Lecture 6: Gain, Phase margins, designing with Bode plots, Compensators

Venkata Sonti

Department of Mechanical Engineering
Indian Institute of Science
Bangalore, India, 560012
This draft: March 12, 2008

1 Gain and Phase Margins

We know that the characteristic equation is
\[ 1 + KGH = 0 \]
written for a particular gain. Then
\[ KGH = -1 \]
so that a) when the amplitude of KGH is 1 or 0 dB and b) phase is 180° then we have the system just become unstable. Both the conditions have to be fulfilled. Hence, if we ask the question: How far is the system from instability? We get two answers. The gain increment required at the frequency where phase is 180° and the phase increment required at the frequency where amplitude is 0 dB. The frequency where phase crosses 180° is called the phase crossover and the frequency where the amplitude crosses the 0dB point is called the gain crossover.

Let us look at the nyquist plot for
\[ GH = \frac{K}{s^3 + 9s^2 + 30s + 40} \]
Figure 1a gives the bode plot for K=50 and 1b gives the corresponding nyquist plot for K=50 and 100. The point in the amplitude bode plot were gain crosses 0dB is at frequency 1.84 rad/s. At this point the phase is -80 degrees. In order to reach 180° I need a phase change of 100 degrees which is the Phase Margin.

The phase becomes -180° degrees at 5.47 rad/s and here the amplitude is less than 1 or negative in dB. I have to raise the gain by 13.24 dB before the system becomes unstable. Hence, that is my Gain Margin.

The corresponding nyquist diagram is shown in figure 1b. The first curve is for K=50 and we can see where the gain becomes 1 or 0dB marked by 1.84 rad/s and where the phase goes to 180° marked by 5.47 rad/s. The long dotted arrow marking the gain crossover at 1.84 rad/s is about 100 degrees behind the negative real axis, i.e., we need to rotate the long arrow by 100 degrees before it coincides with the negative real axis. The second curve is for K=100 and we can see
that the phase =180 point moves leftward towards the \((-1, j0)\) point. Thus, if the gain is further raised, eventually the phase crossover point will intersect the gain crossover point and the system will become unstable.

\[
G_m=13.255 \text{ dB (at 5.4772 rad/sec)}, \quad P_m=100.66 \text{ deg. (at 1.8484 rad/sec)}
\]

Fig. 1: Nyquist and bode plot for \(K/s^3 + 9s^2 + 30s + 40\)

Gain margin is a measure of closeness of phase crossover point to \((-1, j0)\) in \(GH\) plane. At phase crossover \(\omega_c\), \(GH = G(j\omega_c)H(j\omega_c)\).

\[
\text{Gain margin} = 20\log_{10}\frac{1}{|G(j\omega_c)H(j\omega_c)|}
\]

is Usually positive. If gain is increased and \(|GH|\) goes through \((-1, j0)\), the system has gone unstable. If \(GH\) never cuts negative axis the system is forever stable. Gain margin is the amount of gain in \(dB\) that can be allowed to increase in the loop before closed loop system reaches instability.

If \((-1, j0)\) is to the right of phase crossover point \(|GH| > 1\). This means unstable. However first determine if the system is stable or not and then magnitude of stability. Only sign is not enough.

Phase Margin is how much to rotate the gain crossover point so that you go through \((-1, j0)\). Phase margin is defined as the angle in degrees through which the \(G(\omega)H(\omega)\) plot must be rotated about the origin in order that gain crossover point on locus passes through \((-1, j0)\) point.

\[
\phi_M = \angle GH - 180^0.
\]

2 Control design using Bode plots

We keep the following empirical rules in mind while designing control systems using the bode plot. These are rules strictly valid for second order systems. But we can use them for other systems, with caution.
• Bandwidth frequency is defined as the frequency at which closed loop magnitude response is \(-3\) dB from 0 frequency response.

• In designing using bode plots, we look at the frequency where the magnitude is \(-6\) to \(-7.5\) dB.

• Bandwidth\(=\omega_n\) for 2\textsuperscript{nd} order system.

• open \(GH\) loop must be stable for designing via bode plots.

• If for \(GH\), gain crossover \(<\) phase crossover for open loop, closed loop will be stable.

• Closed loop damping ratio is \(\approx \frac{PM \text{ of } GH}{100}\).

• Bandwidth \(-6\) to \(-7.5\) dB of \(GH=\omega_n\) of second order.

2.1 Example 1

Consider the following open loop system with \(GH\) given by

\[ GH = \frac{8}{s + 0.8} \]

We need to design a controller for a step response with the specs

• zero steady state error

• max overshoot (\(Mp\)) less than 40 \%

• settling time less than 2 sec

Lets look at the bode diagram 2. Last lecture we found that \(\lim_{s \to 0} GH = k_p\). From the bode diagram this is just the low frequency asymptote of the amplitude plot. The value is 10 in raw units or 20 dB. We also know that the steady state error for unit step input is \(1/(1+kp) = 0.091\) for this case. The step response shows this value of steady state error. The bandwidth frequency\(=\omega_n\) we said was between -6 to -7.5 dB in bode plots. For this system \(\omega_n\) = is about 10 rad/s. We can now find rise time using \(\omega_n t_r = 1.8\). \(t_r\) is 0.18, which seems alright (remember these are empirical relations for a second order system). The gain crossover occurs at 7.77 rad/s and the phase at this point is -85 degrees. Hence, the phase margin is 95 degrees. From the relation \(\xi = \frac{PM}{100}\) we get \(\xi = 0.95\). This system is very heavily damped which can be seen from the lack of overshoot.

Lets use the formulas we have to find some relevant values. We have, \(Mp\) the overshoot

\[ Mp = e^{-\pi\xi/\sqrt{1-\xi^2}} \]

From here we get

\[ \xi^2 = \frac{ln^2(Mp)}{\pi^2 + ln^2(Mp)}. \]

Using \(Mp=0.4\) (this is the spec above), \(\xi = 0.28\) or we need about a 30 degree phase margin. Since the current phase margin is 95 degrees we need to drop the phase. If we use just an integral controller to remove the steady state error, this gives a denominator zero which will drop the phase by 90 degrees. Lets try this. We add \(Ki/s\) with \(Ki=1\). The PM in figure 3 has dropped to
Figure 2: Bode plot and step response for $8/s + 0.8$. open loop step response.

about 16 degrees. This is too low. We can see with this example why an integral controller will give an overshoot and take the system toward instability. The nyquist plot has been effectively rotated clockwise. Note: Since the root locus begins on the poles of $GH$ and ends on the zeros, by placing a pole right at $s=0$, although the zero does not add anything, the number of poles ($n$) has gone up so that the centre of asymptotes moves to the right making the system relatively more unstable.

Figure 3: Bode plot after integral controller $Ki/s$ for $8/s + 0.8$
We need to add more PM. We do this with a finite zero of the form \( s + 20 \). Now if we put this together the total controller is

\[
\frac{s + 20}{s} = kp + \frac{ki}{s}
\]

which is the PI controller. Let's look at the bode and step response plot figure 4. The response

Figure 4: Bode plot and step response after \((s + 20)/s\) for \(8/s + 0.8\). Closed loop step response satisfies the specs quite well. The overshoot is less than 40%. There is no steady state error. The settling time is about 1 sec.

In the above example, we can understand the effect of adding a zero to \( GH \). A zero behaves as a stabilizer. It adds phase margin. Let's examine the effect of \( K \) on

\[
\frac{s + K}{s} \frac{10}{1.25s + 1}
\]

A small value of \( K \) like 0.1 adds a good deal of phase early and hence can add a large phase margin. This can be seen by bode plots of just \( s + K \). We can even see the effect of \( K \) on the root locus diagram

\[
1 + KG \ast H^* = 1 + \frac{8K}{s^2 + 8.8s}
\]

We can also do

\[
\frac{Ks + 1}{s} \frac{10}{1.25s + 1}
\]

and hence look at the root locus of \( K \) here using

\[
1 + KG \ast H^* = 1 + \frac{8KS}{s^2 + 0.8s + 8}
\]

Finally, for a given zero position such as 1.0 we can look at

\[
1 + KG \ast H^* = 1 + 8K \frac{s + 1}{s^2 + 0.8s}
\]

to understand the over gain effect. We can vary the zero position. The zero behavior using bode and also the root locus plots should give the same physical understanding.
3 Lead and lag compensators

Lead and lag compensators are used quite extensively in control. A lead compensator can increase the stability or speed of response of a system; a lag compensator can reduce (but not eliminate) the steady state error. Depending on the effect desired, one or more lead and lag compensators may be used in various combinations. Lead, lag, and lead/lag compensators are usually designed for a system in transfer function form. The conversions page explains how to convert a state-space model into transfer function form.

3.1 Lead or phase-lead compensator using root locus

A first-order lead compensator can be designed using the root locus. A lead compensator in root locus form is given by

\[ G(s) = Kc \frac{s - Z_o}{s - P_o} \]

where magnitude of \( Z_o \) is less than the magnitude of \( P_o \). A phase-lead compensator tends to shift the root locus toward the left half plane. This results in an improvement in the system’s stability and an increase in the response speed. How is this accomplished? If you recall finding the asymptotes of the root locus that lead to the zeros at infinity, the equation to determine the intersection of the asymptotes along the real axis is:

\[ \alpha = \frac{\Sigma \text{poles} - \Sigma \text{zeros}}{n - m} \]

When a lead compensator is added to a system, the value of this intersection will be a larger negative number than it was before. The net number of zeros and poles will be the same (one zero and one pole are added), but the added pole is a larger negative number than the added zero. Thus, the result of a lead compensator is that the asymptotes’ intersection is moved further into the left half plane, and the entire root locus will be shifted to the left. This can increase the region of stability as well as the response speed. In Matlab a phase lead compensator in root locus form is implemented by using the transfer function in the form

\[
\text{numlead} = kc * [1 \ z];
\]

\[
\text{denlead} = [1 \ p];
\]

and using the conv() function to implement it with the numerator and denominator of the plant

\[
\text{newnum} = \text{conv}(\text{num}, \text{numlead});
\]

\[
\text{newden} = \text{conv}(\text{den}, \text{denlead});
\]

3.2 Lead or phase-lead compensator using frequency response

A first-order phase-lead compensator can be designed using the frequency response. A lead compensator in frequency response form is given by

\[ G(s) = \frac{1 + aTS}{1 + Ts} \quad a > 1 \]

Note that this is equivalent to the root locus form

\[ G(s) = Kc \frac{s - Z_o}{s - P_o} \]
with $P_o = 1/T$, $Z_o = 1/aT$, and $Kc = a$. In frequency response design, the phase-lead compensator adds positive phase to the system over the frequency range $1/aT$ to $1/T$. A bode plot of a phase-lead compensator looks like the following (see figure 5).

![Bode plot for lead compensator](image)

**Figure 5: Bode plot for Lead compensator**

The two corner frequencies are at $1/aT$ and $1/T$; note the positive phase that is added to the system between these two frequencies. Depending on the value of $a$, the maximum added phase can be up to 90 degrees; if you need more than 90 degrees of phase, two lead compensators can be used. The maximum amount of phase is added at the center frequency, which is located at

$$\omega_n = \frac{1}{T\sqrt{a}}$$

The equation which determines the maximum phase is

$$\sin \phi = \frac{a - 1}{a + 1}$$

Additional positive phase increases the phase margin and thus increases the stability of the system. This type of compensator is designed by determining $a$ from the amount of phase needed to satisfy the phase margin requirements, and determining $T$ to place the added phase at the new gain-crossover frequency. Another effect of the lead compensator can be seen in the magnitude plot. The lead compensator increases the gain of the system at high frequencies (the amount of this gain is equal to $a$). This can increase the crossover frequency, which will help to decrease the rise time and settling time of the system. In Matlab, a phase lead compensator in frequency response form is implemented by using the transfer function in the form

$$\text{numlead} = [aT \ 1];$$
$$\text{denlead} = [T \ 1];$$

and using the conv() function to multiply it by the numerator and denominator of the plant.
newnum = conv(num, numlead);
newden = conv(den, denlead);

3.3 Advantages of lead compensators
1. Phase margin is usually improved.
2. This implies more damping and stability.
3. Bandwidth of open and closed loop systems increases.
4. Reduced overshoot.
5. Steady state error is not affected, since the application is at high frequencies. The gain added at low frequencies is zero.

3.4 Disadvantages of lead compensators
1. If the original open loop system is unstable, then one may require a large a which can amplify noise also. Or we may need a series of lead compensators.
2. When the original system has already low stability then either the phase has crossed -180 or is rapidly going toward -180. A lead compensator will not perform well in such situations. (we will see this later. the case with 2 lead compensators.).
3. The phase will rapidly approach -180 if the open loop Tr. Fn. has 2 or more poles close (real or complex) to each other and near the gain cross over.

4 Lag or Phase-Lag Compensator using Root Locus
A first-order lag compensator can be designed using the root locus. A lag compensator in root locus form is given by

\[ G(s) = \frac{s - Z_o}{s - P_o} \]

where the magnitude of \( Z_o \) is greater than the magnitude of \( P_o \). A phase-lag compensator tends to shift the root locus to the right, which is undesirable. For this reason, the pole and zero of a lag compensator must be placed close together (usually near the origin) so they do not appreciably change the transient response or stability characteristics of the system. How does the lag controller shift the root locus to the right? If you recall finding the asymptotes of the root locus that lead to the zeros at infinity, the equation to determine the intersection of the asymptotes along the real axis is:

\[ \alpha = \frac{\Sigma \text{poles} - \Sigma \text{zeros}}{n - m} \]

When a lag compensator is added to a system, the value of this intersection will be a smaller negative number than it was before. The net number of zeros and poles will be the same (one zero and one pole are added), but the added pole is a smaller negative number than the added zero. Thus, the result of a lag compensator is that the asymptotes’ intersection is moved closer
to the right half plane, and the entire root locus will be shifted to the right. It was previously stated that that lag controller should only minimally change the transient response because of its negative effect. If the phase-lag compensator is not supposed to change the transient response noticeably, what is it good for? The answer is that a phase-lag compensator can improve the system’s steady-state response. It works in the following manner. At high frequencies, the lag controller will have unity gain. At low frequencies, the gain will be \( \frac{z_0}{p_0} \) which is greater than 1. This factor \( \frac{z_0}{p_0} \) will multiply the position, velocity, or acceleration constant (Kp, Kv, or Ka), and the steady-state error will thus decrease by the factor \( \frac{z_0}{p_0} \). In Matlab, a phase lead compensator in root locus form is implemented by using the transfer function in the form

\[
\text{numlag} = [1 \ z];
\]
\[
\text{denlag} = [1 \ p];
\]

and using the conv() function to implement it with the numerator and denominator of the plant

\[
\text{newnum} = \text{conv}(\text{num}, \text{numlag});
\]
\[
\text{newden} = \text{conv}(\text{den}, \text{denlag});
\]

4.1 Lag or Phase-Lag Compensator using Frequency Response

A first-order phase-lag compensator can be designed using the frequency response. A lag compensator in frequency response form is given by

\[
G(s) = \frac{1}{a} \left( \frac{1 + aTs}{1 + Ts} \right) \quad a < 1
\]

The phase-lag compensator looks similar to a phase-lead compensator, except that \( a \) is now less than 1. The main difference is that the lag compensator adds negative phase to the system over the specified frequency range, while a lead compensator adds positive phase over the specified frequency. A bode plot of a phase-lag compensator looks like the following (see figure 6)

The two corner frequencies are at \( \frac{1}{T} \) and \( \frac{1}{aT} \). The main effect of the lag compensator is shown in the magnitude plot. The lag compensator adds gain at low frequencies; the magnitude of this gain is equal to \( a \). The effect of this gain is to cause the steady-state error of the closed-loop system to be decreased by a factor of \( a \). Because the gain of the lag compensator is unity at middle and high frequencies, the transient response and stability are not impacted too much. The side effect of the lag compensator is the negative phase that is added to the system between the two corner frequencies. Depending on the value \( a \), up to -90 degrees of phase can be added. Care must be taken that the phase margin of the system with lag compensation is still satisfactory. In Matlab, a phase-lag compensator in frequency response form is implemented by using the transfer function in the form

\[
\text{numlead} = [a * T \ 1];
\]
\[
\text{denlead} = a * [T \ 1];
\]

and using the conv() function to implement it with the numerator and denominator of the plant

\[
\text{newnum} = \text{conv}(\text{num}, \text{numlead});
\]
\[
\text{newden} = \text{conv}(\text{den}, \text{denlead});
\]
4.2 Advantages and disadvantages of lag compensators

1. Lag compensator moves the gain cross over to a lower frequency while keeping the phase curve unchanged.

2. Bandwidth of open and closed loop systems decreases.

3. Rise time increases and system is slower.

4. Since phase curve is untouched, lowering the gain crossover can improve phase margin and gain margin.

4.3 Lead-lag Compensator using either Root Locus or Frequency Response

A lead-lag compensator combines the effects of a lead compensator with those of a lag compensator. The result is a system with improved transient response, stability and steady-state error. To implement a lead-lag compensator, first design the lead compensator to achieve the desired transient response and stability, and then add on a lag compensator to improve the steady-state response.