VOLUME II

FLYING QUALITIES PHASE

CHAPTER 13
FEEDBACK CONTROL THEORY

JANUARY 1988

USAF TEST PILOT SCHOOL
EDWARDS AFB CA

13.1 FUNDAMENTALS OF FEEDBACK CONTROL THEORY

Pilots might be inclined to associate the phrase "control system" with only aircraft flight control systems. Although the control system theory of this course has a large application to flight control systems, this material applies to any process or system in which control is exercised over some output variable. Examples of these controlled variables are: the speed of an automobile, the temperature of a room, the attitude of a spacecraft, ad infinitum.

Feedback control system theory is often called several different things. It might be found under any of the following headings or titles: Control Systems, Automatic Control Systems, Servo-Mechanisms, or our term, Feedback Control Systems.

First, the difference between "open-loop" and "closed-loop" control will be discussed. Consider the roll channel of an aircraft flight control system in which the pilot input is assumed to be a rate command. That is, the pilot commands a roll rate ($\dot{\phi}$) proportional to the stick displacement. Figure 13.1 shows a diagram of this system. The input is a low power input representing a selected value of roll rate.

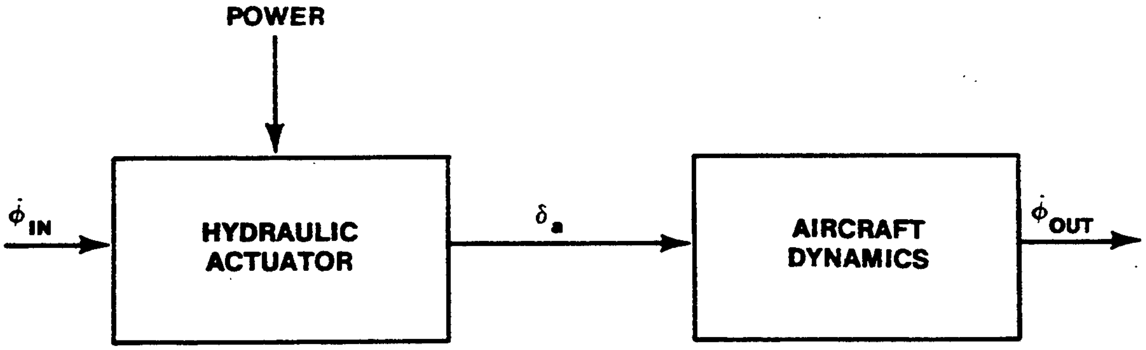


FIGURE 13.1 OPEN-LOOP CONTROL SYSTEM

This mechanical signal is then amplified in a hydraulic control valve/actuator combination to position the ailerons accordingly (δ_a). The deflected ailerons then react with the airstream to produce a roll moment in the required direction. The magnitude of the resultant roll rate ($\dot{\phi}_{out}$) is a function primarily of the dynamic pressure (q) and the moment of inertia about the longitudinal axis (I_x). Both are part of aircraft dynamics.

There are innumerable examples of the "non-feedback" or "open-loop" type of control systems. For instance, a gasoline engine in an automobile has a low power input, the throttle position, which controls the speed of the vehicle expending a large amount of power. In a simple electronic amplifier a very small input signal controls a much larger output signal. In all open-loop control systems the output has no influence on the input whatsoever. The input quantity controls the output only directly through the intermediate components. Referring back to the aircraft roll control system, a lateral stick displacement of a specified amount will not command a constant roll rate under all conditions. As the conditions in the intermediate components change, such as the dynamic pressure, moment of inertia, hydraulic pressure, temperature of hydraulic fluid, condition of hydraulic components, temperature effects on modulus of elasticity of metal components, etc., the resultant roll rate for a specified input will vary.

The performance of any control system with respect to maintaining the output quantity as close as possible to the input quantity can be substantially improved by feeding back the output for comparison with the input. The use of the difference resulting from this comparison as an actuating signal constitutes a feedback or closed-loop control system.

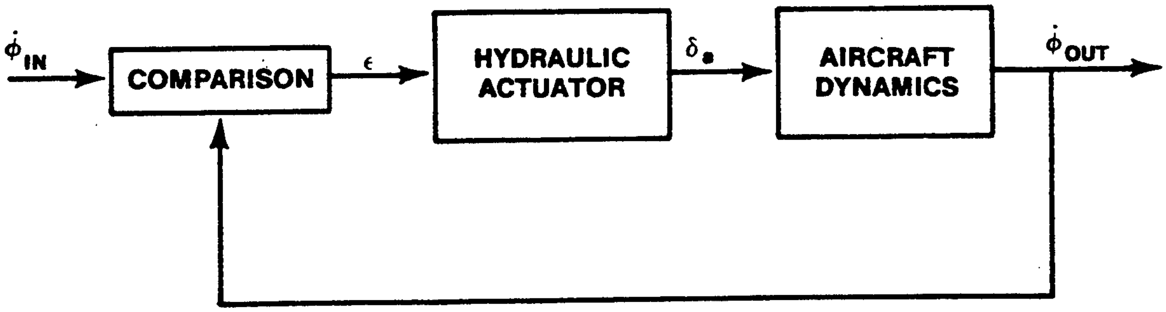


FIGURE 13.2. CLOSED-LOOP CONTROL SYSTEM

Figure 13.2 shows how the open-loop system of Figure 13.1 can be changed to a closed-loop system by the addition of an outer feedback loop to compare the input with the output. Thus, the effect of variations in the intermediate components can be eliminated in that a corrective signal (ε) will continue to exist until the output properly matches the input.

A serious disadvantage of closed-loop control systems, however, is that they can make an otherwise stable system unstable. The possibility of instability is the prime reason for the existence of the science of feedback control system analysis. The first and major effort in control system analysis is the determination of whether or not the closed-loop system is stable. After this fact is established, other response characteristics may be found.

Stability, with respect to control systems is defined as follows: A stable system is a system in which the transients die out with increasing time.

13.2 NOMENCLATURE

The following nomenclature is used in this chapter:

R = input variable

C = output variable

Each of these might represent any quantity depending on the system such as angular or linear position, current, voltage, degrees of temperature, etc, or the time rate of change of those above.

The following symbol represents a summer or differential. It indicates the algebraic summation of the input quantities according to the arrows and

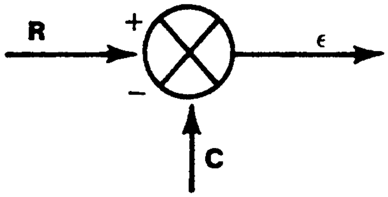


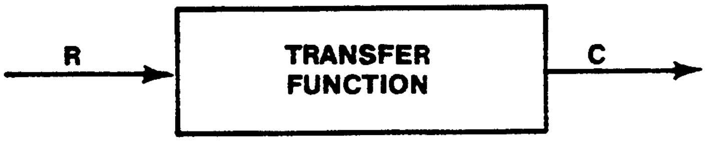
FIGURE 13.3. SUMMER OR DIFFERENTIAL

the signs. The example in Figure 13.3 shows

$$\varepsilon = R - C \tag{13.1}$$

The symbol for gain or amplification factor is K.

Control systems are generally described through the use of block diagrams as in Figures 13.1 and 13.2. However, instead of words to indicate the process or operation occurring within each block, there appears what is called a transfer function (Figure 13.4). The term "transfer function" might be thought of as what is done to the input to produce the output. Although the transfer functions within the blocks are generally written in terms of some operator notation, they are often described graphically, especially for nonlinear systems. A definition of transfer function is: the ratio of the output to the input expressed in Laplace operator notation, assuming zero initial conditions. The transfer function is essentially a mathematical model of the system and embodies all the physical characteristics of the system i.e., mass, damping, etc.



$$C = \left[\begin{array}{c}\text{TRANSFER FUNCTION}\\ \text{(IN OPERATOR}\\ \text{NOTATION)}\end{array} \right] \times R$$

$$\frac{C}{R} = \left[\begin{array}{c}\text{TRANSFER FUNCTION}\\ \text{(IN OPERATOR}\\ \text{NOTATION)}\end{array} \right]$$

FIGURE 13.4. TRANSFER FUNCTION

13.3 DIFFERENTIAL EQUATIONS - CLASSICAL SOLUTIONS

Differential equations for a control system will illustrate the types of responses to be expected from first and second-order systems. These two examples are used throughout this course because higher-order systems produce a transient response consisting of the sum of first and second-order responses.

The reason the transient response is significant rather than the steady state or complete response, concerns the stability of the system. Since positive stability requires that the transients die away with increasing time, the transient solution of the differential equation describing the system is most important to the analysis. The transient solution also provides other important response characteristics.

Figure 13.5 shows a simplified block diagram of VTOL Auto Pitch Control with inertia, I_y, and limited aero-damping proportional to pitch rate. Pitch attitude is maintained by reaction control jets. These jets produce a torque proportional to a valve position. Torque $= \mu\varepsilon$, where μ is the gain of the valve and e the input to the valve. The loop is closed by comparing the output pitch attitude to the commanded pitch attitude.

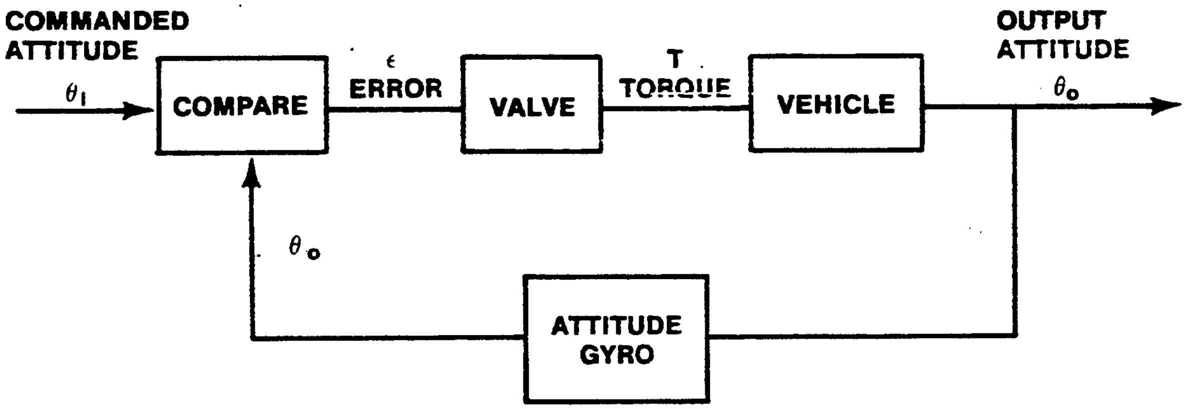


FIGURE 13.5 VTOL AUTOMATIC PITCH CONTROL BLOCK DIAGRAM

This comparison produces the error signal, ε, which is the input to the linear valve. The resulting torque, $\mu\varepsilon$, is applied to the vehicle to change the pitch attitude.

Two situations will be considered in order to simplify the problem. In the first case, only the effect of viscous damping will be considered. This will result in a first-order differential equation. The second case will include both inertia and viscous friction and will result in a second-order system. The first and second-order differential equations will be solved for the transient response.

13.3.1 First-Order System

Using Figure 13.5, including only the effect of damping on the vehicle, the differential equation of the system can be written by equating the applied torque to the absorbed torque. The torque applied by the reaction jets is absorbed by the viscous friction (aero-damping) of the vehicle.

$$\mu\varepsilon = b\dot{\theta}_0$$

The output of the comparator, $\varepsilon = \theta_i - \theta_0$ that produces the system differential equation or the equation of motion

$$\mu(\theta_i - \theta_0) = b\dot{\theta}_0$$

Applied Torque = Absorbed Torque

Using the operator "p" notation (where $p = d/dt$) to determine the system transient response, the homogeneous equation becomes

$$\frac{b}{\mu} p\, \theta_0 + \theta_0 = 0$$

$$\left(\frac{b}{\mu} p + 1\right) \theta_0 = 0$$

the root of the characteristic equation is

$$p = -\frac{\mu}{b}$$

Since the transient response is assumed to be

$$\theta_0(t)_{\text{transient}} = \sum_{i=1}^{n} C_i e^{p_i t} \qquad (13.2)$$

where p_i 's are the roots of the characteristic equation, the transient response for our first-order system is

$$\theta_0(t)_{\text{transient}} = Ce^{-(\mu/b)t} \qquad (13.3)$$

For positive gain μ and damping factor b, $\theta_0(t)$is always stable. Thus, a first-order system has only real roots of the characteristic equation and the transient response is either an exponential increase or decrease depending on the sign on the time constant. The time constant is the reciprocal of the coefficient of t in the exponent of e (b/μ in our case).

The time constant, generally given the symbol τ, can be defined as the value of time that makes the exponent of e equal to -1. In one time constant the exponential $e^{-t/\tau}$ has decreased from the value 1 to the value .368. Figure 13.6 shows a plot of the transient response of a first-order stable system. Time constants are discussed in more detail in Paragraph 13.5.4.

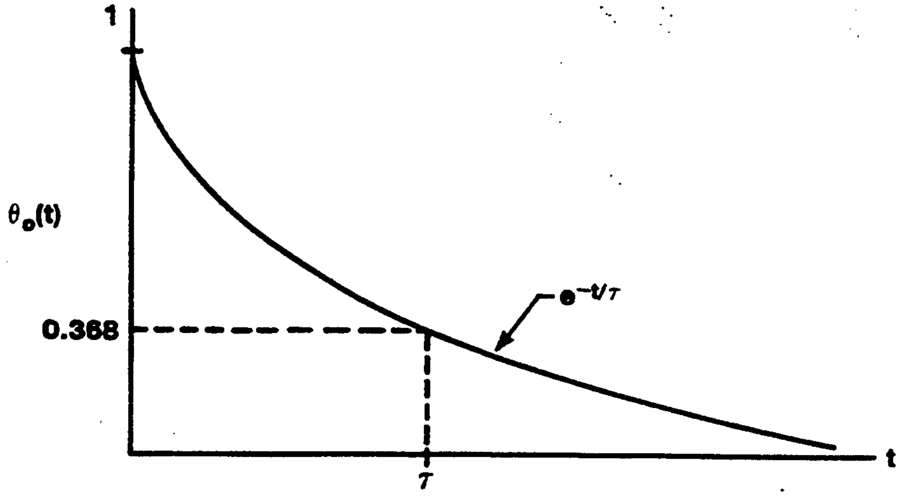


FIGURE 13.6. PLOT OF FIRST ORDER TRANSIENT RESPONSE

A stable transient response requires the root of the characteristic equation be negative.

13.3.2 Second-Order System

Equating applied torque to absorbed torque of Figure 13.5 again, but including the inertia and damping effects yields

$$\mu\varepsilon = b\dot{\theta}_0 + I\ddot{\theta}_0$$

again

$$\varepsilon = \theta_i - \theta_0$$

and

$$\mu(\theta_i - \theta_0) = b\dot{\theta}_0 + I\,\ddot{\theta}_0$$

$$I\ddot{\theta}_0 + b\dot{\theta}_0 + \mu\theta_0 = \mu\theta_i$$

is the equation of motion of the system. The homogeneous equation in operator notation is

$$\theta_0(Ip^2 + bp + \mu) = 0$$

The characteristic equation is

$$Ip^2 + bp + \mu = 0 \qquad (13.4)$$

whose roots are

$$p_{1,2} = \frac{-b \pm \sqrt{b^2 - 4\mu I}}{2I} \qquad (13.5)$$

Depending on the relative magnitudes of the gain, damping factor, and inertia, the roots of the characteristic equation might be real or complex thereby indicating different types of response. If the roots turn out to be real, the transient response is merely the sum of the two resulting first-order exponential terms. If the roots are complex, however, they always appear in complex conjugate pairs in the following form:

$$p_{1,2} = \sigma \pm j\omega_d$$

where σ is the real part and ω_d the imaginary part of the roots. These complex roots yield a solution of the form

$$\theta_0(t)_{\text{transient}} = C_1e^{(\sigma+j\omega_d)t} + C_2e^{(\sigma - j\omega_d)t} \qquad (13.6)$$

After complex variable manipulations, this expression can be shown to be equivalent to

$$\theta_0(t) = Ae^{\sigma t} \cos(\omega_d t + \phi) \qquad (13.7)$$

where A and ϕ are derived from the coefficients C_1 and C_2.

This is the form of the solution whenever the characteristic equation has complex conjugate roots.

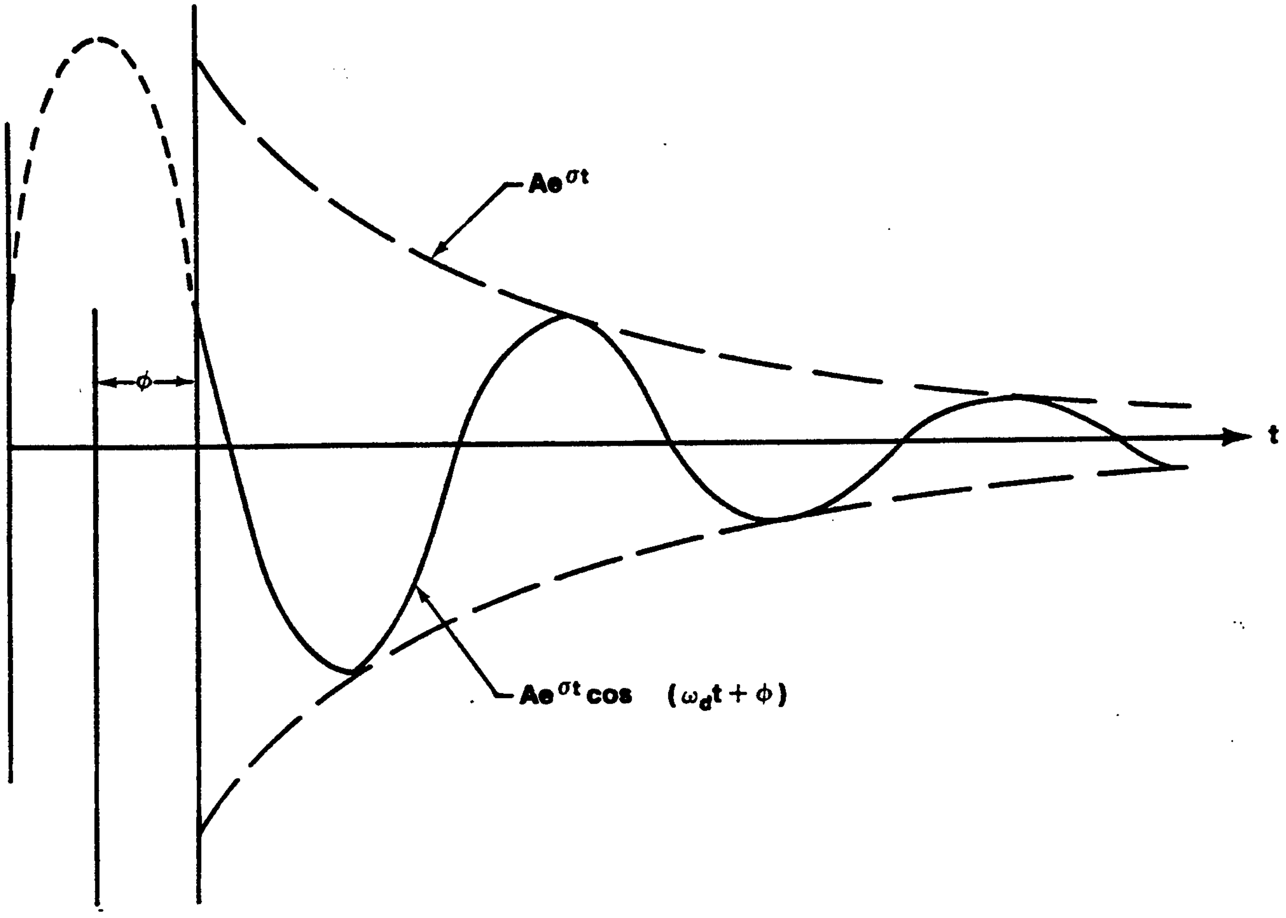


FIGURE 13.7. EXPONENTIALLY DAMPED SINUSOID - TYPICAL SECOND-ORDER SYSTEM RESPONSE

It is called an exponentially damped sinusoid and consists of a sine wave of frequency ω_d whose magnitude is $Ae^{\sigma t}$; that is, it is decreasing exponentially with time if σ is a negative quantity. A typical second-order response is plotted in Figure 13.7.

Referring back to the solution of the characteristic equation, Equation 13.5, the real part can be recognized as the exponent of e

$$\sigma = -\frac{b}{2I} \tag{13.8}$$

and the imaginary part as the frequency of the oscillation of the transient response

$$\omega_d = \sqrt{\frac{4\mu I - b^2}{4I^2}} \tag{13.9}$$

The quantity b represents the effective damping of the system. If b equals $2\sqrt{\mu I}$ the two roots $p_{1,2}$ are equal. This is the critical level of damping and is written $b' = 2\sqrt{\mu I}$.

The damping ratio is defined as the ratio of actual damping to the critical value of damping

$$\zeta = \frac{\text{actual damping}}{\text{critical damping}} = \frac{b}{b'} = \frac{b}{2\sqrt{\mu I}} \tag{13.10}$$

When ζ is greater than zero but less than one the roots are complex and the solution is a damped sinusoid of the form of Figure 13.8 and is called underdamped. When ζ is greater than one the roots are real and the response is overdamped. When ζ is negative, the system is unstable.

The undamped natural frequency, ω_n is defined as the frequency of oscillation of the transient if the damping is zero. From Equation 13.9

$$\omega_n = \sqrt{\frac{\mu}{I}} \tag{13.11}$$

The response in the case of no damping is a sine wave of constant amplitude.

Second-order equations (or factors in more complex systems) are frequently written in terms of the damping ratio and undamped natural frequency. Factoring $_m$ from Equation 13.4 leaves

$$\frac{I}{\mu}p^2 + \frac{b}{\mu}p + 1 = 0 \tag{13.12}$$

where I/μ can be recognized as $1/\omega_n^2$ and b/μ equals $2\zeta/\omega_n$. The characteristic equation becomes

$$\frac{1}{\omega_n^2} p^2 + \frac{2\zeta}{\omega_n} p + 1 = 0$$

Multiplying by ω_n^2 produces the standard form of the second-order system.

$$p^2 + 2\zeta\omega_n p + \omega_n^2 = 0 \qquad (13.13)$$

The roots of this equation are

$$p = \sigma \pm j\omega_d = -\zeta\omega_n \pm j\omega_n \sqrt{1 - \zeta^2} \qquad (13.14)$$

And the transient response in terms of ζ and ω_n is

$$c(t)_{\text{transient}} = Ae^{-\zeta\omega_n t} \quad \cos \quad \omega_n t\sqrt{1 - \zeta^2} + \phi \qquad (13.15)$$

Figure 13.8 shows a family of curves representing the response to a step input of a second-order system as a function of ζ. These curves illustrate the fact that the amount of overshoot and the time to arrive at the input value are a function of ζ.

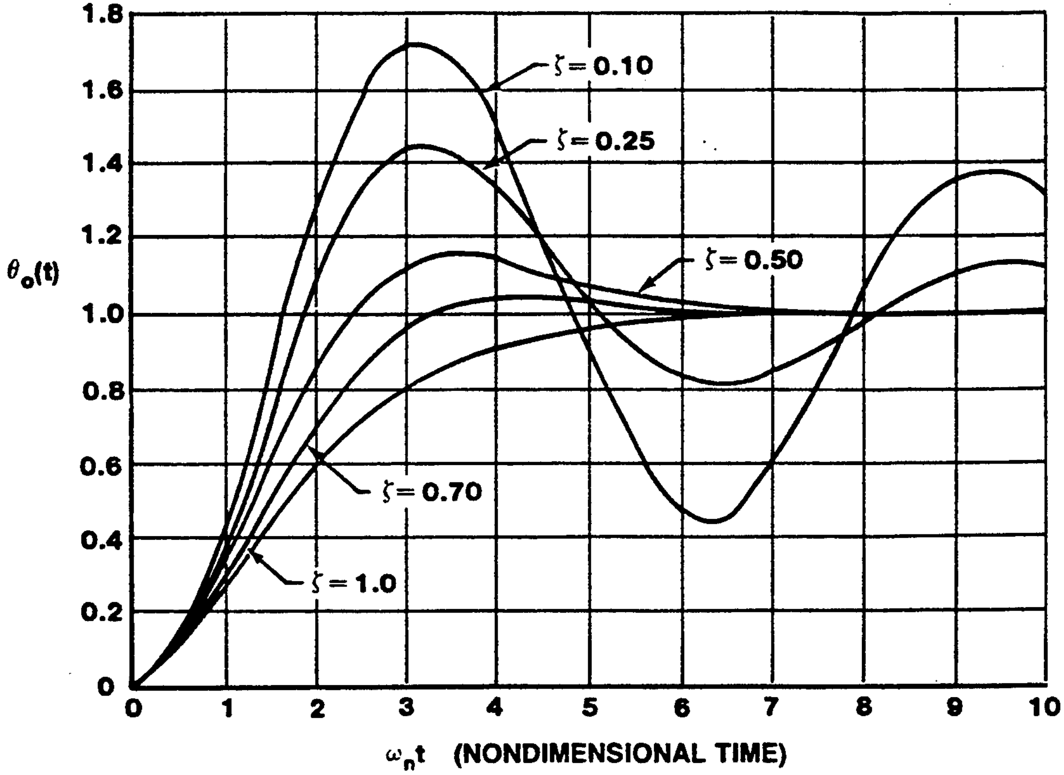


FIGURE 13.8. SECOND-ORDER TRANSIENT RESPONSE VERSUS ζ

13.4 TRANSFER FUNCTIONS

The first and second-order equations derived in the previous paragraph were solved using the classical method to show the types of response to be expected from each type system. In practice, this approach is extremely laborious, if not impossible, for more complicated systems. Therefore, more advanced techniques are used which do not produce the total solution but do indicate whether or not the system is stable; and if not, provide information about how to make the system stable. These sophisticated techniques generally use Laplace transforms. The use of operational calculus offers a definite advantage in that transfer functions can be manipulated using the normal rules of algebra. It also imposes a severe restriction. The systems to be analyzed must be representable by linear differential equations with constant coefficients. The method of determining these transfer functions will now be described.

First, we will consider the system of Figure 13.5 in which only damping was included. Its equation of motion is

$$\frac{b}{\mu}\dot{\theta}_0 + \theta_0 = \theta_i$$

Letting $\tau = b/\mu$ we have

$$\tau\dot{\theta}_0 + \theta_0 = \theta_i$$

Taking the Laplace transform using the notation $\mathcal{L}[\theta_0(t)] = \theta_0(s)$

$$\tau s\theta_0(s) - \tau\theta_0(0^+) + \theta_0(s) = \theta_i(s)$$

where

$\theta_0(0^+)$ is the value of $\theta_0(t)$ at $t = 0^+$

$$\theta_0(s)\ [\tau s + 1] = \theta_i(s) + \tau\theta_0(0^+)$$

$$\theta_0(s) = \frac{\theta_i(s)}{\tau s + 1} + \frac{\tau\theta_0(0^+)}{\tau s + 1}$$

Thus we see that the input to the system is acted upon by the transfer function

$$\frac{1}{\tau s + 1} \tag{13.16}$$

and also the initial condition is acted upon by this transfer function.

From this brief discussion we can see that if we assume all the initial conditions to be zero we obtain the relationship

$$\frac{\theta_0(s)}{\theta_i(s)} = \frac{1}{\tau s + 1} \qquad (13.17)$$

Our first-order system can then be described in the manner of Figure 13.4 where $G(s) = 1/(\tau s + 1)$, the transfer function.

The transfer function of our second-order system in which the inertia and damping were considered will now be determined using the same procedure. From its equation of motion,

$$I \ddot{\theta}_0 + b \dot{\theta}_0 + \mu\theta_0 = \mu\theta_i$$

taking the Laplace transform and assuming all initial conditions to be zero we have

$$\frac{I}{\mu} s^2 \theta_0(s) + \frac{b}{\mu} s \theta_0(s) + \theta_0(s) = \theta_i(s)$$

$$\theta_0(s) \left[\frac{I}{\mu} s^2 + \frac{b}{\mu} s + 1\right] = \theta_i(s)$$

The transfer function is then

$$\frac{\theta_0(s)}{\theta_i(s)} = \frac{1}{\frac{I}{\mu} s^2 + \frac{b}{\mu} s + 1} \qquad (13.18)$$

The transfer functions that have been developed for first and second-order systems (Equations 13.17 and 13.18) are obtained from the equation of motion of the whole system with the feedback loops closed (Figure 13.5). Therefore, they are called closed-loop transfer functions.

The denominator of Equation 13.18 is equivalent to Equation 13.12 and is the characteristic equation of the system.

13.5 TIME DOMAIN ANALYSIS

Much of the work of the control system engineer is done in the s-domain to take advantage of simplicity of solution, but the response of a system is in the time domain. The time response of a system is divided into two parts: (1) the transient response, and (2) the steady-state response.

$$c(t) = c_{trans}(t) + c_{s.s.}(t) \qquad (13.19)$$

In order to analyze a control system, we discuss the performance of the system in terms of time response to a specific input. For a given system, a specific input will result in a predictable transient response and a steady-state error. Control system performance specification can be stated in terms of the transient behavior of the system and the allowable steady-state error. In general, the steady-state error can be a function of time; however, we usually want $\lim_{t \to \infty} e(t)_{ss} = 0$.

In reality, control system specification and obtainable real world solutions are a compromise. The first order of business in analyzing a control system is to determine if the system is stable. If it is stable, then it will be tested to determine if it meets the performance specifications. The response of the system to specific test inputs will provide several measurements of performance.

13.5.1 Typical Time Domain Test Input Signals

13.5.1.1 The step input is the most commonly used test signal. This input is simply an instantaneous change in the reference input variable (Figure 13.9).

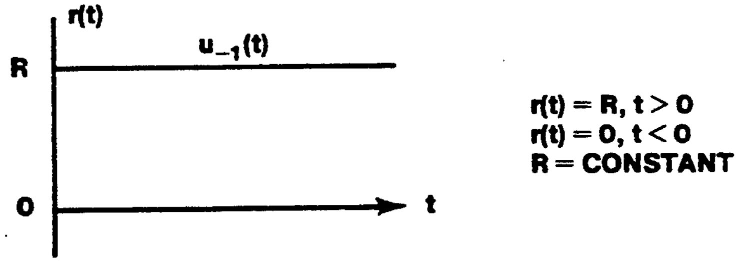


FIGURE 13.9. STEP INPUT

$$r(t) = u_{-1}(t)$$

where $u_{-1}(t)$ is the unit step function. The quantity $r(t)$ is not defined at $t = 0$. The Laplace transform of the unit step is

$$\mathcal{L}\{u_{-1}(t)\} = \frac{1}{s} \qquad (13.20)$$

Therefore the Laplace of $r(t) = Ru_{-1}(t)$ is R/s.

13.5.1.2 Ramp Function. The ramp signal is the integral of the unit step and is often called the velocity input (Figure 13.10).

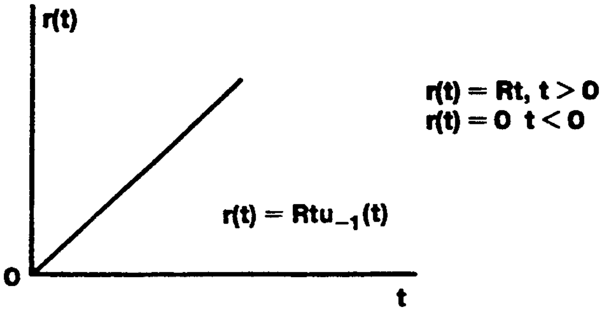


FIGURE 13.10. RAMP INPUT

$$r(t) = Rt\ u_{-1}(t)$$

and

$$\mathcal{L}\{Rt\ u_{-1}(t)\} = \frac{R}{s^2} \qquad (13.21)$$

13.5.1.3 Parabolic Input. The parabolic input signal (Figure 13.11) is the integral of the ramp signal and is often referred to as the acceleration input.

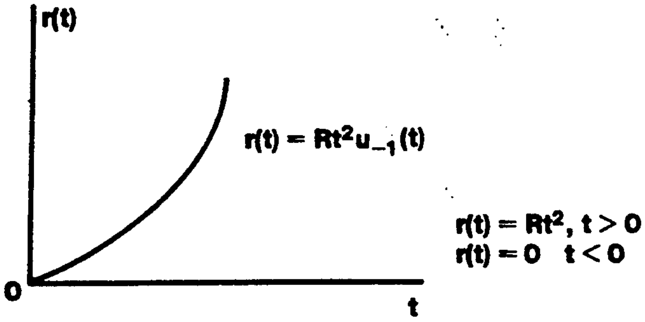


FIGURE 13.11. PARABOLIC INPUT

$$r(t) = Rt^2 u_{-1}(t)$$

and

$$\mathcal{L}\{Rt^2 u_{-1}(t)\} = \frac{2R}{s^3} \qquad (13.22)$$

13.5.1.4 Power Series Input. An input made up of the sums of a step, ramp and a parabola would be a power series of power 2.

$$r(t) = R\left[1 + t + \frac{t^2}{2}\right] u_{-1}(t) \qquad (13.23)$$

13.5.1.5 Unit Impulse. Another useful input is the unit impulse (Figure 13.12).

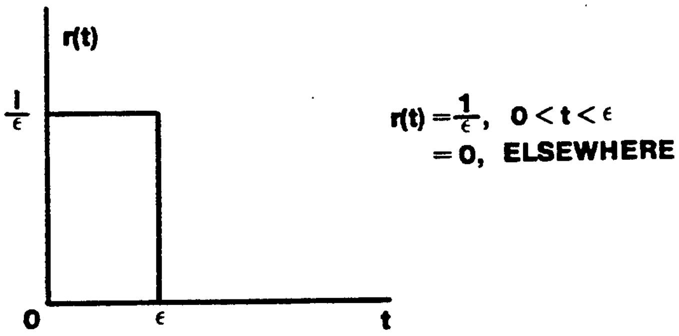


FIGURE 13.12. UNIT IMPULSE

As $\varepsilon \to 0$, the function $r(t)$ approaches the impulse function $\delta(t)$.

$$\mathcal{L}\{\delta(t)\} = 1 \tag{13.24}$$

13.5.2 Time Response of a Second-Order System

Consider the closed loop block diagram in Figure 13.13

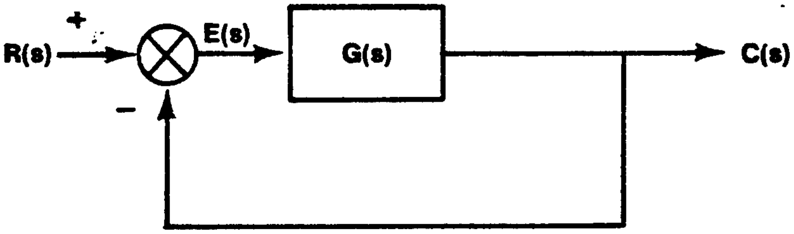


FIGURE 13.13. CLOSED-LOOP CONTROL SYSTEM

where

$$C(s) = \frac{G(s)\ R(s)}{1 + G(s)} \tag{13.25}$$

Let

$$G(s) = \frac{K}{s(s + a)}$$

$$C(s) = \frac{\dfrac{K}{s(s + a)}\ R(s)}{1 + \dfrac{K}{s(s + a)}}$$

$$= \frac{K}{s^2 + as + K}\ R(s)$$

This equation can be generalized in terms of ζ and ω_n.

Let $a = 2\zeta\omega_n$, $K = \omega_n^2$.

Then the control ratio C(s)/R(s) is

$$\frac{C(s)}{R(s)} = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \qquad (13.26)$$

For a unit step input, R(s) = 1/s, and

$$C(s) = \frac{C(s)}{R(s)} R(s)$$

$$C(s) = \frac{\omega_n^2}{s\ (s^2 + 2\zeta\omega_n\ s + \omega_n^2)} \qquad (13.27)$$

Taking the inverse Laplace gives the transient response

$$c(t) = 1- \frac{1}{\sqrt{1-\zeta^2}} \quad e^{-\zeta\omega_n t} \sin\left(\omega_n t\sqrt{1-\zeta^2} + \phi\right) \qquad (13.28)$$

where $\phi = \tan^{-1}\sqrt{1 - \zeta^2}/\zeta$. The transient response of this system varies according to the selected value of ζ. Figure 13.14 depicts this variation.

Several standard performance specification terms common throughout industry are illustrated in Figure 13.15.

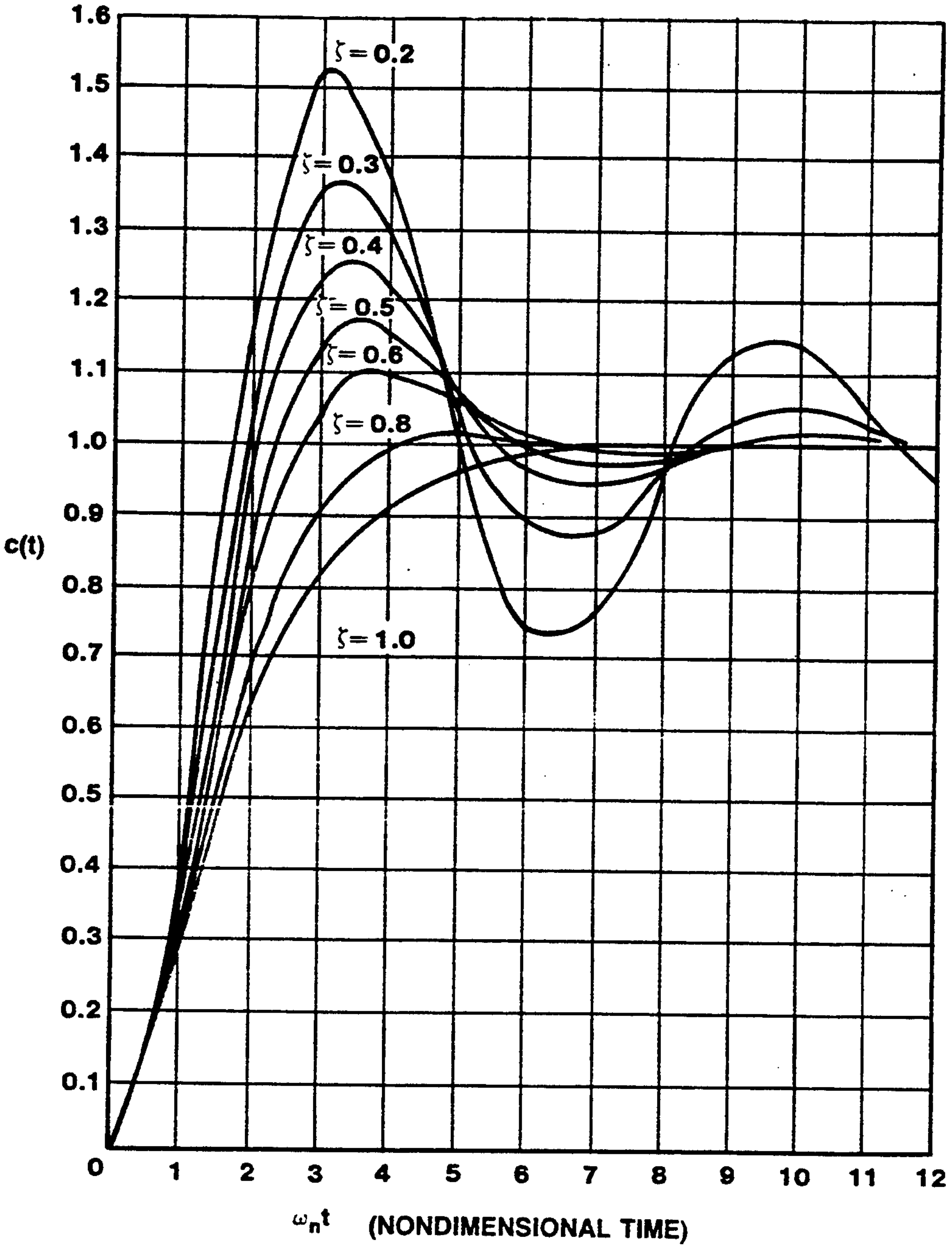


FIGURE 13.14. TRANSIENT RESPONSE OF A SECOND-ORDER SYSTEM TO A STEP INPUT

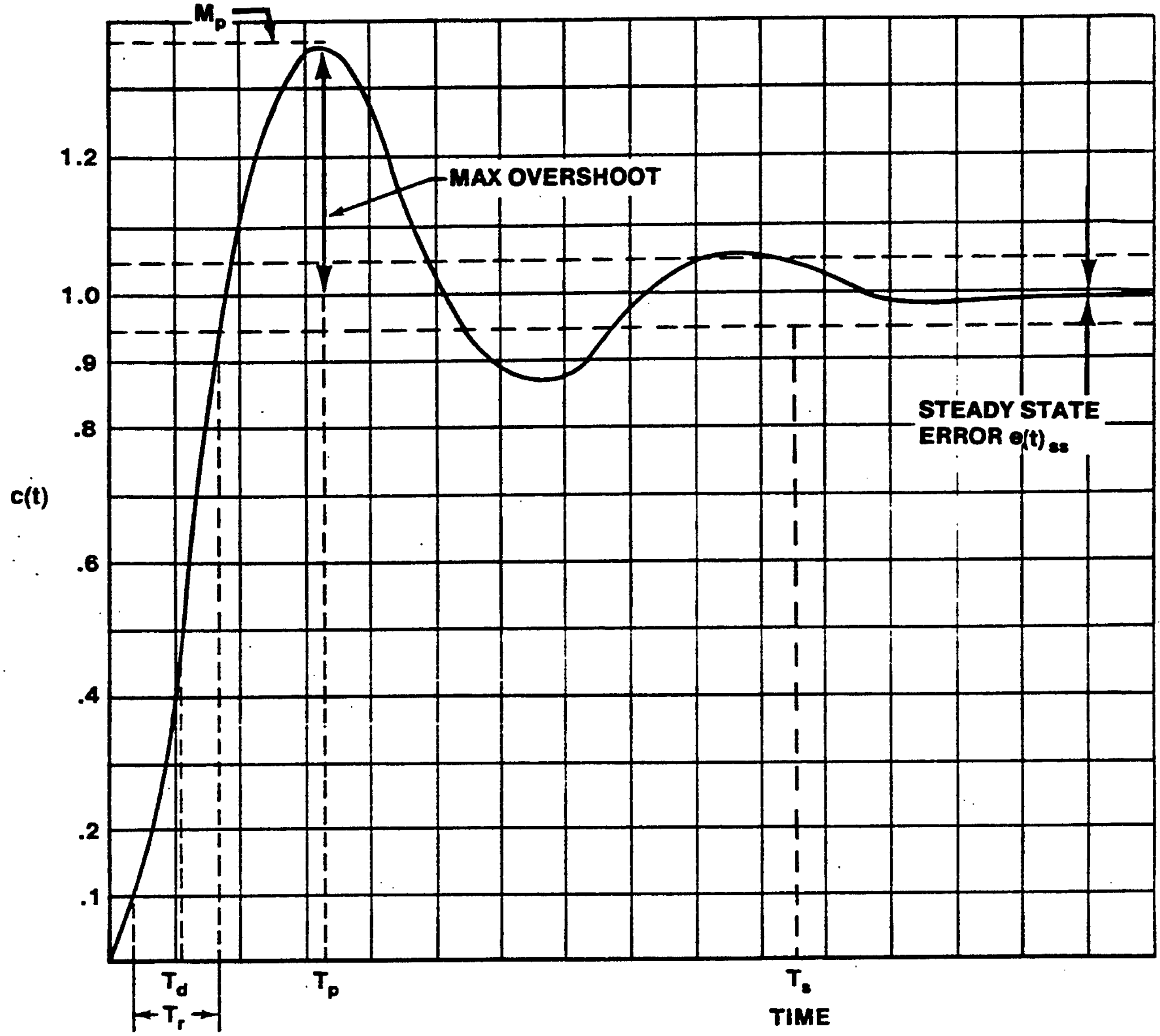


FIGURE 13.15. TIME DOMAIN SPECIFICATIONS

These transient performance specifications are usually defined for a unit step input.

1. Overshoot - indicated by largest error between input and output during the transient state. We can determine the magnitude by using the previously developed equation for a step input to a second-order system.

$$c(t) = 1 - \frac{1}{\sqrt{1-\zeta^2}} e^{-\zeta\omega_n t} \sin\left(\omega_n t \sqrt{1-\zeta^2} + \tan^{-1} \frac{\sqrt{1-\zeta^2}}{\zeta}\right) \quad (13.29)$$

Taking the derivative of this equation and equating to zero yields

$$\frac{dc(t)}{dt} = 0 = \frac{1}{\sqrt{1-\zeta^2}} \zeta\omega_n e^{-\zeta\omega_n t} \sin\left(\omega_n t\sqrt{1-\zeta^2} + \tan^{-1} \frac{\sqrt{1-\zeta^2}}{\zeta}\right)$$

$$- \frac{1}{\sqrt{1-\zeta^2}} e^{-\zeta\omega_n t} \omega_n \sqrt{1-\zeta^2} \cos\left(\omega_n t\sqrt{1-\zeta^2} + \tan^{-1} \frac{\sqrt{1-\zeta^2}}{\zeta}\right) \quad (13.30)$$

This derivative is zero when $\omega_n t \sqrt{1-\zeta^2} = 0, \pi, 2\pi,$ etc.
The peak overshoot occurs at the first value after zero (with initial conditions equal to zero).

Therefore, the time to maximum or peak overshoot is

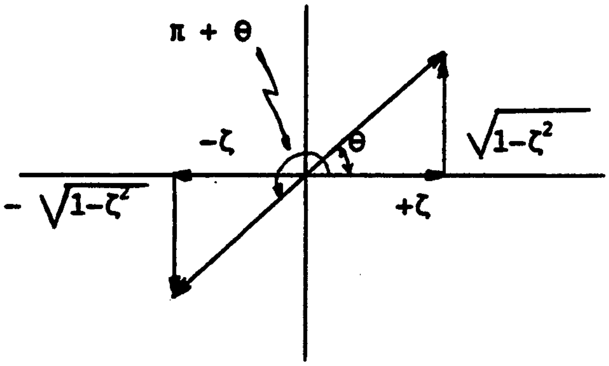
$$\boxed{T_p = \frac{\pi}{\omega_n \sqrt{1-\zeta^2}}} \quad (13.31)$$

Substituting this value into c(t) yields the peak response, M_p.

$$M_p = 1 - \frac{1}{\sqrt{1-\zeta^2}} \exp \frac{-\zeta\omega_n \pi}{\omega_n\sqrt{1-\zeta^2}} \sin\left(\pi + \tan^{-1} \frac{\sqrt{1-\zeta^2}}{\zeta}\right)$$

Note: $$\sin\left(\pi + \tan^{-1} \frac{\sqrt{1 - \zeta^2}}{\zeta}\right) = \sin(\pi + \theta) = -\sin\theta$$

$$= -\sqrt{1 - \zeta^2}$$



$$\theta = \tan^{-1}\frac{\sqrt{1-\zeta^2}}{\zeta}$$

And

$$\boxed{M_p = 1 + \exp\left(-\frac{\zeta\pi}{\sqrt{1 - \zeta^2}}\right)} \qquad (13.32)$$

The overshoot for the unit step input is

$$\text{Overshoot} = M_p - 1 = \exp\left(-\frac{\zeta\pi}{\sqrt{1 - \zeta^2}}\right) \qquad (13.33)$$

and the percent of overshoot

$$\text{P.O.} = \frac{M_p - 1}{1} \times 100\%$$

$$= 100\% \exp\left(-\frac{\zeta\pi}{\sqrt{1 - \zeta^2}}\right) \qquad (13.34)$$

2. Delay Time, T_d, is the time required for the response to a unit step to reach 50% of its final value.

3. Rise Time, T_r, is the time required for the response to a unit step to rise from 10% to 90% of its final value.

4. Settling Time, T_s, is the time required for the response to a unit step to decrease to and to stay within a specific percentage of its final value. Commonly used values are 2% or 5% of the final value.

13.5.3 Higher-Order Systems

The relationships developed in the preceding paragraph using ω_n and ζ apply equally well for each complex-conjugate pair of poles of an n^{th}-order system. The distinction is that the dominant ζ and ω_n apply for that pair of complex-conjugate poles which lie closest to the imaginary axis. The values of ζ and ω_n are dominant because the corresponding transient term has the longest settling time and the largest magnitude. Therefore, the dominant poles primarily determine the shape of the time response, c(t). A nondominant pole(s) has a real axis component that is at least six times further to the left than the corresponding component of the dominant pole(s). Components of c(t) due to nondominant pole(s) die out relatively quickly, and can be neglected. (13.1:245)

13.5.4 Time Constant, τ

The time constant is used as a measure of the exponential decay of a response. For first-order systems, the transient response is an exponential function described by Ae^{-mt}, Figure 13.16.

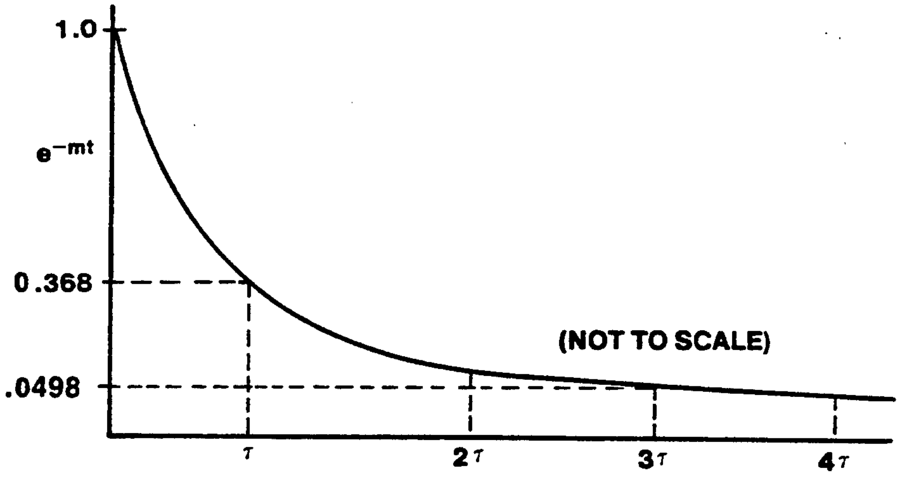


FIGURE 13.16. PLOT OF EXPONENTIAL e^{-mt}

The value of time which makes the exponent of e equal to -1 is defined as the time constant, τ

$$-m\tau = -1$$

$$\tau = \frac{1}{m}$$

In one time constant the exponential e^{-mt} will decrease from a value of 1.0 to a value .368. Table 13.1 shows values for other time constants.

TABLE 13.1

TIME CONSTANT TABLE

t	e^{-mt}
1τ	.368
2τ	.135
3τ	.0498
4τ	.0183

The time constant is another way of specifying settling time. The exponential will decay to 37% of its initial value in τ seconds (one time constant). In 3τ the exponential is within approximately 5% and in 4τ approximately 2% of its final value. For a second-order underdamped system of the form

$$c(t) = 1 - \frac{1}{\beta} e^{-\zeta\omega_n t} \sin(\omega_n \beta t + \phi)$$

the response is bounded by the exponent of the form $(1/\beta)e^{-mt}$. The specifications of ζ and ω_n determine the bounding exponential curve. The time constant, τ, for these systems is

$$\tau = \frac{1}{\zeta\omega_n} \qquad (13.35)$$

13.6 STABILITY DETERMINATION

The most important area of the analysis of a closed-loop control system is the determination of stability. A system is said to be stable if the output of the system corresponds to the input after transients die out. A system is said to be unstable if the transients do not die out or if they grow larger following a disturbance.

Stability is an inherent characteristic of the system and depends only upon the system itself, not upon the input or forcing function. Hence, if a system is unstable, any input will cause the system to diverge. If the system is stable, any bounded input will cause a bounded response.

The problem in determining stability is ascertaining whether or not the transients of a system will die out BEFORE the system is built.

We must determine the conditions under which a system will become unstable and be able to tell when this happens in the analysis of the system. Several methods are available for determining stability: root locus, Bode plot, Routh's stability criterion, and Nyquist Criterion. Only the root locus and the Bode plot methods will be presented in this chapter.

13.6.1 Stability in the s-Plane

Since this course is concerned with linear systems, i.e., those whose differential equations are linear with constant coefficients, the transient response is of the form

$$c(t)_{\text{transient}} = \sum_{i=1}^{n} k_i e^{s_i t}$$

where n is the order of the differential equation and the values of s are the roots of the characteristic equation which are, in general, complex.

$$s = \sigma + j\omega_d \qquad (13.36)$$

σ is the real part of the complex variable s and ω_d is the imaginary part of the complex variable s. The notation used to indicate this is

$$\sigma = \mathrm{Re}\{s\}$$

$$\omega_d = \mathrm{Im}\{s\}$$

Previously, we discussed only a first and second-order system and saw the type of transient response to be expected from each. The characteristic equation of higher-order systems, however, can in theory be factored into the product of several first and second-order factors depending on the order of the equation. This is demonstrated in Equations 13.37 and 13.38

$$A_n s^n + A_{n-1} s^{n-1} + \dots . A_0 = 0 \tag{13.37}$$

Can be expressed

$$(\tau_1 s + 1)(\tau_2 s + 1) \dots \left(\frac{s^2}{\omega_{n_1}^2} + \frac{2\zeta_1}{\omega_{n_1}} s + 1\right)\left(\frac{s^2}{\omega_{n_2}^2} + \frac{2\zeta_2}{\omega_{n_2}} s + 1\right) \dots = 0 \tag{13.38}$$

The transient response of a complex system is the sum of those associated with each of the first and second-order factors. Each root of the characteristic equation must be of one of the forms shown in Figure 13.17. Opposite the possible values of the roots on the left are shown the corresponding transient response components as a function of time. Note that complex or imaginary roots always occur in complex-conjugate pairs. That is, they have imaginary parts of equal magnitude but are opposite in sign.

All the possible values of s can also be described through use of a complex plane -- in this case the s-plane. A complex plane is one in which the value of the real part of the complex variable is the distance along the abscissa and the magnitude of the imaginary part is described along the ordinate. These are called the real and imaginary axes, respectively. The complex variable, s, is then a position vector in the complex s-plane where $\sigma = \mathrm{Re}\{s\}$ is the magnitude of the real component and $\omega_d = I_m\{s\}$ is the imaginary component.

Figure 13.18 shows the s-plane. If the values of s, which are the roots of the characteristic equation, are plotted in the s-plane, they produce a transient solution component as indicated. Areas in the s-plane in which roots produce stable and unstable responses are also shown. Roots yielding an undamped response or neutrally stable output are all on the imaginary axis. Roots associated with non-oscillatory response are all on the real axis. A root of the characteristic equation at the origin ($s = 0$) has a transient solution equal to a constant.

The mathematical definition of a stable system is one in which the roots of the equation have only negative real parts. In other words, where $s = \sigma + j\omega_d$ are the roots of the characteristic equation, $\sigma < 0$ produces a stable system.

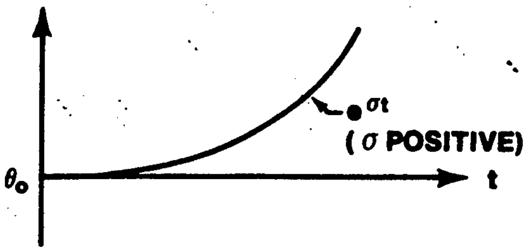
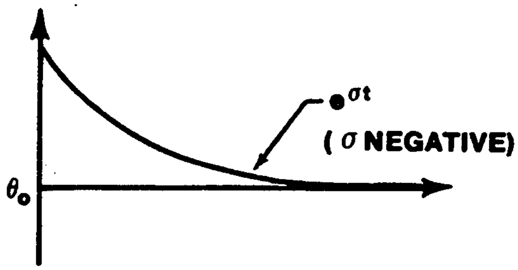
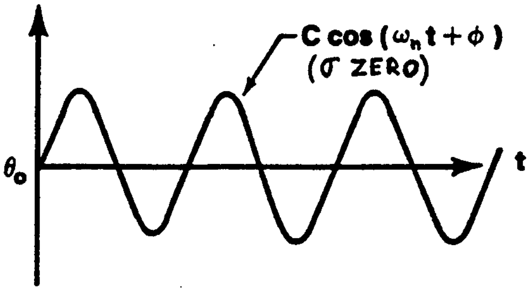
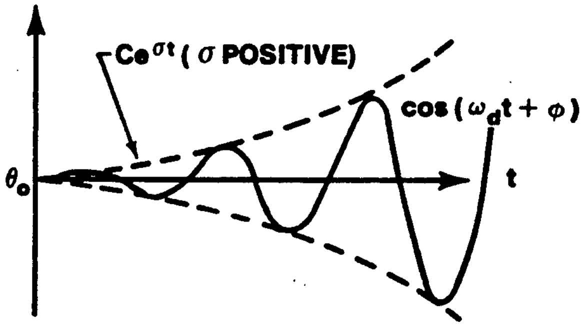
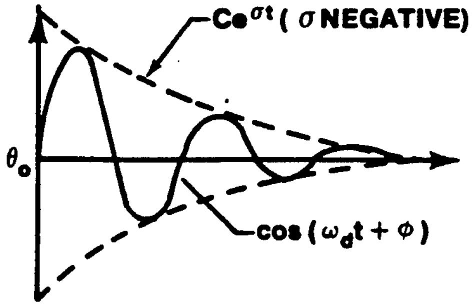
VALUE OF s	TIME RESPONSE	PICTORIAL SOLUTION
σ (POSITIVE)	$Ce^{\sigma t}$	
σ (NEGATIVE)	$Ce^{\sigma t}$	
$\pm j\omega_n$ (σ ZERO)	$C \cos(\omega_n t + \phi)$	
$\sigma \pm j\omega_d$ (σ POSITIVE)	$Ce^{\sigma t} \cos(\omega_d t + \phi)$	
$\sigma \pm j\omega_d$ (σ NEGATIVE)	$Ce^{\sigma t} \cos(\omega_d t + \circ)$	

FIGURE 13.17. TRANSIENT SOLUTION OF LINEAR CONSTANT COEFFICIENT EQUATIONS

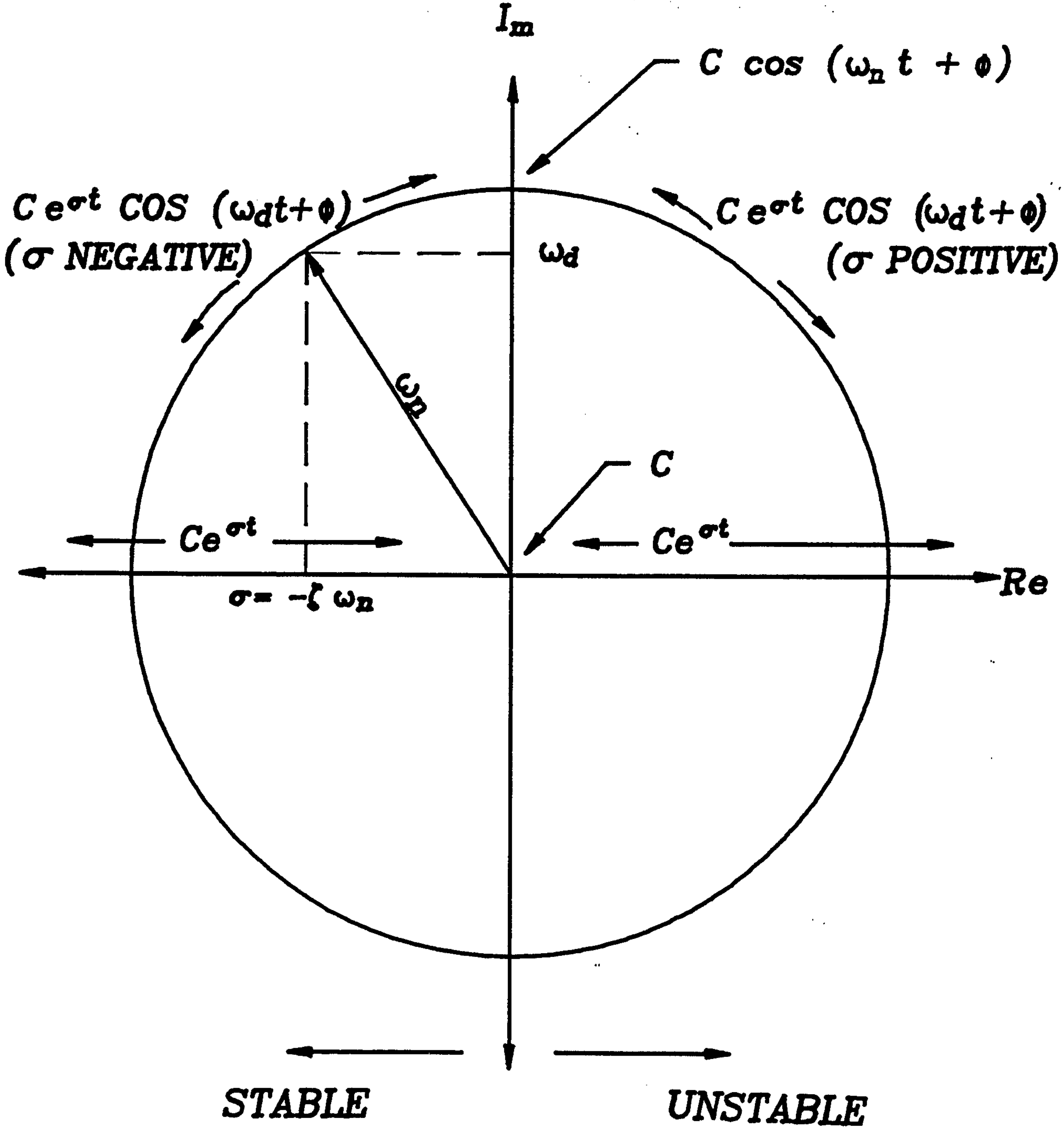


FIGURE 13.18 COMPLEX s-PLANE

13.6.2 Additional Poles and Zeros

The results of adding a real pole or a real zero to the basic second-order control ratio as given by Equation 13.26 will be investigated. When a pair of complex-conjugate poles are dominant, the approximations developed in Paragraph 13.5.2 yield accurate results. The addition of a third real pole to a second-order transfer function can significantly alter the system time response c(t), and the approximations given in Paragraph 13.5.2 no longer give accurate results.

The effects of a third real pole can be seen by considering the control ratio

$$\frac{C(s)}{R(s)} = \frac{K}{\left(s^2 + 2\zeta\omega_n\, s + \omega_n^2\right)\left(s - p_3\right)} \; ; \; K = \left(\omega_n^2 \quad p_3\right) \tag{13.39}$$

The time response resulting from a unit step input is

$$c(t) = 1 + 2\,|A|\; e^{-\zeta\omega_n t} \sin\left(\omega_d t + \phi\right) + Be^{p_3 t} \tag{13.40}$$

$$\phi = \tan^{-1} \frac{\omega_d}{p_3 - \zeta\omega_n} \tag{13.40b}$$

The transient term due to the real pole, p_3, has the form $Be^{p_3 t}$, where B is always negative. Therefore, the peak overshoot is reduced, the settling time, t_s, may be increased or decreased, and the phase angle is decreased. This is the typical effect of adding a third real pole. The further to the left p_3 is, the smaller the magnitude of B and the effect on c(t). Typical time responses as a function of the real pole location are shown in Figure 13.19 (13.1:350).

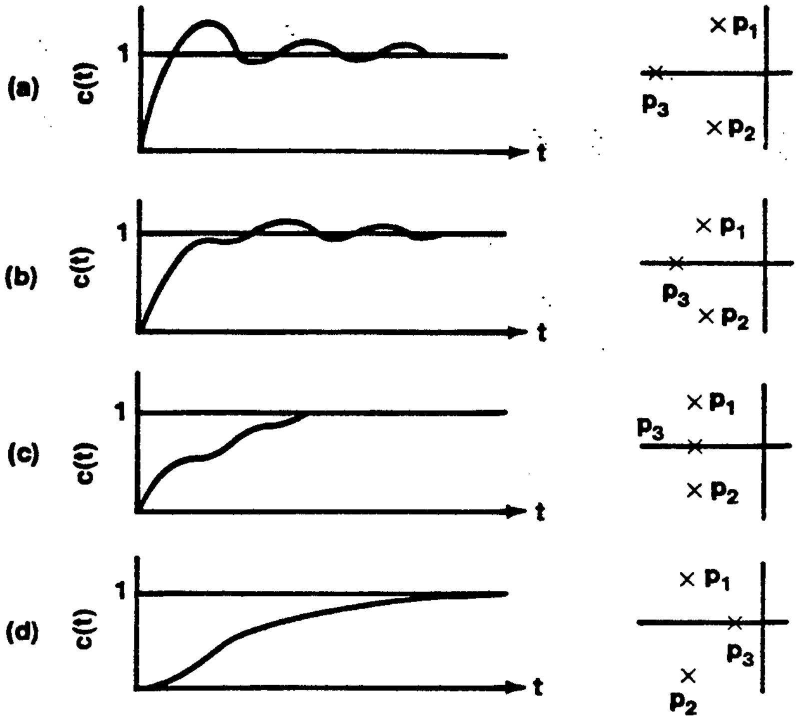


FIGURE 13.19 TIME RESPONSE AS A FUNCTION OF REAL POLE LOCATION

The time response is also altered by the addition of a real zero to the basic second-order transfer function of Equation 13.26. The control ratio now becomes

$$\frac{C(s)}{R(S)} = \frac{K(s - z)}{\left(s^2 + 2\zeta\omega_n s + \omega_n^2\right)}; \quad \left(K = \omega_n^2 /z\right) \tag{13.41}$$

The time response resulting from a unit step input is

$$c(t) = 1 + \frac{1}{\omega_d} \left[(z - \zeta\omega_n)^2 + \omega_d{}^2\right]^{1/2} e^{-\zeta\omega_n t} \sin(\omega_d t + \phi) \tag{13.42}$$

where

$$\phi = \tan^{-1} \frac{\omega_d}{z - \zeta\omega_n} \tag{13.43}$$

From Equations 13.42 and 13.43, it is seen that the addition of a real zero affects both the magnitude and phase of the transient part of c(t). The real zero tends to increase the overshoot and decrease the phase angle of c(t). This effect becomes more dramatic as the zero approaches the imaginary axis. This is illustrated in Figure 13.20 for a stable second-order system with $\zeta\omega_n$ held constant.

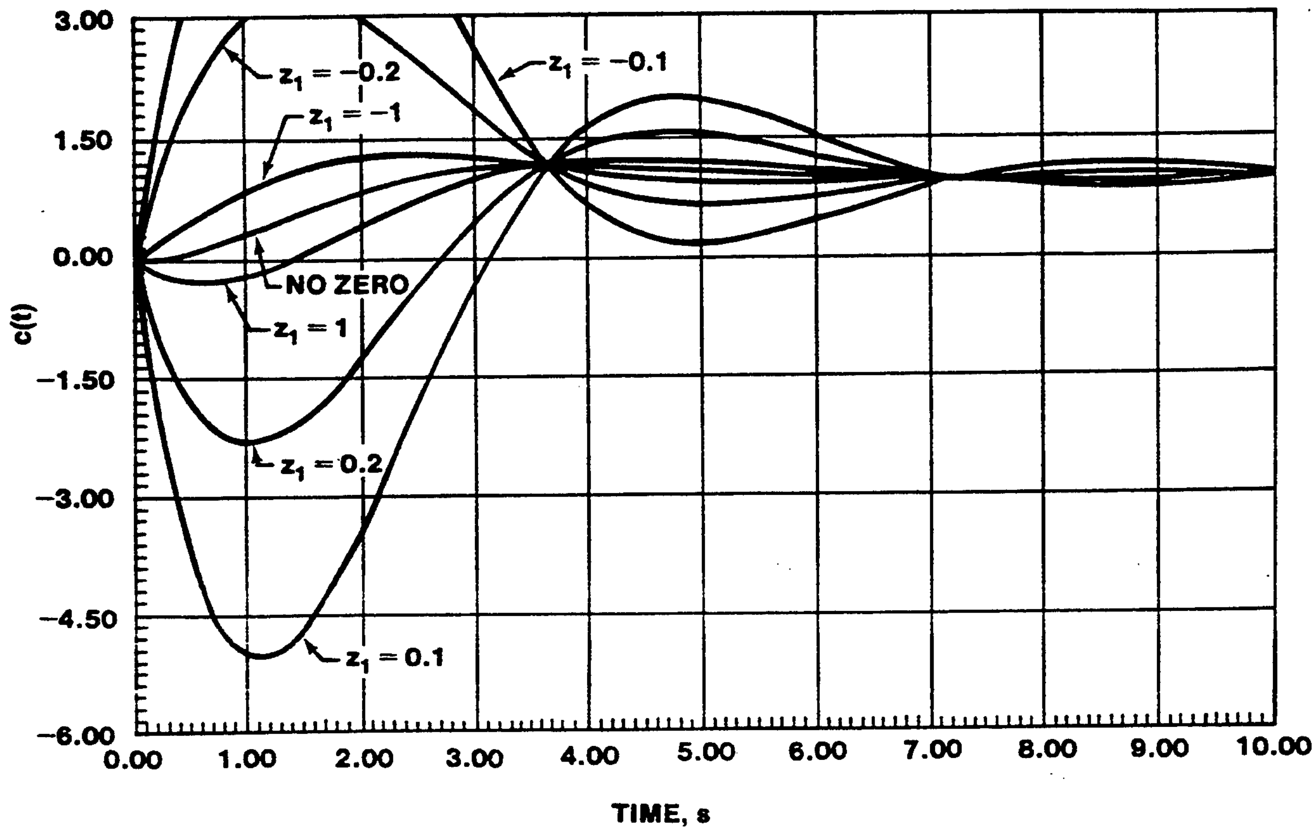


FIGURE 13.20. TIME RESPONSE OF A SECOND-ORDER SYSTEM AS A FUNCTION OF REAL ZERO LOCATION

Several things can be seen by examining the responses in Figure 13.20. First, it should be noted that rise time is decreased and the overshoot is increased with the addition of a real zero. This effect is more noticeable as the zero moves closer to the imaginary axis. This is to be expected, because when the zero is at the origin, it acts as a pure differentiator of the input. Differentiation of the unit step input yields the unit impulse. When the zero is in the right half plane, the response is stable but the direction of the initial response is opposite to the final steady-state value. Additionally, it should be noted that the initial slope of the response is not zero as is true for a second-order system without a zero (13.2:91). The block diagram of a system zero is shown in Figure 13.21.

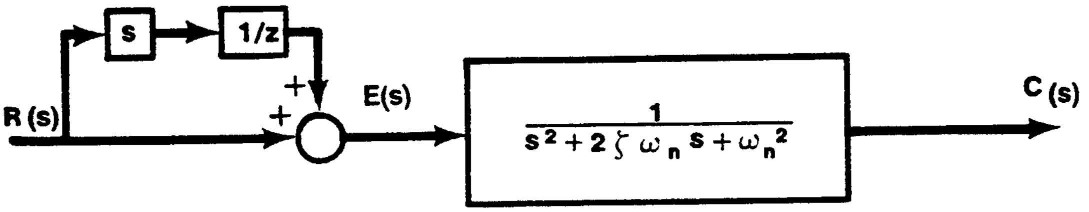


FIGURE 13.21. BLOCK DIAGRAM OF A SECOND-ORDER SYSTEM WITH A REAL ZERO

From Figure 13.21 it can be seen that the zero operates on the input signal to produce a signal proportional to both the magnitude and the derivative (rate of change) of the input signal. Therefore, the system will react not only to the magnitude of input, but also to its rate of change. If R(s) is changing rapidly, then E(s) is large and the system responds faster. (13.1:360).

13.7 STEADY-STATE FREQUENCY RESPONSE

We have looked at the time domain analysis and specifications of control systems. In the time domain analysis, the typical test inputs were the step, ramp, and parabola. The frequency response technique, introduced in this paragraph, is a valuable tool to the control systems engineer and provides a standardized method representing the total performance of a system. The input for steady-state frequency response is the sinusoid

$$r(t) = A_1 \sin \omega t$$

The basis for the frequency response method is that a system's response to a sinusoid will be a sinusoid at the same frequency, but the response will differ in magnitude and phase angle. All that is needed to completely specify the steady-state frequency response is to be able to find the magnitude of the output and the phase angle.

The fact that the output is a sinusoid of the same frequency can be shown by analyzing a sinusoid input to a first-order system described by

$$G(s) = \frac{B}{s + \frac{1}{\tau}} \tag{13.44}$$

The input, $r(t) = A_1 \sin \omega t$ in Laplace transform is

$$R(s) = \frac{A_1 \omega}{s^2 + \omega^2}$$

$$C(s) = G(s)\ R(s)$$

and

$$C(s) = \frac{B}{s + \frac{1}{\tau}} \frac{A_1 \omega}{s^2 + \omega^2}$$

Using partial fraction expansion yields

$$C(s) = \frac{C_1}{s + \frac{1}{\tau}} + \frac{C_2\ s}{s^2 + \omega^2} + \frac{C_3}{s^2 + \omega^2}$$

Finding the coefficients C_1, C_2, and C_3 can be a tedious process. By inspection we can write the form of the solution as

$$c(t) = C_1 e^{-t/\tau} + C_2 \cos \omega t + \frac{C_3}{\omega} \sin \omega t$$

Another form of this equation is

$$c(t) = C_1 e^{-t/\tau} + A_0 \sin (\omega t + \phi) \tag{13.45}$$

The steady-state response can be written as

$$c(t)_{ss} = \underset{\text{Magnitude}}{A_0} \sin (\omega t + \underset{\text{Phase Angle}}{\phi}) \tag{13.46}$$

which tells us that the steady-state response will always have the same frequency as the input but will differ in phase angle and magnitude. The transient response due to the exponential term, $C_1 e^{-t/\tau}$, decays to zero as $t \to \infty$.

The Laplacian operator, s, contains both real and imaginary components and to evaluate coefficient C_3 the complex variable "s" would be selected to be $\pm j\omega$ (i.e. purely imaginary). Thus

$$s = \sigma + j\omega_d$$

and for a constant amplitude-input sinusoid, σ is zero, (Figure 13.22), therefore

$$s = j\omega$$

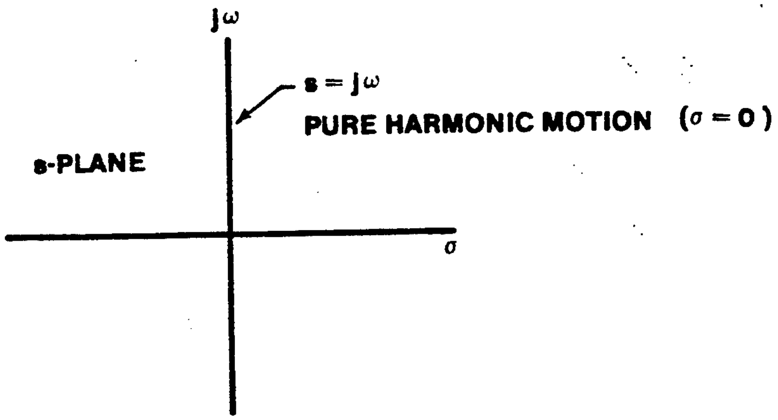


FIGURE 13.22. s-PLANE - PURE HARMONIC MOTION

The frequency response function, G(jω), is defined by replacing s with jω in the system transfer function (Equation 13.44).

$$G(s) = \frac{B}{s + \frac{1}{\tau}}$$

becomes

$$G(j\omega) = \frac{B}{j\omega + \frac{1}{\tau}}$$

It is important to remember that we are talking about the steady-state frequency response only when we replace s with jω.

13.7.1 Complex Numbers

In the study of feedback control systems, the relative magnitude and time relationship between such quantities as position, speed, voltage, current, force, and torque are the items of interest. These are all real physical quantities which behave according to the laws of nature. It is frequently convenient, however, to represent these physical quantities by complex mathematical symbols that indicate more than the information describing the real quantities themselves. The use of complex variables to represent real physical quantities has the advantage of simplifying the mathematical process necessary to solve the prvoblem. On the other hand, it has the disadvantage of obscuring the true value of the real physical quantities.

It is the purpose of this chapter to introduce the complex variable notation which will be used later. Complex quantities are usually expressed in one of four forms:

(a) rectangular

(b) polar

(c) trigonometric

(d) exponential

The equivalence of these four forms will now be demonstrated.

13.7.1.1 Rectangular Form. The complex quantity $\overline{z}$ is drawn on the complex plane in Figure 13.23. It can be thought of as a position vector in the complex plane.

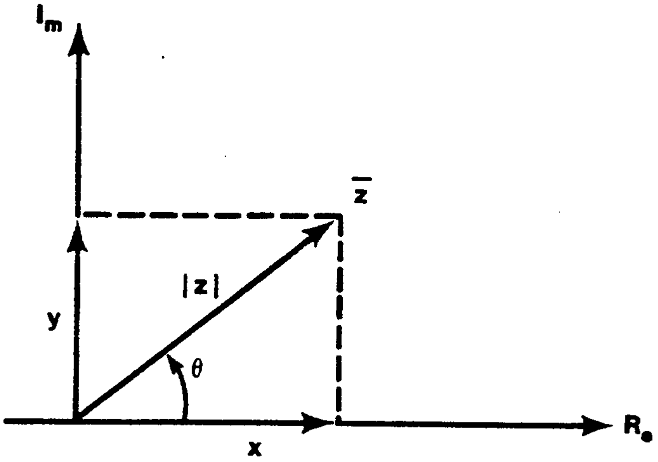


FIGURE 13.23. THE COMPLEX PLANE

The real part is measured along the horizontal or real axis and the imaginary part is measured along the vertical or imaginary axis.

In rectangular form the complex variable, z, is

$$\bar{z} = x + jy$$

where j is the imaginary quantity $\sqrt{-1}$.

13.7.1.2 Polar Form. Any position in the complex plane can also be defined by the angle, θ, of the position vector $\bar{z}$, and its magnitude, $|\bar{z}|$(Figure 13.23). In polar form the complex quantity, $\bar{z}$, is

$$\bar{z} = |\bar{z}| \angle \theta$$

In terms of the rectangular form parameters,

$$|\bar{z}| = \sqrt{x^2 + y^2}$$

$$\theta = \tan^{-1} y/x$$

13.7.1.3 Trigonometric Form. The trigonometric form of the position vector in the complex plane can be written again using Figure 13.23. We see that

$$\cos \theta = \frac{x}{\sqrt{x^2 + y^2}}$$

and

$$\sin \theta = \frac{y}{\sqrt{x^2 + y^2}}$$

Therefore,

$$\cos \theta + j \sin \theta = \frac{x}{\sqrt{x^2 + y^2}} + j \frac{y}{\sqrt{x^2 + y^2}}$$

Multiplying both sides by

$$|\bar{z}| = \sqrt{x^2 + y^2}$$

We have

$$|\bar{z}| \; [\cos\theta + j\sin\theta] = \sqrt{x^2 + y^2}\left[\frac{x}{\sqrt{x^2 + y^2}} + j\frac{y}{\sqrt{x^2 + y^2}}\right]$$

The trigonometric form is then

$$\bar{z} = |\bar{z}| \; [\cos\theta + j\sin\theta]$$

13.7.1.4 Exponential Form. The exponential form of a complex quantity is most convenient for mathematical manipulation. It will be shown equivalent to the trigonometric form.

The MacLaurin series expansion e^x is

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$$

Letting $x = j\theta$

$$e^{j\theta} = 1 + j\theta - \frac{\theta^2}{2!} - j\frac{\theta^3}{3!} + \ldots \qquad (13.47)$$

Now, $\sin\theta$ and $\cos\theta$ can be defined by series expansions as follows:

$$\sin\theta = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + \ldots \qquad (13.48)$$

$$\cos\theta = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \ldots \qquad (13.49)$$

Recalling that $j^2 = -1$, Equation 13.48 and 13.49 may be written

$$\cos\theta = 1 + \frac{(j\theta)^2}{2!} + \frac{(j\theta)^4}{4!} + \frac{(j\theta)^6}{6!} + \ldots \qquad (13.50)$$

$$j \sin \theta = \frac{j\theta}{1!} + \frac{(j\theta)^3}{3!} + \frac{(j\theta)^5}{5!} + \frac{(j\theta)^7}{7!} + \ldots \qquad (13.51)$$

Adding Equation 13.50 to 13.51 yields

$$\cos \theta + j \sin \theta = 1 + \frac{j\theta}{1!} + \frac{(j\theta)^2}{2!} + \frac{(j\theta)^3}{3!} + \ldots \qquad (13.52)$$

The right side of Equation 13.52 is equal to Equation 13.47, therefore

$$e^{j\theta} = \cos \theta + j \sin \theta$$

and finally

$$\bar{z} = |\bar{z}|\, e^{(j\theta)} \qquad (13.53)$$

We have proven the four forms of the complex variable z to be consistent. They are summarized below

Rectangular $\bar{z} = x + jy$

Polar $\bar{z} = |\bar{z}| \angle \theta$

Trigonometric $\bar{z} = |\bar{z}| [\cos \theta + j \sin \theta]$

Exponential $\bar{z} = |\bar{z}|\, e^{j\theta}$

13.7.2 Bode Plotting Technique

With this background in complex notation we will develop the Bode technique of frequency response. Beginning with a generalized transfer function (Equation 13.54), we will manipulate it into the frequency response standard form, sometimes called the Bode form (Equation 13.55).

$$G(s) = \frac{K\, s^m\, (s + a_1)(s + a_2)(s^2 + 2\zeta\omega_{n_1} s + \omega_{n_1}^{\ 2})}{s^n (s + a_3)(s + a_4)(s^2 + 2\zeta\omega_{n_2} s + \omega_{n_2}^{\ 2})} \qquad (13.54)$$

Equation 13.54 must be normalized as follows:

$$G(s) = \frac{K\, a_1\, a_2\, \omega_{n_1}^{\ 2} s^m (\tau_1\ s + 1)\ (\tau_2\ \ s + 1)\left(\frac{s^2}{\omega_{n_1}^{\ 2}} + \frac{2\zeta s}{\omega_{n_1}} + 1\right)}{a_3\ a_4\ \omega_{n_2}^{\ 2}\ s^n\ \ (\tau_3\ s + 1)\ \ (\tau_4\ s + 1)\left(\frac{s^2}{\omega_{n_2}^{\ 2}} + \frac{2\zeta s}{\omega_{n_2}} + 1\right)} \qquad (13.55)$$

where

$$\tau_1 = \frac{1}{a_1},\quad \tau_2 = \frac{1}{a_2},\ \text{etc.}$$

Let

$$K_n = \frac{K\, a_1\ \ a_2\ \ \omega_{n_1}^{\ 2}}{a_3\ \ a_4\ \ \omega_{n_2}^{\ 2}}$$ and K = static loop sensitivity

Substitute $j\omega$ for s in Equation 13.55 and rewriting in the Frequency Response Standard Form

$$G(j\omega) = \frac{K_n (j\omega)^m\ (1 + j\tau_1\omega)\ (1 + j\tau_2\omega)\left(1 - \frac{\omega^2}{\omega_{n_1}^{\ 2}} + j2\zeta\ \frac{\omega}{\omega_{n_1}}\right)}{(j\omega)^n\ (1 + j\tau_3\omega)\ \ (1 + j\tau_4\omega)\left(1 - \frac{\omega^2}{\omega_{n_2}^2} + j2\zeta\ \frac{\omega}{\omega_{n_2}}\right)} \qquad (13.56)$$

Equation 13.56 can be written in polar form

$$G(j\omega) = |G(j\omega)| \angle \phi(\omega)$$

where $|G(j\omega)|$ is of the form $\sqrt{Re^2 + Im^2}$

and $\phi(\omega)$ of the form $\tan^{-1} \frac{Im}{Re}$

Furthermore, $G(j\omega)$ can be written in exponential form where

$$G(j\omega) = |G(j\omega)| \; e^{j\phi(\omega)} \qquad (13.57)$$

To express $G(j\omega)$ in either of these formats will entail finding the magnitude $G(j\omega)$ and the phase angle $\phi(\omega)$. The Bode technique requires taking the log of $G(j\omega)$ to take advantage of addition and subtraction in lieu of multiplication, and division.

Taking the log of Equation 13.57 yields

$$\log \; |G(j\omega)| e^{j\phi(\omega)} = \log \; |G(j\omega)| + \log e^{j\phi(\omega)}$$

$$= \log \; |G(j\omega)| + j \; .4343 \; \phi(\omega)$$

The quantity, $j \; .434 \; \phi(\omega)$, is the imaginary part and in future discussion only the angle $\phi(\omega)$ will be used.

In feedback system work the unit commonly used for the logarithm of the magnitude is the decibel. The logarithm of the magnitude of a transfer function $G(j\omega)$ expressed in decibels is

$$20 \log \; |G(j\omega)| = (\quad) \text{ db} \qquad (13.58)$$

This quantity is often referred to as the log magnitude and is abbreviated Lm where

$$\text{Lm } G(j\omega) = 20 \log \; |G(j\omega)| \text{ db} \qquad (13.59)$$

Now, how does multiplication and division become addition and subtraction for Bode development? We will take the log of Equation 13.56 and multiply this by 20 which will give the amplitude ratio in decibels. The use of logarithms will allow us to add for multiplication and subtract for division.

The Lm of Equation 13.56 becomes

$$\begin{aligned} 20 \log |G(j\omega)| &= 20 \log K_n + 20\, m \log |j\omega| + 20 \log |1 + j\,\tau_1 \omega| \\ &+ 20 \log |1 + j\tau_2 \omega| + 20 \log \left|1 - \frac{\omega^2}{\omega_{n_1}^{\,2}} + j\, \frac{2\zeta\omega}{\omega_{n_1}}\right| \\ &- 20\, n \log |j\omega| - 20 \log |1 + j\tau_3 \omega| \\ &- 20 \log |1 + j\tau_4 \omega| - 20 \log \left|1 - \frac{\omega^2}{\omega_{n_2}^{\,2}} + j\, \frac{2\zeta\omega}{\omega_{n_2}}\right| \end{aligned} \qquad (13.60)$$

The associated phase angle of Equation 13.56 becomes,

$$\begin{aligned} \angle G(j\omega) &= \angle K_n + m\, 90° + \tan^{-1} \omega\tau_1 + \tan^{-1} \omega\tau_2 \\ &+ \tan^{-1} \frac{2\zeta\omega/\omega_{n_1}}{1 - \dfrac{\omega^2}{\omega_{n_1}^{\,2}}} - n\, 90° - \tan^{-1} \omega\tau_3 \\ &- \tan^{-1} \omega\tau_4 - \tan^{-1} \frac{2\zeta\omega/\omega_{n_2}}{1 - \dfrac{\omega^2}{\omega_{n_2}^{\,2}}} \end{aligned}$$

It is immediately obvious that a point by point solution of these equations for varying input frequency, ω, would be very tedious. A technique known as the Bode plot simplifies this process. Notice in Equation 13.60 that there are four types of factors in the open-loop transfer function $G(j\omega)$.

1. Constant term, K_n
2. Pole or zero at origin, $(j\omega)^{\pm n}$
   $(+n = \text{zero}, -n = \text{pole})$
3. Simple pole or zero, $(1 + j\omega\tau)^{\pm n}$
4. Quadratic pole or zero, $\left(1 - \frac{\omega^2}{\omega_n^2} + j2\zeta \frac{\omega}{\omega_n}\right)^{\pm n}$

The Bode plot uses semilog paper. Magnitude and phase angle are represented on the ordinate (linear scale) and frequency along the logarithmic scale as in Figures 13.24 and 13.25. Bode plot technique uses asymptotes and corrections to the asymptotes for each of the four types of factors listed above. All of the factors are individually plotted on the Bode diagram, and then are added and subtracted (taking advantage of logarithms) to achieve the composite curve. We will develop the technique for each of the four types of factors.

Constant term, K_n The magnitude of K_n in db is

$$20 \log |K_n| = \text{constant}$$

and the associated phase angle is

$$\text{Arg } (K_n) = 0^\circ \quad \text{or}$$

$$\text{Arg } (- K_n) = \pm 180^\circ$$

as shown in Figure 13.26.

The magnitude and phase angle are depicted respectively on Figures 13.24 and 13.25.

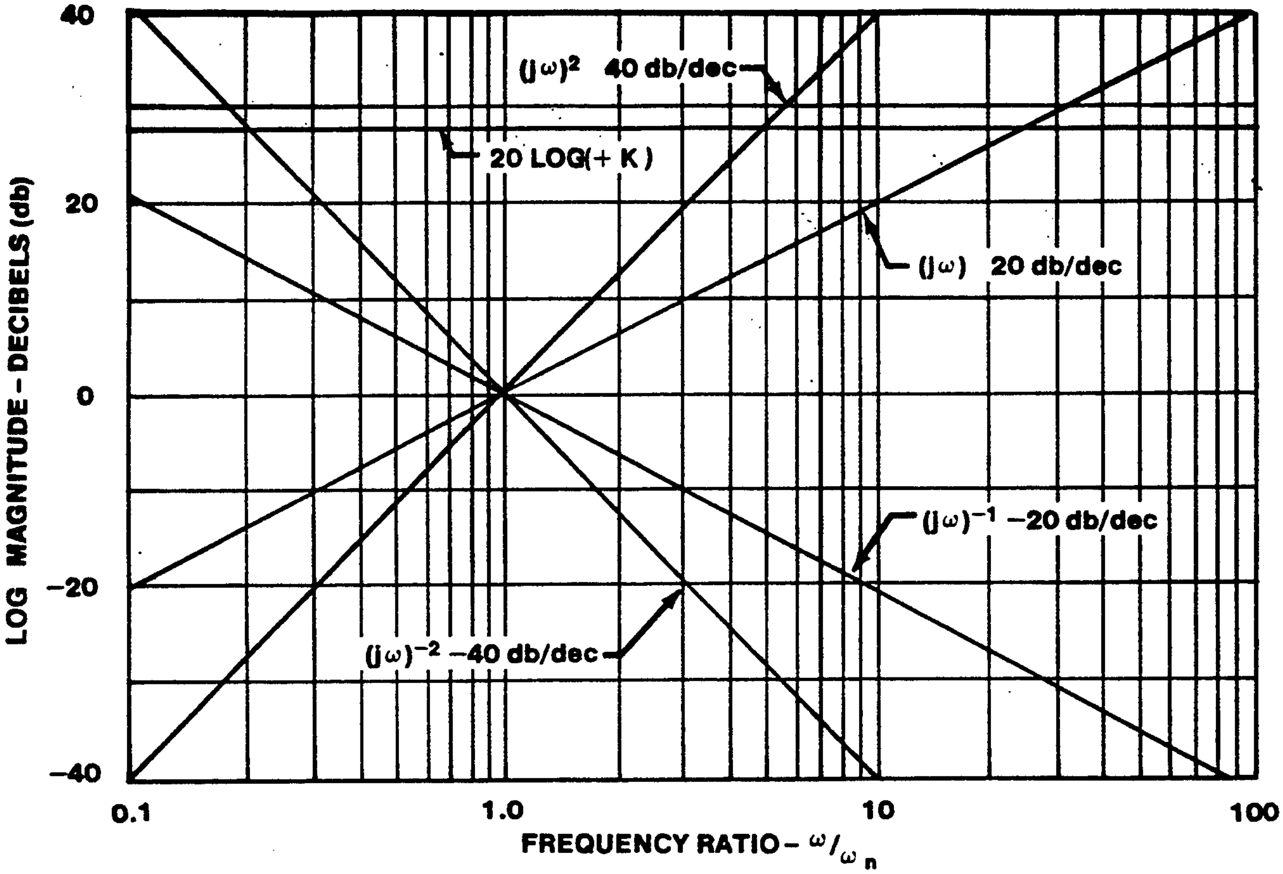


FIGURE 13.24. BODE MAGNITUDE PLOT OF $(j\omega)^{\pm n}$ AND K

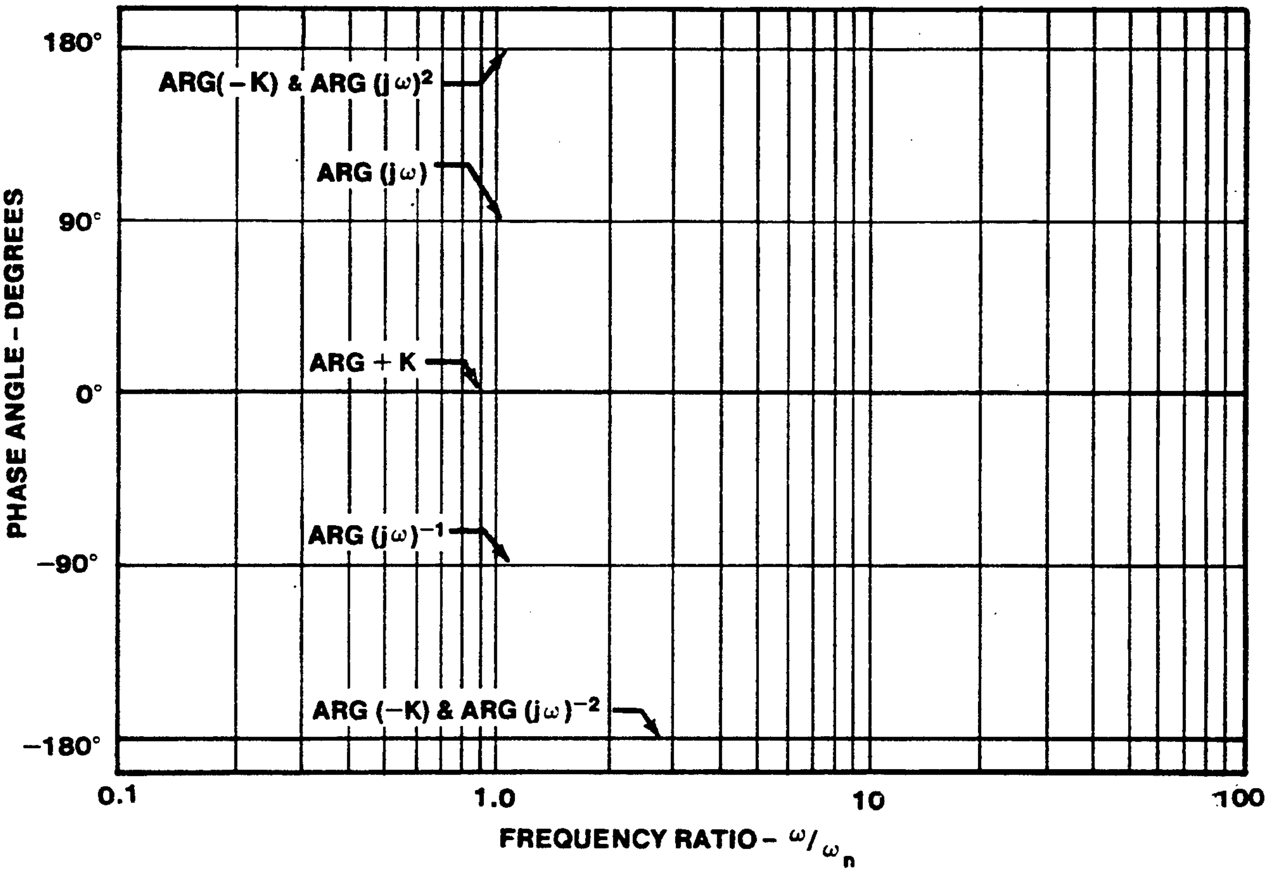


FIGURE 13.25. BODE PHASE ANGLE PLOT OF $(j\omega)^{\pm n}$ AND K

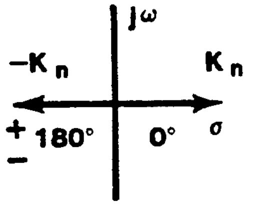


FIGURE 13.26. PLOT OF A CONSTANT ON THE COMPLEX PLANE

Pole or zero at origin, $(j\omega)^{\pm n}$. The magnitude of $(j\omega)^{\pm n}$ is

$$20 \log \left|(j\omega)^{\pm n}\right| = \pm 20 \text{ n} \log \omega_{db}$$

which is the equation of a straight line with slope of ±20 n db/decade.

A decade is a frequency band from f_1 to f_2 where $f_2/f_1 = 10$. If $f_1 = 4$ rad/s, then one decade above f_1 is : $f_2 = 40$ rad/s. The octave is also used as a frequency ratio and is a frequency band from f_1 to f_2 where $f_2/f_1 = 2$. If $f_1 = 4$ rad/s, then one octave above f_1 is : $f_2 = 8$ rad/s. The slope of ±20 n db/decade = ±6 n db/octave.

When $\omega = 1$, the term $20 \log \left|(j\omega)^{\pm n}\right| = 20 \log(1) = 0$. Therefore, the magnitude plot of a pole or zero at the origin is a straight line with a slope = ± 20n dB/decade, passing through the 0 db point at frequency $\omega = 1$ (Figure 13.24). The phase angle of $(j\omega)^{\pm n}$ is

$$\text{Arg}\left\{(j\omega)^{\pm n}\right\} = \pm n\ 90°$$

as shown in Figure 13.25.

Simple Pole or Zero, $(1 + j\omega\tau)^{\pm n}$ the magnitude of this term in db is expressed as

$$20 \log \left|(1 + j\omega\tau)\right|^{\pm n} = \pm 20 \text{ n} \log (1 + j\omega t)$$

at very low frequency (i.e., $\omega\tau << 1$) the magnitude of this curve is 0 db. At frequencies where $\omega\tau >> 1$ the magnitude asymptote has a slope of $\pm 20n$ db/decade. The 0 db asymptote and the $\pm 20n$ db/dec asymptote intersect at the corner frequency, $\omega_c = 1/\tau$, Figure 13.27.

The phase angle is expressed as

$$\text{Arg } (1 + j\omega\tau)^{\pm n} = \pm n \tan^{-1} \omega\tau$$

At $\omega = 0, \phi = 0^{\circ}$ and at

$\omega = \infty, \phi = \pm n\ \pi/2$ radians

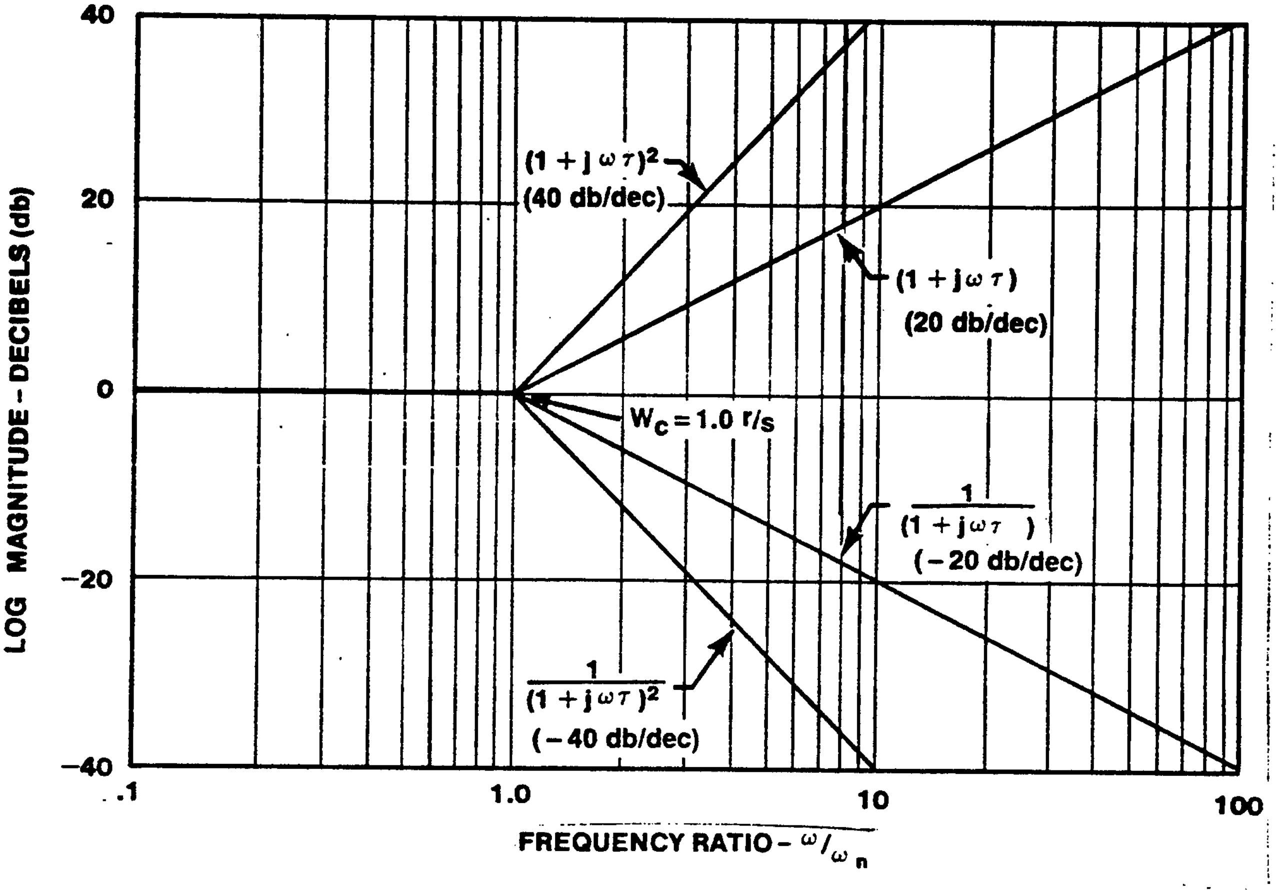


FIGURE 13.27. BODE MAGNITUDE PLOT OF TERM $(1 + j\omega\tau)^{\pm n}$

Table 13.2 shows the variation of the phase angle with normalized frequency ω/ω_c, for n = -1.

TABLE 13.2

PHASE ANGLE VARIATION WITH NORMALIZED FREQUENCY $\left[(1 + j\omega\tau)^{-1}\right]$

$\frac{\omega}{\omega_c}$	$\tan^{-1} \omega\tau$
0	0
.1	$- 5.7^{\circ}$
.5	-26.5°
1.0	-45.0°
2.0	-63.4°
10.0	-84.3°
∞	-90.0°

The following techniques are used to plot the $(1 + j\omega\tau)^{\pm n}$ factor.

1. Locate corner frequency, $\omega_c = 1/\tau$.

2. Draw $\pm$20n db/decade asymptotes through the corner frequency (+20n db/dec for zero terms and -20n db/dec for pole terms).

3. A straight line can be used to approximate the phase shift. The line is drawn from 0° at one decade below the corner frequency to n ($\pm90^{\circ}$) (+ for zero term, - for pole term) at one decade above the corner frequency. The maximum deviation using this approximation is about 6°. The specific phase angle values are shown in Table 13.2 and the appropriate corrections can be applied if desired. These corrections are shown in Figure 13.28.

4. The error to the magnitude curve (created by using the asymptote technique) can be determined analytically. First determine the error at the corner frequency $\omega_c = 1/\tau$.

$$\pm 20n \log \sqrt{1 + \omega_c^2 \tau^2} \text{ becomes } \pm 20n \log \sqrt{1 + \frac{\tau^2}{\tau^2}}$$

and

$$\pm 20n \log \sqrt{2} = \pm 10n \log 2 = \pm 3.01n \text{ db}$$

This shows that the asymptote can be corrected by adding approximately $\pm 3n$ db at the corner frequency. Likewise for a frequency one decade above the corner frequency,

$$\omega = 10\, \omega_c = \frac{10}{\tau}$$

and

$$\pm 20n \log \sqrt{1 + 100\, \tau^2/\tau^2} = \pm 20n \log \sqrt{101}$$

and

$$-10 \log 101 = -20.043 \text{ db (actual, for } n = -1)$$

Our straight line asymptote used -20 db so the total error at one decade is -.043 db. Similarly the error at $\omega_c/10$ can be found. At $\omega = 2\omega_c$ (one octave) $= 2/\tau$, the actual Im for $n = -1$ is

$$-20 \log \sqrt{1 + 4\, \tau^2/\tau^2} = -10 \log 5 = -6.9897 \text{ db}$$

The asymptote method produced a value of -6 db at this point, thus an error -.9897, or approximately -1.0 db.

Therefore, the straight line asymptotes can be made closer to the actual Im curve by applying a $\pm 3n$ db correction at ω_c and a $\pm 1n$ db correction 1 octave above and 1 octave below ω_c.

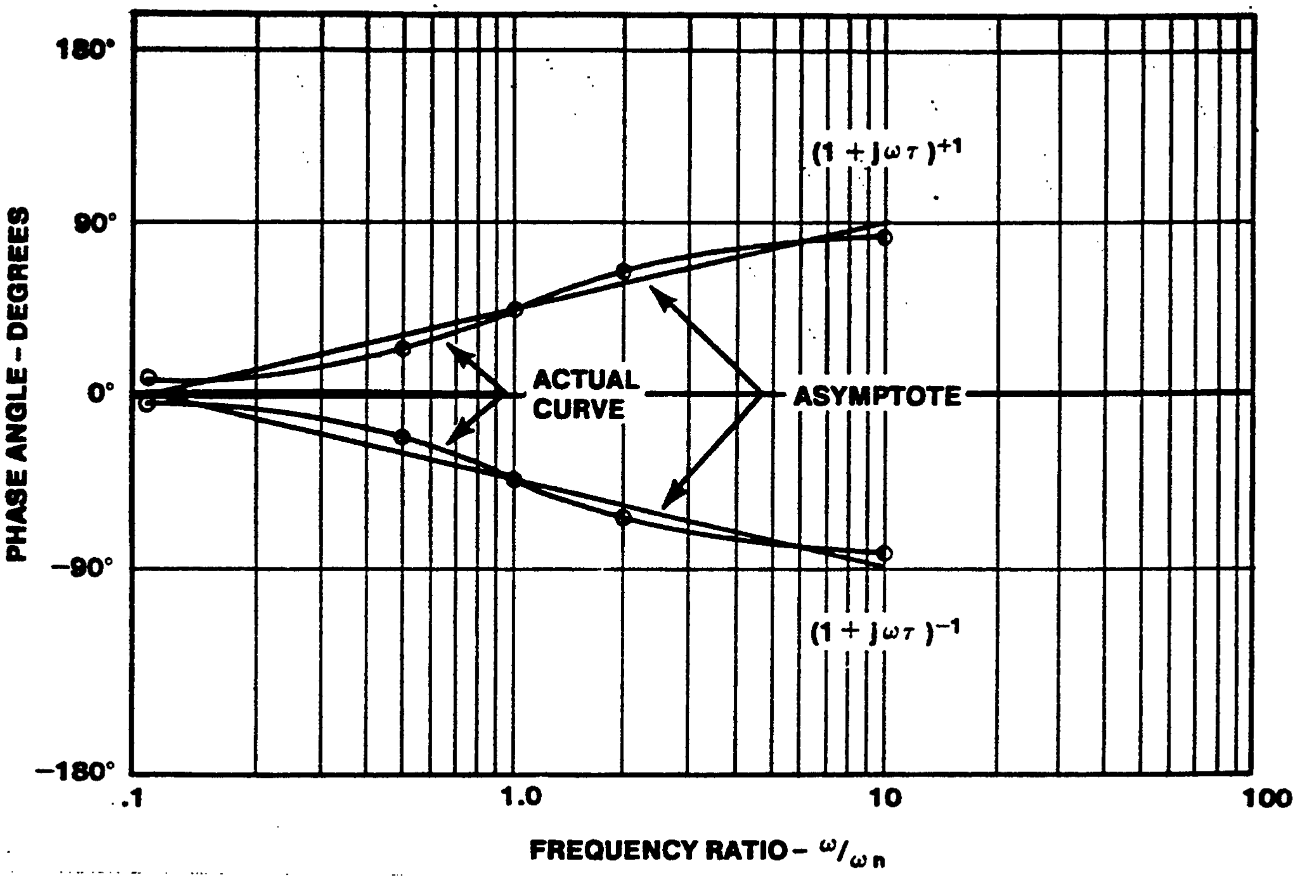


FIGURE 13.28. BODE PHASE ANGLE PLOT OF $(1 + j\omega\tau)^{\pm n}$

<u>Quadratic Term,</u> $[1 + \left(2\zeta/\omega_n\right) j\omega + (j\omega/\omega_n)^2]^{\pm n}$

Consider the quadratic term

$$G(s) = \frac{1}{\left(1 + \frac{2\zeta s}{\omega_n} + \frac{s^2}{\omega_n^{\,2}}\right)^n} \tag{13.61}$$

The log magnitude of $G(j\omega)$, Equation 13.61, in db is

$$20 \log G(j\omega) = -20\ n \log \sqrt{\left[1 - \left(\frac{\omega}{\omega_n}\right)^2\right]^2 + \left[2\zeta\left(\frac{\omega}{\omega_n}\right)\right]^2} \tag{13.62}$$

and the phase angle is

$$\operatorname{Arg} G(j\omega) = -n \tan^{-1} \left[\frac{2\zeta \frac{\omega}{\omega_n}}{1 - \left(\frac{\omega}{\omega_n}\right)^2} \right] \qquad (13.63)$$

If $\zeta > 1$, the quadratic term can be factored into two first-order factors plotted following the technique of the previous section. If $\zeta < 1$ the quadratic factors into a complex-conjugate pair and we plot the entire quadratic without factoring. The influence of the damping ratio, ζ, on the magnitude plot and phase angle plot is illustrated in Figure 13.29. From Figure 13.29 we see that the maximum value of the resonant peak is a function of ζ. The maximum value of the resonant peak is given by

$$M_r = \frac{1}{2\zeta \sqrt{1 - \zeta^2}} \qquad (13.64)$$

and the frequency at which this peak occurs is

$$\omega_r = \omega_n \sqrt{1 - 2\zeta^2} \qquad (13.65)$$

The asymptote technique will provide accurate curves provided corrections are applied.

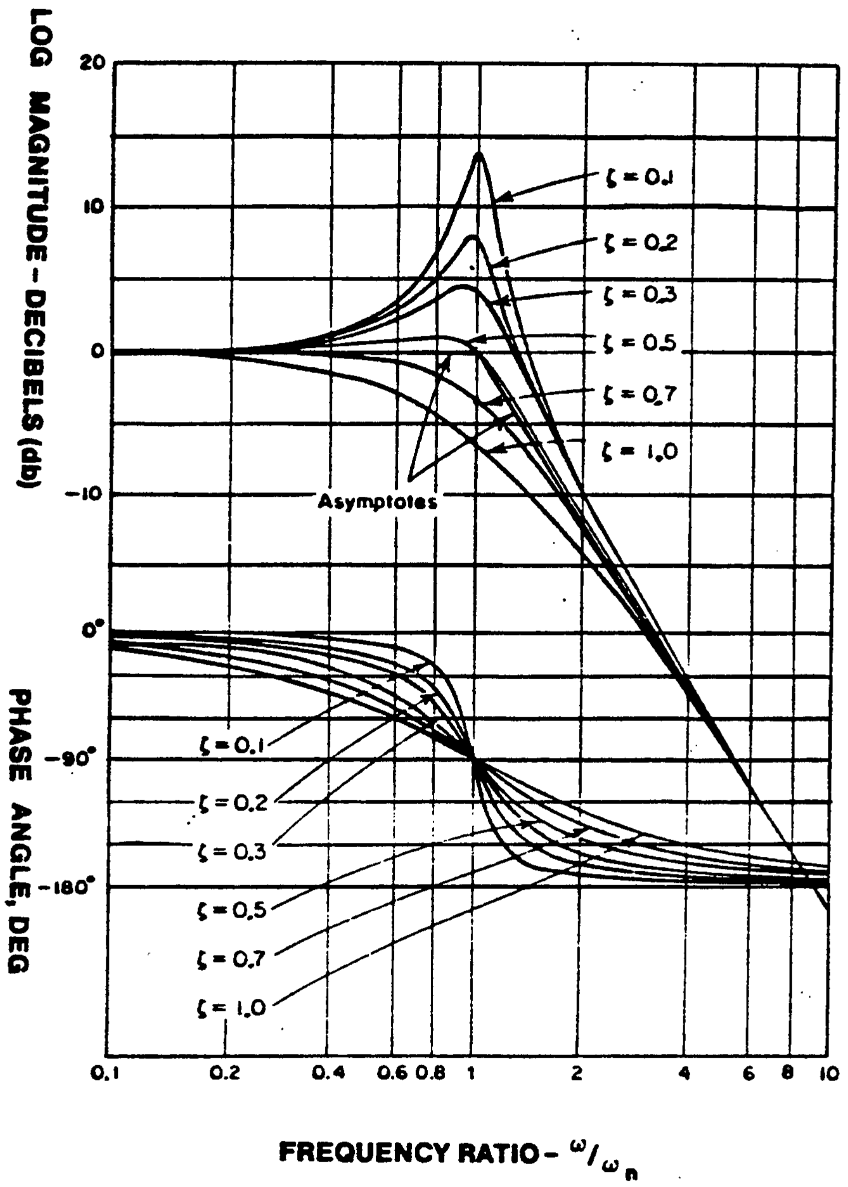


FIGURE 13.29. BODE DIAGRAM FOR $G(j\omega) = \left[1 + (2\zeta/\omega_n)\ j\omega + (j\omega/\omega_n)^2\right]^{-1}$

At very low frequency $\frac{\omega}{\omega_n} << 1$

$$20 \log G(j\omega) \approx -20 n \log 1 = 0 \text{ db}$$

Therefore, at low frequencies the asymptote for the quadratic term is a straight line with a slope of zero. At high frequency $\omega/\omega_n >> 1$, Equation 13.62 can be approximated as

$$20 \log G(j\omega) \approx -20 n \log \sqrt{\left(\frac{\omega}{\omega_n}\right)^4}$$

$$\approx -40 n \log \left(\frac{\omega}{\omega_n}\right) \text{ db}$$

The last equation represents the equation of a straight line with slope of -40n db/dec. If the quadratic is in the numerator (i.e., +n) the slope is positive. The two asymptotes intersect at ω_n, hence, ω_n is considered the corner frequency of the quadratic factor. The actual magnitude plot for the quadratic factor differs strikingly from its asymptotic plot in that the amplitude curve depends not only on ω_n, but also on ζ. From Figure 13.29, several values of Lm around ω_n can be plotted for a specific ζ to obtain an accurate magnitude plot.

The phase angle plot for a quadratic factor can be obtained by locating the ±90°n point at ω_n and obtaining a few points either side of ω_n for a specific value of ζ from Figure 13.29.

To summarize, the procedures for plotting the quadratic term are:

a. Determine the value of ζ and ω_n.

b. Plot the zero db asymptote from low frequencies to ω_n and a $\underline{+}40n$ db/dec asymptote beginning at ω_n.

c. Use the curves presented in Figure 13.29 to correct the asymptotes in the vicinity of the corner frequency.

d. At the corner frequency, ω_n, locate the $\pm 90^{\circ}n$ phase point. Using the curves of Figure 13.29 for the specific ζ, plot enough data points to permit sketching the phase angle curve.

When each of the four types of factors are plotted on the Bode plot, all the magnitude curves and phase angle curves are summed at different frequencies to complete the composite curves. The following problem will illustrate the simplicity of this technique.

<u>Example Problem:</u>

Given: $$G(s) = \frac{640s\ (s + 1000)}{(s + 10)\,(s^2 + 80s + 6400)}$$

where

$$2\zeta\omega_n = 80, \quad \omega_n = 80, \quad \zeta = 0.5$$

1. First put G(s) into the frequency response standard form.

$$G(j\omega) = \frac{(640)\,(1000)\,(j\omega)\,(1 + j\ .001\omega)}{(6400)\,(10)\,(1 + j\ .1\omega)\left(1 - \frac{\omega^2}{6400} + j\,\frac{80\omega}{6400}\right)}$$

where

$$K_n = \frac{(640)\quad(1000)}{(6400)\quad(10)} = 10$$

2. Find the corner frequencies where $\omega_c = 1/\tau$. For a quadratic term the natural frequency ω_n is the corner frequency ω_c.

   Zeros: $\omega_c = 1000$, +20 db/dec

   Poles: $\omega_c = 10$, -20 db/dec; $\omega_c = 80$, -40 db/dec

3. Plot the individual magnitude and phase angle terms on the Bode. Also $20 \log K_n = 20 \log 10 = 20$ db.

4. Apply the appropriate corrections at the corner frequencies.

5. Add the curves.

6. Figures 13.30 and 13.31 show the contribution of the separate factors and the composite curve for the example.

We must emphasize that the development on the Bode plot presented here is based on the steady-state frequency response of an open-loop system to a sinusoidal input. Techniques exist to arrive at the closed-loop frequency response, but these are beyond the scope of our study. The closed-loop frequency response graph is a plot of magnitude ratio, $M(j\omega) = C(j\omega)/R(j\omega)$, and phase angle, ϕ, versus frequency. One method of determining the closed-loop frequency response is by using the Nichol's chart (Reference 13.3).

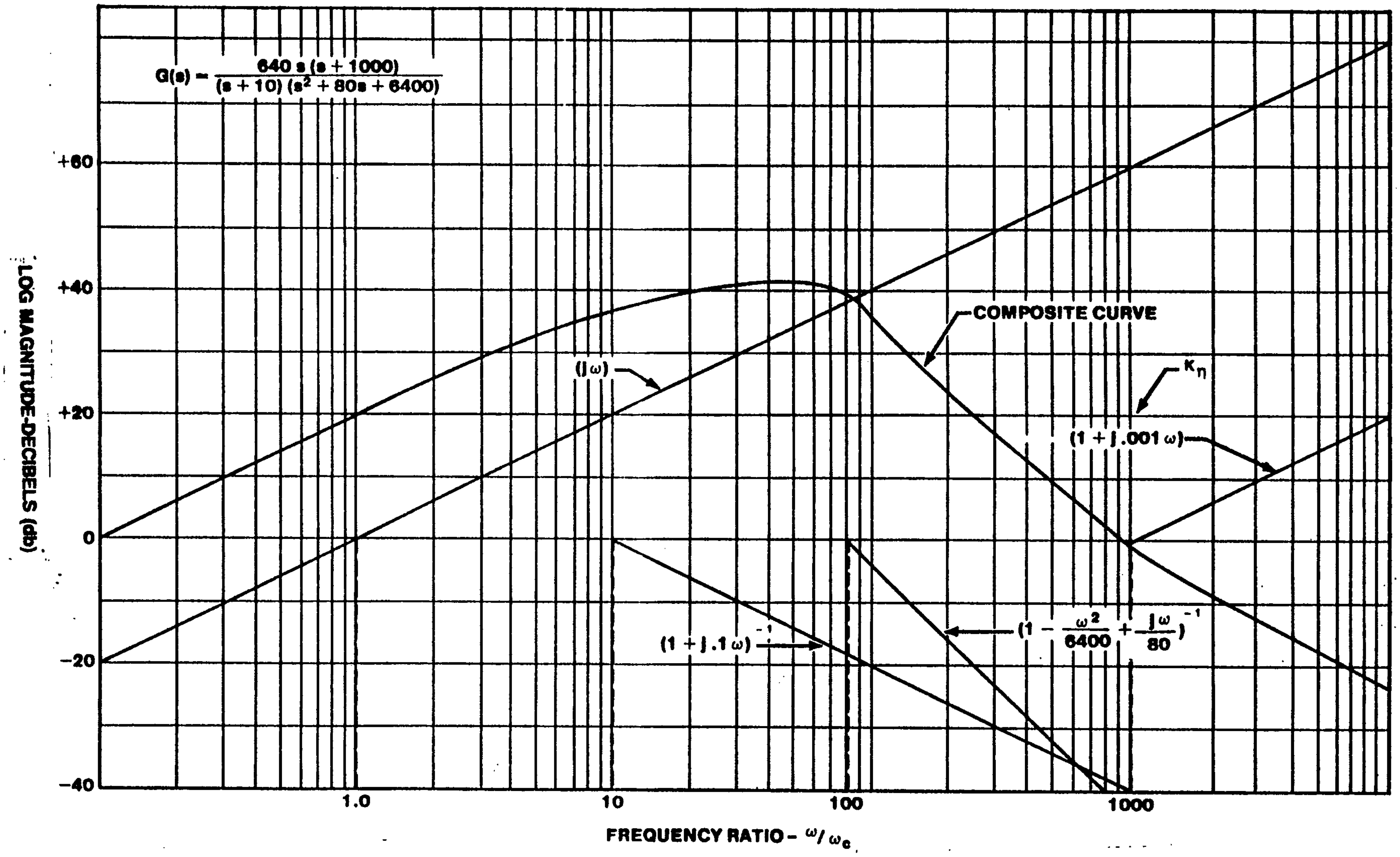


FIGURE 13.30. BODE LOG MAGNITUDE PLOT

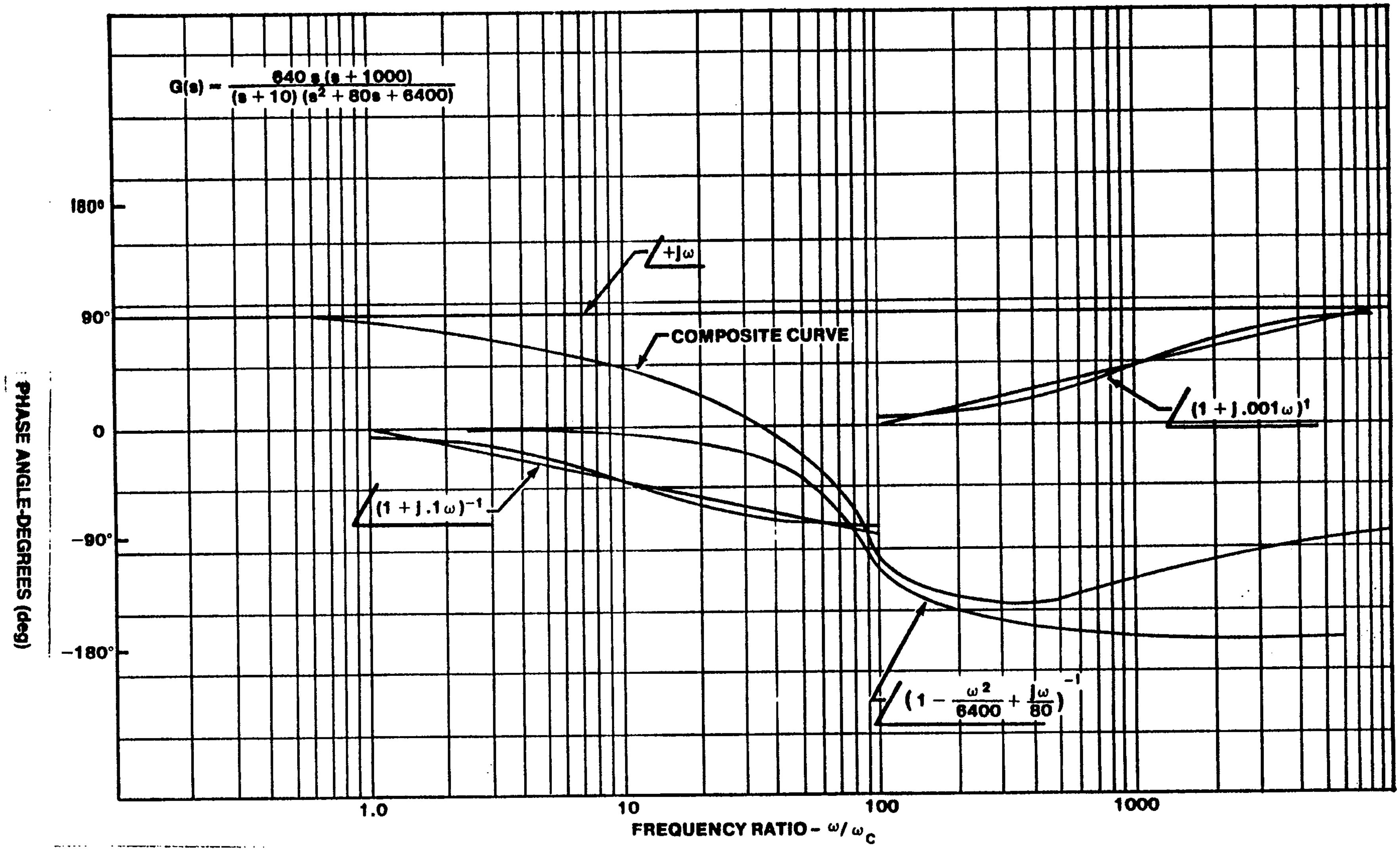


FIGURE 13.31. BODE PHASE ANGLE PLOT

13.7.3 Relative Stability

The relative stability of a closed-loop system can be determined by looking at the Bode plot of the open-loop transfer function, KGH(jω). Several terms are used to relate stability by the Bode Plot. The mathematical basis for these relationships comes from the Nyquist Stability Criteria. The terms are:

13.7.3.1 Gain Margin. Gain margin is the additional amount of gain, measured in decibels, that the magnitude ratio can be increased before the system goes unstable. The gain margin is defined as the reciprocal of the open-loop transfer function, GH(jω), evaluated at the frequency where the phase angle is -180°.

$$\text{Gain Margin} = 20 \log_{10} \frac{1}{GH(j\omega)}$$

This quantity is illustrated in Figure 13.32.

13.7.3.2 Phase Margin. Phase margin is the amount of phase shift, measured in degrees, that the phase angle curve can be displaced to produce instability in the system. Phase margin is measured at the frequency where the Im plot crosses the 0 db line.

$$\text{Phase Margin} = +180^{\circ} + \phi$$

$$(\phi = \text{phase angle measured at 0 db})$$

This quantity is illustrated in Figure 13.32.

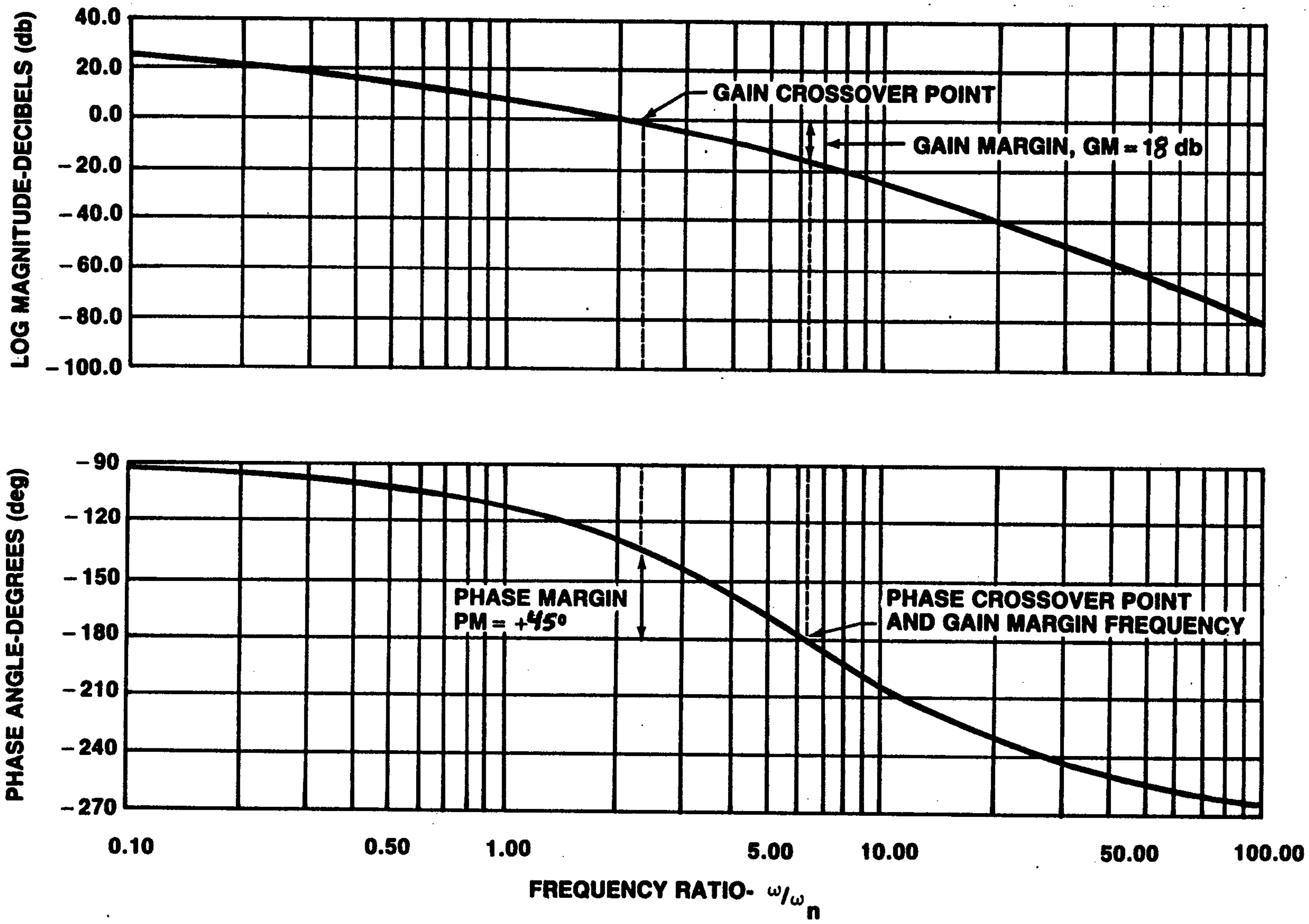


FIGURE 13.32. BODE PLOT RELATIVE STABILITY RELATIONSHIPS

Stability requires that the phase margin be positive, i.e., the phase angle at the 0 db crossover point must be greater than -180°.

13.7.3.3 The Gain Crossover Point is defined as the point or points where the magnitude curve crosses the zero db axis.

13.7.3.4 The Phase Crossover Point is the point on the Bode phase angle plot at which the phase angle is -180°. The frequency at which the phase crossover occurs is called the gain margin frequency.

13.7.4 Frequency Domain Specifications

There are several terms used to express the specifications of systems in the frequency domain. Although these terms are usually used to define the closed-loop response, they can also be used to express characteristics of the open-loop Bode plot.

13.7.4.1 Bandwidth (BW). The definition of bandwidth of a system depends on an accurate description of the problem. Normally the bandwidth is defined as the frequency at which the magnitude ratio $M(j\omega) = C(j\omega)/R(j\omega)$ has dropped to 70.7% of the zero frequency level or 3 db down from the zero frequency level as shown in Figure 13.33. This does not cover all cases in that the magnitude ratio at zero frequency may be low as in Figure 13.34. In this case the bandwidth is defined as the frequency range over which the magnitude ratio does not vary more than -3 db from its value at a specified frequency. For the purposes of this course the bandwidth can also be determined as the -3 db point on the Im plot of the open-loop system. This value should correspond closely to the bandwidth of the closed-loop system.

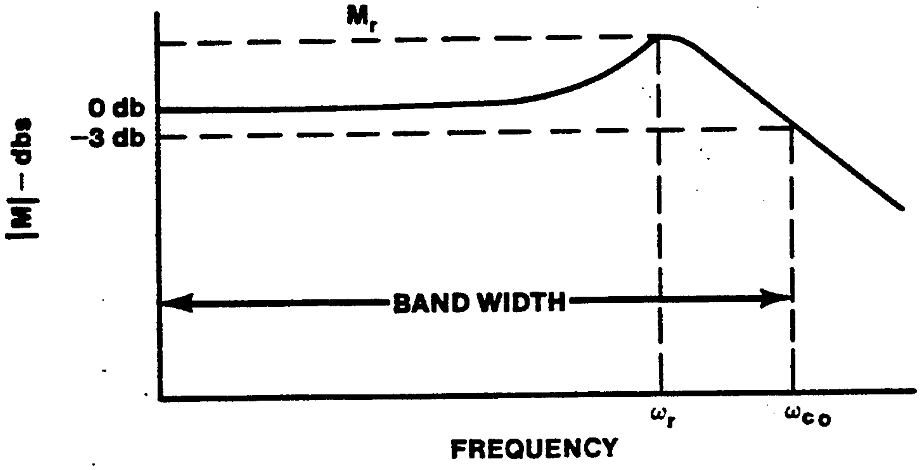


FIGURE 13.33. FREQUENCY DOMAIN CHARACTERISTICS

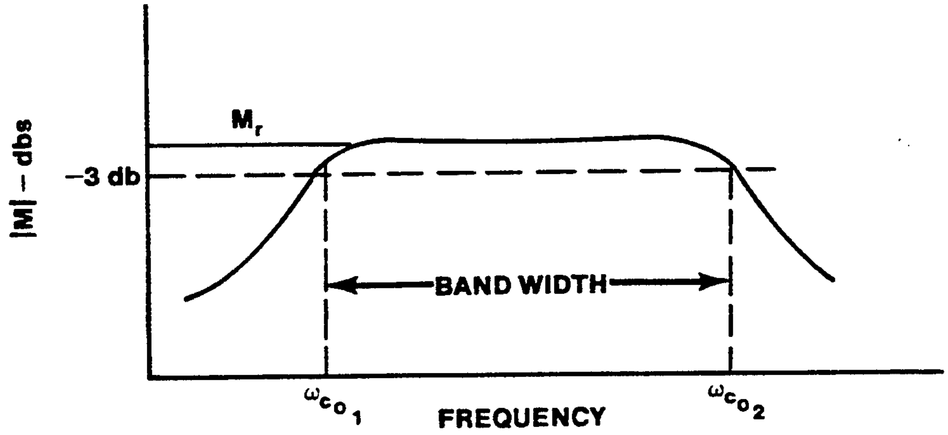


FIGURE 13.34. FREQUENCY DOMAIN CHARACTERISTICS

The frequency at which the -3 db point is reached is called the cutoff frequency, ω_{co}. Bandwidth is important for two reasons. First, it is indicative of the noise filtering characteristics of the system. System noise is always present and the bandwidth and the corresponding cutoff frequency dictate at what frequency the response and thus the noise will be filtered. Secondly, the bandwidth is a measure of the transient response properties of a system. A large bandwidth will allow higher frequencies to pass to the output and the system may be characterized by fast rise time and large overshoots. However, if the bandwidth is narrow, only low frequency signals are passed and the time response will generally be slow and sluggish.

13.7.4.2 Resonant Peak, M_r. If the system is of second-order or higher, it may have a resonant peak, M_r. For a second-order system there exists exact mathematical relationships between ζ, the damping ratio, and ω_r, the frequency at which M_r occurs. A higher-order system can often be approximated by a second-order system to simplify the solution. The resonant peak, M_r, Figure 13.33, is an indication of the relative stability of the system as a high value of M_r corresponds to a large overshoot in the time domain. Typical values of M_r for usable stable systems may vary from 1.1 to 1.5.

13.7.5 Experimental Method of Frequency Response

A Bode plot may be determined experimentally and ultimately will provide the system transfer function. The method depicted in Figure 13.35 will allow for measurement of the magnitude ratio and phase angle versus frequency.

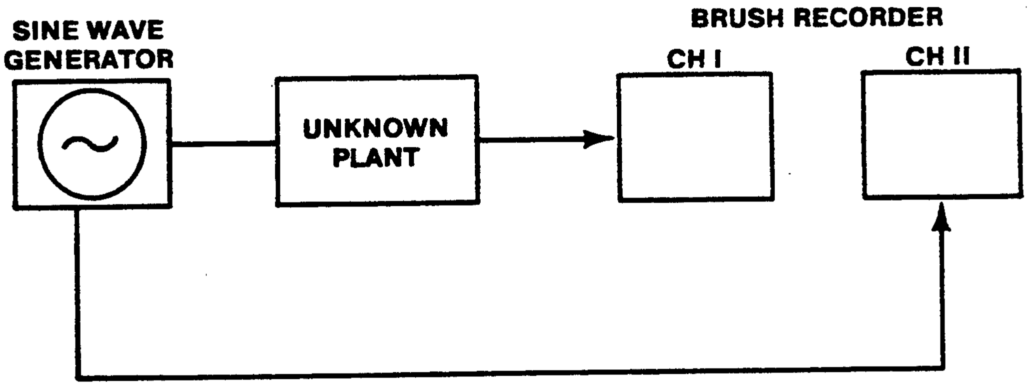


FIGURE 13.35. EXPERIMENTAL BODE TECHNIQUE

Once the Lm and phase angle versus frequency curves are plotted, asymptotes can be fitted on the curve and the corner frequencies determined. If a resonant peak occurs, use the techniques discussed previously to determine ζ and ω_n.

In our discussion, we have been talking about open-loop systems that do not have poles and/or zeros in the right-half of the s-plane (RHP). These systems are known as minimum phase systems. A nonminimum phase system is one which has an open-loop pole and/or zero in the RHP. A nonminimum phase factor is of the form

$$(1-j\zeta\omega)^{\pm n} \quad \text{or} \quad 1 + \frac{\omega}{\omega_n}{}^{2} \pm j2\zeta\frac{\omega}{\omega_n}{}^{\pm n}$$

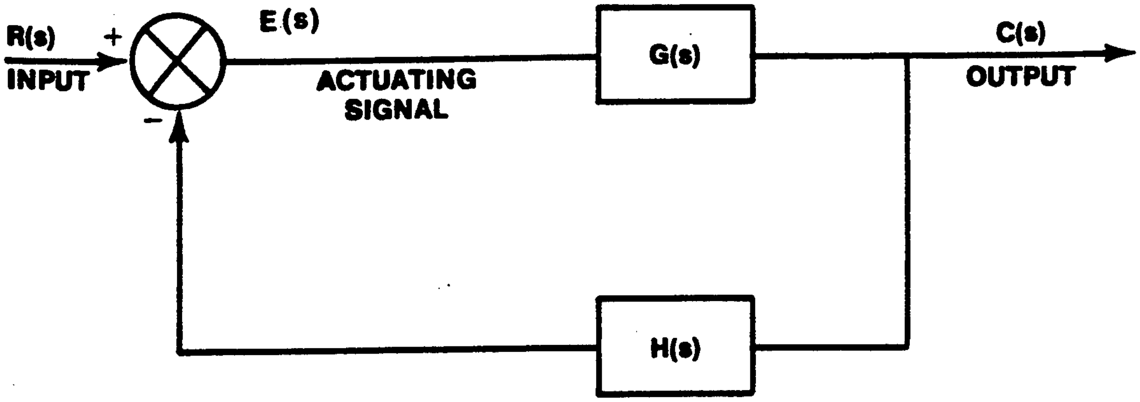
If a system is known to be minimum phase, only the Lm plot is required to fully determine the system transfer function; while both the Lm and phase angle plots are required to determine the system transfer function if the system is nonminimum phase. For example, consider the following transfer functions:

$$G_1(s) = \frac{(s+1)}{(s+10)}, \quad G_2(s) = \frac{(s+1)}{(s-10)}, \quad G_3(s) = \frac{(s-1)}{(s+10)}, \text{ and } G_4(s) = \frac{(s-1)}{(s-10)}$$

All four transfer functions have similar Lm plots, but their phase angle plots are all different ($G_1(s)$ is minimum phase; $G_2(s)$, $G_3(s)$, and $G_4(s)$ are all nonminimum phase).

13.8 CLOSED-LOOP TRANSFER FUNCTION

For reasons to be seen shortly, complex control systems are most often represented by a block diagram in the form of Figure 13.36 where the forward transfer function, G, and the feedback transfer function, H, are expressed as functions of s, the Laplace transform variable. The closed-loop transfer function of Figure 13.36 will now be developed in terms of the forward and feedback transfer functions for our first and second-order systems. No transfer function present in the feedback loop is called unity feedback (in Figure 13.5, H(s) = 1).



FORWARD TRANSFER FUNCTION — G(s)
FEEDBACK TRANSFER FUNCTION — H(s)
OPEN-LOOP TRANSFER FUNCTION (OLTF) — G(s) H(s)
CLOSED-LOOP TRANSFER FUNCTION (CLTF) — C(s)/R(s)
ACTUATING SIGNAL — E(s)

FIGURE 13.36. STANDARD FORM OF FEEDBACK CONTROL SYSTEM

In order to find G(s) of the first-order system (Figure 13.5), equate torques and assume only damping present

$$\mu\varepsilon = b\,\dot{\theta}$$

taking the Laplace transform and noting that G(s) = θ(s)/E(s) we find

$$\mu E(s) = bs\theta(s)$$

$$G(s) = \frac{\theta}{E}(s) = \frac{1}{\frac{b}{\mu}s}$$

again letting

$$\tau = \frac{b}{\mu}$$

$$G(s) = \frac{1}{\tau s} \qquad (13.66)$$

The same procedure for the case including inertia yields the following torque summation:

$$\mu\varepsilon = I\ddot{\theta} + b\dot{\theta}$$

$$E(s) = \frac{I}{\mu} s^2 \theta(s) + \frac{bs}{\mu}\theta(s)$$

$$G(s) = \frac{\theta}{E}(s) = \frac{\mu}{Is^2 + bs} \qquad (13.67)$$

Thus, we have the forward transfer function for our two systems. Referring to Figure 13.36, we will now derive an expression for the closed-loop transfer function in terms of G and H.

$$G(s) = \frac{C(s)}{E(s)}$$

and also

$$E(s) = R(s) - H(s)\,C(s)$$

substituting

$$G(s)\ [R - H(s)\ C(s)] = C(s)$$

$$R(s)\ G(s) = C(s)\ [1 + GH(s)] \qquad (13.68)$$

the system closed-loop transfer function becomes

$$\boxed{\boxed{\frac{C}{R}(s) = \frac{G(s)}{1 + GH(s)}}} \qquad (13.69)$$

<u>This is a very important relationship which should immediately be committed to memory</u>.

As noted previously, the block diagram of Figure 13.36 is the standard form of the feedback control system. When in this form the closed-loop transfer function can be quickly found by Equation 13.69. But most important, the characteristic equation of the system from which the transient response is determined is immediately evident. Referring to Equations 13.68 and 13.69 we will show that the characteristic equation is found from the denominator of the right hand term in Equation 13.69.

$$1 + GH(s) = 0$$

The characteristic equation is merely 1 plus the system open-loop transfer function, GH(s), which is directly available.

But first, applying our closed-loop transfer function expression to the first-order system we use the forward transfer function of Equation 13.66. Since the system has unit feedback, H(s) = 1 and

$$GH(s) = \frac{1}{\tau s}$$

Therefore, using Equation 13.69 yields

$$\frac{C}{R}(s) = \frac{G(s)}{1 + GH(s)} = \frac{\frac{1}{\tau s}}{1 + \frac{1}{\tau s}}$$

$$\frac{C}{R}(s) = \frac{1}{\tau s + 1}$$

which is consistent with Equation 13.16, the transfer function derived from the equation of motion of the entire system.

In the case of the second-order system, since H(s) is unity

$$GH(s) = \frac{\mu}{Is^2 + bs}$$

For this system, using Equation 13.69

$$\frac{C}{R}(s) = \frac{G(s)}{1 + GH(s)} = \frac{\frac{\mu}{Is^2 + bs}}{1 + \frac{\mu}{Is^2 + bs}}$$

$$\frac{C}{R}(s) = \frac{1}{\frac{Is^2}{\mu} + \frac{bs}{\mu} + 1} \qquad (13.70)$$

is again consistent with the more direct method leading to Equation 13.18. The denominator of Equation 13.70 is the characteristic equation introduced by Equation 13.12.

Transfer functions are written to describe either whole systems or parts of systems using the appropriate differential equation. When control systems described by block diagrams, are reduced to some standard form, they quickly yield both the transfer function of the entire closed-loop system and its characteristic equation.

We will now discuss the technique of manipulating control systems in block diagram notation to obtain the desired form.

13.9 BLOCK DIAGRAM ALGEBRA

It was seen that the simplification resulting from the use of operational calculus is further increased when transfer functions and block diagrams are introduced. The special methods of predicting the transient response of a system without solving its equation of motion are most conveniently employed when the block diagram is of the form of Figure 13.36.

In practice, individual transfer functions are written for each integral unit of a more complex system. For example, the system of Figure 13.37 represents the pitch axis of an aircraft autopilot where the input is the commanded pitch attitude and the output the actual aircraft attitude. The autopilot, the elevator servo, and the aircraft itself are described separately in G_1, G_2, and G_3 respectively. As long as it is realized that transformed quantities are used the G(s) can be discarded and only G used.

The system of Figure 13.37 can be simplified by combining the inner loop into a single transfer function. If we let G_4 be the closed-loop transfer function of the inner loop we have

$$G_4 = \frac{G_2}{1 + G_2}$$

Figure 13.37 can then be redrawn as shown in Figure 13.38. This diagram is then further reduced by noting that

$$G_1 = \frac{\delta_i}{E},\ G_4 = \frac{\delta_e}{\delta_i},\ G_3 = \frac{C}{\delta_e}$$

and

$$G_1G_4G_3 = \frac{\delta_i}{E}\ \frac{\delta_e}{\delta_i}\ \frac{C}{\delta_e} = \frac{C}{E}$$

Denoting

$$G_5 = G_1G_4G_3 = \frac{G_1G_2G_3}{1 + G_2}$$

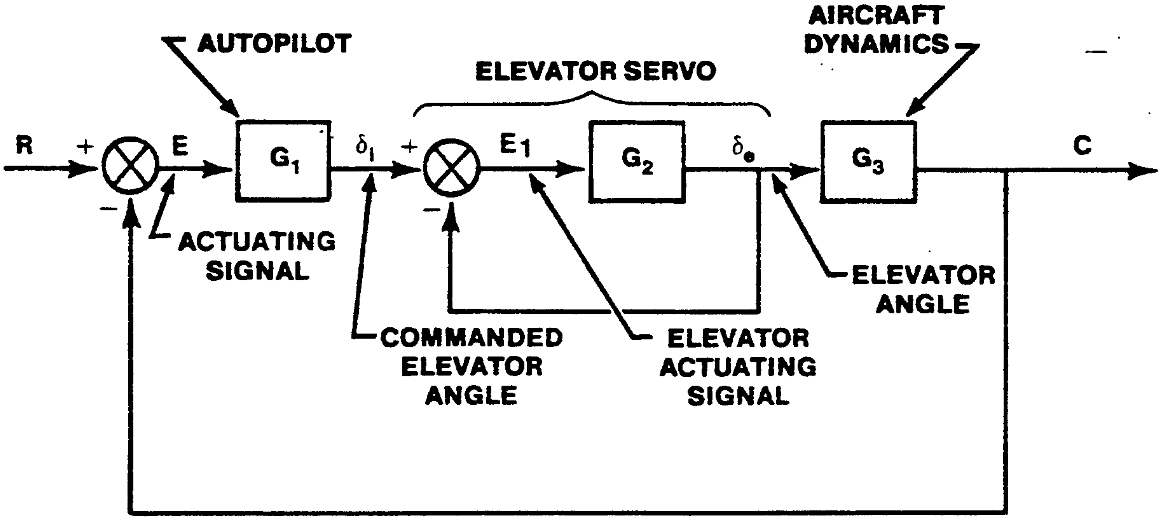


FIGURE 13.37. AIRCRAFT PITCH AXIS CONTROL SYSTEM

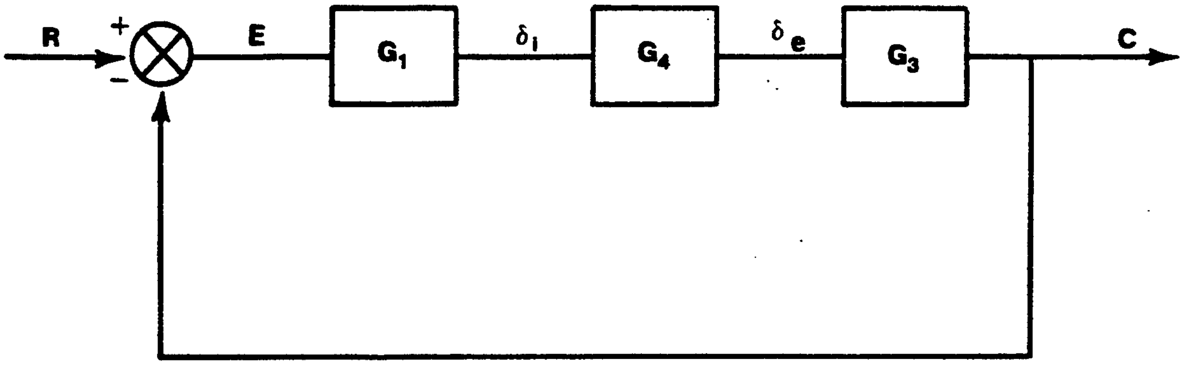


FIGURE 13.38. (FIGURE 13.37 REDUCED)

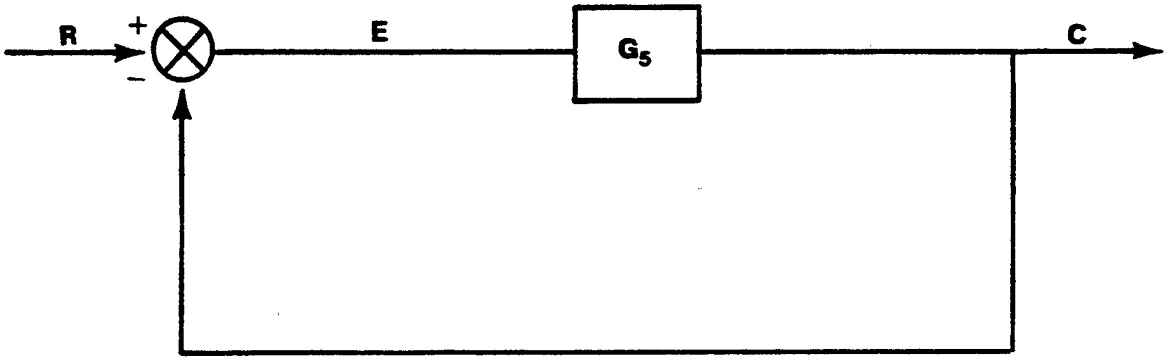
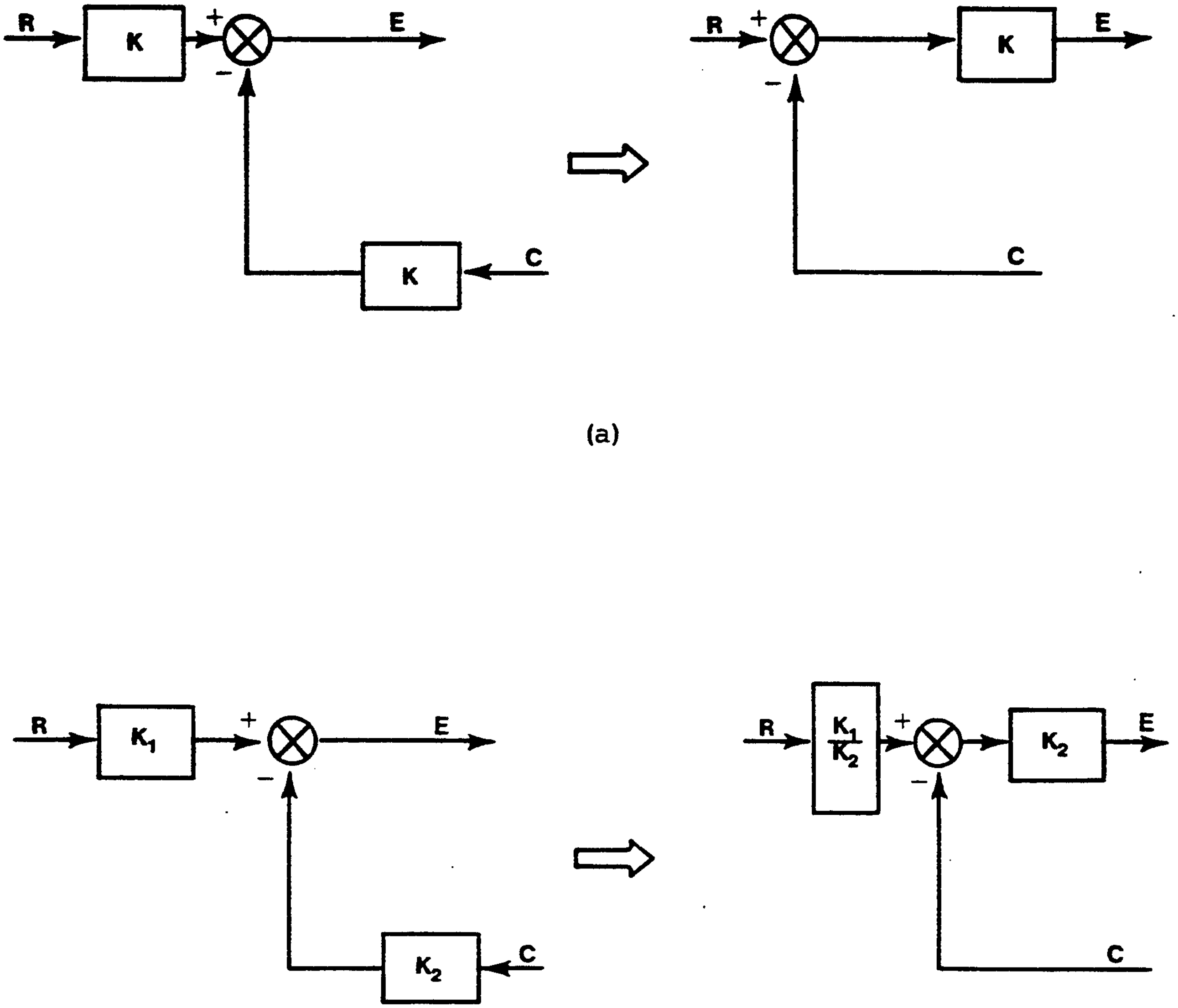


FIGURE 13.39. (FIGURE 13.37 FURTHER REDUCED)

We have, finally, the control system described in the proper form in Figure 13.39.

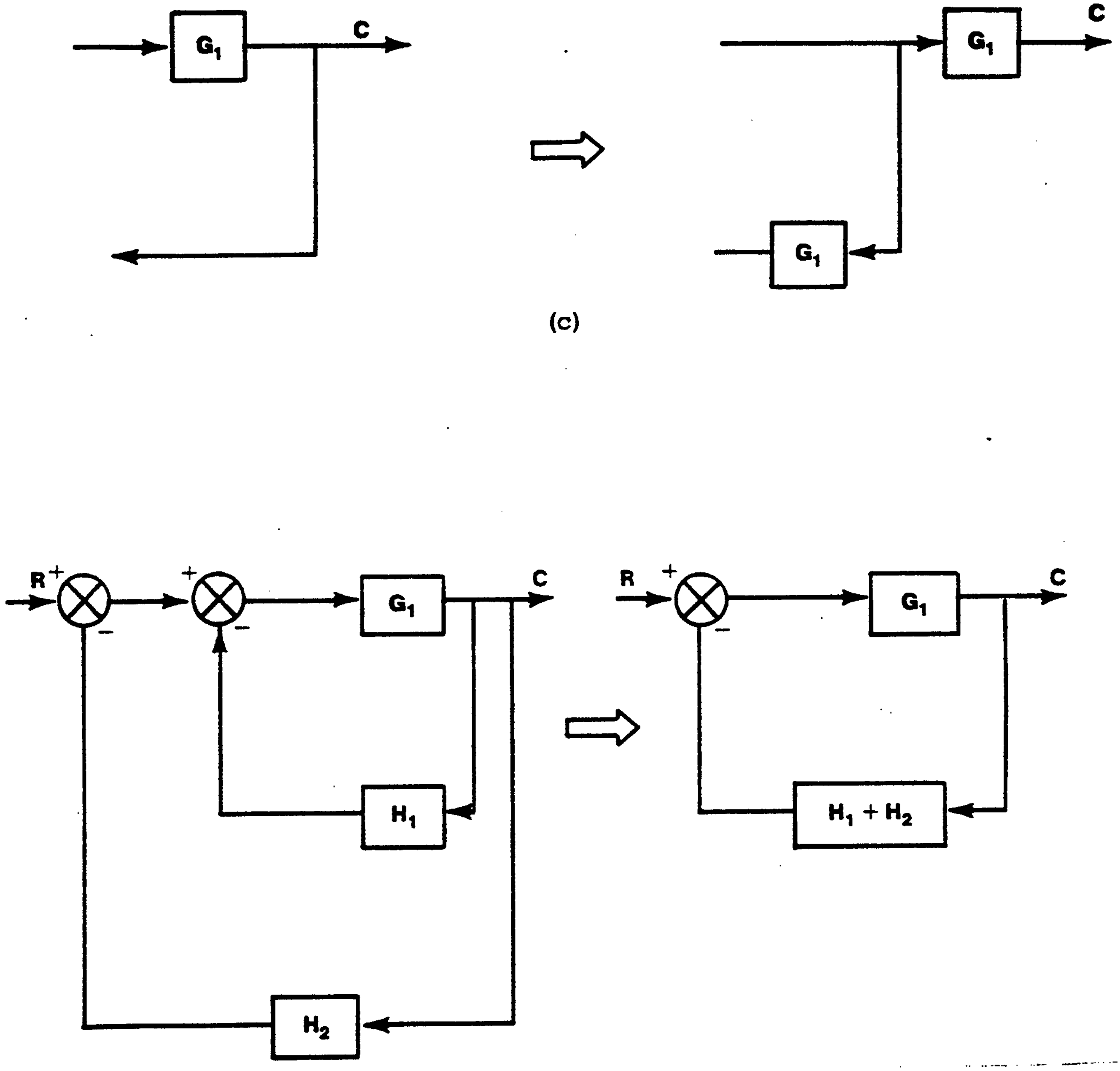
The following block diagram identities, Figure 13.40, will assist in manipulating complex control systems into the standard form for analysis.



(a)

(b)

FIGURE 13.40. BLOCK DIAGRAM IDENTITIES



(c)

(d)

FIGURE 13.40. CONT. BLOCK DIAGRAM IDENTITIES

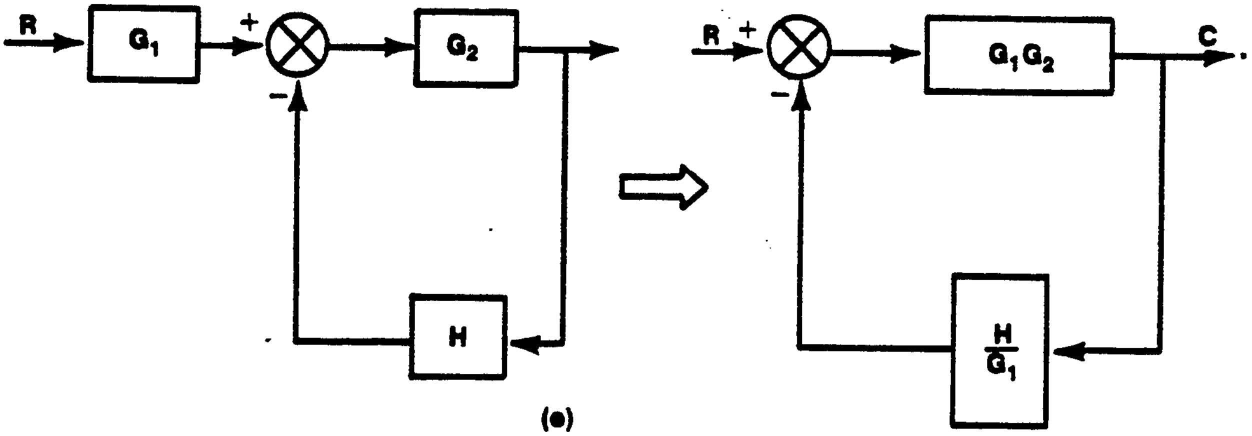


FIGURE 13.40. CONT. BLOCK DIAGRAM IDENTITIES

13.10 STEADY-STATE PERFORMANCE

The steady-state accuracy of a system is of considerable importance and is often related in terms of the steady-state error. Figures of merit for steady-state performance are the error constants, K_p, K_v, and K_a often referred to as the position, velocity and acceleration error constants.

A technique used to indicate the steady-state performance of a system is to classify the system by "Type". The number of free or pure integrators in the forward loop is the Type system (system must be stable and represented by unity feedback). For a specified input function, an n-type system will produce a mathematically predictable steady-state error. Consider the unity feedback system in Figure 13.41. A unity feedback system is used in this development since we will be relating the performance as a function of the steady-state error, $e(t)_{ss}$, where

$$e(t)_{ss} = r(t)_{ss} - c(t)_{ss}$$

For this relationship to be valid in this development, the reference input r(t) and the control variable c(t) must be dimensionally the same and must be to the same scale.

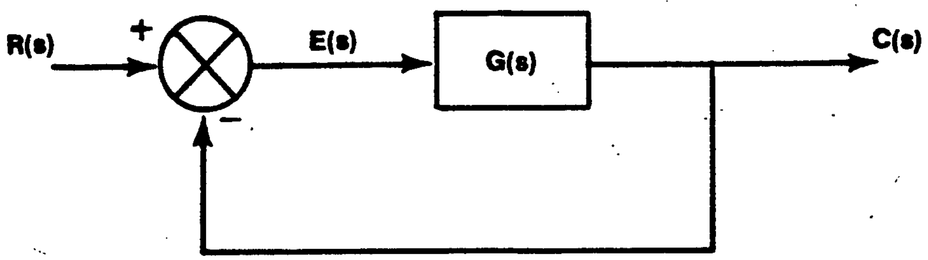


FIGURE 13.41. UNITY FEEDBACK SYSTEM

$$E(s) = R(s) - C(s)$$

and

$$C(s) = E(s)\,G(s)$$

therefore

$$E(s) = R(s) - E(s)\,G(s)$$

$$E(s)\,[1 + G(s)] = R(s)$$

$$\frac{E(s)}{R(s)} = \frac{1}{1 + G(s)}$$

The error signal E(s) is a function of the plant, G(s), and the reference input R(s).

G(s) can be represented by

$$G(s) = \frac{K_n(\tau_1 s + 1)(\tau_2 s + 1) \text{ ----}}{\underset{\uparrow\text{Defines "Type"}}{s^n}\,(\tau_a\, s + 1)(\tau_b\, s + 1)\left(\frac{s^2}{\omega_n^2} + \frac{2\zeta s}{\omega n} + 1\right) \text{ ----}}$$

where $n = 0, 1, 2, \ldots.$ For a Type 0 system $n = 0$; i.e., no free integrators in the forward loop. G(s) must be expressed in the above form to properly evaluate the overall gain, K_n, of the transfer function. This gain is often referred to as the "DC" gain or "type" gain.

We will seek to show the relationships between the inputs, the n-type system and the steady-state error. First consider the general error and apply a step, ramp, and a parabolic input.

$$E(s) = \frac{R(s)}{1 + G(s)}$$

13.10.1 Step Input

Let $r(t) = Ru_{-1}(t)$ and $R(s) = R/s$ and apply the final value theorem. Recall for any function F(s).

$$\lim_{s \to 0} s\,F(s) = \lim_{t \to \infty} f(t) \qquad E(s) = \frac{1}{1 + G(s)} R(s)$$

$$e(t)_{ss} = \lim_{s \to 0} \frac{s\,R/s}{1 + G(s)}$$

$$= \frac{R}{1 + \lim_{s \to 0} G(s)}$$

K_p Position Error Constant. The position error constant, Kp, is defined as

$$K_p = \lim_{s \to 0} G(s)$$

Therefore,

$$e(t)_{ss} = \frac{R}{1 + K_p}$$

For a Type 0 system

$$G(s) = K_0 \frac{(\text{---})(\text{---})}{(\text{---})(\text{---})}$$

$$K_p = \lim_{s \to 0} G(s) = \lim_{s \to 0} K_0 \frac{(\text{---})(\text{---})}{(\text{---})(\text{---})}$$

$K_p = K_0$, the overall gain of the transfer function.

For a Type 0 system a step input yields

$$e(t)_{ss} = \frac{R}{1 + K_0}$$

and the steady-state error is represented graphically in Figure 13.42

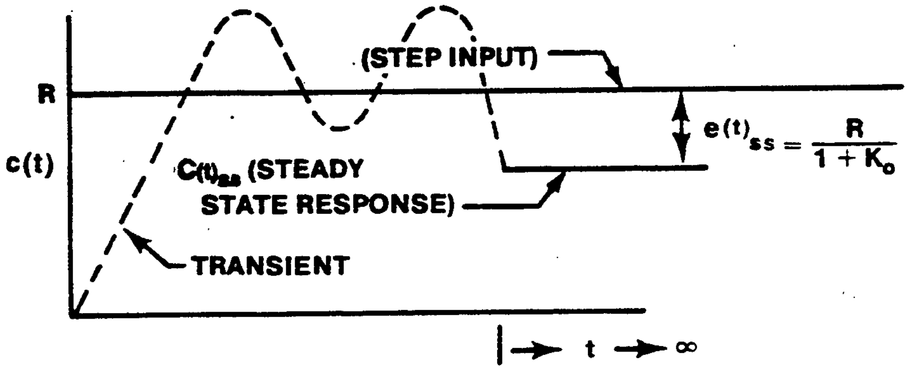


FIGURE 13.42. STEADY-STATE ERROR, TYPE 0 SYSTEM - STEP INPUT

For a Type 1 system

$$G(s) = \frac{K_1(\text{---})(\text{---})}{s\ (\text{---})(\text{---})}$$

$$K_p = \lim_{s \to 0} G(s) = \lim_{s \to 0} \frac{K_1 (---)(---)}{s\ (---)(---)}$$

$$K_p = \infty$$

Therefore,

$$e(t)_{ss} = \frac{R}{1 + \infty} = 0$$

The resulting error from a step input into a Type 1 system is zero. Similarly for a Type 2 and 3 system with a step input, the resulting error is zero.

13.10.2 Ramp Input

Consider a ramp input $r(t) = Ru_{-1}(t)$, $R(s) = R/s^2$

Therefore,

$$E(s) = \frac{R/s^2}{1 + G(s)}$$

$$e(t)_{ss} = \lim_{s \to 0} \frac{s\ R/s^2}{1 + G(s)}$$

$$= \lim_{s \to 0} \frac{R}{s + s\ G(s)}$$

$$= \frac{R}{\lim_{s \to 0} s\ G(s)}$$

K_v, Velocity Error Constant. The velocity error constant, k_v is defined as

$$K_v = \lim_{s \to 0} s\ G(s)$$

Therefore,

$$e(t)_{ss} = \frac{R}{K_v}$$

which is the error in displacement (of the output) due to a ramp input.

For a Type 0 system

$$K_v = \lim_{s \to 0} s\, K_0 = 0,$$

and the resulting steady-state error is infinite (Figure 13.43)

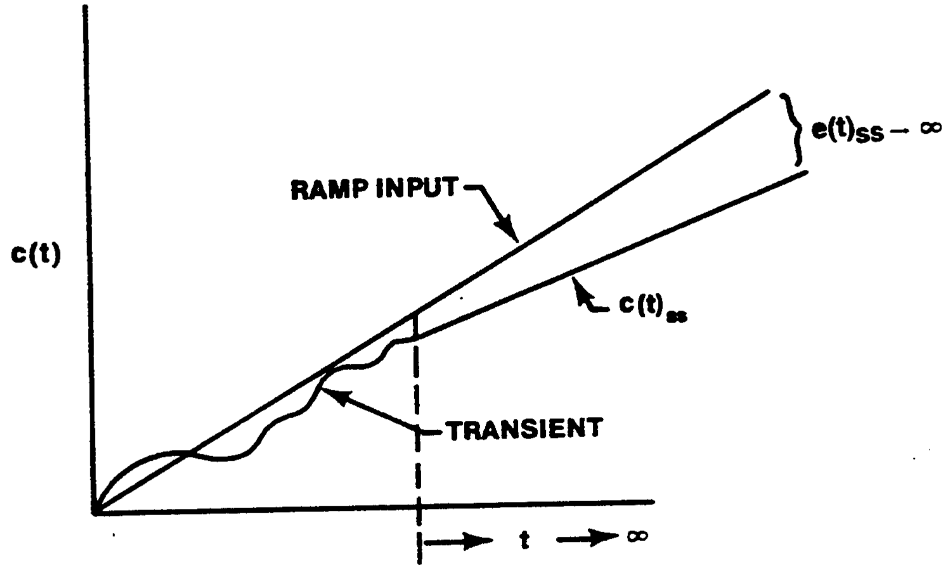


FIGURE 13.43. STEADY-STATE ERROR - TYPE "0" SYSTEM, RAMP INPUT

For a Type 1 system

$$K_v = \lim_{s \to 0} \frac{s\, K_1}{s} = K_1,$$

the overall gain of the transfer function. The resulting $e(t)_{ss} = R/K_1$. Therefore,

$$e(t)_{ss} = R/K_1$$

Figure 13.44 illustrates a Type 1 system with a ramp input.

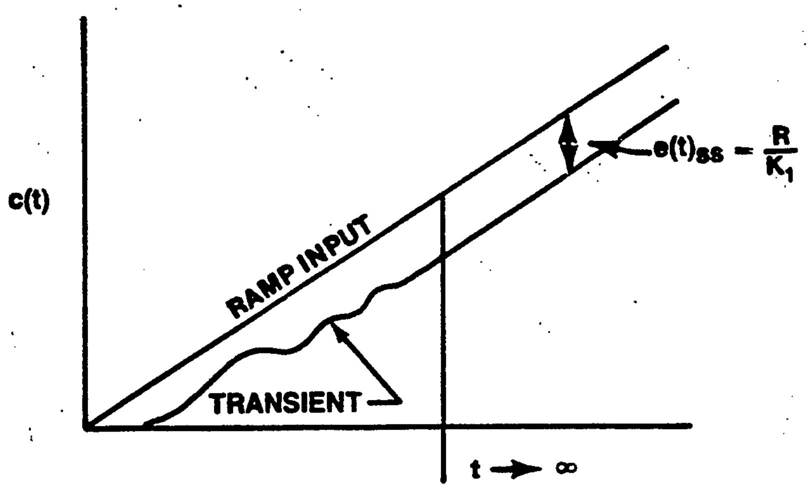


FIGURE 13.44. STEADY-STATE RESPONSE OF A TYPE 1 SYSTEM WITH A RAMP INPUT

For Type 2 and Type 3 systems $K_v = \infty$ and the resulting steady-state error is zero.

13.10.3 Parabolic Input

Consider the input $r(t) = Rt^2 u_{-1}(t)/2$, $R(s) = R/s^3$

$$e(t)_{ss} = \lim_{s\to 0} \frac{s\frac{R}{s^3}}{1 + G(s)}$$

$$= \lim_{s\to 0} \frac{R}{s^2 + s^2 G(s)}$$

$$= \frac{R}{\lim\limits_{s\to 0} s^2 G(s)}$$

K_a, Acceleration Error Constant. The acceleration error constant, K_a, is defined as

$$k_a = \lim_{s\to 0} s^2 G(s)$$

The steady-state error,

$$e(t)_{ss} = \frac{R}{K_a}$$

is the error in displacement (of the output) due to an acceleration type input.

For a Type 0 system

$$K_a = \lim_{s \to 0} s^2 K_0 = 0$$

and for a Type 1 system

$$K_a = \lim_{s \to 0} \frac{s^2 K_1}{s} = 0$$

For Type 0 and Type 1 systems a parabolic input will result in a parabolic output with the steady-state error, $e(t)_{ss}$, increasing to infinity (Figure 13.45).

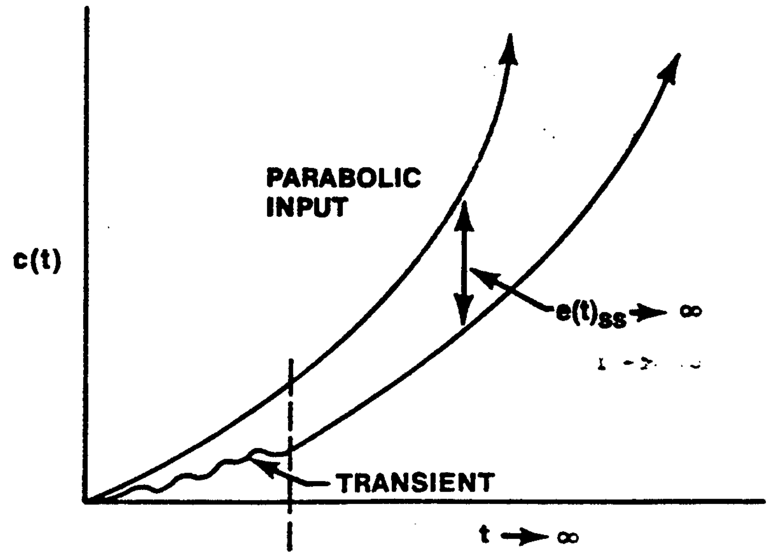


FIGURE 13.45. STEADY-STATE RESPONSE OF TYPE 0 AND TYPE 1 SYSTEMS TO A PARABOLIC INPUT

For a Type 2 system

$$K_a = \lim_{s\to 0} \frac{s^2 K_2}{s^2} = K_2$$

the overall gain of the transfer function.

The steady-state error, $e(t)_{ss}$, is equal to R/K_2 (a constant) (Figure 13.46).

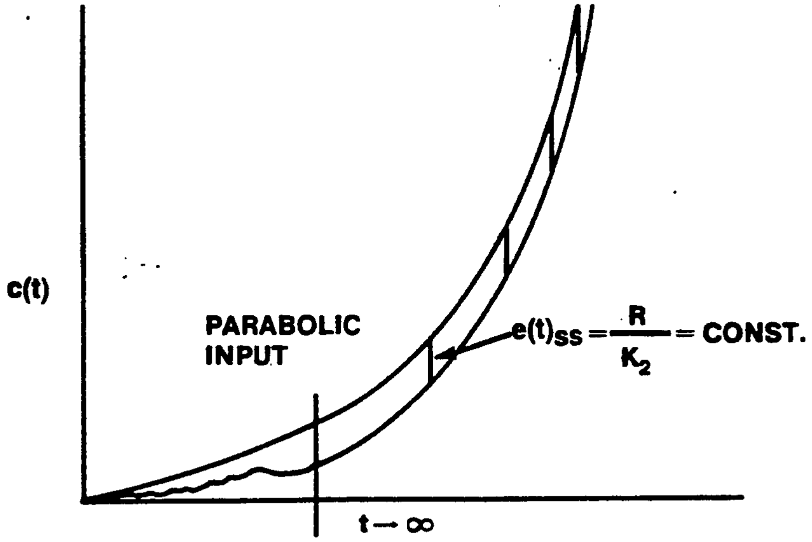


FIGURE 13.46. STEADY-STATE RESPONSE OF A TYPE 2 SYSTEM TO A PARABOLIC INPUT

For a Type 3 system

$$K_a = \lim_{s \to 0} \frac{s^2 K_3}{s^3} = \infty$$

Therefore the steady-state error is zero.

The information that has been developed is presented in tabular form in Table 13.3.

TABLE 13.3

STEADY-STATE ERROR

Type System	Error Constants			Steady-State Error		
	Step	Ramp	Parabolic	Step Input	Ramp Input	Parabolic Input
	$K_p =$	$K_v =$	$K_a =$	$e(t)_{ss} = \frac{R}{1 + K_p}$	$e(t)_{ss} = \frac{R}{K_v}$	$e(t)_{ss} = \frac{R}{K_a}$
0	K_0	0	0	$e(t)_{ss} = \frac{R}{1 + K_0}$	∞	∞
1	∞	K_1	0	= 0	$e(t)_{ss} = \frac{R}{K_1}$	∞
2	∞	∞	K_2	= 0	= 0	$e(t)_{ss} = \frac{R}{K_2}$
3	∞	∞	∞	= 0	= 0	= 0

13.10.4 Steady-State Response of the Control Variables

The foregoing discussion has been looking at the steady-state error, based on a specific input to a known plant, G(s). It is also interesting to look at the steady-state value of the control variable, $c(t)_{ss}$, for a known steady-state error signal, $e(t)_{ss}$.

Consider again the following equation:

$$G(s) = \frac{C(s)}{E(s)} = \frac{K_n\ (\tau_1\ s + 1)(\tau_2 s + 1)\ \ldots}{s^n\ (\tau_a\ s + 1)(\tau_b\ s + 1)\ \ldots}$$

Rewriting yields

$$E(s) = \frac{(\tau_a s + 1)\ (\tau_b s + 1)\ \ldots}{K_n\ (\tau_1\ s + 1)\ (\tau_2\ s + 1)\ \ldots}\ s^n\ C(s)$$

Applying the final value theorem

$$e(t)_{ss} = \lim_{s\to 0} s\,E(s) = \lim_{s\to 0}\left[\frac{s\ (\tau_a\ s + 1)(\tau_b\ s + 1)\ \ldots}{K_n(\tau_1\ s + 1)(\tau_2\ s + 1)\ \ldots}\ s^n\ C(s)\right]$$

$$= \lim_{s\to 0} s\ \frac{[s^n\ C(s)]}{K_n}$$

Recall the differential theorem

$$[D^n\ c(t)] = s^n\ C(s)$$

with initial conditions equal to zero.

Applying the final value theorem to the differential theorem yields

$$\lim_{s\to 0} s\ [s^n\ C(s)] = D^n\ c(t)_{ss}$$

We may now write

$$e(t)_{ss} = \frac{D^n\ c(t)_{ss}}{K_n}$$

or

$$K_n\ e(t)_{ss} = D^n\ c(t)_{ss}$$

From this equation and the characteristics of the systems as shown in Table 13.3, the following conclusions are drawn regarding the steady-state response:

a. A type 0 system is one in which a constant actuating signal maintains a constant value of the output, i.e.,

$$K_0\ e(t)_{ss} = c(t)_{ss}$$

b. A Type 1 system is one in which a constant actuating signal maintains a constant rate of change of the output, i.e.,

$$K_1\ e(t)_{ss} = D\ c(t)_{ss}$$

c. A Type 2 system is one in which the second derivative of the output is maintained constant by a constant actuating (error signal) i.e.,

$$K_2\ e(t)_{ss} = D^2\ c(t)_{ss}$$

13.10.5 Determining System Type and Gain From the Bode Plot

System type and gain can be obtained from a Bode Plot with the system in unity feedback form. The slope of the low frequency portion of the Lm curve determines the system type: a 0 db/dec slope represents a Type 0 system, -20 db/dec a Type 1 system, and a -40 db/dec slope a Type 2 system. The system gain (K_n) is determined from the Lm plot by projecting a vertical line from $\omega = 1.0$ rad/sec to the low frequency asymptote (or its projection) and reading across horizontally the Lm value. This Lm value represents

$$\text{Lm}\ G(j\omega) = 20\ \log\ K_n$$

Depending on system type, $K_n = K_0$, K_1, or K_2. Figure 13.47 illustrates how the Bode Plot is used in the manner described.

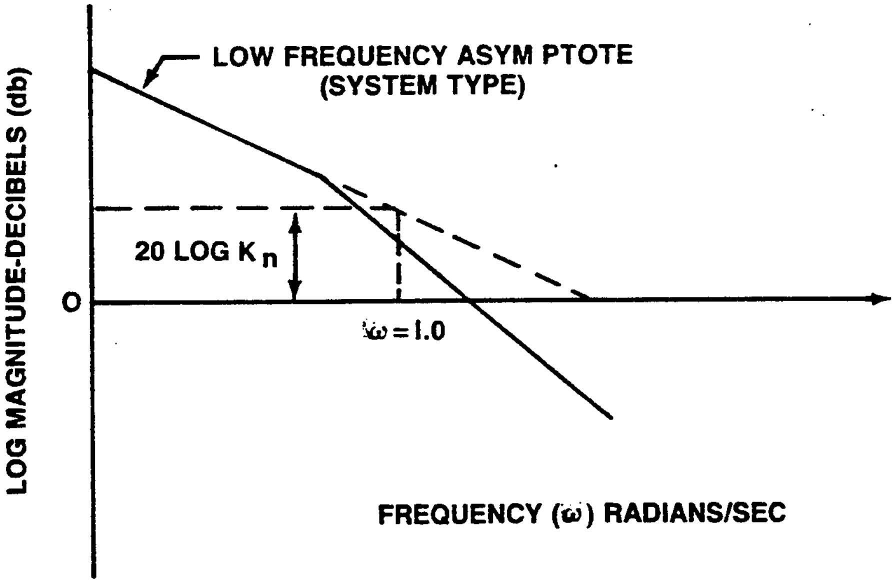


FIGURE 13.47. SYSTEM TYPE AND GAIN FROM A BODE PLOT

13.10.6 Summary

The static error constant can be used to quickly determine the ability of control system to follow a specific input. The application of error constants is not limited to systems with inputs classified as one of the three basic types of test signals. For linear systems, the concept can easily be extended to systems with inputs that can be represented by a polynomial, i.e.,

$$r(t) = R\left(1 + t + \frac{t^2}{2}\right)u_{-1}(t)$$

The steady-state error is

$$e(t)_{ss} = \frac{R}{1 + K_p} + \frac{R}{K_v} + \frac{R}{K_a}$$

a superposition of the errors due to each input signal component acting alone.

The chief advantage to the foregoing approach to steady-state response is the ease and timeliness of arriving at the answer. The chief disadvantage of the error constant approach is that only one of the constants has a finite value which is not zero or infinity for a particular n-type system. In cases where the steady-state error is a function of time, the error constant approach only gives an answer of infinity for $e(t)_{ss}$ and does not provide an indication of how the error varies with time. Even though the steady-state error may turn out to be infinite, for an actual problem, the input may be applied for a finite time, thus the error will be finite. This finite error may be well within the specifications.

It would appear desirable to select a large value of K_n, the overall gain of the transfer function, to minimize the steady-state error; but not without a penalty. Too large a value of K_n may force the system unstable. As we will see when we get to root locus analysis, an adjustment of the system gain effects both the natural frequency and the damping ratio for a closed-loop system with complex poles. In many cases the exact value of K_n which results in unstable system operation may be found by analysis. Routh's criterion, for example, will provide the value and the application of the criterion may be found in the literature (Reference 1).

13.11 ROOT LOCUS

An accurate prediction of the system's performance can be obtained by deriving the differential equation of a control system and then determining its solution. This approach is not feasible, however, for any but the simplest system. Not only is the direct solution method extremely tedious, but if the response does not meet the required specifications, no indication is given of how to improve its performance.

The aim of the design engineer is to predict the performance of the system without solving its equations of motion. Also, he would like the analysis to indicate how to modify the system in order to produce the desired response characteristics. Several methods are available which both predict stability and indicate the type of compensation required. Of those, the Bode plot has already been discussed and the root locus will be discussed next. Another technique is Nyquist criterion. The theory and application of root locus will be described.

Definition: The root locus is a plot of the roots of the characteristic equation of the closed-loop system as the gain is varied from zero to infinity. The definition itself presents the underlying theory of the root locus method. The primary objective is to determine system stability. This leads to another question. What determines stability? The answer, is the transient solution, which is determined from the roots of the characteristic equation, which cannot have positive real parts and be stable.

The general approach used in the development of the root locus technique will be to plot a root locus for a simple system the hard way, i.e., successive analytic solutions for the roots of the characteristic equation for selected values of static loop gain, K. Then the significance of the root locus will be discussed for the simple system, and then for any system. Lastly, some rules will be developed which permit quick plots to be drawn using relatively little labor.

13.11.1 Poles and Zeros

This section will define what is meant by poles and zeros and also discuss their relationship in functions of interest here. Consider the system represented by the block diagram in Figure 13.48.

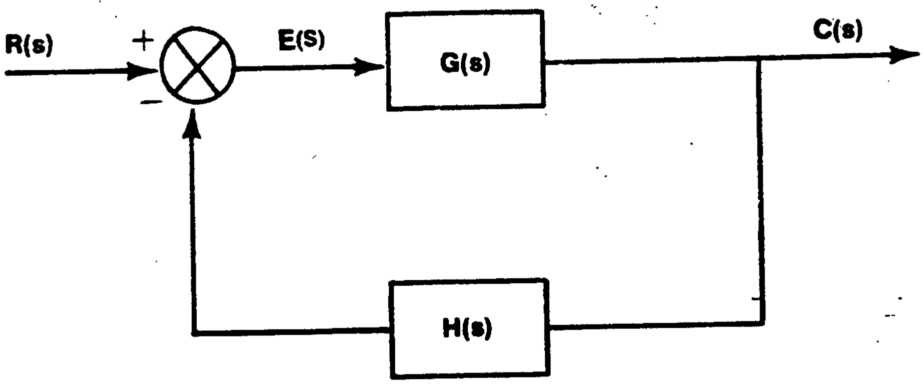


FIGURE 13.48. CLOSED-LOOP SYSTEM

where

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1 + GH(s)}$$

and where in general

$$G(s) = \frac{K_n (\tau_1 s + 1)(\tau_2 s + 1)(\ldots)}{s^n (\tau_a s + 1)(\tau_b s + 1)(\ldots)} = K_n G'(s) = K_n \frac{N_G(s)}{D_G(s)} \qquad (13.71)$$

and

$$H(s) = K_h \frac{(\tau_\alpha s + 1)(\tau_\beta s + 1)(\ldots)}{(\tau_I s + 1)(\tau_{II} s + 1)(\ldots)} = K_h H'(s) = K_h \frac{N_H(s)}{D_H(s)} \qquad (13.72)$$

where the numbers $\tau_1, \tau_2, \ldots; \tau_a, \tau_b, \ldots; \tau_\alpha, \tau_\beta, \ldots; \tau_I, \tau_{II}, \ldots$ may be real, complex or zero.

We now define two new terms:

ZERO — A zero of a function (like G(s)) is a value of s that makes that function zero.

POLE — A pole of a function (like G(s)) is a value of s that makes that function go to infinity.

For Example:

$$s = -\frac{1}{\tau_1} \text{ is a zero of } G(s)$$

$$s = -\frac{1}{\tau_\alpha} \text{ is a zero of } H(s)$$

$$s = -\frac{1}{\tau_a} \text{ is a pole of } G(s)$$

$$s = -\frac{1}{\tau_{II}} \text{ is a pole of } H(s)$$

In terms of the s-plane, (Figure 13.49), this means that there are values that cause G(s) and H(s) and incidentally their product G(s) H(s) to be zero and to be infinite. Figure 13.50 is a plot of the function G(s) H(s).

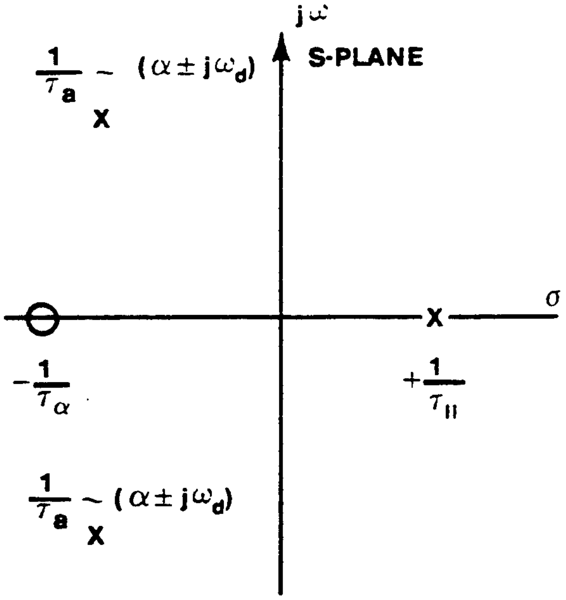


FIGURE 13.49. s-PLANE

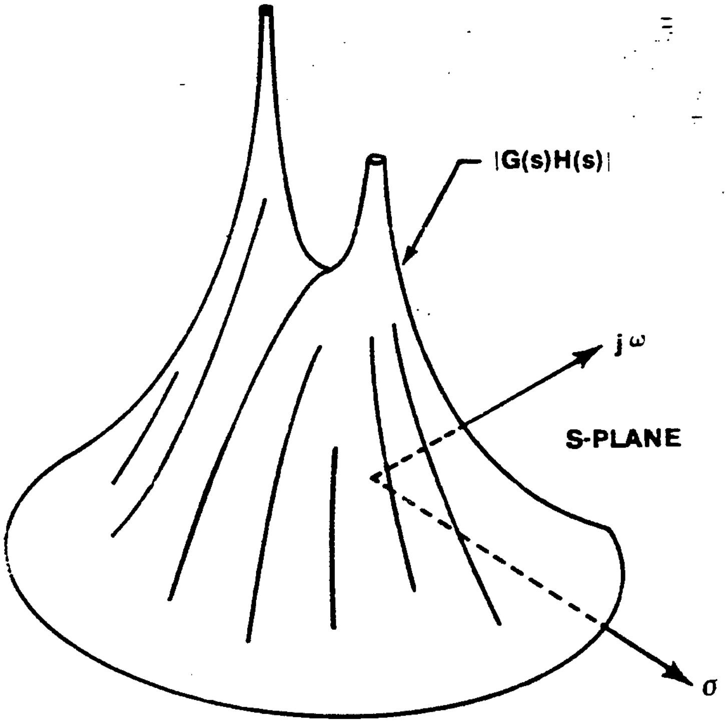


FIGURE 13.50. SURFACE OF G(s) H(s)

The value of s which results in an infinite value of G(s)H(s) is the pole of G(s)H(s).

The pole gets its name from the appearance that a graph of the magnitude of G(s)H(s) makes as "s" assumes values near the point + $1/\tau_{II}$ (Figure 13.51).

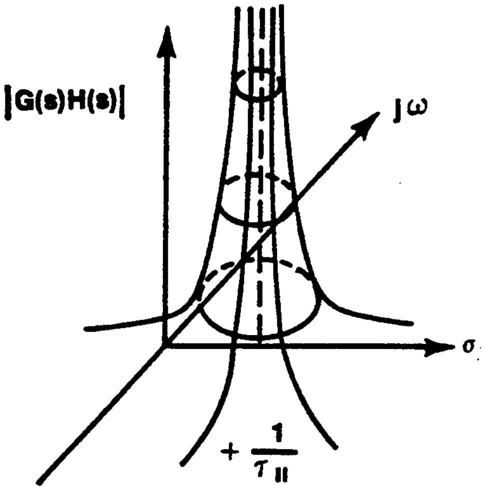


FIGURE 13.51. A POLE OF G(s)H(s)

Note that the poles and zeros of the function G(s)H(s) completely describe the function. When we take the composite function like G(s)/[1 + G(s)H(s)], where presumably we know the poles and zeros of G(s) and H(s), one must exercise caution regarding the transference of this information to the composite function.

For Example:

A zero of G(s) is also a zero of G(s)/[1 + G(s)H(s)]

A pole of G(s) does not result in a pole or zero of G(s)/[1 + G(s)H(s)]

A zero of H(s) does not result in a pole or zero of G(s)/[1 + G(s)H(s)]

A pole of H(s) is a zero of G(s)/[1 + G(s)H(s)]

A zero of 1 + G(s)H(s) is a pole of G(s)/[1 + G(s)H(s)]

Now, since we want to see if the transients die out let us take the expression C(s)/R(s) (Equation 13.69), solve it for C(s) and assume some form of excitation R(s). Actually any form of excitation (sine, unit step, unit ramp, etc.) may be used.

Substituting Equations 13.71 and 13.72 into Equation 13.69 yields

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1 + G(s)H(s)}$$

$$= \frac{K_n \dfrac{N_G(s)}{D_G(s)}}{1 + K_n \dfrac{N_G(s)}{D_G(s)} K_h \dfrac{N_H(s)}{D_H(s)}}$$

$$= \frac{K_n\ N_G(s)\ D_H(s)}{D_G(s)\ D_H(s) + K_n K_h\ N_G(s) N_H(s)} \qquad (13.73)$$

The zeros of $D_G(s)\ D_H(s) + K_n K_h\ N_G(s)\ N_H(s)$ are the same as the zeros of $1 + G(s)H(s)$ and Equation 13.73 can be factored into the form

$$C(s) = \frac{K_n\ N_G(s)\ D_H(s)}{(s - r_1)(s - r_2)(s - r_3)\ \ldots} = \frac{K_n\ N_G(s)\ D_H(s)}{\text{(roots from Root Locus)}} \qquad (13.74)$$

where, for convenience we let $\theta_i(s) = 1$, the unit impulse function. By partial fractions Equation 13.74 can be expanded into the form

$$C(s) = \frac{A_1}{s - r_1} + \frac{A_2}{s - r_2} + \frac{A_3}{s - r_3} + \ldots$$

where r_i are the zeros of $1 + G(s)H(s)$ and the poles of $G(s)/[1 + G(s)H(s)]$. The factors, r_i, may be real or complex and positive or negative. Note that the inverse transform of each element leads to an exponential term. Assuming r_i is real and positive, then ($r_i = + \sigma_i$) and

$$\frac{A_i}{s - r_i} = \frac{A}{s - \sigma_i} \twoheadrightarrow A_i e^{+\sigma_i t} \qquad (13.75)$$

$$\underline{\text{UNSTABLE}}$$

and if r_i is negative then $(r_1 = -\sigma_i)$

$$\frac{A_i}{s - r_i} = \frac{A_i}{s - (-\sigma_i)} = \frac{A_i}{s + \sigma_i} \twoheadrightarrow A_i e^{-\sigma_i t} \qquad (13.76)$$

$$\underline{\text{STABLE}}$$

In the first case (Equation 13.75), the amplitude of the transient term gets large as time gets large because of the $e^{+\sigma_i t}$ term.

In the second case (Equation 13.76), the transient term disappears because as time gets large $e^{-\sigma_i t}$ goes to zero.

Thus, if a system under investigation has any positive real poles of G(s)/[1 + G(s)H(s)] or a positive real zero of 1 + G(s)H(s) then the system is unstable.

Conversely, if the system being investigated has all negative real poles of G(s)/[1 + G(s)H(s)] or all negative real zeros of 1 + G(s)H(s), then the system is stable.

If we assume that r_c is complex then we know there exists another zero of 1 + G(s)H(s) which is the complex conjugate of r_c, namely $\hat{r}_c$.

This pair of zeros of 1 + G(s)H(s) leads to a term in the partial fraction expansion where $r_c = \alpha_c + j\omega_c$ and $\hat{r}_c = \alpha_c - j\omega_c$ of the form

$$\frac{A_c s + A_d}{(s - \alpha_c)^2 + (\omega_c)^2}$$

which has the inverse transform of the form

$$K e^{\alpha_c t} \cos(\omega_c t + \phi_c)$$

where if α_c is positive, $\alpha_c = +\sigma_c$, then we get an exponentially increasing cosine term (Figure 13.52).

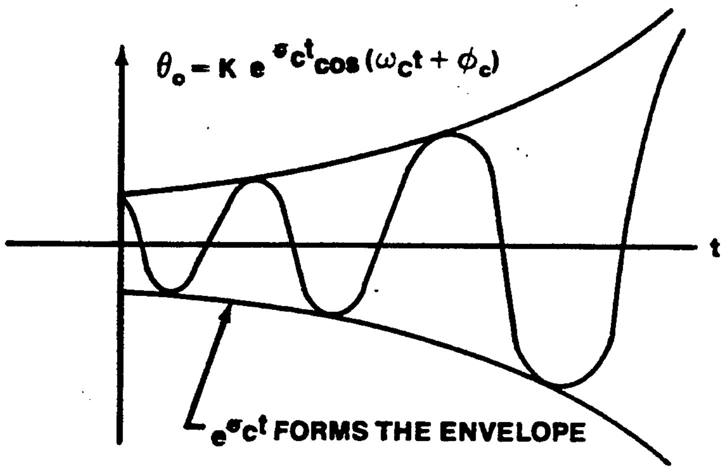


FIGURE 13.52. EXPONENTIALLY INCREASING COSINE TERM

However, if α_C is negative, $\alpha_C = -\sigma_C$. Then the response is of the form $Ke^{-\sigma_C t}$ cos $(\omega_C t + \phi_C)$ which is a cosine term with an envelope that decreases with time (Figure 13.53).

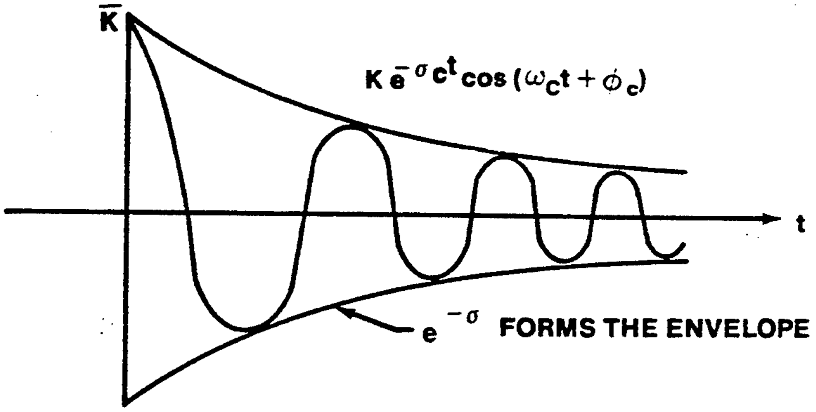


FIGURE 13.53. EXPONENTIALLY DECREASING COSINE TERM

Thus, we conclude that if a complex zero of 1 + G(s)H(s) has a positive real part, $\alpha = +\sigma_C$, then the system is unstable and if a complex zero of 1 + G(s) H(s) has a negative real part, $\alpha = -\sigma_C$, then the system is stable.

Actually the conditions for real zeros and complex zeros are the same:

REAL PART POSITIVE — SYSTEM UNSTABLE

REAL PART NEGATIVE — SYSTEM STABLE.

Now what is the significance of the location of the zeros of 1 + G(s)H(s) upon the s-plane? Looking at the s-plane we find that if ANY zeros of 1 + G(s)H(s) are in the RHP the system is unstable. If ALL zeros of 1 + G(s)H(s) are in the LHP the system is stable.

Now knowing that instability is caused by a zero or zeros of 1 + G(s)H(s) with a positive real part, the problem of determining stability degenerates to the problem of determining whether or not there are any zeros of 1 + G(s)H(s) in the RHP or, equivalently, whether 1 + G(s)H(s) does indeed have a zero or zeros with positive real parts.

13.11.2 Direct Locus Plotting

The example to be used for direct root locus plotting will be the second-order system whose differential equation and transfer functions were derived earlier. The system is shown in Figure 13.54 and Equation 13.67 gives the forward transfer function: H(s) = 1.

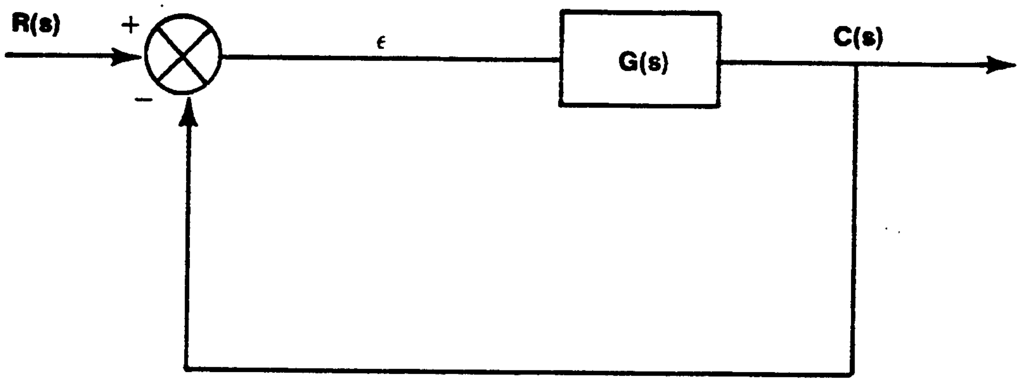


FIGURE 13.54. UNIT FEEDBACK SYSTEM

$$G(s) = \frac{\mu}{Is^2 + bs}$$

The following values will be assumed for the constants

$$I = 1$$

$$b = 2$$

$$\mu = K$$

Equation 13.67 becomes

$$G(s) = \frac{K}{s(s + 2)} \qquad (13.77)$$

The problem is to determine the roots of the characteristic equation for all values of K and to plot these roots in the s-plane.

From Equation 13.69 the system closed-loop transfer function is

$$\frac{C}{R}(s) = \frac{G(s)}{1 + GH(s)}$$

$$\frac{C}{R}(s) = \frac{\frac{K}{s(s + 2)}}{1 + \frac{K}{s(s + 2)}} = \frac{K}{s^2 + 2s + K}$$

The system characteristic equation is

$$s^2 + 2s + K = 0 \qquad (13.78)$$

The roots of Equation 13.78 are

$$s_{1,2} = -1 \pm \sqrt{1 - K}$$

The location of roots for various values of K is shown in Table 13.4.

TABLE 13.4

CLOSED-LOOP ROOT LOCATIONS AS A FUNCTION OF K

K	s_1	s_2
0	$0 + j0$	$-2 - j0$ (open-loop poles)
1/2	$-.3 + j0$	$-1.7 - j0$
1	$-1 + j0$	$-1 - j0$
2	$-1 + j1$	$-1 - j1$
3	$-1 + j\sqrt{2}$	$-1 - j\sqrt{2}$
∞	$-1 + j\infty$	$-1 - j\infty$

The points from Table 13.4 are plotted in the s-plane of Figure 13.55

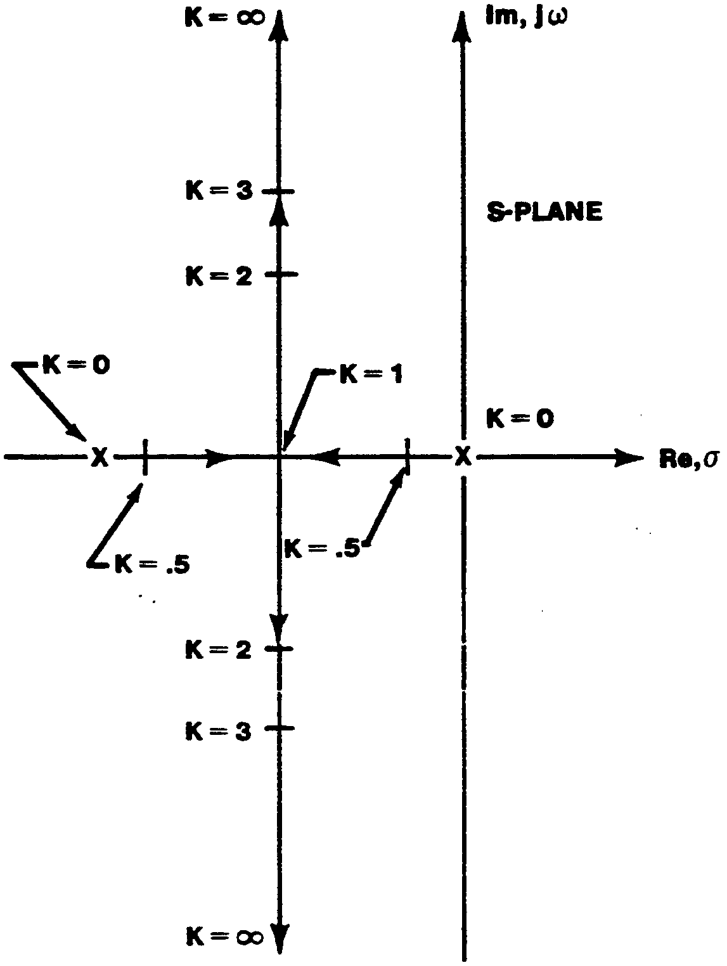


FIGURE 13.55.

Root Locus for $GH(s) = \frac{K}{s(s + 2)}$

The root locus of Figure 13.55, which is the locus of roots of the characteristic equation as a function of gain, quickly indicates whether or not the system is stable, and, also, the form of the transient response for any selected value of K. From the plot, it can be seen that for $0 < K < 1$ the roots are real and negative resulting in exponential decay from each root. For $1 < K < \infty$, however, the roots are complex with the real part negative. The corresponding transient response is oscillatory within an exponentially decaying envelope. For example, for $K = 1.5$ the approximate values of s from the locus are

$$s_1 = -1 + j\,0.5$$

$$s_2 = -1 - j\,0.5$$

The system transient response for $K = 1.5$ is then

$$c(t) = Ce^{-t}\cos(0.5t + \phi)$$

The time involved in constructing a root locus for a complex system in this manner is obviously prohibitive. This difficulty will be overcome later.

To further discuss the significance of the s-plane and the root locus we will consider a second-order characteristic equation in its standard form

$$s^2 + 2\zeta\omega_n s + \omega_n{}^2 = 0$$

The roots of this equation are

$$s_{1,2} = \underbrace{-\zeta\omega_n}_{\sigma} \pm \underbrace{j\omega_n\sqrt{1-\zeta^2}}_{\omega_d} \qquad (13.79)$$

In order to realize what this means in the s-plane, refer to Figure 13.56a. Any arbitrary value of s will have, from Equation 13.79, a real part $\sigma = \zeta\omega_n$ and an imaginary part $\omega_d = \omega_n\sqrt{1-\zeta^2}$. From these values shown in Figure 13.56, and the Pythagorean theorem, the magnitude of the position vector, s is found to be equal ω_n. The angle ϕ is also significant since

$$\cos\phi = \frac{\zeta\omega_n}{\omega_n} = \zeta$$

Figure 13.56a summarizes this information and shows how parameters important to the transient response can be easily obtained from the position of roots in the s-plane. From the root locus, then, the transient response characteristics for all values of gain, K, can be seen at a glance.

13.11.3 Angle and Magnitude Conditions

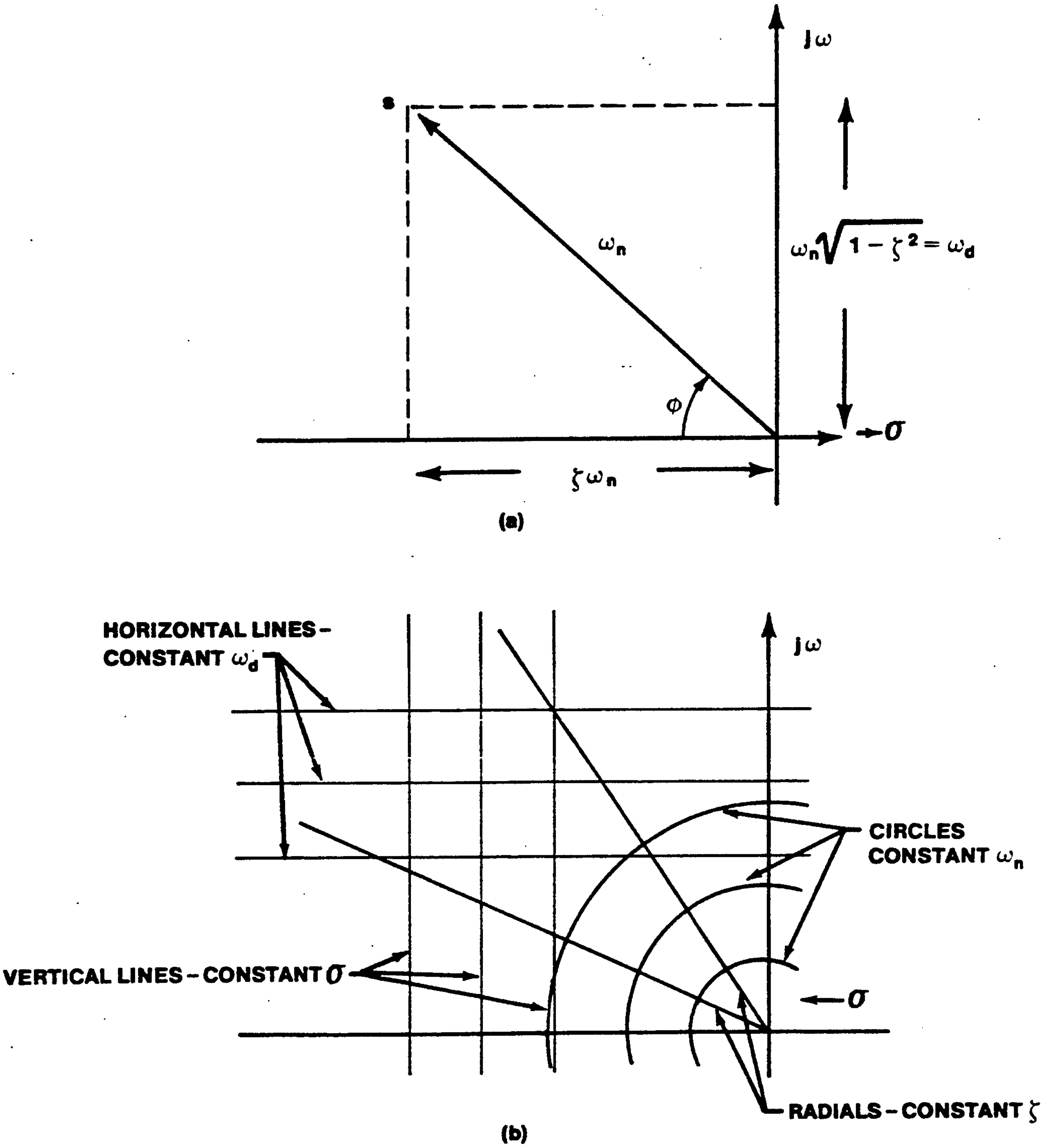
Now that the value of the root locus has been established, the rules will be developed which permit simplified plotting of complex systems. These rules are based on two conditions, the angle condition and the magnitude condition which evolve from the characteristic equation.

As stated many times, the closed-loop control system may be represented by

$$\frac{C}{R}(s) = \frac{G(s)}{1 + GH(s)}$$

from which the characteristic equation is

$$1 + GH(s) = 0$$



(a)

(b)

FIGURE 13.56. SIGNIFICANCE OF s-PLANE PARAMETERS

The system open-loop transfer function can be written in the following form

$$GH(s) = \frac{K(s - z_1)(s - z_2) \ . \ . \ .}{(s - p_1)(s - p_2) \ . \ . \ .} \qquad (13.80)$$

where the static gain, K, is factored out and the z's and p's are the open-loop zeros and poles respectively.

Since the values of s to be determined must satisfy the relationship

$$1 + GH(s) = 0$$

we can say

$$GH(s) = -1 \qquad (13.81)$$

but, since s is, in general, complex, GH(s) is then a function of a complex variable and Equation 13.81 can be written

$$GH(s) = 1e^{j(1 + 2n)\pi} = 1 \ \underline{/(1 + 2n)\pi} \qquad (13.82)$$

where

$$n = 0, \pm 1, \pm 2, \pm 3, \ . \ . \ . \ .$$

Equation 13.82 says that in order for the value of s to be a zero of 1 + GH(s) the magnitude of the complex quantity GH(s) must be equal to 1 and the argument be some odd multiple of π. Hence the

Magnitude condition

$$|GH(s)| = 1 \qquad (13.83)$$

Angle condition

$$\underline{/GH(s)} = (1 + 2n)\pi \qquad (13.84)$$

$$n = 0, \pm 1, \pm 2, \pm 3 \ . \ . \ .$$

Substituting Equation 13.80 into Equation 13.84, the angle condition becomes

$$\angle \frac{K(s - z_1)(s - z_2) \ldots}{(s - p_1)(s - p_2) \ldots} = (1 + 2n)\pi \qquad (13.85)$$

Since each factor of GH(s) can be represented by a vector in the s-plane from the pole or zero to the s-point in question, Equation 13.85 can be written

$$\frac{\angle s - z_1 + \angle s - z_2 + \ldots}{\angle s - p_1 + \angle s - p_2 + \ldots} = (1 + 2n)\pi$$

or

$$\angle s - z_1 + \angle s - z_2 + \ldots - \angle s - p_1 - \angle s - p_2 - \ldots = (1 + 2n)\pi \qquad (13.86)$$

Thus, using the angle condition, any point in the s-plane can be investigated to determine whether or not it is a point on the root locus by measuring the angles of the vectors from the poles and zeros to the point in question, and adding them according to the left side of Equation 13.86. If this sum equals an odd multiple of π, the value of s satisfies the characteristic equation, and is on the root locus.

Figure 13.57 demonstrates the application of the angle condition. The vectors representing GH(s) for $s = s_0$ are shown in Figure 13.57. The angle condition test to determine if s_0 is on the root locus is

$$\phi - \theta_1 - \theta_2 = (1 + 2n)\pi \qquad (13.87)$$

From the figure,

$$\phi = 170^\circ$$

$$\theta_1 = 195^\circ$$

$$\theta_2 = 140^\circ$$

Equation 13.87 becomes

$$170 - 195 - 140 \neq \pm \pi$$

$$-165 \neq \pm \pi$$

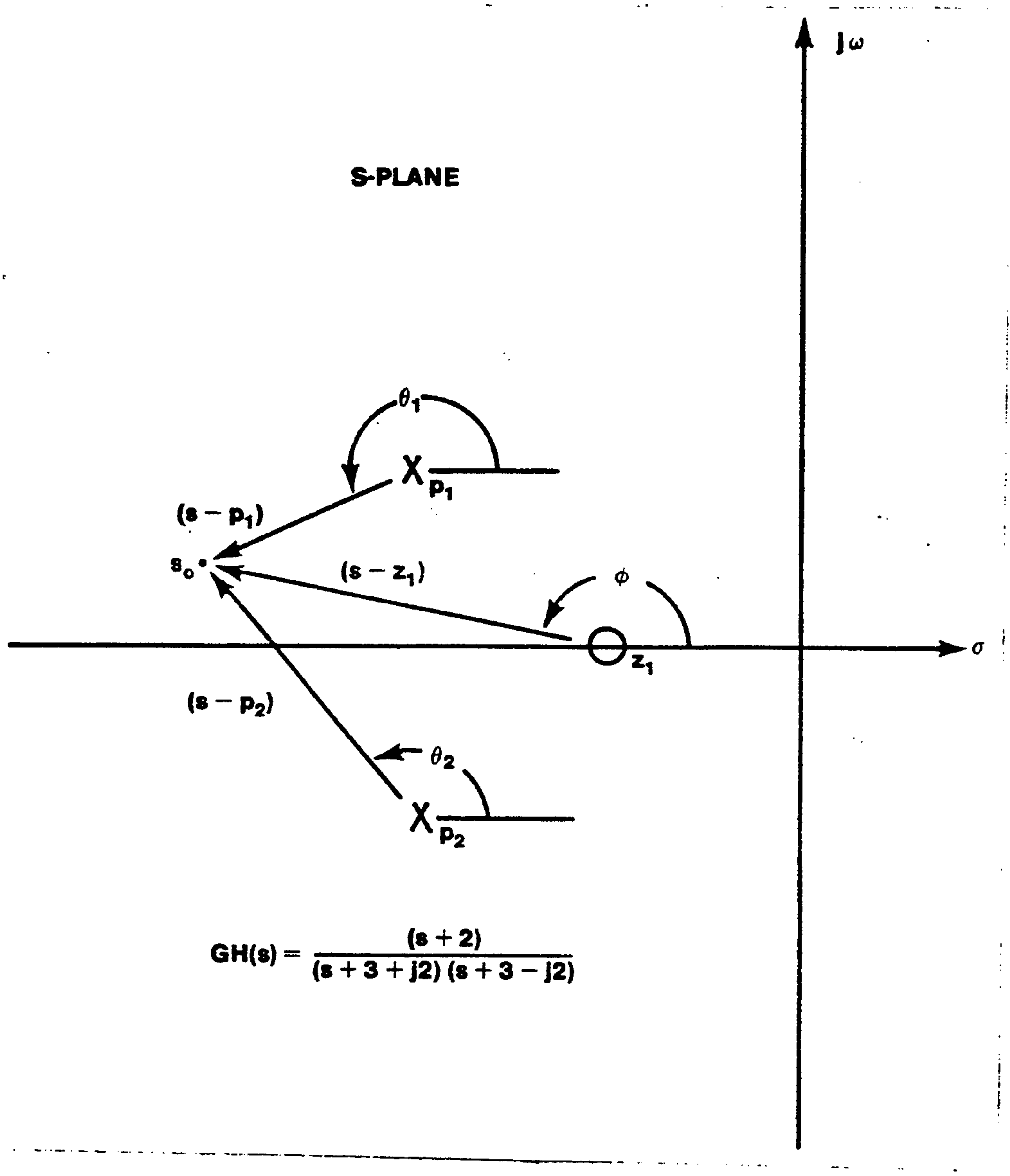


FIGURE 13.57. APPLICATION OF ANGLE CONDITION

Therefore, s_0 is not on the locus but is fairly close. Successive tries will allow converging on the point that satisfies the characteristic equation.

Combining Equation 13.80 and 13.83 the magnitude condition becomes

$$\frac{K \cdot |s - z_1| \cdot |s - z_2| \; . \; . \; .}{|s - p_1| \cdot |s - p_2| \; . \; . \; . \; . \; .} = 1$$

$$K = \frac{|s - p_1| \cdot |s - p_2| \; . \; . \; . \; .}{|s - z_1| \cdot |s - z_2| \; . \; . \; . \; .} \qquad (13.88)$$

Equation 13.88 says that the magnitude of all the vectors from the poles to the s point in question, divided by the magnitudes of the vectors from the zeroes is equal to the gain, K. This condition allows us to determine the value of gain from any point on a locus in the s-plane.

The reader is encouraged to check the example system's root locus (Figure 13.55), to ensure it is consistent with both the angle and magnitude condition.

13.11.4 Rules for Root Locus Construction

The following rules for K > 0 will allow a sketch of the root locus to be drawn quickly. These rules are based upon the angle condition and an analysis of the characteristic equation.

a. The number of branches of the locus is equal to the number of open-loop poles (i.e., the order of the characteristic equation).

$$1 + GH(s) = 1 + \frac{K(s - z_1) \; . \; . \; . \; (s - z_Z)}{(s - p_1) \; . \; . \; . \; (s - p_P)} = 0 \qquad (13.89)$$

$$(s - p_1) \; . \; . \; . (s - p_P) + K \; (s - z_1) \; . \; . \; . \; (s - z_Z) = 0$$

Since we have assumed a rational polynomial, $P \geq Z$, and the highest order of s is P.

P is the number of open loop poles, Z is the number of open loop zeros.

b. The Loci branches begin at the open-loop poles where K = 0. If we write the characteristic equation with the static gain factored out

$$1 + GH(s) = 1 + KGH'(s) = 0$$

then

$$KGH'(s) = -1$$

$$GH'(s) = -1/K \qquad (13.90)$$

for $K = 0$, $GH'(s) = \infty$, which means s is at a pole of GH(s).

c. The branches end at the open-loop zeros where $K = \infty$. From Equation 13.90, when $K = \infty$, $GH'(s) = 0$, which is a zero of GH(s). When $P > Z$ (the usual case) the additional branches end at $s = \infty$ which may also be considered an open-loop zero.

d. The loci branches that do approach $s = \infty$ do so asymptotically to radial lines centered at

$$\sigma = \frac{\Sigma \text{ Re } \{p's\} - \Sigma \text{ Re } \{z's\}}{P - Z}$$

e. The angles of these asymptotes are given by

$$\gamma = \frac{(1 + 2n)\pi}{P - Z}$$

where

$$n = 0, \pm 1, \pm 2 \ldots$$

f. A point on the real axis is on a locus branch if it is to the left of an odd number of open-loop poles and zeros. This can be seen from Figure 13.58 in which it is obvious that the net angular contribution from a pair of complex conjugate poles or zeros to a search point on the real axis is 0°. The real axis poles and zeros contribute zero or π when the search point is to the right or left of the pole or zero respectively. Therefore, the angle condition is satisfied only to the left of an odd number of open-loop poles and zeros.

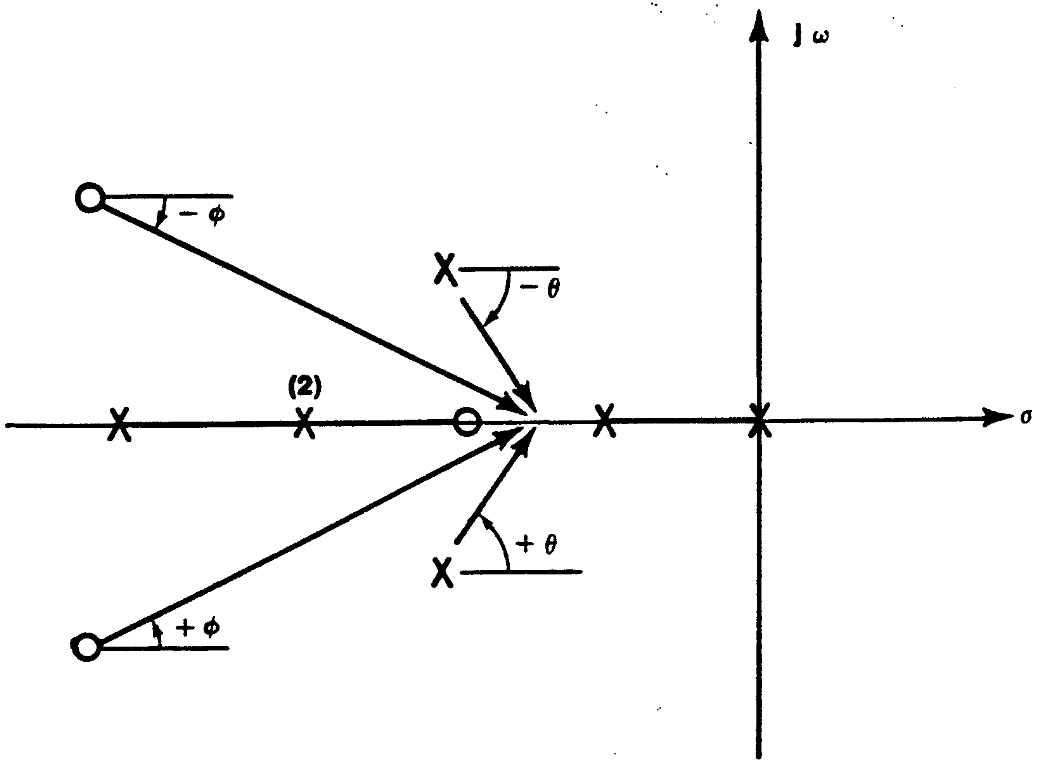


FIGURE 13.58. REAL AXIS LOCI

g. The angle of departure approach of a loci branch from (to) a complex pole (zero) can be found using the angle condition as follows. If a search point is chosen ε distance from a complex pole (zero), the angle of this vector is the departure(approach) angle at that pole (zero).

Consider the situation of Figure 13.59 where it is desired to find the departure angle from p_1. If the magnitude of the vector from p_1 is ε then the angle of the vectors from the other poles and zeros can be measured directly to p_1. Solving the angle condition, Equation 13.86 for $\angle s - P_1$ (an unknown quantity) will yield the departure angle.

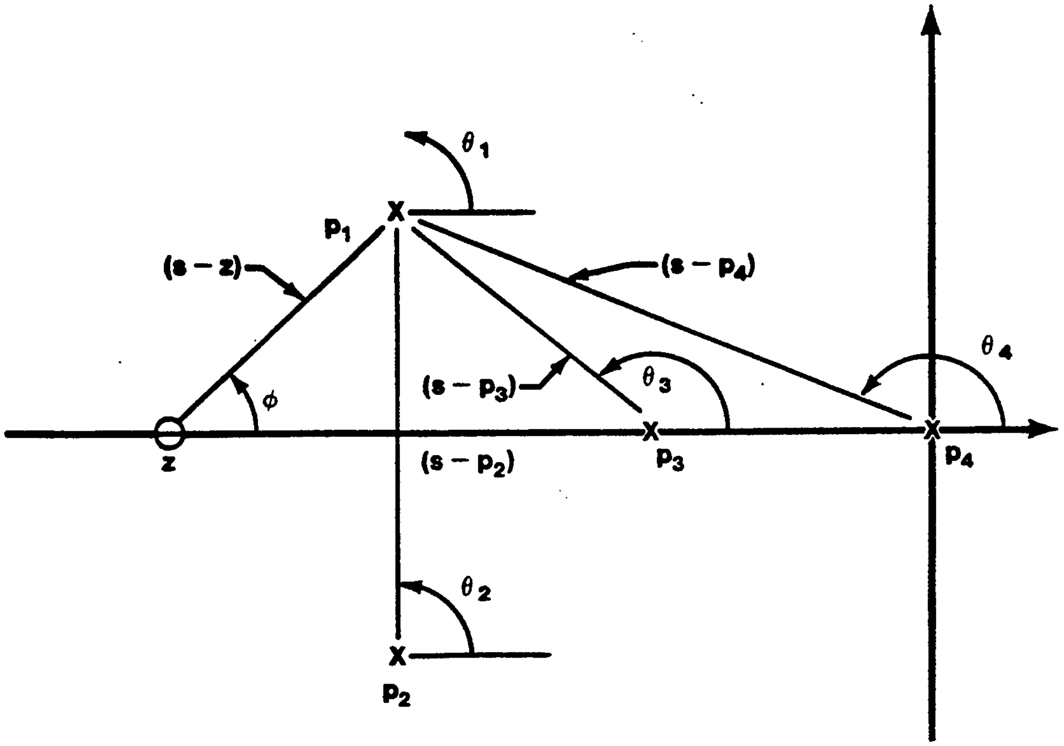


FIGURE 13.59. DEPARTURE ANGLE DETERMINATION

Equation 13.86 says

$$\angle s - z - \angle s - p_1 - \angle s - p_2 - \angle s - p_3 - \angle s - p_4 = (1 + 2n)\pi$$

From Figure 13.59

$$\phi - \theta_1 - \theta_2 - \theta_3 - \theta_4 = (1 + 2n)\pi$$

$$\phi = 45^{\circ}$$

$$\theta_2 = 90^{\circ}$$

$$\theta_3 = 135^{\circ}$$

$$\theta_4 = 155°$$

$$\theta_1 = 45° - 90° - 135° - 155° - \pi$$

Departure angle is

$$\theta_1 = -515° = -155°$$

h. The points at which the loci branches leave or enter the real axis are sometimes called "breakaway" or "breakin" points, respectively.

Figure 13.60 illustrates the computation of a breakaway point for the system whose open-loop transfer function is

$$GH(s) = \frac{K}{s(s+1)(s+2)}$$

From the characteristic equation,

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

$$K = -s(s+1)(s+2)$$

$$K = -s^3 - 3s^2 - 2s$$

To find the point where K is maximum, we can differentiate the expression for K and set it equal to zero. The solution of this equation should produce the desired result.

$$\frac{dK}{ds} = -3s^2 - 6s - 2 = 0$$

$$s_{1,2} = -1 \pm 0.574$$

$$s = -1.574 \text{ or } -0.426$$

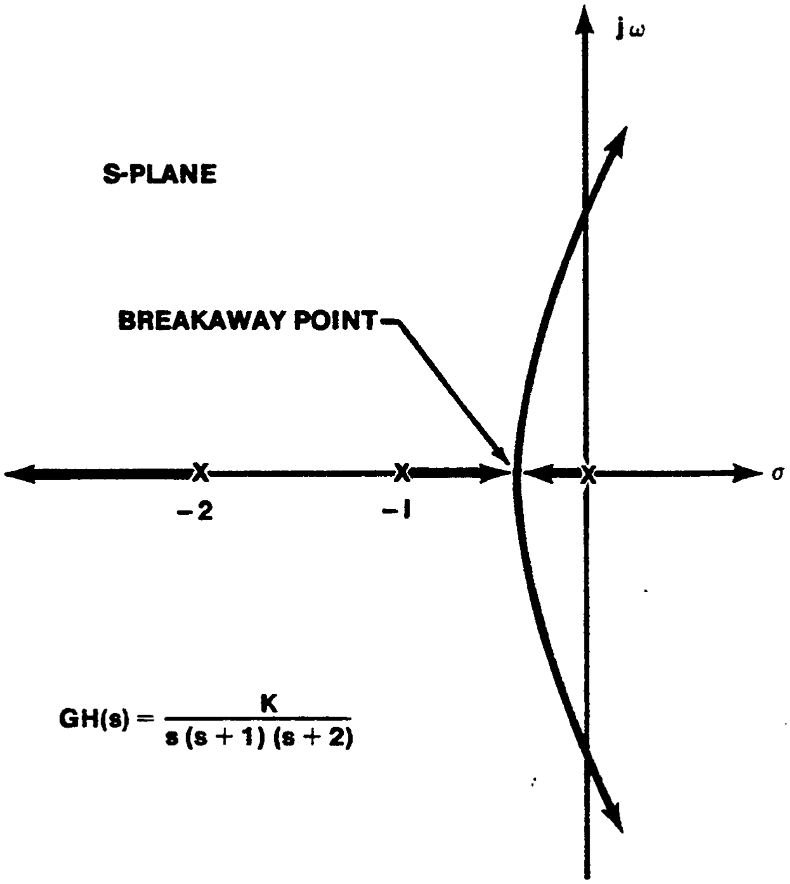


FIGURE 13.60. BREAKAWAY POINT COMPUTATION

It is obvious from Figure 13.60 that the latter solution, s = -0.4257, is the meaningful answer because it appears on a real axis locus. This, then is the breakaway point.

This method of computing breakaway points is restrictive in that more complex systems with higher-order transfer functions are difficult to solve since dK/ds is higher than second-order.

The fact that the root locus is always symmetrical about the real axis is advantageous in that only the upper half of the s-plane need be plotted. The lower half is just a mirror image. The reason for this should be obvious when one considers that complex roots always appear in conjugate pairs.

The rules above will permit a sketch of the root locus very quickly. If more accurate data is required, the branches can be checked using the angle condition and a protractor. After the locus is complete, the values of K that are deemed important can be computed using the magnitude condition.

Root Locus Examples:

Example 1:

$$G(s) = \frac{K'}{s(0.5s + 1)(0.2s + 1)}$$

$$H(s) = 1$$

Rewriting in the more useful form

$$GH(s) = \frac{10K'}{s(s + 2)(s + 5)}$$

Letting 10K' = K, we can plot the locus and calibrate it in gain. In the actual system, however, the gain selected, K', will be a factor of 10 less than that found from the locus plot.

From the previous equation

$$p_1 = 0, \; p_2 = -2, \; p_3 = -5$$

$$P = 3, \quad Z = 0$$

Applying the rules developed in the previous section, the root locus of Figure 13.61 is plotted.

a. The number of branches is 3.

b. The locus branches begin at the poles s = 0, -2, -5, where K = 0.

c. Since there are no open-loop zeros, the 3 branches will end along asymptotes whose real axis intercepts are

$$\sigma = \frac{\Sigma \text{ Re } \{p's\} - \Sigma \text{ Re } \{z's\}}{P - Z}$$

$$\sigma = \frac{[0 - 2 - 5] - [0]}{3 - 0}$$

$$\sigma = - 7/3$$

The asymptotes angles are

$$\gamma = \frac{(1 + 2n)\pi}{P - Z}$$

$$\gamma = \frac{(1 + 2n)\pi}{3} = 60^{\circ}, 180^{\circ}, 300^{\circ}$$

d. Real axis loci exist between s = 0 and s = -2, and from s = -5 to s = -∞.

e. There are no complex poles or zeros so rule (g) does not apply.

f. The breakaway point is found as follows:

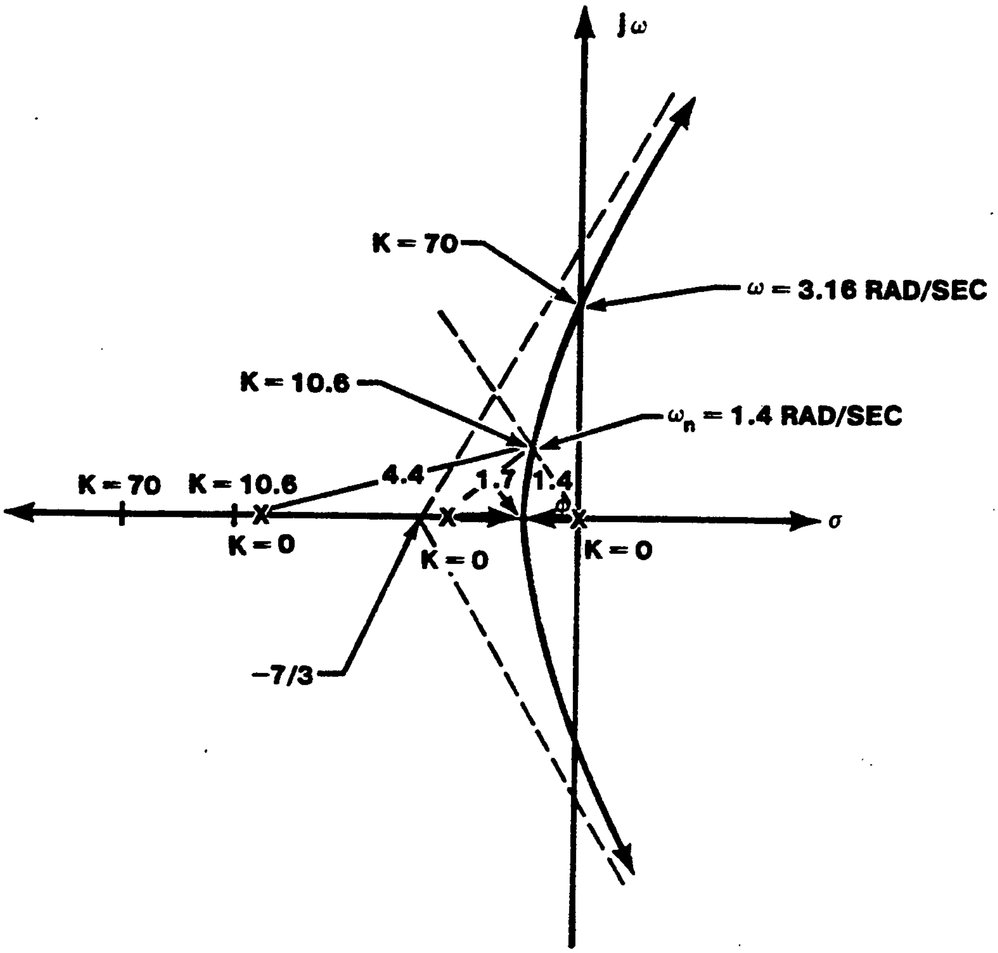


FIGURE 13.61. EXAMPLE OF A ROOT LOCUS

$$GH(s) = \frac{K}{s(s+2)(s+5)} = -1$$

$$K = -s(s+2)(s+5)$$

$$K = -s^3 - 7s^2 - 10s$$

$$\frac{dK}{ds} = -3s^2 - 14s - 10 = 0$$

$$s = -3.79, -0.88$$

Since - 3.79 is not on a locus branch, s = -0.88 is the breakaway point.

Now that the root locus is plotted, the desired values of K can be found. It is generally useful to know the value of K for which the system becomes unstable. By measuring the length of the vectors from each of the poles to the point where the locus crosses the imaginary axis into the LHP, we find

$$K = (3.16)(3.7)(6.0) = 70$$

The system is neutrally stable at K = 70 and the frequency of the oscillation will be 3.16 rad/sec. For K > 70, the roots of the characteristic equation are in the RHP where they have positive real parts and the system is unstable.

If some specific transient response parameter is required, such as $\zeta = 0.5$, the corresponding value of K can be determined.

Since

$$\cos^{-1} \zeta = \phi$$

$$\cos^{-1} 0.5 = 60°$$

Where the radial from the origin 60° from the negative real axis crosses the root locus, the system is at the required operating condition, and the gain for this point is

$$K = (4.4)(1.7)(1.4) = 10.6$$

To find the total transient response, however, it is necessary to find the point on all branches where K = 10.6. Each branch contributes one term to the transient response and they all must have the same value of K. Because the lower half plane is a mirror image, the lower root is the complex conjugate of the upper value. The point on the third branch along the real axis where K = 10.6 is found through trial and error. The transient solution, then, for K = 10.6, $\zeta = 0.5$, has a quickly damped pure exponential term and a dominant more slowly damped exponentially decaying oscillation at $\omega_n = 1.4$.

$$\theta_0(t)_{\text{transient}} = C_1 e^{-5.5t} + C_2 e^{-0.7t} \cos (1.2t + \phi)$$

Example 2:

$$GH(s) = \frac{K(s + 2)}{s(s + 3)(s^2 + 2s + 2)}$$

$$GH(s) = \frac{K(s + 2)}{s(s + 3)(s + 1 + j)(s + 1 - j)}$$

Applying the rules:

a. The number of branches is 4.

b. The locus branches begin on the open-loop poles at $s = 0, -3, -1 \pm j$.

c. One branch will end on the open loop zero at $s = -2$. The remaining 3 branches will end at $s = \infty$ along asymptotes centered at

$$\sigma = \frac{[0 - 3 - 1 - 1] - [-2]}{4 - 1}$$

$$\sigma = -1$$

with asymptote angles

$$\gamma = \frac{(1 + 2n)\pi}{4 - 1}$$

$$\gamma = 60^{\circ},\ 180^{\circ},\ 300^{\circ}$$

d. Real axis loci exist between s = 0 and -2, and form s = -3 to - ∞.

e. The departure angle from the complex poles is found by

$$\theta = 45^{\circ} - 135^{\circ} - 90^{\circ} -26.5^{\circ} + 180^{\circ}$$

$$\theta = - 26.5^{\circ}$$

Once the value of s for the imaginary axis crossing is found, a graphical solution for K using the magnitude condition can be used, i.e.,

$$K = \frac{(3.41)(2.81)(1.60)(1.19)}{2.59} = 7.04$$

We will continue this development to determine the closed-loop response for $\zeta = .5$. As the open-loop transfer function was specified

$$GH = \frac{K(s + 2)}{s(s + 3)(s^2 + 2s + 2)}$$

we shall further specify that

$$G(s) = \frac{K}{s(s^2 + 2s + 2)} \qquad \text{and}$$

$$H(s) = \frac{(s + 2)}{(s + 3)}$$

It can be seen from the root locus plot that the transient response will have two pure exponentially decaying terms from the real axis branches and one oscillatory damped term from the complex loci. It should also be noted that the maximum ζ possible for the oscillatory terms is $\zeta = .7$ established by the open-loop complex poles where K = 0. Although the $\zeta = .5$ radial crosses the complex locus in the vicinity of $s_{1,2} = 0.5 \pm .9j$, the exact crossing can be found using the angle condition. A point on the $\zeta = .5$ radial near the sketched locus is tried as a search point and if the point satisfies the angle condition, the point is on the locus and pins down the exact roots in question. A couple of search point narrows our locus to $s = -.56 \pm .96j$. At this point the angle condition is satisfied, i.e., $\angle s - z_1 - \angle s - p_3 - \angle s - p_2 - \angle s - p_4 - \angle s - p_1 = \pi$, or $33.6^\circ - 120^\circ - 77^\circ - 21.2^\circ - (-5^\circ) = 179.6^\circ \cong \pi$.

Next, the magnitude condition used at this point is $= -.56 + j.96$, to determine the value of gain, K.

$$K = \frac{|s| \quad |s + 3| \quad |s+ 1 +j| \quad |s + 1 -j|}{|s + 2|}$$

$$= \frac{(2.36)\,(2.025)\,(1.12)\,(.45)}{1.74}$$

$$= 1.532.$$

So far we have determined two roots (a complex pair) and the value of gain for system operation. Two other roots lie on the real axis loci (one on each branch) and are found by trial and error by using the magnitude condition and finding a search point that will yield a value of K = 1.532.

A search point at s = -.8 is determined to be a root corresponding to a value of K = 1.532, i.e.,

$$\frac{(.8)\,(2.2)\,(1.025)^2}{(1.2)} = 1.54 = K$$

(close enough for a graphical solution)

Another search point at s = -3.81 also satisfies the magnitude condition, i.e.,

$$\frac{(.81)\,(3.81)\,(2.99)^2}{1.81} = 1.525 = K$$

(close enough for a graphical solution)

Thus we have located the four roots that correspond to the specified requirements. These roots are s = -.8, s = -3.81, s = -.56 + 96j, s = - .56 - .96j, with a corresponding operating gain of 1.53 (Figure 13.62).

Recalling Equations 13.71, 13.72, and 13.73 the closed-loop response can be written using the roots determined from the root locus, i.e.,

$$G(s) = K_n \frac{N_G(s)}{D_G(s)}$$

$$H(s) = K_H \frac{N_H(s)}{D_H(s)}$$

$$\frac{C(s)}{R(s)} = \frac{K_n N_G(s)\ D_H(s)}{D_G(s) D_H(s) + K_n K_h N_G(s) N_H(s)}$$

$$= \frac{K_n N_G(s)\ D_G(s)}{\text{(Roots from Root Locus)}}$$

Substituting in the appropriate quantities yields

$$\frac{C(s)}{R(s)} = \frac{1.53(s + 3)}{(s + 3.81)(s + .8)(s^2 + 1.12s + 1.23)}$$

$$GH = \frac{K\ (s + 2)}{s(s + 3)(s^2 + 2s + 2)}$$

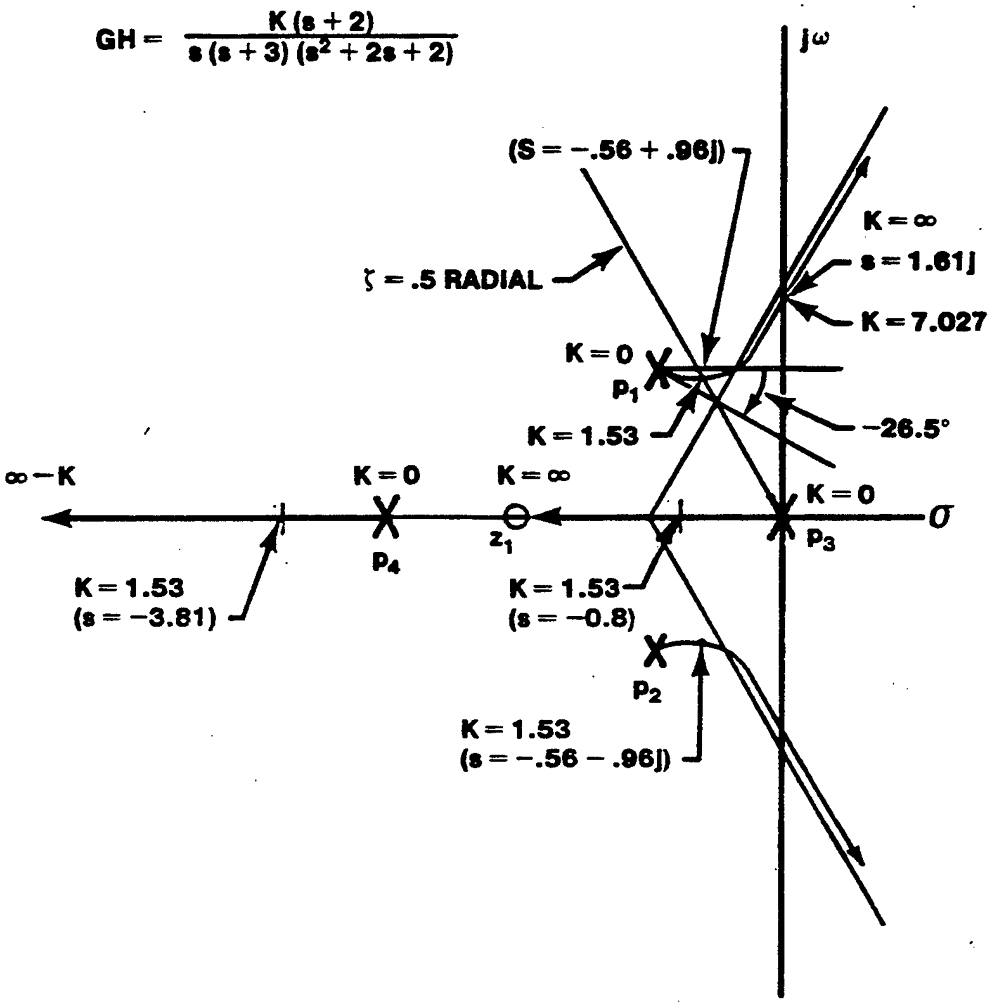


FIGURE 13.62. ROOT LOCUS PLOT

13.12 COMPENSATION TECHNIQUES

You have been introduced to theory and technique used to set up a control system problem and to analyze the results. Unfortunately the system often needs to be "fixed" to meet the desired performance specifications. This "fixing" is called compensation and is used to reshape the root locus to achieve the desired performance specifications. Generally, three performance specifications are changed by compensation: degree of stability, transient response, and steady-state error.

Actual compensation is achieved by addition of electrical network, and/or mechanical devices which may contain levers, springs, dashpots, gyros, etc. There are two positions in the control system where compensation is usually performed. In the feedback loop where it is referred to as feedback compensation and in the forward loop where it is called cascade compensation. In the forward loop the compensator is normally placed in the low energy point so the power dissipation will be small.

Common networks used to achieve compensation are lag, lead, and lead-lag, all of which are passive. Modern control systems often use active networks which modify the systems to ensure the desired specifications are met by cancelling the undesirable characteristics and replacing them with the desired characteristics.

Our approach to compensation will be to look at our basic system then to see the effect of each type of compensation on the shape of the root locus and the steady-state performance.

13.12.1 Feedback Compensation

13.12.1.1 Proportional Control (Unity Feedback). Consider the basic system in Figure 13.63.

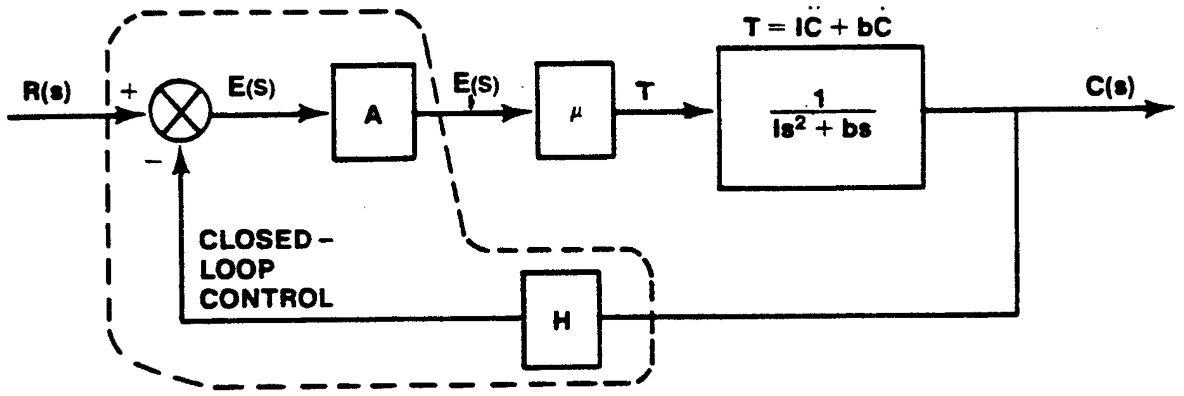


FIGURE 13.63. BASIC SYSTEM

where for the first case

$$A = 1$$

$$G(s) = \frac{\mu}{Is^2 + bs} = \frac{\mu/I}{s(s + b/I)}$$

$$H(s) = 1$$

$$GH(s) = \frac{\mu/I}{s(s + b/I)}$$

The closed-loop transfer function is

$$\frac{C}{R}(s) = \frac{\mu/I}{s^2 + \frac{b}{I}s + \frac{\mu}{I}} \qquad (13.90)$$

and

$$\omega_n = \sqrt{\mu/I}$$

and

$$\zeta = \frac{b}{2\sqrt{\mu I}}$$

Figure 13.64 is the root locus of this basic system.

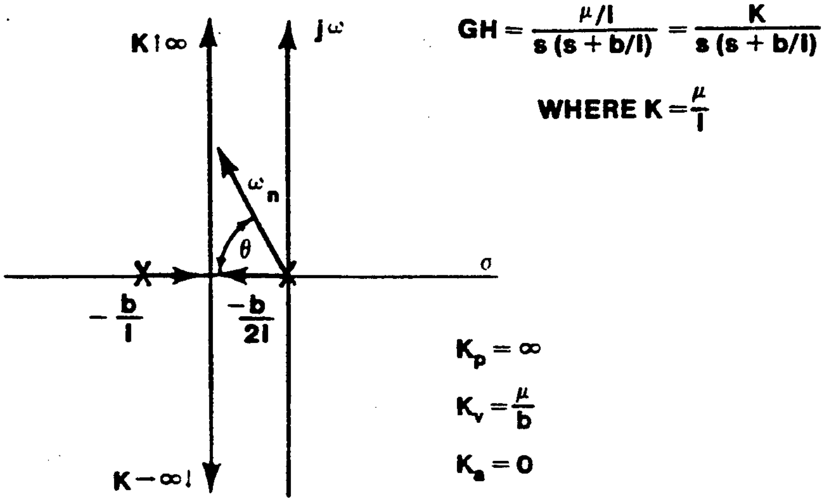


FIGURE 13.64. ROOT LOCUS OF BASIC SYSTEM WITH UNITY FEEDBACK

Application of a step, R(s) = 10/s, results in

$$C(s) = \frac{10}{s}\left(\frac{\mu/I}{s^2 + \frac{b}{I}s + \frac{\mu}{I}}\right) \qquad (13.91)$$

The inverse transform of Equation 13.91 is

$$c(t) = 10 - \frac{10}{\sqrt{1-\zeta^2}} e^{-\zeta\omega_n t} \sin(\omega_d t + \phi)$$

After a finite time the transients die out and the response settles down to the commanded value of 10 units with no steady-state error. If the damping term, b, was zero, the locus would move right to the imaginary axis and pure harmonic motion would result. By looking at the error coefficients K_p, K_v and K_a, we see that this system has zero error for a step, Rb/μ error for a ramp and has infinite error for a parabolic input.

The amount of viscous damping in a practical system is often limited by physical constraints. To cope with a lightly damped system, artificial damping is added. An investigation of Equation 13.90 reveals that the damping of the system can be increased by increasing the coefficient of the s term (θ term) in the denominator.

13.12.1.2 Derivative Control (Rate Feedback). The problem of adding a $K\dot{C}$ term (Ks term) into the forward loop is solved by derivative control or rate feedback. Figure 13.65 shows the introduction of the $K\dot{C}$ term into the block diagram.

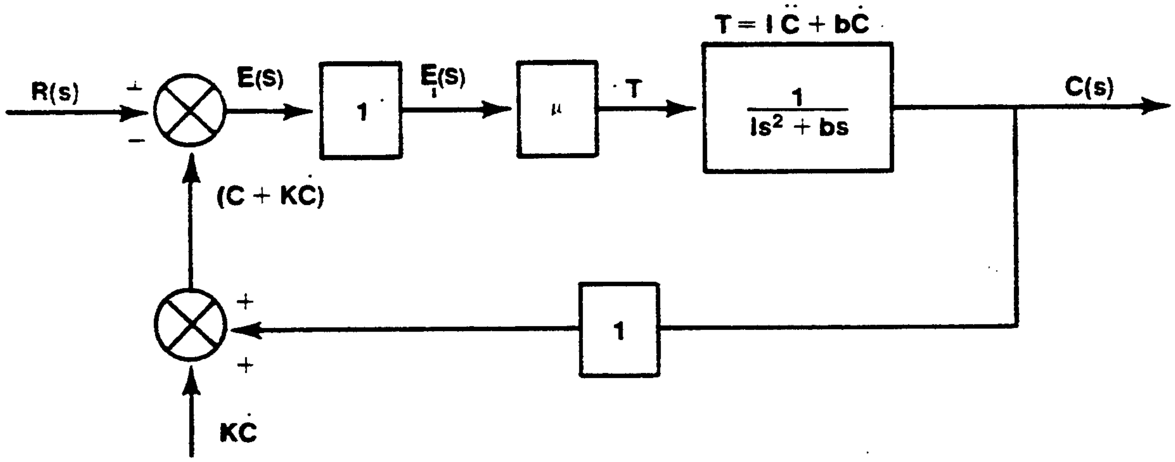


FIGURE 13.65. BASIC SYSTEM WITH INTRODUCTION OF $K\dot{C}$ TERM

The $K\dot{C}$ term can be achieved by taking the derivative of C and is often achieved by means of a rate gyro as shown in Figure 13.66.

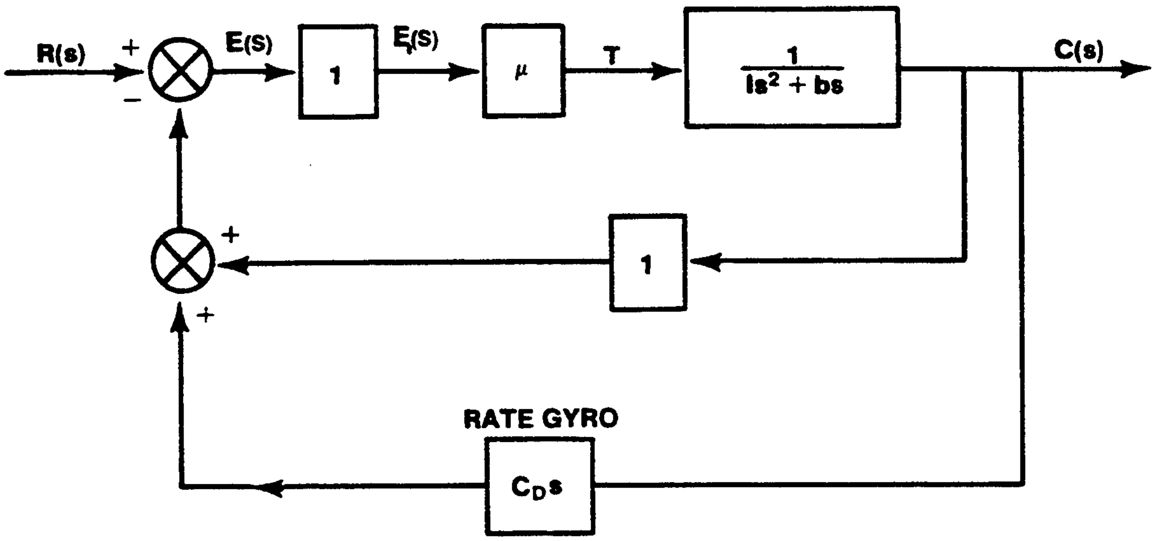


FIGURE 13.66. BASIC SYSTEM WITH RATE GYRO ADDED TO FEEDBACK LOOP

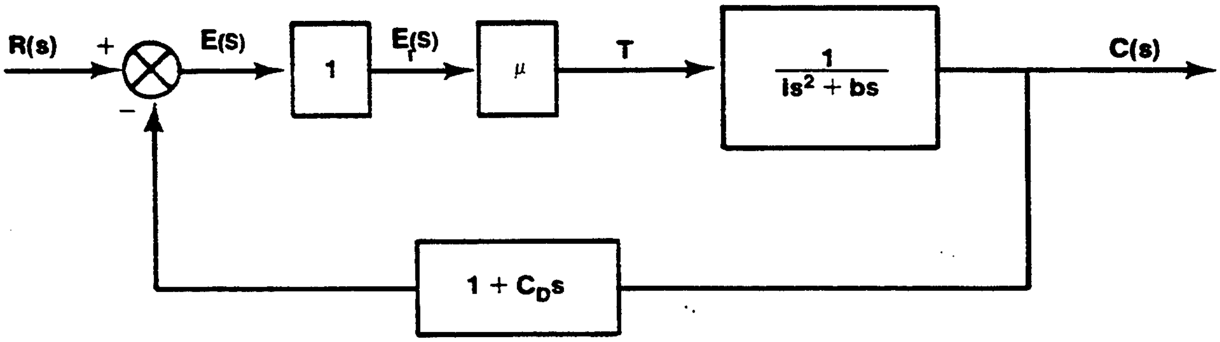


FIGURE 13.67. REDUCTION OF FIGURE 13.66

The control ratio is written for Figure 13.67

$$\frac{C(s)}{R(s)} = \frac{\mu}{Is^2 + bs + \mu (1 + C_D s)}$$

$$= \frac{\mu/I}{s^2 + \left(\frac{b + \mu C_D}{I}\right)s + \frac{\mu}{I}}$$

The natural frequency of the system is unchanged; however the damping has been increased by $\mu C_D/I$.

The effect of derivative control on the root locus is illustrated in Figure 13.68. The forward transfer function becomes

$$G(s) = \frac{\mu/I}{s(s + b/I)}$$

and the feedback transfer function is

$$H(s) = (1 + C_D s)$$

This in effect adds an open-loop zero to our open-loop transfer function.

$$GH(s) = \frac{\frac{\mu}{I}(1 + C_D s)}{s\left(s + \frac{b}{I}\right)} = \frac{\frac{\mu}{I}\ C_D\left(s + \frac{1}{C_D}\right)}{s\left(s + \frac{b}{I}\right)}$$

$$GH(s) = \frac{\frac{\mu}{I} C_D (s + 1/C_D)}{s(s + \frac{b}{I})} = \frac{K'(s + 1/C_D)}{s(s + b/I)}$$

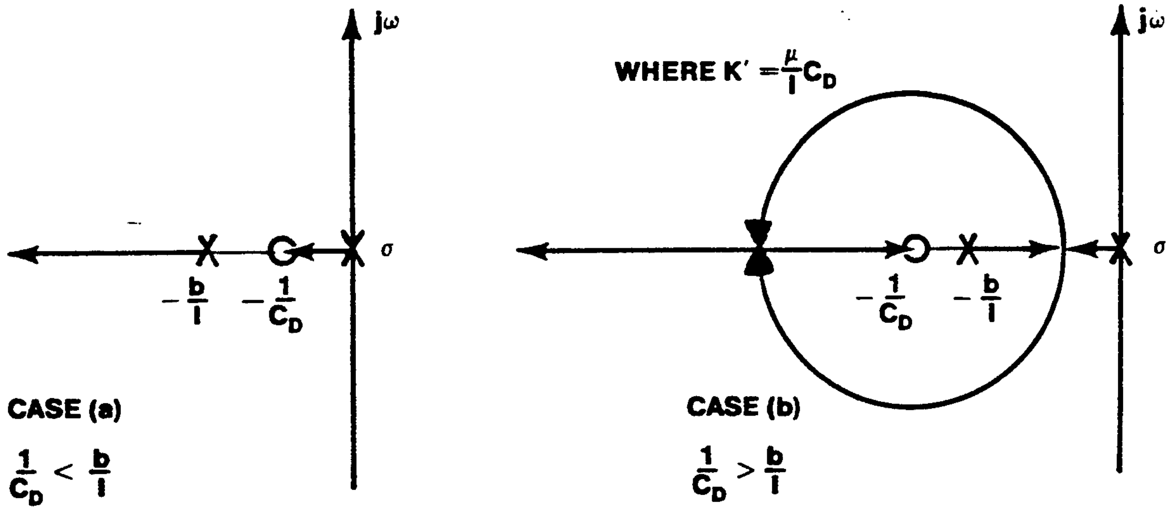


FIGURE 13.68. ROOT LOCUS OF BASIC SYSTEM WITH DERIVATIVE CONTROL

The value of C_D will determine if the response will have an oscillatory term. In case (a) of Figure 13.68, $1/C_D < b/I$ will result in two first order roots and oscillation is not possible. However in case (b) where $1/C_D > b/I$, the locus path allows for an oscillatory response over a wide range of static loop sensitivity ($\mu/I\ C_D$). It is obvious from Figure 13.68 that high damping can be achieved with reasonably high values of static loop sensitivity (gain). Before we can check the steady-state error, we must rearrange the block diagram (Figure 13.69 a, b, c) to unity feedback.

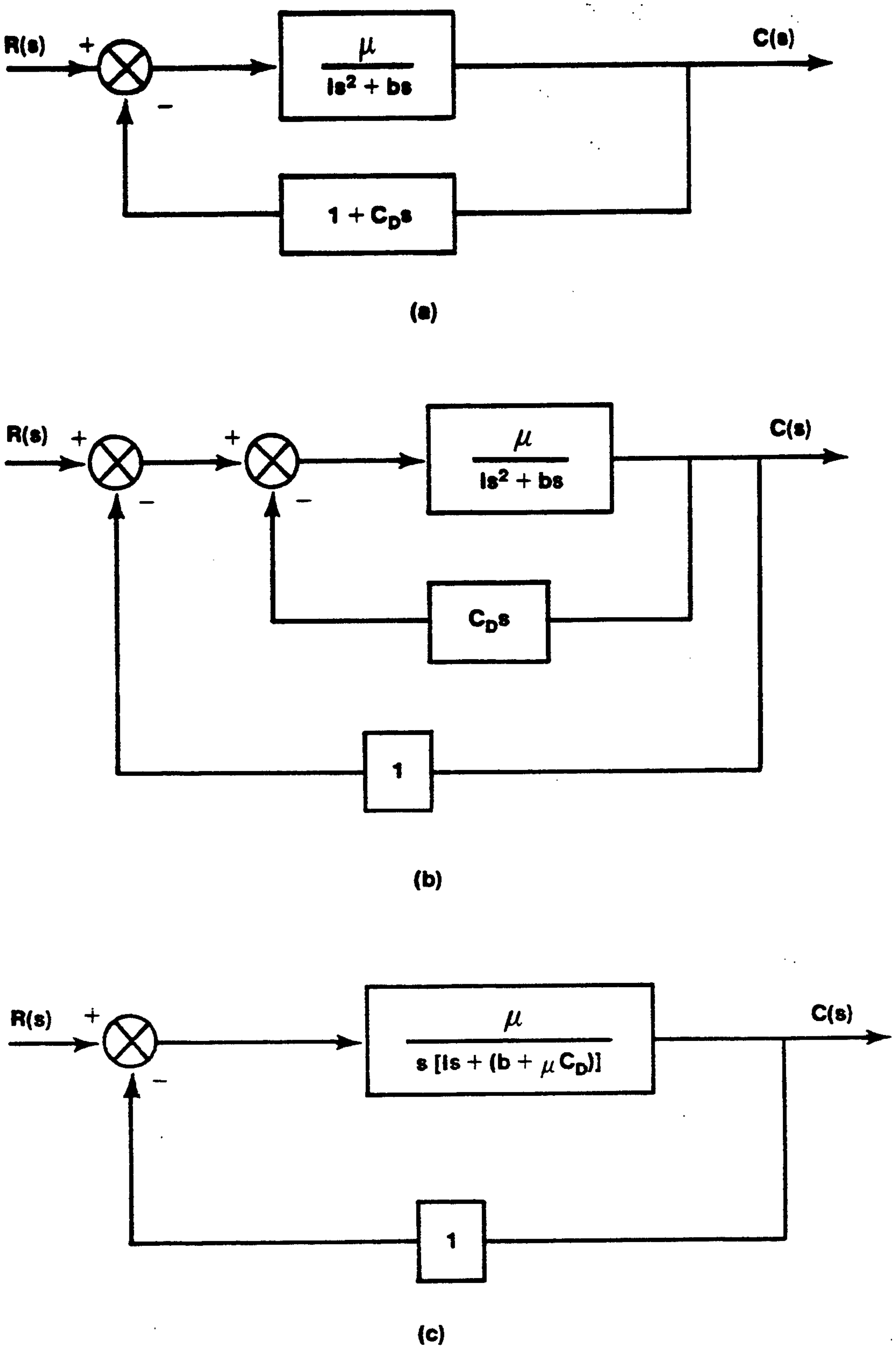


FIGURE 13.69. REDUCTION OF DERIVATIVE CONTROL BLOCK DIAGRAM TO UNITY FEEDBACK

The corresponding error coefficients are

$$K_p = \infty$$

$$K_v = \frac{\mu}{(b + \mu C_D)}$$

$$K_a = 0$$

Therefore we see the steady-state error will increase with the introduction of derivative control (system type remains the same).

We have discussed two types of feedback compensation and have found that the transient response can be improved by addition of derivative control, but only at the expense of the steady-state error. The other compensation techniques to be considered are used in the forward loop.

13.12.2 Cascade Compensation

The first forward compensation technique to be discussed is error rate compensation.

13.12.2.1 Error Rate Compensation. Error rate or ideal derivative compensation is used when the transient response of the system must be improved. This is achieved by reshaping the root locus by moving it to the left. This in effect decreases the system time constant, $\tau = 1/\zeta\omega_n$, thereby speeding up the response. Error rate compensation is achieved by adding the rate of change of the error signal to the error signal (Figure 13.70).

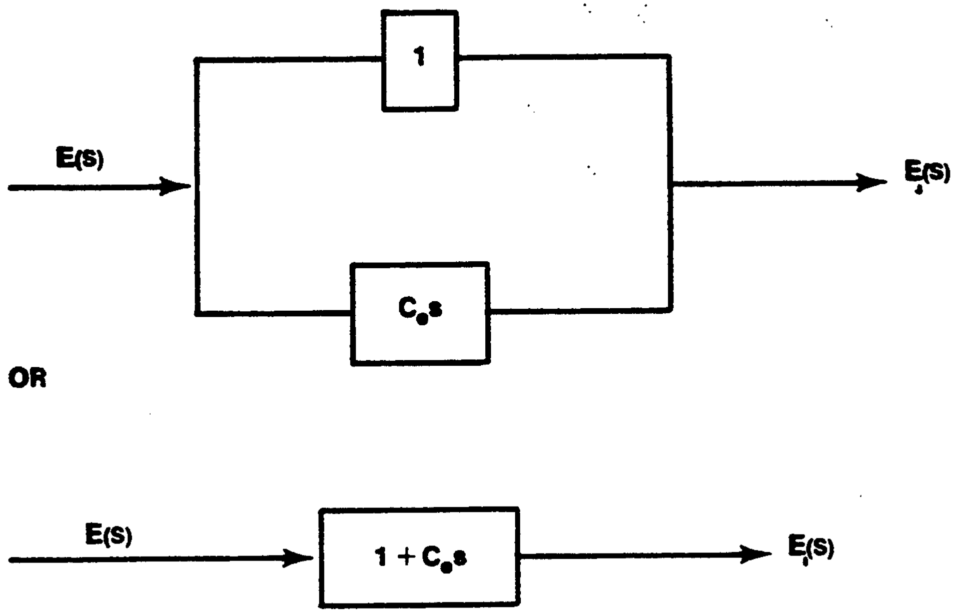


FIGURE 13.70. IDEAL ERROR RATE COMPENSATOR

The physical effect of error rate can be described as introducing anticipation into the system. The system reacts not only to magnitude of the error, but, also to its probable value in the future. If the error is changing rapidly, then $E_1(s)$ is large and the system responds faster. The net result is to speed up the system. Figure 13.71 shows error rate added to the basic system.

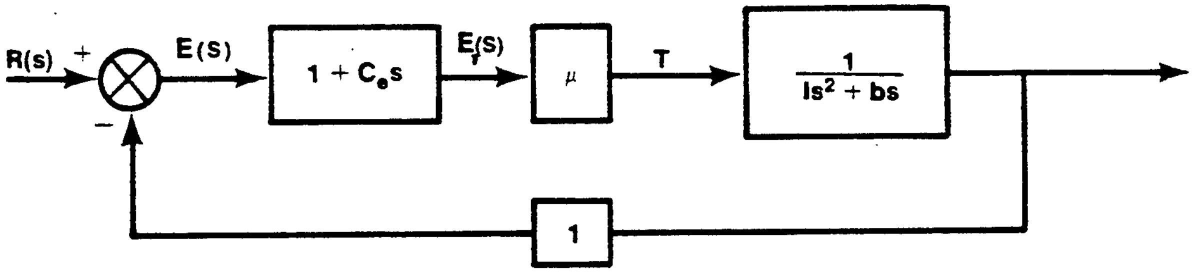


FIGURE 13.71. BASIC SYSTEM WITH ERROR RATE CONTROL

Now the forward transfer function becomes

$$G(s) = \frac{\mu(1 + C_e s)}{Is^2 + bs}$$

$$= \frac{\frac{\mu}{I} C_e \left(s + \frac{1}{C_e}\right)}{s\left(s + \frac{b}{I}\right)}$$

and

$$H(s) = 1$$

Therefore the open-loop transfer function is

$$GH(s) = \frac{\frac{\mu}{I} C_e \left(s + \frac{1}{C_e}\right)}{s\left(s + \frac{b}{I}\right)}, \quad \text{and}$$

$$\frac{C(s)}{R(s)} = \frac{\frac{\mu}{I}\ C_e \left(s + \frac{1}{C_e}\right)}{s^2 + \left(\frac{b + C_e\mu}{I}\right)s + \frac{\mu}{I}}$$

We see that error rate has added a zero to the function as well as achieving artificial damping similar to derivative control. Although the open-loop transfer function, GH(s), of the derivative control and error rate control are similar, the closed-loop transfer functions are different due to the cascade zero of the error rate compensator. The root locus for error rate is illustrated in Figure 13.72. The zero, in effect, draws the locus to the left; thereby speeding up the response and making the system more stable.

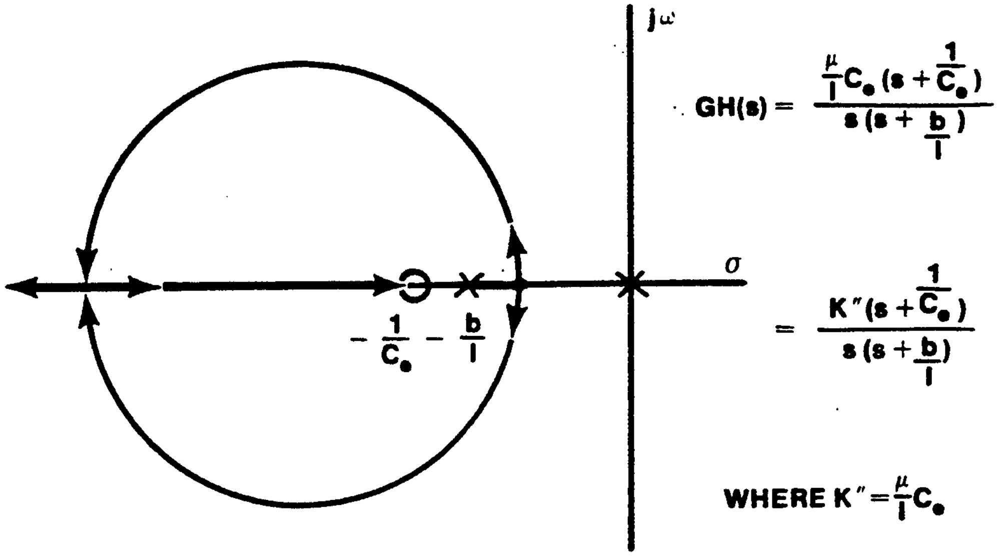


FIGURE 13.72. ROOT LOCUS OF BASIC SYSTEM WITH ERROR RATE COMPENSATION

Looking at the error coefficients we find

$$K_p = \infty$$

$$K_v = \frac{\mu}{b}$$

and

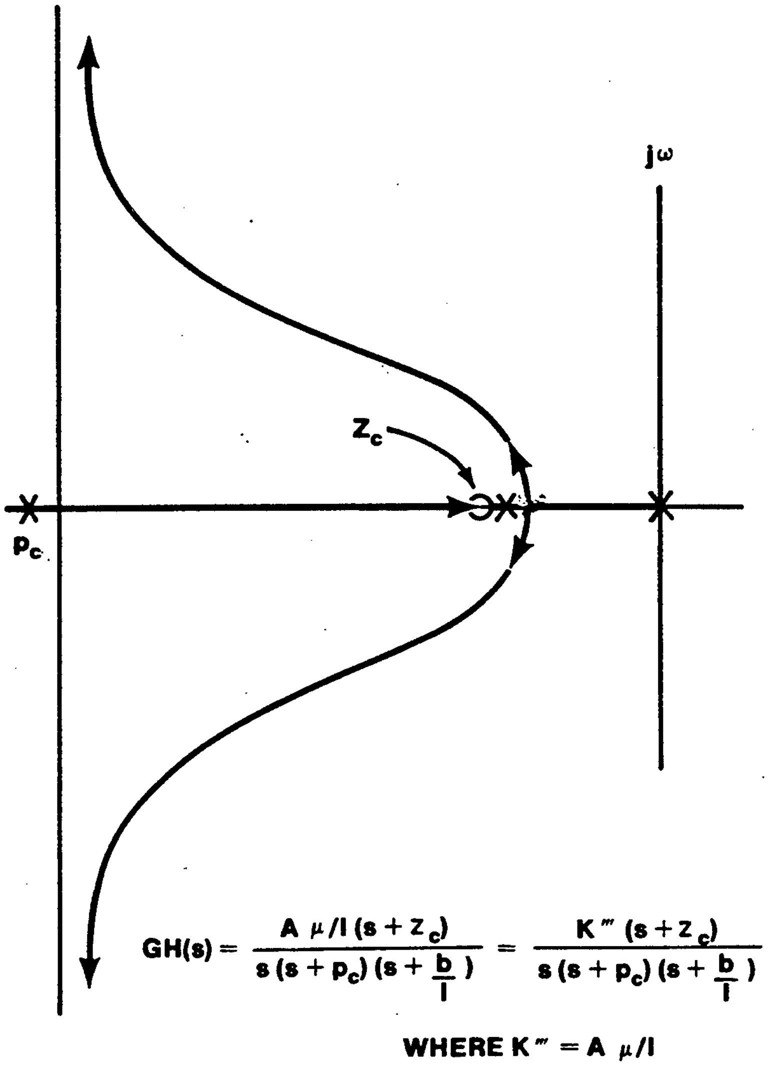
$$K_a = 0$$

Therefore the steady-state performance of the system is unchanged from the basic system.

In the real world an ideal differentiator is difficult to construct and other problems with differentiation of system noise arise. A passive element known as a lead compensator is used to approximate ideal error rate control. The transfer function of this device is

$$G_c(s) = \frac{A\left(s + \frac{1}{\tau}\right)}{\left(s + \frac{1}{\alpha\tau}\right)} = \frac{A(s + z_c)}{(s + p_c)}, \ \alpha < 1$$

The pole p_C, is located far to the left so that the angle of the compensator is nearly all lead due to the zero, z_C, which is placed by trial and error to a point near the original locus. This normally results in a small increase in gain and a large increase in the undamped natural frequency, thereby reducing settling time. Introduction of this lead compensation into the basic system could result in a locus as shown in Figure 13.73.



$$GH(s) = \frac{A\,\mu/l\,(s + Z_c)}{s\,(s + p_c)\,(s + \frac{b}{l})} = \frac{K'''\,(s + Z_c)}{s\,(s + p_c)\,(s + \frac{b}{l})}$$

WHERE $K''' = A\,\mu/l$

FIGURE 13.73. LEAD COMPENSATION APPLIED TO BASIC SYSTEM

13.12.2.2 Integral Control. The second cascade compensation technique is integral control. Often the transient response of a system is satisfactory but the steady-state error is excessive. Integral control produces an actuating signal that is proportional to both the magnitude and the integral of the error signal E(s), Figure 13.74.

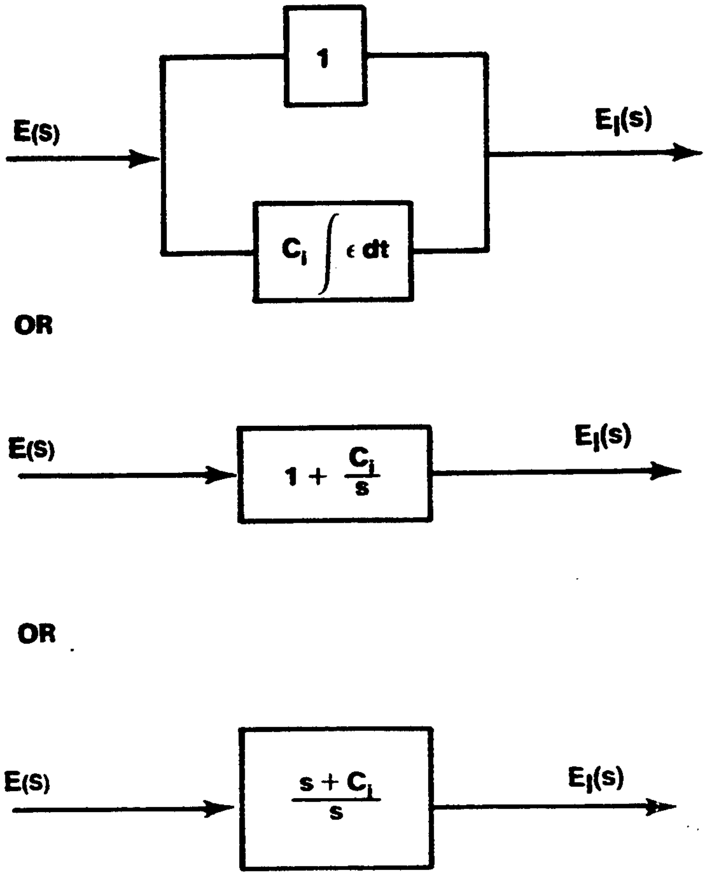


FIGURE 13.74. IDEAL INTEGRAL CONTROL

The net result is the increase in system type. The error $E_1(s)$ continues to increase as long as an error, E(s), is present and eventually becomes large enough to produce an output signal equal to the input. The error, E(s), is then zero. Since we are not interested in changing the time response of the system, the positioning of the zero becomes very important. The pole at the origin has the effect of moving the locus to the right and slowing down the response. The zero must be placed very close to the origin to reduce the effect of the pole on the locus.

Figure 13.75 shows integral control added to out basic system.

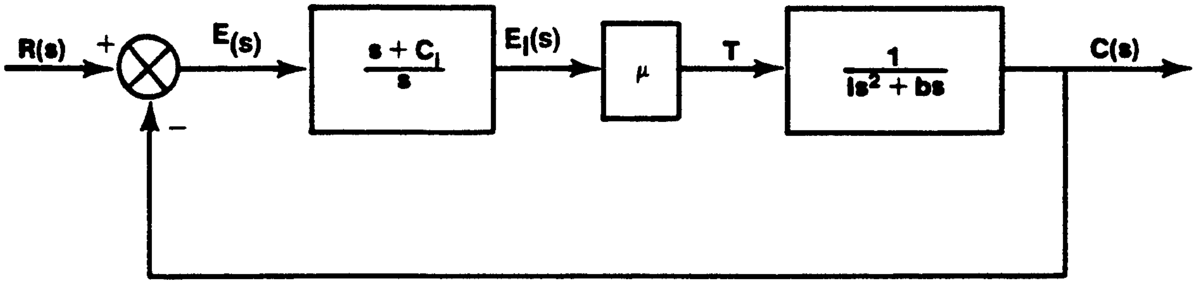


FIGURE 13.75. BASIC SYSTEM WITH INTEGRAL CONTROL

Now the forward transfer function becomes

$$G(s) = \frac{\mu(s + C_i)}{s(Is^2 + bs)} = \frac{\frac{\mu}{I}(s + C_i)}{s^2\left(s + \frac{b}{I}\right)}$$

and

$$H(s) = 1$$

The open-loop transfer function is

$$GH(s) = \frac{\frac{\mu}{I}(s + C_i)}{s^2\left(s + \frac{b}{I}\right)} \qquad (13.92)$$

Figure 13.76 is the root locus of Equation 13.92.

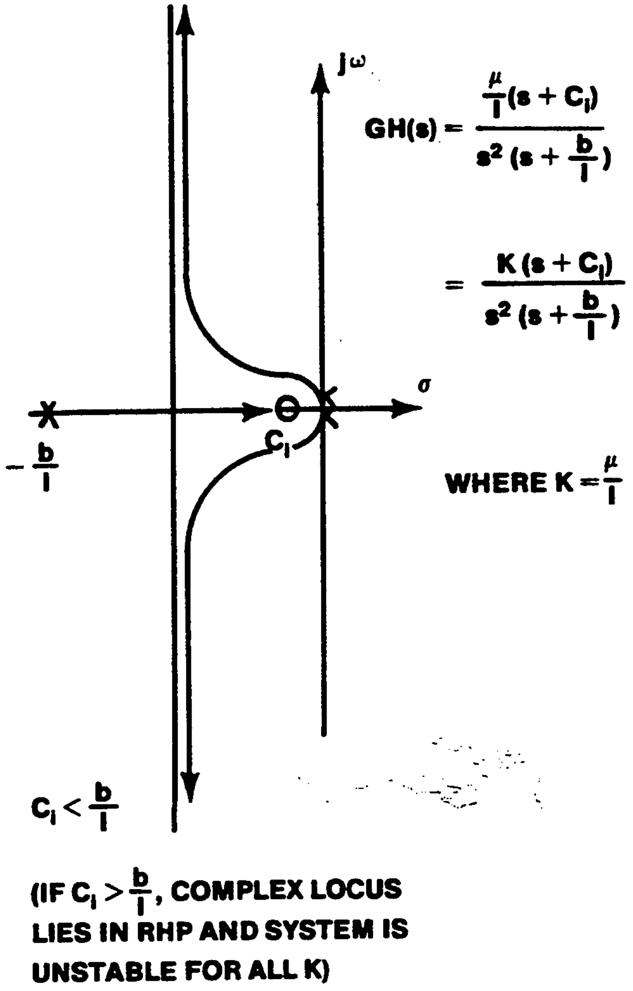


FIGURE 13.76. ROOT LOCUS OF BASIC SYSTEM WITH INTEGRAL CONTROL

The additional pole at the origin increases the system type as well as the scaling of K along the loci. The change in the transient response for the ideal integral control is minimized by placing the zero very close to origin. In a real world system there is a limit to achieving this proximity.

The integral control compensator is achieved electrically by a lag network or mechanically by use of an integrating gyroscope.

The lag network transfer function is of the form

$$G_C(s) = \frac{A}{\alpha} \frac{(s + 1/\tau)}{\left(s + \frac{1}{\alpha\tau}\right)} , \qquad (13.93)$$

where $\alpha > 1$ and is usually about 10; therefore, Equation 13.93 is approximately

$$G_c(s) \cong \frac{A}{\alpha} \frac{(s + 1/\tau)}{s}$$

In design, if the pole and zero are placed very close (at the origin), the net angle contribution at the dominate poles can be kept to less than 5^o and the locus is displaced only slightly. The resulting increased gain of the system and increased system type all result in decreased steady-state error.

The error coefficient for the compensated system became

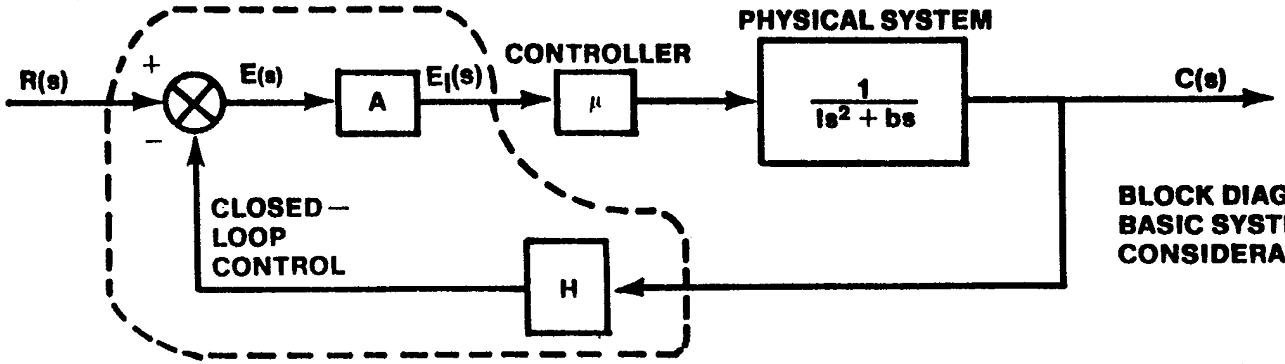
$$K_p = \infty$$

$$K_v = \infty$$

$$K_a = \frac{\mu C_i}{b}$$

and the system will now handle a parabolic input with a $e(t)_{ss} = R\, b/\mu C_i$.

A summary of passive compensations contained in Table 13.5.



BLOCK DIAGRAM OF
BASIC SYSTEM UNDER
CONSIDERATION

Table 13.5
PASSSIVE COMPENSATION

NAME	A	H	ω_n	ζ	$E_I(s)$	STEADY-STATE ERROR $[e(t)_{ss}]$		
						STEP	RAMP	PARABOLIC
PROPORTIONAL	1	1	$\sqrt{\frac{\mu}{I}}$	$\frac{b}{2\sqrt{\mu I}}$	$R - C$	$e(t)_{ss} = 0$	$e(t)_{ss} = \frac{b}{\mu} R$	INFINITE ERROR
DERIVATIVE	1	$(1 + C_D s)$	$\sqrt{\frac{\mu}{I}}$	$\left(\frac{b}{\mu} + C_D\right)\frac{1}{2}\sqrt{\frac{\mu}{I}}$	$R - C - C_D s C$	$e(t)_{ss} = 0$	$e(t)_{ss} = (\frac{b}{\mu} + C_D) R$	INFINITE ERROR
ERROR RATE	$(1 + C_e s)$	1	$\sqrt{\frac{\mu}{I}}$	$\left(\frac{b}{\mu} + C_e\right)\frac{1}{2}\sqrt{\frac{\mu}{I}}$	$E + C_e s E$	$e(t)_{ss} = 0$	$e(t)_{ss} = \frac{b}{\mu} R$	INFINITE ERROR
INTEGRAL	$\left(1 + \frac{C_I}{s}\right)$	1	3RD ORDER	3RD ORDER	$E + C_I \quad E/s$	$e(t)_{ss} = 0$	$e(t)_{ss} = 0$	$e(t)_{ss} = \frac{Rb}{\mu C_I}$

13.13 SUMMARY

In this chapter an attempt has been made to present the fundamentals of control system analysis. Applications of the theory were held to a minimum so that full attention could be devoted to learning the tools and techniques used in this type of analysis.

Once the analysis techniques have been mastered, the more interesting appropriate areas of flight control and handling qualities may be addressed. A knowledge of root locus theory and frequency response is essential in understanding the applications of feedback analysis to flight vehicle systems. Despite the introduction of modern control theory (state variables) and digital flight control, an understanding of these systems is still based in large measure on knowledge of classical feedback control systems.

BIBLIOGRAPHY

13.1. D'Azzo, J. J. and Houpis, C. H. Linear Control System Analysis and Design, Conventional & Modern. McGraw Hill Book Co., New York, N. Y., 2nd Edition, 1981.

13.2. Reid, J. G. Linear System Fundamentals. McGraw Hill Book Co., New York, N. Y., 1983.

13.3. James, H. M., Nichols, M. B. and Phillips, R. S. Theory of Servomechanisms. McGraw Hill Book Co., New York, N. Y., 1947.