method. Reference [20] discusses the derivation of the algebraic Riccati equation for the LQRY problem, while the MATLAB-CST contains the dedicated *lqry* command for solving the same.

### 3.4.2.1 Frozen LQR Design for a Time-Varying Plant

While we have derived the LQR gain matrix for the LTI plant, the same approach can be extended to a slowly time varying plant. In such a case, we allow a variation of the cost coefficients,  $\mathbf{Q}$ ,  $\mathbf{R}$ , with time and express the cost function as follows:

$$\mathcal{J}_\infty(t) = \frac{1}{2} \int_0^\infty [\mathbf{x}^T(\tau)\mathbf{Q}(\tau)\mathbf{x}(\tau) + \mathbf{u}^T(\tau)\mathbf{R}(\tau)\mathbf{u}(\tau)] d\tau, \quad (3.112)$$

subject to the state constraint equation

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t) \quad \mathbf{x}(0) = \mathbf{x}_0, \quad (3.113)$$

and the linear, full-state feedback control law

$$\mathbf{u}(t) = -\mathbf{K}_\infty(t)\mathbf{x}(t). \quad (3.114)$$

The general, linear time-varying optimal control problem results in a matrix Riccati differential equation, which must be integrated in time by a nonlinear programming method (Appendix A). However, the derivation of the quasi-steady gain matrix,  $\mathbf{K}_\infty(t)$ , at a given time involves the assumption that it can stabilize the “time-frozen” plant  $\mathbf{A}(t)$ ,  $\mathbf{B}(t)$ , at each time instant in the same way as the constant feedback gain matrix would stabilize the LTI plant. Therefore, we have

$$\mathbf{K}_\infty(t) = \mathbf{R}^{-1}(t)\mathbf{B}^T(t)\mathbf{P}(t), \quad (3.115)$$

where  $\mathbf{P}_\infty(t)$  is the symmetric, positive semidefinite, quasi-steady solution of the following algebraic Riccati equation:

$$\mathbf{P}_\infty(t)\mathbf{A}(t) + \mathbf{A}^T(t)\mathbf{P}_\infty(t) - \mathbf{P}_\infty(t)\mathbf{B}(t)\mathbf{R}^{-1}(t)\mathbf{B}^T(t)\mathbf{P}_\infty(t) + \mathbf{Q}(t) = \mathbf{0}. \quad (3.116)$$

Thus, the quasi-steady approximation is tantamount to finding a constant, steady state solution to (3.116) at each time instant, provided the variation of  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{Q}$ ,  $\mathbf{R}$  with time is much slower than the settling-time of all the transients of the closed-loop dynamics matrix,  $\mathbf{A} - \mathbf{BK}$ . An efficient choice of  $\mathbf{Q}$ ,  $\mathbf{R}$  in such a case is by constant matrices. We shall have the occasion to apply the frozen LQR design to spacecraft and rocket plants.

Great caution must be exercised while attempting to apply the frozen LQR approach to a rapidly varying plant, as the closed-loop system could be unstable even though the eigenvalues of  $\mathbf{A}(t) - \mathbf{B}(t)\mathbf{K}(t)$  are always in the left-half  $s$ -plane.

### 3.4.3 Linear Observers and Output Feedback Compensators

Control systems with output (rather than state) feedback require observers that can reconstruct the missing information about the system’s states from the input applied to the plant, and the output fed back from the plant. In other words, an observer mimics the plant by generating an estimated state vector,  $\mathbf{x}_o$ , instead of the actual plant state vector,  $\mathbf{x}$ , and supplies it to the regulator. A control system that contains both an observer and a regulator is called a *compensator*. Due to a

decoupling of the observer and plant states in the control system, it is possible to design the regulator and observer separately from each other by what is known as the *separation principle* [20].

### 3.4.3.1 Full-Order Observer

The output equation,

$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u}, \quad (3.117)$$

is used in the design of a *full-order observer* with the following state equation:

$$\dot{\mathbf{x}}_o = (\mathbf{A} - \mathbf{L}\mathbf{C})\mathbf{x}_o + (\mathbf{B} - \mathbf{L}\mathbf{D})\mathbf{u} + \mathbf{L}\mathbf{y}, \quad (3.118)$$

where  $\mathbf{x}_o(t)$  is the estimated state and  $\mathbf{L}$  the observer gain matrix, provided the plant  $(\mathbf{A}, \mathbf{C})$  is observable (Chap. 1). The observer gain matrix,  $\mathbf{L}$ , can be selected in a manner similar to the regulator gain,  $\mathbf{K}$ , by either eigenstructure assignment for the observer dynamics matrix,  $\mathbf{A} - \mathbf{L}\mathbf{C}$ , or linear, quadratic, optimal control where  $\mathbf{A}$  is replaced by  $\mathbf{A}^T$ , and  $\mathbf{B}$  by  $\mathbf{C}^T$ .

The closed-loop control system dynamics with a desired state,  $\mathbf{x}_d(t)$ , and linear feedforward/feedback control with output feedback (Chap. 1),

$$\mathbf{u} = \mathbf{K}_d\mathbf{x}_d + \mathbf{K}(\mathbf{x}_d - \mathbf{x}_o), \quad (3.119)$$

is given by the state equation

$$\begin{Bmatrix} \dot{\mathbf{x}} \\ \dot{\mathbf{x}}_o \end{Bmatrix} = \begin{pmatrix} \mathbf{A} & -\mathbf{B}\mathbf{K} \\ \mathbf{L}\mathbf{C} & \mathbf{A} - \mathbf{B}\mathbf{K} - \mathbf{L}\mathbf{C} \end{pmatrix} \begin{Bmatrix} \mathbf{x} \\ \mathbf{x}_o \end{Bmatrix} + \begin{pmatrix} \mathbf{B}(\mathbf{K} + \mathbf{K}_d) \\ \mathbf{B}(\mathbf{K} + \mathbf{K}_d) \end{pmatrix} \mathbf{x}_d, \quad (3.120)$$

where  $\mathbf{K}, \mathbf{L}$  are separately designed (by the separation principle) and the feedforward gain matrix,  $\mathbf{K}_d$ , is selected to ensure that the closed-loop error dynamics is independent of the desired state,  $\mathbf{x}_d(t)$ , with a given state equation

$$\dot{\mathbf{x}}_d = \mathbf{f}_d(\mathbf{x}_d), \quad (3.121)$$

such that

$$(\mathbf{A} + \mathbf{B}\mathbf{K}_d\mathbf{x}_d) - \mathbf{f}_d(\mathbf{x}_d) = \mathbf{0}. \quad (3.122)$$

Thus, one can design a tracking system for a plant that is both controllable and observable with the available inputs and outputs, respectively, as well as satisfies (3.122) with its desired state vector.

*Example 3.17.* Design a full-order observer for the plant of Example 3.14 with the output coefficient matrix,

$$\mathbf{C} = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

such that the observer's poles are roughly twice as deep in the left-half  $s$ -plane as those of the regulator.

The design is carried out by the MATLAB-CST command *lqr* with  $\mathbf{Q} = 20\mathbf{I}_{6 \times 6}$ ,  $\mathbf{R} = \mathbf{I}_{2 \times 2}$  as follows:

```
>> C=[-1 0 0 0 0 0;0 0 0 0 0 1];
>> rank(ctrb(A',C')) % rank of the observability test matrix

ans = 6

% LQR design of full-order observer:
>> [Lp,S,E]=lqr(A',C',10*eye(6),eye(2)); L=Lp'

L =
-10.2781    1.2284
 13.5273   -11.2823
   7.2149    2.9955
 -33.3984   14.1360
 -54.1484    1.7663
 -1.2284    7.4850

>> eig(A-L*C) % poles of the observer

ans =
-3.6643 + 1.4221i
-3.6643 - 1.4221i
-3.6780
-2.1044 + 1.5877i
-2.1044 - 1.5877i
-1.0478
```

MATLAB-CST has the command *lqe* is dedicated to observer design by the LQR method. The observer designed by LQR method is also referred to as a *Kalman filter*.

### 3.4.3.2 Reduced-Order Observer

When a part of the plant's state vector can be directly obtained from the output vector, it is unnecessary to estimate the entire state vector by a full-order observer. Consider a plant whose state vector is partitioned as follows:

$$\mathbf{x} = (\mathbf{x}_1^T, \mathbf{x}_2^T)^T, \quad (3.123)$$

such that

$$\begin{aligned} \dot{\mathbf{x}}_1 &= \mathbf{A}_{11}\mathbf{x}_1 + \mathbf{A}_{12}\mathbf{x}_2 + \mathbf{B}_1\mathbf{u} \\ \dot{\mathbf{x}}_2 &= \mathbf{A}_{21}\mathbf{x}_1 + \mathbf{A}_{22}\mathbf{x}_2 + \mathbf{B}_2\mathbf{u}. \end{aligned} \quad (3.124)$$

The measurable part of the state vector,  $\mathbf{x}_1$ , can be directly obtained by inversion of the output equation with a square coefficient matrix,  $\mathbf{C}$ :

$$\mathbf{y} = \mathbf{C}\mathbf{x}_1. \quad (3.125)$$

The unmeasurable part,  $\mathbf{x}_2$ , needs estimation by a *reduced-order observer*, and can be expressed as follows:

$$\mathbf{x}_{o2} = \mathbf{L}\mathbf{y} + \mathbf{z}, \quad (3.126)$$

where  $\mathbf{z}$  is the state vector of the reduced-order observer with the following state equation:

$$\dot{\mathbf{z}} = \mathbf{F}\mathbf{z} + \mathbf{H}\mathbf{u} + \mathbf{G}\mathbf{y}, \quad (3.127)$$

whose coefficient matrices are determined from the requirement that the estimation error,  $\mathbf{e}_o = \mathbf{x}_2 - \mathbf{x}_{o2}$ , should go to zero in the steady state, irrespective of the control input and the output [20]:

$$\begin{aligned} \mathbf{F} &= \mathbf{A}_{22} - \mathbf{L}\mathbf{C}\mathbf{A}_{12} \\ \mathbf{G} &= \mathbf{F}\mathbf{L} + (\mathbf{A}_{21} - \mathbf{L}\mathbf{C}\mathbf{A}_{11})\mathbf{C}^{-1} \\ \mathbf{H} &= \mathbf{B}_2 - \mathbf{L}\mathbf{C}\mathbf{B}_1, \end{aligned} \quad (3.128)$$

with the observer gain  $\mathbf{L}$  selected by eigenstructure assignment or Kalman filter (see later) approach, such that all the eigenvalues of the observer dynamics matrix,  $\mathbf{F}$ , are in the left-half  $s$ -plane.

*Example 3.18.* Design a reduced-order observer for the plant of Example 3.17.

We begin by partitioning the state-space coefficient matrices and continue with the aforementioned design steps as follows:

```
>> C=[C(1,1) C(1,6); C(2,1) C(2,6)]

C =
    -1     0
     0     1

>> A11=[A(1,1) A(1,6); A(6,1) A(6,6)]

A11 =
     0    -2
     0     1

>> A12=[A(1,2:5); A(6,2:5)]

A12 =
    -1     3     0     1
    -1     0     0     0

>> A21=[A(2:5,1) A(2:5,6)]

A21 =
    -1     1
     0     1
     0     1
    -1     0

>> A22=A(2:5,2:5)
```

```

A22 =
    0.5000         0   -1.0000         0
    1.0000   -1.0000    2.0000   -1.0000
   -2.0000   -1.0000         0    0.5000
         0         0         0    1.0000

>> B1=[B(1,:);B(2,:)]

B1 =
   -1.0000         0
         0   -0.5000

>> B2=B(2:5,:)

B2 =
         0   -0.5000
    1.0000         0
         0   -1.0000
    2.0000         0

% Reduced-order observer design by LQR method:
>> Lp=lqr(A22',A12'*C',eye(4),eye(2));L=Lp'

L =
    1.6824   -1.7849
    1.6870    0.0196
   -5.5490    1.6157
  -10.2132   -0.1612

>> F=A22-L*C*A12

F =
   -2.9673    5.0473   -1.0000    1.6824
   -0.6674    4.0609    2.0000    0.6870
    5.1648   -17.6471         0   -5.0490
   10.0521   -30.6396         0   -9.2132

>> eig(F) % Poles of reduced-order observer

ans =
   -3.0003
  -2.0452 + 1.4449i
  -2.0452 - 1.4449i
   -1.0290

>> G=F*L+(A21-L*C*A11)*inv(C)

G =
   -7.1115    2.9282
  -12.3865    1.9979
   30.4858    1.7323
   60.3197    3.5311

>> H=B2-L*C*B1

H =
   -1.6824   -1.3924
   -0.6870    0.0098
    5.5490   -0.1921
   12.2132   -0.0806

```

### 3.4.4 Linear, Quadratic, Gaussian (LQG) Compensator

For a multivariable plant, it is often advantageous to employ linear optimal control (Appendix A) for designing both the regulator and the observer. The resulting design is called a linear, quadratic, Gaussian (LQG) compensator, and is known for its good robustness properties. The plant linearized about the nominal trajectory is described by:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{F}\mathbf{p}, \quad (3.129)$$

$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u} + \mathbf{w}, \quad (3.130)$$

where  $\mathbf{y}(t)$  is the measured output,  $\mathbf{p}(t)$  the process noise, and  $\mathbf{w}(t)$  the measurement noise. A full-order observer designed by the steady state, linear optimal control<sup>4</sup> for minimizing a quadratic cost function of white process noise spectral density matrix,  $\bar{\mathbf{P}}$ , measurement noise spectral density matrix,  $\mathbf{W}$ , and cross-spectral density matrix between process and measurement noise,  $\Psi$ , is called a *Kalman filter* [20] and has the following state equation:

$$\dot{\mathbf{x}}_o = (\mathbf{A} - \mathbf{L}\mathbf{C})\mathbf{x}_o + \mathbf{B}\mathbf{u} + \mathbf{L}\mathbf{y}. \quad (3.131)$$

The steady state Kalman filter gain matrix,  $\mathbf{L}$ , is given by:

$$\mathbf{L} = (\mathbf{V}\mathbf{C}^T + \mathbf{F}\Psi)\mathbf{W}^{-1}, \quad (3.132)$$

where  $\mathbf{V}$  is the *optimal covariance matrix* [20], the unique, symmetric and positive semidefinite solution to the following steady state, algebraic Riccati equation:

$$\mathbf{0} = \mathbf{A}_G\mathbf{V} + \mathbf{V}\mathbf{A}_G^T - \mathbf{V}\mathbf{C}^T\mathbf{W}^{-1}\mathbf{C}\mathbf{V} + \mathbf{F}\mathbf{P}_G\mathbf{F}^T, \quad (3.133)$$

and

$$\begin{aligned} \mathbf{A}_G &= \mathbf{A} - \mathbf{F}\Psi\mathbf{W}^{-1}\mathbf{C} \\ \mathbf{P}_G &= \bar{\mathbf{P}} - \Psi\mathbf{W}^{-1}\Psi^T. \end{aligned} \quad (3.134)$$

Note the similarity between the Kalman filter and the LQRY regulator problem. When the process and measurement white noise are uncorrelated (as they usually are), we have  $\Psi = \mathbf{0}$  and a considerable simplification takes place in the Kalman filter derivation. A sufficient condition for the existence of a unique, positive semidefinite solution to (3.133) is that the plant  $(\mathbf{A}, \mathbf{C})$  is observable,  $\mathbf{P}_G$  is positive semidefinite, and  $\mathbf{W}$  is positive definite. Generally, the matrices  $\bar{\mathbf{P}}, \mathbf{W}, \Psi$  are treated as design parameters for achieving a desired robustness through loop transfer recovery (LTR) [20]. Alternative design techniques such as  $\mathcal{H}_\infty$ -control and  $\mu$ -synthesis [20] also achieve robustness objectives in a manner similar to the LQG/LTR method.

---

<sup>4</sup>The extension to a reduced-order observer is easily carried out; thus, there is no loss of generality.

### 3.5 Digital Control System

While we have assumed a continuous time (or *analog*) control system in the foregoing analysis, whenever a control system is implemented in practice, a time delay invariably occurs in processing the feedback signal and generating the corresponding control input by a controller. It is as if the output is held constant for an instant before being fed back to the plant in the form of the control input. Thus, we have a digitized (or *digital*) control system that works in discrete – rather than continuous – time steps. The process of handling continuous time signals in discrete time steps is called *analog-to-digital* (A/D) conversion, and consists of sampling and holding the data for a brief time interval. Every manned or automatic control system is inherently digital, even though the plant and controller themselves are analog systems. A modern digital computer with a fast processor and memory can quickly handle signals in a way that may appear to be continuous due to the very small time steps involved. On the other hand, a human operator who takes a few seconds to process a plant's output is very obviously a digital controller. The only difference between the two is the sampling rate.

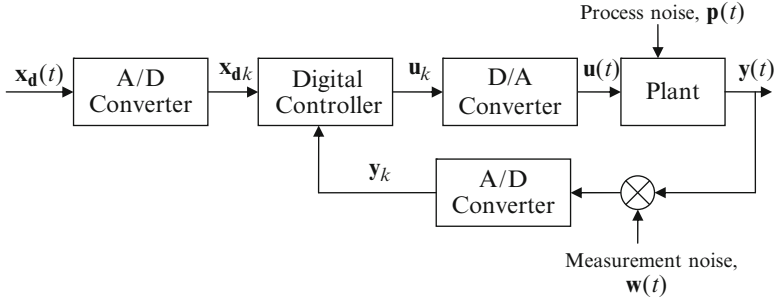
The sampling rate at which an analog signal is processed is crucial for the successful implementation of a digital control system. The sampling rate must not be smaller than the bandwidth (i.e., the range of frequencies in the power spectrum) of the analog signal, otherwise a distorted digital output will be produced. In order to reduce signal distortion, the sampling rate should be at least twice the bandwidth of the analog signal. One half of the sampling rate is known as the *Nyquist frequency*, and is the approximate bandwidth of the digital system. Hence, we may regard the A/D converter as a *low-pass filter* [20] with the Nyquist frequency as the cut-off frequency.

When a digital controller sends a discrete time input to an analog plant, it must be first converted into a continuous signal by a *digital-to-analog* (D/A) convertor, which is the reverse of the sampling and hold process. An example of D/A convertor is playing a series of still photographs on a movie projector. By spinning the wheel of the projector, the time interval between any two shots is reduced to almost zero (compared with the time taken by the human brain to process individual pictures), thereby producing the sensation of continuous motion. The block-diagram of a digital control system is shown in Fig. 3.16.

Consider a first-order digital system with the following discrete time governing equation:

$$x_{k+1} = ax_k + bu_k, \quad (3.135)$$

where the subscript  $k$  denotes the state and input variables at the  $k$ th time step, while the subscript  $k+1$  refers to the variables at the next time step, and  $a, b$  are constant coefficients. Clearly, (3.135) describes the evolution of the system in time (like the differential equation of a continuous time system) and is called a *difference equation*. If, instead of the subscripts, we express the system's future state by an



**Fig. 3.16** A typical digital control system

operator,  $\mathcal{Z}(\cdot)$ , that operates on the present state such that  $\mathcal{Z}(x_{k+1}) = zx$ , where  $z$  is a complex variable, the difference equation is said to be *z-transformed* as follows:

$$zx = ax + bu, \quad (3.136)$$

or

$$x = \frac{bu}{z - a}. \quad (3.137)$$

Clearly, the  $z$ -transform for a digital system is equivalent to Laplace transform for a continuous time system, and is useful in predicting the future state. A single-input, single-output (SISO) digital system is described by a *pulse transfer function*,  $G(z)$ , which is analogous to the transfer function of a continuous time system. Analog-to-digital conversion of a SISO system with transfer function  $G(s)$  thus consists of taking a  $z$ -transform,  $\mathcal{Z}(G(s)) = G(z)$ . For more on  $z$ -transform and pulse transfer function refer to Chap. 8 of [20].

In order to derive difference equations of a multivariable plant, consider the LTI continuous system solution (Chap. 1) held for a sampling interval,  $\Delta t = T$ :

$$\mathbf{x}(t + T) = e^{AT} \mathbf{x}(t) + \int_t^{t+T} e^{A(t+T-\tau)} \mathbf{B} \mathbf{u}(\tau) d\tau, \quad (3.138)$$

or, because  $\mathbf{u}(t)$  is held during the sampling interval,

$$\mathbf{x}(t + T) = e^{AT} \mathbf{x}(t) + e^{AT} \left( \int_0^T e^{-A\tau} d\tau \right) \mathbf{B} \mathbf{u}(t). \quad (3.139)$$

If one assumes that the interval  $T$  is small enough for the integrand,  $e^{-A\tau} d\tau$  to be constant during the sampling process, we have

$$\mathbf{x}(t + T) = e^{AT} \mathbf{x}(t) + T \mathbf{B} \mathbf{u}(t). \quad (3.140)$$

In any case, we can write the following difference equation:

$$\mathbf{x}_{k+1} = \bar{\mathbf{A}}\mathbf{x}_k + \bar{\mathbf{B}}\mathbf{u}_k, \quad (3.141)$$

with

$$\bar{\mathbf{A}} = e^{\mathbf{A}T}; \quad \bar{\mathbf{B}} = e^{\mathbf{A}T} \left( \int_0^T e^{-\mathbf{A}\tau} d\tau \right) \mathbf{B}. \quad (3.142)$$

The derivation of the digital state-space coefficient matrices from the analog ones is thus given by (3.142), which is also programmed in the inbuilt MATLAB-CST functions called *c2d* and *c2dm*. After conversion of the system into a digital form, one can apply techniques similar to those used for the continuous time systems for designing feedback regulators and observers [20].

*Example 3.19.* Derive the digital equivalent of the continuous time plant of Example 3.17 with  $\mathbf{D} = \mathbf{0}$  using a zero-order hold and a sampling interval of  $T = 0.1$  s.

We employ the inbuilt MATLAB-CST function *c2dm* as follows:

```
>> A=[0 -1 3 0 1 -2; -1 0.5 0 -1 0 1; 0 1 -1 2 -1 1;
      0 -2 -1 0 0.5 1; -1 0 0 0 1 0; 0 -1 0 0 0 1];
>> B=[-1 0;0 -0.5;1 0;0 -1;2 0;0 0.5];
>> C=[-1 0 0 0 0 0;0 0 0 0 0 1];D=[0 0;0 0];
>> [Ad,Bd,Cd,Dd] = c2dm(A,B,C,D,0.1,'zoh')
```

```
Ad =
    0.9996    -0.0799    0.2844    0.0333    0.0908   -0.1987
   -0.1026    1.0610   -0.0098   -0.1035   -0.0076    0.1130
    0.0008    0.0724    0.8957    0.1856   -0.0949    0.1139
    0.0077   -0.2150   -0.0944    1.0009    0.0578    0.0888
   -0.1052    0.0045   -0.0150   -0.0012    1.1003    0.0103
    0.0053   -0.1081    0.0003    0.0053    0.0003    1.0996
```

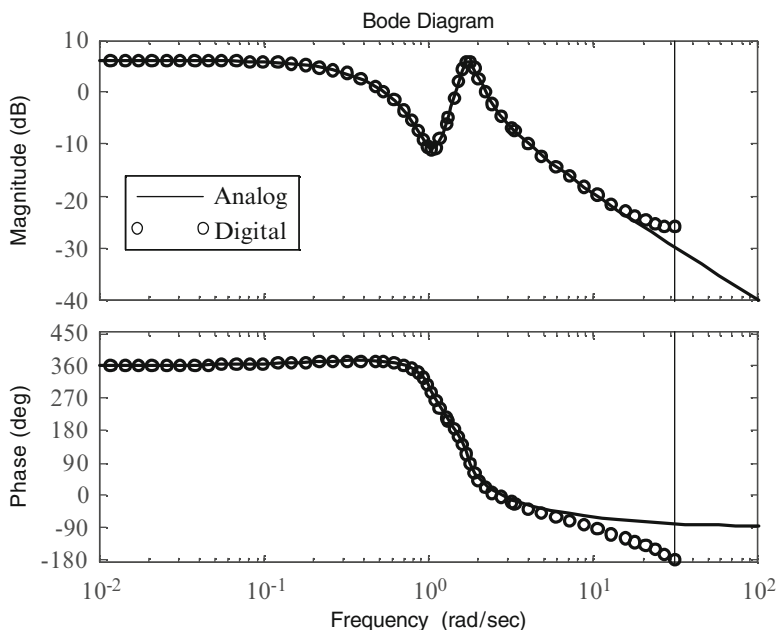
```
Bd =
   -0.0761   -0.0040
    0.0043   -0.0436
    0.0852   -0.0088
    0.0004   -0.0925
    0.2147    0.0001
   -0.0002    0.0550
```

```
Cd =
   -1     0     0     0     0     0
    0     0     0     0     0     1
```

```
Dd =
    0     0
    0     0
```

A digital Bode plot of the analog (solid line) and corresponding digital (circle) systems' transfer function from the first input to the first output – obtained as follows – is shown in Fig. 3.17:

```
>> bode(A,B(:,1),C(1,:),D(1,1),0.1) % Bode plot of analog system
>> hold on,dbode(Ad,Bd(:,1),Cd(1,:),Dd(1,1),0.1)%Bode plot of digital system
```



**Fig. 3.17** Bode plots from the first input to the first output for analog (*solid line*) and digital (*circle*) systems

Note the vertical line at  $\omega = \pi/T = 31.4159$  rad/s (or 5 Hz) indicating the cut-off (Nyquist) frequency. Thus, the digital system's response is limited to below 5 Hz. Furthermore, at the Nyquist frequency, the digital system's phase is shifted to  $-180^\circ$  from the  $-90^\circ$  for the analog system. On the other hand, the gain of the digital system at the Nyquist frequency is increased by about 5 dB compared with that of the analog system.

### 3.6 Summary

Flight control systems are often linear and can be either of single- or multivariable type, which are designed by the transfer function or state-space methods, respectively. While the step response can provide a measure of the single-input, single-output (SISO) system's performance, the frequency response reveals its stability and robustness. The practical SISO designs include the PID and feedforward/feedback tracking systems, which can be compensated for actuator dynamics by using lead or lag compensators. Multivariable regulators and observers can be designed separately by either eigenstructure assignment or the linear, optimal

(LQR) method that requires solving an algebraic Riccati equation. Linear observers designed by the linear, optimal control method are called Kalman filters, and have a guaranteed robustness for white noise disturbances. Digital control systems involve time discretization and require difference equations and  $z$ -transform for design and analysis.

## Exercises

**3.1.** Derive an expression for the impulse response of a spring, mass, and damper system described by the differential equation,

$$m\ddot{y} + c\dot{y} + ky = u,$$

where  $u(t)(=\delta(t))$  is the applied input and  $y(t)(=g(t))$  the output, for underdamped ( $0 \leq \zeta < 1$ ), critically damped, ( $\zeta = 1$ ), and overdamped ( $\zeta > 1$ ) cases.

**3.2.** Derive the ramp response for the system of Exercise 3.1 for underdamped ( $0 \leq \zeta < 1$ ), critically damped, ( $\zeta = 1$ ), and overdamped ( $\zeta > 1$ ) cases.

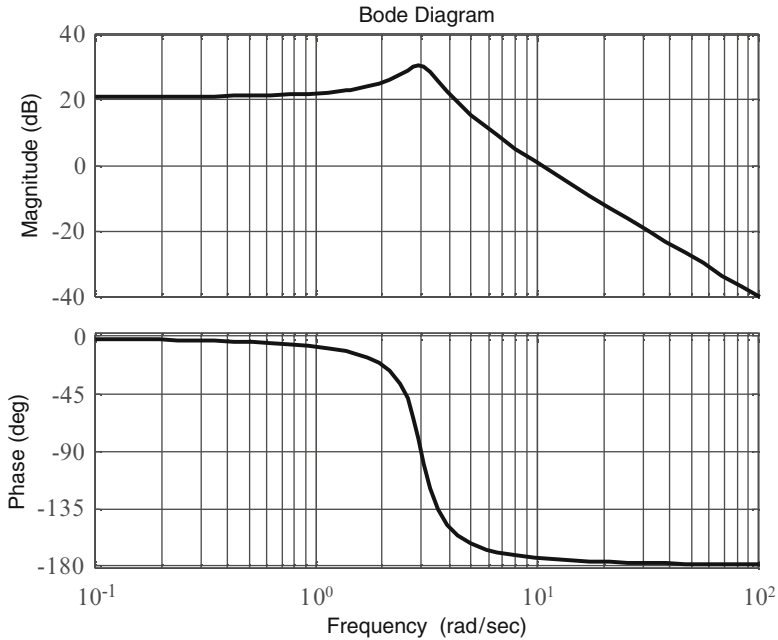
**3.3.** A plant has the following transfer function:

$$G(s) = \frac{100}{s - 4}.$$

- (a) Derive an expression for the plant's step response. Is the plant stable?
- (b) When the desired output is a unit step function, design a closed-loop controller for the plant such that the output has a settling-time of 1 second, damping ratio 0.7, and a zero steady-state error.
- (c) Derive the expression for the closed-loop step response for the system designed in Part (b).
- (d) Derive the expression for the closed-loop frequency response for the system designed in Part (b).
- (e) Find a state-space representation for the plant.

**3.4.** Design a feedback compensator for the roll dynamics of a fighter aircraft described by the transfer function,  $G(s) = 1,000/[s(s + 5)]$ , such that closed-loop step response settles down in 1 second with a zero steady-state error, and a damping ratio of 0.7.

**3.5.** For the aircraft longitudinal dynamics transfer function given in Example 3.9, find the approximate transfer function by neglecting the higher-frequency mode. Compare the frequency response and step response of the approximate transfer function with those given in the example.



**Fig. 3.18** Second-order, open-loop frequency response for the control system of Exercise 3.7

**3.6.** A closed-loop control system of Fig. 3.8 without any process or measurement noise has the following plant's transfer function:

$$G(s) = \frac{2s - 1}{s^2 + 2s - 3}.$$

- Can a compensator with transfer function,  $H(s) = (s - 1)/(s + 1)$ , stabilize this plant?
- If the compensator of Part (a) is used, what are the locations of the closed-loop poles?
- Can a proportional-integral-derivative (PID) compensator transfer function,  $H(s)$ , be selected such that the closed-loop step response settles down with a zero steady-state error? Why?

**3.7.** A closed-loop control system of Fig. 3.8 without any process or measurement noise has a second-order, open-loop frequency response plotted in Fig. 3.18. Determine the following:

- The open-loop transfer function.
- The gain and phase margins of the closed-loop system.

- (c) The steady-state error of the closed-loop system to a unit step desired output,  $Y_d(s)$ .
- (d) What modification do you propose in the control system in order to make the steady-state error in Part (c) zero?

**3.8.** Consider the control system shown in Fig. 3.8, where the plant's transfer function is  $G(s) = 1/(s^2 + 0.3s + 0.02)$ . It is intended to use a PID compensator for controlling this plant.

- (a) What are the values of the PID compensator constants,  $k_1, k_2, k_3$ , for achieving a zero steady-state error with a closed-loop pole at  $s = -1$ , and a pair of complex conjugate poles with damping ratio 0.707, and natural frequency, 1 rad/s?
- (b) Derive the closed-loop transfer function,  $Y(s)/Y_d(s)$ , with the compensator in part (a), and compute the closed-loop system step response. What is the maximum percentage overshoot, settling-time, and steady-state error of the closed-loop step response?
- (c) Plot the locus of the closed-loop poles as the integral gain,  $k_3$ , varies from 0 to 10, with  $k_1, k_2$  remaining constant at the values calculated in part(a). What are the values of  $k_3$  for which the closed-loop system is stable?
- (d) Draw a Bode plot of the closed-loop system of part (a), and determine the gain and phase margins, as well as the respective crossover frequencies.

**3.9.** Consider the following transfer function of an aircraft's longitudinal dynamics:

$$G(s) = -\frac{s^2 + 0.6s + 0.01}{s^4 + s^3 + s^2 + 0.01s + 0.009}.$$

- (a) What are the poles of the system? Is the system stable?
- (b) Using modal decomposition, resolve the denominator polynomial of  $G(s)$  into two quadratic factors (modes), each corresponding to a pair of complex poles determined in Part (a). What are the respective natural frequencies and damping ratios of the two modes?
- (c) Determine the step response by the MATLAB-CST command *step*.
- (d) Make a Bode plot of the system using the MATLAB-CST command *bode*.

**3.10.** A system is described by the following transfer function:

$$G(s) = \frac{10(s + 1)}{s^2(s - 2)}.$$

- (a) Is the system stable?
- (b) Design a state feedback regulator for the system such that the closed-loop poles satisfy the characteristic equation,

$$s^3 + 2s^2 + 2s = -1.$$

**3.11.** Is it possible to design a tracking system for the plant given by Exercise 1.8 such that the desired state satisfies the following dynamic equation:

$$\dot{\mathbf{x}}_d = \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 0 \\ 3 & -2 & -6 \end{pmatrix} \mathbf{x}_d?$$

**3.12.** Consider an LTI system with the following state-space coefficient matrices:

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}; \quad \mathbf{B} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & -1 \end{pmatrix}.$$

- Design a full-order observer for this plant using  $\mathbf{C} = (1, 0, -1)$  with poles at  $s_1 = -2$ ,  $s_{2,3} = -2 \pm 2i$ .
- Based upon the direct measurement of any two state variables, design a reduced-order observer for this plant with pole at  $s = -2$ .
- Can a tracking system be designed for this plant to track a constant desired state of  $\mathbf{x}_d = (0, 1, 0)^T$ ? Explain.

**3.13.** Design a linear, quadratic, regulator for a spring-mass system with the following governing equation:

$$\ddot{x} + 100x = u,$$

where  $x(t)$  is the displacement from equilibrium in meters, and  $u(t)$  is the applied input acceleration in  $\text{m/s}^2$ . The optimal regulator must bring the displacement to zero with a settling-time of one second in the presence of an initial disturbance of  $\dot{x}(0) = 1 \text{ m/s}$ .

**3.14.** For an aircraft with the following lateral dynamics:

$$\mathbf{A} = \begin{pmatrix} -15 & 0 & -15 & 0 & 0 \\ 0 & -0.8 & 10 & 0 & 0 \\ 0 & -1 & -0.8 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}; \quad \mathbf{B} = \begin{pmatrix} 25 & 3 \\ 0 & -3.5 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Design a bi-input linear, quadratic regulator with  $\mathbf{Q} = \mathbf{I}_{5 \times 5}$  and  $\mathbf{Q} = \mathbf{I}_{2 \times 2}$ . Calculate and plot the initial response of the regulated system for the initial condition,  $\mathbf{x}(0) = (0.5, 0, 0, 0, 0)^T$ . What are the settling-time and the maximum overshoot of the closed-loop initial response? What are the largest input magnitudes required for the closed-loop initial response?

**3.15.** Using any two state variables as the outputs, design a Kalman filter for the aircraft lateral dynamics of Exercise 3.14 with  $\mathbf{F} = \mathbf{I}$ ,  $\mathbf{W} = \mathbf{C}\mathbf{C}^T$ ,  $\Psi = 0$ , and  $\bar{\mathbf{P}} = 0.001\mathbf{I}$ . What are the poles of the Kalman filter?

Automatic Control of Atmospheric and Space Flight  
Vehicles

Design and Analysis with MATLAB® and Simulink®

Tewari, A.

2011, XIV, 374 p. 163 illus., Hardcover

ISBN: 978-0-8176-4863-3

A product of Birkhäuser Basel