Nyquist/Bode Frequency Response Plots and System Stability

8

8.1 INTRODUCTION

In the preceding chapters, we have presented techniques for the analysis and design of feedback control systems based on pole-zero (root locus) formalism. An overview of the control-system design procedure is as follows.

Step 1: Construct transfer function models for the controlled process, the actuator, and the sensor. Validate the models with experimental data where possible.

Step 2: Investigate the stability properties of a feedback structure of the form shown in Fig. 8.1.
The stability as we know, is dictated by the poles of the closed-loop transfer function

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

Given the open-loop transfer function $G(s)H(s)$, the stability of the closed-loop transfer function can be determined using Routh stability criterion, without actually computing the closed-loop poles. The root locus plots are more informative.

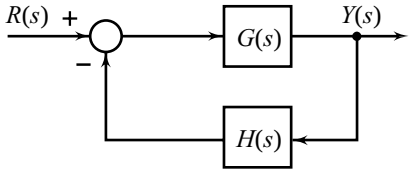


Fig. 8.1 A feedback control system

Step 3: Translate the performance requirements on time response into pole-zero specifications.
A typical result of this step is a requirement that the system has a step response with specified limits on rise time, peak overshoot, and settling time. This requirement is translated into a pair of dominant poles in the specified region of the s-plane. The maximum allowable steady-state error in tracking standard test signals is also specified.

Step 4: Shape the root locus plot by simple cascade compensators and/or minor-loop feedback compensators. Try to meet the specifications on dominant closed-loop poles and steady-state accuracy.
Compare the trial-and-error compensators with respect to parametric sensitivity, and the effects of sensor noise and external disturbances. If a design seems satisfactory, go to step 6; otherwise try step 5.

Step 5: Re-evaluate the specifications, the physical configuration of the process, and the actuator and sensor selections in the light of the current design and go to step 1.

Step 6: Build a computer model and compute (simulate) the performance of the design (Appendix A). The computer model of the system must include important nonlinearities, and other modelling uncertainties. Such a simulation of the design will confirm robustness, and allow one to predict the true performance to be expected from the system. If the performance is not satisfactory, go to step 1.

Step 7: Build a prototype.
As a final test before production, a prototype can be built and tested. After these tests, one might want to reconsider the sensor, actuator, and process, and conceivably return to step 1—unless time, money, or ideas have run out.

The root locus design technique, developed by Walter R. Evans in 1948, is highly attractive because it offers the designer the advantage of dealing directly with the poles and zeros of a closed-loop system, enabling him/her to exert direct influence on the dynamic behaviour of the system. By the judicious location of poles and zeros of compensation devices, the designer can see almost at a glance the manner in which the transient response of the closed-loop system is affected. Construction of the root locus plots to acceptable engineering accuracy is straightforward. The success of the method is, however, dependent upon the availability of reasonably accurate pole-zero models of the process, the actuator, and the sensor. Robustness against modelling uncertainties may not be satisfactory.

We now turn our attention to a design technique that uses only the information of $G(s)$ and $H(s)$ along the positive imaginary axis, that is, $G(j\omega)$ and $H(j\omega)$ for all $\omega \geq 0$; and is called the *frequency domain technique*. Chronologically, the frequency-domain technique was the first method developed to design control systems. Harry Nyquist in 1932 published his study of stability theory, which is the foundation of the frequency-domain approach to control system analysis and design. The Nyquist criterion is a graphical procedure for illustrating system stability, or lack thereof. The required Nyquist plot can usually be sketched quickly without extensive mathematical calculations. System stability and the margins by which that stability is obtained (relative stability) are readily apparent from the sketch.

The general objective in all the frequency-domain design procedures is to shape the Nyquist plot to achieve acceptable closed-loop response characteristics. Direct use of the Nyquist plot for design is not particularly convenient, since changes in parameters other than gain require extensive plot revisions. The work of Hendrik W. Bode during the 1930s led to a more efficient design procedure: the general characteristics of the Nyquist plot can be visualized with reasonable accuracy in most cases of interest from the Bode plot and the Bode plot is easily constructed and modified. The frequency-domain design procedures are reduced to orderly graphical trial-and-error processes carried out using the Bode plot. Thus, although the Nyquist plot characteristics form the basis for all of our designs, the mechanics of the design procedures are readily established on the Bode plot without requiring the construction of Nyquist plot.

A preview of the frequency-domain design procedure is as follows.

Step 1: Construct transfer function models for the controlled process, the actuator, and the sensor. Validate the models $G(s)$ and $H(s)$ (refer Fig. 8.1) with experimental data where possible. Compute $G(j\omega)$ and $H(j\omega)$; $\omega \geq 0$.

Note that the frequency-domain methods use only $G(j\omega)$ and $H(j\omega)$ which, for stable devices, can be obtained by direct measurement. Once $G(j\omega)$ and $H(j\omega)$ are measured, we may proceed directly to the design without computing the transfer functions $G(s)$ and $H(s)$.

Step 2: Investigate stability properties of a feedback structure of the form shown in Fig. 8.1. Given the data $G(j\omega)$ and $H(j\omega)$; $\omega \geq 0$, computed from $G(s)$ and $H(s)$ with s replaced by $j\omega$, or obtained by experimental measurement, the stability of the closed-loop system can be determined by the Nyquist stability criterion using Nyquist plots/Bode plots.

Step 3: Translate the performance requirements into frequency-domain specifications. There is a definite correlation between the time-domain and frequency-domain modes of behaviour. A common procedure is to interpret desired time-domain behaviour (step response with specified limits on rise time, peak overshoot, and settling time) in terms of frequency-domain characteristics. Design is carried out in the frequency domain and is translated back into the time domain. There is a constant interplay of characteristics between the two domains.

Thus the correlation between the time-domain and frequency-domain modes of behaviour plays a major role in the design procedure. Unfortunately, this correlation is not a simple one mathematically. Accordingly, in dealing with the problem of stabilizing and compensating control systems by using the frequency-domain techniques, control over the time-domain behaviour is most conveniently secured in terms of appropriate figures of merit such as gain margin, phase margin, bandwidth, etc., which although defined in frequency domain, are used as indicators of the performance in time domain. These figures of merit provide only an approximate correlation between the time-domain and frequency-domain characteristics. This is probably the major limitation of the frequency-domain approach; it sacrifices direct control on the time-domain performance. The major gain is the robustness against modelling uncertainties and external disturbances.

Step 4: Shape the Bode plot by simple cascade compensators and/or minor-loop feedback compensators. Try to meet the specifications on transient response (given in terms of frequency-domain measures) and steady-state accuracy.

Compare the trial-and-error compensators with respect to parameteric sensitivity, and the effects of sensor noise and external disturbances. If a design seems satisfactory, go to step 6; otherwise try step 5.

Step 5: Re-evaluate the specifications, the physical configuration of the process, and the actuator and sensor selections in the light of the current design and go to step 1.

Step 6: Build a computer model and compute (simulate) the performance of the design. The computer model must include important nonlinearities, and other modelling uncertainties (Appendix A).

Since the preliminary design is based on the approximate correlations between time-domain and frequency-domain modes of behaviour, the time-domain specifications must be checked before final parameter choices are made. A simulation study is generally the most appropriate method for this purpose.

Step 7: Build a prototype.

It may be noted that each design method, root locus or Bode plot, has its own particular use and advantage in a particular situation. As the reader becomes fully acquainted with these methods, he/she will become aware of the potentialities of each and will know when each should be used. We normally attempt the design using both the time-domain and frequency-domain approaches; this way one can gain an additional perspective to the complex analysis and design problems of feedback control systems.

The following strengths of the frequency-domain approach will become clear as we take up the subject.

1. The method does not need precise mathematical description of systems. The root locus method needs reasonably accurate mathematical description. The frequency-domain approach leads to robust design.
2. The method is independent of the complexity of systems and is applicable to systems containing time-delay elements. The root locus method, as we know, needs an approximate description of time-delay elements.
3. The method gives a simple and orderly approach for trial-and-error design.
4. It provides a very good indication of the system bandwidth which often appears explicitly in the specifications, and can only be approximated with the root-locus method of design.

The major limitation is that the direct control on the time-domain performance is lost.

The present chapter is concerned with the stability analysis in frequency domain. A frequency-domain stability criterion was developed by H. Nyquist in 1932 and remains a fundamental approach to the investigation of stability of linear control systems. In addition to answering the question of absolute stability, this criterion also gives some useful results on relative stability. The relative stability measures given by the Nyquist stability criterion are, in fact, central to the great importance of frequency-domain design methods.

We first develop the Nyquist stability criterion. Two common types of frequency-domain plots (Nyquist plots, Bode plots) are then introduced, and using these plots, absolute and relative stability are investigated on the basis of the Nyquist stability criterion. The use of the Nyquist criterion in frequency-domain design procedures will appear in the next two chapters.

8.2 DEVELOPMENT OF THE NYQUIST STABILITY CRITERION

In order to investigate the stability of a control system, we consider the closed-loop transfer function

$$\frac{Y(s)}{R(s)} = \frac{G(s)}{1 + G(s)H(s)} \tag{8.1}$$

keeping in mind that the transfer functions of both the single-loop and the multiple-loop control systems can be expressed in this form. The characteristic equation of the closed-loop system is obtained by setting the denominator of $Y(s)/R(s)$ to zero, which is same as setting the numerator of $1 + G(s)H(s)$ to zero. Thus, the roots of the characteristic equation must satisfy

$$1 + G(s)H(s) = 0 \tag{8.2}$$

We assume at this point that $G(s)H(s)$ can be expressed by a ratio of finite algebraic polynomials in s. This assumption has been made for convenience; the Nyquist criterion applies to more general situations, as shown subsequently. Let

$$G(s)H(s) = \frac{K(s + z_1')(s + z_2') \cdots (s + z_m')}{(s + p_1)(s + p_2) \cdots (s + p_n)}; \; m \le n \tag{8.3}$$

$G(s)H(s)$ is the product of plant, compensator, and sensor transfer functions; its pole and zero locations are assumed to be known since these transfer functions are generally available in factored form. Substituting for $G(s)H(s)$ from Eqn. (8.3) into Eqn. (8.2), we obtain

$$\begin{aligned} 1 + G(s)H(s) &= 1 + \frac{K(s + z_1')(s + z_2') \cdots (s + z_m')}{(s + p_1)(s + p_2) \cdots (s + p_n)} \\ &= \frac{(s + p_1)(s + p_2) \cdots (s + p_n) + K(s + z_1')(s + z_2') \cdots (s + z_m')}{(s + p_1)(s + p_2) \cdots (s + p_n)} \\ &= \frac{(s + z_1)(s + z_2) \cdots (s + z_n)}{(s + p_1)(s + p_2) \cdots (s + p_n)} \end{aligned} \tag{8.4}$$

It is apparent from Eqn. (8.4) that the poles of $1 + G(s)H(s)$ are identical to those of $G(s)H(s)$, i.e., the open-loop poles of the system; and the zeros of $1 + G(s)H(s)$ are identical to the roots of the characteristic equation, i.e., the closed-loop poles of the system. For the closed-loop system to be stable, the zeros of $1 + G(s)H(s)$ must lie in the left half of the s-plane. It is important to note that even if some of the open-loop poles lie in the right half s-plane, all the zeros of $1 + G(s)H(s)$, i.e., the closed-loop poles may lie in the left half s-plane, meaning thereby that an open-loop unstable system may lead to closed-loop stable operation.

Recall that we have introduced two methods of checking whether or not all zeros of $1 + G(s)H(s)$ have negative real parts. The zeros of $1 + G(s)H(s)$ are roots of the polynomial

$$(s + p_1)(s + p_2) \cdots (s + p_n) + K(s + z_1')(s + z_2') \cdots (z + z_m')$$

and we may apply the Routh test. Another method is to plot the root loci of

$$\frac{K(s + z_1')(s + z_2') \cdots (s + z_m')}{(s + p_1)(s + p_2) \cdots (s + p_n)} = -1$$

as a function of K. In this section, we shall introduce yet another method of checking whether or not all zeros of $1 + G(s)H(s)$ lie inside the open left half s-plane. The method, called the *Nyquist stability criterion*, is based on the *principle of argument* in the theory of complex variables [32]. The basic concept used in the Nyquist criterion is explained below (more details in Review Example 8.1).

Consider a rational function

$$1 + G(s)H(s) = Q(s) = \frac{s + z_1}{s + p_1}$$

The pole-zero map of $Q(s)$ is shown in Fig. 8.2a. This figure also shows a closed contour Γ_1. A point or an area is said to be *enclosed* by a closed path if it is found to lie to the right of the path when the path is traversed in the clockwise direction. The pole $-p_1$ and the zero $-z_1$ are therefore enclosed by the contour Γ_1 in Fig. 8.2a.

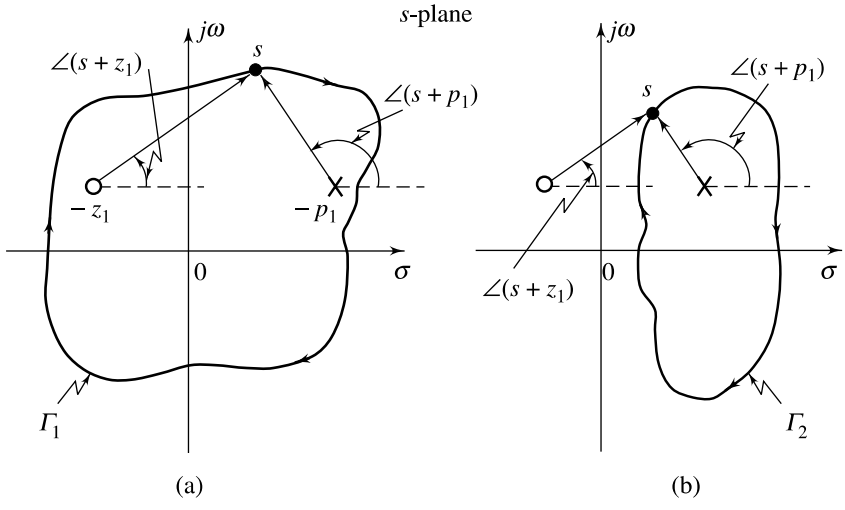


Fig. 8.2 (a) A contour which encloses both the pole and the zero
(b) A contour which encloses only the pole

Now we see that as the point s follows the prescribed path (i.e., clockwise direction) on the s-plane contour Γ_1, both $\angle(s + z_1)$ and $\angle(s + p_1)$ decrease continuously. For one clockwise traversal of Γ_1, we have

$$\delta_{\Gamma_1}\angle(s + z_1) = -2\pi$$
$$\delta_{\Gamma_1}\angle(s + p_1) = -2\pi$$

where $\delta_{\Gamma_1}\angle$ indicates the change in angle as Γ_1 is traversed.

$$\delta_{\Gamma_1}\angle Q(s) = \delta_{\Gamma_1}\angle(s+z_1) - \delta_{\Gamma_1}\angle(s+p_1)$$
$$= -2\pi - (-2\pi) = 0$$

Thus the change in $\angle Q(s)$ as s traverses Γ_1 (which is a completely general contour in the s-plane except that it encloses both $-z_1$ and $-p_1$) once in clockwise direction, is zero. Using these arguments on the contour Γ_2 in the s-plane which encloses only $-p_1$ and not $-z_1$ (Fig. 8.2b), we find that (for the zero $s = -z_1$ not enclosed by the contour Γ_2, $\delta_{\Gamma_2}\angle(s+z_1) = 0$ for one clockwise traversal of Γ_2)

$$\delta_{\Gamma_2}\angle Q(s) = 0 - (-2\pi) = 2\pi$$

We thus see that a change in $\angle Q(s)$ that one gets in traversing a contour in the s-plane, is strictly a function of how many poles and zeros of $Q(s)$ are enclosed by the s-plane contour. Let us consider a general $Q(s) = 1 + G(s)H(s)$, given by Eqn. (8.4), for which a typical pole-zero plot might be made as shown in Fig. 8.3. The figure also shows an s-plane contour Γ_s which encloses Z zeros and P poles of $Q(s)$. Note that Γ_s does not go through any of the poles or zeros of $Q(s)$. If we make one clockwise traversal of the contour Γ_s, it may be seen that

$$\delta_{\Gamma_s}\angle Q(s) = Z(-2\pi) - P(-2\pi)$$
$$= (P - Z)\,2\pi \qquad (8.5)$$

Another way of looking at the result of Eqn. (8.5) is to consider $Q(s)$ given by Eqn. (8.4) evaluated on the s-plane contour Γ_s of Fig. 8.3 and plotted in the $Q(s)$-plane. For every point $s = \sigma + j\omega$ on the s-plane contour Γ_s, we obtain $Q(s) = \mathrm{Re}Q + j\mathrm{Im}Q$. Alternatively, it can be stated that the function $Q(s)$ *maps* the point $\sigma + j\omega$ in the s-plane into the point $\mathrm{Re}Q + j\mathrm{Im}Q$ in the $Q(s)$-plane. It follows that for the closed contour Γ_s in the s-plane, there corresponds a closed contour Γ_Q in the $Q(s)$-plane. A typical Γ_Q is shown in Fig. 8.4. The arrowheads on this contour indicate the direction that $Q(s)$ takes as s moves on Γ_s in the clockwise direction. Now as we traverse Γ_Q once in the direction indicated by arrowheads, contour Γ_s in the s-plane is traversed once in the clockwise direction. For the example of Fig. 8.4, traversing Γ_Q once in the clockwise direction gives a change in $\angle Q(s)$ of -4π, since the origin is encircled twice in the negative direction. Thus Eqn. (8.5) gives

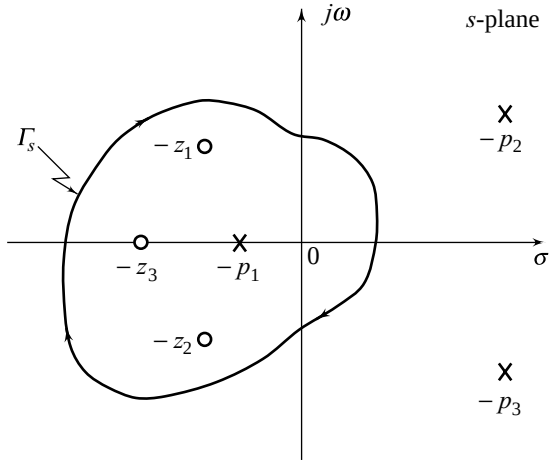


Fig. 8.3 A general pole-zero plot with contour Γ_s which encloses Z zeros and P poles

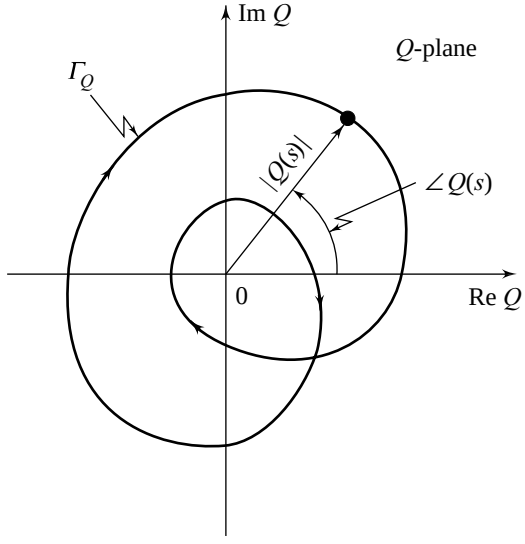


Fig. 8.4 Q(s) evaluated on the contour Γ_s in Fig. 8.3

$$\delta_{\Gamma_s} \angle Q(s) = 2\pi(P - Z) = -4\pi$$

or $$Z - P = 2$$

In general,

$$Z - P = N \tag{8.6}$$

where N is the number of clockwise encirclements of the origin that Γ_Q makes in the $Q(s)$-plane. It may be noted that we are not interested in the exact shape of the $Q(s)$-plane contour. The important fact that concerns us is the encirclements of the origin by the $Q(s)$-plane contour.

The relation (8.6) between the enclosure of poles and zeros of $Q(s)$ by the s-plane contour and the encirclements of the origin by the $Q(s)$-plane contour is commonly known as the *principle of argument* which may be stated as follows.

Let $Q(s)$ be a ratio of polynomials in s. Let P be the number of poles and Z be the number of zeros of $Q(s)$ which are enclosed by a simple closed contour in the s-plane, multiplicity accounted for. Let the closed contour be such that it does not pass through any poles or zeros of $Q(s)$. The s-plane contour then maps into the $Q(s)$-plane contour as a closed curve.

The number N of clockwise encirclements of the origin of the $Q(s)$-plane, as a representative point s traces out the entire contour in the s-plane in the clockwise direction, is equal to $Z - P$.

Since $$Q(s) = 1 + G(s)H(s),$$

we can obtain $Q(s)$-plane contour from the $G(s)H(s)$-plane contour simply by adding +1 to each point of $G(s)H(s)$-plane contour. Figure 8.5 shows a typical $G(s)H(s)$-plane contour Γ_{GH}. The effect of adding +1 to each point of Γ_{GH} to obtain $Q(s)$-plane contour Γ_Q is accomplished simply by adding +1 to the scale of the real axis, as shown by the numbers in parentheses in Fig. 8.5. It is seen that $-1+j0$ point of the $G(s)H(s)$ map Γ_{GH} corresponds to the origin of the $Q(s)$ map Γ_Q. For convenience, we designate the $-1+j0$ point of the $G(s)H(s)$-plane as the *critical point*. Thus the encirclements of the origin by the contour Γ_Q is equivalent to the encirclements of the critical point $-1+j0$ by the contour Γ_{GH}. In the light of these observations, we can express Eqn. (8.6) as follows:

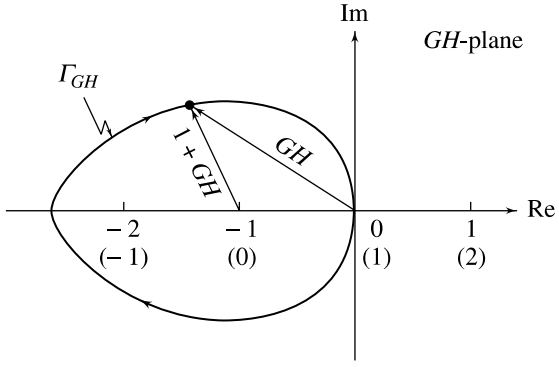


Fig. 8.5 Evaluating $1+G(s)H(s)$ from the $G(s)H(s)$ map

$$Z - P = N \tag{8.7}$$

where

Z = number of zeros of $1+G(s)H(s)$ enclosed by the s-plane contour Γ_s;

P = number of poles of $G(s)H(s)$ enclosed by the s-plane contour Γ_s; and

N = number of clockwise encirclements of the critical point $-1+j0$ made by the $G(s)H(s)$-plane contour Γ_{GH}.

In general N can be positive ($Z > P$), zero ($Z = P$), or negative ($Z < P$). $N > 0$ corresponds to N *net* encirclements of the critical point $-1+j0$ in clockwise direction by the Γ_{GH} contour. $N = 0$ indicates zero *net* encirclements of the critical point by the Γ_{GH} contour. $N < 0$ corresponds to N *net* encirclements of the critical point in counterclockwise direction by the Γ_{GH} contour.

A convenient way of determining N with respect to the critical point $-1+j0$ of the $G(s)H(s)$-plane is to draw a radial line from this point. The number of *net* intersections of the radial line with the Γ_{GH} contour gives the magnitude of N. Figure 8.6 gives several examples of the method of determining N.

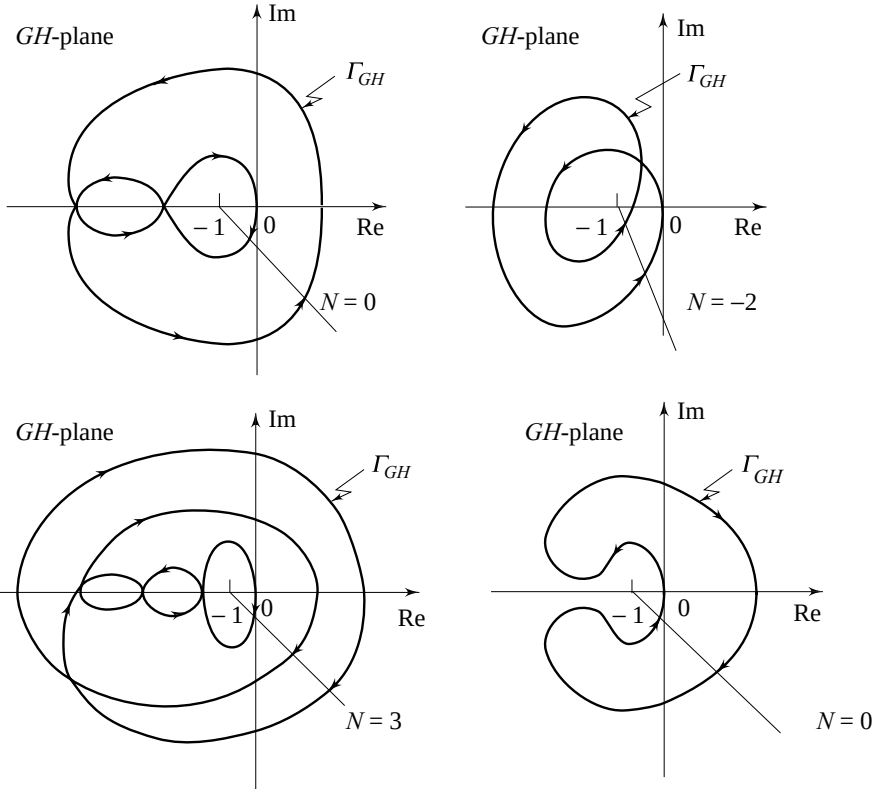


Fig. 8.6 Examples of determination of N

8.2.1 The Nyquist Contour

At this point the reader may place himself/herself in the position of Nyquist many years ago, confronted with the problem of stability of the closed-loop system that has the transfer function of Eqn. (8.1), which is equivalent to determining whether or not the function $1+G(s)H(s)$ has zeros in the right half s-plane. Apparently, Nyquist discovered that the principle of argument of the complex-variable theory could be applied to solve the stability problem if the s-plane contour Γ_s is taken to be one that encloses the entire right half of the s-plane. Of course, as an alternative, Γ_s can be chosen to enclose the entire left half of the s-plane, as the solution is a relative one.

Figure 8.7 illustrates a Γ_s contour that encloses the entire right half of the s-plane. Such a contour is called the *Nyquist contour*. It is directed clockwise and comprises of an infinite line segment C_1 along the $j\omega$-axis and arc C_2 of infinite radius.

Along C_1, $s = j\omega$ with s varying from $-j\infty$ to $+j\infty$.

Along C_2, $s = R\,e^{j\theta}$ with $R \to \infty$ and θ varying from $+90°$ through $0°$ to $-90°$.

As the s-plane contour Γ_s must avoid all poles of $1+G(s)H(s)$, modifications in the Nyquist contour defined in Fig. 8.7 are required when $G(s)H(s)$, and therefore $1+G(s)H(s)$, has one or more poles on the

imaginary axis (the zeros of $1 + G(s)H(s)$ are unknown to us; we will shortly see the implications of s-plane contour passing through the zeros of $1 + G(s)\,H(s)$). The basic trick, of course, is to take a small detour around the imaginary axis poles. Figure 8.8 illustrates a modified Nyquist contour when $G(s)H(s)$ has a pole at $s = 0$. Along the semicircular indent around the pole at the origin, $s = \rho e^{j\phi}$ with $\rho \to 0$ and ϕ varying from $-90°$ through 0° to + 90°. The indented Nyquist contour in Fig. 8.8 does not enclose the pole at the origin. Of course, as an alternative, the Nyquist contour may be indented to enclose the pole at the origin.

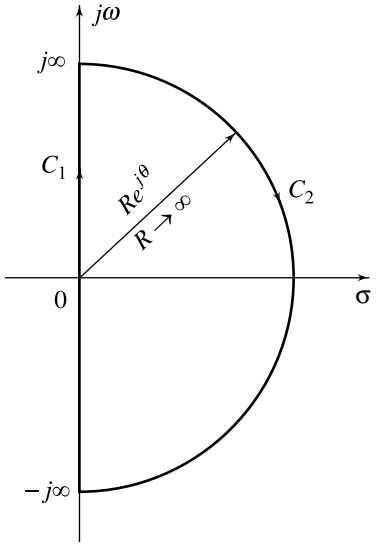


Fig. 8.7 The Nyquist contour

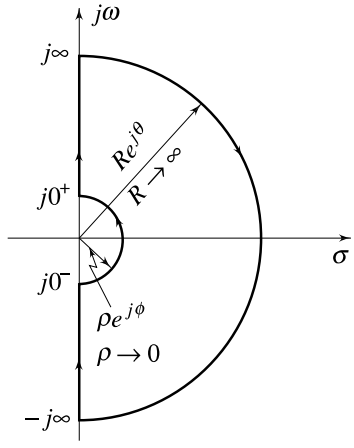


Fig. 8.8 Indented Nyquist contour

8.2.2 The Nyquist Plot

The Nyquist stability criterion is a direct application of the principle of argument when the s-plane contour Γ_s is the Nyquist contour. In principle, once the Nyquist contour is specified, the stability of a closed-loop system can be determined by plotting the $G(s)H(s)$ locus when s takes on values along the Nyquist contour, and investigating the behaviour of the $G(s)H(s)$ plot with respect to the critical point $-1+j0$. The $G(s)H(s)$ plot that corresponds to the Nyquist contour is called the *Nyquist plot* of $G(s)H(s)$.

With the additional dimension added to the stability problem, we define N, P, and Z as follows.

N = number of clockwise encirclements of the critical point $-1+j0$ made by the $G(s)H(s)$ locus of the Nyquist plot.

P = number of poles of $G(s)H(s)$, and therefore of $1+ G(s)H(s)$, enclosed by the Nyquist contour.

Then (refer Eqn. (8.7))

Z = number of zeros of $1+ G(s)H(s)$ enclosed by the Nyquist contour

$$= N + P \tag{8.8a}$$

Closed-loop stability requires that

$$Z = 0 \tag{8.8b}$$

This condition is met if

$$N = -P \tag{8.8c}$$

In the special case[1] (this is generally the case in most single-loop practical systems) of $P = 0$ (i.e., the open-loop transfer function $G(s)H(s)$ has no poles in right half s-plane), the closed-loop system is stable if

$$N = 0 \tag{8.8d}$$

8.2.3 The Nyquist Criterion

We can now state the Nyquist stability criterion as follows:

If the Nyquist plot of the open-loop transfer function G(s)H(s) corresponding to the Nyquist contour in the s-plane encircles the critical point $-1+j0$ in the counterclockwise direction as many times as the number of right half s-plane poles of G(s)H(s), the closed-loop system is stable.

In the commonly occurring case of G(s)H(s) with no poles in right half s-plane, the closed-loop system is stable if the Nyquist plot of G(s)H(s) does not encircle the $-1+j0$ point.

When the Nyquist plot of $G(s)H(s)$ passes through $-1+j0$ point, the number of encirclements N is indeterminate. This corresponds to the condition where $1+G(s)H(s)$ has zeros on the imaginary axis. A necessary condition for applying the Nyquist criterion is that the Nyquist contour must not pass through any poles or zeros of $1+G(s)H(s)$. When this condition is violated, the value for N becomes indeterminate and the Nyquist stability criterion cannot be applied.

8.3 SELECTED ILLUSTRATIVE NYQUIST PLOTS

The following examples serve to illustrate the application of the Nyquist criterion to the stability study of control systems.

Example 8.1 Consider a single-loop feedback control system with the open-loop transfer function given by

$$G(s)H(s) = \frac{K}{(\tau_1 s + 1)(\tau_2 s + 1)};\ \tau_1,\ \tau_2 > 0,\ K > 0 \tag{8.9}$$

$G(s)H(s)$ has no poles in the right-half s-plane; therefore, stability is assured if the Nyquist plot of $G(s)H(s)$ does not encircle the $-1+j0$ point.

The Nyquist plot of $G(s)H(s)$, as we know, is the mapping of the Nyquist contour in the s-plane onto $G(s)H(s)$-plane. In Fig. 8.9a, the Nyquist contour has been divided into three sections: C_1, C_2, and C_3. Section C_1 is defined by $s = j\omega$, $0 \le \omega < \infty$; section C_2 is defined by $s = j\omega$, $-\infty < \omega \le 0$; and section C_3 is defined by $s = Re^{j\theta}$, $R \to \infty$, and θ varies from $+90°$ through $0°$ to $-90°$.

Mapping of Section C_1 onto G(s)H(s)-plane

Substituting $s = j\omega$ into $G(s)H(s)$, we obtain

$$G(j\omega)H(j\omega) = \frac{K}{(j\omega\tau_1 + 1)(j\omega\tau_2 + 1)} \tag{8.10}$$

A plot of $G(j\omega)H(j\omega)$ on polar coordinates as ω is varied from 0 to ∞ is the map of section C_1 on the $G(s)H(s)$-plane. This plot is called the *polar plot* of sinusoidal transfer function $G(j\omega)H(j\omega)$.

[1]Example of a physical system whose transfer function model has a pole in right-half s-plane is given in Figs 12.14.

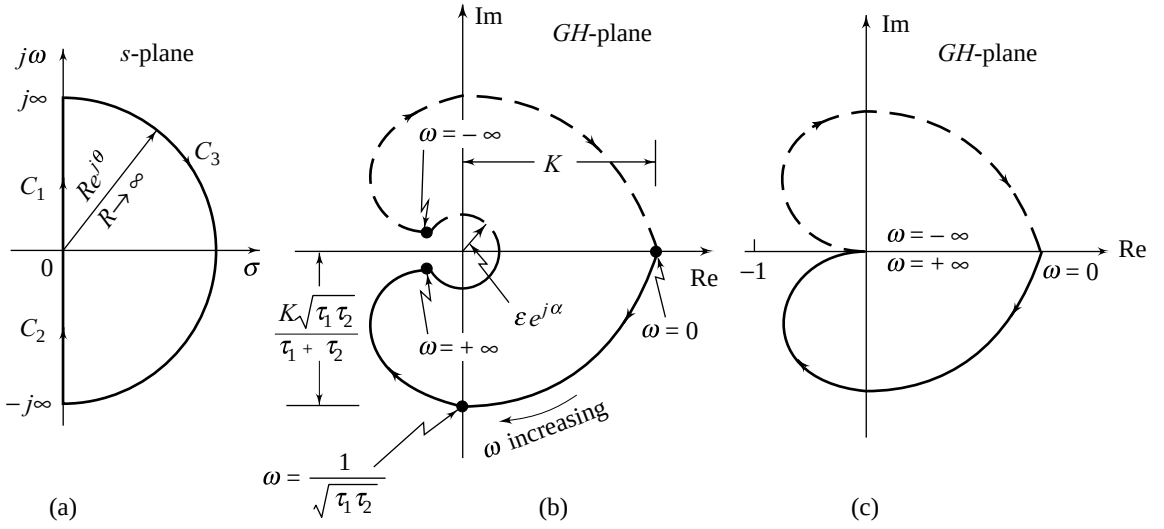


Fig. 8.9 (a) Nyquist contour; (b) Nyquist plot of $G(s)H(s)$ of Eqn. (8.9); (c) Simplified Nyquist plot

For application of the Nyquist stability criterion, an exact polar plot of $G(j\omega)H(j\omega)$ is not essential. Often a rough sketch is adequate for stability analysis. The general shape of the polar plot of $G(j\omega)H(j\omega)$ may be determined from the following information.

1. The behaviour of the magnitude and phase of $G(j\omega)H(j\omega)$ at $\omega = 0$ and $\omega = \infty$.
2. The points of intersection of the polar plot with the real and imaginary axes, and the values of ω at these intersections.
3. A few on-locus points in second and third quadrants in the vicinity of the critical point $-1+j0$.

For $G(j\omega)H(j\omega)$ given by Eqn. (8.10), the magnitude and phase at $\omega = 0$ and $\omega = \infty$ are calculated as follows:

$$\left|G(j\omega)H(j\omega)\right|_{\omega=0} = K;\ \angle G(j\omega)H(j\omega)\Big|_{\omega=0} = 0°$$

$$\left|G(j\omega)H(j\omega)\right|_{\omega\to\infty} = \left|\frac{K}{j^2\omega^2\tau_1\tau_2}\right|_{\omega\to\infty} = 0$$

$$\angle G(j\omega)H(j\omega)\Big|_{\omega\to\infty} = \angle\left(\frac{K}{j^2\omega^2\tau_1\tau_2}\right) = -180°$$

The intersections of the polar plot with the axes of $G(s)H(s)$-plane can easily be ascertained by identifying the real and imaginary parts of $G(j\omega)H(j\omega)$.

$$G(j\omega)H(j\omega) = \frac{K}{(1+j\omega\tau_1)(1+j\omega\tau_2)}\left[\frac{(1-j\omega\tau_1)(1-j\omega\tau_2)}{(1-j\omega\tau_1)(1-j\omega\tau_2)}\right]$$

$$= \frac{K\left[(1-\omega^2\tau_1\tau_2) - j\omega(\tau_1+\tau_2)\right]}{(1+\omega^2\tau_1^2)(1+\omega^2\tau_2^2)}$$

$$= \frac{K(1-\omega^2\tau_1\tau_2)}{\omega^4\tau_1^2\tau_2^2+\omega^2(\tau_1^2+\tau_2^2)+1} - j\frac{K\omega(\tau_1+\tau_2)}{\omega^4\tau_1^2\tau_2^2+\omega^2(\tau_1^2+\tau_2^2)+1}$$

$$= \text{Re}[G(j\omega)H(j\omega)] + j\text{Im}[G(j\omega)H(j\omega)]$$

When we let $\text{Im}[G(j\omega)H(j\omega)]$ to zero, we get $\omega = 0$, meaning that the polar plot intersects the real axis only at $\omega = 0$. Similarly, the intersection of $G(j\omega)H(j\omega)$-plot with the imaginary axis is found by setting $\text{Re}[G(j\omega)H(j\omega)]$ to zero, which gives

$$\omega = \frac{1}{\sqrt{\tau_1\tau_2}}$$

$$\left|G(j\omega)H(j\omega)\right|_{\omega=\frac{1}{\sqrt{\tau_1\tau_2}}} = \left.\frac{K\omega(\tau_1+\tau_2)}{\omega^4\tau_1^2\tau_2^2+\omega^2(\tau_1^2+\tau_2^2)+1}\right|_{\omega=\frac{1}{\sqrt{\tau_1\tau_2}}}$$

$$= \frac{K\sqrt{\tau_1\tau_2}}{\tau_1+\tau_2}$$

Based on this information, a rough sketch of the polar plot can easily be made, as shown in Fig. 8.9b (the portion of the locus from $\omega = 0$ to $\omega = +\infty$).

Mapping of Section C_2 onto G(s)H(s)-plane

Given $G(j\omega)H(j\omega)$, $0 \le \omega < \infty$, $G(j\omega)H(j\omega)$ for $-\infty < \omega \le 0$ is constructed by realizing

$$G(-j\omega)H(-j\omega) = [G(j\omega)H(j\omega)]^*$$

where * denotes conjugate. Thus, given the $G(j\omega)H(j\omega)$-locus for $0 \le \omega < \infty$, we get the $G(j\omega)H(j\omega)$-locus for the negative $j\omega$-axis by simply taking the mirror image about the real axis of the locus for positive ω. The dashed portion of the locus in Fig. 8.9b from $\omega = -\infty$ to $\omega = 0$ is thus the map of section C_2 of Nyquist contour.

Mapping of Section C_3 onto G(s)H(s)-plane

Along section C_3, $s = Re^{j\theta}$; $R \to \infty$ and θ varies from + 90° through 0° to − 90°. Substituting into $G(s)H(s)$, we obtain

$$G(s)H(s)\Big|_{s=Re^{j\theta}} = \frac{K}{(\tau_1 Re^{j\theta}+1)(\tau_2 Re^{j\theta}+1)}$$

$$= \frac{K}{\tau_1\tau_2 R^2 e^{j2\theta}} = \frac{K}{\tau_1\tau_2 R^2}e^{-j2\theta}$$

The infinitesimal semicircular locus around the origin (refer Fig. 8.9b), represented by

$$G(s)H(s) = \varepsilon e^{j\alpha}$$

with $\varepsilon \to 0$ and α varying from −180° through 0° to + 180°, is thus the map of section C_3 of Nyquist contour.

For application of the Nyquist stability criterion, exact shape of the Nyquist plot around the origin of $G(s)H(s)$-plane is of no interest to us. We can therefore take $\varepsilon = 0$; the resulting Nyquist plot is shown in Fig. 8.9c. This plot does not encircle the critical point $-1+j0$ for any positive values of K, τ_1, and τ_2. Therefore, the system is stable for all positive values of K, τ_1, and τ_2.

Example 8.2 Consider now a feedback system whose open-loop transfer function is given by

$$G(s)H(s) = \frac{K}{s(\tau s + 1)}; \; K > 0, \; \tau > 0 \tag{8.11}$$

We note here that for this type-1 system, there is one pole on the imaginary axis, precisely at the origin. The Nyquist contour is shown in Fig. 8.10a, where a semicircular detour about the origin is indicated. The Nyquist contour has been divided into four sections: C_1, C_2, C_3 and C_4. Section C_1 is defined by $s = j\omega$, $0 < \omega < \infty$; section C_2 is defined by $s = j\omega$, $-\infty < \omega < 0$; section C_3 is defined by $s = Re^{j\theta}$, $R \to \infty$, and θ varies from + 90° through 0° to – 90°; and section C_4 is defined by $s = \rho e^{j\phi}$, $\rho \to 0$, and ϕ varies from – 90° through 0° to + 90°.

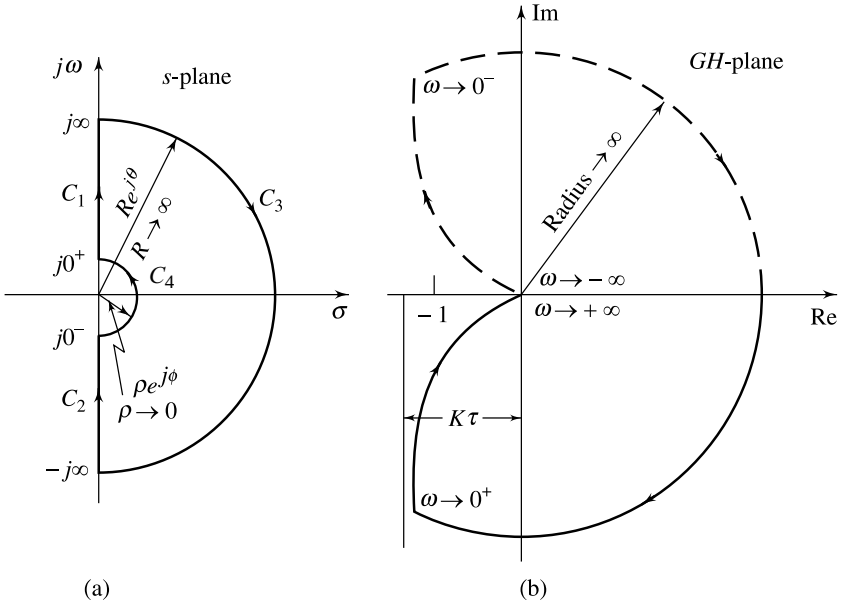


Fig. 8.10 (a) Nyquist contour; (b) Nyquist plot of $G(s)H(s)$ of Eqn. (8.11)

Mapping of section C_1 onto the $G(s)H(s)$-plane is given by the polar plot of sinusoidal transfer function

$$G(j\omega)H(j\omega) = \frac{K}{j\omega(j\omega\tau + 1)} \tag{8.12}$$

This sinusoidal transfer function can be expressed as

$$G(j\omega)H(j\omega) = -\frac{K\tau}{1 + \omega^2\tau^2} - j\frac{K}{\omega(1 + \omega^2\tau^2)}$$

The low-frequency portion of the polar plot becomes

$$\lim_{\omega \to 0} G(j\omega)H(j\omega) = -K\tau - j\infty$$

and the high-frequency portion is (refer Eqn. (8.12))

$$\lim_{\omega \to \infty} G(j\omega)H(j\omega) = \frac{K}{\omega^2\tau} \angle \left(\frac{K}{j^2\omega^2\tau} \right) \Bigg|_{\omega \to \infty} = 0 \angle -180°$$

The general shape of the polar plot is shown in Fig. 8.10b. The $G(j\omega)H(j\omega)$-plot is asymptotic to the vertical line passing through the point $-K\tau + j0$.

Mapping of section C_2 onto $G(s)H(s)$-plane is obtained simply by taking the mirror image about the real axis of the polar plot of $G(j\omega)H(j\omega)$.

Mapping of section C_3 onto the $G(s)H(s)$-plane is obtained as follows.

$$G(s)H(s)\Big|_{s=Re^{j\theta}} = \frac{K}{Re^{j\theta}(\tau\, Re^{j\theta}+1)} = \frac{K}{\tau R^2}e^{-j2\theta}$$

The $G(s)H(s)$-locus thus turns at the origin with zero radius from $-180°$ through $0°$ to $+180°$.

Mapping of section C_4 onto the $G(s)H(s)$-plane is obtained as follows. Along C_4, $s = \rho e^{j\phi}$; $\rho \to 0$, and ϕ varies from $-90°$ through $0°$ to $+90°$.

Substituting into $G(s)H(s)$, we obtain

$$G(s)\,H(s)\Big|_{s=\rho e^{j\phi}} = \frac{K}{\rho e^{j\phi}(\tau\rho e^{j\phi}+1)} = \frac{K}{\rho}e^{-j\phi}$$

The value K/ρ approaches infinity as $\rho \to 0$, and $-\phi$ varies from $+90°$ through $0°$ to $-90°$ as s moves along section C_4 of the Nyquist contour. Thus the infinitesimal semicircular indent around the origin in the s-plane maps into a semicircular arc of infinite radius on the $G(s)H(s)$-plane as shown in Fig. 8.10b.

The complete Nyquist plot for $G(s)H(s)$ given by Eqn. (8.11) is shown in Fig. 8.10b. In order to investigate the stability of this second-order system, we first note that the number of poles of $G(s)H(s)$ in the right-half s-plane is zero. Therefore, for this system to be stable, we require that the Nyquist plot of $G(s)H(s)$ does not encircle the critical point $-1+j0$. Examining Fig. 8.10b, we find that irrespective of the value of the gain K and the time-constant τ, the Nyquist plot does not encircle the critical point, and the system is always stable.

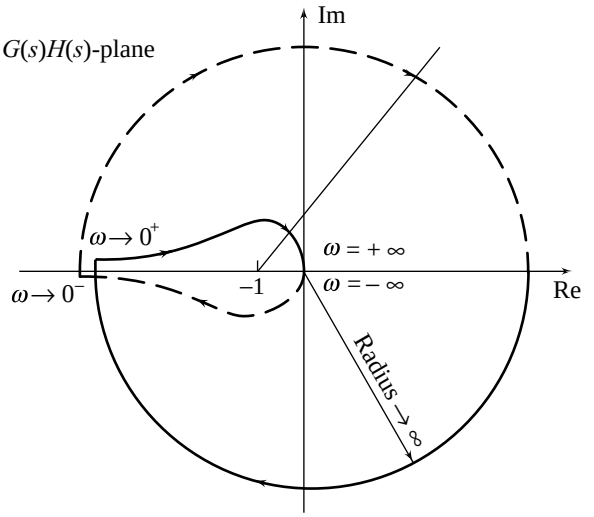


Fig. 8.11 Nyquist plot[2] for $G(s)H(s)$ of Eqn. (8.13)

Example 8.3 We now consider a type-2 system with open-loop transfer function

$$G(s)H(s) = \frac{K}{s^2(\tau s+1)};\ K>0,\ \tau>0 \qquad (8.13)$$

The Nyquist contour is shown in Fig. 8.10a. The Nyquist plot is shown in Fig. 8.11. Clearly the portion of the $G(s)H(s)$-locus from $\omega = 0^+$ to $\omega = +\infty$ is simply the polar plot of

$$G(j\omega)H(j\omega) = \frac{K}{j^2\omega^2(j\omega\tau+1)} \qquad (8.14)$$

[2] For $K = 1$, $\tau = 1$;

ω	0.01	0.05	0.1	0.5	1
$\lvert GH\rvert$	4×10^6	600	98.51	3.57	0.7
$\angle GH$	179.97	177.68	174.2	153.4	135

Note that

$$G(j\omega)H(j\omega)\Big|_{\omega\to 0^+} = \left(\frac{K}{\omega^2}\Big|_{\omega\to 0^+}\right)\angle -180^\circ = \infty \angle \pm 180^\circ$$

$$G(j\omega)H(j\omega)\Big|_{\omega\to +\infty} = \left(\frac{K}{\omega^3\tau}\Big|_{\omega\to +\infty}\right)\angle -270^\circ = 0 \angle -270^\circ \text{ or } +90^\circ$$

It can easily be examined that the locus of $G(j\omega)H(j\omega)$ does not intersect the real or imaginary axis; it lies in the second quadrant for all values of ω.

The $G(s)H(s)$-locus turns at the origin with zero radius from -270° through 0° to $+270^\circ$. This corresponds to the locus from $\omega = +\infty$ to $\omega = -\infty$. The portion of the $G(s)H(s)$-locus from $\omega = -\infty$ to $\omega = 0^-$ is the mirror image about the real axis of the locus for positive ω. From $\omega = 0^-$ to $\omega = 0^+$, $s = \rho e^{j\phi}$, $\rho \to 0$ and ϕ varies from -90° through 0°, to $+90^\circ$. This gives

$$G(s)H(s)\Big|_{s=\rho e^{j\phi}} = \frac{K}{\rho^2 e^{j2\phi}(\tau\rho e^{j\phi}+1)} = \frac{K}{\rho^2}e^{-j2\phi}$$

The portion of the $G(s)H(s)$ locus from $\omega = 0^-$ to $\omega = 0^+$ is a circular arc of infinite radius ranging from $+180^\circ$ at $\omega = 0^-$ through 0° to -180° at $\omega = 0^+$ ($+180^\circ \to +90^\circ \to 0^\circ \to -90^\circ \to -180^\circ$).

From Fig. 8.11, we observe that the Nyquist plot encircles the critical point $-1+j0$ twice in the clockwise direction, i.e., $N = 2$. Since P = number of poles of $G(s)H(s)$ in right-half s-plane = 0, we have from Eqn. (8.8a), $Z = 2$; i.e, there are two roots of the closed-loop system in the right-half s-plane and the system, irrespective of the gain K and time-constant τ, is unstable.

Example 8.4 Consider now an open-loop system with the transfer function

$$G(s)H(s) = \frac{s+1}{s^2(s-2)} \tag{8.15}$$

Let us determine whether the system is stable when the feedback path is closed.

From the transfer function of the open-loop system it is observed that there is one open-loop pole in the right half s-plane. Therefore $P = 1$.

The Nyquist contour is shown in Fig. 8.10a. The Nyquist plot is shown in Fig. 8.12. Clearly, the portion of the $G(s)H(s)$-locus from $\omega = 0^+$ to $\omega = +\infty$ is simply the polar plot of

$$G(j\omega)H(j\omega) = \frac{j\omega+1}{j^2\omega^2(j\omega-2)} \tag{8.16}$$

The following points were used to construct the polar plot:

$$G(j\omega)H(j\omega)\Big|_{\omega\to 0^+} \to \infty\angle 0^\circ$$

$$G(j\omega)H(j\omega)\Big|_{\omega=0.1} = 50\angle 9^\circ$$

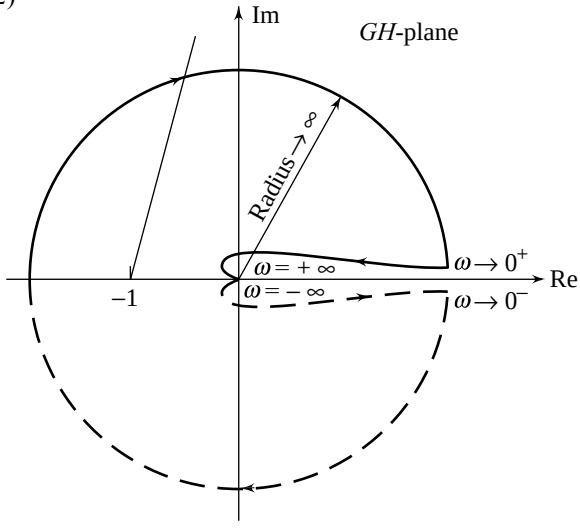


Fig. 8.12 Nyquist plot for $G(s)H(s)$ of Eqn. (8.15)

$$G(j\omega)H(j\omega)\big|_{\omega=1} = 0.6\angle 71°;\ G(j\omega)H(j\omega)\big|_{\omega=10} = 0.01\angle 162°$$

$$G(j\omega)H(j\omega)\big|_{\omega\to+\infty} \to 0\angle -180°$$

The plot of $G(j\omega)H(j\omega)$ for $\omega < 0$ is the reflection, with respect to the real axis, of the plot for $\omega > 0$. Every point on the large semicircle in Fig. 8.10a, is mapped by $G(s)H(s)$ into the origin of the $G(s)H(s)$-plane. A point $s = \rho e^{j\phi}$; $\rho \to 0$, ϕ varying from $-90°$ through $0°$ to $+90°$ is mapped by $G(s)H(s)$ into

$$G(s)H(s)\big|_{s=\rho e^{j\phi}} = \frac{s+1}{s^2(s-2)}\bigg|_{s=\rho e^{j\phi}} = -\frac{1}{2\rho^2 e^{j2\phi}} = \frac{1}{2\rho^2}\angle(180° - 2\phi)$$

The small semicircle in Fig. 8.10a is mapped by $G(s)\,H(s)$ into a circular arc of infinite radius ranging from $360°$ at $\omega = 0^-$ through $180°$ to $0°$ at $\omega = 0^+$ ($360° \to 270° \to 180° \to 90° \to 0°$).

Figure 8.12 indicates that the critical point $-1+j0$ is encircled by the Nyquist plot once in the clockwise direction. Therefore $N = 1$, and (refer Eqn. (8.8a))

Z = number of zeros of $1+G(s)H(s)$ enclosed by the Nyquist contour

$= N + P = 2$

Hence the feedback system is unstable with two poles in the right half s-plane.

Example 8.5 Consider a feedback system with the following open-loop transfer function:

$$G(s)H(s) = \frac{K}{s(s+3)(s+5)} \tag{8.17}$$

Let us investigate the stability of this system for various values of K.

First set $K = 1$ and sketch the Nyquist plot for the system using the contour shown in Fig. 8.10a. For all points on the imaginary axis,

$$G(j\omega)H(j\omega) = \frac{K}{s(s+3)(s+5)}\bigg|_{\substack{K=1\\ s=j\omega}} = \frac{-8\omega^2 - j(15\omega - \omega^3)}{64\omega^4 + \omega^2(15-\omega^2)^2} \tag{8.18}$$

At $\omega = 0$, $G(j\omega)H(j\omega) = -0.0356 - j\infty$.

Next, find the point where the Nyquist plot intersects the negative real axis. Setting the imaginary part of Eqn. (8.18) equal to zero, we find $\omega = \sqrt{15}$. Substituting this value of ω back into Eqn. (8.18), yields the real part of -0.0083.

Finally, at $\omega = \infty$, $G(j\omega)H(j\omega) = G(s)H(s)\big|_{s\to j\infty} = \dfrac{1}{(j\omega)^3} = 0\angle -270°$.

The Nyquist plot for $G(s)H(s)$ of Eqn. (8.17) with $K = 1$ is shown in Fig. 8.13. Application of the Nyquist criterion shows that with $K = 1$, the closed-loop system is stable.

If we were to increase the gain by a factor $1/0.0083 = 120.48$, all points on the $G(j\omega)H(j\omega)$-locus would increase in magnitude by this factor along radial lines centred at the origin. The length of the vector $G(j\sqrt{15})H(j\sqrt{15})$ would be $0.0083\ (1/0.0083) = 1$, and therefore the curve $G(j\omega)H(j\omega)$ would go through the critical point; the closed-loop system would be at the limit of instability. Hence, for stability, K must be less than 120.48, i.e., the stability range is $0 < K < 120.48$.

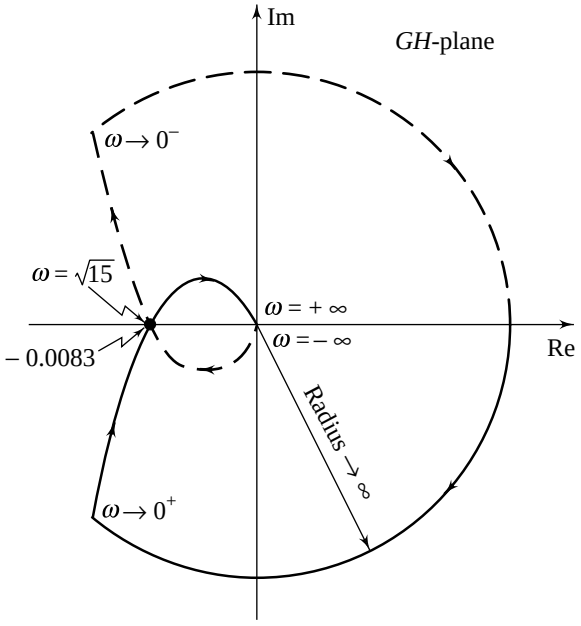


Fig. 8.13 Nyquist plot for $G(s)H(s)$ of Eqn. (8.17)

Example 8.6 Let us consider the control system shown in Fig. 8.14 and determine its stability. The characteristic equation of the system is

$$s(s-1) + K_A(1 + K_t s) = 0$$

We rearrange it in the form convenient for Nyquist plotting:

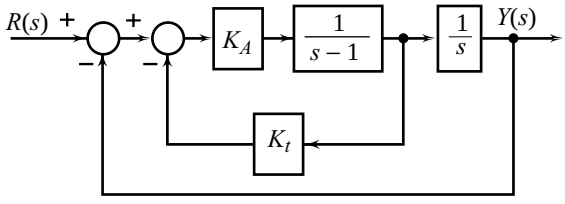


Fig. 8.14 A feedback control system

$$1 + \frac{K_A(1 + K_t s)}{s(s-1)} = 0$$

or

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

$$G(s) = \frac{K_A(1 + K_t s)}{s(s-1)} \tag{8.19}$$

The open-loop transfer function has one pole in the right half s-plane, and therefore $P = 1$. In order for the closed-loop system to be stable, we require $N = -1$ (i.e., the Nyquist plot must encircle the critical point once in the counterclockwise direction).

Several important values of $G(j\omega)$-locus are given below:

$$G(j\omega)\Big|_{\omega \to 0^+} = \infty \angle +90°;\ G(j\omega)\Big|_{\omega \to +\infty} = 0\angle -90°$$

$$G(j\omega) = \frac{K_A(1 + K_t j\omega)}{-\omega^2 - j\omega}$$

$$= \frac{-K_A(\omega^2 + \omega^2 K_t) + j(\omega - K_t\omega^3)K_A}{\omega^2 + \omega^4} \quad (8.20)$$

$$G(j\omega)\Big|_{\omega^2 = 1/K_t} = -K_A K_t$$

At the semicircular detour at the origin in the s-plane (refer Fig. 8.10a), we let $s = \rho e^{j\phi}$; $\rho \to 0$ and ϕ varies from – 90° through 0° to + 90°.

$$G(s)\Big|_{s=\rho e^{j\phi}} = \frac{K_A}{-\rho e^{j\phi}} = \frac{K_A}{\rho}\angle(-180° - \phi)$$

Therefore the semicircular detour in the Nyquist contour is mapped into a semicircle of infinite radius in the left half of $G(s)$-plane, as shown in Fig. 8.15. When $-K_A K_t < -1$ or $K_A K_t > 1$, the Nyquist plot encircles the $-1 + j0$ point once in counterclockwise direction. Thus the system is stable when $K_A K_t > 1$.

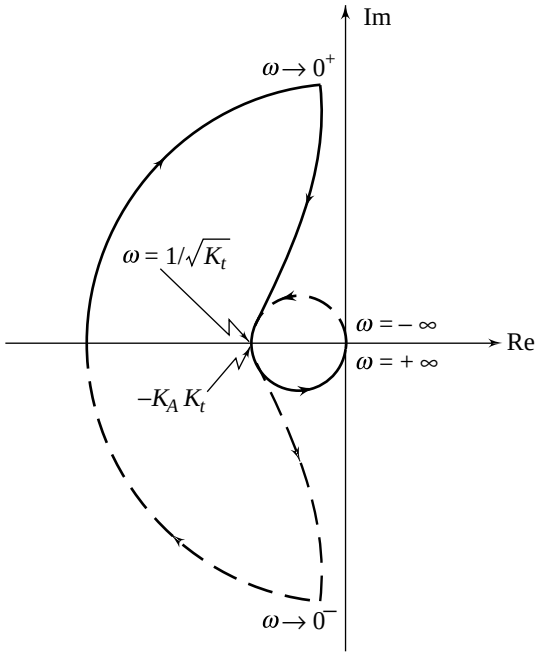


Fig. 8.15 Nyquist plot[3] for G(s) of Eqn. (8.19)

8.4 STABILITY MARGINS

In the preceding sections, the Nyquist plot was used to detemine if a system was stable or unstable. Of course, in general, a usable system must be stable. However, there are concerns beyond simple stability, for two reasons. First, a stable system must also have, among other characteristics, an acceptable transient response. Also, the model that is used in the analysis and design of a control system is never exact. Hence, the model may indicate that a system is stable, whereas in fact the physical system is unstable. Generally we require not only that a system be stable but also that it be stable by some margin of safety.

Measures of degree of stability of a closed-loop system that has open-loop transfer function with no poles in right half s-plane, can be conveniently created through Nyquist plot. The stability information for such a system becomes obvious by the inspection of the Nyquist plot of its open-loop transfer function $G(s)H(s)$, since the stability criterion is merely the non-encirclement of the critical point $-1 + j0$. It can be intuitively imagined that as the Nyquist plot gets closer to the critical point, the system tends towards instability.

Consider two different systems whose dominant closed-loop poles are shown on the s-plane in Figs 8.16a and 8.16b. Obviously, system A is 'more stable' than system B, since its dominant closed-loop poles are located comparatively away to the left from the $j\omega$-axis. The Nyquist plots of open-loop transfer functions of systems A and B are shown in Figs 8.16c and 8.16d, respectively. Comparison of the closed-loop pole locations of these systems with their corresponding Nyquist plots reveals that as a Nyquist plot moves closer

[3] for $K_A = 1$, $K_t = 3$;

ω	0.02	0.05	0.09	1
$\lvert G\rvert$	146.3	74	37.16	2.23
$\angle G$	91.57	93.1	96.14	153.56

to the critical point $-1+j0$, the system closed-loop poles move closer to the $j\omega$-axis and hence the system becomes relatively less stable and *vice versa*.

Thus, the distance between the Nyquist plot of open-loop transfer function $G(s)H(s)$ and the critical point $-1+j0$ can be used as a measure of degree of stability of the closed-loop system. Generally, the larger the distance, the more stable the system. The distance can be found by drawing a circle touching the Nyquist plot as shown in Figs 8.16c and 8.16d. Such a distance, however, is not convenient for analysis and design. The measures gain margin and phase margin, are commonly used to quantify the closeness of a Nyquist plot to the critical point.

8.4.1 Gain Margin

Figure 8.17 shows a 'typical' Nyquist plot of $G(j\omega)H(j\omega)$; $\omega \geq 0$. $G(s)H(s)$ is known to have no poles in the right half s-plane. A closed-loop system with the $G(s)H(s)$ as open-loop transfer function, is stable.

The point at which the Nyquist plot crosses the negative real axis is marked A, and the frequency at that point is designated ω_ϕ. The point A is thus the tip of the vector $G(j\omega_\phi)H(j\omega_\phi)$ with magnitude $|G(j\omega_\phi)H(j\omega_\phi)|$ and angle 180°. It is seen from Fig. 8.17 that $|G(j\omega_\phi)H(j\omega_\phi)| < 1$, and therefore the tip of the vector $G(j\omega_\phi)H(j\omega_\phi)$ does not reach the critical point $-1+j0$.

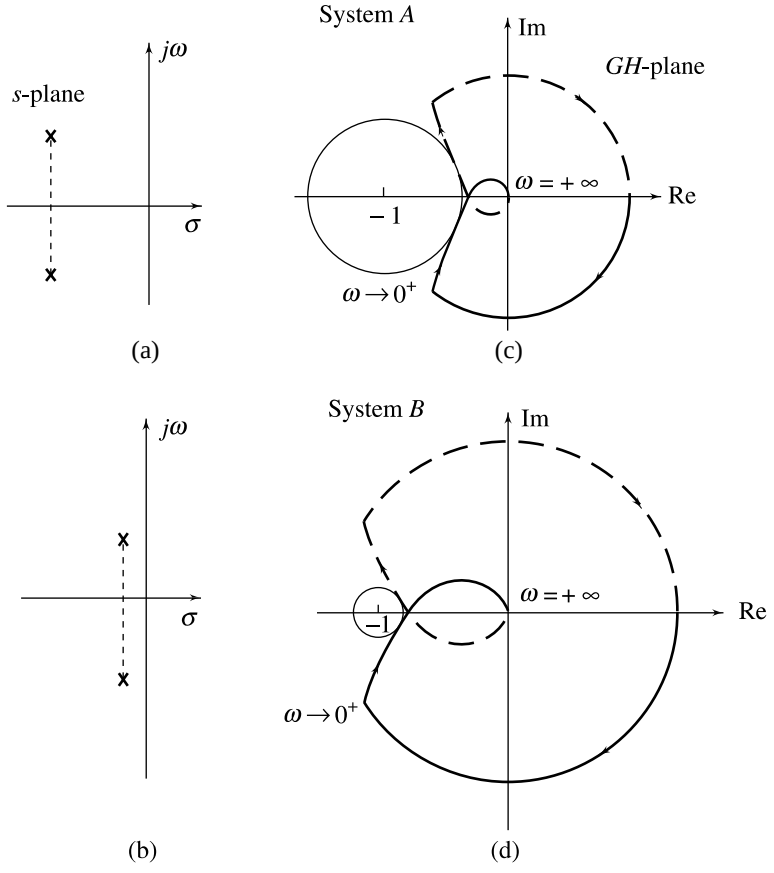


Fig. 8.16 Correlation between the closed-loop s-plane root locations and open-loop frequency-response curves

If we were to increase the gain of $G(j\omega)H(j\omega)$ by a factor $1/|OA|$ without altering its phase, all points on the Nyquist plot of $G(j\omega)H(j\omega)$ would increase in magnitude by this factor along radial lines centred at the origin. The length of the vector $G(j\omega_\phi)H(j\omega_\phi)$ would be $|OA|(1/|OA|) = 1$, and therefore the curve $G(j\omega)H(j\omega)$ would go through the critical point; the closed-loop system would be at the limit of stability. We call the multiplying factor $1/|OA|$ the *gain margin* because it is the factor by which the gain can be increased to drive the system to the verge of instability.

From the graphical description of Fig. 8.17, we now make the following definitions.

Phase crossover point It is a point on the $G(s)H(s)$-plane at which the Nyquist $G(j\omega)H(j\omega)$-plot intersects the negative real axis.

Phase crossover frequency ω_ϕ It is the frequency at the phase crossover point, or where

$$\angle G(j\omega_\phi)H(j\omega_\phi) = -180° \tag{8.21}$$

Gain margin The gain margin GM of the closed-loop system that has $G(s)H(s)$ as its open-loop transfer function is defined as the number

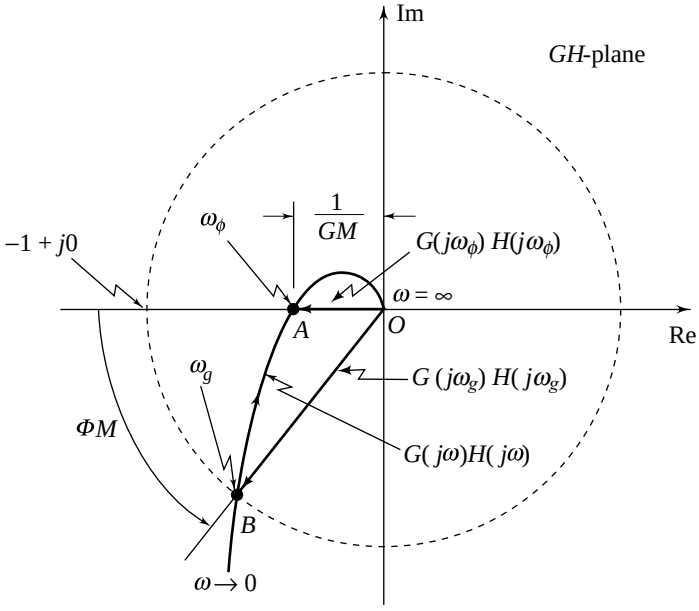


Fig. 8.17 Typical Nyquist plot of $G(j\omega)H(j\omega)$

$$GM = \frac{1}{|G(j\omega_\phi)H(j\omega_\phi)|} \tag{8.22}$$

For example, if $|G(j\omega_\phi)H(j\omega_\phi)| = 0.5$, then the $GM = 2$; the loop gain magnitude could be increased by a factor of 2 before the closed-loop system would go unstable. For stable systems, GM is always a number greater than one.

8.4.2 Phase Margin

A unit circle centred at the origin has been drawn in Fig. 8.17 in order to identify point B at which Nyquist $G(j\omega)H(j\omega)$-plot has unity magnitude; the frequency at point B has been designated ω_g. The point B is thus

the tip of the vector $G(j\omega_g)H(j\omega_g)$ with magnitude $|G(j\omega_g)H(j\omega_g)| = 1$ and angle $\angle G(j\omega_g)H(j\omega_g)$. It is seen from Fig. 8.17 that the tip of the vector $G(j\omega_g)H(j\omega_g)$ does not reach the critical point $-1+j0$.

If we were to decrease the phase of $G(j\omega)H(j\omega)$ without altering its gain, all points on the Nyquist $G(j\omega)H(j\omega)$-plot would rotate clockwise about the origin an angular amount equal to the increase in phase lag. In the light of Fig. 8.17, we see that the tip of the vector $G(j\omega_g)H(j\omega_g)$ would be placed at the critical point if the increase in phase lag of $G(j\omega)H(j\omega)$ were made equal to the angle ΦM. We call the angle ΦM, *measured positively in the counterclockwise direction from the negative real axis to the vector* $G(j\omega_g)H(j\omega_g)$, *the phase margin* because it is the angle by which the phase of $G(j\omega)H(j\omega)$ can be decreased to drive the system to the verge of instability.

From the graphical description of Fig. 8.17, we now make the following definitions.

Gain crossover point It is a point on the $G(s)H(s)$-plane at which the Nyquist $G(j\omega)H(j\omega)$-plot has unity magnitude.

Gain crossover frequency ω_g It is the frequency at the gain crossover point, or where

$$|G(j\omega_g)H(j\omega_g)| = 1 \tag{8.23}$$

Phase margin The phase margin ΦM of the closed-loop system that has $G(s)H(s)$ as its open-loop transfer function, is defined as

$$\Phi M = 180° + \angle G(j\omega_g)H(j\omega_g) \tag{8.24}$$

For example, if $\angle G(j\omega_g)H(j\omega_g) = -135°$, then $\Phi M = 45°$; an additional phase lag of 45° could be associated with $G(j\omega)H(j\omega)$ before the closed-loop system would go unstable. For stable systems, ΦM is always positive.

Consider now an unstable closed-loop system with Nyquist $G(j\omega)H(j\omega)$-plot of Fig. 8.18 ($G(s)H(s)$ is known to have no poles in the right half s-plane). Applying the definition of gain margin to this plot we find that $G(j\omega_\phi)H(j\omega_\phi)$ is a number greater than one and therefore GM is a number less than one. Applying the definition of phase margin, we observe that the vector $G(j\omega_g)H(j\omega_g)$ lies in the second quadrant. The angle measured from negative real axis to this vector is negative (counterclockwise measurement is taken as positive by definition) and therefore the phase margin ΦM is negative.

In general, the gain margin and the phase margin are mutually independent. For example, it is possible for $G(j\omega)H(j\omega)$ to reflect an excellent gain margin but a poor phase margin. Such is the case depicted in Fig. 8.19, where the $GM = \infty$ but the $\Phi M < 15°$. In such a case, the phase crossover frequency ω_ϕ is undefined. Conversely, $G(j\omega)H(j\omega)$ can reflect a poor gain margin but an excellent phase margin, as is the case depicted in Fig. 8.20, where the $\Phi M = \infty$ but the GM is only slightly greater than 1. In such a case, the gain crossover frequency ω_g is undefined.

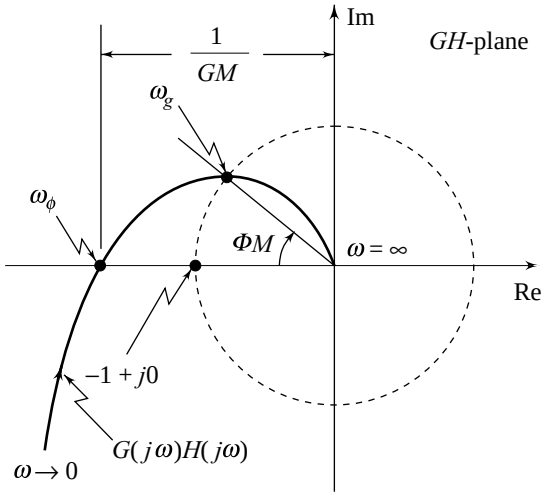


Fig. 8.18 Nyquist plot of an unstable closed-loop system

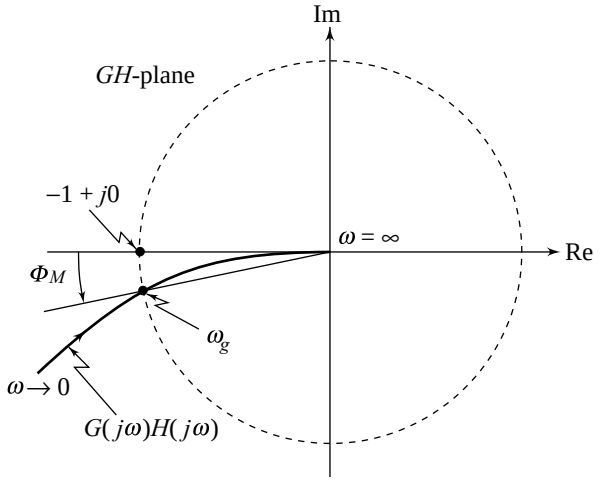


Fig. 8.19 Infinite GM, poor ΦM

In the light of these observations, gain and phase margins should be used in conjunction with one another when determining the degree of stability of a closed-loop system. Together, they represent a measure of the distance of $G(j\omega)H(j\omega)$ from the critical point $-1+j0$.

8.4.3 Some Constraints and Cautions

Gain margin and phase margin are valid measures of relative stability of a closed-loop system only if its open-loop transfer function $G(s)H(s)$ has no poles in the right half s-plane. If the open-loop transfer function has poles in the right half s-plane, it is safer to examine the complete Nyquist plot and/or root locus for interpreting relative stability of the closed-loop system.

It should be observed that the sketches of Figs 8.17–8.20 each cross the negative real axis or the unit circle atmost once. For systems with numerous zeros and poles, the Nyquist plot may cross the negative real axis several times, or it may cross the unit circle more than once. For such curves, several values of gain margin and phase margin are defined, as shown in Figs 8.21 (a) and (b). Caution is suggested when interpreting the stability of such systems. It is safer not to depend on the gain margin and phase margin for relative stability analysis.

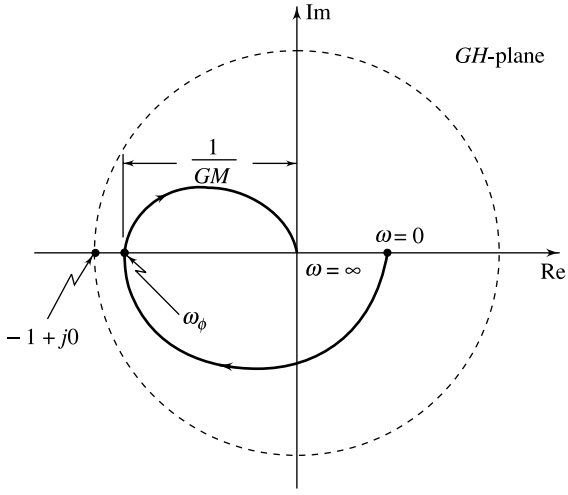


Fig. 8.20 Poor GM, infinite ΦM

It should also be observed that the sketches of Figs 8.12–8.20 correspond to systems where increasing gain leads to instability. There are certain practical situations wherein increase in gain can make the system stable. Consider, for example, unity-feedback system with open-loop transfer function

$$G(s) = \frac{K(s+10)^2}{s^3}$$

This is a system for which increasing gain causes a transition from instability to stability (The Routh criterion shows that the closed-loop system is unstable for $K < 5$ and stable for $K > 5$). The Nyquist plot in Fig. 8.21c has been drawn for the stable value $K = 7$. From this figure we observe that the phase margin is positive and the gain margin is less than one. According to the rules for stability discussed earlier, these two margins yield conflicting signals on system stability. Caution is suggested when interpreting the stability of such systems. It is safer not to depend on the gain margin and phase margin for relative stability analysis.

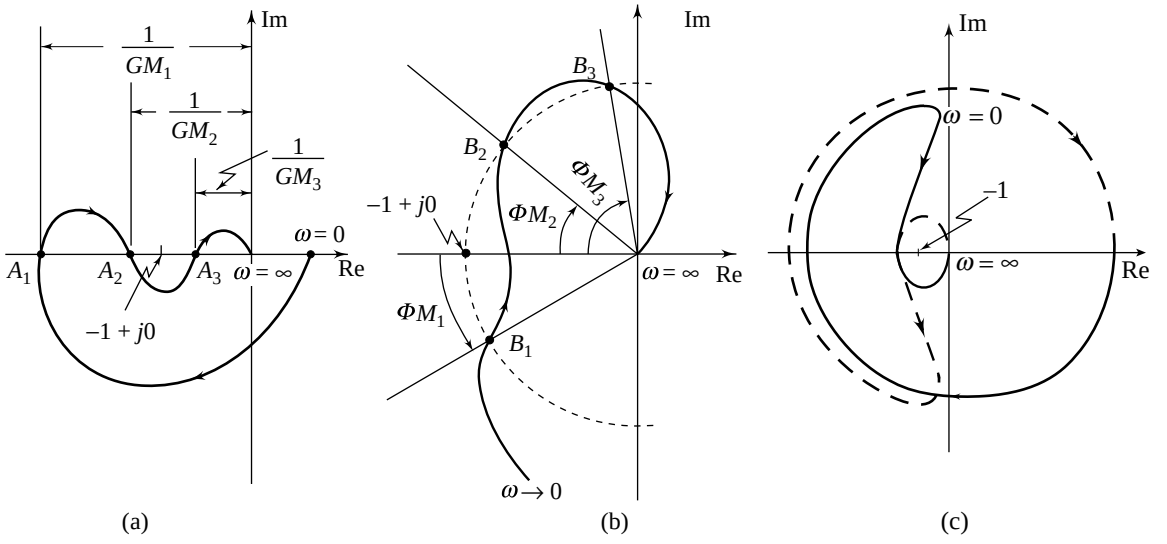


Fig. 8.21 Nyquist plots requiring careful interpretation of GM and ΦM

Example 8.7 A unity-feedback system has open-loop transfer function

$$G(s) = \frac{6}{(s^2 + 2s + 2)(s + 2)} \tag{8.25}$$

Since the open-loop poles are only in the left half s-plane, the Nyquist criterion tells us that we want no encirclement of $-1+j0$ point for stability. In such cases, closed-loop stability can be evaluated from Nyquist $G(j\omega)$-plot.

$$G(j\omega)|_{\omega=0} = 1.5\angle 0°;\ G(j\omega)|_{\omega\to\infty} = 0\angle 90°$$

$$G(j\omega) = \left.\frac{6}{(s^2 + 2s + 2)(s + 2)}\right|_{s=j\omega}$$

$$= \frac{6[4(1-\omega^2) - j\omega(6-\omega^2)]}{16(1-\omega^2)^2 + \omega^2(6-\omega^2)^2} \tag{8.26}$$

Setting the imaginary part to zero, we find the phase crossover frequency $\omega_\phi = \sqrt{6}$ rad/sec. At this frequency, the real part of $G(j\omega)$ in Eqn. (8.26) is calculated to be -0.3 (refer Fig. 8.22). Thus the gain can be increased by (1/0.3) before the real part becomes -1. The gain margin is $GM = 3.33$.

To find the phase margin, first find the gain crossover frequency ω_g: the frequency for which the magnitude of $G(j\omega)$ in Eqn. (8.26) is unity. As the problem stands, this calculation requires computational tools. Later in this chapter, we will simplify the calculation process by using Bode plots.

The gain crossover frequency $\omega_g = 1.253$ rad/sec. At this frequency, the phase angle of $G(j\omega)$ is $-112.33°$. The difference between this angle and 180° is 67.67°, which is the phase margin ΦM.

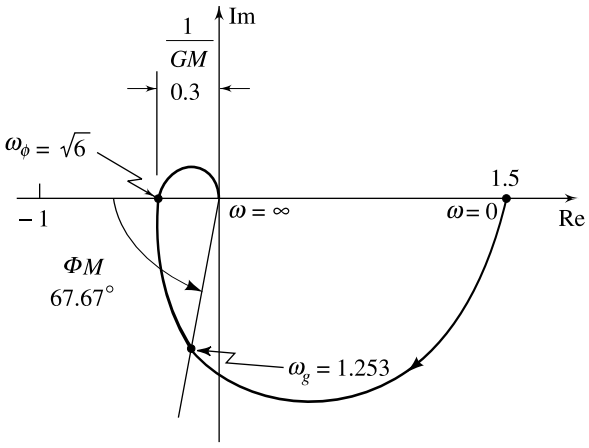


Fig. 8.22 Nyquist plot for $G(j\omega)$ of Eqn. (8.26)

8.5 THE BODE PLOTS

A sinusoidal transfer function may be represented by two separate plots; one giving the magnitude *versus* frequency, and the other the phase angle *versus* frequency. A *Bode plot* (named after Hendrick W. Bode) consists of two graphs: one is a plot of the logarithm of the magnitude of the sinusoidal transfer function, and the other is a plot of the phase angle in degrees; both are plotted against the frequency in logarithmic scale.

In a Bode plot, the logarithmic magnitude of sinusoidal transfer function $G(j\omega)$ is represented[4] as $20 \log |G(j\omega)|$, where the base of the logarithm is 10. The unit used in this representation of the magnitude is *decibel* (one-tenth of a 'Bel': a unit named after Alexander Graham Bell), usually abbreviated as dB.

The main advantage of using the logarithmic plot is that the multiplication of magnitudes can be converted into addition. Consider a typical rational transfer function $G(s)$ factored in the time-constant form:

$$G(s) = \frac{K(1+s\tau_a)(1+s\tau_b)\cdots}{s^N(1+s\tau_1)(1+s\tau_2)\cdots[1+(2\zeta/\omega_n)s+s^2/\omega_n^2]\cdots} \tag{8.27}$$

This transfer function has zeros at $s = -1/\tau_a, -1/\tau_b, \ldots$, N poles at the origin, real poles at $s = -1/\tau_1$, $-1/\tau_2, \ldots$, and complex poles at $-\zeta\omega_n \pm j\omega_n\sqrt{1-\zeta^2}$. We note that the constant multiplier $K = \lim_{s\to 0} s^N G(s)$, where N is the type number for the system with open-loop transfer function $G(s)$ given by Eqn. (8.27). Thus for a type-0 system, $K_p = K$. For a type-1 system, $K_v = K$, and for a type-2 system, $K_a = K$. Any rational $G(s)$ can be factored in the form of Eqn. (8.27). If $G(s)$ has complex zeros, then even though it is not shown in Eqn. (8.27), a quadratic term of the form $[1 + (2\zeta/\omega_n)s + s^2/\omega^2_n]$ will also appear in the numerator.

The magnitude of $G(j\omega)$ in dB is obtained by multiplying the logarithm to the base 10 of $|G(j\omega)|$ by 20.

[4] There is little reason for this choice; the choice dates from Alexander Graham Bell's study of the response of the human ear. In electromagnetics and some other applications, the natural logarithm is used; the unit then is *neper*.

$$
\begin{aligned}
|G(j\omega)|_{\text{dB}} &= 20 \log |G(j\omega)| \\
&= 20 \log K + 20 \log |1 + j\omega\tau_a| + 20 \log |1 + j\omega\tau_b| + \cdots \\
&\quad - 20N \log |j\omega| - 20 \log |1 + j\omega\tau_1| - 20 \log |1 + j\omega\tau_2| - \cdots \\
&\quad - 20 \log |1 + j2\zeta\omega/\omega_n - \omega^2/\omega_n{}^2| - \cdots \qquad (8.28a)
\end{aligned}
$$

The phase of $G(j\omega)$ is

$$
\begin{aligned}
\angle G(j\omega) &= \tan^{-1}\omega\tau_a + \tan^{-1}\omega\tau_b + \cdots - N(90°) - \tan^{-1}\omega\tau_1 - \tan^{-1}\omega\tau_2 - \cdots \\
&\quad - \tan^{-1} 2\zeta\omega\omega_n/(\omega_n{}^2 - \omega^2) - \cdots \qquad (8.28b)
\end{aligned}
$$

Equations (8.28) express the magnitude and the phase of $G(j\omega)$ as linear combinations of relatively simple terms. The curves for each term can be added together graphically to get the curves for the complete transfer function.

The logarithmic scale used for frequency in Bode plots, has some interesting properties. The frequency in a typical control system application varies over many powers of ten so that most information would be compressed near the origin if a linear scale were used. The logarithmic scale is nonlinear, that is, the distance between 1 and 2 is greater than the distance between 2 and 3 and so on. As a result, use of this scale enables us to cover a greater range of frequencies; both low and high frequency behaviour of a system can be adequately displayed in one plot. The other interesting property of the use of log scale for frequency, as we shall see shortly, is that the straight-line approximation of the Bode plots becomes possible. This property allows the plotting by hand that is quick and yet sufficiently accurate for control-system design. Most control-system designers will have access to computer programs that will diminish the need for hand plotting; however, it is still important to develop good intuition so that erroneous computer results are quickly identified, and for this one needs the ability to check results by hand plotting.

The phase angle graph of the Bode plot, $\angle G(j\omega)$ *versus* $\log\omega$, has a linear scale for phase in degrees and a logarithmic scale for frequency ω. The magnitude graph[5] of the Bode plot, $|G(j\omega)|_{\text{dB}}$ versus $\log\omega$, has a linear scale for magnitude in dB and a logarithmic scale for frequency ω.

In general, a rational transfer function $G(s)$ can contain just four simple types of factors:

1. Real constant: $K > 0$
2. Poles or zeros at the origin of order N: $(s)^{\mp N}$
3. Poles or zeros at $s = -1/\tau$ of order q: $(1 + s\tau)^{\mp q}$
4. Complex poles or zeros of order r: $(1 + 2\zeta s/\omega_n + s^2/\omega_n{}^2)^{\mp r}$

Equations (8.28) indicate one of the unique characteristics of Bode plot—each of the four types of factors listed can be considered as a separate plot; the individual plots are then algebraically added to yield the total plot of the given transfer function. Bode plots of the four types of factors are, therefore, the basic building blocks for the construction of the Bode plot of the given transfer function.

8.5.1 Magnitude Plot: Straight-Line Approximation

We shall now discuss how the $|G(j\omega)|_{\text{dB}}$ *versus* $\log\omega$ graph can be sketched with very little effort using straight-line approximation. We first examine the basic building blocks, and then use these blocks to build the magnitude plot of the given transfer function.

[5]An alternative presentation of the magnitude graph uses logarithm scale for frequency ω and logarithmic scale for magnitude $|G(j\omega)|$, resulting in $\log |G(j\omega)|$ *versus* $\log\omega$ locus.

The dB *versus* log ω graph can be sketched on a linear rectangular coordinate graph paper. However, use of semilog graph paper is more convenient as it eliminates the need to take logarithm of many numbers. The logarithmic scale of the semilog graph paper is used for frequency and the linear scale is used for dB.

The units used to express frequency bands or frequency ratios are the *octave* and the *decade*. An octave is a frequency band from ω_1 to ω_2 where $\omega_2/\omega_1 = 2$. There is an increase of one decade from ω_1 to ω_2 when $\omega_2/\omega_1 = 10$. Semilog graph papers come in two, three, four or five cycles, indicating the range of coverage in decades (refer Fig. 8.23). It may be noted that we cannot locate the point $\omega = 0$ on the log scale since log $0 = -\infty$.

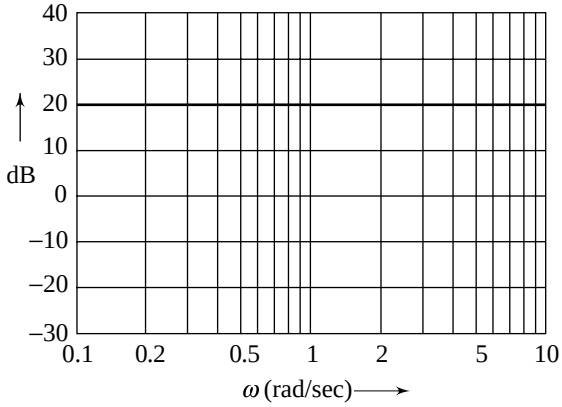


Fig. 8.23 Bode magnitude plot of *K*

Real constant K Since the constant K is frequency-invariant, the plot of

$$\text{dB} = 20 \log K \tag{8.29}$$

is a horizontal straight line. The magnitude plot for $K = 10$ is shown in Fig. 8.23 in semilog coordinates. The horizontal axis is log ω; labeled ω on logarithmic scale. The vertical axis is magnitude in decibels; labeled dB on linear scale.

Poles or zeros at the origin The factor $1/(j\omega)$ appearing in a transfer function $G(j\omega)$ has the magnitude

$$\text{dB} = 20 \log \left| \frac{1}{j\omega} \right| = -20 \log \omega \tag{8.30}$$

The dB versus log ω plot of Eqn. (8.30) is a straight line with a slope of -20 dB per unit change in log ω. A unit change in log ω means

$$\log(\omega_2/\omega_1) = 1 \text{ or } \omega_2 = 10\omega_1$$

This range of frequencies is called a *decade*. Thus the slope of Eqn. (8.30) is -20 dB/decade. The range of frequencies $\omega_2 = 2\omega_1$ is called an *octave*. Since $-20 \log 2 = -6$ dB, the slope of Eqn. (8.30) could also be expressed as -6 dB/octave. The plot of Eqn. (8.30) is shown in Fig. 8.24a; it intersects the 0-dB axis at $\omega = 1$.

For an Nth-order pole at the origin, the magnitude is

$$\text{dB} = 20 \log \left| \frac{1}{(j\omega)^N} \right| = -20N \log \omega \tag{8.31}$$

The magnitude plot is still a straight line that intersects the 0-dB axis at $\omega = 1$, but the slope is now $-20N$ dB/decade.

For the case that a transfer function has a zero at the origin, the magnitude of the term is given by

$$\text{dB} = 20 \log |j\omega| = 20 \log \omega \tag{8.32}$$

and the plot is the negative of that for a pole at the origin; it is a straight line with a slope of + 20 dB/decade that intersects the 0-dB axis at $\omega = 1$ (Fig. 8.24b). For the case of an Nth-order zero at the origin, it is seen that the plot is the negative of that for the Nth-order pole.

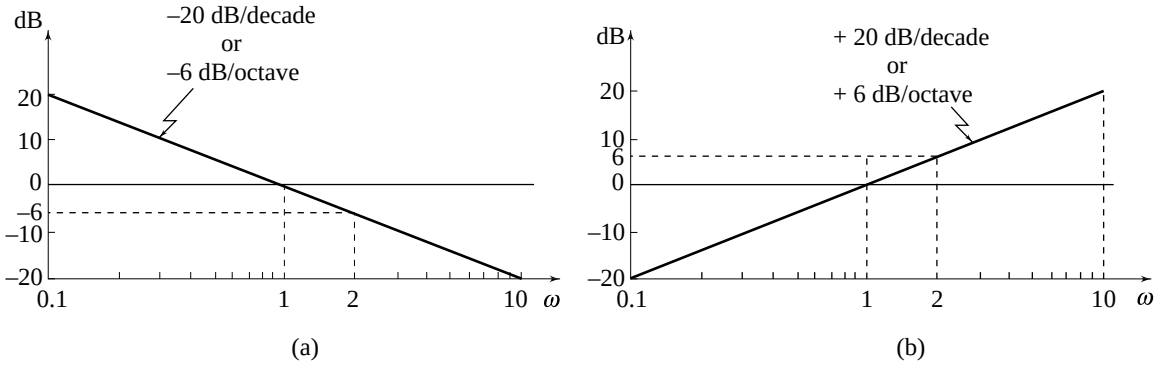


Fig. 8.24 Bode magnitude plot of (a) pole (b) zero at the origin

Example 8.8 Consider the transfer function

$$G(s) = \frac{K}{s^N}$$

There are two terms in $G(s)$: the constant K, and the repeated pole at the origin. We could add the Bode magnitude plots of the two terms to get the Bode magnitude plot of $G(s)$. However, treating K/s^N as a single term turns out to be more convenient, as is seen below.

$$\text{dB} = 20 \log \left| \frac{K}{(j\omega)^N} \right| = -20N \log \omega + 20 \log K \tag{8.33}$$

With $\log \omega$ as abscissa, the plot of Eqn. (8.33) is a straight line having a slope $-20N$ dB/decade and passing through $20 \log K$ dB when $\log \omega = 0$, i.e., when $\omega = 1$, as shown in Fig. 8.25. Further, the plot has a value of 0 dB at the frequency of

$$20N \log \omega = 20 \log K \quad \text{or} \quad \omega = (K)^{1/N}$$

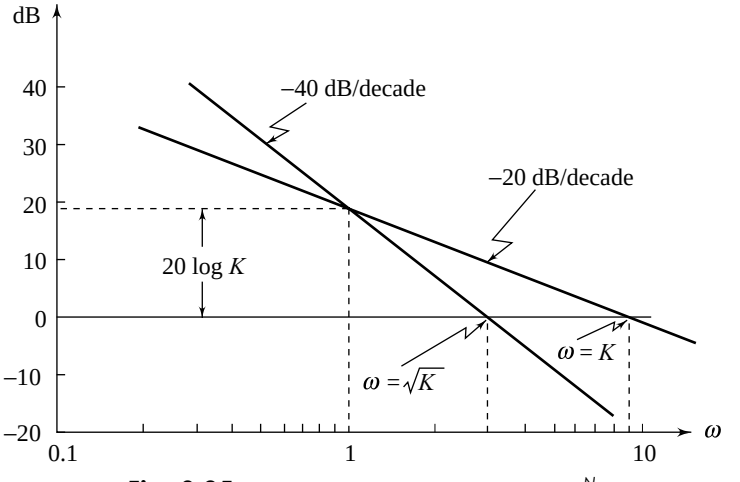


Fig. 8.25 Bode magnitude plot of K/s^N

Poles or zeros at s = –1/τ The factor $1/(1 + j\omega\tau)$ appearing in a transfer function $G(j\omega)$ has magnitude

$$\text{dB} = 20 \log \left| \frac{1}{1 + j\omega\tau} \right| = -20 \log \sqrt{1 + \omega^2\tau^2} \tag{8.34}$$

For low frequencies, such that $\omega << 1/\tau$, the magnitude may be approximated by

$$\text{dB} = -20 \log 1 = 0 \tag{8.35}$$

For high frequencies, such that $\omega >> 1/\tau$, the magnitude may be approximated by

$$\text{dB} = -20 \log(\omega\tau) = -20 \log \omega - 20 \log \tau \tag{8.36}$$

The plot of Eqn. (8.35) is a straight line coincident with the 0-dB axis. The plot of Eqn. (8.36) is also a straight line having a slope of –20 dB/decade intersecting the 0-dB axis at $\omega = 1/\tau$, as shown in Fig. 8.26. We call the straight line approximations the *asymptotes*. The low-frequency approximation is called the *low-frequency asymptote* and the high-frequency approximation is called the *high-frequency asymptote*. The frequency $\omega = 1/\tau$ at which the two asymptotes meet is called the *corner frequency* or *break frequency*.

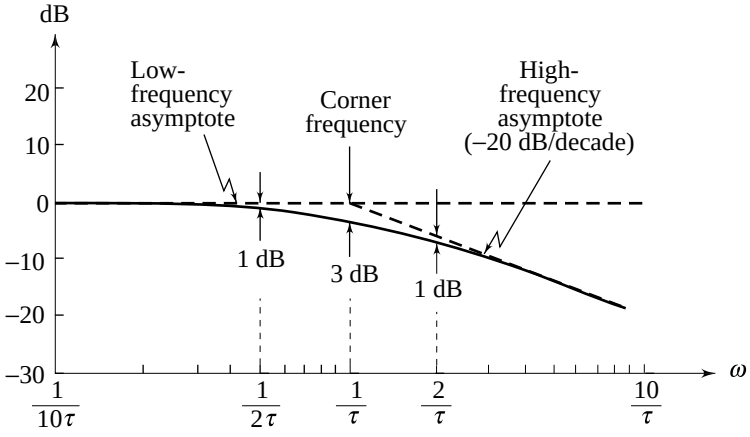


Fig. 8.26 Bode magnitude plot of $1/(1 + j\omega\tau)$

Though the asymptotic approximations hold good for $\omega << 1/\tau$ and $\omega >> 1/\tau$, with some loss of accuracy these could be extended for $\omega \leq 1/\tau$ and $\omega \geq 1/\tau$. Therefore, the dB *versus* log ω curve of $1/(1 + j\omega\tau)$ can be approximated by two asymptotes, one a straight line at 0 dB for the frequency range $0 < \omega \leq 1/\tau$ and the other, a straight line with a slope –20 dB/decade (or – 6 dB/octave) for the frequency range $1/\tau \leq \omega < \infty$. The corner frequency divides the plot into two regions, a low-frequency region and a high-frequency region.

The error in the magnitude plot caused by the use of asymptotes can easily be calculated. The maximum error occurs at the corner frequency $\omega = 1/\tau$ and is given by (refer Eqns (8.34) and (8.35))

$$-20 \log \sqrt{1 + 1} + 20 \log 1 = -3.01 \text{ dB}$$

The error at the frequency one octave below the corner frequency ($\omega = 1/(2\tau)$) is (refer Eqns (8.34) and (8.35))

$$-20 \log \sqrt{1 + \tfrac{1}{4}} + 20 \log 1 = -0.97 \text{ dB}$$

The error at the frequency one octave above the corner frequency ($\omega = 2/\tau$) is (refer Eqns (8.34) and (8.36))

$$-20 \log \sqrt{1+4} + 20 \log 2 = -0.97 \text{ dB}$$

The error at the frequency one decade below the corner frequency ($\omega = 1/(10\tau)$) is

$$-20 \log \sqrt{1+\tfrac{1}{100}} + 20 \log 1 = -0.04 \text{ dB}$$

The error at the frequency one decade above the corner frequency ($\omega = 10/\tau$) is

$$-20 \log \sqrt{1+100} + 20 \log 10 = -0.04 \text{ dB}$$

Similarly, the errors at other frequencies may be calculated.

In practice, a sufficiently accurate magnitude plot is obtained by correcting the asymptotic plot by -3 dB at the corner frequency and by -1 dB one octave below and one octave above the corner frequency, and then drawing a smooth curve through these points approaching the low and high frequency asymptotes, as shown in Fig. 8.26.

It is seen from Eqn. (8.28a) that the magnitude plot of the factor $1 + j\omega\tau$ is exactly of the same form as that of the factor $1/(1 + j\omega\tau)$, but with the opposite sign. The approximate plot consists of two straight line asymptotes, one a straight line at 0 dB for the frequency range $0 < \omega \le 1/\tau$, and the other a straight line with a slope of +20 dB/decade (or + 6 dB/octave) for the frequency range $1/\tau \le \omega < \infty$. The frequency $\omega = 1/\tau$ is the corner frequency. A sufficiently accurate plot is obtained by correcting the asymptotic plot by + 3 dB at the corner frequency and by + 1 dB one octave below and one octave above the corner frequency, and then drawing a smooth curve through these points approaching the low and high frequency asymptotes, as shown in Fig. 8.27.

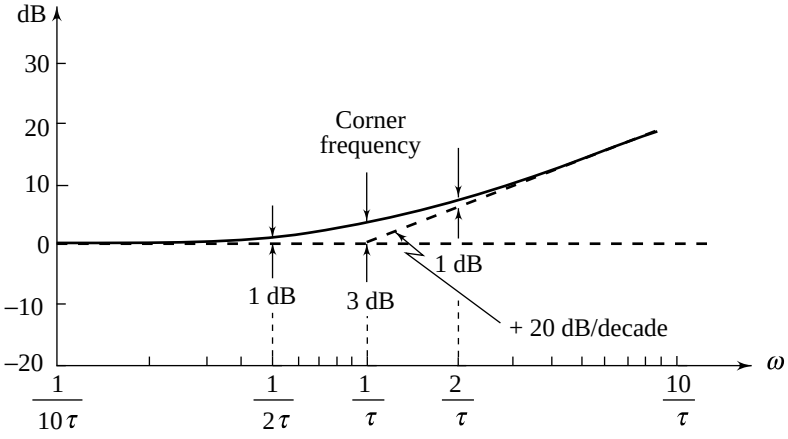


Fig. 8.27 Bode magnitude plot of $(1 + j\omega\tau)$

For the case when a given transfer function involves terms like $(1+j\omega\tau)^{\mp q}$, a similar asymptotic construction may be made. The corner frequency is still at $\omega = 1/\tau$ and the asymptotes are straight lines. The low-frequency asymptote is a straight line at 0 dB and the high-frequency asymptote has a slope of $\mp 20q$ dB/decade (or $\mp$ $6q$ dB/octave). The error involved in asymptotic approximation is q times that for $(1 + j\omega\tau)^{\mp 1}$; $\mp$ $3q$ dB at the corner frequency and $\mp$ q dB one octave below and one octave above the corner frequency.

Example 8.9 Let us draw the Bode magnitude plot for the transfer function

$$G(s) = \frac{200(s+1)}{(s+10)^2}$$

The rearrangement of the transfer function in the time-constant form gives

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

Therefore, the sinusoidal transfer function in the time-constant form is given by

$$G(j\omega) = \frac{2(1+j\omega)}{(1+j\omega/10)^2}$$

Our approach is to first construct an asymptotic plot and then apply corrections to it to get an accurate plot. The corner frequencies of the asymptotic plot in order of their occurrence as frequency increases, are

(i) $\omega_{c1} = 1$, due to zero at $s = -1$;

(ii) $\omega_{c2} = 10$, due to double pole at $s = -10$

At frequencies less than ω_{c1}, the first corner frequency, only the factor $K = 2$ is effective.

Pertinent characteristics of each factor of the transfer function are given in Table 8.1. By sketching the asymptotes for each factor and then algebraically adding them, we obtain the asymptotic plot for the transfer function $G(j\omega)$. We can, however, draw the composite asymptotic plot directly, as is outlined below.

Step 1: We start with the factor $K = 2$. Its magnitude plot is the asymptote 1; a horizontal straight line at the magnitude of 6 dB.

Step 2: Let us now add to the asymptote 1, the plot of the factor $(1 + j\omega)$ corresponding to the lowest corner frequency $\omega_{c1} = 1$. Since this factor contributes zero dB for $\omega \le \omega_{c1} = 1$, the resultant plot up to $\omega = 1$ is the same as that of the asymptote 1. For $\omega > \omega_{c1} = 1$, this factor contributes +20 dB/decade such that the resultant plot of the two factors is the asymptote 2 of slope + 20 dB/decade passing through (6 dB, 1 rad/sec) point. At $\omega = \omega_{c2} = 10$, the resultant plot has a magnitude of 26 dB as shown in Fig. 8.28.

Table 8.1 Asymptotic approximation table for Bode magnitude plot of $2(1 + j\omega)/(1 + j\omega/10)^2$

Factor	*Corner frequency*	*Asymptotic magnitude characteristic*
2	None	Constant magnitude of + 6 dB
$1 + j\omega$	$\omega_{c1} = 1$	Straight line of 0 dB for $\omega \le \omega_{c1}$; straight line of +20 dB/decade for $\omega \ge \omega_{c1}$
$1/(1 + j\omega/10)^2$	$\omega_{c2} = 10$	Straight line of 0 dB for $\omega \le \omega_{c2}$; straight line of – 40 dB/decade for $\omega \ge \omega_{c2}$

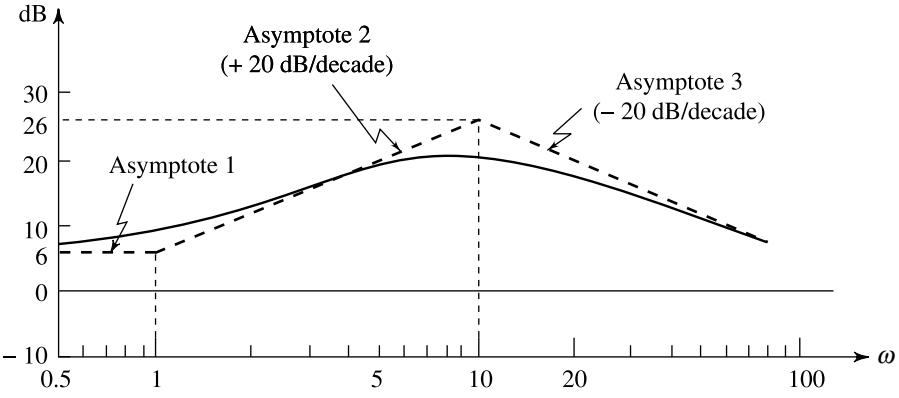


Fig. 8.28 Bode magnitude plot for Example 8.9

Step 3: We now add to the resultant plot of step 2, the plot of the factor $1/(1 + j\omega/10)^2$ corresponding to the corner frequency $\omega_{c2} = 10$. Since this factor contributes 0 dB for $\omega \le \omega_{c2} = 10$, the resultant plot up to $\omega = 10$ is the same as that of step 2. For $\omega > \omega_{c2} = 10$, this factor contributes – 40 dB/decade such that the resultant plot of the three factors is the asymptote 3 of slope (+ 20) + (– 40) = – 20 dB/decade passing through (26 dB, 10 rad/sec) point. Figure 8.28 shows the asymptotic magnitude plot of given $G(j\omega)$.

To the asymptotic plot thus obtained, corrections are to be applied. The corrections at each corner frequency and at an octave above and below the corner frequency are usually sufficient. The corner frequency $\omega_{c1} = 1$ corresponds to the first-order factor $(1 + j\omega)$: the corrections are + 3dB at $\omega = 1$, +1dB at $\omega = 0.5$ and + 1dB at $\omega = 2$. The corner frequency $\omega_{c2} = 10$ corresponds to second-order factor $1/(1 + j\omega/10)^2$: the corrections are – 6dB at $\omega = 10$, – 2 dB at $\omega = 5$ and – 2 dB at $\omega = 20$. Table 8.2 lists the net corrections.

Table 8.2 Corrections to the asymptotic magnitude plot of $2(1 + j\omega)/(1 + j\omega/10)^2$

Frequency, ω	0.5	1	2	5	10	20
Net correction	+1 dB	+3 dB	+1 dB	–2 dB	–6 dB	–2 dB

The corrected asymptotic magnitude plot of the given transfer function is shown in Fig. 8.28.

Example 8.10 Consider the transfer function $G(j\omega) = \dfrac{10(1 + j\omega/2)}{(j\omega)^2(1 + j\omega)}$

The corner frequencies of the asymptotic Bode magnitude plot of $G(j\omega)$ in order of their occurrence as frequency increases, are

(i) $\omega_{c1} = 1$, due to simple pole,
(ii) $\omega_{c2} = 2$, due to simple zero.

At frequencies less than ω_{c1}, only the factor $10/(j\omega)^2$ is effective.

Asymptotic magnitude plot of $G(j\omega)$ is shown in Fig. 8.29. The plot is obtained following the steps given below:

Step 1: We start with the factor $10/(j\omega)^2$ corresponding to double pole at the origin. Its magnitude plot is the asymptote 1, having a slope of -40 dB/decade and passing through the point 20 log 10 = 20 dB at $\omega = 1$. Asymptote 1 intersects the 0-dB line at $\omega = \sqrt{10}$.

Step 2: Let us now add to the asymptote 1, the plot of the factor $1/(1 + j\omega)$ corresponding to the lowest corner frequency $\omega_{c1} = 1$. Since this factor contributes zero dB for $\omega \leq 1$, the resultant plot up to $\omega = 1$ is the same as that of asymptote 1. For $\omega > 1$ this factor contributes -20 dB/decade such that the resultant plot of the two factors is the asymptote 2 of slope $(-40) + (-20) = -60$ dB/decade passing through (20 dB, 1 rad/sec) point. At $\omega = \omega_{c2} = 2$, the resultant plot has a magnitude of 2 dB as shown in Fig. 8.29.

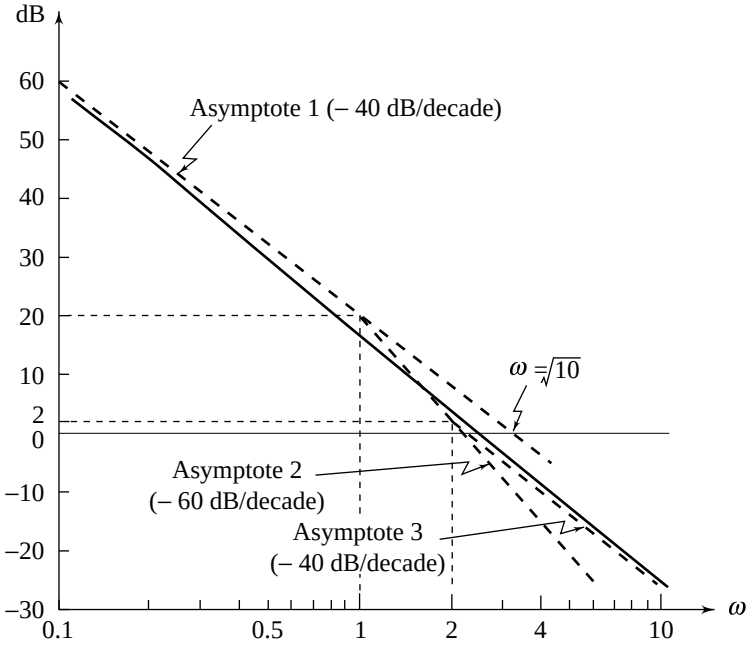


Fig. 8.29 Bode magnitude plot for Example 8.10

Step 3: We now add to the resultant plot of step 2, the plot of the factor $(1 + j\omega/2)$ corresponding to the corner frequency $\omega_{c2} = 2$. This gives rise to a straight line of slope + 20 dB/decade for $\omega > 2$, which when added to asymptote 2 results in asymptote 3 of slope $(-60) + (+20) = -40$ dB/decade passing through (2 dB, 2 rad/sec) point. Figure 8.29 shows the asymptotic magnitude plot of given $G(j\omega)$.

To the asymptotic plot thus obtained, corrections are to be applied. The corrections due to asymptotic approximation of magnitude plot of the pole factor are: -3 dB at $\omega = 1$, -1 dB at $\omega = 0.5$ and -1 dB at $\omega = 2$. To these corrections, we algebraically add the corrections due to the asymptotic approximation of magnitude plot of the zero factor: + 3 dB at $\omega = 2$, + 1 dB at $\omega = 1$, and +1 dB at $\omega = 4$. Table 8.3 lists the net corrections.

Table 8.3 Corrections to the asymptotic magnitude plot of $10(1+j\omega/2)/(j\omega)^2(1+j\omega)$

Frequency, ω	0.5	1	2	4
Net correction	–1 dB	–2 dB	+2 dB	+1 dB

The corrected asymptotic magnitude plot of the given transfer function is shown in Fig. 8.29.

Complex Poles or Zeros Transfer functions of control systems often possess quadratic factors of the form

$$\frac{1}{1+j(2\zeta/\omega_n)\omega-\omega^2/\omega_n^2};\ 0<\zeta\le 1 \tag{8.37a}$$

This term is slightly more complicated since besides being a function of ω as before, it is also a function of the variable ζ. As may be expected, the shape of the Bode plot depends strongly upon what value of damping ratio ζ is being considered.

In normalized form, the quadratic factor (8.37a) may be written as

$$\frac{1}{(1+j2\zeta u-u^2)} \tag{8.37b}$$

where $u=\omega/\omega_n$ is the normalized frequency. The magnitude of this factor is

$$\text{dB}=20\log\left|\frac{1}{1-u^2+j2\zeta u}\right|=-20\log\sqrt{(1-u^2)^2+(2\zeta u)^2} \tag{8.38}$$

For low frequencies, such that $u \ll 1$, the magnitude may be approximated by

$$\text{dB}=-20\log 1=0 \tag{8.39}$$

For high frequencies, such that $u \gg 1$, the magnitude may be approximated by

$$\text{dB}=-20\log(u^2)=-40\log u \tag{8.40}$$

Therefore an approximate magnitude plot of the quadratic factor (8.37b) consists of two straight line asymptotes, one horizontal line at 0 dB for $u \le 1$ and the other, a line with a slope -40 dB/decade for $u \ge 1$. The two asymptotes meet on 0-dB line at $u = 1$, i.e., $\omega = \omega_n$, which is the corner frequency of the plot. The asymptotic plot is shown in Fig. 8.30.

The error in the magnitude plot caused by the use of asymptotes can be calculated as follows.

For $0 < u \le 1$, the error is (refer Eqns (8.38) and (8.39))

$$-20\log\sqrt{(1-u^2)^2+(2\zeta u)^2}+20\log 1 \tag{8.41a}$$

and for $1 \le u < \infty$, the error is (refer Eqns (8.38) and (8.40))

$$-20\log\sqrt{(1-u^2)^2+(2\zeta u)^2}+40\log u \tag{8.41b}$$

The error at the corner frequency ($u = 1$ or $\omega = \omega_n$) is

$$-20\log 2\zeta \tag{8.42}$$

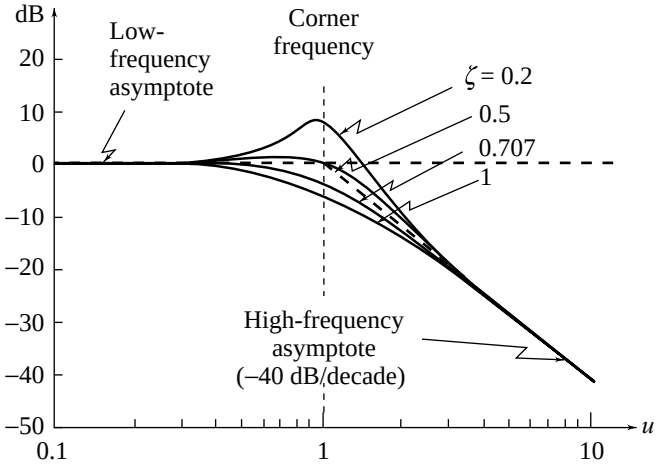


Fig. 8.30 Bode magnitude plot of $1/(1 + j2\zeta u - u^2)$

For $\zeta = 1$, the error at the corner frequency is $-20 \log 2 \cong -6$ dB. Note that for $\zeta = 1$, Eqn. (8.37) has two equal real poles:

$$\left.\frac{1}{(1 + j2\zeta u - u^2)}\right|_{\zeta=1} = \frac{1}{(1 + ju)^2} = \frac{1}{(1 + j\omega/\omega_n)^2}$$

From expression (8.42) we see that at the corner frequency, the error is negative for $\zeta > 0.5$ and positive for $\zeta < 0.5$. The corrected magnitude curves for various values of ζ are shown in Fig. 8.30. From this figure, it is seen that for small values of ζ, the magnitude curves have pronounced peak at a frequency slightly less than the corner frequency. In fact, smaller the value of ζ the higher the peak. We shall discuss this peak in some detail.

Consider Fig. 8.31, which shows what might be a typical plot of (refer Eqn. (8.38))

$$\text{dB} = -20 \log \sqrt{(1 - u^2)^2 + (2\zeta u)^2}$$

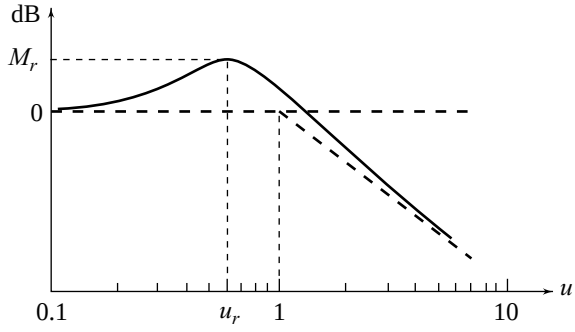


Fig. 8.31 A typical magnitude curve of $1/(1 + j2\zeta u - u^2)$

The peak value occurs at normalized frequency which is labeled u_r ($u_r = \omega_r/\omega_n$) and the height of the peak is M_r. This peak will occur at a u where $\sqrt{(1-u^2)^2 + (2\zeta u)^2}$ is at a minimum. The minimum can be found by simple differentiation.

$$\frac{d}{du}\left[\sqrt{(1-u^2)^2 + (2\zeta u)^2}\right]\Bigg|_{u=u_r} = \frac{\frac{1}{2}\left[-4(1-u_r^2)u_r + 8\zeta^2 u_r\right]}{[(1-u_r^2)^2 + (2\zeta u_r)^2]^{1/2}} = 0$$

which gives

$$4u_r^3 - 4u_r + 8\zeta^2 u_r = 0$$

or

$$u_r = \sqrt{1-2\zeta^2} \tag{8.43}$$

The corresponding peak value M_r is then

$$M_r = -20 \log \sqrt{(1-u_r^2)^2 + (2\zeta u_r)^2}$$

$$= -20 \log 2\zeta\sqrt{1-\zeta^2} \text{ dB} \tag{8.44}$$

Note that as ζ goes to zero, the peak value M_r goes off to infinity (refer Eqn. (8.44)) and frequency u_r approaches the corner frequency (refer Eqn. (8.43)). For $0 < \zeta < 1/\sqrt{2}$, u_r is less than the corner frequency. For $\zeta \geq 1/\sqrt{2}$, the errors given by expressions (8.41) are negative for all frequencies and therefore magnitude curves do not have a peak.

For given values of ζ and ω_n, an exact dB versus $\log\omega$ curve is obtained from the following expression (refer Eqn. (8.38)).

$$\text{dB} = -20 \log \sqrt{\left[1-\left(\frac{\omega}{\omega_n}\right)^2\right]^2 + \left(\frac{2\zeta\omega}{\omega_n}\right)^2} \tag{8.45}$$

Most control-system designers will have access to computer programs to do this job. The following procedure is useful for hand plotting.

Draw two straight line asymptotes: one a 0-dB line for $0 < \omega \leq \omega_n$ and the other a -40 dB/decade line for $\omega_n \leq \omega < \infty$. The two asymptotes meet at the corner frequency $\omega = \omega_n$. Correct the asymptotic plot by M_r dB at the frequency ω_r where (refer Eqns (8.44) and (8.43))

$$M_r = -20 \log 2\zeta\sqrt{1-\zeta^2} \text{ dB} \tag{8.46a}$$

$$\omega_r = \omega_n \sqrt{1-2\zeta^2} \text{ rad/sec} \tag{8.46b}$$

Of course, this correction is applied when $\zeta < 1/\sqrt{2}$. Correct the asymptotic plot by $-20 \log (2\zeta)$ dB at the corner frequency $\omega = \omega_n$ (refer expression (8.42)). Draw a smooth curve through these points approaching the low and high frequency asymptotes.

The accuracy of the plot can be improved by correcting the asymptotic plot at some other frequencies also; say one octave below and above the corner frequency. The errors are calculated from the following expressions (refer expressions (8.41)).

$$-20\log\sqrt{\left[1-\left(\frac{\omega}{\omega_n}\right)^2\right]^2+\left[2\zeta\left(\frac{\omega}{\omega_n}\right)\right]^2} \quad \text{for } 0<\omega\le\omega_n \tag{8.47a}$$

$$-20\log\sqrt{\left[1-\left(\frac{\omega}{\omega_n}\right)^2\right]^2+\left[2\zeta\left(\frac{\omega}{\omega_n}\right)\right]^2}+40\log\left(\frac{\omega}{\omega_n}\right) \quad \text{for } \omega_n\le\omega<\infty \tag{8.47b}$$

The magnitude plot of the factor $[(1+j(2\zeta/\omega_n)\omega-\omega^2/\omega^2{}_n]$ is exactly of the same form as that of the factor $1/[1+j(2\zeta/\omega_n)\omega-\omega^2/\omega_n{}^2]$, but with the opposite sign. Similar asymptotic construction may be made for the case when a given transfer function involves terms like $[1+j\,(2\zeta/\omega_n)\omega-\omega^2/\omega_n{}^2]^{\mp r}$.

Example 8.11 Let us draw the Bode magnitude plot of the transfer function

$$G(s)=\frac{200(s+2)}{s(s^2+10\,s+100)}$$

The rearrangement of the transfer function in time-constant form gives

$$G(s)=\frac{4(1+s/2)}{s[1+s/10+(s/10)^2]}$$

Therefore, the sinusoidal transfer function in time-constant form is given by

$$G(j\omega)=\frac{4(1+j\omega/2)}{j\omega(1+j\omega/10-\omega^2/100)} \tag{8.48}$$

The corner frequencies of the asymptotic plot of $G(j\omega)$, in order of their occurrence as frequency increases, are

(i) $\omega_{c1}=2$, due to zero at $s=-2$;

(ii) $\omega_{c2}=10$, due to pair of complex conjugate poles with $\zeta=0.5$ and $\omega_n=10$.

At frequencies less than ω_{c1}, only the factor $4/(j\omega)$ is effective.

Asymptotic magnitude plot of $G(j\omega)$ is drawn in Fig. 8.32. The plot is obtained following the steps given below.

Step 1: We start with the factor $4/(j\omega)$ corresponding to the pole at the origin. Its magnitude plot is the asymptote 1 having a slope of –20 dB/decade passing through the point 20 log 4 = 12 dB at $\omega = 1$. Asymptote 1 intersects the 0-dB line at $\omega=4$. At $\omega=\omega_{c1}=2$, the magnitude is 6 dB.

Step 2: We now add to the asymptote 1, the plot of the factor $(1+j\omega/2)$ corresponding to the lowest corner frequency $\omega_{c1}=2$. Since this factor contributes 0 dB for $\omega\le 2$, the resultant plot up to $\omega=2$ is the same as that of asymptote 1. For $\omega>2$, this factor contributes +20 dB/decade such that the resultant plot of the two factors is the asymptote 2 of slope (– 20) + (+ 20) = 0 dB/decade passing through (6 dB, 2 rad/sec) point. At $\omega=\omega_{c2}=10$, the resultant plot has a magnitude of 6 dB as shown in Fig. 8.32.

Step 3: We now add to the resultant plot of step 2, the plot of the factor $1/(1+j\omega/10-\omega^2/100)$ corresponding to the corner frequency $\omega_{c2}=\omega_n=10$. This gives rise to a straight line of slope – 40 dB/decade for $\omega>10$ which when added to asymptote 2 results in asymptote 3 with a slope 0 + (– 40) = – 40 dB/decade passing through the point (6 dB, 10 rad/sec).

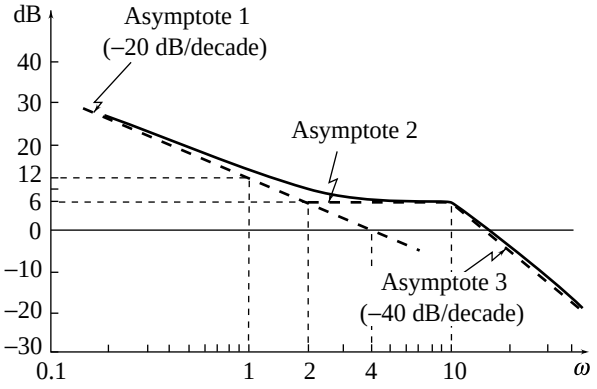


Fig. 8.32 Bode magnitude plot for Example 8.11

To the asymptotic plot thus obtained, corrections are to be applied. Table 8.4 lists the corrections.

Table 8.4 Corrections to the asymptotic plot of $4\,(1 + j\omega/2)/(j\omega)(1 + j\omega/10 - \omega^2/100)$

Frequency	*Correction*	*Remarks*
$\omega = 2$	+ 3 dB	
$\omega = 1$	+ 1 dB	
$\omega = 4$	+ 1 dB	
$\omega = 10$	0 dB	Error = $-20 \log 2\zeta = 0$
$\omega = 10/\sqrt{2}$	$-20 \log (\sqrt{3}/2)$ dB	$M_r = -20 \log 2\zeta\sqrt{1-\zeta^2}$

The corrected asymptotic plot of the given transfer function is shown in Fig. 8.32.

8.5.2 Phase Plot

The phase angle at any frequency can be obtained as the algebraic sum of the phase angles due to various factors in the transfer function, in a manner similar to that used for obtaining magnitude plots (refer Eqns (8.28)). To illustrate the procedure, we consider the transfer function

$$G(j\omega) = \frac{K(1 + j\omega\tau_1)}{j\omega(1 + j\omega\tau_2)(1 + (j2\zeta/\omega_n)\omega - \omega^2/\omega_n^2)} \tag{8.49}$$

The phase of $G(j\omega)$ is

$$\begin{aligned} \angle G(j\omega) &= \tan^{-1}\omega\tau_1 - 90° - \tan^{-1}\omega\tau_2 - \tan^{-1}2\zeta\omega\omega_n/(\omega_n^2 - \omega^2) \\ &= \phi_1 + \phi_2 + \phi_3 + \phi_4 \end{aligned} \tag{8.50}$$

The first term on the right-hand side of Eqn. (8.50) is due to the real zero, and is given by

$$\phi_1 = \tan^{-1}\omega\tau_1$$

The values of ϕ_1 for various values of ω are given in Table 8.5, and these values are plotted in Fig. 8.33. We can approximate the exact curve with the straight line construction shown in Fig. 8.33. The straight line approximation for the phase characteristic breaks from 0° at the frequency $\omega = 1/10\,\tau_1$, and breaks again to a

constant value of 90° at $\omega = 10/\tau_1$. The maximum error in the approximation is 5.7°, and occurs at the two corners $1/10\tau_1$ and $10/\tau_1$.

The second term on the right-hand side in Eqn. (8.50) is due to the pole at the origin and is identically equal to $-90°$. The third term in Eqn. (8.50) is due to the real pole and is similar to the first term but for reversal in sign. We shall now look at the fourth term, which is due to the complex-conjugate pair of poles, and is given by

$$\phi_4 = -\tan^{-1}\frac{2\zeta\omega_n\omega}{\omega_n^2 - \omega^2}$$

Table 8.5 Phase characteristics of $(1 + j\omega\tau_1)$

ω (rad/sec)	*Exact value*	*Straight line approximation*
$0.05/\tau_1$	2.9°	0°
$0.1/\tau_1$	5.7°	0°
$0.2/\tau_1$	11.3°	13.5°
$0.5/\tau_1$	26.6°	31.5°
$0.8/\tau_1$	38.7°	40.6°
$1/\tau_1$	45°	45°
$2/\tau_1$	63.4°	58.5°
$5/\tau_1$	78.7°	76.5°
$8/\tau_1$	82.9°	85.6°
$10/\tau_1$	84.3°	90°
$20/\tau_1$	87.1°	90°

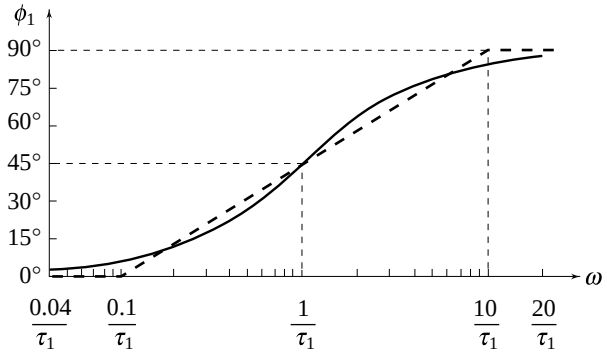


Fig. 8.33 Phase characteristics of $(1 + j\omega\tau_1)$

For $\omega << 0.1\omega_n$, we have $\phi_4 = 0$, and for $\omega >> 10\omega_n$, we get $\phi_4 = 180°$. At the intermediate frequency $\omega = \omega_n$, we have $\phi_4 = 90°$. The plots of the asymptotic approximation and the actual curve for various values of ζ are shown in Fig. 8.34. It is seen that for small values of ζ, the actual phase angle changes very rapidly for ω near the undamped natural frequency ω_n.

The overall phase angle is now obtained as the algebraic sum of the phase angle due to each factor. The addition, however, is not as simple as in the case of the magnitude plot due to the fact that the straight line approximation consists of three parts over four decades. The only exception arises when the various poles and zeros have 'large' separation, so that frequency range of phase angle plot of one factor does not overlap with the frequency range of phase angle plot of any other factor. In this case, the phase angle plots for the various factors can be superimposed on each other to obtain the overall phase.

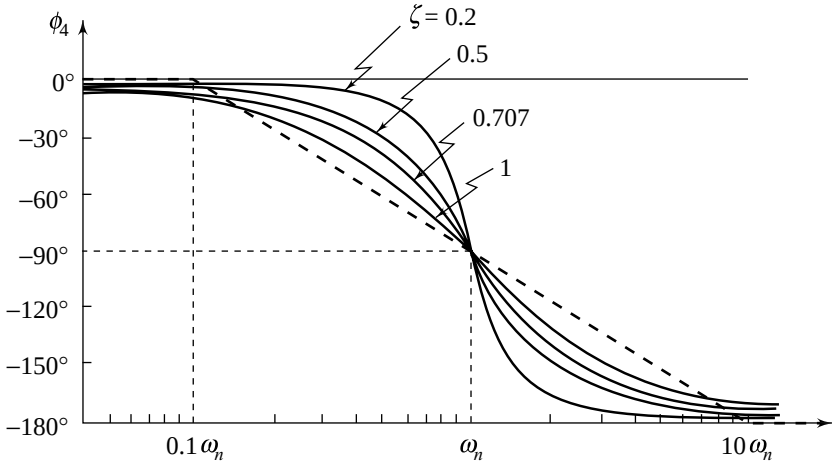


Fig. 8.34 Phase characteristics of $1/[1 + (j\,2\zeta/\omega_n)\omega - \omega^2/\omega_n{}^2\ ]$

For a general case, it is usually advisable to make a table of phase angle against frequency for each factor and then obtain the total phase as the algebraic sum of these. Better accuracy is obtained by using the exact values in the table instead of those obtained through the asymptotic approximation. The computational effort to achieve this accuracy is not heavy.

Example 8.12 Consider the transfer function given in Eqn. (8.48) for which the magnitude plot was obtained in Fig. 8.32. We shall now plot the phase curve.

The phase angle due to each factor for different values of ω is given in Table 8.6. The plot of the total phase is shown in Fig. 8.35.

Table 8.6 Phase angle due to different factors of the transfer function given by Eqn. (8.48)

ω (rad/sec)	*Poles at the origin*	*Zero at* $s = -2$	*Complex conjugate poles*	*Total phase angle,* ϕ
0.1	– 90°	2.86°	– 0.57°	– 87.71°
0.2	– 90°	5.71°	– 1.15°	– 85.44°
0.5	– 90°	14.04°	– 2.87°	– 78.33°
1	– 90°	26.57°	– 5.77°	– 69.20°
2	– 90°	45.00°	– 11.77°	– 56.77°
5	– 90°	68.20°	– 33.69°	– 55.49°
10	– 90°	78.69°	– 90.00°	– 101.32°
20	– 90°	84.29°	– 146.31°	– 152.02°
50	– 90°	87.71°	– 168.23°	– 170.52°
100	– 90°	88.85°	– 174.22°	– 175.68°

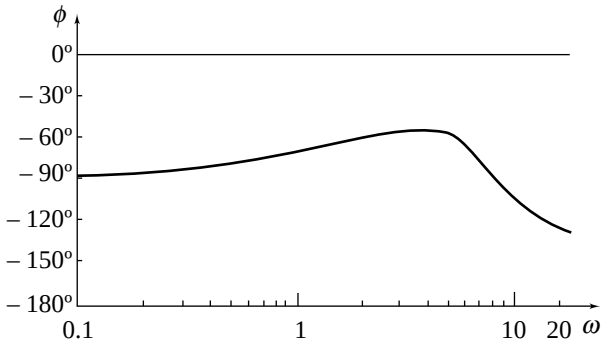


Fig. 8.35 Bode phase plot for Example 8.12

8.6 STABILITY MARGINS ON THE BODE PLOTS

The general objective of the frequency-domain design procedures is to shape the Nyquist plot such that the critical point $-1+j0$ is avoided with some reasonable stability margins, thereby providing acceptable closed-loop response characteristics. Direct use of the Nyquist plot for design is not particularly convenient, since changes in parameters other than gain require extensive plot revisions. Fortunately, general characteristics of Nyquist plots can be visualized with reasonable accuracy in most cases of interest from Bode plots, and Bode plots are easily constructed and modified.

Since the Bode plot corresponds to only the positive portion of the $j\omega$-axis from $\omega = 0$ to $\omega = \infty$ in the s-plane, its use in stability studies is limited to the determination of gain and phase crossover points and the corresponding phase and gain margins. In practice, the Bode plot is more convenient to apply when the system has open-loop transfer function with no poles in right half s-plane.

Since the straight line approximation of the Bode plot is relatively easier to construct, the data necessary for the other frequency-domain plots, such as the Nyquist plots, can be easily generated from the Bode plot.

Example 8.13 A unity-feedback system has open-loop transfer function

$$G(s) = \frac{K}{(s+2)(s+4)(s+5)}; K = 200 \tag{8.51a}$$

Let us determine the stability of this system implementing the Nyquist stability criterion using Bode plots.

We begin by sketching the Bode magnitude and phase plots. In the time-constant form, the transfer function $G(s)$ is expressed as

$$G(s) = \frac{5}{(1+s/2)(1+s/4)(1+s/5)} \tag{8.51b}$$

The sinusoidal transfer function

$$G(j\omega) = \frac{5}{(1+j\omega/2)(1+j\omega/4)(1+j\omega/5)}$$

Bode magnitude and phase plots for $G(j\omega)$ are shown in Fig. 8.36.

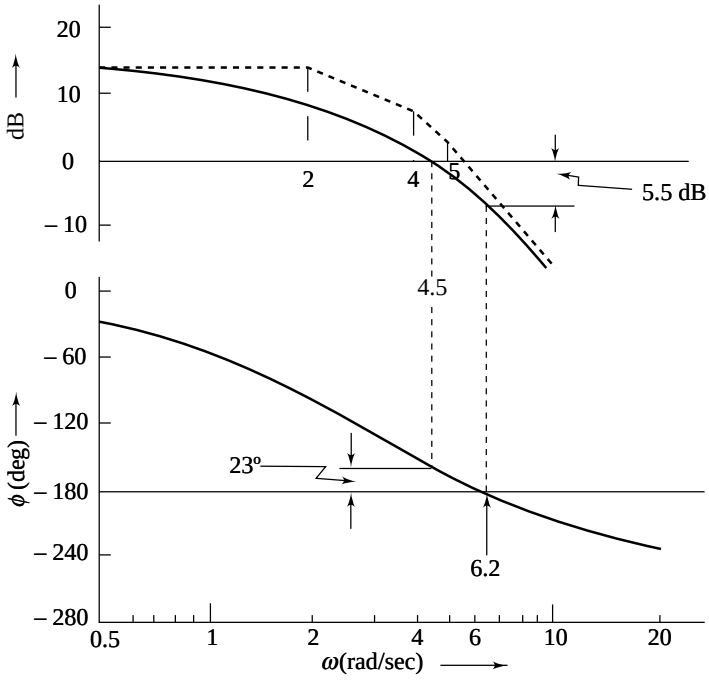


Fig. 8.36 Bode plot for Example 8.13

The Nyquist criterion for this example tells us that we want zero encirclements of the $-1+j0$ point for stability. Thus, we recognise that on Bode magnitude plot, the magnitude must be less than 0 dB (which corresponds to unity gain of $G(j\omega)$ on polar plane) at that frequency where the phase is $-180°$. Accordingly, we see that at a frequency of 6.2 rad/sec where the phase is $-180°$, the magnitude is -5.5 dB. Therefore, an increase in gain of +5.5 dB is possible before the system becomes unstable. Since the magnitude plot was drawn for a gain of 200 (refer Eqn. (8.51a)), + 5.5 dB (20 log 1.885 ≅ 5.5) represents the allowed increase in gain above 200. Hence the gain for instability is 200 × 1.885 = 377. The closed-loop system is stable for $0 < K < 377$.

Next we show how to evaluate the gain and phase margins using Bode plots. The gain margin is found by using the phase plot to find the phase crossover frequency ω_ϕ where the phase angle is $-180°$. At this frequency, look at the magnitude plot to determine the gain margin, GM, which is the gain required to raise the magnitude curve to 0 dB. From Fig. 8.36, we find that

$$\omega_\phi = 6.2 \text{ rad/sec}; \; GM = 5.5 \text{ dB}$$

The phase margin is found by using the magnitude curve to find the gain crossover frequency ω_g, where the gain is 0 dB. On the phase curve at that frequency, the phase margin, ΦM, is the difference between the phase value and $-180°$. From Fig. 8.36, we find that

$$\omega_g = 4.5 \text{ rad/sec}; \; \Phi M = -157° - (-180°) = 23°$$

Example 8.14 Consider now a unity-feedback system with open-loop transfer function

$$G(s) = \frac{K}{s(1+0.1s)(1+0.2s)}; \; K = 100$$

The Bode plot of $G(j\omega)$ is shown in Fig. 8.37. From this figure we find that

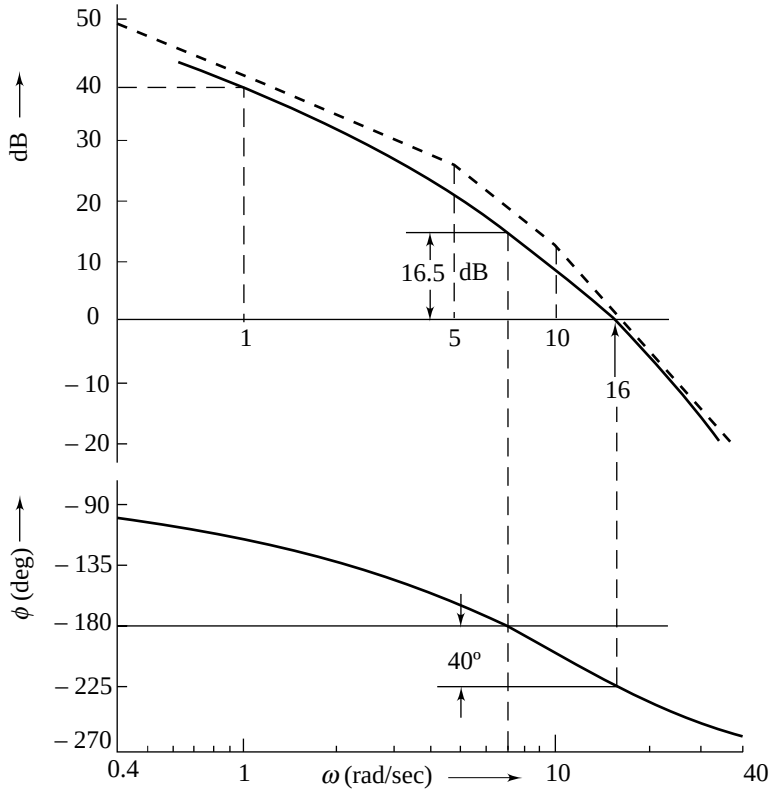


Fig. 8.37 Bode plot for Example 8.14

ω_ϕ = phase crossover frequency = 7 rad/sec

GM = gain margin = – 16.5 dB

ω_g = gain crossover frequency = 16 rad/sec

ΦM = phase margin = – 40°

The closed-loop system is thus unstable.

If we were to decrease the gain by 16.5 dB, all points on dB vs ω plot would decrease in magnitude by this amount without affecting the phase plot. The closed-loop system would be at the limit of instability. The gain K must be less than $100 \times x$ ($16.5 = 20 \log x$), i.e., the stability range is $0 < K < 668$.

8.7 STABILITY ANALYSIS OF SYSTEMS WITH DEAD-TIME

Figure 8.38 shows the block diagram of a system with dead-time elements in the loop. Open-loop transfer function of the system is

$$G(s)H(s) = G_1(s)e^{-s\tau_D} \tag{8.52}$$

where $G_1(s)$ is a rational function.

Root locus analysis of systems with dead-time elements is possible, but the method is quite complex. In Section 7.11, root locus analysis was carried out using the following approximation:

$$e^{-s\tau_D} \cong \frac{1 - s\tau_D/2}{1 + s\tau_D/2}$$

This approximation works fairly well as long as the dead-time τ_D is small in comparison to the system time-constants.

The frequency-domain graphical methods provide a simple yet exact approach to handle the dead-time problem since the factor $e^{-s\tau_D}$ is readily interpreted in terms of either the Nyquist or the Bode plot.

8.7.1 Nyquist Plots of Systems with Dead-Time

Assume that $G_1(s)$ in Fig. 8.38 is an integrator. Then the open-loop transfer function of the system becomes

$$G(s)H(s) = \frac{e^{-s\tau_D}}{s}$$

The sinusoidal transfer function is

$$G(j\omega)H(j\omega) = \frac{e^{-j\omega\tau_D}}{j\omega}$$

Clearly,

$$|G(j\omega)H(j\omega)| = 1/\omega;\ \angle G(j\omega)H(j\omega) = -\omega\tau_D - \pi/2$$

$$G(j\omega)H(j\omega) = \frac{1}{j\omega}(\cos\omega\tau_D - j\sin\omega\tau_D)$$

$$= -\frac{\sin\omega\tau_D}{\omega} - j\frac{\cos\omega\tau_D}{\omega}$$

$$\lim_{\omega\to 0}[G(j\omega)H(j\omega)] = -\tau_D - j\infty$$

Since the magnitude decreases monotonically, and the phase angle also decreases monotonically indefinitely, the polar plot of the given transfer function will spiral into the $\omega \to \infty$ point at the origin, as shown in Fig. 8.39. A curious reader may like to find some of the intersections with real and imaginary axes. For the first intersection with the real axis, set

$$\text{Im}\,[G(j\omega)H(j\omega)] = 0 \text{ to obtain } \omega\tau_D = \pi/2.$$

$$\text{Re}[G(j\omega)H(j\omega)]\Big|_{\omega=\frac{\pi}{2\tau_D}} = -\frac{1}{\omega}\Big|_{\omega=\frac{\pi}{2\tau_D}} = -2\tau_D/\pi$$

Once the Nyquist plot is constructed, the stability of the closed-loop system is determined in the usual manner. From Fig. 8.39, we observe that relative stability of the system reduces as τ_D increases. Sufficiently large values of τ_D may drive the system to instability.

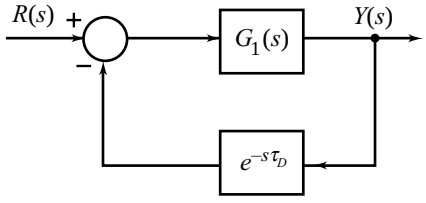


Fig. 8.38 A closed-loop system with dead-time elements in the loop

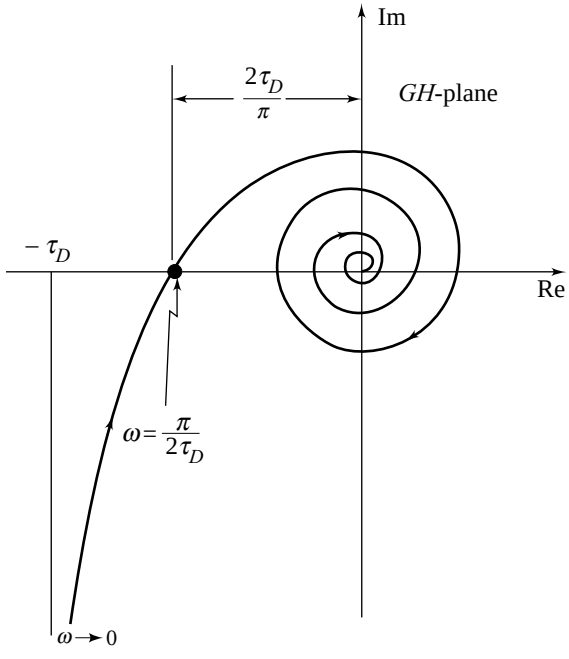


Fig. 8.39 Polar plot of $e^{-s\tau_D}/s$

The Critical Trajectory Thus far, we have used the point $-1+j0$ of the $G(s)H(s)$-plane as the critical point for stability analysis using Nyquist criterion. We can extend the critical-point idea into a trajectory if necessary.

The roots of the characteristic equation of the system shown in Fig. 8.38 satisfy (refer Eqn. (8.52))

$$1 + G_1(s)\, e^{-s\tau_D} = 0$$

or

$$G_1(s)\, e^{-s\tau_D} = -1 \tag{8.53}$$

The right-hand side of the last equation points to the fact that $-1+j0$ is the critical point for stability analysis of the closed-loop system. Equation (8.53) can be written as

$$G_1(s) = -e^{+s\tau_D} \tag{8.54}$$

The corresponding condition for the system without dead-time is

$$G_1(s) = -1 \tag{8.55}$$

Comparison of Eqns (8.54) and (8.55) reveals that the effect of dead-time is simply to shift the critical stability point $-1+j0$ to $-e^{s\tau_D}$, which describes a *critical trajectory*. When $\omega\tau_D = 0$, the trajectory starts at the point $-1+j0$, and as $\omega\tau_D$ increases, the critical point traces out a circle with unit radius centred at the origin of $G_1(s)$-plane, in the counterclockwise direction.

Example 8.15 Let us determine, with the help of Nyquist stability criterion, the maximum value of τ_D for stability of the closed-loop system of Fig. 8.38 with

$$G_1(s) = \frac{1}{s(s+1)(s+2)}$$

From Eqn. (8.54), we have with $s = j\omega$,

$$G_1(j\omega) = \frac{1}{j\omega(j\omega+1)(j\omega+2)} = -e^{+j\omega\tau_D}$$

Figure 8.40 shows the $G_1(j\omega)$ locus together with the critical trajectory of $-e^{+j\omega\tau_D}$. The frequency at which the $G_1(j\omega)$ locus intersects the critical trajectory is found by setting the magnitude of $G_1(j\omega)$ to unity, i.e.,

$$|G_1(j\omega)| = \left|\frac{1}{-3\omega^2 + j\omega(2-\omega^2)}\right| = 1$$

which gives $\omega = 0.446$ rad/sec.

Since

$$\angle G_1(j\omega)\big|_{\omega=0.446} = \angle -e^{j\omega\tau_D}\big|_{\omega=0.446}$$

we obtain (refer Fig. 8.40)

$$53.4\ (\pi/180) = 0.932 \text{ rad} = 0.446\ \tau_D \text{ rad}$$

which gives

$$\tau_D = 2.09 \text{ sec}$$

From Fig. 8.40, we observe that the critical point on the critical trajectory is encircled by the $G_1(j\omega)$ locus for $\omega\tau_D > 0.932$ ($\tau_D > 2.09$) and is not encircled for $\omega\tau_D < 0.932$ ($\tau_D < 2.09$); and hence we conclude that the system under discussion is stable if $\tau_D < 2.09$ sec.

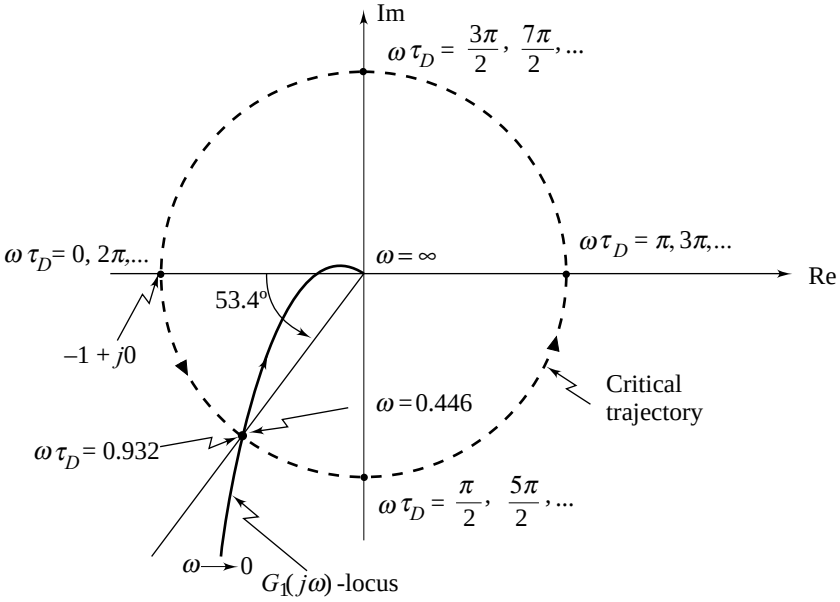


Fig. 8.40 Critical trajectory and $G_1(j\omega)$-locus for Example 8.15

8.7.2 Bode Plots of Systems with Dead-Time

The stability analysis of systems with dead-time can be conducted easily using the Bode plots. Since $20 \log |e^{-j\omega\tau_D}| = 0$, the magnitude plot of a system is unaffected by the presence of dead-time. The dead-time, of course, contributes a phase angle of $-(\omega\tau_D \times 180°)/\pi$, thereby causing the modification of the phase plot. An illustrative example follows.

Example 8.16 Consider the closed-loop system of Fig. 8.38 with

$$G_1(s) = \frac{K}{s(s+1)(s+2)} = \frac{K/2}{s(1+s)(1+s/2)}$$

Figure 8.41 shows the Bode plot of $G_1(j\omega)e^{-j\omega\tau_D}$ with $K = 1$ and $\tau_D = 0$. We find from this figure that

ω_g = gain crossover frequency = 0.446 rad/sec
ΦM = phase margin = 53.4°
ω_ϕ = phase crossover frequency = 1.4 rad/sec
GM = gain margin = 16 dB

The effect of dead-time is to add the phase shift of $-(\omega\tau_D \times 180)/\pi$ degrees to the phase curve while not affecting the magnitude curve. The adverse effect of dead-time on stability is apparent since the negative phase shift caused by the dead-time increases rapidly with the increase in ω. Let us set $\tau_D = 1$ sec and find the critical value of K for stability. Figure 8.41 also shows Bode plot of $G_1(j\omega)e^{-j\omega\tau_D}$ with $K = 1$ and $\tau_D = 1$. The magnitude curve is unchanged; the phase curve drops with the increase in ω, and the phase crossover frequency is now at 0.66 rad/sec. The gain margin is 4.5 dB. Note that 20 log 1.67 = 4.5 dB; therefore the critical value of K for stability when $\tau_D = 1$ is 1.67.

Let us now determine the critical value of dead-time, τ_D, for stability with $K = 1$. Since the phase margin with $\tau_D = 0$ and $K = 1$ is 53.4°, a phase lag of 53.4° can be introduced by dead-time before instability sets in. The critical value of dead-time is, therefore, given by the relation

$$\omega_g \tau_D \times 180/\pi = 53.4$$

which gives

$$\tau_D = 2.09 \text{ sec}$$

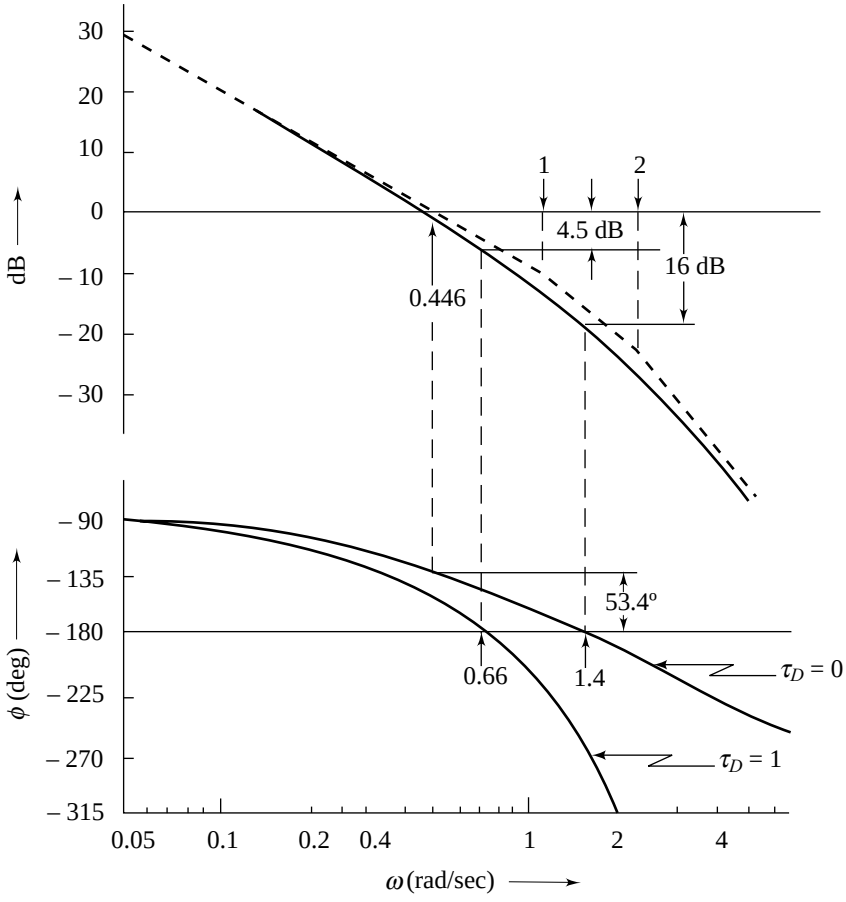


Fig. 8.41 Bode plots of Example 8.16

8.8 FREQUENCY RESPONSE MEASUREMENTS

In this section we will describe how a system responds to sinusoidal input—called the system's *frequency response*—and how the frequency response measurements can be used for design purposes.

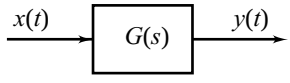


Fig. 8.42 A stable linear time-invariant system

To introduce the ideas, we consider a *stable* linear time-invariant system shown in Fig. 8.42. The input and output of the system are described by $x(t)$ and $y(t)$ respectively. The response of the system to sinusoidal input

$$x(t) = X \sin \omega t$$

will consist of two components: the transient component and the steady-state component (refer Section 2.9). The transient component will disappear as time goes by. The steady-state output that results after the transient dies out, is of the form

$$y(t) = A(\omega) \times \sin(\omega t + \phi(\omega))$$

which is a sine wave of the same frequency as the input; it differs from the sinusoidal input only in the amplitude and the phase angle. Both the amplitude and the phase angle change when we use a different frequency ω.

The output-input amplitude ratio $A(\omega)$, and the phase angle $\phi(\omega)$ between the output and input sinusoids, are given by the relations

$$A(\omega) = \frac{\text{Amplitude of the sinusoidal output}}{\text{Amplitude of sinusoidal input}} = |G(s = j\omega)|$$

$\phi(\omega)$ = Phase angle of the sinusoidal output – phase angle of the sinusoidal input

$$= \angle G(s = j\omega)$$

The variation of $A(\omega)$ and $\phi(\omega)$ with frequency ω is, by definition, the *frequency response* of the system.

If the system equations are known, complete with numerical values, we can derive its transfer function $G(s)$ and therefrom the *sinusoidal transfer function* $G(j\omega)$. Frequency response of the system can be computed from $G(j\omega)$.

If the equations and/or parameter values are not known, but the physical system or its components exist and are available for test, we can generate the frequency-response data by experimental measurements:

(i) A sinusoidal input of known amplitude and frequency is applied to the system.
(ii) The system output is allowed to 'settle' into a steady-state pattern.
(iii) The amplitude and relative phase of the sinusoidal output are measured and recorded.
(iv) This procedure is repeated for values of ω spanning the frequency range of interest.

Obviously, no steady-state response exists for an unstable system; so a frequency-response test will directly indicate instability.

Raw measurements of the output amplitude and phase of a stable plant undergoing a sinusoidal input are sufficient to analyse feedback system stability. No intermediate processing of the data is necessary because Nyquist plot/Bode plot can be sketched directly from experimental measurements.

Sometimes it is desirable to obtain an approximate model, in terms of a transfer function, directly from the frequency-response measurements. Determination of a transfer function $G(s)$ from measured data is an *identification problem*.

For a stable system, $|G(j\omega)|$ and $\angle G(j\omega)$ can be obtained by measurement. This is also possible if the system is marginally stable with a pole at $s = 0$. Experimental data are used to obtain the exact log-magnitude and phase angle curves of the Bode plot. Asymptotes are drawn on the exact log-magnitude curves by utilizing the fact that their slopes must be multiples of ± 20 dB/decade. From these asymptotes, the system type and approximate time-constants are determined.

Careful interpretation of the phase angle curve is necessary to identify whether the transfer function is a *minimum phase* or a *nonminimum phase* transfer function. Let us first define these terms.

Consider the following transfer functions:

$$G_1(j\omega) = \frac{1 - j\omega\tau}{(1 + j\omega\tau_1)(1 + j\omega\tau_2)}\ ;\ G_2(j\omega) = \frac{1 + j\omega\tau}{(1 + j\omega\tau_1)(1 + j\omega\tau_2)}$$

It can easily be deduced that the two transfer functions have the same magnitude characteristics:

$$|G_1(j\omega)| = |G_2(j\omega)|;\ \omega \geq 0$$

However, the phase characteristics are different for the two cases as illustrated in Fig. 8.43. The transfer function with a zero in the right half s-plane undergoes a net change in phase, when evaluated for frequency inputs between zero and infinity, which is greater than that for the transfer function with the zero in the left half s-plane; magnitude plots of two transfer functions being the same. Based on the phase shift characteristics, transfer functions are classified as minimum phase and nonminimum phase.

The range of phase shift of a *minimum phase transfer function* is the least possible corresponding to a given magnitude curve, whereas the range of phase shift of *nonminimum phase transfer function* is greater than the minimum possible for the given magnitude curve.

It can easily be established that a proper rational transfer function is a minimum phase transfer function if all of its zeros lie in the left half s-plane. It is a nonminimum phase transfer function if it has one or more zeros in the right half s-plane[6].

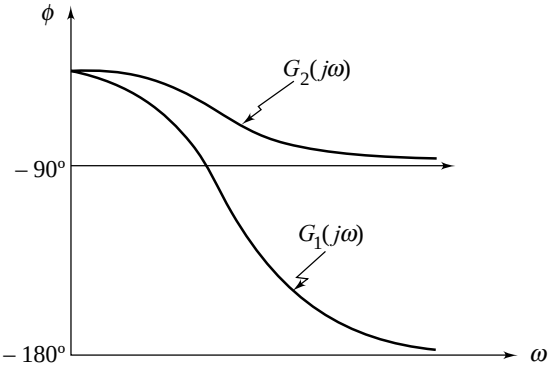


Fig. 8.43 Phase angle characteristics of minimum phase and nonminimum phase transfer functions

Using the basic definition, we can easily establish the minimum phase or nonminimum phase character of irrational transfer functions as well. A nonminimum phase irrational transfer function of practical importance is one with $e^{-\tau_D s}$ as its factor, e.g.,

$$G(j\omega) = \frac{(1 + j\omega\tau)\, e^{-j\omega\tau_D}}{(1 + j\omega\tau_1)(1 + j\omega\tau_2)}$$

With these observations, we can go back to the identification problem. If the Bode plot is obtained by frequency response measurements, then, except for a possible pole at $s = 0$, all the poles must lie in the left half s-plane. However, zeros in the right half s-plane and/or presence of dead-time elements cannot be ruled out. Therefore, care must be taken in interpreting the phase angle plot to detect the presence of right half plane zeros and/or dead-time elements. We illustrate the identification procedure through examples.

Example 8.17 Table 8.7 gives experimentally-obtained frequency-response data of a system. From this data, let us determine the approximate transfer function model of the system.

Bode plot shown in Fig. 8.44 is drawn using the data of Table 8.7. First we approximate the magnitude plot by the three straight dashed lines shown. They intersect at $\omega = 15$ and $\omega = 150$. We begin with the leftmost part of the magnitude plot. The low frequency asymptote is a horizontal line at 25 dB. It indicates that the transfer function represents a type-0 system with a gain K given by

$$20 \log K = 25 \text{ or } K = 17.8$$

[6] Nonminimum Phase response (also called *inverse response*) is exhibited by some processes, e.g., liquid level in a drum boiler. The dynamic behavior of such processes deviates drastically from what we have seen so far; initially the response is in the opposite (inverse) direction to what it eventually ends up. Such a behavior is the net result of two opposing effects. When a system possesses an inverse response, its transfer function has a zero in the right-half plane. Nonminimum phase systems are particularly difficult to control and require special attention.

Table 8.7 Experimental frequency response data

f(Hz)	*ω(rad/sec)*	*Gain(dB)*	*Phase(deg)*
60	377	– 7.75	– 155
50	314	– 4.3	– 150
40	251	– 0.2	– 145
35	219	0.75	– 140
25	157	5.16	– 135
20	126	7.97	– 120
16	100	10.5	– 110
10	63	15.0	– 100
7	44	16.9	– 85
2.5	16	20.4	– 45
1.3	8	21.6	– 30
0.22	1.38	24.0	– 5
0.16	1.0	24.1	0

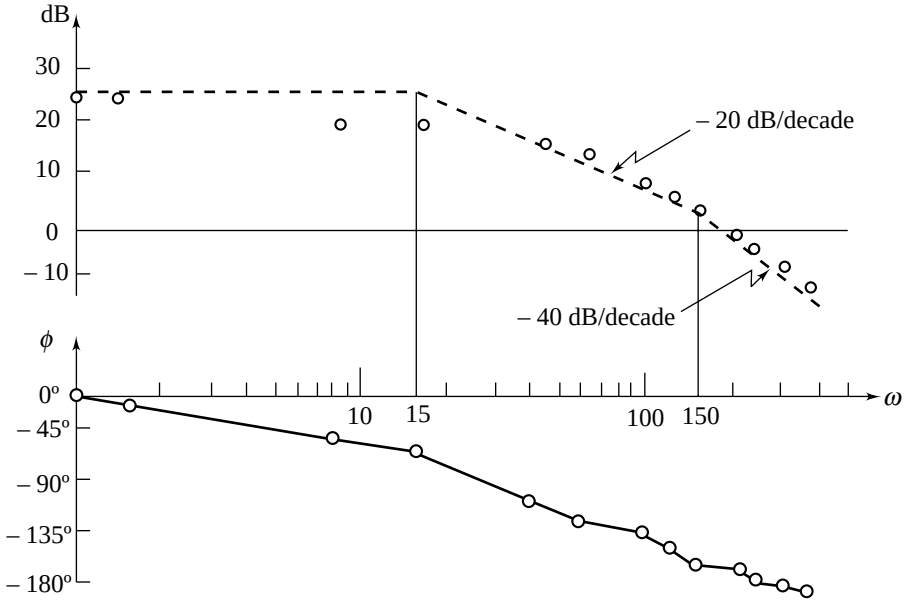


Fig. 8.44 Experimental data in Bode coordinate system

At $\omega = 15$, the slope becomes –20 dB/decade. Thus there is a pole with corner frequency 15. At $\omega = 150$, the slope becomes –40 dB/decade, a decrease of 20 dB/decade. Therefore, there is another pole with corner frequency 150.

The transfer function must be of the form: $G(s) = \dfrac{17.8}{(1 + s/15)(1 + s/150)}$.

The phase characteristics calculated from this transfer function are in fair agreement with experimentally obtained characteristics shown in Fig. 8.44. Dead-time element is therefore, not present.

Example 8.18 Let us find the transfer function of the experimentally obtained Bode plot shown in Fig. 8.45. First we approximate the magnitude plot by the three dashed lines shown. They intersect at

$\omega = 1$ and $\omega = 10$. In the low-frequency range, there is an asymptote with slope -20 dB/decade; it indicates the presence of a factor of the form $K/(j\omega)$ in the transfer function. At $\omega = 1$, the magnitude of the asymptote is 40 dB. Therefore

$$20 \log K = 40$$

This gives $K = 100$

At $\omega = 1$, the slope becomes -40 dB/decade, a decrease of 20 dB/decade; thus there is a pole with corner frequency $\omega = 1$. At $\omega = 10$, the slope becomes -20 dB/decade, an increase of 20 dB/decade; there is a zero with corner frequency $\omega = 10$. The transfer function must be of the form

$$G(s) = \frac{100\left(1 \pm \frac{s}{10}\right)}{s\,(1+s)}$$

We have assumed that the Bode plot of Fig. 8.45 is obtained by frequency response measurements; the pole in the right half s-plane is, therefore, not possible. However, a zero in the right half s-plane may exist, i.e., the system under experimental analysis may be a nonminimum-phase system.

Now we use the phase plot to determine the sign of the zero factor. If the sign of zero $(1 \pm 0.1s)$ is positive, the zero will introduce positive phase into $G(s)$ or, equivalently, the phase of $G(s)$ will increase as ω passes through the corner frequency at 10. This is not the case as seen from Fig. 8.45; thus we have $(1 - 0.1s)$, and the transfer function of the Bode plot is

$$G(s) = \frac{100\,(1 - 0.1s)}{s\,(1+s)}$$

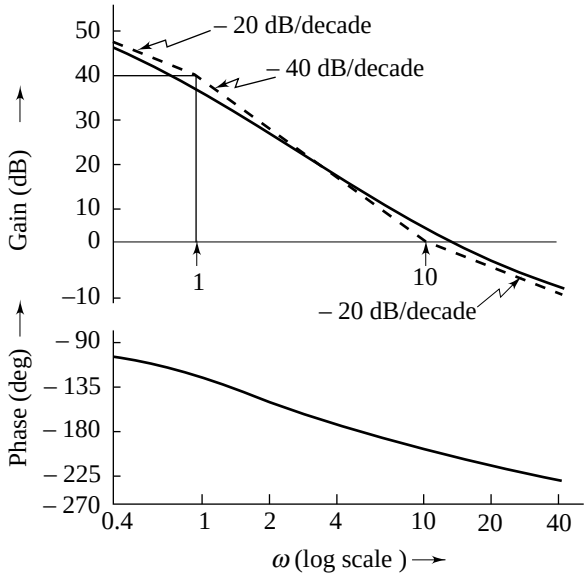


Fig. 8.45 Experimentally obtained Bode plot

The phase characteristics calculated from this transfer function are in fair agreement with experimentally obtained characteristics shown in Fig. 8.45.

Review Examples

Review Example 8.1 A heuristic derivation of a simple equation based on the *principle of argument* is as follows [64].
Consider the function

$$Q(s) = \frac{s - 0.5}{s(s-1)(s+4)} \tag{8.56}$$

Suppose $Q(s)$ is evaluated along the simple, circular contour Γ_s of radius 2.0 in the s-plane (shown in Fig. 8.46 a). We start the evaluation from the point $s = 2\angle 0°$ on the contour Γ_s. From Eqn. (8.56), we get

$$Q(s)|_{s=2\angle 0°} = \frac{|s-0.5|}{|s||s-1||s+4|}\angle[\angle(s-0.5) - \angle s - \angle(s-1) - \angle(s+4)]|_{s=2\angle 0°}$$

$$= \frac{1.5}{2\times 1\times 6}\angle 0° = 0.125\angle 0°$$

From Fig. 8.46a, we observe that the vectors $(s - 0.5)$, s, $(s - 1)$, and $(s + 4)$ have magnitudes 1.5, 2, 1, and 6, respectively, at $s = 2\angle 0°$. Each vector contributes an angle of 0° to $Q(s)$ at $s = 2\angle 0°$.

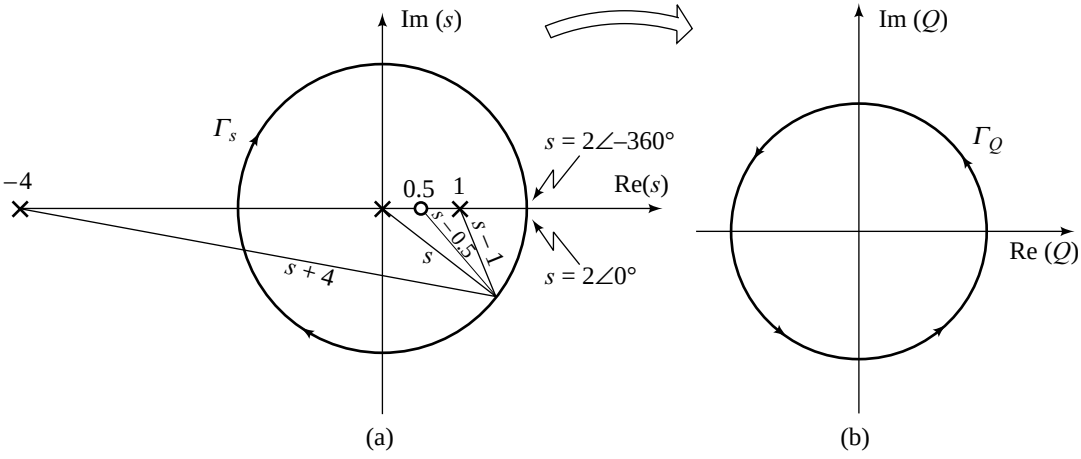


Fig. 8.46(a) Contour Γ_s in the s-plane, and (b) resulting contour Γ_Q in the Q(s)-plane

Contour Γ_s in the s-plane is a closed contour. Let us traverse the contour in the clockwise direction. A point or an area is said to be *enclosed* by a closed path if it is found to lie to the right of the path when the path is traversed in the clockwise direction. The closed path given by contour Γ_s in Fig. 8.46a encloses the circular region in the s-plane.

We started evaluation of $Q(s)$ from the point $s = 2\angle 0°$. Now we see that as the point s follows the prescribed path (i.e., clockwise direction) on the s-plane contour Γ_s, evaluation of $Q(s)$ at each point on Γ_s generates the closed contour Γ_Q in the $Q(s)$-plane, shown in Fig. 8.46b. The contour Γ_Q will be smooth because $Q(s)$ is a simple ratio of polynomials in the complex variable s that is well defined along Γ_s. It will become clear as the discussion proceeds that it is only the general shape of Γ_Q that is important.

Therefore, *exact* evaluation of $Q(s)$ along all the points of the contour Γ_s is not required. We evaluate $Q(s)$ at some key points along the contour Γ_s; the closed contour Γ_Q could then be approximated by simply plotting and connecting these values of $Q(s)$ (Γ_Q shown in Fig. 8.46b is not drawn to scale). Several features of Fig. 8.46 deserve comment.

1. As the evaluation of $Q(s)$ moves from $s = 2\angle 0°$ to $s = 2\angle -360°$ in the clockwise direction along the Γ_s contour, each of the vectors *inside* Γ_s rotates clockwise through 360°.
2. During this evaluation,
   (i) the vectors s and $(s - 1)$ in the denominator of $Q(s)$ will each contribute $-(-360°)$ of phase to $Q(s)$;
   (ii) the vector $(s - 0.5)$ in the numerator of $Q(s)$ will contribute $-360°$ of phase of $Q(s)$; and
   (iii) the vector $(s + 4)$ in the denominator of $Q(s)$ will contribute *no* net phase to $Q(s)$.
3. The net change in the phase of $Q(s)$ evaluated along Γ_s, traversing it once in the clockwise direction is 360°; or one *encirclement* of the origin of the $Q(s)$-plane in counterclockwise direction (refer Fig. 8.46 b).

We thus see that a change in phase of $Q(s)$ that one gets in traversing a contour Γ_s in the s-plane, is strictly a function of how many poles and zeros of $Q(s)$ are enclosed by the s-plane contour. Let us consider a general $Q(s)$ with m zeros and n poles. Assume that the s-plane contour Γ_s encloses Z zeros and P poles of $Q(s)$; and it does not go through any of the poles or zeros of $Q(s)$ (the poles of $Q(s)$ in s-plane are mapped to ∞ in $Q(s)$-plane, and the zeros are mapped to the origin). If we make one clockwise traversal of the contour Γ_s, it may be seen that:

$$\begin{aligned} \textit{net} \text{ change in phase of } Q(s) &= Z(-2\pi) - P(-2\pi) \\ &= (P - Z)2\pi \end{aligned}$$

A change in phase of $Q(s)$ by 2π leads to one counterclockwise encirclement of the origin in the $Q(s)$-plane by the Γ_Q contour. Therefore, net change in phase of $Q(s)$ by $(P - Z)\,2\pi$ as we traverse the contour Γ_s once in clockwise direction, will lead to $(P - Z)$ encirclements of the origin of the $Q(s)$-plane in counterclockwise direction by the Γ_Q contour.

In the case of simple function being considered here (Eqn. (8.56)), $P = 2$, $Z = 1$; and there is one encirclement of the origin by Γ_Q contour in counterclockwise direction.

Having made this qualitative analysis, it is possible to formulate a simple equation based on the contours Γ_s and Γ_Q, that can be applied to any function $Q(s)$ expressible as a ratio of polynomials in s with real coefficients. Let

N = number of *clockwise* encirclements of the origin of the $Q(s)$-plane by the contour Γ_Q;

P = number of poles of $Q(s)$ enclosed by the Γ_s contour in the s-plane, when traversed *clockwise*; and

Z = number of zeros of $Q(s)$ enclosed by the Γ_s contour in the s-plane, when traversed *clockwise*.

Then

$$N = Z - P \tag{8.57}$$

The relation (8.57) between the enclosure of poles and zeros of $Q(s)$ by the s-plane contour and the encirclement of the origin by the $Q(s)$-plane contour is commonly known as the *principle of argument*.

Equation (8.57) appears to be rather trivial. However, Nyquist managed to turn it into a very powerful tool for analyzing the stability of closed-loop control systems by carefully choosing the contour Γ_s and the function $Q(s)$.

Review Example 8.2 Consider a feedback system with an open-loop transfer function

$$G(s)H(s) = \frac{1 + 4s}{s^2(1 + s)(1 + 2s)}$$

which has a double pole at the origin. The indented Nyquist contour shown in Fig. 8.47a does not enclose these poles. The Nyquist plot shown in Fig. 8.47b is obtained as follows.

The portion of the $G(s)H(s)$-locus from $\omega = 0^+$ to $\omega = +\infty$ is simply the polar plot of

$$G(j\omega)H(j\omega) = \frac{1 + j4\omega}{(j\omega)^2 (1 + j\omega)(1 + j2\omega)}$$

The following points were used to construct the polar plot:

(i) $G(j\omega)H(j\omega)|_{\omega \to 0^+} \to \infty \angle -180°$

(ii) The $G(j\omega)H(j\omega)$-locus intersects the real axis at a point where

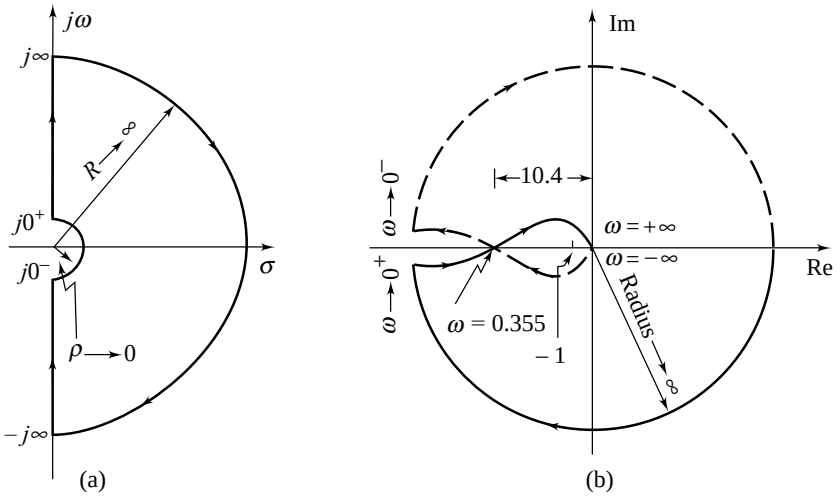


Fig. 8.47 Nyquist contour and the corresponding mapping[7] for $G(s)H(s) = (4s + 1)/[s^2(s + 1)(2s + 1)]$

$$\angle G(j\omega)H(j\omega) = -180°$$

or $$-180° - \tan^{-1}\omega - \tan^{-1}2\omega + \tan^{-1}4\omega = -180°$$

or $$\tan^{-1}4\omega - \tan^{-1}\omega = \tan^{-1}2\omega$$

Therefore $$\tan(\tan^{-1}4\omega - \tan^{-1}\omega) = \tan(\tan^{-1}2\omega)$$

or $$\frac{4\omega - \omega}{1 + 4\omega^2} = 2\omega$$

This gives $$\omega = \frac{1}{2\sqrt{2}} = 0.355 \text{ rad/sec}$$

$$|G(j\omega)H(j\omega)|_{\omega = 0.355} = 10.4$$

(iii) $G(j\omega)H(j\omega)\big|_{\omega \to +\infty} \to 0 \angle -270°$

7

ω	0.01	0.1	0.35	0.36
$\lvert GH\rvert$	4×10^6	104.09	11.4	10.3
$\angle GH$	− 179.97	− 175.2	− 179.35	179.64

The plot of $G(j\omega)H(j\omega)$ for $\omega < 0$ is the reflection with respect to the real axis, of the plot for $\omega > 0$. Every point on the large semicircle in Fig. 8.47a, is mapped by $G(s)H(s)$ into the origin of the $G(s)H(s)$-plane. A point $s = \rho e^{j\phi}$; $\rho \to 0$, ϕ varying from $-90°$, through $0°$ to $+90°$, is mapped by $G(s)H(s)$ into

$$G(s)H(s)\Big|_{s=\rho e^{j\phi}} = \frac{1+4s}{s^2(1+s)(1+2s)}\Big|_{s=\rho e^{j\phi}} = \frac{1}{\rho^2 e^{j2\phi}} = \frac{1}{\rho^2}\angle -2\phi$$

The small semicircle in Fig. 8.47a is mapped by $G(s)H(s)$ into a circular arc of infinite radius ranging from $180°$ at $\omega = 0^-$ through $0°$ to $-180°$ at $\omega = 0^+$ ($180° \to 90° \to 0° \to -90° \to -180°$).

Figure 8.47b indicates that the critical point $-1 + j0$ is encircled by the Nyquist plot twice in the clockwise direction. Therefore $N = 2$. From the given transfer function it is seen that no pole of $G(s)H(s)$ lies in the right half s-plane, i.e., $P = 0$. Therefore (refer Eqn. (8.8a))

Z = number of zeros of $1 + G(s)H(s)$ enclosed by the Nyquist contour
$= N + P = 2$

Hence the feedback system is unstable with two poles in the right half s-plane.

Review Example 8.3 The frequency response test data for the forward-path elements of a unity-feedback control system are given in Table 8.8. These data are sufficient to analyse feedback system stability, as is seen below.

Figure 8.48 shows the frequency response data of Table 8.8 on Bode magnitude and phase plots. From these plots we find that

phase crossover frequency $\omega_\phi = 48$ rad/sec,
gain crossover frequency $\omega_g = 3$ rad/sec,
gain margin $GM = 28$ dB, and
phase margin $\Phi M = 49°$.

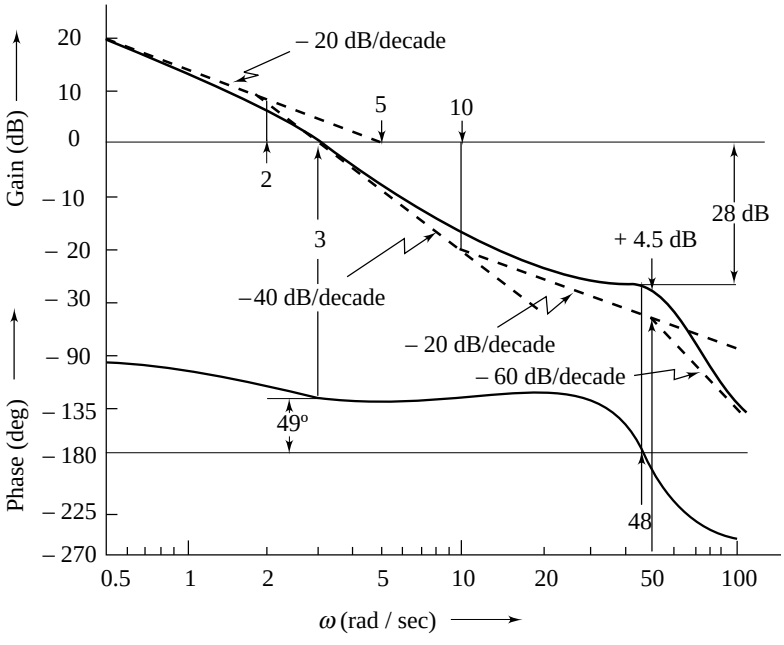


Fig. 8.48 Experimentally obtained magnitude and phase characteristics

Table 8.8 Experimental frequency response data

ω(rad/sec)	0.5	1	2	4	10	20	30
Gain (dB)	19.75	13	5	– 4	– 17	– 24	– 26
Phase (deg)	– 101	– 111	– 125	– 134	– 130	– 127	– 134
ω(rad/sec)	40	50	60	70	80	90	100
Gain (dB)	– 27	– 29	– 34	– 39	– 43	– 47	– 50
Phase (deg)	– 154	– 189	– 218	– 235	– 243	– 249	– 252

The feedback system is therefore stable.

It may be desirable to obtain a transfer function that approximates the experimental amplitude and phase characteristics. A series of straight-line asymptotes fitted to the experimentally obtained magnitude curve are shown in Fig. 8.48. The low frequency asymptote has a slope of –20 dB/decade and when extended, intersects the 0-dB axis at $\omega = 5$. Therefore, the asymptote is a plot of the factor $5/(j\omega)$. The corner frequencies are found to be located at $\omega = 2$, $\omega = 10$ and $\omega = 50$. At the first corner frequency, the slope of the curve changes by – 20 dB/decade and at the second corner frequency, it changes by + 20 dB/decade. Therefore, the transfer function has factors $1/(1 + j\omega/2)$ and $(1 + j\omega/10)$ corresponding to these corner frequencies. At $\omega = 50$, the curve changes by a slope of – 40 dB/decade. At this frequency, the error between actual and approximate plots is + 4.5 dB. The change of slope by – 40 dB/decade indicates the presence of either double pole on real axis or a pair of complex conjugate poles. The error of + 4.5 dB at the corner frequency and the peak occurring at a frequency less than the corner frequency indicate the presence of a quadratic factor with $\zeta < 0.5$. In fact, we can calculate the approximate value of ζ using expression (8.42):

$$-20 \log 2\zeta = 4.5$$

This gives $$\zeta = 0.3$$

Therefore, the transfer function has a quadratic factor

$$\frac{1}{1 + j2\zeta\,(\omega/50) + (j\omega/50)^2} \quad \text{where } \zeta = 0.3$$

Thus an approximate transfer function of the forward-path elements of the feedback system becomes

$$G(j\omega) = \frac{5\,(1 \pm j\omega/10)}{j\omega\,(1 + j\omega/2)\left[1 + j\,0.6\,(\omega/50) + (j\omega/50)^2\right]}$$

Now we use the phase plot to determine the sign of the zero factor. The phase characteristics calculated from the transfer function are in fair agreement with experimentally obtained characteristics shown in Fig. 8.48, when the sign of the zero factor is positive. Therefore, the transfer function of the Bode plot of Fig. 8.48 is

$$G(s) = \frac{5\,(1 + s/10)}{s\,(1 + s/2)\left[1 + 0.6s/50 + (s/50)^2\right]}$$

Review Questions

8.1 (a) Using the principle of argument, derive the Nyquist stability criterion.
(b) What is the interpretation of stability analysis when Nyquist plot crosses the $-1 + j0$ point?

8.2 (a) Give an account of meaning of the terms 'gain margin' and 'phase margin' with reference to Nyquist plots.

(b) How can a frequency response be represented by a Bode plot? Indicate what gain and phase margins are in this context.

(c) Why do we use logarithmic scale for frequency in Bode plots?

8.3 Give the properties of minimum-phase and nonminimum-phase transfer functions.
Describe a method of identification of transfer function model of a system using frequency-response measurements.

8.4 Give examples of Nyquist plots with

(i) poor ΦM, infinite GM (ii) poor GM, infinite ΦM

Problems

8.1 Sketch the general shapes of the polar plots for

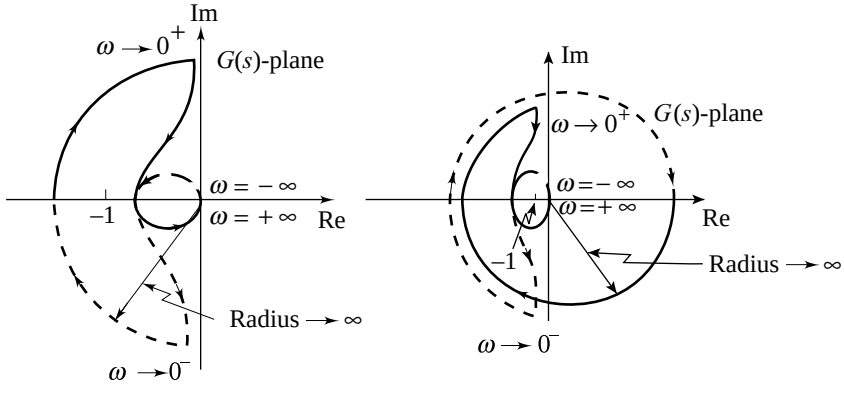
(a) $G(j\omega) = \dfrac{1}{(1 + j\omega\tau_1)(1 + j\omega\tau_2)}$ (b) $G(j\omega) = \dfrac{1}{j\omega(1 + j\omega\tau)}$

(c) $G(j\omega) = \dfrac{1}{(j\omega)^2(1 + j\omega\tau)}$ (d) $G(j\omega) = \dfrac{1}{j\omega(1 + j\omega\tau_1)(1 + j\omega\tau_2)}$

(e) $G(j\omega) = \dfrac{1}{(j\omega)^2(1 + j\omega\tau_1)(1 + j\omega\tau_2)}$

based on magnitude and phase calculations at (i) $\omega = 0$, (ii) $\omega = \infty$, (iii) the point of intersection (if any) with the real axis, and (iv) the point of intersection (if any) with the imaginary axis.

8.2 Check the stability of unity-feedback systems whose Nyquist plots are shown in Fig. P8.2. $G(s)$ represents open-loop transfer function.



Number of poles of $G(s)$ in right half s-plane = 1

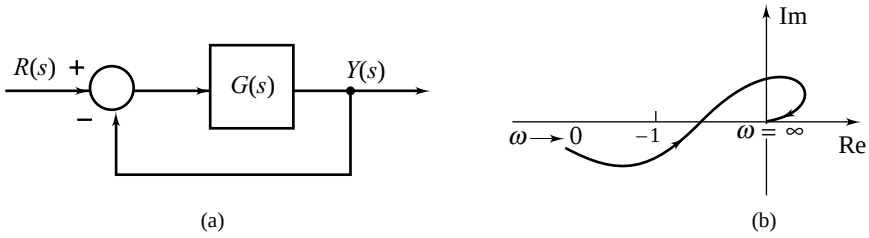
(a)

Number of poles of $G(s)$ in right half s-plane = 0

(b)

Fig. P8.2

8.3 Consider the unity-feedback system shown in Fig. P8.3a. The polar plot of $G(s)$ is of the form shown in Fig. P8.3b. Assuming that the Nyquist contour in s-plane encloses the entire right half s-plane, draw a complete Nyquist plot in the $G(s)$-plane. Then answer the following questions.



(a) (b)

Fig. P8.3

(i) If $G(s)$ has no poles and no zeros in the right half s-plane, is the closed-loop system stable?
(ii) If $G(s)$ has one pole and no zeros in the right half s-plane, is the closed-loop system stable?
(iii) If $G(s)$ has one zero and no poles in the right half s-plane, is the closed-loop system stable?

8.4 By use of the Nyquist criterion, determine whether the closed-loop systems having the following open-loop transfer functions are stable or not. If not, how many closed-loop poles lie in the right half s-plane?

(a) $G(s)H(s) = \dfrac{180}{(s+1)(s+2)(s+5)}$ (b) $G(s)H(s) = \dfrac{2}{s(s+1)(2s+1)}$

(c) $G(s)H(s) = \dfrac{s+2}{(s+1)(s-1)}$ (d) $G(s)H(s) = \dfrac{s+2}{s^2}$

(e) $G(s)H(s) = \dfrac{1}{s^2+100}$ (f) $G(s)H(s) = \dfrac{2(s+3)}{s(s-1)}$

MATLAB Exercise After completing the hand sketches, verify your results using MATLAB. Turn in your hand sketches and MATLAB-based plots on the same scale.

8.5 The stability of a closed-loop system with open-loop transfer function

$$G(s)H(s) = \frac{K(\tau_2 s+1)}{s^2(\tau_1 s+1)};\ K,\ \tau_1,\ \tau_2 > 0$$

depends on the relative magnitudes of τ_1 and τ_2. Draw Nyquist plots and therefrom determine stability of the system when (i) $\tau_1 < \tau_2$, and (ii) $\tau_1 > \tau_2$.

8.6 Use the Nyquist criterion to determine the range of values of $K > 0$ for the stability of the system in Fig. P8.6 with

(a) $G(s) = \dfrac{8}{(s+1)(s^2+2s+2)}$

(b) $G(s) = \dfrac{4(1+s)}{s^2(1+0.1s)}$

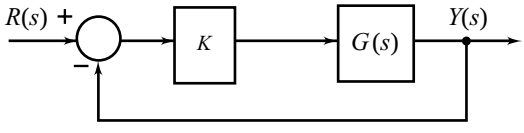


Fig. P8.6

(c) $G(s) = \dfrac{4(1+0.1s)}{s^2(1+s)}$

(d) $G(s) = \dfrac{e^{-0.8s}}{s+1}$

MATLAB Exercise (i)After completing the hand sketches, verify your results using MATLAB. Turn in your hand sketches and MATLAB-based plots on the same scale.(ii) Also verify your results using MATLAB-based root locus plots. Make suitable approximation for the deadtime in Problem 8.3d.

8.7 Sketch the Nyquist plot for a feedback system with open-loop transfer function

$$G(s)H(s) = \frac{K(s+3)(s+5)}{(s-2)(s-4)};\ K>0$$

Find the range of values of K for which the system is stable.

8.8 Sketch the Nyquist plot for a feedback system with open-loop transfer function

$$G(s)H(s) = \frac{K(1+0.5s)(s+1)}{(1+10s)(s-1)};\ K>0$$

Find the range of values of K for which the system is stable.

8.9 Consider the control system shown in Fig. P8.9. Using the Nyquist criterion, determine the range of gain $K > 0$ for stability of the system.

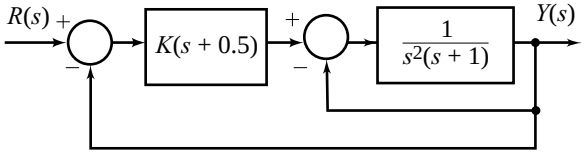


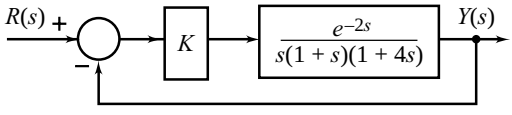
Fig. P8.9

MATLAB Exercise Make a rough sketch to decide on appropriate frequency range. Then generate a MATLAB-based plot over this frequency range. Also verify your result using MATLAB-based root locus plot.

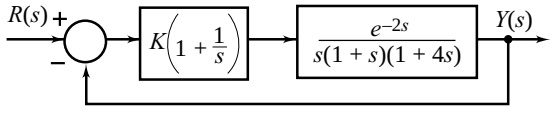
8.10 Consider the feedback systems shown in Figs P8.10a and P8.10b. Sketch the Nyquist plot in each case and therefrom determine the maximum value of K for stability.

8.11 A unity-feedback system has open-loop transfer function

$$G(s) = \frac{e^{-s\tau_D}}{s(s+1)}$$



(a)



(b)

Fig. P8.10

Draw the Nyquist plot of $G_1(j\omega) = \dfrac{1}{j\omega\,(j\omega + 1)}$, together with the critical trajectory of $e^{-j\omega\tau_D}$. Using these curves, obtain the maximum value of deadtime τ_D in seconds for the closed-loop system to be stable.

8.12 Using the Nyquist plot, determine gain crossover frequency, phase crossover frequency, gain margin, and phase margin of feedback system with open-loop transfer function

$$G(s)H(s) = \frac{10}{s\,(1 + 0.2s)\,(1 + 0.02s)}$$

MATLAB Exercise Using MATLAB dialogues, sketch the Nyquist plot with selection of appropriate frequency range from the Bode plot of the system. From the Nyquist plot, determine the required performance indices. Also determine the indices using Bode plot.

8.13 Sketch the Nyquist plot for a feedback system with open-loop transfer function

$$G(s)H(s) = \frac{K\,(s + 10)^2}{s^3};\ K > 0$$

(a) Show that the feedback system is stable for $K > 5$.
(b) Determine the gain margin and phase margin when $K = 7$.

8.14 Sketch the Bode asymptotic plots showing the magnitude in dB as a function of log frequency for the transfer functions given below. Determine the gain crossover frequency in each case.

(a) $G(s) = \dfrac{25}{(s + 1)\,(0.1s + 1)\,(0.05s + 1)}$

(b) $G(s) = \dfrac{50\,(0.2s + 1)}{s\,(s + 1)\,(0.02s + 1)}$

(c) $G(s) = \dfrac{500\,(0.2s + 1)\,(0.1s + 1)}{s^2\,(s + 1)\,(0.02s + 1)}$

(d) $G(s) = \dfrac{50\,(0.05s + 1)}{s\,(0.1s + 1)\,(0.02s + 1)\left((s/200)^2 + (0.02s/200) + 1\right)}$

8.15 Using Bode plot, determine gain crossover frequency, phase crossover frequency, gain margin, and phase margin of a feedback system with open-loop transfer function

(a) $G(s) = \dfrac{10}{s\,(0.1s + 1)}$

(b) $G(s) = \dfrac{10}{s\,(0.1s + 1)^2}$

(c) $G(s) = \dfrac{20\,(0.2s + 1)}{s\,(0.5s + 1)}$

(d) $G(s) = \dfrac{20\,(0.2s + 1)\,e^{-0.1s}}{s\,(0.5s + 1)}$

(e) $G(s) = \dfrac{40}{(s + 2)\,(s + 4)\,(s + 5)}$

(f) $G(s) = \dfrac{10}{s^2\,(0.2s + 1)}$

MATLAB Exercise Generate Bode plots in MATLAB window and determine the performance indices.

8.16 Use Bode plots to determine the stability of the system shown in Fig. P8.16 for the two cases: (i) $K = 10$, and (ii) $K = 100$.

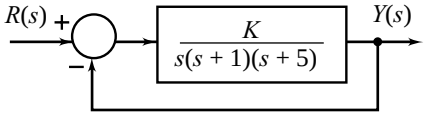


Fig. P8.16

MATLAB Exercise Determine the stability using both the Bode plots and the root locus plots, generated using MATLAB.

8.17 Use Bode plot to determine the range of K within which a unity-feedback system with open-loop transfer function $G(s)$ is stable. Given:

(a) $G(s) = \dfrac{K}{(s+2)(s+4)(s+5)}$ (b) $G(s) = \dfrac{K}{s(1+0.2s)(1+0.02s)}$

(c) $G(s) = \dfrac{Ke^{-s}}{s(s+1)(s+2)}$

MATLAB Exercise Determine the range using both the Bode plots and the root locus plots, generated using MATLAB. Note that the transfer function for Problem 8.9c has deadtime; it may be approximated by a rational function for generating root locus plot.

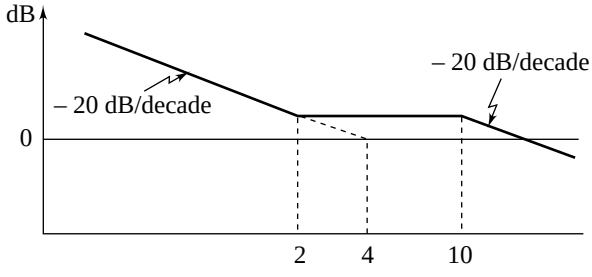
8.18 Open-loop transfer function of a closed-loop system is

$$G(s)H(s) = \frac{10e^{-s\tau_D}}{s(0.1s+1)(0.05s+1)}$$

(a) Find the gain margin and phase margin when $\tau_D = 0$.
(b) Find the gain margin and phase margin when $\tau_D = 0.04$ sec. Comment upon the effect of dead-time.
(c) Determine the maximum value of τ_D for the closed-loop system to be stable.

MATLAB Exercise Use MATLAB environment in dialogue mode to solve this problem.

8.19 The experimental frequency response data of certain systems presented on Bode plots and asymptotically approximated are shown in Fig. P8.19. Find the transfer function in each case (systems are known to have minimum-phase characteristics).



(a)

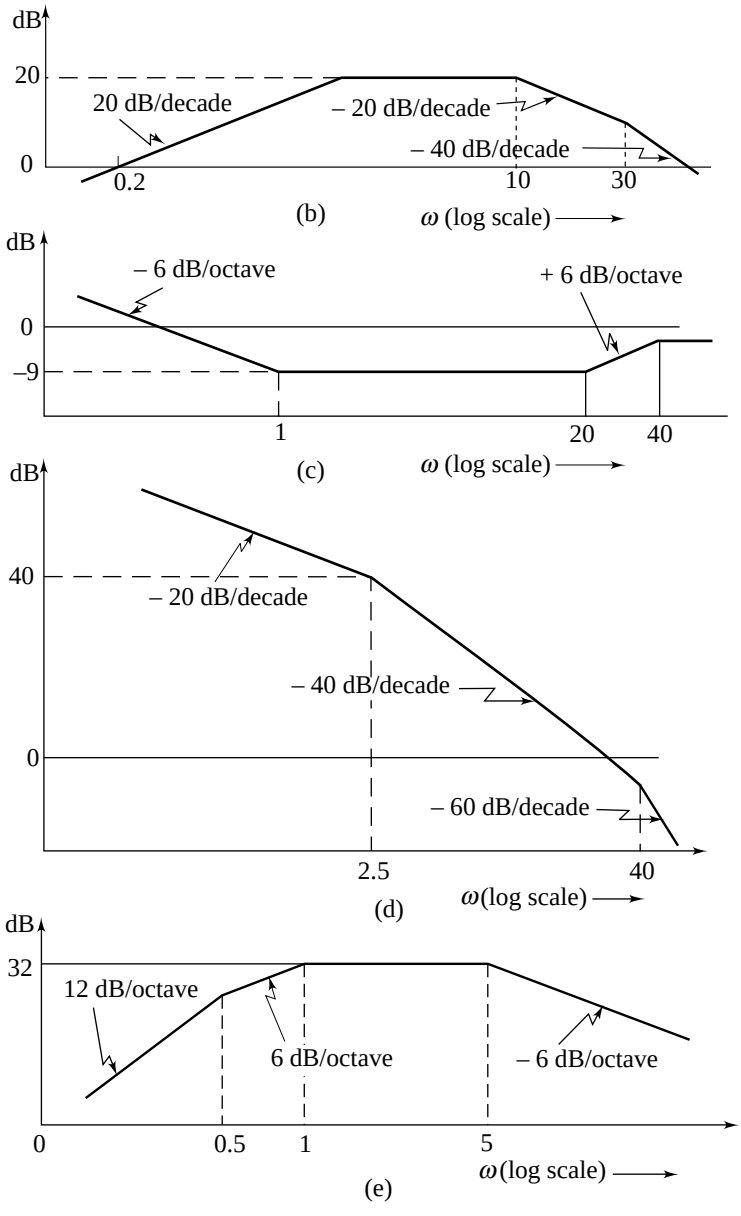


Fig. P8.19

8.20 Consider a minimum-phase system whose asymptotic amplitude frequency response is depicted in Fig. P8.20.

(a) Determine the transfer function $G(s)$ of the system.

(b) Determine the two gain crossover frequencies ω_{g1} and ω_{g2}.

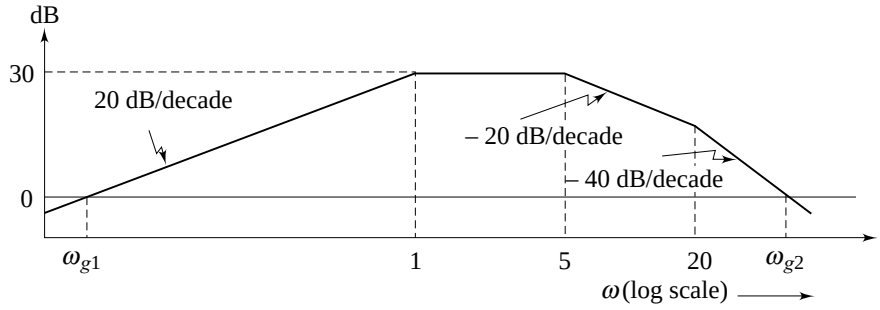


Fig. P8.20

8.21 The asymptotic amplitude frequency response of the open path of a feedback system is shown in Fig. P8.21. Determine the gain margin of the system. The system is known to have minimum-phase characteristics.

8.22 The following frequency response test data were obtained for a system known to have minimum-phase characteristics. Plot the data on semilog graph paper and determine the transfer function of the system using asymptotic approximation.

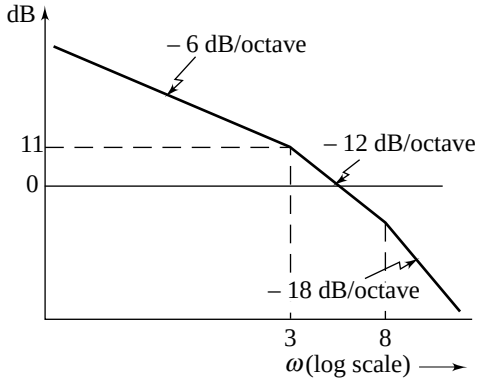


Fig. P8.21

Gain (dB)	34	28	24.6	14.2	8	1.5	– 3.5	– 7.2
Frequency (rad/sec)	0.1	0.2	0.3	0.7	1.0	1.5	2.0	2.5
Gain (dB)	– 12.5	– 14.7	– 16.0	– 17.5	– 17.5	– 17.5		
Frequency (rad/sec)	4.0	5.0	6.0	9.0	20	35		

8.23 Determine an approximate transfer function model of a system for which experimental frequency response data is given in the following table.

ω	*dB*	*Phase*
0.1	– 20	0
0.5	– 21	0
1.0	– 21	– 9°
2.0	– 22	– 54°
3.0	– 24	– 90°
5.0	– 28	– 135°
10.0	– 40	– 170°
30.0	– 60	– 178°
100.0	– 84	– 180°