CHAPTER 15
CONTROLLER DESIGN

Thomas Peter Neal
Consultant
Lake View, NY

1 INTRODUCTION **620**

2 FUNDAMENTALS OF CLOSED-LOOP PERFORMANCE **621**
2.1 Accuracy and Loop Gain 621
2.2 Dynamic Response and Stability 625

3 FREQUENCY COMPENSATION TO IMPROVE OVERALL PERFORMANCE **632**
3.1 Well-Damped Systems 633
3.2 Poorly Damped Systems 637
3.3 Higher Order Effects 641

4 INNER FEEDBACK LOOPS **645**
4.1 Derivatives of the Controlled Variable 646
4.2 Alternative Inner Loop Variables 649
4.3 Nonelectronic Inner Loops 650

5 PREFILTERS AND FEEDFORWARD **651**
5.1 Lag Prefilters 651
5.2 Lead Prefilters 652
5.3 Feedforward 654

6 PID CONTROLLERS **656**
6.1 Equivalence to Frequency Compensation 656
6.2 Systems Having No Inherent Integration 659
6.3 Systems Having an Inherent Integration 660

7 EFFECTS OF NONLINEARITIES **660**
7.1 Simple Nonlinearities 663
7.2 Complex Nonlinearities 668
7.3 Computer Simulation 668

8 CONTROLLER IMPLEMENTATION **669**
8.1 Analog Controllers 669
8.2 Hard-Wired Digital Controllers 672
8.3 Computer-Based Digital Controllers 673

REFERENCES **677**

1 INTRODUCTION

The purpose of this chapter is to provide a basis for the specification and functional design of electronic servocontrollers. No attempt is made to treat the subject of electronic circuit design. Instead, the goal is to aid the engineer in selecting and applying a suitable off-the-shelf controller or in specifying the controller requirements to a circuit designer. The emphasis is on position, velocity, or force control of mechanical loads, although many of the techniques are applicable to controller design in general. Specialized subjects such as multiaxis control and adaptive control are beyond the scope of this chapter.

As a starting point, it is presumed that a servoactuator has been selected and mounted, together with a suitable power supply, drive amplifier, and mechanical drive mechanism. In addition, the primary feedback transducer has been chosen, and a simple loop closure has been analyzed to determine whether the specified closed-loop performance can be obtained.

Reprinted from *Instrumentation and Control,* Wiley, New York, 1990, by permission of the publisher.

The process of accomplishing these tasks is treated in Chapter 14. If a simple loop closure provides adequate performance, the controller design problem primarily consists of making some basic decisions concerning electronic implementation.

In many applications, a simple loop closure is inadequate, and more elaborate controller functions are required. These latter cases are the primary subject of this chapter. The thrust of the discussion is synthesis of the controller function, rather than analysis of an existing design. For this reason, continuous-time frequency-domain techniques will be used extensively, all starting from block diagrams and transfer functions based on the Laplace operator.[1–3] These techniques, many of which are graphical, are particularly useful in the early design stages. Since most people intuitively relate to time-domain responses, relationships between the frequency-domain results and time histories will be discussed as appropriate.

If the control system is to be implemented in digital hardware or software, it is certainly possible to handle the entire design task using the mathematics of sampled-data systems.[4–7] This approach has not been used here because the so-called classical techniques based on the Laplace transform are more illustrative, and design trade-offs are easier to evaluate. For the control system applications being considered here, it is generally necessary to keep the resolution high and the sampling interval small. In this case, controller characteristics described as transfer functions in Laplace form can be accurately transformed into various mathematical forms appropriate for digital implementation, for example the *Z*-transform.

2 FUNDAMENTALS OF CLOSED-LOOP PERFORMANCE

To properly design a servocontroller, it is necessary to maintain a clear picture of the desired end result, namely, the achievement of some predetermined performance goals. These performance specifications should be established early in the design. As a minimum, they should define the desired static and dynamic accuracy, bandwidth (response time), and stability. The following sections offer a brief review of these important factors and how they relate to basic loop parameters.

2.1 Accuracy and Loop Gain

The most basic requirement of a servomechanism is probably static accuracy; that is, the controlled variable must accurately hold the command set point. Referring to the generalized block diagram of Fig. 1, several sources of inaccuracy can be described. An external disturbance can cause the load to move without any change in the command signal. The load will continue to move until the resulting error signal causes the actuator to balance the

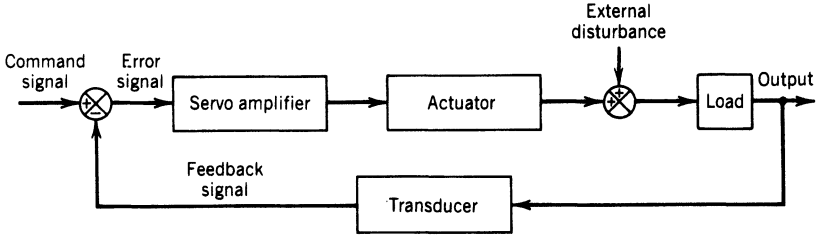


Figure 1 Generalized servomechanism.

disturbance. Anomalies in the actuator and load must also be offset by a finite error signal. Examples are temperature-induced null shifts, hysteresis, threshold, friction, and lost motion. The magnitude of these error signals is minimized if the amplifier gain is high.

Ideally, the amplifier gain would be set high enough that the accuracy of the servo becomes dependent only upon the accuracy of the transducer itself. In practice, however, the amplifier gain is limited by stability considerations. Therefore, it is desirable to provide high gains at low frequencies for accuracy and low gains at high frequencies to minimize stability problems. Since the rate of amplitude roll-off with frequency is directly related to phase lag, excessive roll-off can create more stability problems than it solves. A good compromise is to make the entire forward path look like an integrator over the frequency range of interest (a type I system). This technique is very commonly used to give nearly infinite static gain and a linear gain roll-off with frequency, at the cost of 90° phase lag.

It is important to note that some servoloops contain an inherent integrator, which complicates the accuracy-versus-stability problem. For example, many actuators are inherently rate devices when operated open loop, so that a steady input results in a proportional velocity output. A velocity servo using such an actuator will inherently have a proportional forward loop, and an integrating servoamplifier can be used (Fig. 2). However, the corresponding position servo will inherently have an integration in the forward loop, as shown in Fig. 3. In this latter case, the use of an integrating servoamplifier can cause severe stability problems. From these figures, transfer functions can be written for the closed-loop responses to command and disturbance inputs. Note that the dynamic response characteristics of all elements have been neglected, and it is assumed that the load includes no spring to ground:

$$\frac{V}{e_c} = \frac{1}{K_v} \frac{1}{s/K_{vv} + 1} \tag{1}$$

$$\frac{V}{F_d} = K_4 \frac{s}{s + K_{vv}} \tag{2}$$

$$\frac{X}{e_c} = \frac{1}{K_x} \frac{1}{s/K_{vx} + 1} \tag{3}$$

$$\frac{X}{F_d} = \frac{K_4}{K_{vx}} \frac{1}{s/K_{vx} + 1} \tag{4}$$

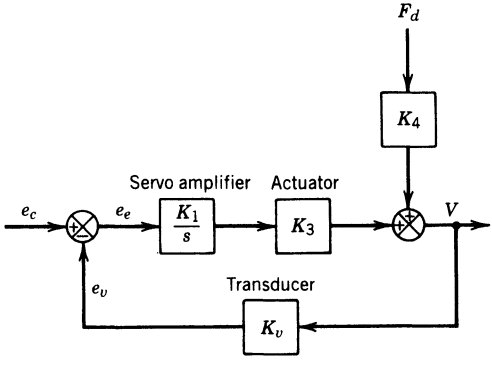


Figure 2 Simplified velocity servo.

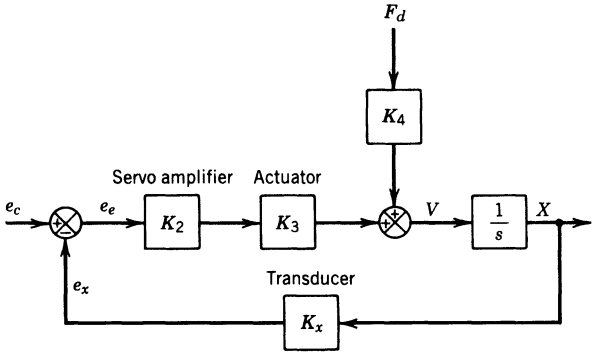


Figure 3 Simplified position servo.

where e_c = command signal V
e_e = error signal, V
e_v = velocity feedback signal, V
e_x = position feedback signal, V
F_d = disturbance force applied to the load, N
K_1 = integrating servoamplifier gain, (V/s)/V
K_2 = proportional servoamplifier gain, V/V
K_3 = actuator gain, (mm/s)/V
K_4 = actuator velocity droop due to force disturbance, (mm/s)/N
K_v = velocity transducer gain, volts/(mm/s)
K_x = position transducer gain, volts/mm
K_{vv} = open-loop gain of velocity servo = $K_1K_3K_v$, s^{-1}
K_{vx} = open-loop gain of position servo = $K_2K_3K_x$, s^{-1}
X = load position, mm
$V = \dot{X}$, mm/s

As shown in Fig. 4, the velocity and position responses to commands are both characterized by a first-order lag having a break frequency equal to the open-loop gain. However, the responses to disturbance forces are quite different in the two cases (Fig. 5). When the disturbance is downstream of the integrator (velocity servo), the servo error is K_4F_d at high frequencies but rolls off at frequencies (in radians per second) below K_{vv} and is zero statically. When the disturbance is upstream of the integrator (position servo), there is a static error inversely proportional to the open-loop gain, which rolls off at frequencies above K_{vx}. Note that Eqs. (1)–(4) remain reasonably valid when the dynamic response characteristics of the various open-loop elements are considered, for those cases in which K_{vv} (or K_{vx}) is well below the lowest break frequencies of those elements. At higher loop gains, the closed-loop dynamics can change considerably, as shown in the following discussion.

The conclusions regarding the effects of disturbance forces on servo accuracy can be generalized to any forward-loop offset or uncertainty. Referring again to Fig. 2, it can be seen that the integrating amplifier will compensate for any forward-loop offset downstream of the integrator, so that the static errors are zero. From Fig. 3, it is apparent that static errors due to offsets upstream of the integration can be quantified as

$$\frac{X_e}{V_0} = \frac{1}{K_{vx}} \tag{5}$$

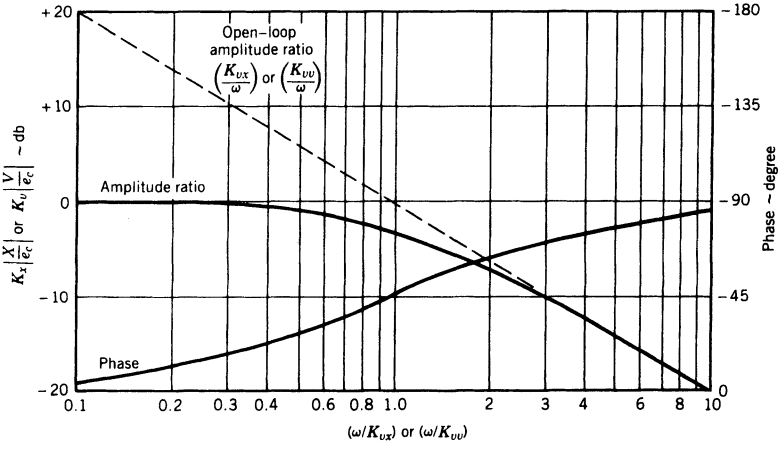


Figure 4 Simplified response to commands.

where X_e = output error = $X_c - X$, mm
V_0 = forward-loop offset, converted to an equivalent open-loop offset in volts, mm/s
X_c = position command = e_c/K_x, mm

Even when no forward-loop offsets or disturbances are present, servos exhibit following errors, sometimes called tracking errors. In servos having forward-loop integrations, these quasi-static errors result whenever the command signal changes at a constant rate, as in a position-tracking servo. For the position servo of Fig. 3, the following error is

$$\frac{e_e}{\dot{e}_c} = \frac{X_e}{\dot{X}_c} = \frac{1}{K_{vx}} \tag{6}$$

For the velocity servo of Fig. 2, the following error is

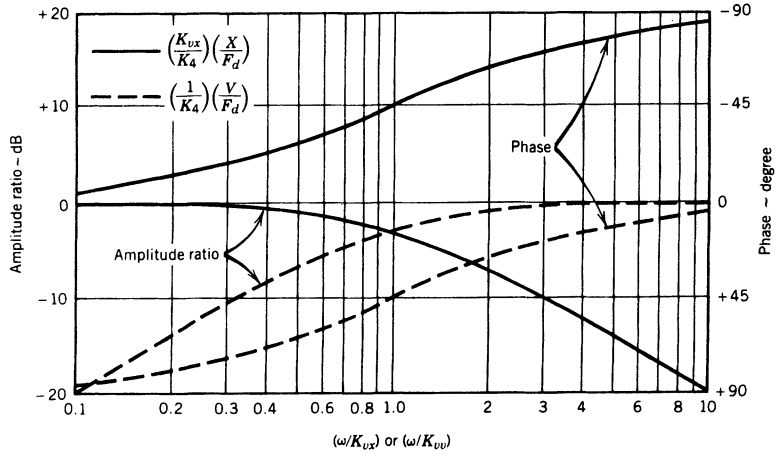


Figure 5 Simplified response to disturbance.

$$\frac{e_e}{\dot{e}_c} = \frac{V_e}{\dot{V}_c} = \frac{1}{K_{vv}} \tag{7}$$

where V_e = output error = $V_c - V$, mm/s
V_c = velocity command = e_c/K_v, mm/s

Note that the following error for the position servo is the position error resulting from a steady rate of change of position command. For the velocity servo, the following error is the velocity error resulting from a steady rate of change of velocity command.

The servo errors discussed thus far have been those that can be minimized by a tight servoloop (high loop gain). To these must be added errors in the transducer mechanism. Even if infinite loop gain were achievable, the servo can be no more accurate than the transducer itself. The most important types of transducer inaccuracies are repeatability, resolution, and linearity. Errors due to transducer location and mounting geometry must also be taken into account.

Many of the foregoing concepts can be applied to servos in general. For example, a force or pressure servo working against a spring load is similar to a position servo in the sense that output force is proportional to actuator position. Also, temperature control servos tend to behave like position servos since the controlling device tends to provide heat flow proportional to temperature error, and thermal loads tend to produce temperature rate of change proportional to heat flow.

2.2 Dynamic Response and Stability

As discussed in Section 2.1, open-loop gain has a strong influence on servo accuracy. High loop gains also provide fast dynamic response in most cases. However, stability considerations will limit the maximum useful loop gain. The dynamic response and stability of a servo are determined by the dynamic characteristics of the various loop components. In many situations, the forward-loop dynamics are dominated by a relatively small number of low-frequency lag elements, and the transducer dynamics are negligible. In these cases, it is often possible to obtain an adequate estimate of servo performance by approximating the combined forward-loop characteristics with an integrator plus a first-order or second-order lag. The adequacy of this approximation can be determined by the match of the frequency response gain and phase for the frequency range in which the phase lag is less than 180°. Using these two rather basic dynamic forms, the relationships among loop gain, stability, and dynamic response are easily seen.

A block diagram using the basic dynamic forms is shown in Fig. 6,

where U = generalized controlled variable
U_c = generalized command input
D = generalized disturbance input
G_d = open-loop response of U to D

The first-order lag is typical of simple temperature control systems and dc servos having short electrical time constants. The second-order lag is often representative of electrohydraulic servos and dc servos having long electrical time constants. The closed-loop responses to command inputs are

$$\frac{U_1}{U_c} = \frac{1}{(\tau_1/K_{u1})s^2 + (1/K_{u1})s + 1} \tag{8}$$

and

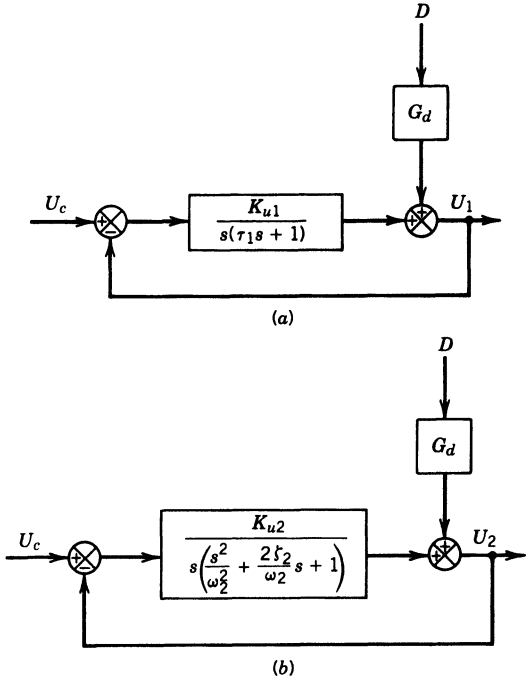


Figure 6 Basic dynamic configurations: (*a*) first-order lag; (*b*) second-order lag.

$$\frac{U_2}{U_c} = \frac{1}{s^3/K_{u2}\omega_2^2 + (2\zeta_2/K_{u2}\omega_2)s^2 + (1/K_{u2})s + 1} \tag{9}$$

Representative root loci are shown in Fig. 7 for both forms. In Fig. 7*a*, the lag combines with the integrator to produce second-order closed-loop poles. The closed-loop natural frequency increases with 1oop gain, while the damping ratio decreases. In the case of Fig. 7*b*,

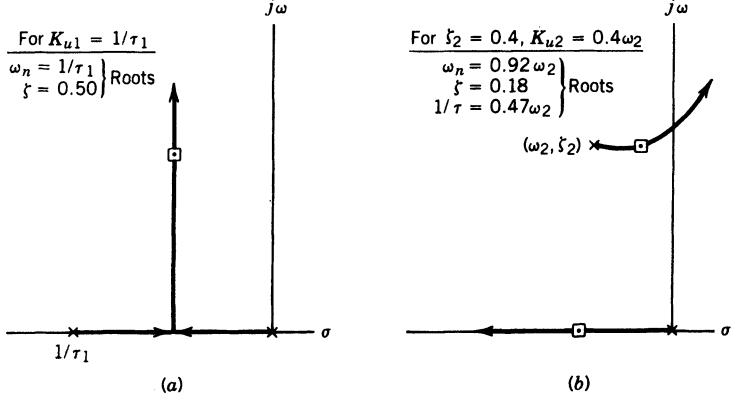


Figure 7 Root loci for basic dynamic configurations: (*a*) first-order lag; (*b*) second-order lag.

the closed-loop transfer function consists of a first-order and a second-order lag. The break frequency of the first-order lag increases with 1oop gain, while the second-order damping ratio rapidly decreases. In both loop closures, there are clearly trade-offs between closed-loop bandwidth and stability.

There are numerous methods for quantifying the relationships between bandwidth and stability. Closed-loop frequency responses to command inputs are shown for both basic forms in Figs. 8 and 9, while Figs. 10 and 11 present the corresponding step responses. Useful numerical measures of stability are phase margin, gain margin, and damping ratio of the closed-loop complex pair. These are given in Figs. 12 and 13. Note that the gain margin for Fig. 12 is infinite.

Referring to Fig. 6, closed-loop responses to disturbance inputs can be written as

$$\frac{U_1}{D} = \frac{G_d}{K_{u1}} \frac{s(\tau_1 s + 1)}{(\tau_1/K_{u1})s^2 + (1/K_{u1})s + 1} \tag{10}$$

and

$$\frac{U_2}{D} = \frac{G_d}{K_{u2}} \frac{s[s^2/\omega_2^2 + (2\zeta_2/\omega_2)s + 1]}{(s^3/K_{u2}\omega_2^2) + (2\zeta_2/K_{u2}\omega_2)s^2 + (1/K_{u2})s + 1} \tag{11}$$

To determine the final dynamic form of these responses, it is necessary to have a transfer function for G_d. This transfer function can be obtained from the physical model of the system, by deriving the response of the controlled variable to the disturbance input with the controller output equal to zero. As an example, consider a dc motor driving an inertial load and having a short electrical time constant. For an integrating velocity loop, the disturbance transfer function has the form

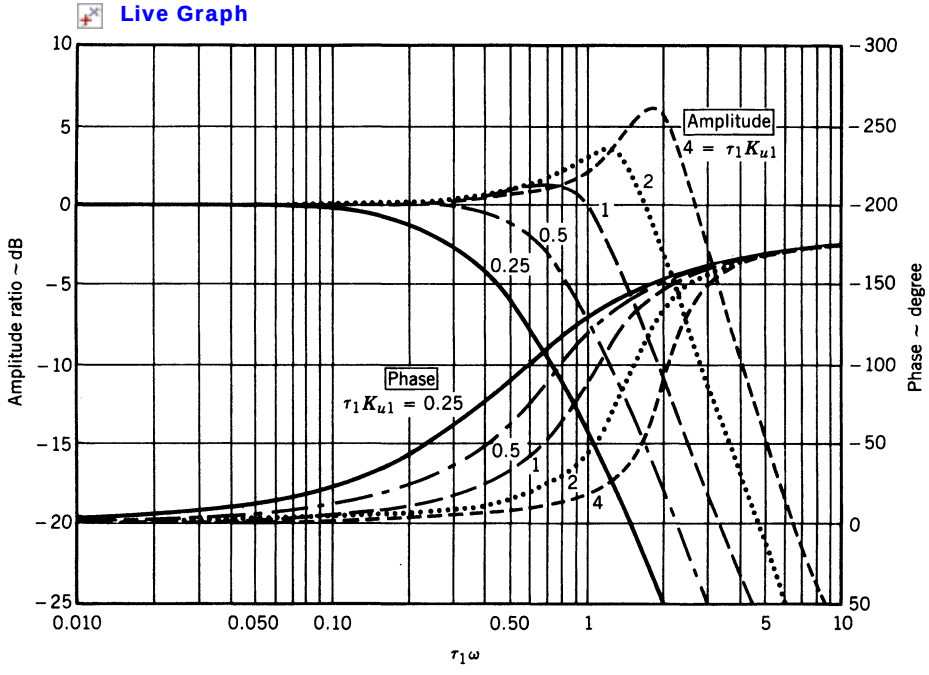
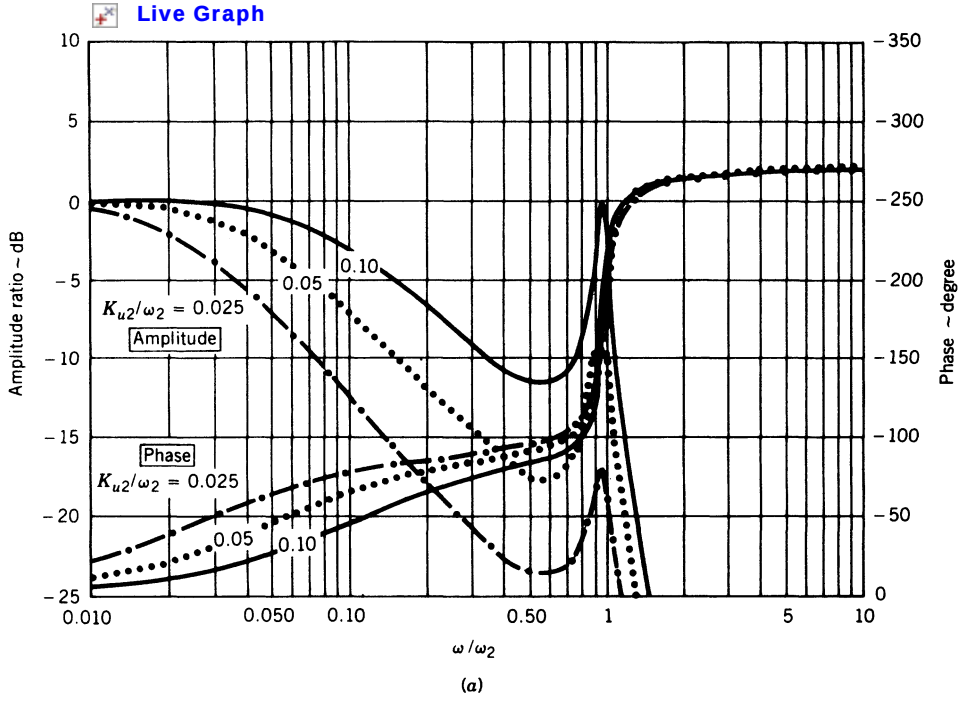


Figure 8 Closed-loop frequency responses for U_1/U_c.



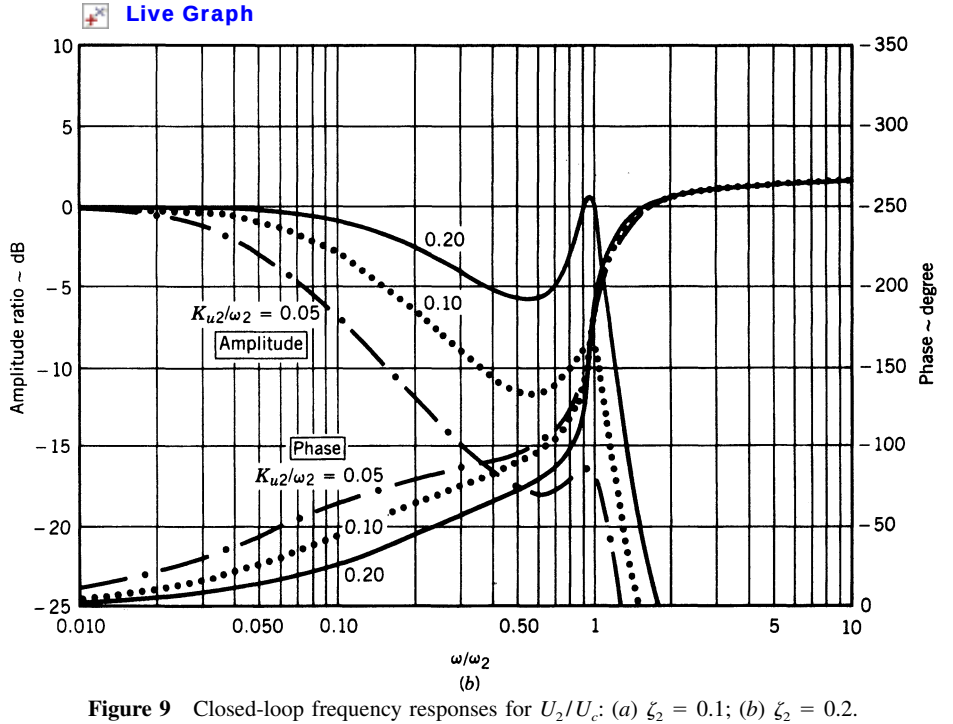
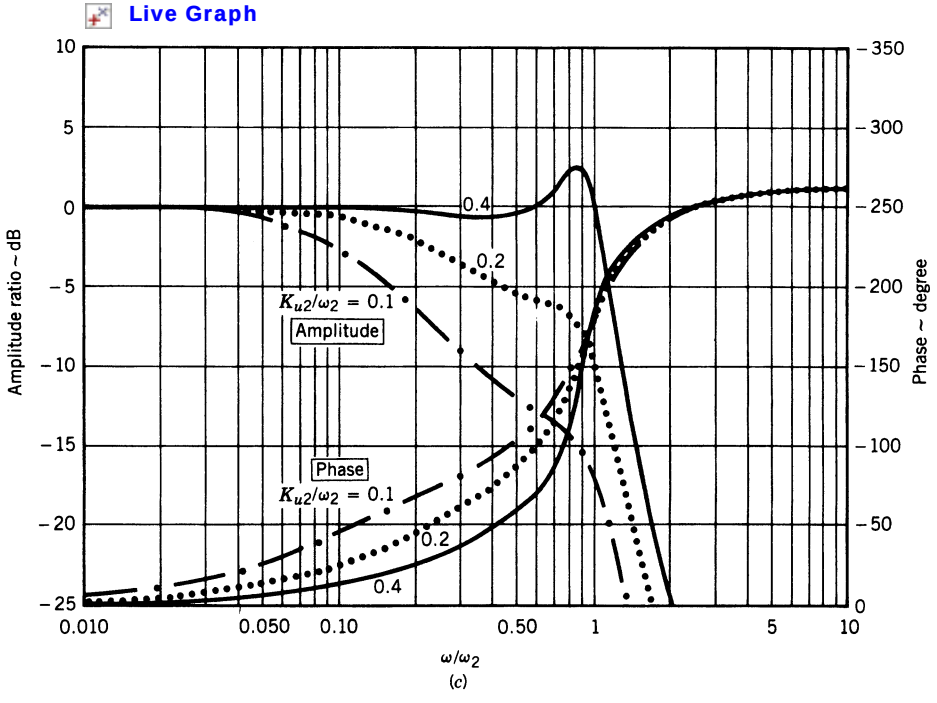


Figure 9 Closed-loop frequency responses for U_2/U_c: (*a*) $\zeta_2 = 0.1$; (*b*) $\zeta_2 = 0.2$.



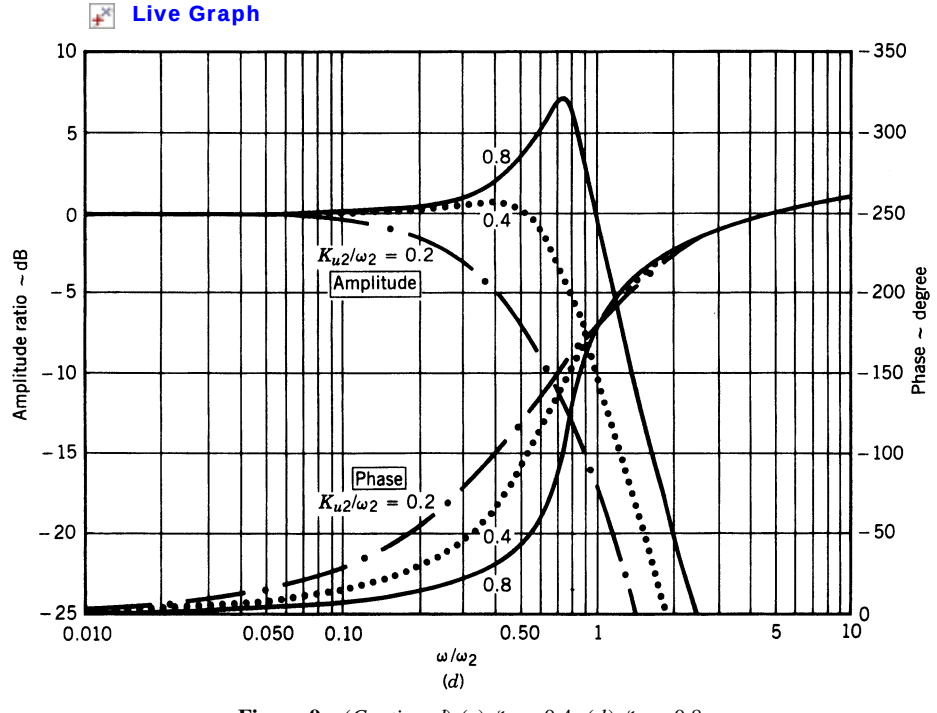


Figure 9 (*Continued*) (*c*) $\zeta_2 = 0.4$; (*d*) $\zeta_2 = 0.8$.

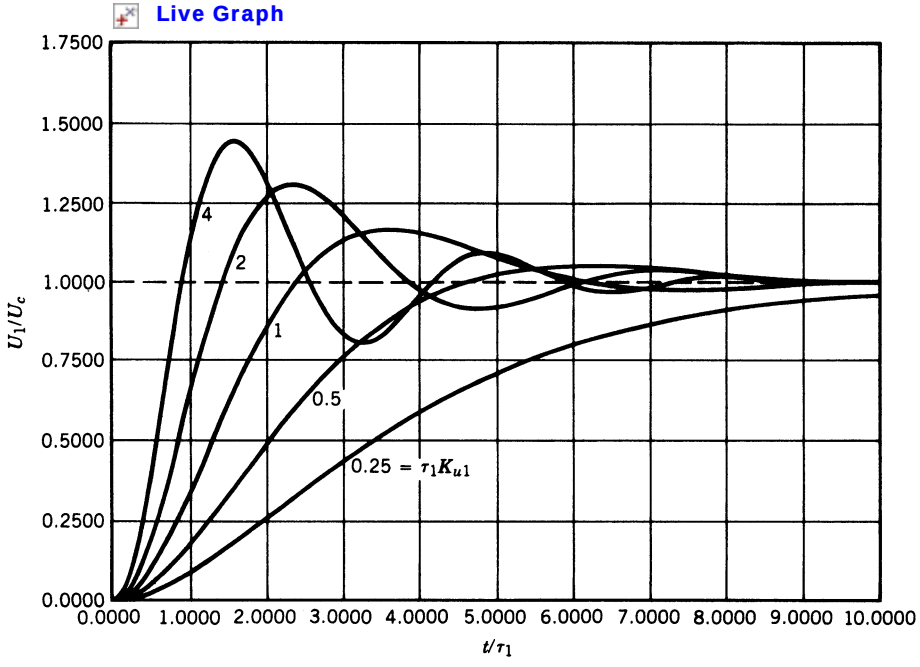


Figure 10 Closed-loop step responses for U_1/U_c.

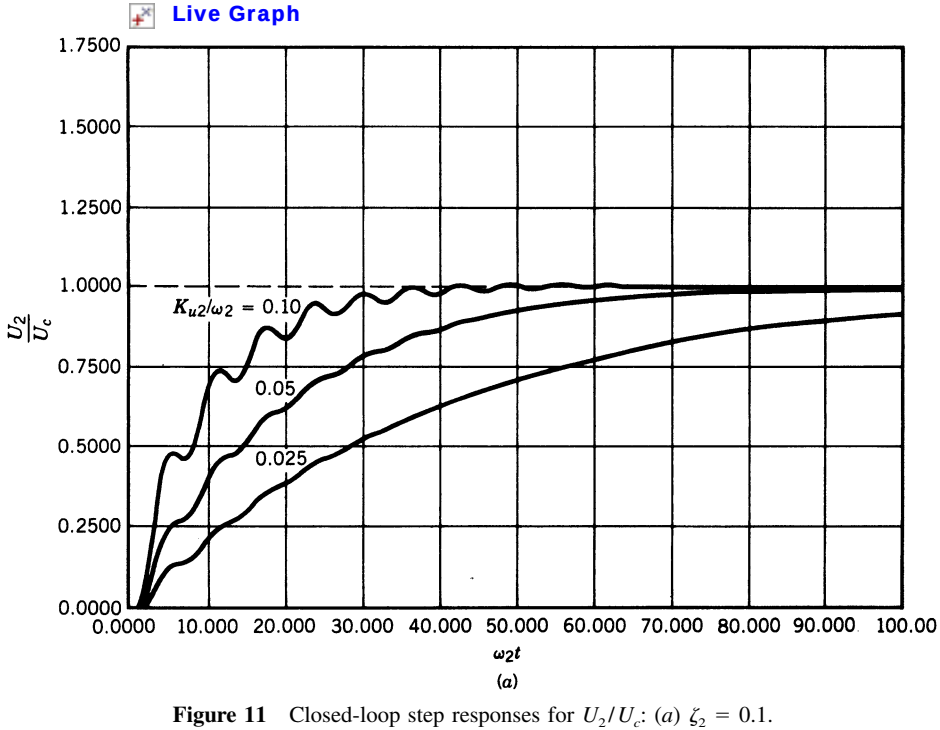
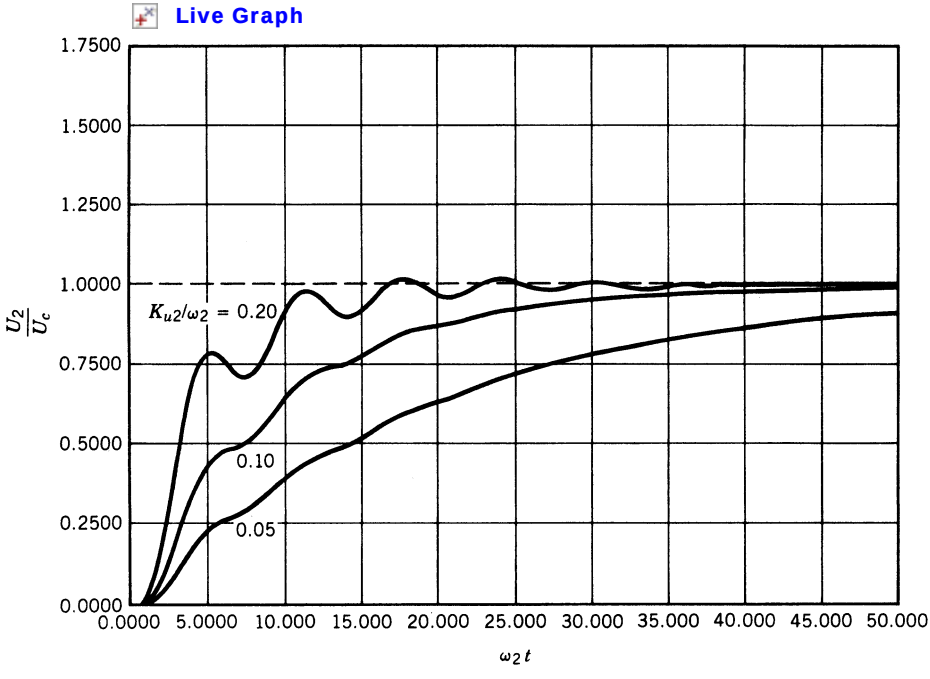
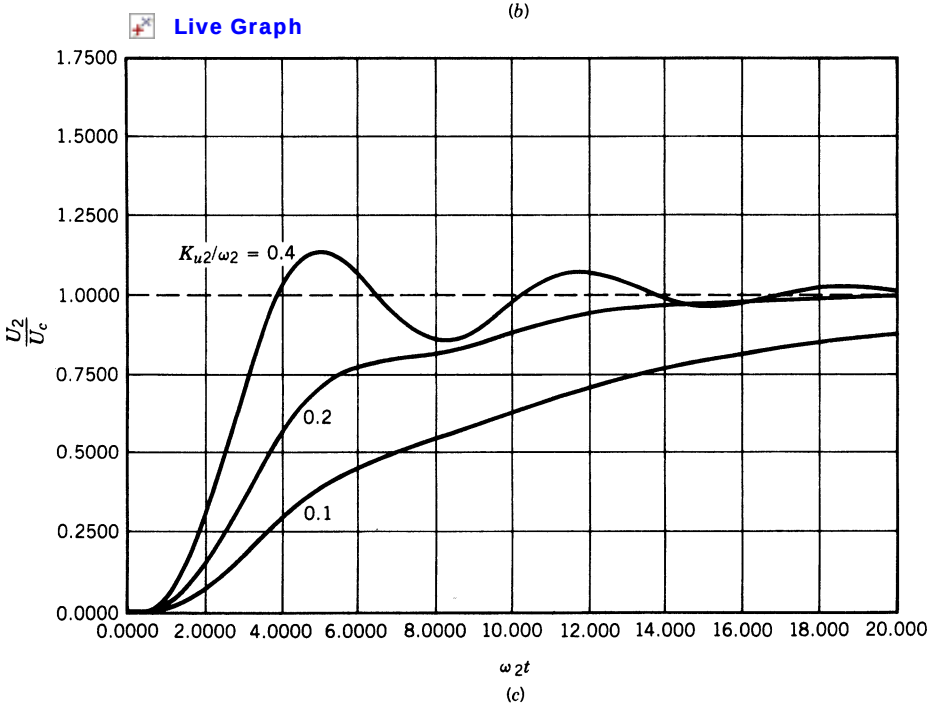


Figure 11 Closed-loop step responses for U_2/U_c: (*a*) $\zeta_2 = 0.1$.



(*b*)



(*c*)

Figure 11 (*Continued*) (*b*) $\zeta_2 = 0.2$; (*c*) $\zeta_2 = 0.4$.

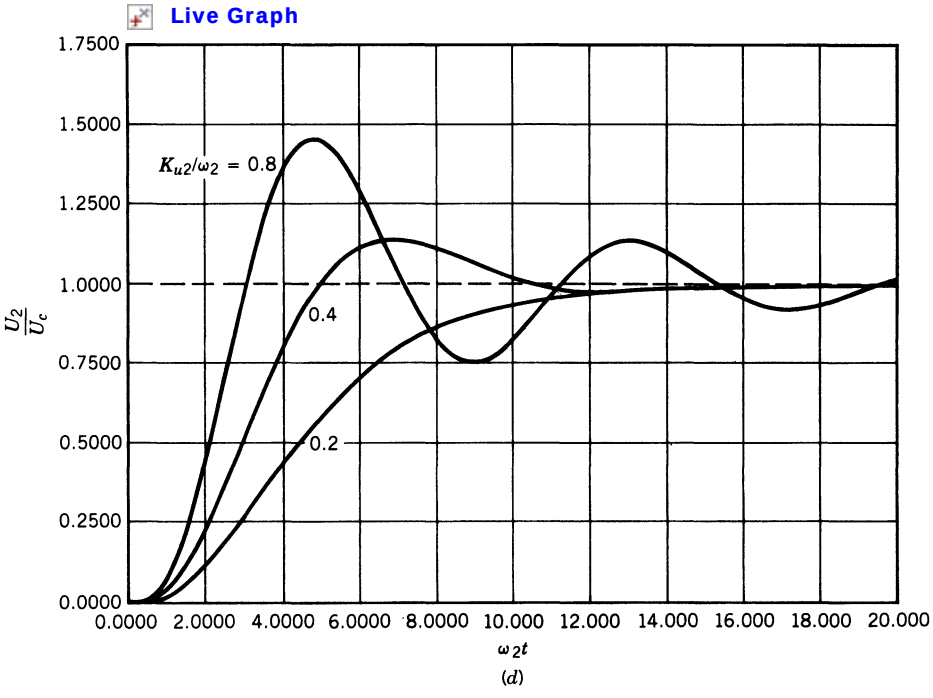


Figure 11 (*Continued*) (*d*) $\zeta_2 = 0.8$.

$$G_d = \frac{K_d}{\tau_1 s + 1} \tag{12}$$

and

$$\frac{U_1}{D} = \frac{K_d}{K_{u1}} \frac{s}{(\tau_1/K_{u1})s^2 + (1/K_{u1})s + 1} \tag{13}$$

For a position loop

$$G_d = \frac{K_d}{s(\tau_1 s + 1)} \tag{14}$$

and

$$\frac{U_1}{D} = \frac{K_d}{K_{u1}} \frac{1}{(\tau_1/K_{u1})s^2 + (1/K_{u1})s + 1} \tag{15}$$

It is instructive to compare Eqs. (13) and (15) with Eqs. (2) and (4), respectively.

3 FREQUENCY COMPENSATION TO IMPROVE OVERALL PERFORMANCE

In Section 2, it is clear that open-loop dynamic characteristics impose profound limitations upon closed-loop performance. However, it is often possible to extend these limitations by

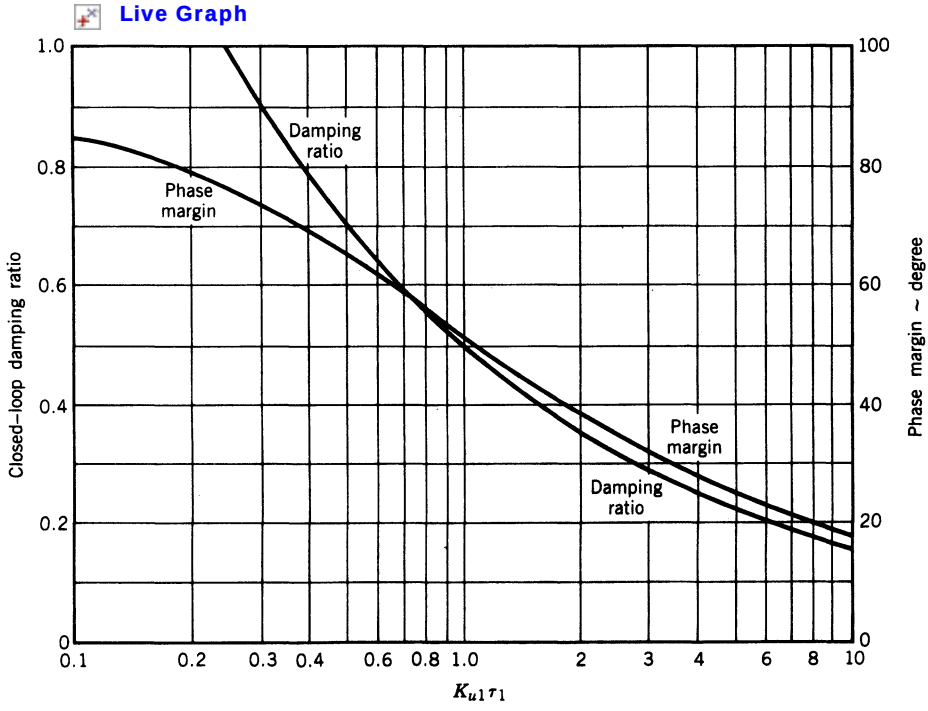


Figure 12 Stability parameters for U_1/U_c.

modifying the inherent open-loop dynamics with frequency compensation (shaping). There are many techniques for designing compensators. The best technique to use is a function of the particular open-loop dynamics under consideration, as well as closed-loop performance goals. The following sections describe some techniques that are useful in various commonly encountered situations.

3.1 Well-Damped Systems

As mentioned in Section 2.1, an ideal form for the combined forward-loop transfer function is an integrator. This ensures very high gains at low frequencies, a linear gain roll-off with frequency, and only 90° of phase lag. For systems in which the dominant open-loop poles are reasonably well damped, loop gain is usually limited by phase lag. This is clearly illustrated by Figs. 12 and 13, in which phase margins deteriorate faster than gain margins (except when ζ_2 is low).

An obvious way to improve phase margins is to make the open-loop transfer function look like an integrator out to higher frequencies. This can be accomplished by using a lead compensator, whose zeros are identical to the dominant forward-loop poles. To make the compensator physically realizable, it must have at least as many poles as zeros, but these poles can be placed at higher frequencies. The net effect of such a lead compensator is to move the break frequencies of the forward-loop poles to higher frequencies. For the example of Fig. 6*a*, the form of the compensator would be

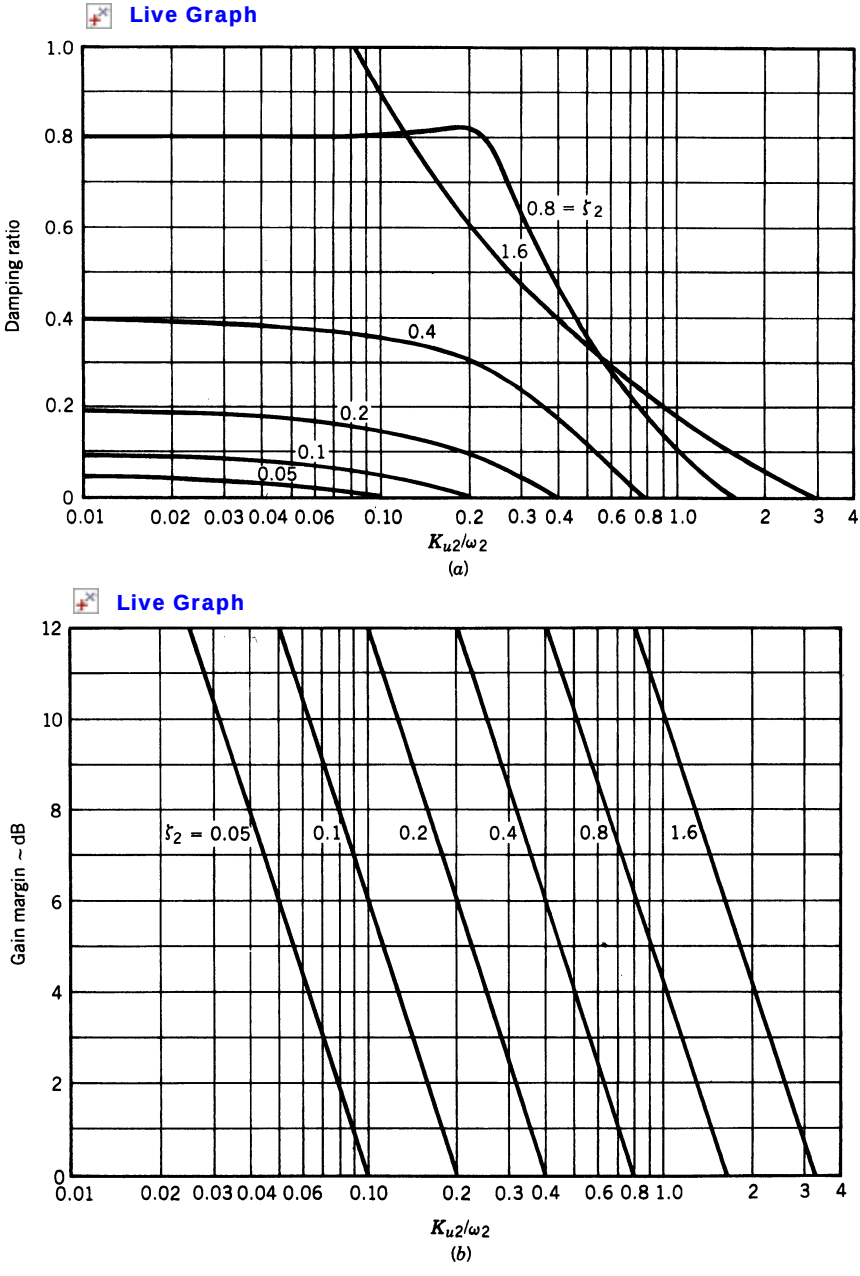


Figure 13 Stability parameters for U_2/U_c: (*a*) closed-loop damping ratio; (*b*) gain margin.

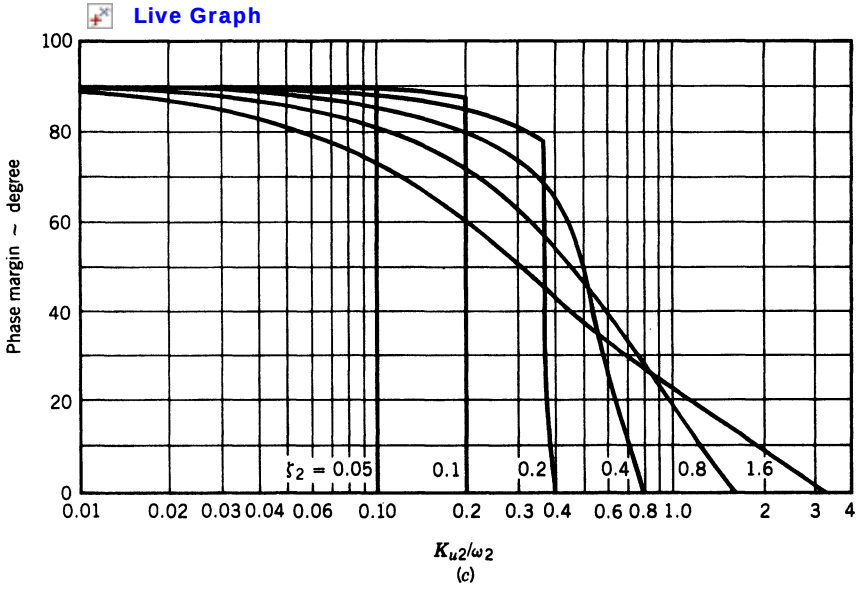


Figure 13 (*Continued*) (*c*) phase margin.

$$G_{c1} = \frac{\tau_{cz}s + 1}{\tau_{cp}s + 1} \tag{16}$$

where $\tau_{cz} = \tau_1$. For the example of Fig. 6*b*, the form would be

$$G_{c2} = \frac{s^2/\omega_{cz}^2 + (2\zeta_{cz}/\omega_{cz})s + 1}{s^2/\omega_{cp}^2 + (2\zeta_{cp}/\omega_{cp})s + 1} \tag{17}$$

where $\omega_{cz} = \omega_2$ and $\zeta_{cz} = \zeta_2$. In both cases, the closed-loop performance is now determined by the compensator poles since the original poles are canceled by the compensator zeros. This augmented performance can be quantified by using Figs. 8–13, with τ_{cp}, ω_{cp}, ζ_{cp} substituted for τ_1, ω_2, ζ_2, respectively.

There are a number of practical limitations on the use of lead compensation, primarily related to the large high-frequency gain of the compensator itself. For the examples of Eqs. (16) and (17), the high-frequency gains are τ_{cp}/τ_{cz} and $(\omega_{cp}/\omega_{cz})^2$, respectively. As a minimum, this characteristic will amplify any high-frequency electrical noise in the system. With reasonable care in the electrical design, high-frequency compensator gains of 10 or more are often practical. For a first-order compensator, this means that the forward-loop break frequencies can be boosted by a factor of 10, while the boost is only the square root of this factor for a second-order compensator. Another problem associated with the gain boost of a lead compensator is that poorly damped high-frequency modes can be excited or even destabilized. This latter effect will be further discussed in Section 3.3. The practicality of lead compensation in any given application can be best determined experimentally (additional high-frequency lags are sometimes needed).

The high-frequency noise situation is improved considerably for systems in which the integrator is electronic (e.g., the velocity servo described in Section 2.1). In this case, the integrator can replace one of the compensator poles so that Eqs. (16) and (17) are replaced by

$$G_{c3} = \frac{\tau_{z3}s + 1}{s} \tag{18}$$

and

$$G_{c4} = \frac{s^2/\omega_{z4}^2 + (2\zeta_{z4}/\omega_{z4})s + 1}{s(\tau_{p4}s + 1)} \tag{19}$$

Because of the noise-attenuating effect of the electronic integrator, $1/\tau_{p4}$ can often be set a factor of 10 greater than ω_{z4}. As previously discussed, the ratio of ω_{cp}/ω_{cz} in Eq. (17) is often limited to $\sqrt{10}$. As shown in the frequency responses of Figs. 14 and 15, the open-loop characteristics of a lead-compensated system will exhibit substantially improved phase characteristics when the integration is electronic rather than inherent. This makes it possible to greatly improve the closed-loop bandwidth of the system for a given level of stability.

In some cases, the open-loop dynamics may be dominated by a low-frequency lead term rather than a lag. When this occurs, a canceling lag compensator can often prevent instabilities due to poorly damped high-frequency modes. Lag compensation is normally well behaved and suffers none of the noise problems that limit the use of lead compensation.

Transfer functions describing the system open-loop characteristics are not always available. Sometimes the only available system description is an experimental frequency response. When this is the case, a compensator can often be designed by graphically subtracting the experimental amplitude ratio (in decibels) and phase from those of an integrator having the same low-frequency gain. This "ideal" compensator can then be approximated with an appropriate transfer function.

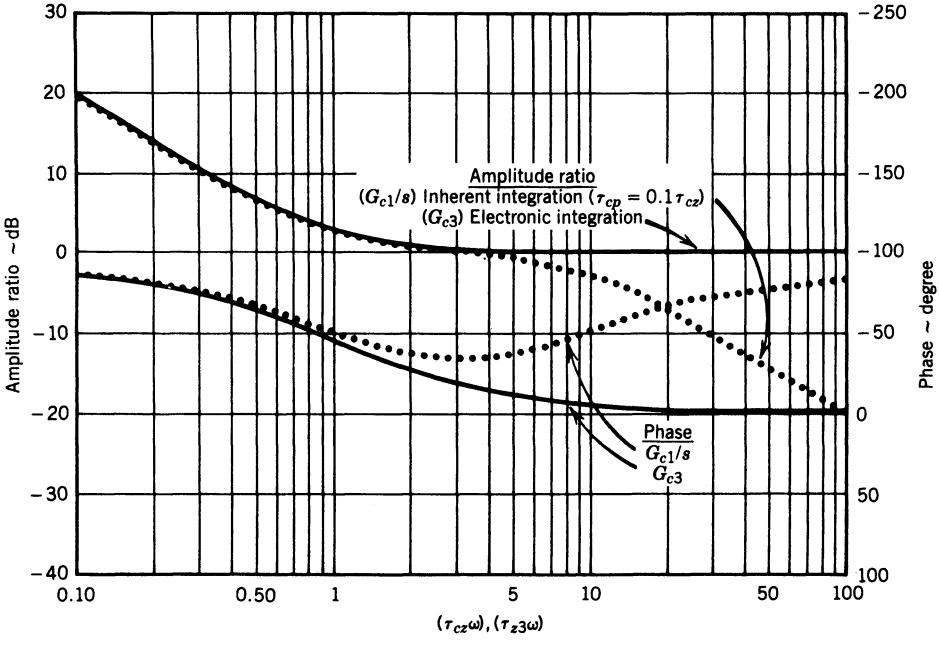


Figure 14 First-order lead compensation.

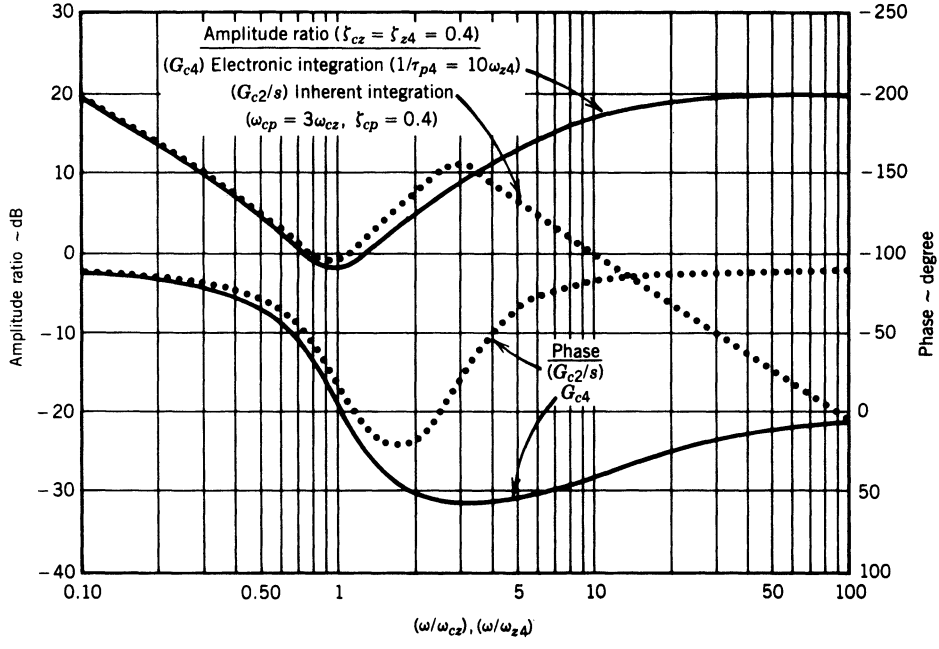


Figure 15 Second-order lead compensation.

3.2 Poorly Damped Systems

Section 3.1 discussed the benefits of lead compensation whose zeros are identical to the dominant forward-loop poles of the system. Theoretically, this technique can be used for any forward-loop transfer function. However, when the dominant forward-loop poles are second order and poorly damped, practical considerations render the technique highly risky in many cases. The basic problem is that the amplitude and phase of a poorly damped pair of poles change very rapidly with frequency in the vicinity of the resonant peak. If the compensator zeros are not precisely matched to the poles, the combined forward-loop transfer function can easily exhibit 180° of phase lag in the vicinity of substantial local peaking.

To illustrate the potential stability problem, consider the system of Fig. 6*b*, with $\zeta_2 = 0.10$. Suppose that the lead compensator of Eq. (7) is added with $\omega_{cz} = \omega_2$, $\zeta_{cz} = 0.10$, $\omega_{cp} = 3\omega_{cz}$, and $\zeta_{cp} = 0.80$. Theoretically, the forward loop is now dominated by the integrator and the compensator poles. Referring to Fig. 9*d*, it can be seen that a well-behaved response can be obtained with a loop gain $K_{u2} = 1.2\omega_{cz}$. However, suppose that ω_2 shifts to a lower value. For example, with an electrohydraulic servoactuator driving an inertial load, the "hydraulic resonance" can change 50% or more over the stroke range of the cylinder. Figure 16 shows how the closed-loop roots and the open-loop frequency response change with variations in ω_2.

Note that a reduction of ω_2 to $0.89\omega_{cz}$ will cause the closed-loop system to become unstable (0 dB at −180° phase). Even if the natural frequency of the forward-loop poles does not change at all, the poles remain poorly damped in the closed-loop transfer function. Although they are masked by the compensator zeros with regard to command inputs, they may be excited by disturbance inputs to the system. The same general comments apply to

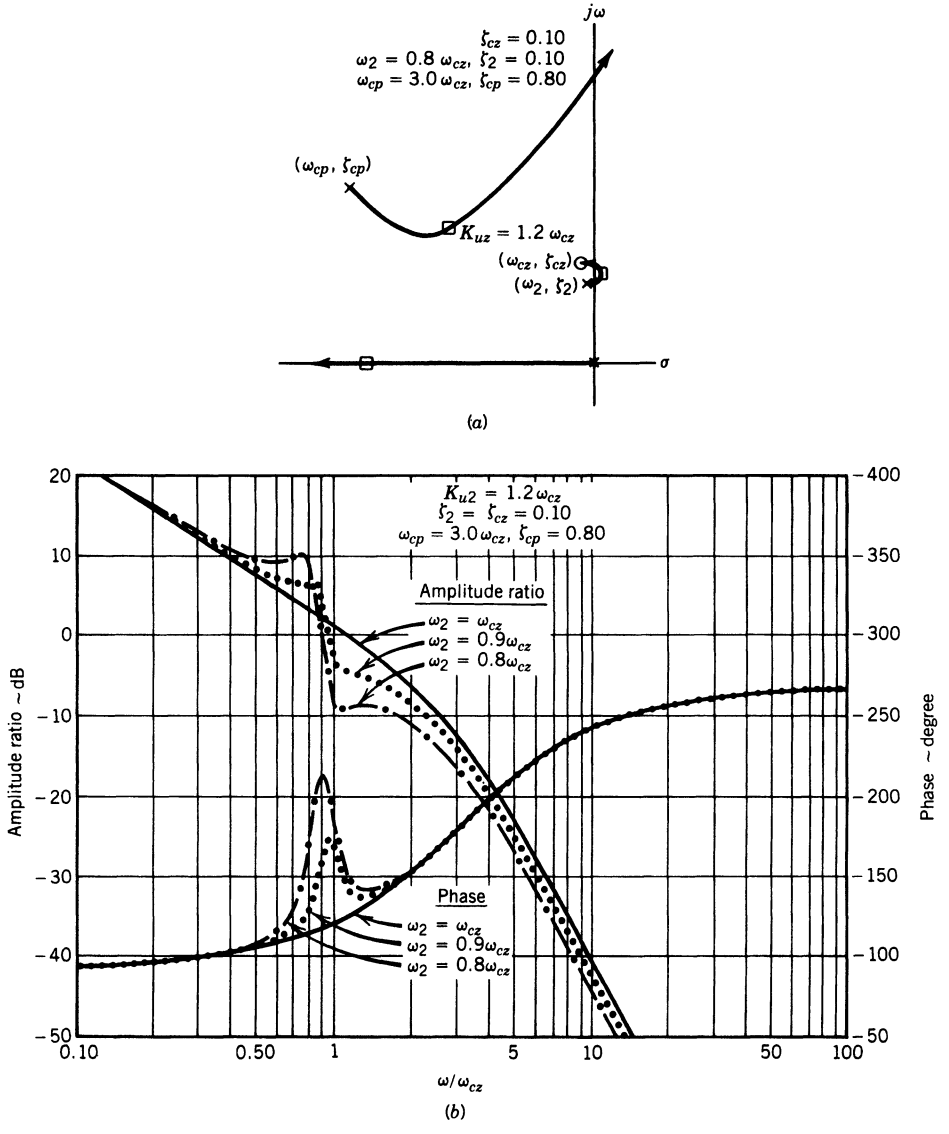


Figure 16 Effects of variations in forward-loop poles on stability of a lead-compensated loop: (*a*) root locus; (*b*) open-loop frequency response.

the use of notch filters. The notch is intended to attenuate the resonant peak, without the bandwidth boost of the lead compensator described in the present example ($\omega_{cp} = \omega_{cz}$, instead of $3\omega_{cz}$).

However, it is possible to design a compensator that will improve the damping of these poles. To accomplish this, it is useful to recall that for any system in which the number of poles exceeds the number of zeros by two or more, the sum of the real parts of all the poles is not changed when the loop is closed.[1] In this case, closing the loop will cause some roots to become more stable and others to become less stable (usually those that were poorly

damped to begin with). On the other hand, if the number of poles exceeds the number of zeros by one or less, it is possible for the loop closure to improve the stability of all the roots. Therefore, to be effective in damping a poorly damped dynamic mode, lead compensation of sufficiently high order is required. To illustrate this concept, again consider the system of Fig. 6*b*, with $\zeta_2 = 0.1$. Even if we consider ideal lead compensators having no poles, Fig. 17 makes it clear that damping cannot be improved unless the order of the lead is 2 or more.

As mentioned in Section 3.1, there are important practical constraints on the use of second-order lead compensation. If the forward-loop integrator is electronic in nature, making it practical to achieve a ratio of compensator pole-zero break frequencies on the order of 10, the improvements in damping indicated by Fig. 17*c* are indeed possible. Such a compensator is defined by Eq. (19), with $\omega_{z4} = \omega_2$, $\zeta_{z4} = 0.4$, and $1/\tau_{p4} = 10\omega_{z4}$. The resulting root locus, shown in Fig. 18*a*, indicates that open-loop gain on the order of $8\omega_2$ is possible with good stability (gain margin 6 dB and phase margin 50°). To achieve the same gain margin without compensation, the open-loop gain would be limited to $0.1\omega_2$. Closed-loop frequency responses to commands for both cases are given in Fig. 18*b*.

If the forward-loop integration is inherent rather than electronic, the ratio of compensator pole–zero break frequencies may be limited to values as low as 3 (Section 3.1). Using the compensator of Eq. (17), a good compromise is $\omega_{cz} = \omega_2$, $\zeta_{cz} = 0.4$, $\omega_{cp} = 3\omega_{cz}$, and $\zeta_{cp} = 0.4$. The resulting impact on performance is shown in Fig. 19. To maintain a 6-dB gain margin, the open-loop gain must be reduced to $1.0\omega_2$. Even with this reduced gain, the phase margin is only 30°. If the compensator is located in the forward 1oop, the closed-loop response to commands exhibits considerable peaking, as shown in Fig. 19*b*. This figure also illustrates that if the compensator is located in the feedback loop, the compensator zeros do not appear in the closed-loop response to commands, and the response exhibits less peaking with more phase lag.

A more straightforward method for improving the performance of poorly damped systems is by the use of lag compensation. A first-order low-pass filter in the forward loop will slow the degradation of closed-loop damping ratio as loop gain is increased. In addition, the lag compensator will attenuate abrupt command or disturbance inputs to the system, thereby reducing their ability to excite the poorly damped mode. Figure 20 shows the effects of a lag compensator optimized for the system of Fig. 6 with $\zeta_2 = 0.10$. The effects of an optimized lag compensator on various stability parameters are illustrated in Fig. 21 for several values of ζ_2. Comparisons with Fig. 13 show that lag compensation can provide a

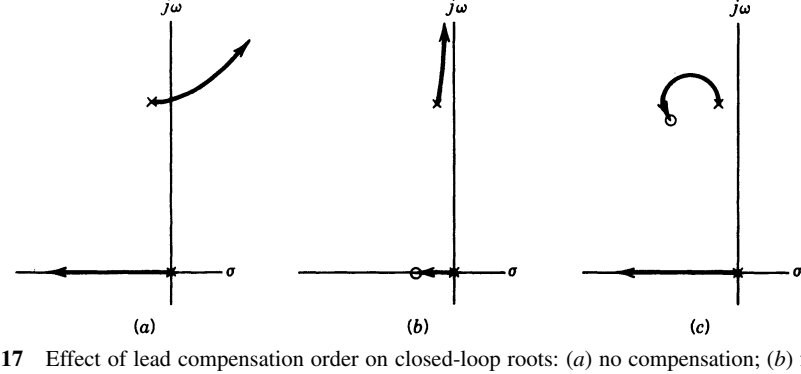


Figure 17 Effect of lead compensation order on closed-loop roots: (*a*) no compensation; (*b*) first-order lead; (*c*) second-order lead.

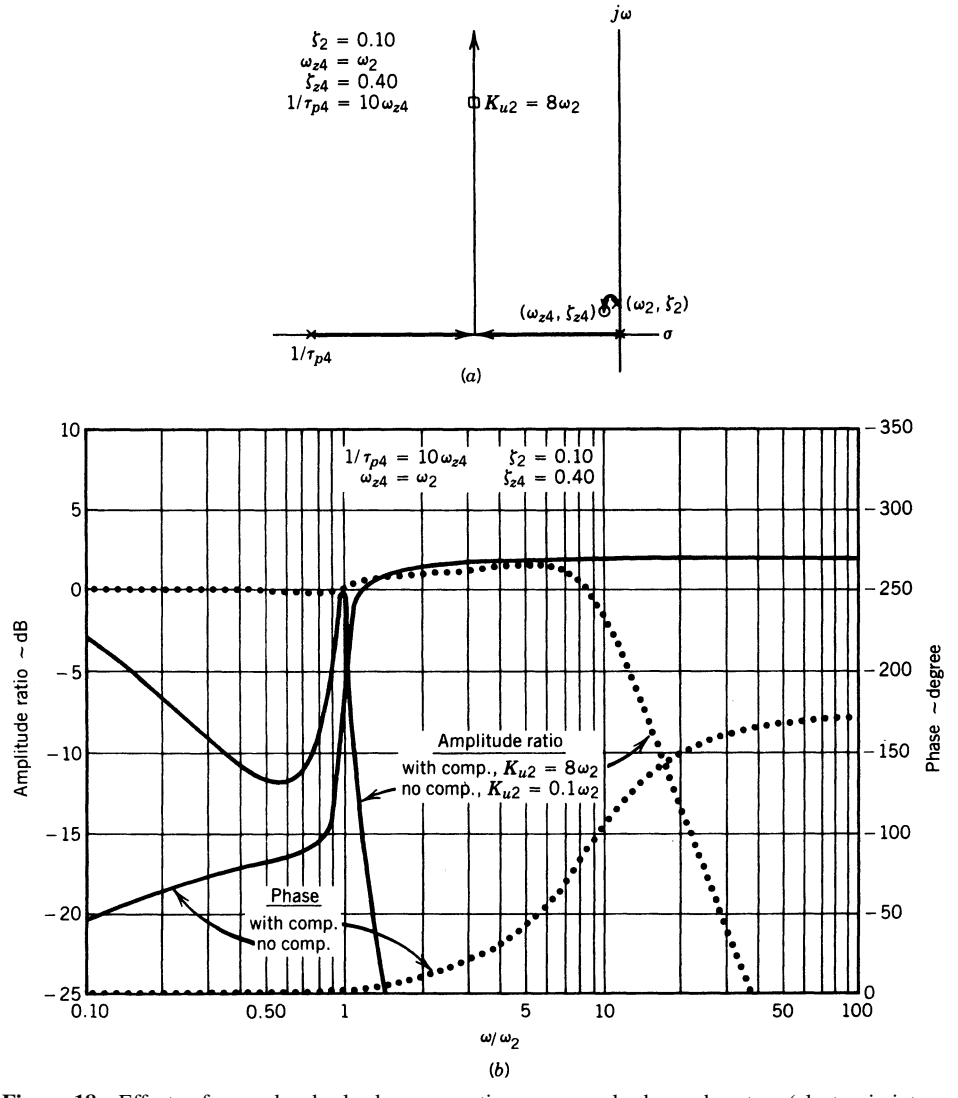


Figure 18 Effects of second-order lead compensation on a poorly damped system (electronic integration): (*a*) root locus; (*b*) closed-loop frequency response.

substantial improvement in loop gain for low ζ_2 but is of little use for $\zeta_2 > 0.3$. For $\zeta_2 = 0.10$, lag compensation allows K_{u2} to be increased from 0.10 to 0.27 s^{-1} with comparable levels of stability (gain margin 6 dB). It should be noted that very little advantage is provided by the use of a higher order low-pass filter instead of the first-order type.

In summary, the use of lead compensation to improve the damping of a poorly damped system is not straightforward, and the end result may be a marginal improvement in closed-loop performance. Lag compensation is more straightforward but offers only a modest improvement in performance. If substantial improvement in system damping and performance

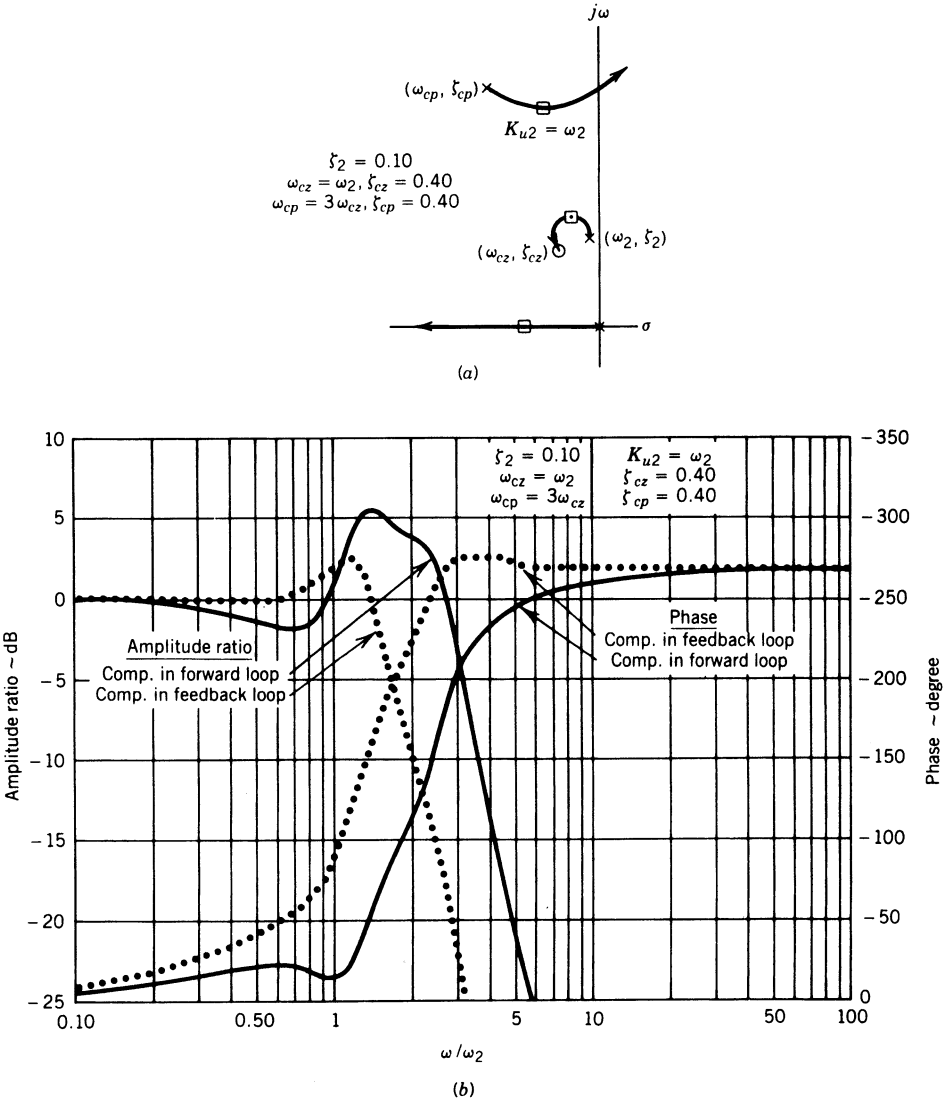


Figure 19 Effects of second-order lead compensation on a poorly damped system (inherent integration): (*a*) root locus; (*b*) closed-loop frequency response.

is required, the use of inner feedback loops is generally more effective, as explained in Section 4.

3.3 Higher Order Effects

In the foregoing sections, the discussion has centered around systems whose forward-loop dynamics can be approximated by relatively simple lag elements. While this approach is often entirely adequate for controller design and performance estimation, the designer should

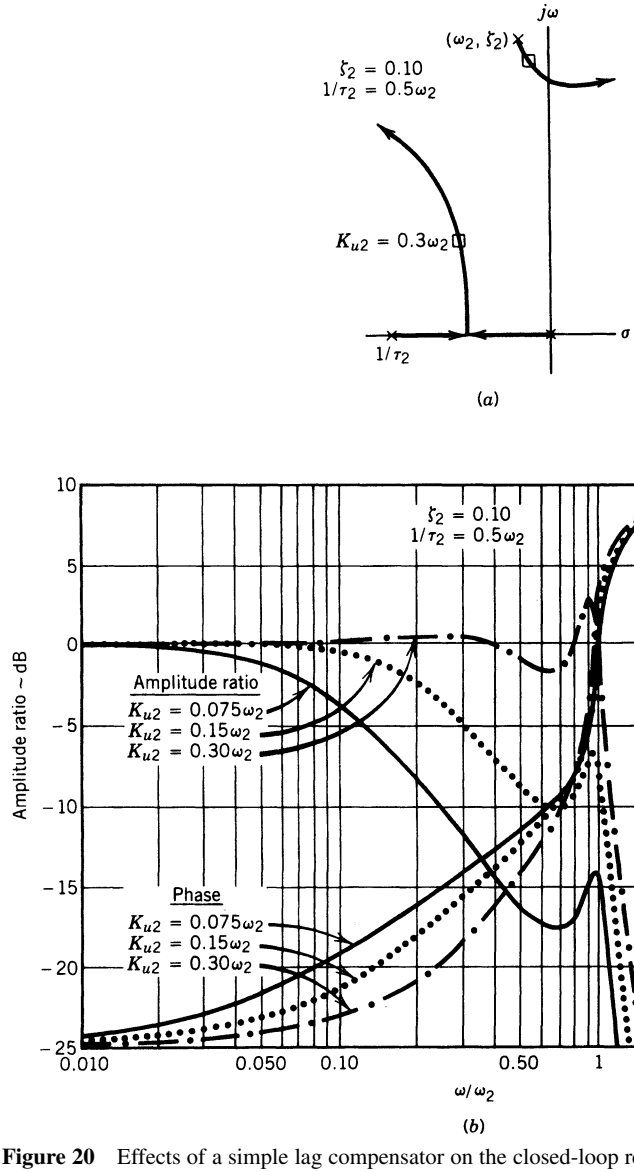


Figure 20 Effects of a simple lag compensator on the closed-loop response of a poorly damped system: (*a*) root locus; (*b*) frequency response.

be aware of several potential limitations. These limitations usually involve the higher order, higher frequency dynamic modes that were unknown or neglected in the early stages of the design. Examples are structural modes, transducer dynamics, and drive amplifier dynamics.

Higher order modes associated with actuator mounting structure, actuator–load mechanical connections, and transducer mounting are usually second order and poorly damped. Often these modes are characterized as poorly damped pole–zero combinations. In either

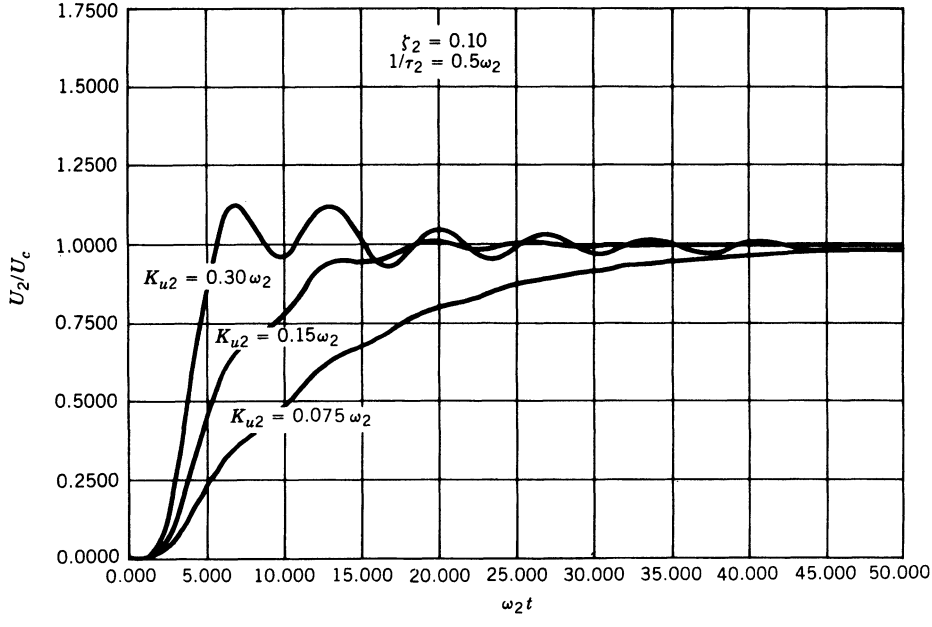


Figure 20 (*Continued*) (*c*) Step response.

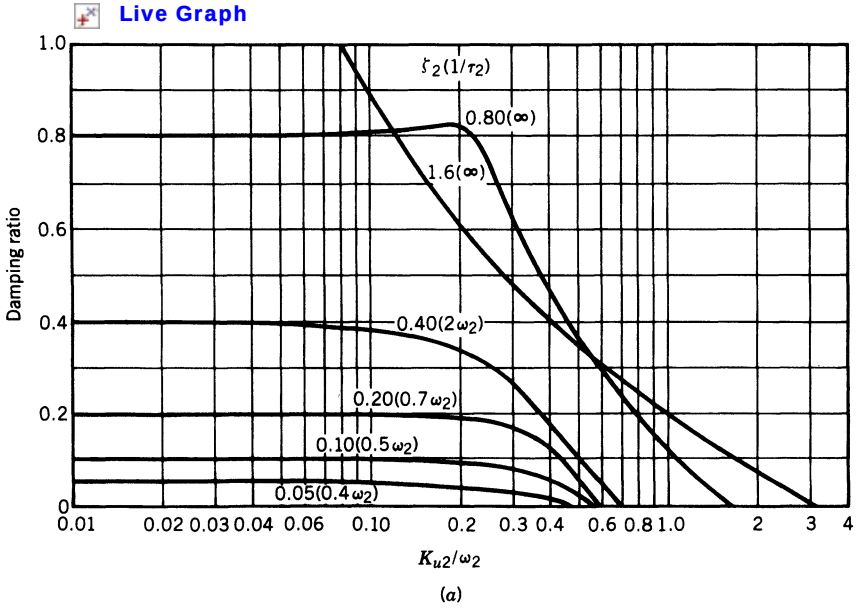
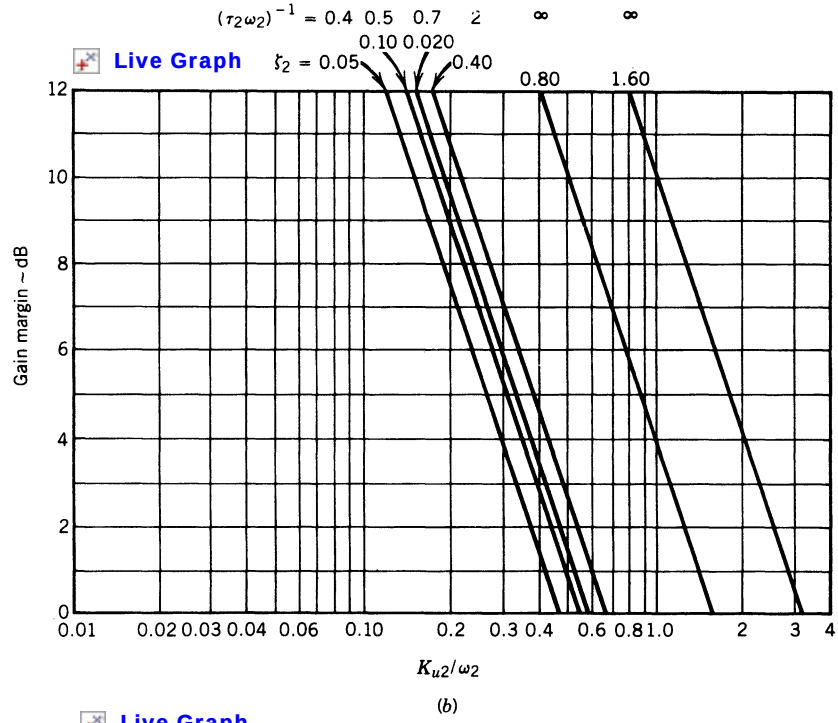


Figure 21 Effects of a simple lag compensator on the stability parameters of a poorly damped system: (*a*) closed-loop damping ratio.



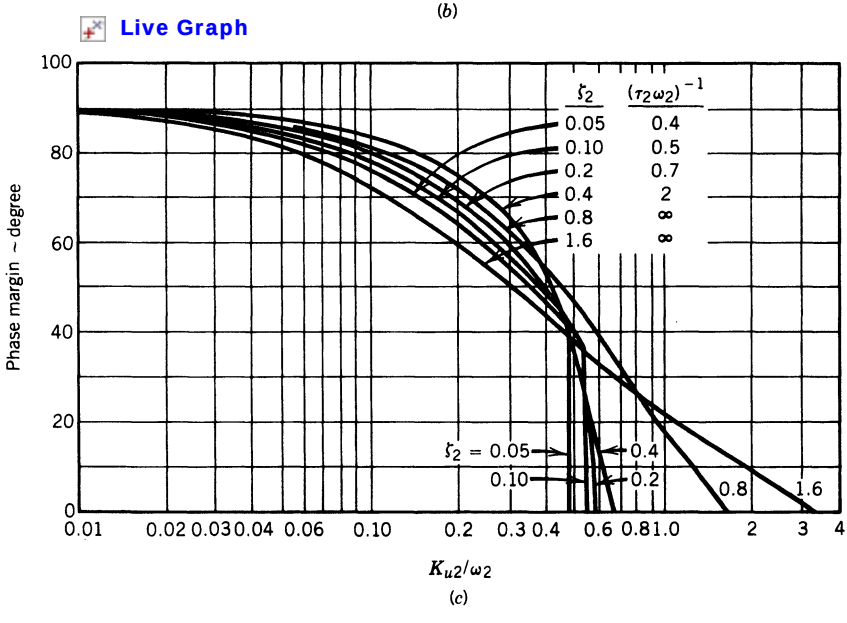


Figure 21 (*Continued*) (*b*) Gain margin; (*c*) phase margin.

case, loop gains selected to obtain adequate performance and stability from the dominant low-frequency modes could be high enough to drive the poorly damped high-frequency modes unstable. This is particularly likely if lead compensation is used to improve the dynamics of the dominant modes. High-frequency stability is probably best assessed by a root-locus plot or Bode plot that includes estimates of all the system dynamic modes. Often, accurate description of the higher order modes is difficult at the design stage. For this reason, it is worthwhile investigating the effects of various high-frequency compensation techniques that can be added when the system is first evaluated experimentally. For example, notch filters are often useful in reducing the effects of high-frequency structural modes at the cost of some low-frequency phase lag.

Many higher order modes are reasonably well damped. The dynamics of drive amplifiers, transducer signal conditioners, and notch filters are typical examples. The primary influence of such modes is that they introduce phase lags that can destabilize the low-frequency dominant modes or limit the effectiveness of lead compensation used to improve low-frequency behavior. A useful approximation of these effects can be made if there is reasonable frequency separation of the higher order modes from the dominant lower frequency modes (a factor of 5 or more). In this case, the system's low-frequency dynamic behavior can usually be assessed by replacing the high-frequency modes with a single first-order lag. The time constant of this first-order approximation, τ_3, can be determined as follows:

$$\tau_3 = \frac{\phi_3}{57.3\omega_3} \tag{20}$$

where ϕ_3 is the net phase lag (degrees) of all the combined high-frequency modes, measured at ω_3 (radians per second). The frequency ω_3 should be approximately equal to the highest natural frequency of the lower frequency dominant modes. It is probably best not to use this simplification if ϕ_3 approaches 30°.

If frequency response data are available, they can be used to estimate ϕ_3. However, at the design stage, it is likely that rough estimates of higher order natural frequencies and damping ratios are the only information available. In this case, ϕ_3 can be determined by adding the phase-lag contributions of the individual high-frequency modes:

$$\phi_3 = \sum \phi_4 + \sum \phi_5 \tag{21}$$

$$\phi_4 = 57.3\tau_4\omega_3 \tag{22}$$

$$\phi_5 = 115\zeta_5 \frac{\omega_3}{\omega_5} \tag{23}$$

where τ_4 is the time constant of a first-order mode, while ω_5 and ζ_5 are the natural frequency and damping ratio of a second-order mode (ϕ_4 and ϕ_5 due to lag terms add to ϕ_3, while lead terms subtract from ϕ_3). The approximations of Eqs. (22) and (23) are accurate to one degree or better for $1/\tau_4 > 3\omega_3$ and $\omega_5 > 3\omega_3$, respectively.

4 INNER FEEDBACK LOOPS

If the desired servo performance cannot be achieved using frequency compensation, as described in Section 3, the addition of inner feedback loops can often provide the needed improvement. This is particularly true when the open-loop damping ratio is poor and the forward loop has an inherent integration. Although inner feedback loops require additional

transducers, they offer more flexibility in modifying the servo dynamics than does frequency compensation alone. The following sections discuss the merits of feeding back derivatives of the controlled variable, feeding back variables dynamically different than the controlled variable, and nonelectronic mechanizations of inner loops.

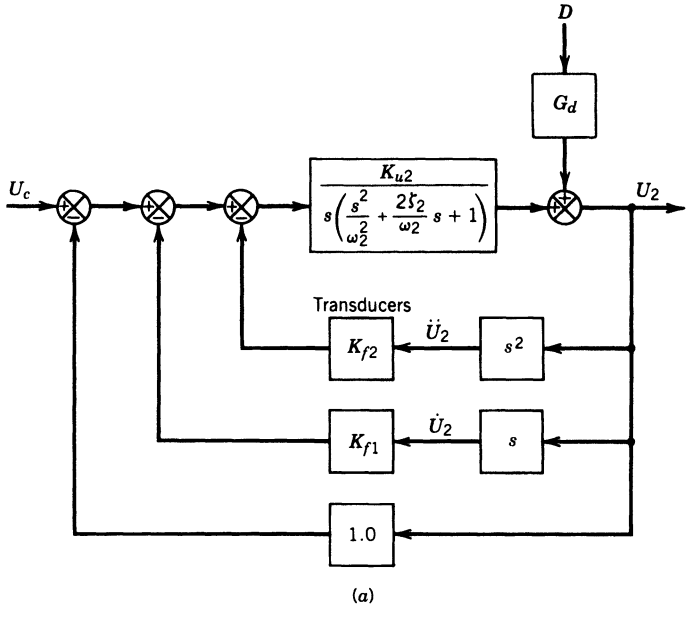
4.1 Derivatives of the Controlled Variable

Section 3 illustrates the benefits of lead compensation but also shows that its effectiveness is limited by the high-frequency gain amplification of the compensator itself. This problem can be alleviated by the use of transducers that directly measure derivatives of the controlled variable. To illustrate the potential benefits, consider again the example of Fig. 6*b* with additional feedback of the first and second derivatives of the controlled variable, as illustrated in Fig. 22*a*. As shown in Fig. 22*b*, the three feedbacks can be mathematically combined into a single 1oop having a pure second-order pair of zeros, without the added lag normally associated with lead compensators. The result is that much higher forward-loop gains can be achieved for a given level of closed-loop stability. Of course, each transducer will introduce higher frequency dynamic effects that will eventually limit the maximum forward-loop gain, but these effects are usually less restrictive than those imposed by electronic lead compensators. Note that the natural frequency and damping ratio of the feedback zeros in Fig. 22 are determined by the magnitude of the derivative feedback gains relative to the primary feedback gain, which is 1.0 in this case.

There are a variety of uses for derivative feedback loops, including improved closed-loop accuracy, bandwidth, and stability, as well as reduced sensitivity to changes in system parameters. These uses can be illustrated with the aid of Figs. 23 and 24, which show the effects of various sets of feedback zeros on the closed-loop dynamics of Fig. 22*b*. Figure 23 gives root loci for the different zero locations. Because electrical noise and excitation of high-frequency dynamic modes usually limit the gain in the highest derivative loop, a forward-loop gain is selected for each locus that holds $K_{u2}K_{f2}$ constant. Closed-loop frequency responses are then given in Fig. 24. It can be seen that closed-loop damping is improved in all cases. Placement of the feedback zeros at high frequency (low derivative feedback gains) yields high closed-loop bandwidth, but the second-order poles can easily be destabilized by a reduction in forward-loop gain or the presence of the inevitable higher order modes described in Section 3.3. On the other hand, low-frequency feedback zeros (high derivative feedback gains) result in lower closed-loop bandwidth but offer greatly reduced gain sensitivity and are little affected by high-frequency modes. Also in the latter case, the closed-loop response is dominated by the low-frequency second-order poles, which are nearly the same as the feedback zeros. Therefore, the closed-loop characteristics do not vary significantly with large changes in the plant parameters (K_{u2}, ω_2, ζ_2). A good compromise between bandwidth and parameter sensitivity is to set ω_{f2} approximately equal to ω_2.

Comparison of Figs. 23 and 24 with Figs. 18 and 19 show that derivative feedback offers better stability and more flexibility than lead compensation, particularly when the forward loop has an inherent integration. In addition, higher forward-loop gains can generally be used with derivative feedback, which improves the static accuracy of the system. However, forward-loop lead compensation can often produce higher closed-loop bandwidth if the forward-loop integration is electronic in nature, as illustrated in Fig. 18.

A useful way to optimize the system of Fig. 22*a* is to first establish rough estimates of the feedback and forward-loop gains using the combined feedback approach of Figs. 22*b* and 23. It is often helpful to include a first-order approximation of the phase lags caused by the higher order modes [Eq. (20)]. Once the approximate gains are established, the closed-



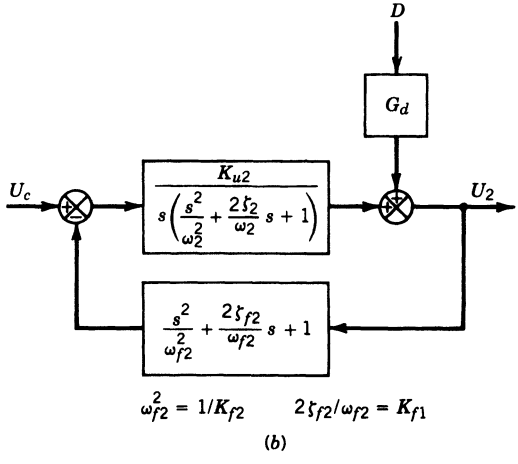


Figure 22 Feedback of controlled-variable derivatives: (*a*) inner loops to improve damping; (*b*) combined feedback loops.

loop response and stability can be checked by analysis of a complete multiloop model, with the higher order modes described more completely and placed in the appropriate loops.

If the forward-loop integrator is electronic in nature, the benefits of derivative feedback can be achieved with one less derivative than if the integrator is inherent. This is accomplished by feeding back the inner 1oops downstream of the integrator, as illustrated in Fig. 25. The static and dynamic characteristics of Fig. 25 are entirely equivalent to those of Fig. l7.22*a* and can also be mathematically reduced to the single-loop configuration of Fig. 22*b*.

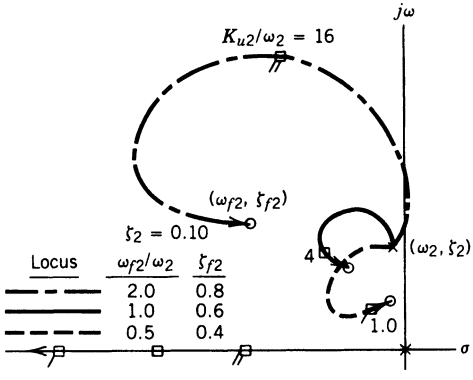


Figure 23 Effects of derivative feedback on closed-loop roots.

Note, however, that the practical implementation of Fig. 25 requires one less transducer than Fig. 22*a*.

Another way in which derivative feedback can be useful is to provide very smooth and repeatable dynamic response and high static accuracy when the primary control loop has an inherent integration. This is particularly useful when closed-loop bandwidth is not a major concern. The technique involves closing a tight integrating loop around the first derivative of the controlled variable, as illustrated in Fig. 26. This loop submerges the effects of forward-loop gain variations, static offsets, and external disturbances. The primary control loop gain can then be set at relatively low levels to ensure smooth, repeatable dynamic

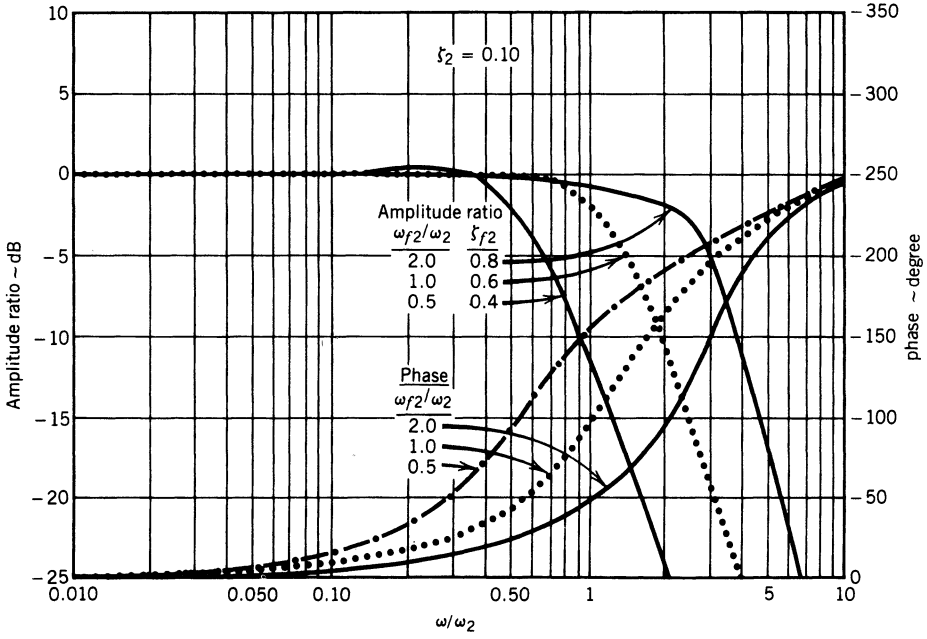


Figure 24 Effects of derivative feedback on closed-loop frequency reqponse of U_2/U_c.

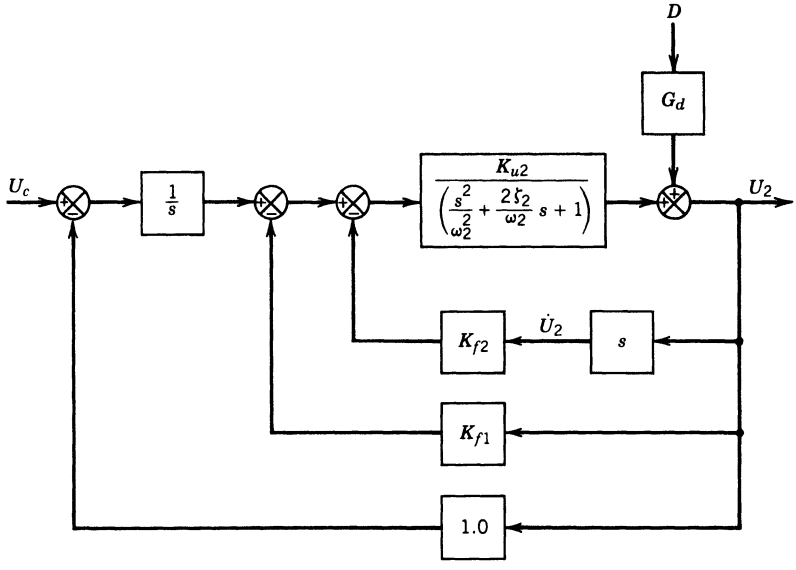


Figure 25 Rearrangement of Fig. 22*a* inner loops (electronic integration).

response without the usual concerns about reduced static accuracy. This technique is also useful if the mounting arrangement of the primary transducer results in gain-limiting higher-order dynamic characteristics in the outer feedback loop.

4.2 Alternative Inner Loop Variables

Sometimes it is advantageous to close inner feedback loops around variables that are dynamically different than derivatives of the primary controlled variable. This might be done because of practical problems related to accurately transducing derivatives of the controlled

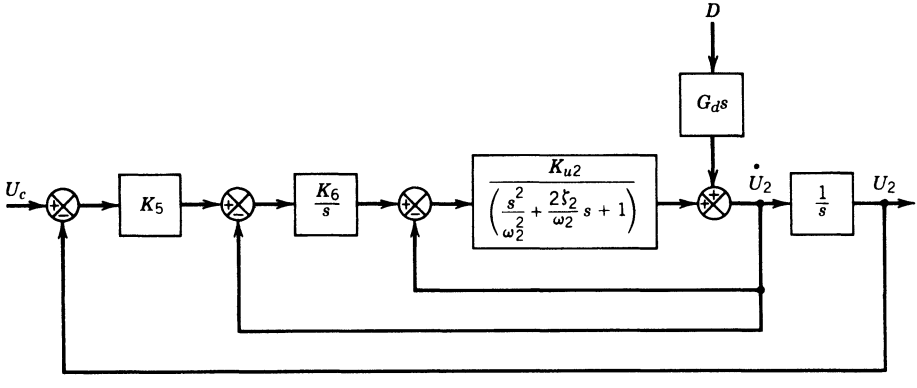


Figure 26 Rearrangement of Fig. 6 to give smooth, repeatable, and accurate response (inherent integration).

variable or because it offers an inherent advantage related to feedback dynamic characteristics. Usually, there are advantages and disadvantages of using alternative feedback loops, as illustrated in the following example.

Consider an electrohydraulic position servo that has rather demanding requirements for dynamic response. Suppose that envelope or environmental constraints make it very difficult to mount velocity and acceleration transducers or that the mounting arrangement itself introduces undesirable higher order dynamics. In this case, consideration might be given to transducing cylinder differential pressure in place of the velocity or acceleration feedbacks. If the load can be represented as primarily a mass, cylinder pressure will be proportional to load acceleration. If cylinder or load friction is large, it may be necessary to use a load cell rather than a pressure transducer. However, the load is often more complex than a simple mass. For example, the load might also have substantial stiffness to ground and viscous damping. In this case, cylinder pressure would have components proportional to load acceleration, velocity, and position. If these components were of the proper relative size and not highly variable, a pressure transducer alone could replace the acceleration and velocity transducers. However, if the load dynamics were complex or highly variable, the use of cylinder pressure as an inner loop might do more harm than good.

Another potential problem associated with the use of alternative inner feedback 1oops is the influence of external disturbances. Consider again the example of the electrohydraulic position servo. Feedback of either velocity or acceleration will have no effect on closed-loop static stiffness because, by definition, all derivatives of position are zero in the steady state. However, an external force applied to the load will change the cylinder pressure, even in the steady state. Therefore, pressure feedback will reduce closed-loop stiffness unless the pressure feedback signal is high passed, which introduces its own set of dynamic characteristics. Of course, there are applications in which closed-loop stiffness is not critical in the first place.

It should also be mentioned that the mounting of the primary transducer can introduce its own dynamic peculiarities. For example, the controlled variable might be load position relative to ground, and the transducer might be integrally mounted within the servoactuator assembly. If there were substantial compliance in the structure to which the actuator is mounted, the position feedback loop would contain structural zeros. An integral velocity transducer would have the same problem, but a load-mounted accelerometer would not. In this case and in others where the derivative feedback loops are dynamically different from one another, the simplified techniques of Section 4.1 are of limited use, and the complete multiloop model must be analyzed directly. Again it should be noted that these more complex derivative feedback loops may result in better or poorer closed-loop performance than comparable inner 1oops that feed back pure derivatives of the controlled variable. Proper assessment of these trade-offs requires a good physical model of the system, showing the proper relationships of all the feedback variables being considered.

4.3 Nonelectronic Inner Loops

Occasionally, it is useful to implement inner loop feedbacks by mechanical design rather than electronic means. For example, it may be possible to mount the servoactuator or primary transducer so that structural deflections under load produce favorable feedback zeros that improve closed-loop damping. In an electrohydraulic servo, improved damping can often be obtained by using hydraulic pressure feedback, implemented with a cross-port orifice or laminar leakage path. Both of the schemes will suffer loss of closed-loop stiffness. Sometimes it is possible to mount an electrohydraulic servoactuator so that its rod attaches to the mount-

ing structure and its body attaches to the load. If a mass is then attached to the control valve spool and the spool is aligned with the actuator centerline, a form of acceleration feedback can be achieved (valve porting must be arranged to give proper feedback polarity). Mechanical feedback schemes offer the potential advantages of reduced costs and complexity, as well as alleviating the need for high open-loop bandwidth in the actuator and controller. However, the design of servos using these techniques often requires manipulation of rather complex physical models, which is beyond the scope of this chapter.

5 PREFILTERS AND FEEDFORWARD

As can be seen from Sections 3 and 4, the business of obtaining good closed-loop performance can become rather complex. Sometimes feedback loops and frequency compensation are optimized to achieve the desired stability and accuracy, but the closed-loop response to commands is not particularly desirable. Rather than compromise stability or accuracy by altering the servoloop characteristics, it is often easier to shape the command signal before it enters the servoloop. The following sections discuss some commonly used techniques for accomplishing this, together with their limitations.

5.1 Lag Prefilters

High-gain servoloops are required to achieve static accuracy and rejection of load transient disturbances, but rapid response to commands is often unnecessary or undesirable. In this case, the addition of a simple lag prefilter will often provide the desired result:

$$G_{\text{pf}} = \frac{1}{\tau_{\text{pf}} s + 1} \tag{24}$$

If this prefilter is placed in the command path, as illustrated in Fig. 27, τ_{pf} can be made large enough that G_{pf} dominates the U/U_c response, with an appropriate rise time.

Another way that lag prefilters can be used is to provide a particular set of dynamics for U/U_c, which are easily settable and do not change with variations in the forward-loop parameters. Often called "model following," this technique requires the use of high-bandpass servoloops, so that the dynamics of the prefilter model (G_{pf}) dominate the U/U_c response. In concept, this makes it very easy to obtain any desired U/U_c. transfer function by simply changing the electronic prefilter model. However, to have U/U_c faithfully reflect the model dynamics, it is often necessary for the bandpass of the servoloop to be an order of magnitude higher than the highest frequency singularity in the model transfer function. This is often impractical. If inner feedback loops are needed to achieve the desired servoloop bandpass, it may be more effective to tailor feedback 1oops to provide a combined feedback transfer

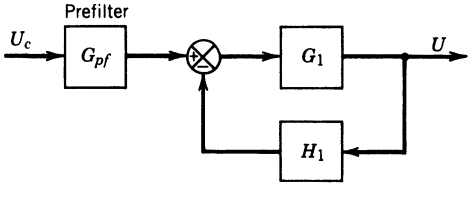


Figure 27 Generalized use of prefilters.

function that is the inverse of the desired U/U_c transfer function. This can be seen with the aid of Fig. 27 by eliminating the prefilter. If the forward-loop gain is high enough, $U/U_c = 1/H_1$. This technique is further discussed in Section 4.1, and an example is given in Fig. 22.

5.2 Lead Prefilters

Since the closed-loop response characteristics of most servoloops are dominated by lag elements, lead prefilters are often used to improve the response to command inputs. Theoretically, this can be accomplished by simply making the prefilter transfer function equal to the reciprocal of the servoloop closed-loop transfer function. For the generalized system illustrated in Fig. 27, the ideal prefilter would be

$$G_{\text{pf}} = \frac{1 + G_1H_1}{G_1} \tag{25}$$

Unfortunately, the lead required to accomplish this will be limited by its associated poles, which must be selected to prevent excessive electrical noise (as discussed in Section 3.1). Also lead prefilters can accentuate the oscillatory tendencies of a poorly damped servoloop. Even for a well-damped servoloop, overshooting response can occur if the servoloop parameters vary substantially over the operating envelope. In many cases, the lead network may be more effective if it is moved to the forward loop so that higher servoloop gains can be used.

To illustrate the effects of lead prefilters, consider a servoloop of the configuration shown in Fig. 6*a*. Since the closed-loop transfer function will be a second-order lag, as shown by Eq. (8), a second-order prefilter lead would be appropriate. The use of such a prefilter is illustrated in Fig. 28. Assume that the forward-loop gain is selected to give a closed-loop damping ratio $\zeta_6 = 0.50$. This would require $K_{u1} = 1/\tau_1$, as determined from Fig. 12. The closed-loop natural frequency can then be determined from Eq. (8), as follows:

$$\omega_6 = \sqrt{\frac{K_{u1}}{\tau_1}} = \frac{1}{\tau_1} \tag{26}$$

The prefilter zeros can now be set to cancel the closed-loop lag: $\omega_{\text{pz}} = 1/\tau_1$ and $\zeta_{\text{pz}} = 0.50$. As discussed in Section 3.1, a practical upper limit on the high-frequency gain amplification of a lead network is approximately a factor of 10. Therefore, the prefilter poles can be set to $\omega_{\text{pp}} = 3/\tau_1$ and $\zeta_{\text{pp}} = 0.50$. The net effect of the prefilter is an effective boost in closed-loop natural frequency by a factor of 3.

It should be noted that the use of lead prefilters can cause some peculiar effects when variations in the forward-loop characteristics are considered. For the example of Fig. 28, the effects of forward-loop gain variations are illustrated in Figs. 29 and 30. Because the closed-loop peaking at high gains is accentuated by the prefilter, it is usually best to optimize the prefilter for the maximum-gain situation and accept the degraded response at low gains. If

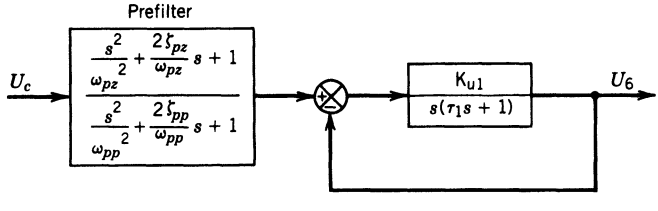


Figure 28 Example of a lead prefilter.

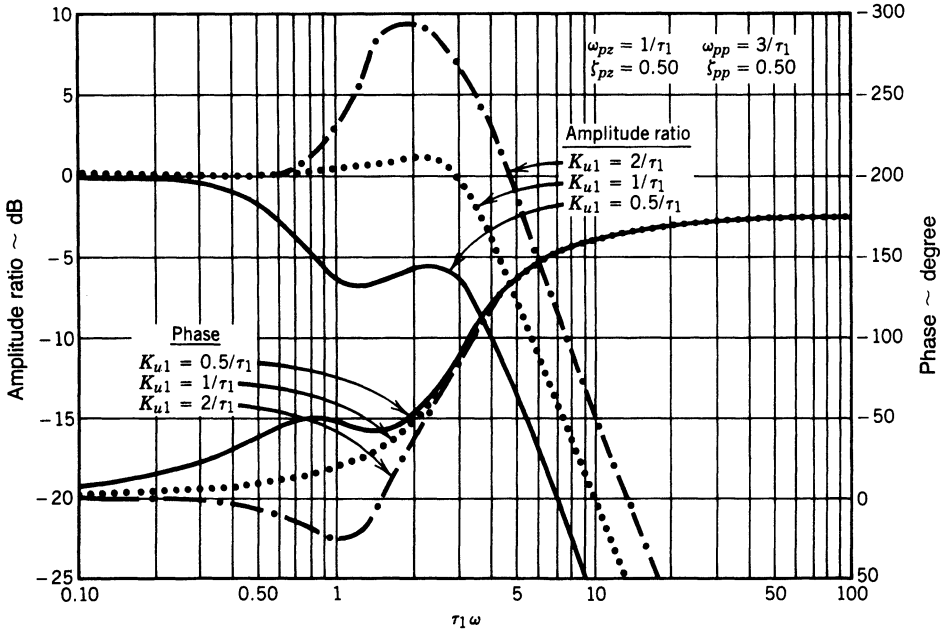


Figure 29 Effects of loop gain variations of U_6/U_c frequency response.

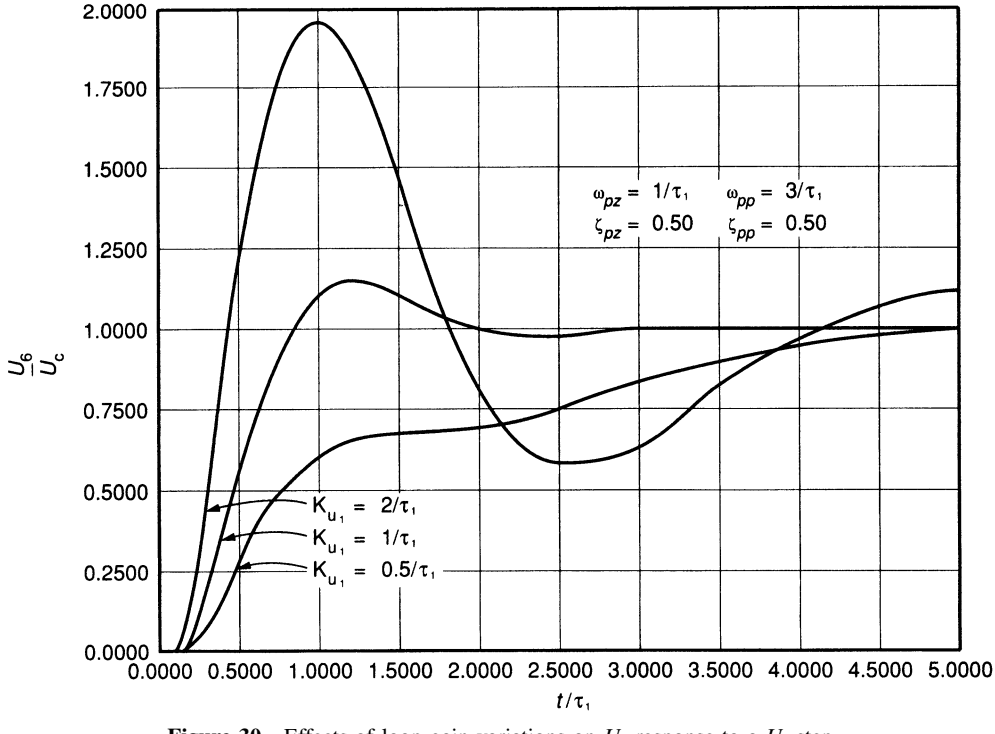


Figure 30 Effects of loop gain variations on U_6 response to a U_c step.

the variations in forward-loop characteristics are large, it may be necessary to use inner feedback loops to minimize the variations in closed-loop response, as explained in Section 4.1.

5.3 Feedforward

As explained in Section 5.2, the effectiveness of lead prefilters is limited by the fact that the command signal must be differentiated. In a system whose forward loop contains electronic lags, the number of command differentiations can often be reduced by the use of feedforward techniques. Feedforward can also be used to reduce the following errors associated with a steady rate of change of the command signal. However, these techniques must be carefully applied to prevent adverse effects on system dynamic response.

The principles of feedforward are illustrated in Fig. 31*a*. If the electronic shaping in the feedforward path approximates the reciprocal of the nonelectronic forward-loop elements, the command signal will nominally be reproduced at the output. Follow-up by the feedback loop will then try to minimize the effects of inaccuracies in the feedforward signal. The use of a prefilter matched to the feedback path further improves the overall response to commands. The net effect of the feedforward configuration can be seen by rearranging the block diagram into an equivalent one that has only a prefilter, as shown in Fig. 31*b*. From this figure, the system response can be written

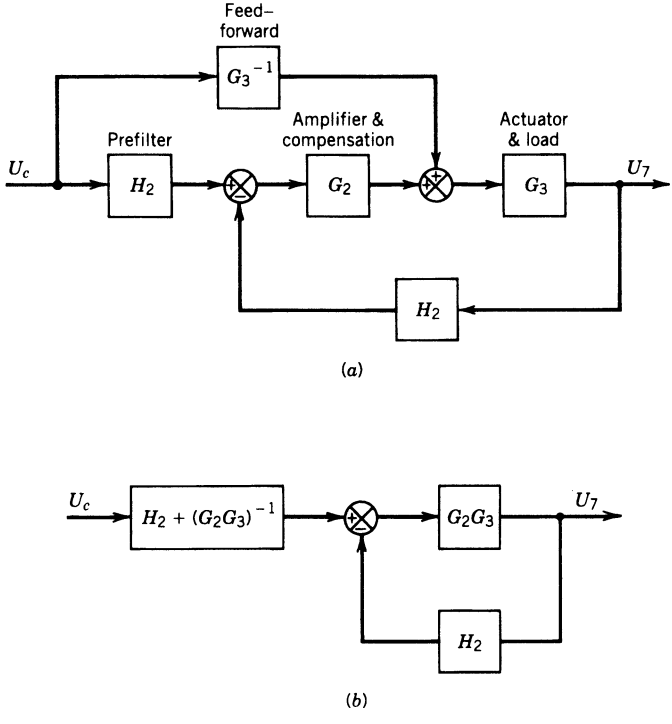


Figure 31 Idealized use of feedforward: (*a*) feedforward model; (*b*) equivalent prefilter model.

$$\frac{U_7}{U_c} = [H_2 + (G_2G_3)^{-1}] \frac{G_2G_3}{1 + G_2G_3H_2}$$

$$= \frac{G_2G_3H_2 + 1}{G_2G_3} \frac{G_2G_3}{1 + G_2G_3H_2} = 1.0 \tag{27}$$

Of course, it is not possible to actually achieve the ideal result given by Eq. (27) because of the lags associated with the lead network in the feedforward path. To illustrate the practical aspects of feedforward, it is useful to reexamine the example of Fig. 28. If the forward-loop integrator is electronic in nature, the use of feedforward offers some advantages over the straight prefilter. Figure 32*a* shows the appropriate form for the feedforward model, and Fig. 32*b* shows it reduced to an equivalent prefilter model. Notice that Fig. 32*a* requires only one differentiation of the command signal, while Fig. 28 requires two. The result is that its equivalent prefilter (Fig. 32*b*) includes a first-order lag rather than a second-order lag (Fig. 28).

To allow direct comparison of the Figs. 28 and 32 examples, the same loop gain is used in each case ($K_{u1} = 1/\tau_1$). Noting that τ_8 will be approximately equal to τ_1 and that the high-frequency gain amplification of the feedforward network should be limited to 10, it is sensible to set $\tau_7 = 0.1\tau_1$. Since the closed-loop poles are given by Eq. (8), the feedforward parameters are selected as follows:

$$\frac{K_8\tau_8}{K_{u1}} = \frac{\tau_1}{K_{u1}} \tag{28}$$

$$\frac{K_8}{K_{u1}} + \tau_7 = \frac{1}{K_{u1}} \tag{29}$$

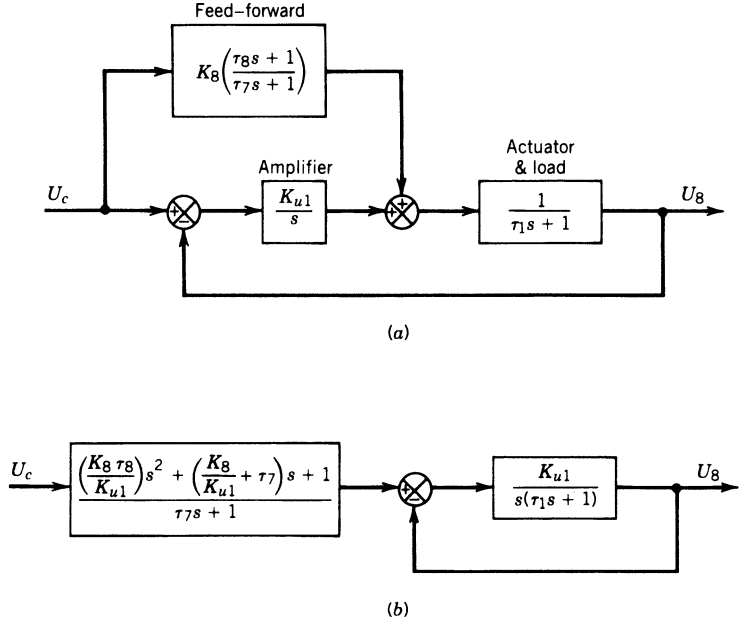


Figure 32 Example of feedforward: (*a*) feedforward model; (*b*) equivalent prefilter model.

Solving these equations after the appropriate substitutions, the feedforward parameters are $K_8 = 0.90$, $\tau_8 = 1.11\tau_1$, $\tau_7 = 0.1\tau_1$. Using these parameters, U_8/U_c frequency and step responses are computed and presented in Figs. 33 and 34, along with the corresponding U_6/U_c responses. In this example, the use of feedforward techniques offers a substantial improvement in system bandwidth, by comparison with a prefilter alone. Of course, the feedforward scheme suffers from sensitivity to forward-loop gain variations similar to those of the prefilter scheme, as illustrated in Figs. 29 and 30.

As previously mentioned, feedforward techniques can also reduce following errors. In general, the following error U_e can be calculated as

$$\frac{U_e}{\dot{U}_c} = \frac{1}{s}\left(1 - \frac{U}{U_c}\right) \tag{30}$$

For the example of Fig. 32, in which the closed-loop transfer function is determined by the feed-forward pole, the following error is determined from

$$\frac{U_{8e}}{\dot{U}_c} = \frac{1}{s}\left(1 - \frac{1}{\tau_7 s + 1}\right) = \frac{\tau_7}{\tau_7 s + 1} \tag{31}$$

For steady command rates, the error is $\tau_7 \dot{U}_c$. Without feedforward, the following error can be determined from Eq. (8):

$$\frac{U_{1e}}{\dot{U}_c} = \frac{1}{s}\left\{1 - \left[\frac{1}{(\tau_1/K_{u1})s^2 + (1/K_{u1})s + 1}\right]\right\}$$

$$= \frac{(\tau_1/K_{u1})s + (1/K_{u1})}{\tau_1/K_{u1})s^2 + (1/K_{u1})s + 1} \tag{32}$$

For steady command rates, the error is $\dot{U}_c/K_{u1}$. For this example, $\tau_7 = 0.1/K_{u1}$. In this case, the proper use of feedforward has reduced the following errors by a factor of 10.

6 PID CONTROLLERS

A very popular form of controller is called PID (proportional-integral-differential). It is very simple in concept and is relatively easy to mechanize. Many essays have been written that describe rules of thumb for ''tuning'' the controller. Unfortunately these tuning procedures can be rather tedious and are usually applicable for only very simple actuator–load dynamics. The purpose of this section is to offer a unified rationale for applying PID controllers that is useful in synthesizing a control system. This rationale also provides insight for systematically adjusting PID parameters on actual hardware.

6.1 Equivalence to Frequency Compensation

The basis for the ensuing discussion is that a PID controller is simply a particular form of forward-loop frequency compensation. This can be seen from the generalized controller shown in Fig. 35. Note that the differential path is filtered to limit the high-frequency amplitude ratio. Combining the parallel paths of Fig. 35, a single transfer function for the controller can be written:

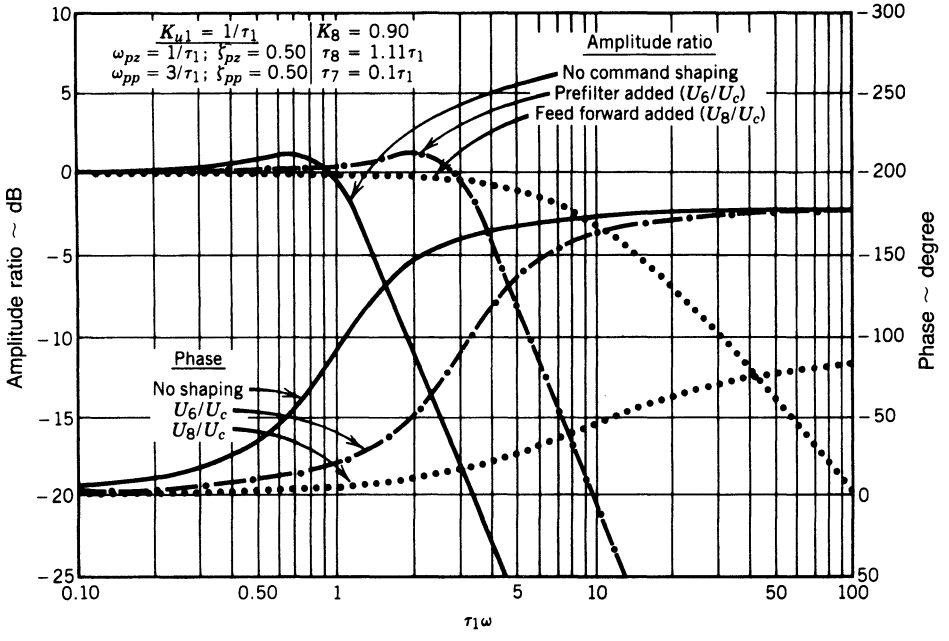


Figure 33 Comparison of closed-loop frequency responses for the prefilter and feedforward examples.

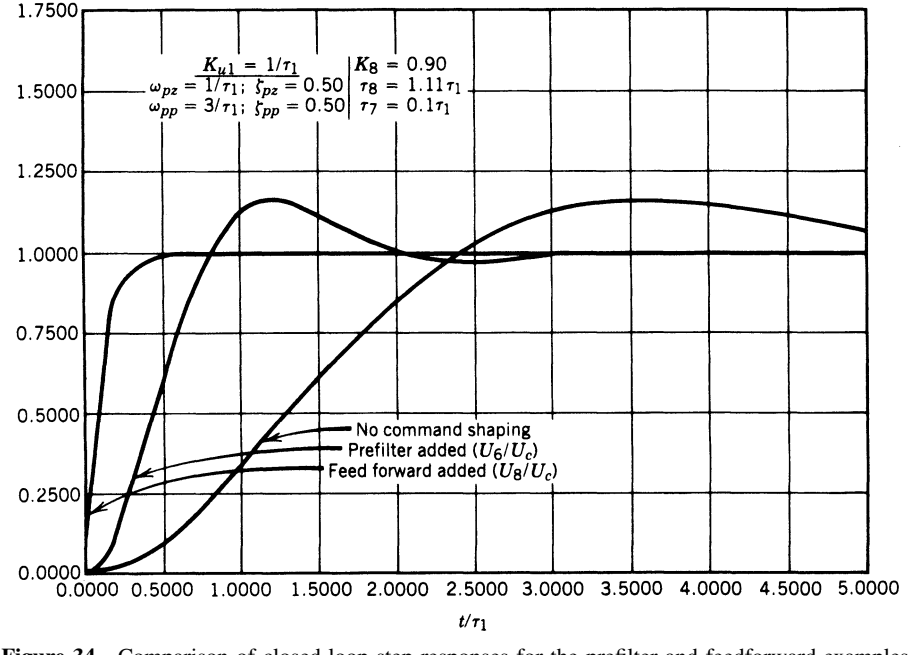


Figure 34 Comparison of closed-loop step responses for the prefilter and feedforward examples.

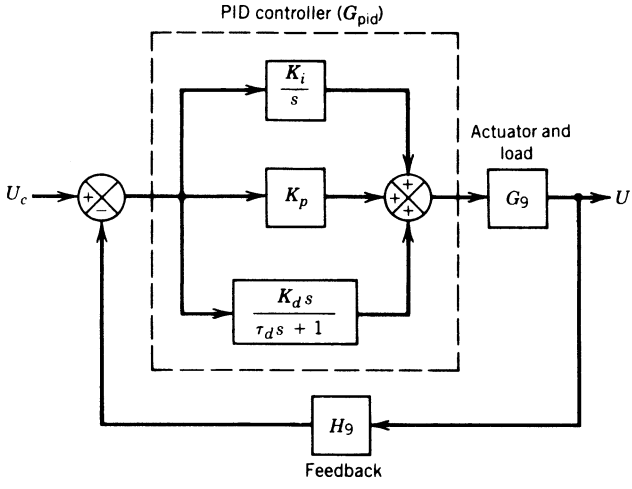


Figure 35 Generalized PID controller.

$$G_{\text{pid}} = \left(\frac{K_d s}{\tau_d s + 1}\right) + K_p + \left(\frac{K_i}{s}\right)$$

$$= \left(\frac{K_i}{s}\right)\left\{\frac{[K_d/K_i + (K_p/K_i)\tau_d]s^2 + [K_p/K_i + \tau_d]s + 1}{\tau_d s + 1}\right\}$$

$$= \left(\frac{K_i}{s}\right)\left[\frac{s^2/\omega_{\text{pid}}^2 + (2\zeta_{\text{pid}}/\omega_{\text{pid}})s + 1}{\tau_d s + 1}\right] \tag{33}$$

Note that Eq. (33) is simply the transfer function of a lead compensator combined with an integrator, as given by Eq. (19). As discussed in Section 3.1, it is usually possible to place the lag break frequency a factor of 10 above the lead natural frequency. In this case, the τ_d terms in the numerator are usually small:

$$\omega_{\text{pid}} = \sqrt{\frac{K_i}{K_d}} \qquad \text{for } \tau_d \ll \frac{K_d}{K_p} \tag{34}$$

$$2\zeta_{\text{pid}}\omega_{\text{pid}} = \left(\frac{K_p}{K_d}\right) \qquad \text{for } \tau_d \ll \frac{K_p}{K_i} \tag{35}$$

In some applications, second-order lead compensation is not required. In such cases, a simplified version of the PID controller can often be useful. This so-called proportional-integral (PI) controller is formed by setting K_d to zero. The resulting transfer function can then be derived from Eq. (33):

$$G_{\text{pi}} = \left(K_p + \frac{K_i}{s}\right) = \left(\frac{K_i}{s}\right)\left(\frac{K_p}{K_i}s + 1\right)$$

$$= \left(\frac{K_i}{s}\right)(\tau_{\text{pi}}s + 1) \tag{36}$$

This result is similar to Eq. (18).

Another simplified version of the PID controller is obtained by setting $K_i = 0$. The transfer function of this so-called proportional-differential (PD) controller can also be derived from Eq. (33):

$$G_{\text{pd}} = \left(\frac{K_d s}{\tau_d s + 1}\right) + K_p$$

$$= K_p\left[\frac{(K_d/K_p + \tau_d)s + 1}{\tau_d s + 1}\right]$$

$$= K_p\left[\frac{\tau_{\text{pd}} s + 1}{\tau_d s + 1}\right] \tag{37}$$

This transfer function is the same as the first-order lead compensator of Eq. (16) and is normally used in systems having an inherent integration. Note that the lead break frequency can be a factor of 10 lower than the lag break frequency. In this case the τ_d term in the numerator of Eq. (37) is small:

$$\tau_{\text{pd}} = \left(\frac{K_d}{K_p}\right) \qquad \text{for } \tau_d \ll \frac{K_d}{K_p} \tag{38}$$

Frequency responses of representative PID, PI, and PD controllers are given in Fig. 36, which also shows the effects of the various controller parameters.

6.2 Systems Having No Inherent Integration

Electronic integrators are normally used in the forward loops of systems having no inherent integrations, as explained in Section 2.1. Section 3 explains the various ways in which lead

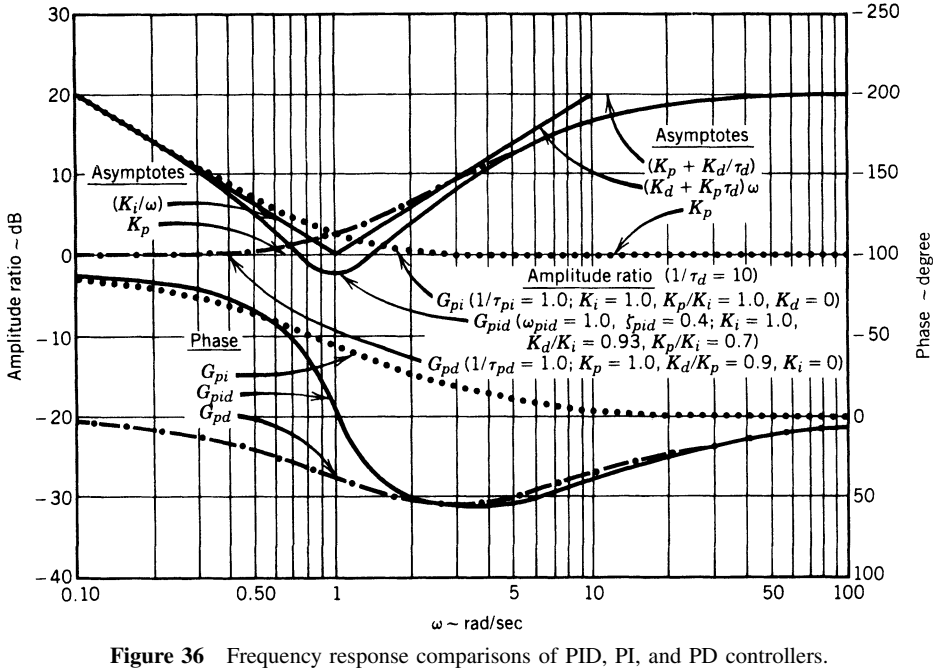


Figure 36 Frequency response comparisons of PID, PI, and PD controllers.

compensation can be usefully applied. A PID controller offers a convenient method for combining the electronic integration with lead compensation. As discussed in Section 6.1, the PI scheme provides first-order lead, while the complete PID scheme offers second-order lead.

6.3 Systems Having an Inherent Integration

Section 2.1 explains that an electronic integrator in the forward loop minimizes static servo errors. However, the addition of an electronic integrator to a system that already has an inherent integration will usually cause dynamic instability. To prevent this type of instability, lead compensation must be combined with the electronic integrator. This can be accomplished with a PI controller.

As shown in Fig. 36, a PI controller contributes nearly 90° of phase lag at low frequencies. If this is added to the 90° of lag already contributed by the integrator inherent in the system, the total low-frequency lag approaches 180°. Since the other system dynamics will add even more phase lag at high frequencies, the PI break frequency must be set low enough to ensure an intermediate frequency range over which the phase lag is reduced. The open-loop frequency responses of Fig. 37*a* illustrates this effect for a system whose inherent characteristics consist of an integrator and a second-order lag. Generally, a PI break frequency greater than 10% of the system's lowest lag frequency will substantially reduce closed-loop stability. This is shown by the phase plots of Fig. 37*a*, by the root loci of Fig. 37*b*, and by the closed-loop frequency responses of 37*c*.

Using the 10% rule of thumb for the PI controller, it is interesting to examine some time histories of the closed-loop system. Figure 38 shows time responses to step and ramp commands for the system of Fig. 37. As illustrated in Fig. 38*a*, the PI controller degrades the step response. Figure 38*b* shows that the PI's double integration at low frequencies eliminates the following error but causes larger overshoot of the steady state.

The 10% rule also applies to systems whose inherent characteristics consist of an integrator and a first-order lag. Frequency responses, root loci, and time histories for such a system are given in Fig. 39. It should also be noted that the effective lag break frequency of the system can be increased by using lead compensation techniques, as described in Section 3. For the system of Fig. 39, this can be accomplished by using a PID, rather than a PI, controller. Reexamining the example of Fig. 39 using the PID characteristics of Fig. 36, a 10-fold increase in effective system bandwidth can be achieved by setting $\omega_{pid} = 1/\tau_1$ and $\zeta_{pid} = 1.0$. In this case, one of the two PID lead terms cancels the system lag at $1/\tau_1$. The resulting open-loop frequency response is shown in Fig. 40.

7 EFFECTS OF NONLINEARITIES

The previous sections have concentrated on the design of controllers for linear systems. In practice, physical systems are never truly linear. If the nonlinearities are not large, design of the controller using linear techniques is very useful. However, it is important to understand the limitations of this approach. The following sections discuss these limitations and offer several approaches to dealing with nonlinearities.

Figure 41 illustrates idealized forms of several nonlinearities that are commonly encountered in systems controlling mechanical loads. Saturating nonlinearities can occur in transducers, electronics, and the servoactuator itself. Deadzone is the lack of output for small changes in input and is generally most significant in servoactuators and transducers. Reso-

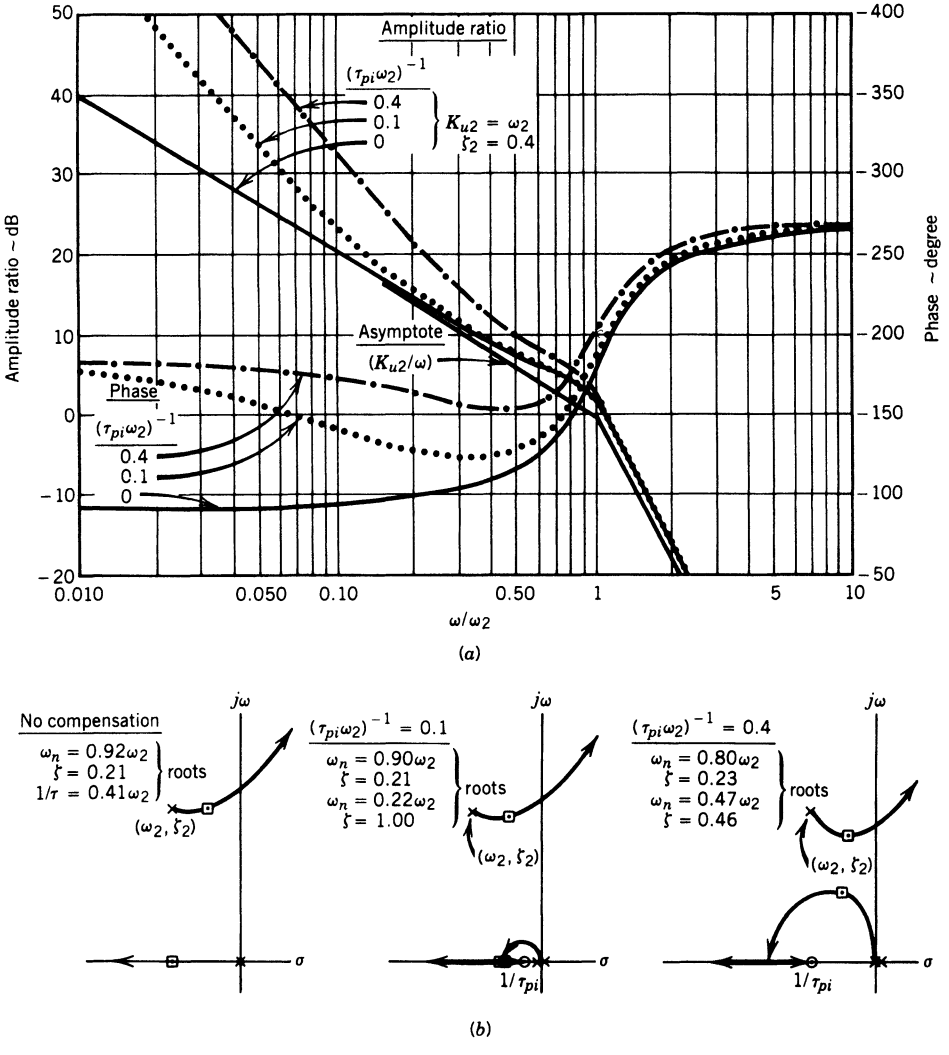


Figure 37 PI controller added to the system of Fig. 6*b*—effects of lead break frequency: (*a*) open-loop frequency response; (*b*) root loci ($\zeta_2 = 0.4$, $K_{u2} = 0.35\omega_2$).

lution is the availability of a limited number of output values and is typical of digital electronics and many types of transducers, including encoders and wire-wound potentiometers. Most servoactuators provide output velocities that are force dependent or output forces that are velocity dependent. The load–velocity curves shown in Fig. 41*d* are typical of an electrohydraulic servoactuator and are highly nonlinear. Coulomb friction is a constant force that always opposes motion. Static friction (stiction) is often larger than Coulomb friction but is very difficult to model.

Mechanical backlash is motion lost when the direction of motion is reversed, as in gear trains and bearings. Since most servoactuators make use of electromagnetic elements, mag-

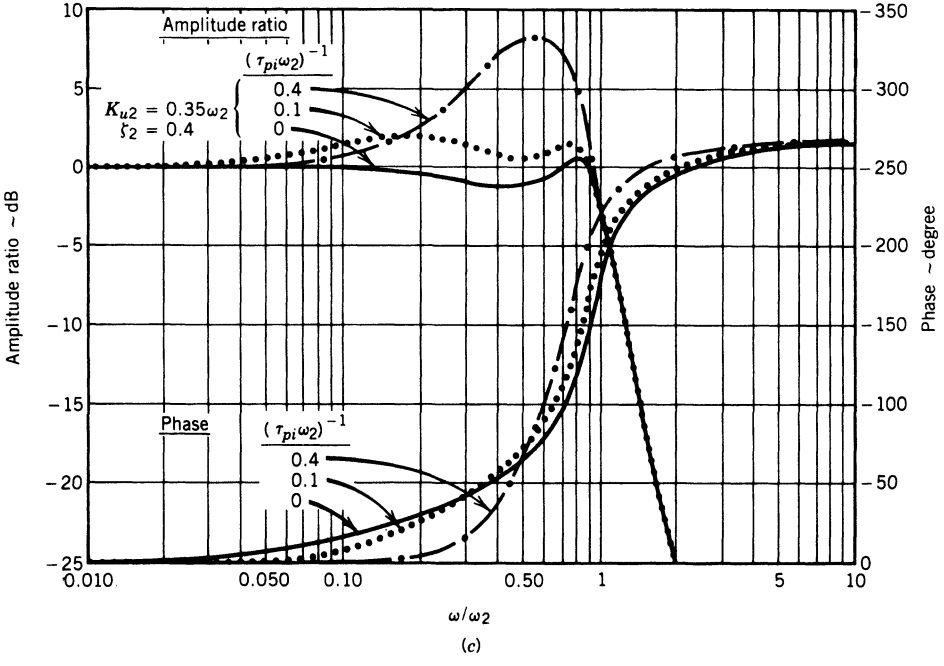


Figure 37 (*Continued*) (*c*) Closed-loop frequency response.

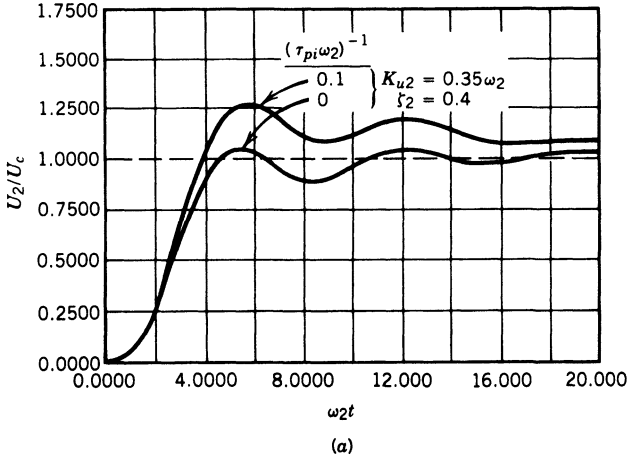


Figure 38 PI controller added to the system of Fig. 6*b*—time histories: (*a*) step response.

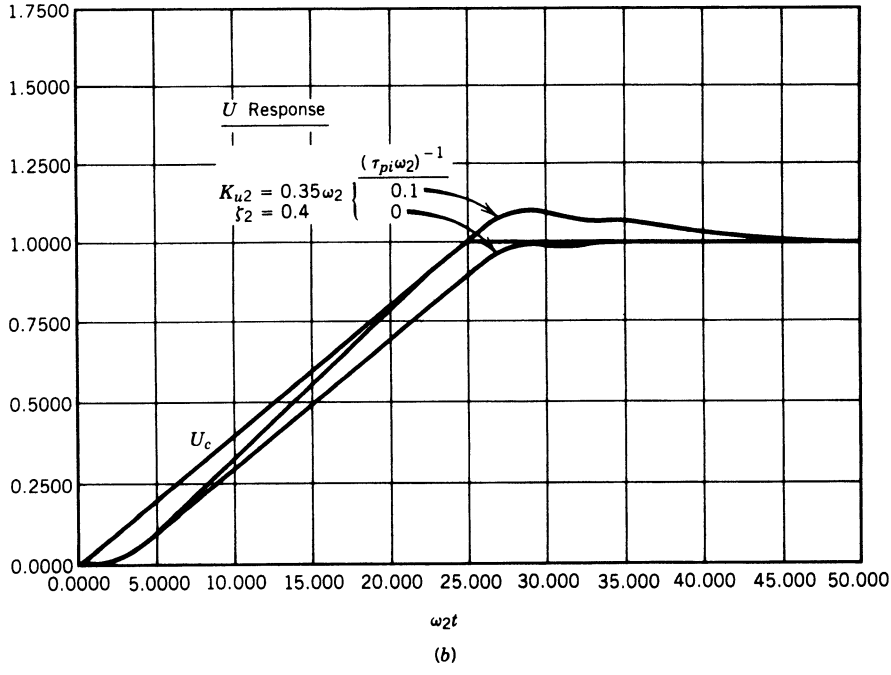


Figure 38 (*Continued*) (*b*) Ramp response.

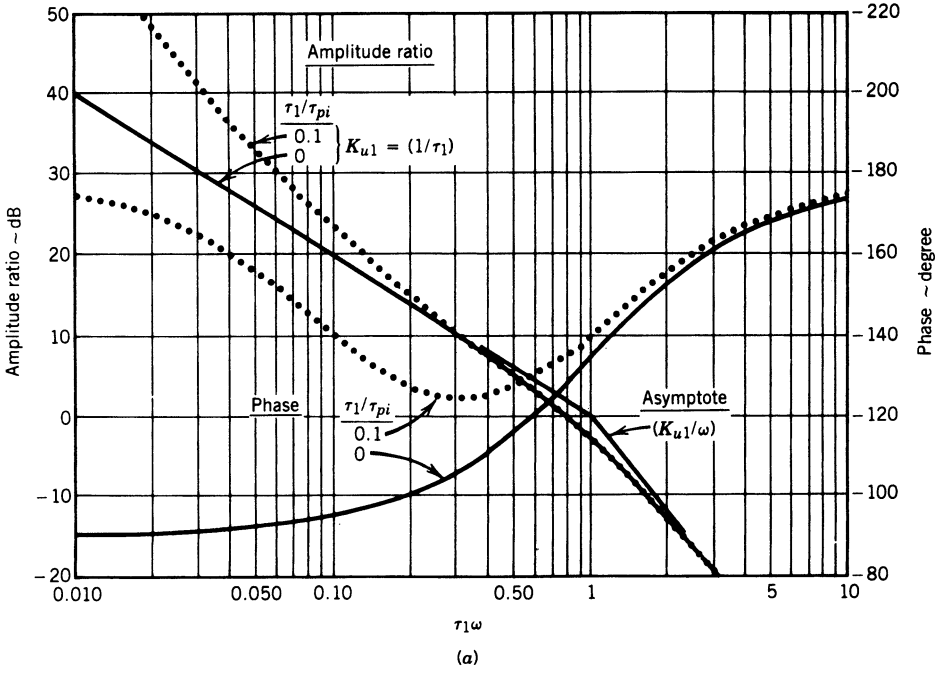
netic hysteresis effects can cause some system performance anomalies. The width of the hysteresis band is dependent upon the amplitude of the input signal (the output is a function of the input's prior history as well as its present value).

7.1 Simple Nonlinearities

When the output of a nonlinear element depends only upon the present value of its input, the element can often be described by a simple relationship between input and output amplitude. If this function is single valued, it is often possible to assess its effect on the system by using linear approximations. One useful technique is to examine the small-perturbation behavior of the system at a series of operating points along the input–output curve by performing a linear analysis using the local slope at each operating point.

Another technique is describing function analysis, which is useful in estimating the response of nonlinear systems to sinusoidal inputs. In general, a describing function is an amplitude-dependent, frequency-dependent transfer function of a nonlinear element which allows the system to be analyzed by conventional frequency-domain techniques. It is derived from a Fourier analysis of the output of the nonlinear element to a sinusoidal input.[1] For simple nonlinearities that can be described by a single-valued output amplitude versus input amplitude, the describing function is a simple gain that varies with input amplitude. In concept, this gain is the average slope of the input–output curve for the particular input amplitude being considered.

Saturation and deadzone are two of the most common nonlinearities encountered in control of mechanical systems. Referring to Fig. 41*a*, an operating-point analysis would



(*a*)

Figure 39 PI controlled added to the system of Fig. 6*a*: (*a*) open-loop frequency response.

idealize the nonlinearity as a simple gain when no saturation takes place and as zero gain when fully in the saturation region. A similar rationale can be applied to Fig. 41*b*. Generally speaking, linear techniques can be used to ensure system stability by analyzing the system with a range of gains determined by the minimum and maximum slopes of the nonlinear amplitude curve. Small-perturbation step response and frequency response of the system around an operating point can also be determined by linear analysis.

Describing-function analysis can provide useful insight into the behavior of the combined deadzone–saturation nonlinearity shown in Fig. 42. Its describing function is a gain that is zero in the deadzone region, increases to a maximum as the input amplitude approaches saturation, then decreases again as the input pushes well into the saturation region. If linear analysis predicts instability at the maximum value of the nonlinear gain, this type of describing function will result in a sustained oscillation at an amplitude corresponding to maximum gain. This behavior is called a stable limit cycle because any tendency of the oscillation to diverge will result in lower gain, which will reduce the tendency to oscillate.

Stable limit cycles can also result from deadzone in a system that is marginally stable at low gains. For example, Section 6.3 explains that a PI compensator used in a system having an inherent integration can exhibit 180° of phase lag at low gains, become stable at intermediate gains, then become unstable at high gains. In this case, a low-frequency oscillation can develop whose amplitude will grow until the describing-function gain is high enough to produce a stable limit cycle.

The effects of saturation and deadzone on system stability are generally straightforward to analyze by operating-point analysis or describing functions, as long as the system consists of single control loops. However, when multiple feedback and feedforward loops are present, the linearized analysis must be performed very carefully. For example, when an inner feed-

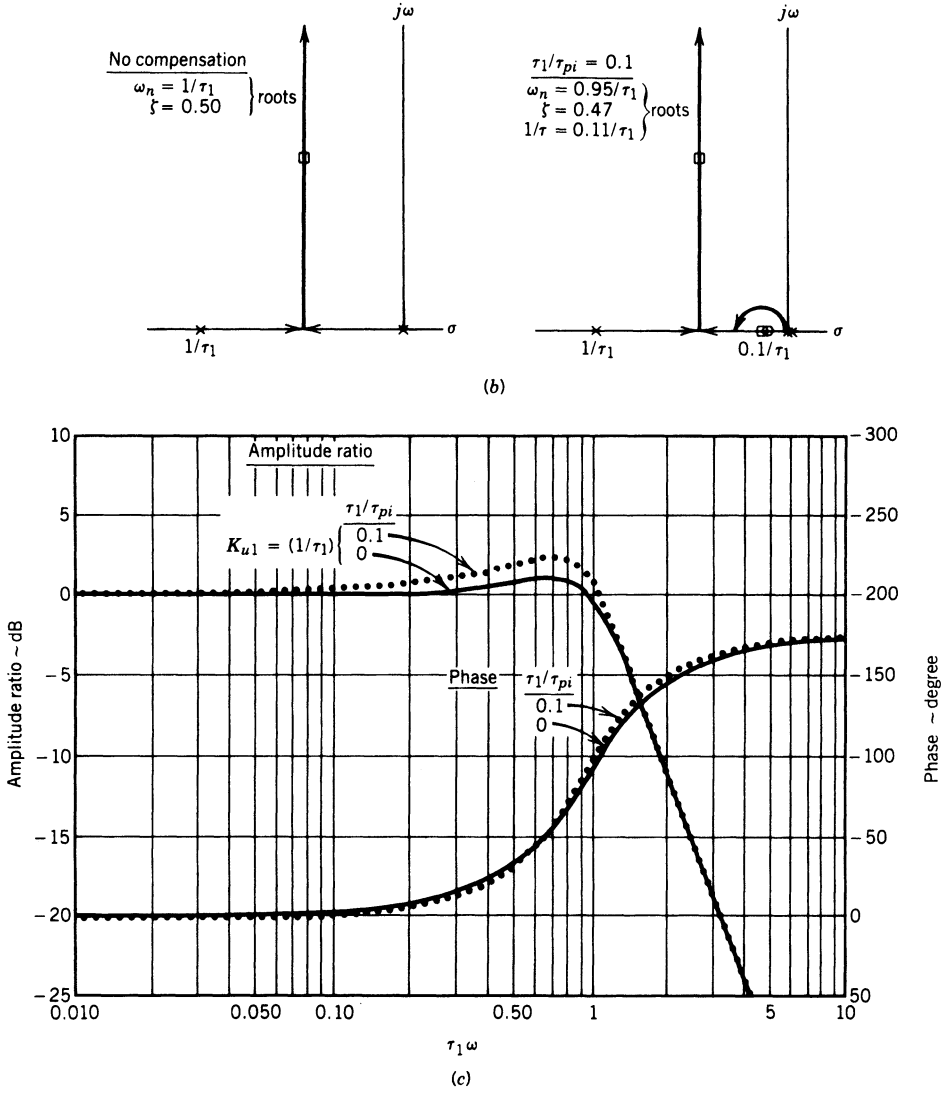
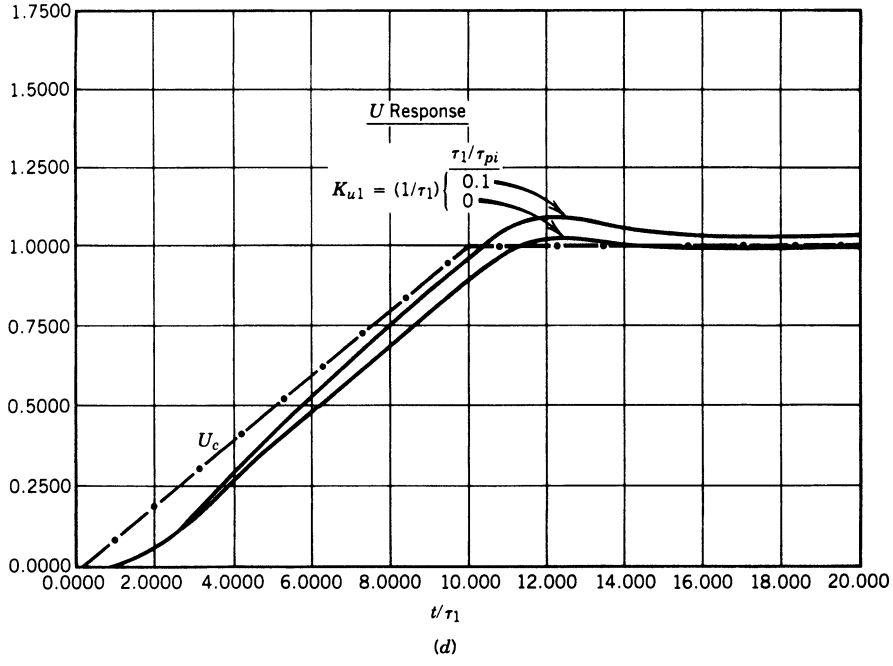


Figure 39 (*Continued*) (*b*) Root loci ($K_{u1} = 1/\tau_1$); (*c*) closed-loop frequency response.

back loop is used to damp an open-loop resonant mode so that higher gains can be achieved in the outer feedback loop, hard saturation or deadzone in the inner feedback path can cause the outer loop to become unstable. Similarly, saturation in a feedforward path or in a lead network can greatly reduce the stabilizing effects they were designed to provide. Another type of saturation is acceleration limiting. Even if it does not create any stability problems, acceleration limiting can cause large overshoots when a position servo decelerates into final position following a large step command.

The resolution of digital systems and certain transducers creates its own set of problems. As shown in Fig. 41*c*, resolution nonlinearities can be described as alternating regions of zero gain and infinite gain. Therefore, resolution will cause most feedback systems to exhibit



(*d*)

Figure 39 (*Continued*) (*d*) Ramp response.

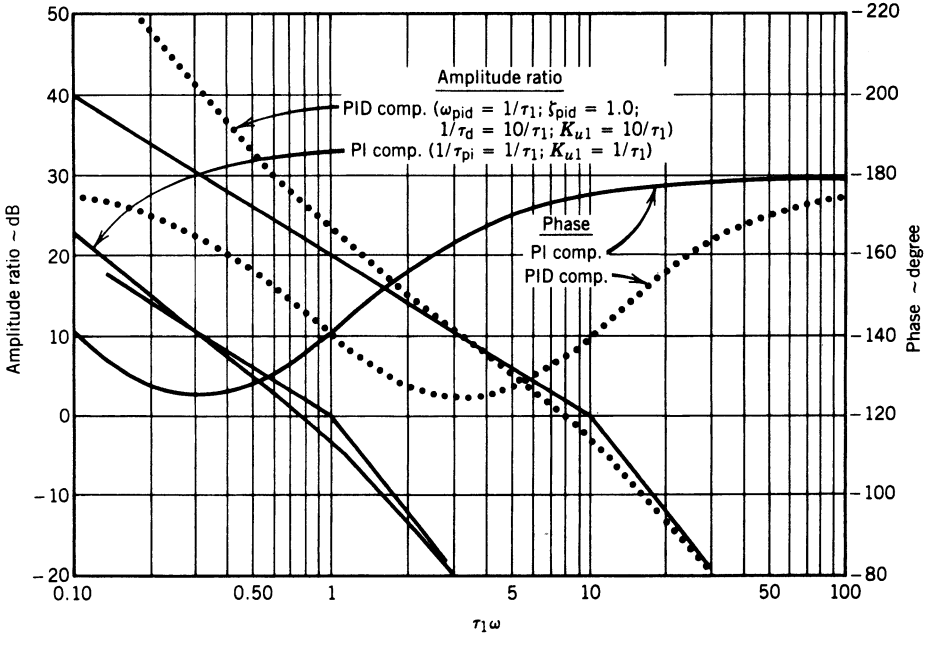


Figure 40 Effects of lead compensation added to the system of Fig. 39.

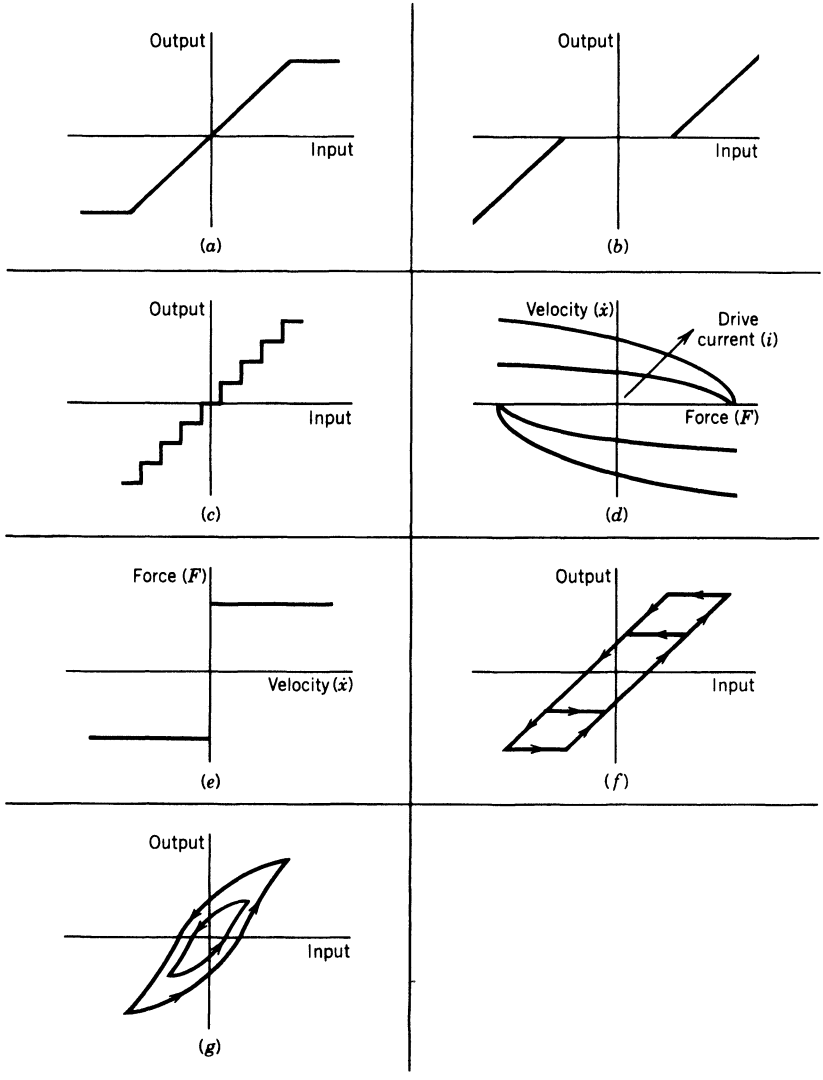


Figure 41 Common nonlinearities: (*a*) saturation; (*b*) deadzone or threshold; (*c*) resolution; (*d*) load–velocity curves; (*e*) Coulomb friction; (*f*) mechanical backlash; (*g*) magnetic hysteresis.

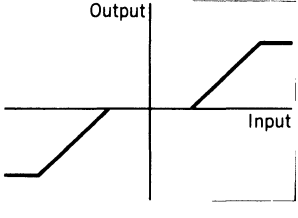


Figure 42 Combined deadzone and saturation.

continuous stable limit cycles with an amplitude corresponding to the resolution increment (least significant bit in a digital system). In a high-resolution system, the magnitude of this limit cycle may be so small that it is not noticeable. Similarly, Coulomb friction (Fig. 41*e*) exhibits infinite gain around zero and a saturating behavior as amplitude increases. In this case, however, the nonlinearity is usually a feedback loop around a mechanical load. When the load is primarily inertial and has no backlash, friction may actually improve system stability rather than decrease it. Of course, friction will also decrease system accuracy.

7.2 Complex Nonlinearities

As nonlinear elements become more complex, linear analysis becomes more complicated and less realistic. However, linear techniques may still be of some use in estimating system stability. For example, a servoactuator's output velocity may be a nonlinear function of output force as well as input drive current. For the example of Fig. 41*d*, system stability can be explored by using linearized characteristics at selected operating points:

$$\Delta\dot{X} = \left(\frac{\partial \dot{X}}{\partial i}\right) \Delta i + \left(\frac{\partial \dot{X}}{\partial F}\right) \Delta F \tag{39}$$

The two derivatives in Eq. (39) can be used in a conventional linear model showing velocity as a function of current input and load force feedback (Fig. 43). Note that the derivative of velocity with respect to force is negative in this case.

Some nonlinear elements such as hysteresis and backlash cannot be approximated by a simple relationship between input and output amplitude. Instead, the output depends upon the history of the input as well as its present value. The describing functions of such elements are typically frequency dependent as well as amplitude dependent.[1] Describing-function analysis with such nonlinearities can become rather complicated and is beyond the scope of this chapter. Also, it can be argued that computer simulation yields more realistic results without much additional effort. This is particularly true if the control system has multiple nonlinearities that are significant.

7.3 Computer Simulation

Unless the control system is extremely complex or highly nonlinear, the use of a simplified linear model is usually the best way to synthesize the basic function of the control system and to perform preliminary performance estimates. The linear techniques described in the body of this chapter are typically faster than simulation, are less prone to major errors, and promote physical understanding of the system's behavior. It is true that simplifying assumptions must be made very carefully, but this process also promotes improved understanding of the system. With the basic system function defined, simulation can then be used to evaluate

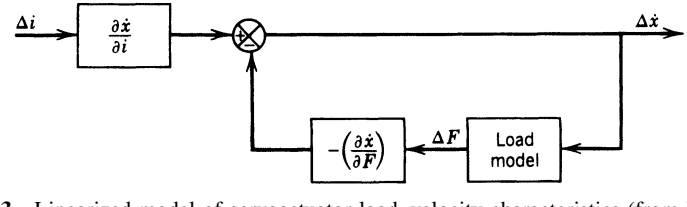


Figure 43 Linearized model of servoactuator load–velocity characteristics (from Fig. 41*d*).

the simplifying assumptions of the linear analysis, ''fine tune'' the system design, and generate detailed performance data over a wide range of operating conditions.

Early computer simulation was accomplished using analog computers, which offered unlimited opportunity for online operator interaction and real-time operation. Real-time capability meant that there was no waiting for data, and system development work could be accomplished using a combination of simulation and real hardware (''hardware in the loop''). Unfortunately, considerable setup time was required because of the need for patching and scaling, and there were severe restrictions on the size of the simulation that could be handled. These limitations have caused analog computers to become virtually extinct. In their place, a variety of real-time digital and hybrid simulation tools have been developed. These typically offer convenient programming, fast run times, the ability to handle large simulations, and hardware-in-the-loop capability. Their disadvantages include the high cost and maintenance problems associated with specialized computer hardware, inability to time share, inability to use for other kinds of engineering problems, and obsolescence.

To avoid the problems associated with specialized computer hardware, general-purpose simulation software is now readily available which can be run on modern PCs and workstations. If used properly, it can provide realistic results for very complicated, highly nonlinear systems. Programming is typically accomplished quickly and easily. However, such programs can have several drawbacks, such as limited ability for online operator interaction and excessive time required to generate output data, especially frequency responses. Furthermore, most general-purpose programs are carefully designed to minimize limitations on what can be programmed. As a result, they will happily violate the laws of physics without complaint. As with all computer tools, it is good practice to check out all critical program functions prior to generating data and to spot check the early data against the original design calculations wherever possible. An efficient means to accomplish this process is to start with simplified simulation models, then add complexity in layers, as each previous version is checked out.

8 CONTROLLER IMPLEMENTATION

As mentioned in the introduction, it is not the intent of this chapter to address the electronic design of a servocontroller. Nevertheless, some basic understanding of controller implementation is required to properly specify and select a controller. The following discussion describes several basic implementation approaches, together with their relative advantages and disadvantages.

Since control of a mechanical load is inherently a continuous process, the use of a dc analog controller typically provides high servo bandwidth and smooth operation. Furthermore, basic servoloops can be implemented with very simple circuits. Hard-wired digital controllers offer the potential for increased overall accuracy at the cost of degraded resolution and more complex electronic hardware. Microprocessor-based digital controllers offer increased flexibility, versatility, accuracy, and computer interface capability, but maintaining adequate resolution and sampling rates can create throughput problems. The electronic hardware associated with processor-based controllers is typically more complex than a simple analog controller. However, in complex control applications, large amounts of analog circuitry can be replaced with software.

8.1 Analog Controllers

Analog controllers are typically used in servoloops for which high closed-loop bandwidth and smooth operation are required. Modern operational amplifiers have bandwidths of several

hundred kilohertz and virtually infinite resolution. Because of this, they are free of the sampling delays, phase lags, and resolution problems associated with microprocessor-based controllers. Furthermore, they are relatively tolerant of electrical noise, and troubleshooting can be accomplished with simple equipment. On the other hand, analog implementation of complex controller functions such as nonlinearities, automatic gain changing, elaborate command processing, and complex failure detection can cause the electronic circuitry to become extremely complicated. In addition, analog controllers typically require periodic adjustment and calibration.

The basic functional elements of a typical analog controller are shown in Fig. 44. The functions of prefilters and compensation networks have been discussed previously in this chapter. Some form of signal conditioning is usually required for transducers. In the case of a simple dc transducer such as a potentiometer, the conditioning may consist of dc excitation together with an output buffer amplifier. Buffer amplifiers can be designed to protect against large voltages erroneously connected to the electronics, to provide consistent loading of the transducer output, to reject electrical cabling noise (electromagnetic interference), and to filter transducer ripple. Low-output transducers, such as those that employ strain gage elements, require high-gain, low-drift amplifiers. Linear variable differential transformers (LVDTs), resolvers, and other ac transducers require ac excitation, demodulation, and filtering to remove ripple. Greatly increased servo accuracy can be obtained with a combination of ac command generation and ac feedback, such as when synchros are utilized. However, this approach has largely been replaced by the use of digital controllers and transducers. In any case, transducer specifications should be carefully studied to determine the proper signal conditioner characteristics. The dynamic characteristics of the signal conditioners, as well as the transducers themselves, can have significant impact on the stability and performance of the servoloops, as explained previously in this chapter.

There are many sources of information concerning the design of analog controllers. References 8 and 9 are excellent sources for the design of operational amplifier circuits for a wide variety of purposes, including compensation and signal conditioning. Furthermore, most manufacturers of operational amplifiers publish useful application handbooks. Also, application literature from transducer manufacturers often discusses signal-conditioning techniques in some detail.

The design of power amplifiers varies widely with the type of servoactuator being driven. In the case of conventional electrohydraulic servoactuators and other types requiring low-power electrical inputs, the power amplifier can be a very simple linear (proportional) circuit. Typically, it utilizes an operational amplifier and a power boost stage consisting of a complementary pair of transistors. Some operational amplifiers have enough output capability to provide the required electrical input directly. In the case of electromechanical devices that must provide a direct electrical-to-mechanical energy conversion, such as brushless servomotors, the power amplifier can become very complex. In this case, its design should be left to an experienced electronics engineer. Linear amplifiers are still the most straightforward and offer the best servo performance but are severely limited in the size of the motor they can control because of the large amount of heat they must dissipate. Since switching transistors typically generate little heat in their full-on or full-off states, various time-modulated on–off power drivers have been developed. The most popular type for servocontrol applications seems to be pulse width modulation (PWM). In this approach, the power devices are switched on and off at a very high fixed frequency. The percent on-time during each cycle is proportional to the dc input voltage from the upstream analog controller circuitry. If the PWM frequency is high enough, only the average cycle voltage will affect the servoactuator output, thereby resulting in nearly proportional control. To accomplish this, the PWM frequency must usually be at least one order of magnitude higher than the bandwidth of the

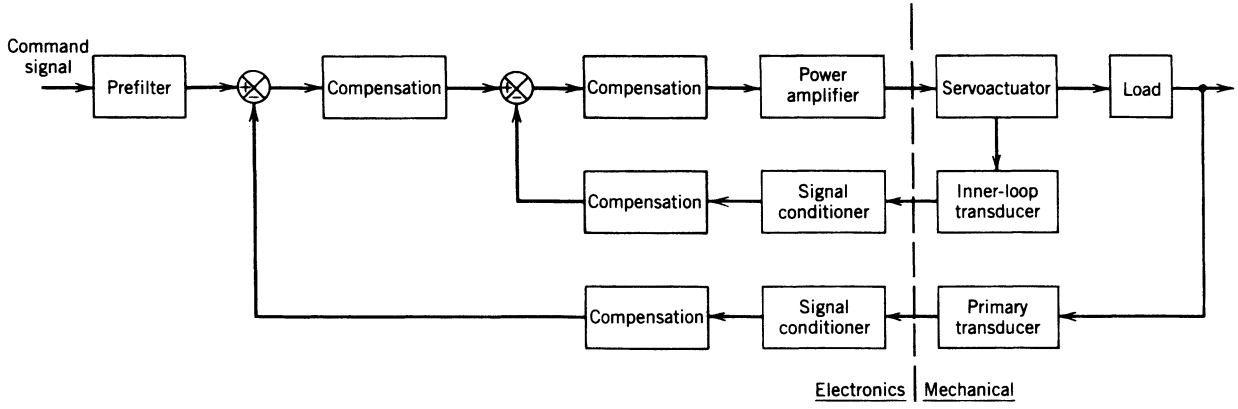


Figure 44 Typical analog controller.

innermost feedback loop (typically a current feedback loop around the power amplifier), or two to three orders of magnitude higher than the bandwidth of the primary control loop. If PWM frequencies of this magnitude are impractical, it may be necessary to reduce bandwidths of the various control loops to prevent unacceptable servo output at the PWM frequency. Design of a PWM amplifier can be difficult, particularly with regard to polarity switching around zero.

It should be noted that servoactuators requiring high-power electrical inputs can create problems related to voltage saturation in the drive electronics. The reason for this is that the input inductive characteristics of many actuators tend to cause long L/R time constants. Current feedback is often used to improve the current response of the actuator, and this leads to large transient voltages for abrupt inputs to the power amplifier. Practical limits on amplifier voltage capability can result in voltage saturation during transients. It is often necessary to design an active voltage-limiting circuit to prevent stability problems during saturation. Careful attention must be devoted to tailoring the current feedback loop and voltage limiter circuits to the servoactuator's electrical dynamics, if amplifier stability and performance problems are to be avoided.

Multiturn potentiometers are usually included in analog controller circuitry to allow proper calibration of transducer scale factors, compensate for electrical and mechanical offsets, and adjust loop gains. The desire for system accuracy often drives the designer to providing many adjustments, but this practice can greatly complicate maintenance procedures. If adequate performance cannot be achieved with a limited number of well-placed adjustments, then serious consideration should be given to the use of a digital controller.

The use of integrating amplifiers in control loops requires some special consideration. First, pure integrators often cause low-frequency oscillations or ''hunting'' when backlash, deadzone, or friction exists in the system. This behavior can often be controlled by adding a large resistance across the integrator's feedback capacitor. This limits the amplifier's gain and makes it look proportional at low frequencies, while preserving an integrating characteristic in the crossover frequency range. If it is possible for an integrating amplifier to saturate during abrupt commands, it may ''latch up'' and exhibit a long recovery period, which can result in large servo overshoots. This behavior can be prevented by proper gain distribution in the servoloops or by providing the amplifier with a diode limiter. Of course, an integrating amplifier can drift into saturation if it is powered up before the servoactuator is allowed to move. For example, electronics are often powered up prior to releasing a mechanical brake or applying hydraulic power. In this case, integrator saturation can cause a large engagement transient. This can be prevented by shorting the integrator's feedback capacitor with a relay contact or electronic switch. The short is then opened when the actuator is mechanically or hydraulically engaged.

8.2 Hard-Wired Digital Controllers

The overall accuracy of a servomechanism can be greatly enhanced by the use of a digital transducer and a digital controller. Furthermore, the need for periodic calibration and adjustment can be virtually eliminated. Hard-wired digital electronics can provide this improved accuracy with bandwidths comparable to analog electronics. However, these digital circuits are considerably more complex than comparable analog circuits, are more susceptible to electrical noise, and have finite resolution. For these reasons, hard-wired digital electronics are usually used only in the primary control loop (accuracy is typically not critical in the inner loops). Furthermore, frequency compensation is difficult to implement in digital hardware and is usually left to analog circuitry.

Hard-wired digital electronics are commonly used with high-resolution incremental encoders, as illustrated in Fig. 45*a*. Two pulse trains are generated by the encoder 90° out of phase with one another. Pulse-conditioning circuitry squares up the incoming pulses, determines the transducer's direction of motion, and often increases the resolution by a factor of 4. The asynchronous counter is incremented up or down by each feedback pulse, depending upon the direction of motion. Similarly, command pulses also increment the counter up or down. The net count at any particular time represents the difference between the number of command and feedback pulses since the counter was initialized. After digital-to-analog conversion, this count becomes the error signal transmitted to the analog electronics.

The command pulse train can be generated by additional digital hardware or by a computer. However, if the transducer resolution is high and the desired maximum command rate is high, a computer may be hard pressed to provide the required pulse rates. In this case, hardware comparators and rate multipliers may prove more satisfactory. It should be noted that an incremental system has no inherent knowledge of its absolute position. Therefore, power shutdowns and electrical noise can cause such a system to lose track of where it is. For this reason a ''marker pulse'' is often provided at some known position to reinitialize the counter periodically. Alternatively, the servo occasionally can be commanded to a mechanical ''home'' position. As with most transducers having finite resolution, encoders will usually cause limit cycling with an amplitude equal to the least significant bit (Section 7.1).

Absolute digital systems can also be implemented in electronic hardware, as shown in Fig. 45*b*. The encoding transducer outputs a digital word that represents its absolute position at all times. For an optical encoder, the resolution is typically between 12 and 24 bits. After buffering, the feedback word is digitally subtracted from a digital command, and the resulting error is converted to an analog signal that is transmitted to the analog electronics. The digital summing junction can be implemented in a number of ways, including the use of an arithmetic logic unit (ALU), which operates at very high speeds. The digital command and feedback information can be transmitted to the summing junction as parallel digital words or as serial data that must be multiplexed and then decoded. The command information can be generated from a computer or digital thumbwheels.

The absolute system is less susceptible to loss of position information than the incremental system, but the transducers are considerably more expensive and less reliable, and a wire is required for each bit. To improve reliability and reduce cost, resolvers or other sine/cosine output devices are often used, together with a resolver-to-digital (R/D) converter. The penalty is reduced overall accuracy, although this can be improved by using coarse/fine resolvers and appropriate additional hardware logic.

8.3 Computer-Based Digital Controllers

The hard-wired digital controllers of Fig. 45 have limited flexibility and functional capability. The use of a microprocessor may reduce electronic hardware complexity when elaborate system functions are required. There are many such functions that are well suited to microprocessor implementation:

- Command processing (nonlinear functions, limiting, switching, and communication with other computers)
- Redundancy management (fault detection, isolation, and reconfiguration)
- Adaptive control (self-adjustment of control loop parameters as operating conditions or environmental factors change)
- Built-in test (BIT) features

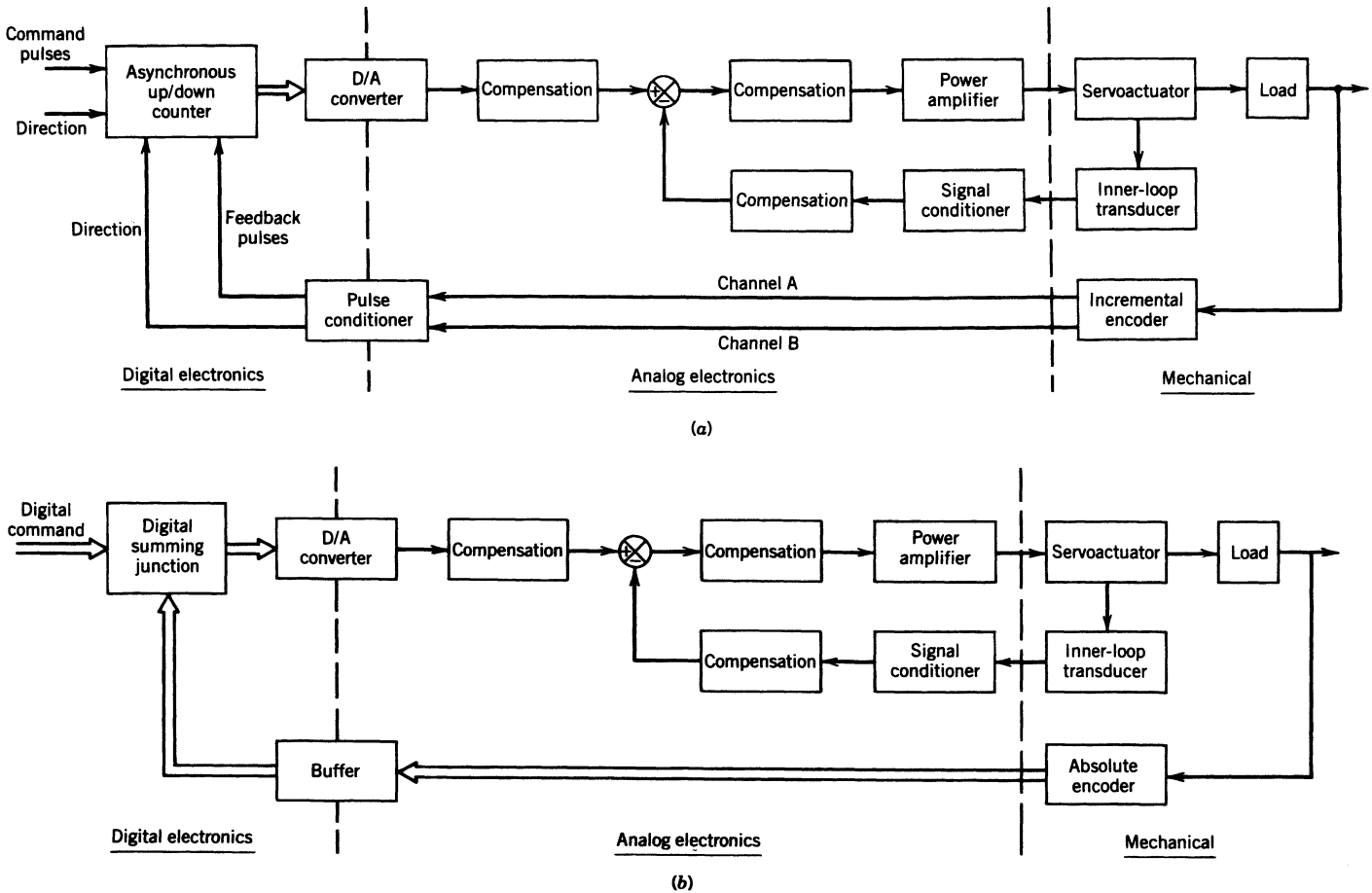


Figure 45 Typical hard-wired digital controller: (*a*) incremental system; (*b*) absolute system.

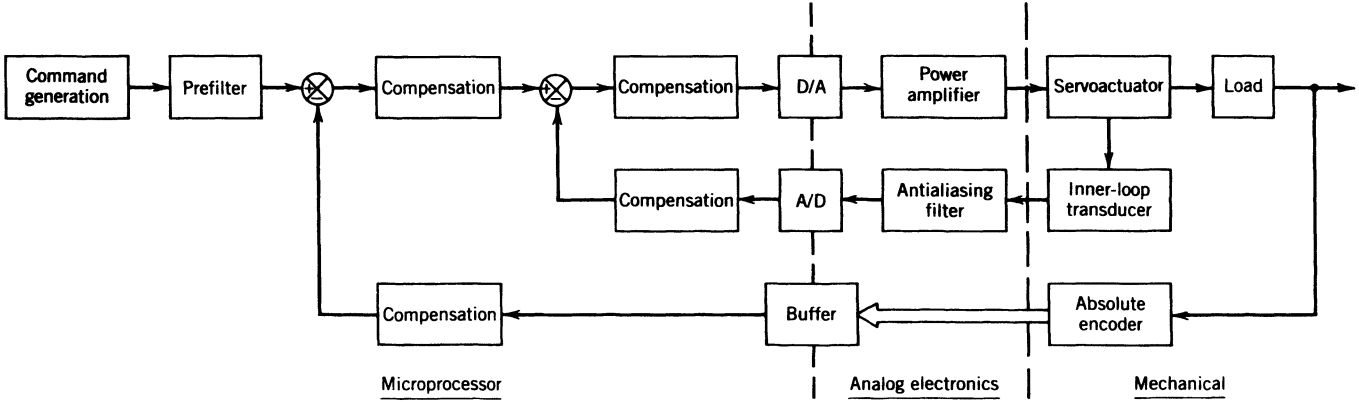


Figure 46 Typical computer-based controller.

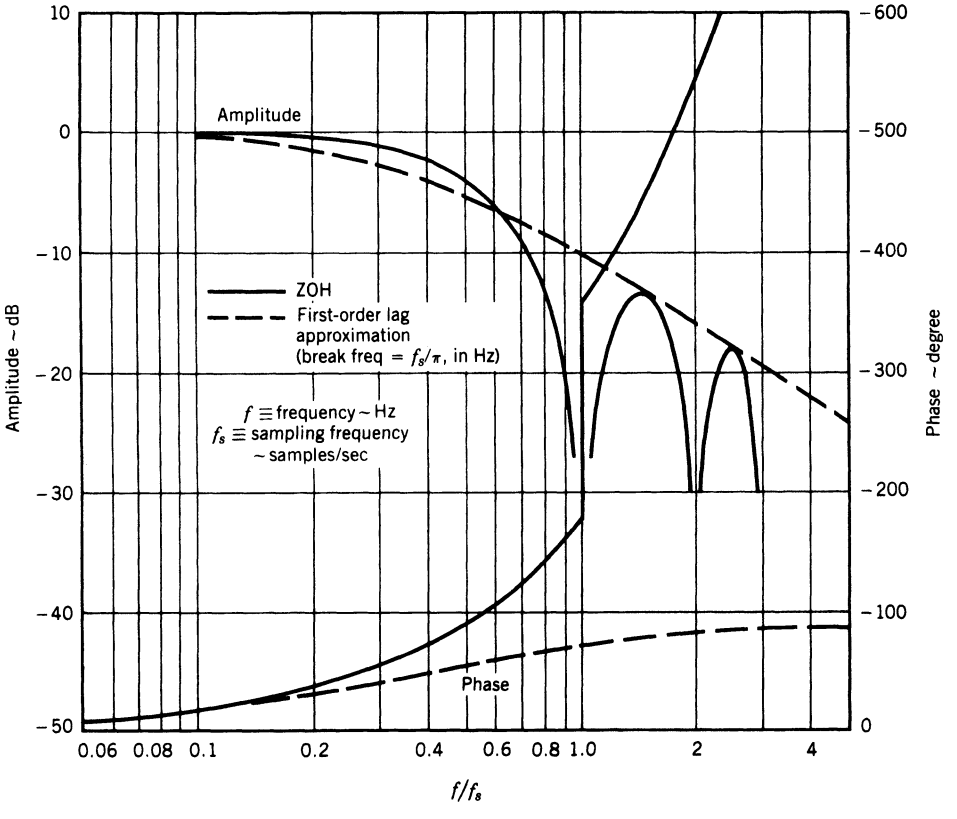


Figure 47 Frequency response characteristics of a zero-order hold.

Once the need for a microprocessor is established, it may also become reasonable to implement the servoloops in software. Note that if the application requires only the closure of simple servoloops without the need for elaborate additional functions, the use of a microprocessor will usually result in more complex hardware than an all-analog system and may be more complex than a hard-wired digital system.

The loop closure architecture of a microprocessor-based controller is illustrated in Fig. 46. This block diagram implements virtually all the loop functions in software, including frequency compensation. Of course, it is not necessary to use a digital outer loop transducer, but use of an analog transducer limits the potential accuracy advantages of the digital controller. Several methods can be used to generate the software compensator designs. Perhaps the most straightforward technique is to construct Laplace transfer functions using the continuous frequency-domain techniques outlined in Sections 1–6. These transfer functions can then be converted to equivalent *Z*-transforms, from which difference equations can be generated for implementation in software.[7]

It should be noted that a microprocessor-based controller is a sampled-data system, and the digital-to-analog converter usually operates as a zero-order hold (ZOH). The sampling nature of the system, together with computation times, introduces time delays into the control loops which can have a profound influence on system performance and stability. Figure 47 shows a frequency response of a ZOH operated in a sampled-data system. If the system has

been designed using frequency-domain techniques, this figure represents the additional phase lag and amplitude that will result from computer implementation of the design. The phase lag is linear with frequency, and an approximate transfer function is a first-order lag with a break frequency equal to f_s/π, where f_s is the sampling frequency of the computer, in hertz. This approximation is very accurate out to a frequency of $f_s/2\pi$. Note that 10° of phase lag exist at a frequency of $f_s/18$. This suggests that the sampling frequency should be at least 20 times the crossover frequency (in hertz) of the loop being implemented if the impact on phase margin is to be minimized. Furthermore, smooth operation of the servo may require heavy filtering at the output of the digital-to-analog converter to reduce sampling-induced ripple. This will add additional phase lag in the servoloops.

The need to minimize phase lags can place severe restrictions on the complexity of computations that the microprocessor can handle in one sampling interval. This problem can be partially overcome by using a separate processor to perform loop closure computations or by implementing compensators in the analog circuitry. Alternatively, the need for high sampling rates can be reduced by implementing high-gain inner feedback loops in analog circuitry, as shown in Fig. 45*b*. If the inner loops utilize analog transducers, the problem of aliasing[4,5,10] adds another reason for using analog electronics. To properly utilize the output of an analog transducer in the computer, an antialiasing filter is required at the input to the analog-to-digital converter. These filters are often first order with a break frequency equal to f_s/π. This doubles the effective phase lag of the computer.

REFERENCES

1. J. J. D'Azzo, and C. H. Houpis, *Feedback Control System Analysis and Synthesis,* McGraw-Hill, New York, 1966.

2. B. C. Kuo, *Automatic Control Systems,* Prentice-Hall, Englewood Cliffs, NJ, 1982.

3. E. O. Doebelin, *Dynamic Analysis and Feedback Control,* McGraw-Hill, New York, 1962.

4. G. F. Franklin, J. D. Powell, and M. L. Workman, *Digital Control of Dynamic Systems,* Addison-Wesley, Reading, MA, 1990.

5. K. J. Astrom and B. Wittenmark, *Computer-Controlled Systems,* Prentice-Hall, Englewood Cliffs, NJ, 1990, 1997.

6. B. C. Kuo, *Digital Control Systems,* Holt, Rinehart and Winston, New York, 1980.

7. J. A. Cadzow, and H. R. Martens, *Discrete-Time and Computer Control Systems,* Prentice-Hall, Englewood Cliffs, NJ, 1970.

8. J. G. Graeme, G. E. Tobey, and L. P. Huelsman, *Operational Amplifiers, Design and Applications,* McGraw-Hill, New York, 1971.

9. J. G. Graeme, *Amplifier Applications,* McGraw-Hill, New York, 1999.

10. E. O. Doebelin, *System Modeling and Response,* Wiley, New York, 1980.