Feedback Systems

An Introduction for Scientists and Engineers
SECOND EDITION

Karl Johan Åström
Richard M. Murray

Version v3.0j (2019-08-17)

Chapter Twelve

Frequency Domain Design

Sensitivity improvements in one frequency range must be paid for with sensitivity deteriorations in another frequency range, and the price is higher if the plant is open-loop unstable. This applies to every controller, no matter how it was designed.

Gunter Stein in the inaugural IEEE Bode Lecture, 1989 [Ste03].

In this chapter we continue to explore the use of frequency domain techniques with a focus on the design of feedback systems. We begin with a more thorough description of the performance specifications for control systems and then introduce the concept of "loop shaping" as a mechanism for designing controllers in the frequency domain. Additional techniques discussed in this chapter include feedforward compensation, the root locus method, and nested controller design.

12.1 SENSITIVITY FUNCTIONS

In the previous chapter, we used proportional-integral-derivative (PID) feedback as a mechanism for designing a feedback controller for a given process. In this chapter we will expand our approach to include a richer repertoire of controllers and tools for shaping the frequency response of the closed loop system.

One of the key ideas in this chapter is that we can design the behavior of the closed loop system by focusing on the open loop transfer function. This same approach was used in studying stability using the Nyquist criterion: we plotted the Nyquist plot for the *open* loop transfer function to determine the stability of the *closed* loop system. From a design perspective, the use of loop analysis tools is very powerful: since the loop transfer function is $L = PC$, if we can specify the desired performance in terms of properties of L, we can directly see the impact of changes in the controller C. This is much easier, for example, than trying to reason directly about the tracking response of the closed loop system, whose transfer function is given by $G_{yr} = PC/(1 + PC)$.

We will start by investigating some key properties of a closed loop control system. A block diagram of a basic two degree-of-freedom control system is shown in Figure 12.1. The system loop is composed of two components: the process and the controller. The two degree-of-freedom controller itself has two blocks: the feedback block C and the feedforward block F. There are two disturbances acting on the process, the load disturbance v and the measurement noise w. The load disturbance represents disturbances that drive the process away from its desired behavior, while the measurement noise represents disturbances that corrupt information about the process given by the sensors. For example, in a cruise control system the major load disturbances are changes in the slope of the road, and measurement noise is
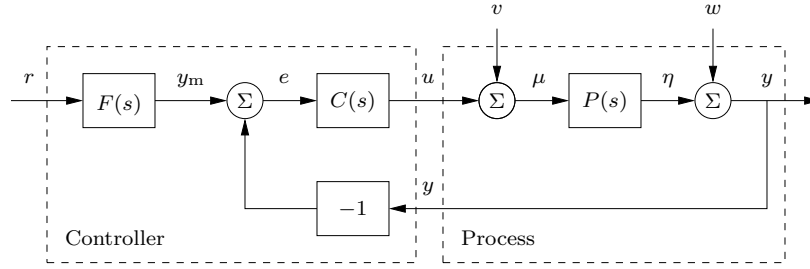
Figure 12.1: Block diagram of a control system with two degrees of freedom. The controller has a feedback block C and a feedforward block F. The external signals are the reference signal r, the load disturbance v, and the measurement noise w. The measured output is y, and the control signal is u.

caused by the electronics that convert pulses measured on a rotating shaft to a velocity signal. The load disturbances typically have low frequencies, lower than the controller bandwidth, while measurement noise typically has higher frequencies. It is assumed that load disturbances enter at the process input and that the measurement noise acts at the process output. This is a simplification since disturbances may enter the process in many different ways and that there may be dynamics in the sensors. These assumptions allow us to streamline the presentation without significant loss of generality.

The process output η is the variable that we want to control, and our ultimate goal is to make η track a reference signal r. To shape the response to reference signals, it is common to use a feedforward block to generate a desired (or model) reference signal y_{m} that represents the actual signal we attempt to track. Control is based on the difference between the model reference y_{m} and the measured signal y, where the measurements are corrupted by measurement noise w. The process is influenced by the controller via the control variable u. The process is thus a system with three inputs (the control variable u, the load disturbance v, and the measurement noise w) and one output (the measured signal y). The controller is a system with two inputs (the measured signal y and the reference signal r) and one output (the control signal u). Note that the control signal u is an input to the process and the output of the controller, and that the measured signal y is the output of the process and an input to the controller.

Since the control system in Figure 12.1 is composed of linear elements, the relations between the signals in the diagram can be expressed in terms of the transfer functions. The overall system has three external inputs: the reference r, the load disturbance v, and the measurement noise w. Any of the remaining signals can be relevant for design, but the most common ones are the error e, the control input u, and the output y. In addition, the process input and output, μ and ν, are also useful. Table 12.1 summarizes the transfer functions between the external inputs (rows) and remaining signals (columns).

Although there are 15 entries in the table, many transfer functions appear more than once. For most control designs we focus on the following subset, which we call

Table 12.1: Transfer functions relating the signals of the control system in Figure 12.1. The external inputs are the load disturbance v, measurement noise w, and the reference r, represented by each row. The columns represent the measured signal y, control input u, error e, process input μ, and process output η that are most relevant for system performance.

y	u	e	μ	η	
$\dfrac{PCF}{1+PC}$	$\dfrac{CF}{1+PC}$	$\dfrac{F}{1+PC}$	$\dfrac{CF}{1+PC}$	$\dfrac{PCF}{1+PC}$	r
$\dfrac{P}{1+PC}$	$\dfrac{-PC}{1+PC}$	$\dfrac{-P}{1+PC}$	$\dfrac{1}{1+PC}$	$\dfrac{P}{1+PC}$	v
$\dfrac{1}{1+PC}$	$\dfrac{-C}{1+PC}$	$\dfrac{-1}{1+PC}$	$\dfrac{-C}{1+PC}$	$\dfrac{-PC}{1+PC}$	w

the *Gang of Six*:

$$G_{yr} = \frac{PCF}{1+PC}, \qquad -G_{uv} = \frac{PC}{1+PC}, \qquad G_{yv} = \frac{P}{1+PC},$$
$$G_{ur} = \frac{CF}{1+PC}, \qquad -G_{uw} = \frac{C}{1+PC}, \qquad G_{yw} = \frac{1}{1+PC}. \tag{12.1}$$

The transfer functions in the first column of equation (12.1) give the responses of the process output y and the control signal u to the reference signal r. The second column gives the responses of the control variable u to the load disturbance v and the measurement noise w, and the final column gives the responses of the measured signal y to those two inputs. (Note that the sign convention in equation (12.1) is chosen for later convenience and does not affect the magnitude of the Gang of Six transfer functions.)

The response of the system to load disturbances and measurement noise is of particular importance and these transfer functions are referred to as *sensitivity functions*. They represent the sensitivity of the system to the various inputs, and they have specific names:

$$S = \frac{1}{1+PC} \quad \begin{matrix} \text{sensitivity} \\ \text{function} \end{matrix} \qquad PS = \frac{P}{1+PC} \quad \begin{matrix} \text{load (or input)} \\ \text{sensitivity} \\ \text{function} \end{matrix}$$

$$T = \frac{PC}{1+PC} \quad \begin{matrix} \text{complementary} \\ \text{sensitivity} \\ \text{function} \end{matrix} \qquad CS = \frac{C}{1+PC} \quad \begin{matrix} \text{noise (or output)} \\ \text{sensitivity} \\ \text{function} \end{matrix} \tag{12.2}$$

Because these transfer functions are particularly important in feedback control design, they are called the *Gang of Four*, and they have many interesting properties that will be discussed in detail in the rest of the chapter. Good insight into these properties is essential in understanding the performance of feedback systems for the purposes of both analysis and design.

While the Gang of Four capture the response to disturbances, we are also interested in the response of the system to the reference signal r. The remaining two elements in the full Gang of Six capture the relationship between the reference

signal and the measured output y plus the control input u:

$$G_{yr} = \frac{PCF}{1+PC}, \qquad G_{ur} = \frac{CF}{1+PC}.$$

We see that F can be used to design these responses and provides a second degree of freedom in addition to the feedback controller C. In practice, it is common to first design the feedback controller C using the Gang of Four to provide good response with respect to load disturbances and measurement noise, and then use F and the remaining transfer functions as part of the full Gang of Six to obtain good reference tracking.

In addition to the Gang of Six, one other signal that can be important is the error between the reference r and the process output η (prior to the addition of measurement noise), which satisfies

$$\begin{aligned}\epsilon = r - \eta &= \left(1 - \frac{PCF}{1+PC}\right) r - \frac{P}{1+PC} v - \frac{PC}{1+PC} w \\ &= (1 - TF)r - PSv - Tw.\end{aligned}$$

The signal ϵ is not actually present in our diagram, but is the true error that represents the tracking deviation. We see that it consists of a particular combination of transfer functions chosen from the Gang of Six.

The special case of $F = 1$ is called a system with (pure) *error feedback* because all control actions are based on feedback from the error. In this case the transfer functions given by equations (12.1) and (12.2) are the same and the system is completely characterized by the Gang of Four. In addition, the true tracking error becomes

$$\epsilon = Sr - PSv - Tw.$$

Notice that we have less freedom in design of a system with error feedback because the feedback controller C must now deal with both disturbance attenuation, robustness, and reference signal tracking.

The transfer functions in equation (12.2) have many interesting properties. For example, it follows from equation (12.2) that $S + T = 1$, which explains why T is called the complementary sensitivity function. The loop transfer function PC will typically go to zero for large s, which implies that T goes to zero and S goes to one as s goes to infinity. Thus, it will not be possible to track very high-frequency reference signals ($|G_{yr}| = |FT| \to 0$) and any high-frequency noise will propagate unfiltered to the error ($|G_{ew}| = |S| \to 1$). For controllers with integral action and processes with non-vanishing zero frequency gain, the loop transfer function PC goes to infinity for small s, which implies that S goes to zero and T goes to one as s goes to zero. low-frequency signals are thus tracked well ($|G_{yr}| = |FT| \to 0$), and low-frequency disturbances can be completely attenuated ($|G_{ev}| = |PS| \to 0$). Many more properties of the sensitivity functions will be discussed in detail later in this chapter and in Chapters 13 and 14. Good insight into these properties is essential in understanding the performance of feedback systems for the purposes of both analysis and design. The transfer functions are also used to formulate specifications on control systems.

In Chapter 10 we focused on the loop transfer function, and we found that its properties gave useful insights into the properties of a system. The loop transfer function does not, however, always give a complete characterization of the closed
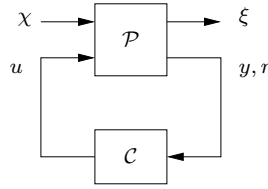
Figure 12.2: A more general representation of a feedback system. The process input u represents the control signal, which can be manipulated, and the process input χ represents the other signals that influence the process. The process output consists of the measured variable(s) y, the reference signal r, and the signal vector ξ representing the other signals of interest in the control design.

loop system. In particular, it can happen that there are pole/zero cancellations in the product of P and C such that $1 + PC$ has no unstable poles, but one of the other Gang of Four transfer functions might be unstable. The follow example illustrates this difficulty.

Example 12.1 The loop transfer function gives only limited insight
Consider a process with the transfer function $P(s) = 1/(s - a)$ controlled by a PI controller with error feedback having the transfer function $C(s) = k(s - a)/s$. The loop transfer function is $L = k/s$, and the sensitivity functions are

$$S = \frac{1}{1 + PC} = \frac{s}{s + k}, \qquad PS = \frac{P}{1 + PC} = \frac{s}{(s - a)(s + k)},$$
$$CS = \frac{C}{1 + PC} = \frac{k(s - a)}{s + k}, \qquad T = \frac{PC}{1 + PC} = \frac{k}{s + k}.$$

Notice that the factor $s - a$ is canceled when computing the loop transfer function and that this factor also does not appear in the sensitivity functions S and T. However, cancellation of the factor is very serious if $a > 0$ since the transfer function PS relating load disturbances to process output is then unstable. A small disturbance v then leads to an unbounded output, which is clearly not desirable.

∇

If all four of the transfer functions in equation (12.2) are stable we say that the feedback system is *internally stable*. In addition, if there is a feedforward controller F then it should also be stable in order for the full system to be internally stable. For more general systems, which may contain additional transfer functions and feedback loops, the system is internally stable if all possible input/output transfer functions are stable. For simplicity we will often say that a closed loop system is stable when we mean that it is internally stable.

As mentioned previously, the system in Figure 12.1 represents a special case because it is assumed that the load disturbance enters at the process input and that the measured output is the sum of the process variable and the measurement noise. Disturbances can enter in many different ways, and the sensors may have dynamics. A more abstract way to capture the general case is shown in Figure 12.2, which has only two blocks representing the process ($\mathcal{P}$) and the controller ($\mathcal{C}$). The process has two inputs, the control signal u and a vector of disturbances χ, and three outputs, the measured signal y, the reference signal r and a vector of signals ξ that is used

(a) Frequency response of $S(i\omega)$ (b) Frequency response of $G_{yr}(i\omega)$
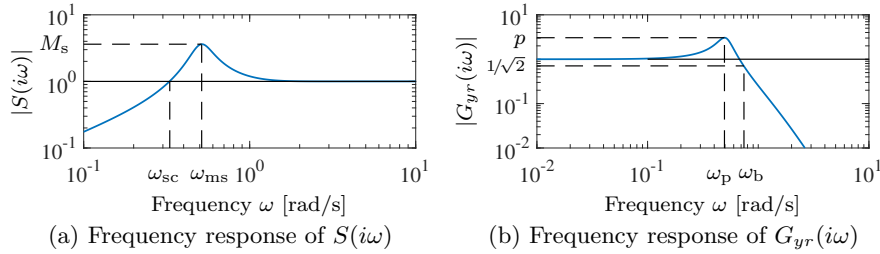
Figure 12.3: Illustration of specifications in frequency domain. (a) Gain curve of sensitivity function; the maximum sensitivity M_s is a robustness measure. (b) Gain curve of the transfer function G_{yr} with peak value p, peak frequency ω_{pc}, and bandwidth ω_b.

to specify performance. The system in Figure 12.1 can be captured by choosing $\chi = (r, v, w)$ and $\xi = (e, \mu, \eta, \epsilon)$. The process transfer function $\mathcal{P}$ describes the effect of χ and u on ξ, y, and r, and the controller transfer function $\mathcal{C}$ describes how u is related to y and r (see Exercise 12.2). Restricting the signal ξ to contain the errors e and ϵ, the control problem can be formulated as finding a controller $\mathcal{C}$ so that the gain of the transfer function from the disturbance $\chi = (r, v, w)$ to the generalized control error $\xi = (e, \epsilon)$ has the smallest gain (discussed further in Section 13.4).

Processes with multiple inputs and outputs can be handled by regarding u and y as vectors. Representations at these higher levels of abstraction are useful for the development of theory because they make it possible to focus on fundamentals and to solve general problems with a wide range of applications. However, care must be exercised to maintain the coupling to the real-world control problems we intend to solve and we must keep in mind that matrix multiplication is not commutative.

12.2   PERFORMANCE SPECIFICATIONS

A key element of the control design process is how we specify the desired performance of a system. Specifications capture robustness to process variations as well performance in terms of the ability to follow reference signals and attenuate load disturbances without injecting too much measurement noise. The specifications are expressed in terms of transfer functions such as the Gang of Six and the loop transfer function, and are often represented by features of the transfer functions or their time and frequency responses.

Robustness to process variations was discussed extensively in Section 10.3, where we introduced gain margin g_m, phase margins φ_m, and stability margin s_m, as shown in Figure 10.11. The largest value of the sensitivity function $M_s = 1/s_m$ is another robustness measure, as illustrated in Figure 12.3a.

To provide specifications it is desirable to capture the characteristic properties of a system with a few parameters. Features of step responses that we have already seen are overshoot, rise time, and settling time, as shown in Figure 6.9. Common features of frequency responses include peak value(s), peak frequency, gain crossover frequency, and bandwidth. Other features of the frequency response include the maximum value of sensitivity function M_s (occurring at frequency ω_{ms})

and the maximum value of the complementary sensitivity function M_t (occurring at frequency ω_{mt}). The *sensitivity crossover frequency* ω_{sc} is defined as the frequency where the magnitude of the sensitivity function $S(j\omega)$ is 1. The various crossover frequencies and the bandwidth are only well defined if the curves are monotone; if this is not the case the lowest such frequency is typically used.

There are interesting relations between specifications in the time and frequency domains. Roughly speaking, the behavior of time responses for short times is related to the behavior of frequency responses at high frequencies, and vice versa. The precise relations are given by the Laplace transform. There are also useful relations between features in the time and frequency domain; typical examples are given in Tables 7.1 and 7.2 in Section 7.3.

In the remainder of this section we consider the different types of responses that are commonly used in control design and describe the types of specifications that are relevant for each.

Response to Reference Signals

Consider the basic feedback loop in Figure 12.1. The responses of the output y and the control signal u to the reference r are described by the transfer functions $G_{yr} = PCF/(1 + PC)$ and $G_{ur} = CF/(1 + PC)$ ($F = 1$ for systems with pure error feedback). Specifications can be expressed in terms of features of the transfer function G_{yr}, such as peak value M_p, peak frequency ω_p, and bandwidth ω_b, as shown in Figure 12.3b.

In the special case where $F = 1$, the transfer function G_{yr} is equal to the complementary sensitivity function T. However, in many cases it is useful to retain the ability to shape the input/output response by using $F \neq 1$. This distinction is captured in the use of the full Gang of Six rather than just the Gang of Four.

The transfer function G_{yr} typically has unit zero frequency gain because we want to design the system so that the response to a step input has zero steady-state gain. The behavior of the transfer function at low frequencies determines the tracking error for slow reference signals. We can capture this analytically by making the following series expansion for small s:

$$G_{yr}(s) \approx 1 - e_1 s - e_2 s^2 - \cdots,$$

where the coefficients e_k are called *error coefficients*. If the reference signal is $r(t)$, the tracking error is then

$$e(t) = r(t) - y(t) = e_1 \frac{dr}{dt} + e_2 \frac{d^2 r}{d^2 t} + \cdots$$

It follows that a ramp input $r(t) = v_0 t$ gives a steady-state tracking error $v_0 e_1$, and we can conclude that the steady-state tracking error is zero if $e_1 = 0$. A system with $e_1 = 0$ has the steady-state error $e(t) = 2ae_2$ for the input $r(t) = a_0 t^2$.

It has been a long practice to focus on the output when we give specifications. However, it is useful to also consider the response of the control signal because this allows us to judge the magnitude and rate of the control signal required to obtain the output response. This is illustrated in the following example.

Example 12.2 Reference signal tracking for a third-order system
Consider a process with the transfer function $P(s) = (s + 1)^{-3}$ and a PI controller
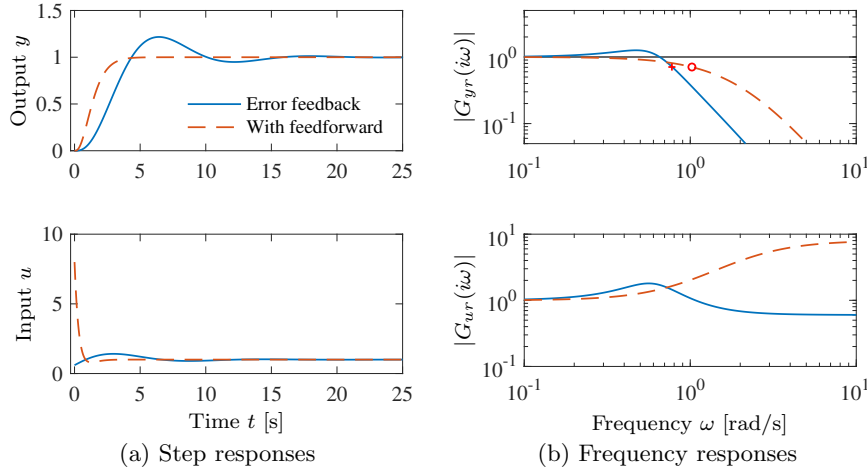
Figure 12.4: Reference signal responses. The responses in process output y and control signal u to a unit step in the reference signal r are shown in (a), and the gain curves of G_{yr} and G_{ur} are shown in (b). Results with PI control with error feedback are shown by solid lines, and the dashed lines show results for a controller with a feedforward compensator.

with error feedback having the gains $k_{\mathrm{p}} = 0.6$ and $k_{\mathrm{i}} = 0.5$. The responses are illustrated in Figure 12.4. The solid lines show results for a proportional-integral (PI) controller with error feedback. The dashed lines show results for a controller with feedforward controller

$$F = \frac{G_{yr}(1 + PC)}{PC} = \frac{2s^4 + 6s^3 + 6s^2 + 3.2s + 1}{0.15s^4 + 1.025s^3 + 2.55s^2 + 2.7s + 1},$$

designed to give the closed loop transfer function $G_{yr} = (0.5s+1)^{-3}$. Looking at the time responses, we find that the controller with feedforward gives a faster response with no overshoot. However, much larger control signals are required to obtain the fast response. The initial value of the control signal for the controller with feedforward is 13.3, compared to 0.6 for the regular PI controller. The controller with feedforward has a larger bandwidth (marked with ∘) and no resonant peak. The transfer function G_{ur} also has higher gain at high frequencies. $\qquad \nabla$

We can get some insight into the relations between time and frequency responses from Figure 12.4. The figures in the top row show the unit step response and the frequency response for the transfer function G_{yr}, and the lower plots show the same quantities for G_{ur}. The dashed time and frequency responses have no peaks while the solid responses have peaks. The peaks are related in the sense that a large overshoot in the time response corresponds to a large resonant peak in the frequency response. The time responses in the bottom plot of Figure 12.4 have the initial values 8 (dashed) and 6 (solid), and the frequency responses have the same final values. In general, it can be shown using the Laplace transform (or appropriate exponential responses) and the initial and final value theorems that for a unit command signal $u(t) \to G(i\infty)$ as $t \to \infty$ and if $x(0) = 0$ then $u(0) = G_{ur}(0)$.

The dashed time response is faster than the solid time response and the dashed frequency response has larger bandwidth than the solid frequency response. The
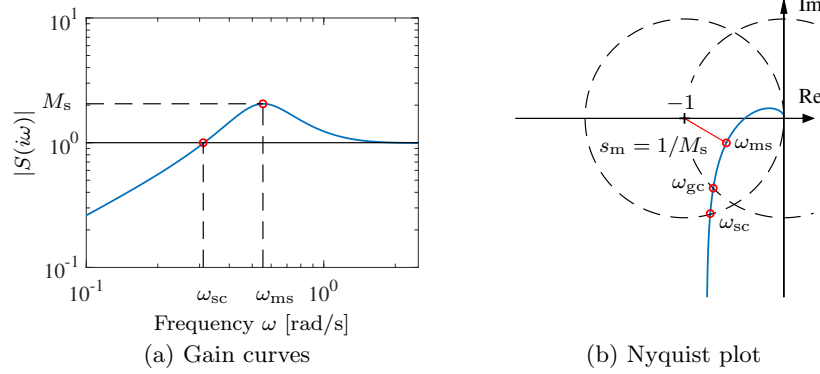
(a) Gain curves   (b) Nyquist plot

Figure 12.5: Illustration of sensitivity to disturbances. The gain curves of the sensitivity function S and the loop transfer function L are shown in (a). The Nyquist plot of the loop transfer function L is shown in (b). Disturbances with frequencies less than the sensitivity crossover frequency, to the left of ω_{sc} in (a) and inside the dashed circle in (b), are attenuated by feedback. Disturbances with frequencies higher than ω_{sc} are amplified. The largest amplification occurs for the frequency ω_{ms}, where the sensitivity has its largest value M_s, the point where the Nyquist curve is closest to the critical point -1 in (b).

product of the rise time of the unit step response and the bandwidth of a transfer function (the *rise time-bandwidth product*) is a dimension-free variable that is useful characteristic. The time responses in Figure 12.4 have rise times of $T_r = 1.7$ (dashed) and 3.0 (solid), and the corresponding bandwidths are $\omega_b = 1.9$ (dashed) and 0.8 (solid), which gives the products $T_r\,\omega_b = 3.2$ (dashed) and 2.4 (solid). A similar observation can be made from Tables 7.1 and 7.2 in Section 7.3, which gives $T_r\,\omega_b \approx 2.7$–2.8. It thus appears that the product of the rise time of the step response and the bandwidth of the frequency response is approximately constant ($T_r\,\omega_b \approx 3$). It can be shown that the rise time-bandwidth product increases if the frequency response has a faster roll-off (see Exercise 12.6, which uses a slightly different definition of bandwidth).

Response to Load Disturbances and Measurement Noise

A simple criterion for disturbance attenuation is to compare the output of the closed loop system in Figure 12.1 with the output of the corresponding open loop system, which can be obtained by setting $C = 0$ in the figure. With identical disturbances for the open and closed loop systems, the output of the closed loop system can be obtained simply by sending the open loop output through a system with the transfer function S (Exercise 12.8). The sensitivity S function thus directly shows how feedback influences the response of the output to disturbances both in the form of load disturbances and measurement noise. Disturbances with frequencies such that $|S(i\omega)| < 1$ are attenuated, but disturbances with frequencies such that $|S(i\omega)| > 1$ are amplified by feedback. The *sensitivity crossover frequency* ω_{sc} is the (lowest) frequency where $|S(i\omega)| = 1$, as shown in Figure 12.5a.

Since the sensitivity function is related to the loop transfer function by $S = 1/(1 + L)$, disturbance attenuation can be visualized graphically by the Nyquist plot of the loop transfer function, as shown in Figure 12.5b. The complex number

$1 + L(i\omega)$, which is the inverse of the sensitivity function $S(i\omega)$, can be represented as the vector from the point -1 to the point $L(i\omega)$ on the Nyquist curve. The sensitivity is thus less than 1 for all points outside a circle with radius 1 and center at -1. Disturbances with frequencies in this range are attenuated by the feedback, while disturbances with frequencies inside the circle are amplified.

The maximum sensitivity M_s, which occurs at the frequency ω_ms, is a measure of the largest amplification of the disturbances. The sensitivity crossover frequency ω_sc and the maximum sensitivity M_s are two parameters that give a gross characterization of load disturbance attenuation. For systems where the phase margin is $\varphi_\mathrm{m} = 60°$, it can be shown that the sensitivity crossover frequency ω_sc is equal to the gain crossover frequency ω_gc and the complementary sensitivity function crossover frequency ω_tc (Exercise 12.19). Notice that the maximum magnitude of $1/(1 + L(i\omega))$ corresponds to the minimum of $|1 + L(i\omega)|$, which is the stability margin s_m defined in Section 10.3, so that $M_\mathrm{s} = 1/s_\mathrm{m}$. The maximum sensitivity is therefore also a robustness measure.

The transfer function G_{yv} from load disturbance v to process output y for the system in Figure 12.1 is

$$G_{yv} = \frac{P}{1 + PC} = PS = \frac{T}{C} \approx \frac{1}{C}. \tag{12.3}$$

Load disturbances typically have low frequencies. For small s (low frequencies) we have $T \approx 1$ which gives the approximation above. For processes with $P(0) \neq 0$ and controllers with integral action we have $C(s) \approx k_\mathrm{i}/s$ for small s. A controller with integral action thus attenuates disturbances with low frequencies effectively, and the integral gain k_i is a measure of disturbance attenuation. For high frequencies we have $S \approx 1$ which implies that $G_{yv} \approx P$ for large s.

Measurement noise, which typically has high frequencies, generates rapid variations in the control variable that are detrimental because they cause wear in the actuators and can even saturate an actuator. It is thus important to keep variations in the control signal due to measurement noise at reasonable levels—a typical requirement is that the variations are only a fraction of the allowable range of the control signal. The effects of measurement noise are captured by the transfer function from the measurement noise to the control signal,

$$-G_{uw} = \frac{C}{1 + PC} = \frac{T}{P} = CS \approx C. \tag{12.4}$$

The approximation is valid for large s (high frequencies), which is appropriate for measurement noise. The formula clearly shows it is useful to filter the derivative so that the transfer function $C(s)$ goes to zero for large s (high-frequency roll-off).

Example 12.3 Disturbance attenuation for a third-order system
Consider a process with the transfer function $P(s) = (s+1)^{-3}$ and a proportional-integral-derivative (PID) controller with gains $k_\mathrm{p} = 2$, $k_\mathrm{i} = 1.5$, and $k_\mathrm{d} = 2.0$. We augment the controller with a second-order noise filter with damping ratio $1/\sqrt{2}$ and $T_\mathrm{f} = 0.1$. The controller transfer function then becomes

$$C(s) = \frac{k_\mathrm{d}s^2 + k_\mathrm{p}s + k_\mathrm{i}}{s(s^2 T_\mathrm{f}^2/2 + sT_\mathrm{f} + 1)}. \tag{12.5}$$
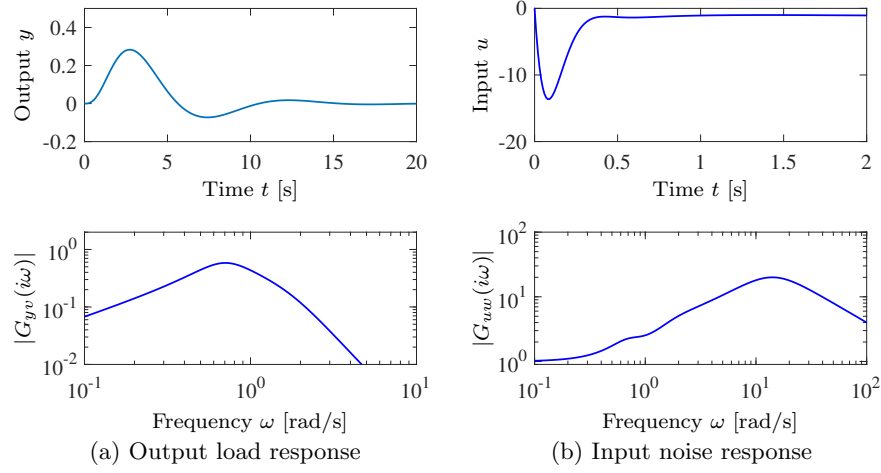
Figure 12.6: Closed loop disturbance responses for Example 12.3. The closed loop unit step response and frequency response for the transfer function G_{yv} from load disturbance v to process output y are shown in (a) and the corresponding responses for the transfer function G_{uw} from measurement noise w to the control signal u are shown in (b).

The closed loop system responses are illustrated in Figure 12.6. The closed loop response of the output y to a unit step in the load disturbance v in the upper part of Figure 12.6a has a peak of 0.28 at time $t = 2.73$ s. The frequency response in Figure 12.6a shows that the gain has a maximum of 0.58 at $\omega = 0.7$ rad/s.

The closed loop response of the control signal u to a step in measurement noise w is shown in Figure 12.6b. The high-frequency roll-off of the transfer function $G_{uw}(i\omega)$ is due to filtering; without it the gain curve in Figure 12.6b would continue to rise after 20 rad/s. The step response has a valley of -14 at $t = 0.08$ s. The frequency response has a peak of 20 at $\omega = 14$ rad/s. Notice that the peak occurs at a frequency far above the peak of the response to load disturbances and far above the gain crossover frequency $\omega_{\mathrm{gc}} = 0.78$ rad/s. An approximation derived in Exercise 12.10 gives $\max |CS(i\omega)| \approx k_{\mathrm{d}}/T_{\mathrm{f}} = 20$ for $\omega = \sqrt{2}/T_{\mathrm{d}} = 14.1$ rad/s. $\quad\nabla$

Figure 12.6 also gives insight into the relation between the time and frequency responses. The frequency response of the transfer functions G_{yv} and G_{uw} have band-pass characteristics and their gains go to zero for high and low frequencies. A consequence is that the corresponding step responses are zero both for small and large times. The frequency response G_{yv} in Figure 12.6a has a peak of 0.6 for $\omega_{\mathrm{p}} = 0.7$ and the time response has a peak of 0.3 for $t_{\mathrm{p}} = 2.7$, hence $\omega_{\mathrm{p}} t_{\mathrm{p}} = 1.9$. Figure 12.6b shows that the low-frequency gain of the transfer function G_{uw} and steady-state time response are both 1, and the time response starts at zero because the frequency response goes to zero at high frequencies. The frequency response has a peak of 20 for $\omega_{\mathrm{p}} = 14$ and the time response has a peak of 14 for $t_{\mathrm{p}} = 0.08$, hence $\omega_{\mathrm{p}} t_{\mathrm{p}} = 1.1$. These observations support the simple rules for transfer functions with a band-pass character: the product of the peak time of the step response and the resonant peak of the frequency response is in the range of 1 to 2 (Exercise 12.7).
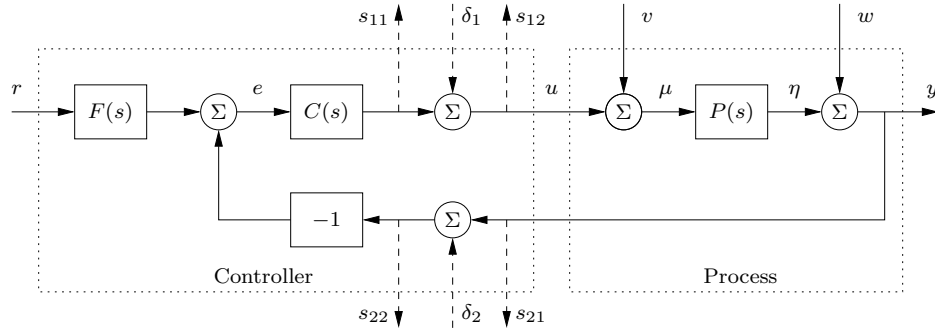
Figure 12.7: Specifications can be tested by injecting signals at test points δ_k and measuring responses at s_{ij}. Compare with Figure 12.1

.

Measuring Specifications

Many specifications are expressed in terms of properties of the transfer functions in the Gang of Six and they can easily be checked simply by computing the transfer functions numerically. To test a real system is is necessary to provide the controller with test points for injecting and measuring signals. Some possible test points are shown in Figure 12.7. As an example, the transfer function G_{yv}, which characterizes response of process output to load a disturbance, can be found by injecting a signal at δ_1 and measuring the output s_{21}. A frequency analyzer that measures the transfer function directly is very convenient for such a test. By measuring the transfer functions we can ensure that robustness and performance are maintained during the design phase and operation of a system.

12.3 FEEDBACK DESIGN VIA LOOP SHAPING

One advantage of the Nyquist stability theorem is that it is based on the loop transfer function $L = PC$, which is the product of the transfer functions of the process and the controller. It is thus easy to see how the controller influences the loop transfer function. For example, to make an unstable system stable we simply have to bend the Nyquist curve away from the critical point. This simple idea is the basis of several different design methods collectively called *loop shaping*. These methods are based on choosing a compensator that gives a loop transfer function with a desired shape. One possibility is to determine a loop transfer function that gives a closed loop system with the desired properties and to compute the controller as $C = L/P$. This approach may lead to controllers of high order and there are limits if the process transfer function has poles and zeros in the right half-plane, as discussed briefly in Section 12.4 and in more detail in Section 14.3. Another possibility is to start with the process transfer function, change its gain to obtain the desired bandwidth, and then add poles and zeros until the desired shape is obtained. In this section we will explore different loop-shaping methods for control law design.
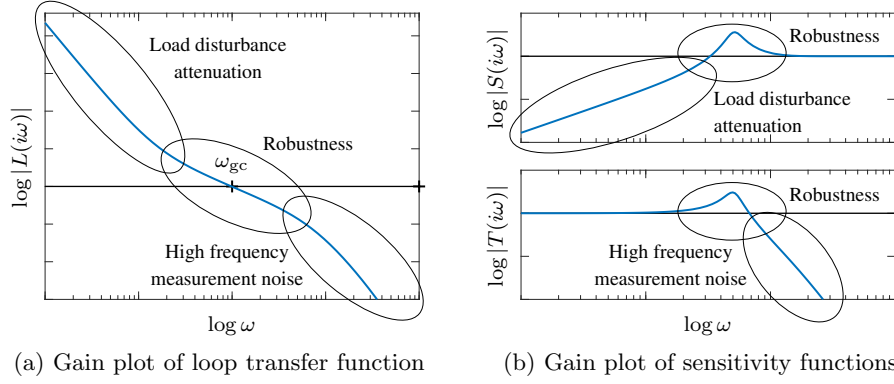
(a) Gain plot of loop transfer function   (b) Gain plot of sensitivity functions

Figure 12.8: Gain plots of the loop transfer function (a) and the sensitivity functions (b) for typical loop transfer functions. The plot on the left shows the gain curve and the plots on the right show the sensitivity function and complementary sensitivity function. The crossover frequency ω_{gc} determines the attenuation of load disturbances, bandwidth, and response time of the closed loop system. The slope n_{gc} of the gain curve of $L(s)$ at the gain crossover frequency ω_{gc} determines the robustness of the closed loop systems (equation (12.6)). At low frequencies, a large magnitude of L provides good load disturbance rejection and reference tracking, while at high frequencies a small loop gain avoids injecting too much measurement noise.

Design Considerations

We will first discuss a suitable shape for the loop transfer function that gives good performance and good stability margins. Figure 12.8 shows a typical loop transfer function. Good performance requires that the loop transfer function is large for frequencies where we desire good tracking of reference signals and good attenuation of low-frequency load disturbances. Since $S = 1/(1 + L)$, it follows that for frequencies where $|L| > 100$ disturbances will be attenuated by approximately a factor of 100 or more and the tracking error is less than 1%. The transfer function from measurement noise to control action is $CS = C/(1 + L)$. To avoid injecting too much measurement noise, which can create undesirable control actions, the controller transfer function should have low gain at high frequencies, a property called *high-frequency roll-off*. The loop transfer function should thus have roughly the shape shown in Figure 12.8. It has unit gain at the gain crossover frequency ($|L(i\omega_{gc})| = 1$), large gain for lower frequencies, and small gain for higher frequencies.

Robustness is determined by the shape of the loop transfer function around the crossover frequency. Good robustness requires good stability margins, which imposes requirements on the loop transfer function around the gain crossover frequency ω_{gc}. It would be desirable to transition from high loop gain $|L(i\omega))|$ at low frequencies to low loop gain as quickly as possible, but robustness requirements expressed via Bode's relations (Section 10.4) impose restrictions on how fast the gain can decrease. Equation (10.9) implies that the slope of the gain curve at ω_{gc} cannot be too steep. If the gain curve has a constant slope around ω_{gc}, we have

the following relation between slope n_{gc} and phase margin φ_{m} (in degrees):

$$n_{\mathrm{gc}} \approx -2 + \frac{\varphi_{\mathrm{m}}}{90}, \qquad (12.6)$$

for a minimum-phase system. A steeper slope thus gives a smaller phase margin. The equation is a reasonable approximation when the gain curve does not deviate too much from a straight line. It follows from equation (12.6) that the phase margins $30°$, $45°$, and $60°$ correspond to the slopes $-5/3$, $-3/2$, and $-4/3$, with a steeper slope giving smaller phase margin. Time delays and poles and zeros in the right half-plane impose further restrictions as will be discussed in Chapter 14.

Loop shaping is a trial-and-error procedure. We typically start with a Bode plot of the process transfer function. Choosing the gain crossover frequency ω_{gc} is a major design decision and is a compromise between attenuation of load disturbances and injection of measurement noise. Notice that the gain crossover frequency and the sensitivity crossover frequencies are the same if the phase margin is $\varphi_{\mathrm{m}} = 60°$, while for smaller phase margins we have $\omega_{\mathrm{gc}} < \omega_{\mathrm{sc}}$. Having determined the gain crossover frequency we then attempt to shape the loop transfer function by changing the controller gain and adding poles and zeros to the controller transfer function. As we shall see, the controller gain at low frequencies can be increased by so-called "lag compensation", and the behavior around the crossover frequency can be changed by so-called "lead compensation." Different performance specifications are evaluated for each controller as we attempt to balance many different requirements by adjusting controller parameters and complexity.

Loop shaping is straightforward to apply to single-input, single-output systems. It can also be applied to systems with one input and many outputs by closing the loops one at a time. The only limitation for minimum phase systems is that large phase leads and high controller gains may be required to obtain closed loop systems with a fast response. Many specific procedures are available: they all require experience, but they also give good insight into the conflicting specifications. There are fundamental limits to what can be achieved for systems that are not minimum phase; they will be discussed in Section 14.3.

Lead and Lag Compensation

A simple way to do loop shaping is to start with the transfer function of the process and add simple compensators with transfer function

$$C(s) = k\frac{s+a}{s+b}, \quad a > 0, b > 0. \qquad (12.7)$$

The compensator is called a *lead compensator* if $a < b$, and a *lag compensator* if $a > b$. The PI controller is a special case of a lag compensator with $b = 0$. A lead compensator is essentially the same as a PD controller with filtering. As described in Section 11.5, we often use a filter for the derivative action of a PID controller to limit the high-frequency gain. This same effect is present in a lead compensator through the pole at $s = b$. Equation (12.7) is a first-order compensator and can provide up to $90°$ of phase lead. Larger phase lead can be obtained by using a
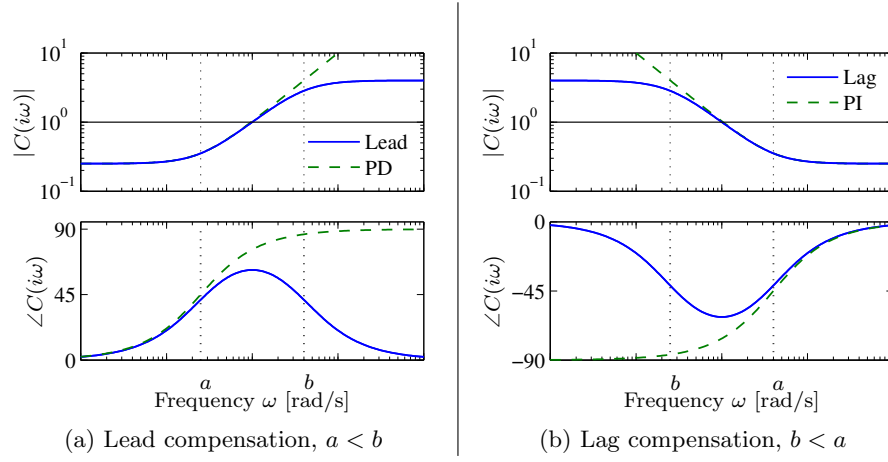
(a) Lead compensation, $a < b$ (b) Lag compensation, $b < a$

Figure 12.9: Frequency response for lead and lag compensators $C(s) = k(s + a)/(s + b)$. Lead compensation (a) occurs when $a < b$ and provides phase lead between $\omega = a$ and $\omega = b$. Lag compensation (b) corresponds to $a > b$ and provides low-frequency gain. PI control is a special case of lag compensation and PD control is a special case of lead compensation. PI/PD frequency responses are shown by dashed curves. The parameters are $a = 0.25$, $b = 4$, $k = 16$ in (a) and $a = 4$, $b = 0.25$, $k = 1$ in (b).

higher-order lead compensator (Exercise 12.16):

$$C(s) = k\frac{(s + a)^n}{(s + b)^n}, \qquad a < b.$$

Bode plots of lead and lag compensators are shown in Figure 12.9. Lag compensation, which increases the gain at low frequencies, is typically used to improve tracking performance and disturbance attenuation at low frequencies. Lead compensation is typically used to improve phase margin. If we set $a < b$ in equation (12.7), we add phase lead in the frequency range between the pole/zero pair (and extending approximately $10\times$ in frequency in each direction). By appropriately choosing the location of this phase lead, we can provide additional phase margin at the gain crossover frequency.

Lead compensation is associated with an increase of the high-frequency gain. Let $G(s)$ be a transfer function with $G(0) > 0$, with no poles and zeros in the right half plane, and assume that $\lim_{s\to\infty} G(s) = G(\infty) > 0$. Then

$$\log\frac{G(\infty)}{G(0)} = \frac{2}{\pi}\int_0^\infty \arg G(i\omega)\, d\log\omega = \frac{2}{\pi}\int_{-\infty}^\infty \arg G(ie^u)\, du. \tag{12.8}$$

This formula, which we call *Bode's phase area formula*, implies that the logarithm of the gain ratio $G(\infty)/G(0)$ for a transfer function is proportional to the area of the phase curve in the Bode plot. The equation was derived by Bode [Bod45, page 286] using the theory of complex variables. Lead compensation thus requires high gain at high frequencies and increases the sensitivity to measurement noise.

Lead and lag compensators can also be combined to form a lead-lag compensator (Exercise 12.11). Compensators that are tailored to specific disturbances can be
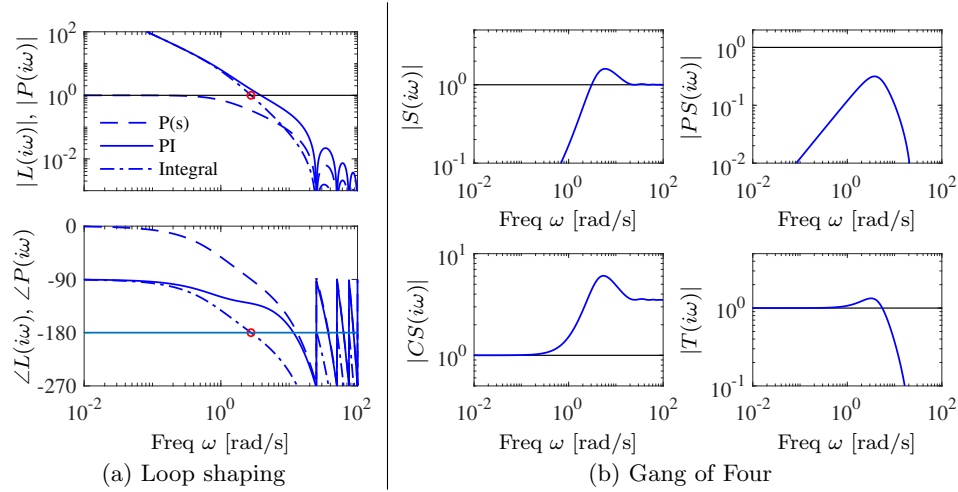
Figure 12.10: Loop-shaping design of a controller for an atomic force microscope in tapping mode. (a) Bode plots of the process (dashed), the loop transfer function for an integral controller with critical gain (dash-dotted) and a PI controller (solid) adjusted to give reasonable robustness. (b) Gain curves for the Gang of Four for the system.

also designed, as shown in Exercise 12.12. The following examples illustrate the use of lag compensation (via PI control) and lead compensation (to increase phase margin).

Example 12.4 Atomic force microscope in tapping mode
A simple model of the dynamics of the vertical motion of an atomic force microscope in tapping mode was given in Exercise 10.2. The transfer function for the system dynamics is

$$P(s) = \frac{a(1 - e^{-s\tau})}{s\tau(s + a)},$$

and the parameters $a = \zeta\omega_0$, $\tau = 2\pi n/\omega_0$ are explained in Example 11.2. The gain has been normalized to 1. A Bode plot of this transfer function for the parameters $a = 1$ and $\tau = 0.25$ is shown using dashed curves in Figure 12.10a. To improve the attenuation of load disturbances we increase the low-frequency gain by introducing an integral controller. The loop transfer function then becomes $L = k_i P(s)/s$, and we start by adjusting the gain k_i so that the closed loop system is marginally stable, giving $k_i = 8.3$. The Bode plot is shown by the dash-dotted line in Figure 12.10a, where the critical point is indicated by $\circ$. Notice the increase of the gain at low frequencies. To obtain a reasonable phase margin we introduce proportional action and we increase the proportional gain k_p gradually until reasonable values of the sensitivities are obtained. The value $k_p = 3.5$ gives maximum sensitivity $M_s = 1.6$ and maximum complementary sensitivity $M_t = 1.3$. The loop transfer function is shown in solid lines in Figure 12.10a. Notice the significant increase of the phase margin compared with the purely integral controller (dash-dotted line).

   To evaluate the design we also compute the gain curves of the transfer functions in the Gang of Four. They are shown in Figure 12.10b. The peaks of the sensitivity curves are reasonable, and the plot of PS shows that the largest value of PS is 0.3,
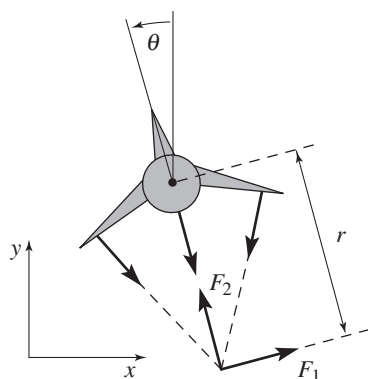
(a) Simplified model

Symbol	Description	Value
m	Vehicle mass	4.0 kg
J	Vehicle inertia, φ_3 axis	0.0475 kg m^2
r	Force moment arm	25.0 cm
c	Damping coefficient	0.05 kg m/s
g	Gravitational constant	9.8 m/s^2

(b) Parameter values

Figure 12.11: Roll control of a vectored thrust aircraft. (a) The roll angle θ is controlled by applying maneuvering thrusters, resulting in a moment generated by F_1. (b) The table lists the parameter values for a laboratory version of the system.

which implies that the load disturbances are well attenuated. The plot of CS shows that the largest noise gain $|C(i\omega)S(i\omega)|$ is 6. The controller has a gain $k_{\mathrm{p}} = 3.5$ at high frequencies, and hence we may consider adding high-frequency roll-off to make CS smaller at high frequencies. ▽

Example 12.5 Roll control for a vectored thrust aircraft
Consider the control of the roll of a vectored thrust aircraft such as the one illustrated in Figure 12.11. Following Exercise 9.11, we model the system with a second-order transfer function of the form

$$P(s) = \frac{r}{Js^2},$$

with the parameters given in Figure 12.11b. We take as our performance specification that we would like less than 1% error in steady state and less than 10% tracking error up to 10 rad/s.

The open loop transfer function from F_1 to θ is shown in Figure 12.12a. To achieve our performance specification, we would like to have a gain of at least 10 at a frequency of 10 rad/s, requiring the gain crossover frequency to be at a higher frequency. We see from the loop shape that in order to achieve the desired performance we cannot simply increase the gain since this would give a very low phase margin. Instead, we must increase the phase at the desired crossover frequency.

To accomplish this, we use a lead compensator (12.7) with $a = 2$, $b = 50$, and $k = 200$. We then set the gain of the system to provide a large loop gain up to the desired bandwidth, as shown in Figure 12.12b. We see that this system has a gain of greater than 10 at all frequencies up to 10 rad/s and that it has more than 60° of phase margin. ▽
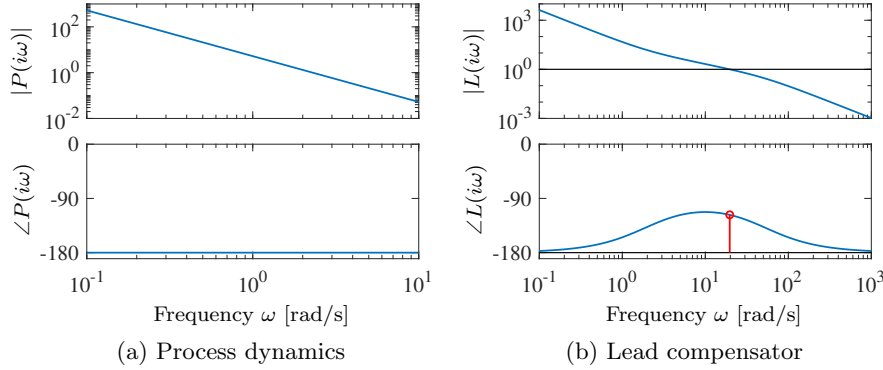
(a) Process dynamics   (b) Lead compensator

Figure 12.12: Control design for a vectored thrust aircraft using lead compensation. The Bode plot for the open loop process P is shown in (a) and (b) shows the Bode plot for the loop transfer function $L = PC$, where C is the lead given by equation (12.7) with $a = 2$, $b = 50$, and $k = 200$. Note the phase lead in the crossover region near $\omega = 20$ rad/s.

12.4   FEEDFORWARD DESIGN

Feedforward is a simple and powerful technique that complements feedback. It can be used both to improve the response to reference signals and to reduce the effect of measurable disturbances. Design of feedforward for controllers based on state feedback and observers was developed in Section 8.5 (Figure 8.11). Section 11.5 presented setpoint weighting as simple form of feedforward for PID controllers (equation (11.15)). In this section we will use transfer functions to develop more advanced methods for feedforward design.

Combining Feedforward and Feedback

Figure 12.13 shows a block diagram of a system with feedback and feedforward control. The process dynamics are separated into two blocks $P_1(s)$ and $P_2(s)$, where the measured disturbance v enters at the input of the block P_2, and we define $P(s) = P_1(s)P_2(s)$. The transfer function F_{m} represents the desired (model) response to reference signals. There are two feedforward blocks with the transfer functions F_r and F_v to deal with the reference signal r and the measured disturbances v.

A major advantage of controllers with two degrees of freedom that combine feedback and feedforward is that the control design problem can be split in two parts. The feedback transfer function C can be designed to give good robustness and effective disturbance attenuation, and the feedforward transfer functions F_r and F_v can be designed independently to give the desired responses to reference signals and to reduce effects of measured disturbances.

We will first explore the response to reference signals. The transfer function $G_{yr}(s)$ from reference input r to process output y in Figure 12.13 is

$$\begin{aligned} G_{yr}(s) &= \frac{P(CF_{\mathrm{m}} + F_r)}{1 + PC} = F_{\mathrm{m}} + \frac{PF_r - F_{\mathrm{m}}}{1 + PC} \\ &= F_{\mathrm{m}} + S(PF_r - F_{\mathrm{m}}) = TF_{\mathrm{m}} + SPF_r, \end{aligned} \tag{12.9}$$
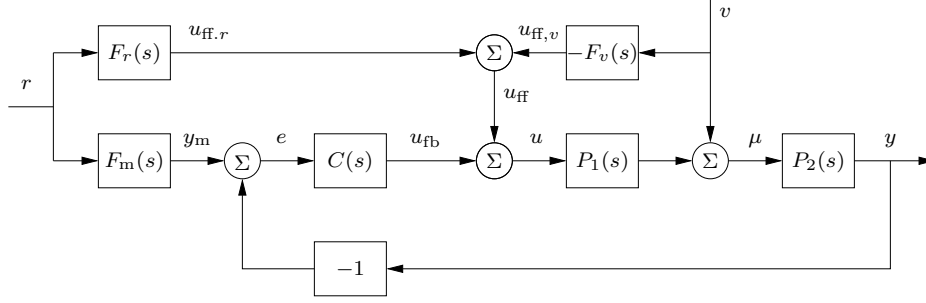
Figure 12.13: Block diagram of a system with feedforward compensation for improved response to reference signals and measured disturbances (2 degree-of-freedom system). Three feedforward elements are present: $F_\mathrm{m}(s)$ sets the desired output value, $F_r(s)$ generates the feedforward command u_fr to improve reference signal response and $F_v(s)$ generates the feedforward signal u_fv that reduces the effect of the measured disturbance v.

where S is the sensitivity function and T the complementary sensitivity function (equation (12.2)). We can make G_{yr} close to the desired transfer function F_m in two different ways: by choosing the feedforward transfer function F_r so that $PF_r - F_\mathrm{m}$ is small, or by choosing the feedback transfer function C so that the sensitivity $S = 1/(1 + PC)$ is small. Perfect feedforward compensation is obtained by choosing

$$F_r = \frac{F_\mathrm{m}}{P_1 P_2} = \frac{F_\mathrm{m}}{P}, \tag{12.10}$$

which gives $G_{yr} = F_\mathrm{m}$. Notice that the feedforward compensator F_r contains an inverse model of the process dynamics.

Next we will consider attenuation of disturbances that can be measured. The transfer function from load disturbance v to process output y is given by

$$G_{yv} = \frac{P_2(1 - P_1 F_v)}{1 + PC} = P_2 S(1 - P_1 F_v). \tag{12.11}$$

The transfer function G_{yv} can be made small in two different ways: by choosing the feedforward transfer function F_v so that $1 - P_1 F_v$ is small, or by choosing the feedback transfer function C so that the sensitivity $S = 1/(1+PC)$ is small. Perfect compensation is obtained by choosing

$$F_v = \frac{1}{P_1}. \tag{12.12}$$

Design of feedforward to improve responses to reference signals and disturbances using transfer functions is thus a simple task, but it requires inversion of process models. We illustrate with an example.

Example 12.6 Vehicle steering
A linearized model for vehicle steering was given in Example 7.4. The normalized transfer function from steering angle δ to lateral deviation y is $P(s) = (\gamma s + 1)/s^2$. For a lane transfer system we would like to have a nice response without overshoot, and we therefore choose the desired response as $F_\mathrm{m}(s) = \omega_\mathrm{c}^2/(s + \omega_\mathrm{c})^2$, where the

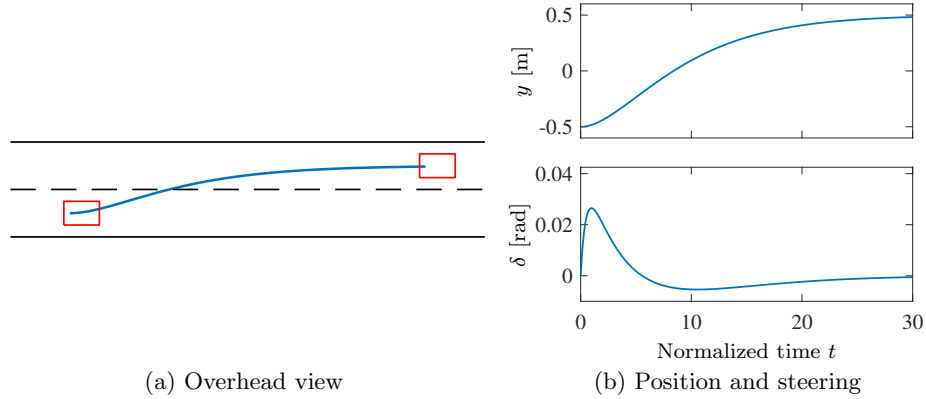(a) Overhead view   (b) Position and steering

Figure 12.14: Feedforward control for vehicle steering. The plot on the left shows the trajectory generated by the controller for changing lanes. The plots on the right show the lateral deviation y (top) and the steering angle δ (bottom) for a smooth lane change control using feedforward (based on the linearized model).

response speed or aggressiveness of the steering is governed by the parameter ω_c. Equation (12.10) gives

$$F_r = \frac{F_m}{P} = \frac{\omega_c^2 s^2}{(\gamma s + 1)(s + \omega_c)^2},$$

which is a stable transfer function as long as $\gamma > 0$. Figure 12.14 shows the responses of the system for $\omega_c = 0.2$. The figure shows that a lane change is accomplished in about 20 vehicle lengths with smooth steering angles. The largest steering angle is slightly larger than 0.2 rad ($12°$). Using the scaled variables, the curve showing lateral deviations (y as a function of t) can also be interpreted as the vehicle path (y as a function of x) with the vehicle length as the length unit. ∇

Difficulties with Feedforward

The ideal feedforward compensators for Figure 12.13 are given by

$$F_r = \frac{F_m}{P_1 P_2}, \qquad F_v = \frac{1}{P_1}. \tag{12.13}$$

Both transfer functions require inversion of process transfer functions and there can be problems with inversion if the process transfer function has time delays, right half-plane zeros, or high pole excess. Inversion of time delays requires prediction, which cannot be done perfectly except in the situation when the command signal is known in advance. If the process transfer functions has zeros in the right half-plane, the inverse process transfer function is unstable and approximate inverses may have to be used. Finally, if the pole excess of the process transfer function is greater than zero, then the inverse requires differentiation. In this case the reference signal must then be sufficiently smooth and there may also be problems with noise.

There is some extra freedom when finding the transfer function F_r because it also contains the transfer function F_m, which specifies the ideal behavior. A stable
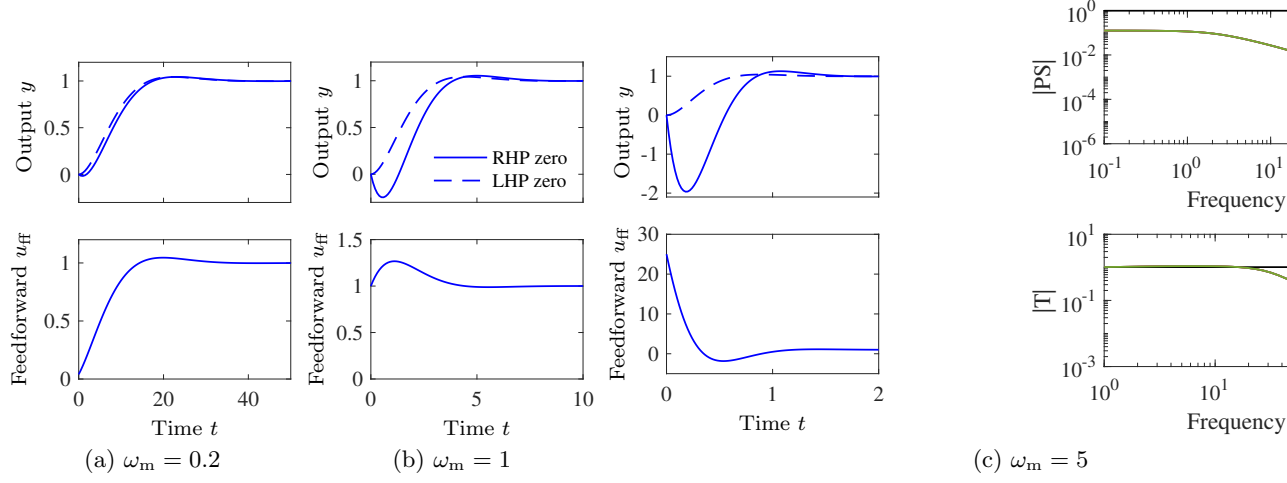
Figure 12.15: Outputs y (upper plots) and feedforward signals u_{ff} (lower plots) for a unit step command signal. The design parameter has the values $\omega_m = 0.2$, 1, and 0.5 for a unit step command in the reference signal. The dashed curve shows the response that could be achieved if the process did not have the right half-plane zero.

feedforward transfer function can be obtained if F_m has the same time delays and right half-plane zeros as the process. We illustrate with an example.

Example 12.7 Feedforward for a process with a right half-plane zero
Let the process and the desired response have the transfer functions

$$P(s) = \frac{1-s}{(s+1)^2}, \qquad F_m(s) = \frac{\omega_m^2(1-s)}{s^2 + 2\zeta_c\omega_m s + \omega_m^2}.$$

Since the process has a right half-plane zero at $s = 1$, the desired transfer function $F_m(s)$ must have the same zeros to avoid having an unstable feedforward transfer function F_r. Equation (12.10) gives the feedforward transfer function:

$$F_r(s) = \frac{\omega_m^2(s+1)^2}{s^2 + 2\zeta_c\omega_m s + \omega_m^2}. \tag{12.14}$$

Figure 12.15 shows the outputs y and the feedforward signals u_{ff} for different values of ω_m. The response to the reference signal goes in the wrong direction initially because of the right half-plane zero at $s = 1$. This effect, called *inverse response*, is barely noticeable if the response is slow ($\omega_m = 1$) but it increases with increasing response speed. For $\omega_m = 5$ the undershoot is more than 200%. The large undershoot is an indication that a right half-plane zero limits the achievable bandwidth, as will be discussed in depth in Chapter 14. A reasonable choice of ω_m is in the range 0.2 to 0.5. Notice that the same feedforward transfer function (12.14) is obtained if the process and the desired model have the transfer functions

$$P(s) = \frac{1}{(s+1)^2}, \qquad F_m(s) = \frac{\omega_m^2}{s^2 + 2\zeta_c\omega_m s + \omega_m^2}.$$

The corresponding responses are shown as dashed lines in Figure 12.15. When there is no right half-plane zero it is thus possible to obtain well-behaved, fast responses.

The control signals for different values of ω_{m} differ significantly, as shown in the bottom row of plots in Figure 12.15. Since $r = 1$ and the zero frequency gain of the feedforward transfer function is $F_{\mathrm{r}}(0) = 1$, the control signal goes to 1 as time goes to infinity in all cases. The feedforward transfer function also has constant gain $F_{\mathrm{r}}(\infty) = \omega_{\mathrm{m}}^2$ for high frequencies, which means that gain for high-frequency signals is ω_{m}^2 and this can be undesirable if ω_{m} is large. The initial response to a unit step signal is then $u_{\mathrm{ff}}(0) = F_{\mathrm{r}}(\infty) = \omega_{\mathrm{m}}^2$ (using the initial value theorem). For $\omega_{\mathrm{m}} = 0.2$ the control signal grows from 0.04 to the final value 1 with a small overshoot. For $\omega_{\mathrm{m}} = 1$ the control signal starts from 1, has an overshoot and then settles on the final value 1. For $\omega_{\mathrm{m}} = 5$ the control signal starts at 25 and decays towards the final value 1 with an undershoot. ∇

Approximate Inverses

Processes with right half-plane zeros do not have stable inverses. To design feedforward compensators for such systems we need to use approximate inverses that are stable. The following theorem, which is presented without proof, provides a means of constructing such approximate inverses.

Theorem 12.1 (Approximate inverse). *Let the rational transfer function $G(s)$ have all its poles in the left half-plane and no zeros on the imaginary axis. Factor the transfer function as $G(s) = G^+(s)G^-(s)$, where $G^+(s)$ has all its zeros in the left half-plane and $G^-(s)$ has all its zeros in the right half-plane. An approximate stable inverse of $G(s)$ that minimizes the mean square error for a step input is*

$$G^\dagger(s) = \frac{1}{G^+(s)G^-(-s)}. \tag{12.15}$$

We illustrate the theorem with an example.

Example 12.8 Approximate inverse for a system with a right half-plane zero

Let the transfer functions of the process and the reference model (desired response) be

$$P(s) = \frac{1 - s}{(s + 1)^2}, \qquad F_{\mathrm{m}}(s) = \frac{\omega_m^2}{s^2 + 2\zeta_c\omega_m s + \omega_m^2}.$$

Note that in comparison to Example 12.7, we do not include the right half-plane zero in F_{m}. The process transfer function can be factored as

$$P^-(s) = 1 - s, \quad P^+(s) = \frac{1}{(s + 1)^2}.$$

Theorem 12.1 then gives the following approximate inverse:

$$P^\dagger(s) = \frac{1}{P^+(s)P^-(-s)} = \frac{(s + 1)^2}{1 + s} = s + 1.$$
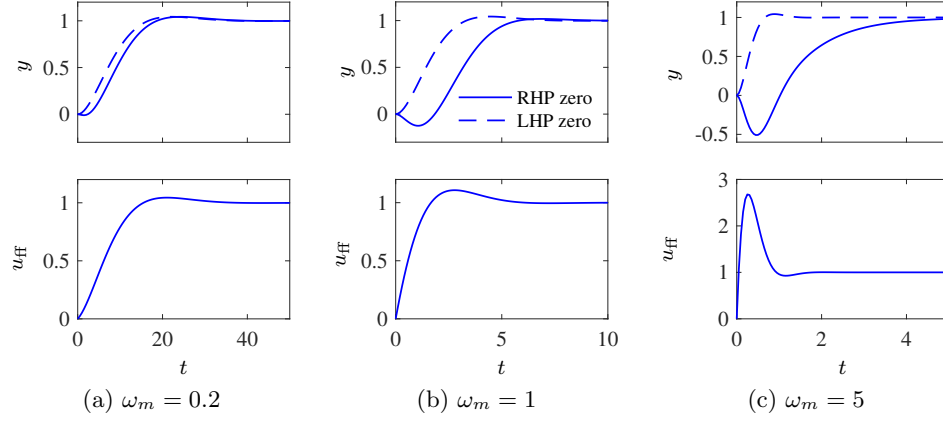
Figure 12.16: Feedforward design based on an approximate inverse. Outputs y (upper plots) and feedforward signals u_{ff} (lower plots) for a unit step command signal. The design parameter has the values $\omega_m = 0.2$, 1, and 0.5 for a unit step command in the reference signal. The dashed curves show the responses for a process without the right half-plane zero.

The feedforward transfer function is then

$$F_r(s) = F_{\text{m}}(s)P^{\dagger}(s) = \frac{\omega_m^2(s+1)}{s^2 + 2\zeta_c\omega_m s + \omega_m^2},$$

which is similar to equation (12.14) but no longer relies on cancellation of the right half-plane zero to obtain a stable feedforward transfer function. The transfer function from reference r to output y is then

$$G_{yr}(s) = P(s)F_r(s) = \frac{1-s}{(s^2 + 2\zeta_c\omega_m s + \omega_m^2)(s+1)}.$$

Figure 12.16 shows the step responses for different values of ω_m.

Comparing Figures 12.15 and 12.16 we find that there are small differences for $\omega_m = 0.2$, but large differences for $\omega_m = 5$. Notice in particular the shapes of the feedforward signals u_{ff}. The design based on the approximate inverse has smaller undershoot but the time responses have somewhat longer settling times. ∇

In summary, we see that feedforward can be used to improve the response to reference signals and to reduce the effects of load disturbances that can be measured. There are limits if the process has time delays, right half-plane zeros, or high pole excess. The zeros depend on the sensors and we can change them by moving or adding sensors. In addition, we will see in Chapter 13 that feedforward controllers can be sensitive to model uncertainty (Section 13.3 and Exercise 13.7), and hence feedforward control is usually combined with feedback control to obtain robust performance.

12.5   THE ROOT LOCUS METHOD

In design methods such as eigenvalue assignment, discussed in Sections 7.2 and 8.3, we designed controllers that give desired closed loop poles. The controllers were sufficiently complex so that all closed loop poles could be specified. The complexity of the controller is thus directly related to the complexity of the process. In practice we may have to use a simple controller for a complex process, and it is then not possible to find a controller that gives all closed poles their desired values. It is interesting to explore what can be done with a controller having restricted complexity as was the case for PID control in Chapter 11 and loop shaping in Section 12.3. The simplest case with only one selectable controller parameter can be investigated with the *root locus method*. The *root locus* is a graph of the roots of the characteristic polynomial as a function of a parameter, and the method gives insight into the effects of the controller parameter. It is straightforward to obtain the root locus by finding the roots of the closed-loop characteristic polynomial for different values of the parameter. There are also good computer tools for generating the root locus. Of greater interest is the fact that the general shape of the root locus can be obtained with very little effort, and that it often gives considerable insight.

To illustrate the root locus method we consider a process with the transfer function

$$P(s) = \frac{b(s)}{a(s)} = \frac{b_0 s^m + b_1 s^{m-1} + \cdots b_m}{s^n + a_1 s^{n-1} + \cdots a_n} = b_0 \frac{(s - z_1)(s - z_2) \ldots (s - z_m)}{(s - p_1)(s - p_2) \cdots (s - p_n)}.$$

The polynomial $a(s)$ has degree n and the polynomial $b(s)$ has degree m. We assume that pole excess $n_{\mathrm{pe}} = n - m$ is positive or zero. The controller is assumed to be a proportional controller with the transfer function $C(s) = k$. We will explore the poles of the closed loop system when the gain k of the proportional controller ranges from 0 to ∞.

The closed loop characteristic polynomial is

$$a_{\mathrm{cl}}(s) = a(s) + kb(s) \tag{12.16}$$

and the closed loop poles are the roots of $a_{\mathrm{cl}}(s)$. The root locus is a graph of the roots of $a_{\mathrm{cl}}(s)$ as the gain k is varied from 0 to ∞. Since the polynomial $a_{\mathrm{cl}}(s)$ has degree n, the plot will have n branches.

When the gain k is zero we have $a_{\mathrm{cl}}(s) = a(s)$ and the closed loop poles are equal to the open loop poles. When there are open loop poles at $s = p_l$ with multiplicity m, the characteristic equation can be written as

$$(s - p_l)^m \tilde{a}(s) + kb(s) \approx (s - p_l)^m \tilde{a}(p_l) + kb(p_l) = 0,$$

where $\tilde{a}(s)$ represents the polynomial $a(s)$ with the poles at $s = p_l$ factored out. For small values of k the roots of this equation are given by $s = p_l + \sqrt[m]{-kb(p_l)/\tilde{a}(p_l)}$. The root locus thus has a star pattern with m branches emanating from the open loop pole $s = p_l$. The angle between two neighboring branches is $2\pi/m$.

To explore what happens for large gain we approximate the characteristic poly-
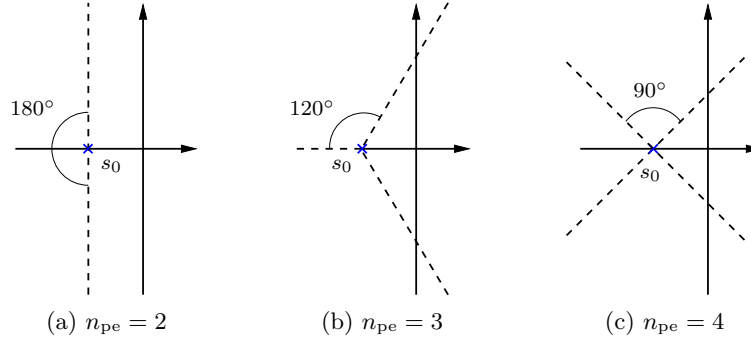
(a) $n_{\mathrm{pe}} = 2$ (b) $n_{\mathrm{pe}} = 3$ (c) $n_{\mathrm{pe}} = 4$

Figure 12.17: Asymptotes of of root locus for systems with pole excess $n_{\mathrm{pe}} = 2$, 3, and 4. There are n_{pe} asymptotes radiate from the point $s = s_0$ given by equation (12.18), and the angles between the asymptotes are $360°/n_{\mathrm{pe}}$.

nomial (12.16) for large s and k, which gives

$$a_{\mathrm{cl}}(s) = b(s)\left(\frac{a(s)}{b(s)} + k\right) \approx b(s)\left(\frac{s^{n_{\mathrm{pe}}}}{b_0} + k\right). \tag{12.17}$$

For large k the closed loop poles are approximately the roots of $b(s)$ and $\sqrt[n_{\mathrm{pe}}]{-b_0 k}$. A better approximation of equation (12.17) is

$$s = s_0 + \sqrt[n_{\mathrm{pe}}]{-kb_0}, \qquad s_0 = \frac{1}{n_{\mathrm{pe}}}\left(\sum_{k=1}^{n} p_k - \sum_{k=1}^{m} z_k\right) \tag{12.18}$$

(Exercise 12.14). The asymptotes are thus n_{pe} lines that radiate from $s = s_0$, the center of mass of poles and zeros. When $b_0 k > 0$ the lines have the angles $(\pi + 2l\pi)/n_{\mathrm{pe}}$, $l = 1, \cdots, n_{\mathrm{pe}}$ with respect to the real line. Figure 12.17 shows the asymptotes of the root locus for large gain for different values of the pole excess n_{pe}.

Summarizing, we find that the root locus plot with the loop gain as the varied parameter has n branches that start at the open loop poles and end either at the open loop zeros or at infinity. The branches that end at infinity have star-patterned asymptotes given by equation (12.18). An immediate consequence is that open loop systems with right half-plane zeros or a pole excess larger than 2 will always be unstable for sufficiently large gains.

There are simple rules for sketching the root locus. We describe here a few of them. As discussed already, the root locus has a (locally) symmetric star pattern at points where there are multiple roots; the number of branches depend on the multiplicity of the roots. For systems with $kb_0 > 0$ the root locus has segments on the real line where there are odd numbers of real poles and zeros to the right of the segment (Exercise 12.15). It is also straightforward to find directions where a branch of the root locus leaves a pole, as discussed in Exercise 12.18.
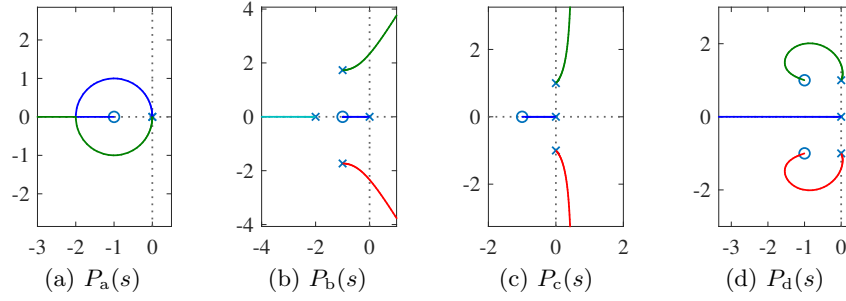
(a) $P_a(s)$ (b) $P_b(s)$ (c) $P_c(s)$ (d) $P_d(s)$

Figure 12.18: Examples of root loci for processes with the transfer functions $P_a(s)$, $P_b(s)$, $P_c(s)$, and $P_d(s)$ given by equation (12.19).

Figure 12.18 shows root loci for systems with $k > 0$ and the transfer functions

$$P_a(s) = k\,\frac{s+1}{s^2}, \qquad P_b(s) = k\,\frac{s+1}{s(s+2)(s^2+2s+4)},$$
$$P_c(s) = k\,\frac{s+1}{s(s^2+1)}, \qquad P_d(s) = k\,\frac{s^2+2s+2}{s(s^2+1)}. \tag{12.19}$$

The locus of $P_a(s)$ in Figure 12.18a starts with two roots at the origin and the pattern locally has the star configuration with $n^* = 2$. As the gain increases the locus bends because of the attraction of the zero. In this particular case the locus is actually a circle around the zero $s = -1$. Two roots meet at the real axis and depart forming a star pattern. One root goes towards the zero and the other one goes to infinity along the negative real axis as the gain k increases. The root locus thus has the segment $(-\infty, -1]$ on the real axis. The locus in Figure 12.18b starts at the open loop poles $s = -2$, 0, and $-1 \pm i\sqrt{3}$. The pole excess is $n_{\mathrm{pe}} = 3$ and the asymptotes that originate from $s_0 = -1$ have the corresponding pattern. The locus in Figure 12.18c has vertical asymptotes since $n_{\mathrm{pe}} = 2$ (see Figure 12.17). The asymptotes originate from $s_0 = 0.5$. The root locus has the segment $[-1, 0]$ on the real line. The locus in Figure 12.18d has three branches: one is the segment $(-\infty, 0]$ on the real line and the other two originate on the complex open loop poles and end at the open loop zeros.
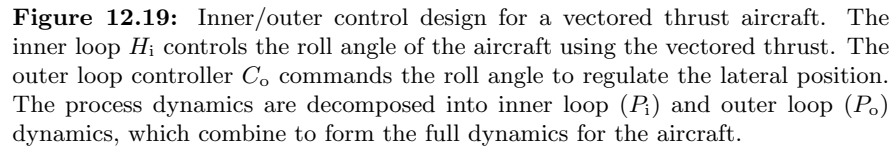
The root locus can also be used for design. Consider for example the system in Figure 12.18c, which can represent PI control of a system with the transfer functions

$$P(s) = \frac{1}{s^2+1}, \qquad C(s) = k\,\frac{s+2}{s}.$$

The root locus in Figure 12.18c shows that the system is unstable for all values of the controller gain and we can immediately conclude that the process cannot be stabilized with a PI controller. To obtain a stable closed loop system we can attempt to choose a PID controller with zeros to the left of the undamped poles, for example

$$C(s) = k\,\frac{s^2+2s+2}{s}.$$

The root locus obtained with this controller is shown in Figure 12.18d. We see that this system is stable for $k > 0$ and we can choose k to place the poles in reasonable

Figure 12.19: Inner/outer control design for a vectored thrust aircraft. The inner loop H_i controls the roll angle of the aircraft using the vectored thrust. The outer loop controller C_o commands the roll angle to regulate the lateral position. The process dynamics are decomposed into inner loop (P_i) and outer loop (P_o) dynamics, which combine to form the full dynamics for the aircraft.

.

locations.

We have illustrated the root locus with a closed loop system with a proportional controller where the parameter is the gain. The root locus can also be used to find the effects of other parameters, as was illustrated in Example 5.17.

12.6   DESIGN EXAMPLE

In this final section we present a detailed example that illustrates the main design techniques described in this chapter.

Example 12.9 Lateral control of a vectored thrust aircraft
The problem of controlling the motion of a vertical takeoff and landing (VTOL) aircraft was introduced in Example 3.12 and in Example 12.5, where we designed a controller for the roll dynamics. We now wish to control the position of the aircraft, a problem that requires stabilization of the attitude.

To control the lateral dynamics of the vectored thrust aircraft, we make use of an "inner/outer" loop design methodology, as illustrated in Figure 12.19. This diagram shows the process dynamics and controller divided into two components: an *inner loop* consisting of the roll dynamics and controller and an *outer loop* consisting of the lateral position dynamics and controller. This decomposition follows the block diagram representation of the dynamics given in Exercise 9.11.

The approach that we take is to design a controller C_i for the inner loop so that the resulting closed loop system H_i assures that the roll angle θ follows its reference θ_r quickly and accurately. We then design a controller for the lateral position y that uses the approximation that we can directly control the roll angle as an input θ to the dynamics controlling the position. Under the assumption that the dynamics of the roll controller are fast relative to the desired bandwidth of the lateral position control, we can then combine the inner and outer loop controllers to get a single controller for the entire system. As a performance specification for the entire system, we would like to have zero steady-state error in the lateral position,
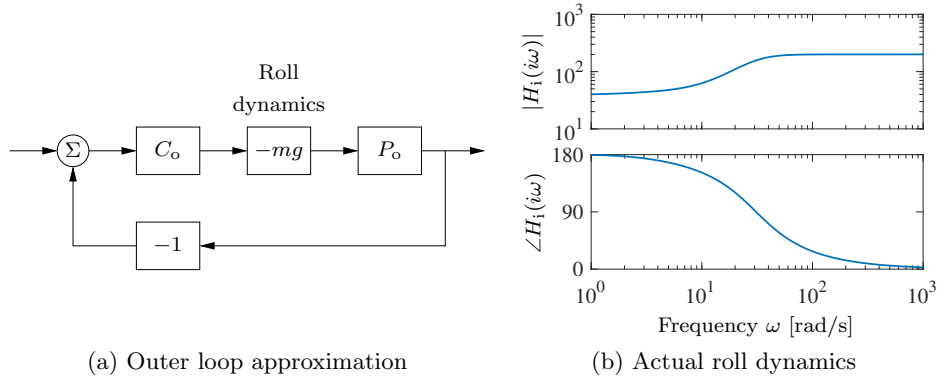
(a) Outer loop approximation   (b) Actual roll dynamics

Figure 12.20: Outer loop control design for a vectored thrust aircraft. (a) The outer loop approximates the roll dynamics as a state gain $-mg$. (b) The Bode plot for the roll dynamics, indicating that this approximation is accurate up to approximately 10 rad/s.

a bandwidth of approximately 1 rad/s, and a phase margin of 45°.

For the inner loop, we choose our design specification to provide the outer loop with accurate and fast control of the roll. The inner loop dynamics are given by

$$P_i(s) = H_{\theta u_1}(s) = \frac{r}{Js^2}.$$

We choose the desired bandwidth to be 10 rad/s (10 times that of the outer loop) and the low-frequency error to be no more than 5%. This specification is satisfied using the lead compensator of Example 12.5 designed previously, so we choose

$$C_i(s) = k\frac{s+a}{s+b}, \qquad a = 2, \quad b = 50, \quad k = 200.$$

The closed loop dynamics for the system satisfy

$$H_i = \frac{C_i}{1 + C_i P_i} - mg\frac{C_i P_i}{1 + C_i P_i} = \frac{C_i(1 - mgP_i)}{1 + C_i P_i}.$$

A plot of the magnitude of this transfer function is shown in Figure 12.20, and we see that $H_i \approx -mg = -39.2$ is a good approximation up to 10 rad/s.

To design the outer loop controller, we assume the inner loop roll control is perfect, so that we can take θ_r as the input to our lateral dynamics. Following the diagram shown in Exercise 9.11, the outer loop dynamics can be written as

$$P(s) = H_i(0)P_o(s) = \frac{H_i(0)}{ms^2 + cs},$$

where we replace $H_i(s)$ with $H_i(0)$ to reflect our approximation that the inner loop will eventually track our commanded input. Of course, this approximation may not be valid, and so we must verify this when we complete our design.

Our control goal is now to design a controller that gives zero steady-state error in y for a step input and has a bandwidth of 1 rad/s. The outer loop process dynamics are given by a double integrator, and we can again use a simple lead
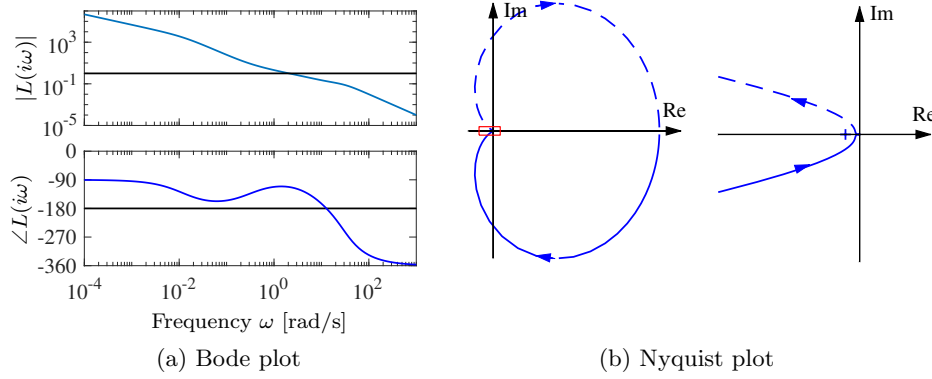
(a) Bode plot   (b) Nyquist plot

Figure 12.21: Inner/outer loop controller for a vectored thrust aircraft. Bode plot (a) and Nyquist plot (b) for the loop transfer function cut at θ_r, for the complete system. The system has a phase margin of $68°$ and a gain margin of 6.2.

compensator to satisfy the specifications. We also choose the design such that the loop transfer function for the outer loop has $|L_o| < 0.1$ for $\omega > 10$ rad/s, so that the H_i high-frequency dynamics can be neglected. We choose the controller to be of the form

$$C_o(s) = -k_o \frac{s + a_o}{s + b_o},$$

with the negative sign to cancel the negative sign in the process dynamics. To find the location of the poles, we note that the phase lead flattens out at approximately $b_o/10$. We desire phase lead at crossover, and we desire the crossover at $\omega_{gc} = 1$ rad/s, so this gives $b_o = 10$. To ensure that we have adequate phase lead, we must choose a_o such that $b_o/10 < 10a_o < b_o$, which implies that a_o should be between 0.1 and 1. We choose $a_o = 0.3$. Finally, we need to set the gain of the system such that at the desired crossover frequency the loop gain has magnitude 1 or more. A simple calculation shows that $k_o = 2$ satisfies this objective. Thus, the final outer loop controller becomes

$$C_o(s) = -2\frac{s + 0.3}{s + 10}.$$

Finally, we can combine the inner and outer loop controllers and verify that the system has the desired closed loop performance. The Bode and Nyquist plots corresponding to Figure 12.19 with inner and outer loop controllers are shown in Figure 12.21, and we see that the specifications are satisfied. In addition, we show the gain curves of the Gang of Four in Figure 12.22, and we see that the transfer functions between all inputs and outputs are reasonable. The sensitivity to load disturbances PS is large at low frequency because the controller does not have integral action.

The approach of splitting the dynamics into an inner and an outer loop is common in many control applications and can lead to simpler designs for complex systems. Indeed, for the aircraft dynamics studied in this example, it is very challenging to directly design a controller from the lateral position y to the input u_1. The use of the additional measurement of θ greatly simplifies the design because it
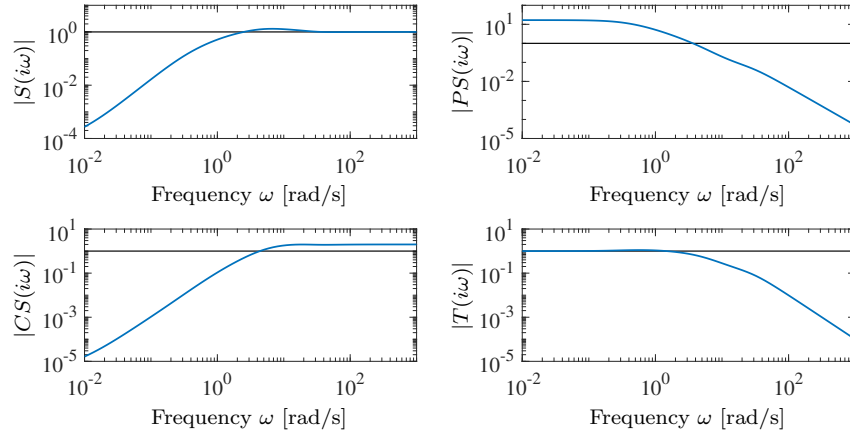
Figure 12.22: Gain curves for the Gang of Four for the vectored thrust aircraft system.

can be broken up into simpler pieces.   ∇

12.7   FURTHER READING

Loop shaping design emerged at Bell Labs in connection with the development of Black's [Bla34] electronic amplifier with negative feedback. Nyquist [Nyq32] derived his stability criterion to understand and avoid instabilities or "singing," as it was called at the time. Bode [Bod40] used the theory of complex variables to establish important fundamental results such as the relation between amplitude and phase for a minimum phase system, the ideal loop transfer functions, and the phase area formula. His results are nicely summarized in the book [Bod45]. Design by loop shaping became a key element in the early development of control, and many design methods were developed; see James, Nichols and Phillips [JNP47], Chestnut and Mayer [CM51], Truxal [Tru55], and Thaler [Tha89]. Loop shaping is also treated in standard textbooks such as Franklin, Powell, and Emami-Naeini [FPEN05], Dorf and Bishop [DB04], Kuo and Golnaraghi [KG02], and Ogata [Oga01]. Horowitz [Hor63] developed the notion of systems with two degrees of freedom. Much of the early work was based on the loop transfer function; the importance of the sensitivity functions appeared in connection with developments in the 1980s that resulted in H_∞ design methods. A compact presentation is given in the texts by Doyle, Francis, and Tannenbaum [DFT92] and Zhou, Doyle, and Glover [ZDG96]. Loop shaping was integrated with the robust control theory in McFarlane and Glover [MG90] and Vinnicombe [Vin01]. Comprehensive treatments of control system design are given in Maciejowski [Mac89] and Goodwin, Graebe, and Salgado [GGS01]. There are fundamental limits to what can be achieved given by nonlinearities of the process and the poles and zeros. These will be discussed in Chapter 14.

EXERCISES

12.1 Consider the system in Example 12.1, where the process and controller transfer functions are given by

$$P(s) = 1/(s-a), \qquad C(s) = k(s-a)/s.$$

Choose the parameter $a = -1$ and compute the time (step) and frequency responses for all the transfer functions in the Gang of Four for controllers with $k = 0.2$ and $k = 5$.

12.2 (Equivalence of Figures 12.1 and 12.2) Consider the system in Figure 12.1 and let the outputs of interest be $\xi = (\mu, \eta)$ and the major disturbances be $\chi = (w, v)$. Show that the system can be represented by Figure 12.2 and give the matrix transfer functions $\mathcal{P}$ and $\mathcal{C}$. Verify that the elements of the closed loop transfer function $H_{\xi\chi}$ are the Gang of Four.

12.3 Consider the spring–mass system given by equation (3.16), which has the transfer function

$$P(s) = \frac{1}{ms^2 + cs + k}.$$

Design a feedforward compensator that gives a response with critical damping ($\zeta = 1$).

12.4 (Equivalence of controllers with two degrees of freedom) Show that the systems in Figures 12.1 and 12.13 give the same responses to command signals if $F_{\mathrm{m}}C + F_u = CF$.

12.5 (Web server control) Feedback and feedforward are increasingly used for complex computer systems such as web servers. Control of a single server is an example. A model for a virtual server is given by equation (3.32),

$$\frac{dx}{dt} = \lambda - \mu,$$

where x is the queue length, λ is the arrival rate, and μ is the server rate. The objective of control is to maintain a given queue length. The service rate μ can be changed by dynamic voltage and frequency scaling (DVFS). Determine a PI controller that gives a closed loop system with the characteristic polynomial $s^2 + 4s + 4$. Use feedforward in the form of setpoint weighting to reduce the overshoot for step changes in reference signals; simulate the closed loop system to determine the setpoint weighting.

12.6 (Rise time-bandwidth product)   Consider a stable system with the transfer function $G(s)$ whose frequency response is an ideal low-pass filter with $|G(i\omega)| = 1$ for $\omega \le \omega_{\mathrm{b}}$ and $|G(i\omega)| = 0$ for $\omega > \omega_{\mathrm{b}}$ and which has low-pass character. Define the rise time T_{r} as the inverse of the largest slope of the unit step response and the bandwidth as $\widetilde{\omega}_{\mathrm{b}} = \int_0^\infty |G(i\omega)|/G(0)\, d\omega$. Show that with this definition of the bandwidth the *rise time-bandwidth product* satisfies $T_{\mathrm{r}}\,\widetilde{\omega}_{\mathrm{b}} \ge \pi$.

12.7 (Peak frequency-peak time product) Consider the transfer function for a sec-

ond order system
$$G(s) = \frac{\omega_0 s}{s^2 + 2\zeta\omega_0 s + \omega_0^2},$$

which has the unit step response

$$y(t) = \frac{1}{\sqrt{1-\zeta^2}} e^{-\zeta\omega_0 t} \sin \omega_0 t \sqrt{1-\zeta^2}.$$

Let $M_\mathrm{p} = \max_\omega |G(i\omega)|$ be the largest gain of $G(s)$, which is assumed to occur at ω_p, and let $y_\mathrm{p} = \max_t y(t)$ be the largest value of $y(t)$, which is assumed to occur at t_p. Show that

$$t_p \omega_p = \frac{\arccos \zeta}{\sqrt{1-\zeta^2}}, \qquad \frac{y_p}{M_\mathrm{p}} = 2\zeta e^{-\zeta\varphi},$$

and evaluate the right-hand sides of the above equations for $\zeta = 0.5$, 0.707, and 1.0.

12.8 (Disturbance attenuation) Consider the feedback system shown in Figure 12.1. Assume that the reference signal is constant. Let y_ol be the measured output when there is no feedback and y_cl be the output with feedback. Show that $Y_\mathrm{cl}(s) = S(s)Y_\mathrm{ol}(s)$, where Y_cl and Y_ol are exponential signals and S is the sensitivity function.

12.9 (Disturbance reduction through feedback) Consider a problem in which an output variable has been measured to estimate the potential for disturbance attenuation by feedback. Suppose an analysis shows that it is possible to design a closed loop system with the sensitivity function

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

Estimate the possible disturbance reduction when the measured disturbance response is

$$y(t) = 5\sin(0.1\,t) + 3\sin(0.17\,t) + 0.5\cos(0.9\,t) + 0.1\,t.$$

12.10 (Approximate expression for noise sensitivity) Show that the effect of high-frequency measurement noise on the control signal for the system in Example 12.3 can be approximated by

$$CS \approx C = \frac{k_\mathrm{d} s}{(sT_\mathrm{f})^2/2 + sT_\mathrm{f} + 1},$$

and that the largest value of $|CS(i\omega)|$ is $k_\mathrm{d}/T_\mathrm{f}$ which occurs for $\omega = \sqrt{2}/T_\mathrm{f}$.

12.11 (Lead-lag compensation) Lead and lag compensators can be combined into a lead-lag compensator that has the transfer function

$$C(s) = k\frac{(s+a_1)(s+a_2)}{(s+b_1)(s+b_2)}.$$

Show that the controller reduces to a PID controller with special choice of parameters and give the relations between the parameters.

12.12 (Attenuation of low-frequency sinusoidal disturbances) Integral action eliminates constant disturbances and reduces low-frequency disturbances because the controller gain is infinite at zero frequency. A similar idea can be used to reduce the effects of sinusoidal disturbances of known frequency ω_0 by using the controller

$$C(s) = k_{\mathrm{p}} + \frac{k_s s}{s^2 + 2\zeta\omega_0 s + \omega_0^2}.$$

This controller has the gain $C_s(i\omega_0) = k_{\mathrm{p}} + k_s/(2\zeta)$ for the frequency ω_0, which can be large by choosing a small value of ζ. Assume that the process has the transfer function $P(s) = 1/s$. Determine the Bode plot of the loop transfer function and simulate the system. Compare the results with PI control.

12.13 (Bode's formula) Consider the lead compensator

$$G(s) = 16\frac{s + 0.25}{s + 4}.$$

Verify Bode's phase area formula (12.8) and show that $G(\infty) = 16G(0)$ by numerical integration.

12.14 (Asymptotes of root locus) Consider proportional control of a system with the transfer function

$$P(s) = \frac{b(s)}{a(s)} = \frac{b_0 s^m + b_1 s^{m-1} + \cdots b_m}{s^n + a_1 s^{n-1} + \cdots a_n} = b_0 \frac{(s - z_1)(s - z_2)\ldots(s - z_m)}{(s - p_1)(s - p_2)\cdots(s - p_n)}.$$

Show that the root locus has asymptotes that are straight lines that emerge from the point

$$s_0 = \frac{1}{n_{\mathrm{e}}}\left(\sum_{k=1}^{n} p_k - \sum_{k=1}^{m} z_k\right),$$

where $n_{\mathrm{e}} = n - m$ is the pole excess of the transfer function.

12.15 (Real line segments of root locus) Consider proportional control of a process with a rational transfer function. Assuming that $b_0 k > 0$, show that the root locus has segments on the real line where there are an odd number of real poles and zeros to the right of the segment.

12.16 Consider a lead compensator with the transfer function

$$C_n(s) = \left(\frac{s\sqrt[n]{k} + a}{s + a}\right)^n,$$

which has zero frequency gain $C(0) = 1$ and high-frequency gain $C(\infty) = k$. Show that the gain required to provide a given phase lead φ is

$$k = \left(1 + 2\tan^2(\varphi/n) + 2\tan(\varphi/n)\sqrt{1 + \tan^2(\varphi/n)}\right)^n,$$

and that $\lim_{n\to\infty} k = e^{2\varphi}$.

12.17 (Phase margin formulas)

12.18 (Initial direction of root locus) Consider proportional control of a system with the transfer function

$$P(s) = \frac{b(s)}{a(s)} = \frac{b_0 s^m + b_1 s^{m-1} + \cdots b_m}{s^n + a_1 s^{n-1} + \cdots a_n} = b_0 \frac{(s - z_1)(s - z_2) \ldots (s - z_m)}{(s - p_1)(s - p_2) \cdots (s - p_n)}.$$

Let p_j be an isolated pole and assume that $k b_0 > 0$. Show that the root locus starting at p_j has the initial direction.

$$\angle(s - p_j) = \pi + \Sigma_{k=1}^{m} \angle(p_j - s_k) - \Sigma_{k \neq j} \angle(p_j - p_k).$$

Give a geometric interpretation of the result.

12.19 (Gain crossover frequency properties) Consider a system where the loop transfer function has monotone gain and phase.

a) Show that the gain crossover frequency ω_{gc}, the sensitivity crossover frequency ω_{sc}, and the crossover frequency ω_{tc} for the complementary sensitivity function $T(s)$ are equal if the phase margin $\varphi_m = 60°$.

b) Show that $\omega_{sc} < \omega_{gc} < \omega_{tc}$ if $\varphi_{,m} < 60°$ and that $\omega_{sc} > \omega_{gc}$ if $\varphi_m > 60°$.

c) Show that $\omega_{sc} \approx P(0)k_i$ for a process with integral action.

d) Find an approximate relation between the bandwidth ω_{bw} and ω_{tc}.