Chapter Twelve
Robust Performance

However, by building an amplifier whose gain is deliberately made, say 40 decibels higher than necessary (10000 fold excess on energy basis), and then feeding the output back on the input in such a way as to throw away that excess gain, it has been found possible to effect extraordinary improvement in constancy of amplification and freedom from non-linearity.

Harold S. Black, "Stabilized Feedback Amplifiers," 1934 [Bla34].

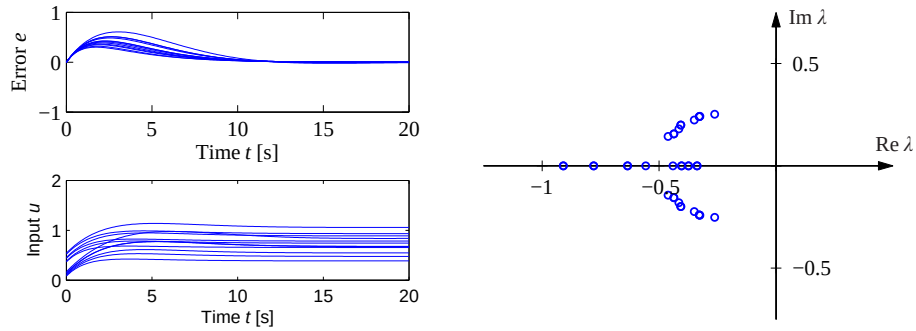
This chapter focuses on the analysis of robustness of feedback systems, a vast topic for which we provide only an introduction to some of the key concepts. We consider the stability and performance of systems whose process dynamics are uncertain and derive fundamental limits for robust stability and performance. To do this we develop ways to describe uncertainty, both in the form of parameter variations and in the form of neglected dynamics. We also briefly mention some methods for designing controllers to achieve robust performance.

12.1 Modeling Uncertainty

Harold Black's quote above illustrates that one of the key uses of feedback is to provide robustness to uncertainty ("constancy of amplification"). It is one of the most useful properties of feedback and is what makes it possible to design feedback systems based on strongly simplified models.

One form of uncertainty in dynamical systems is *parametric uncertainty* in which the parameters describing the system are unknown. A typical example is the variation of the mass of a car, which changes with the number of passengers and the weight of the baggage. When linearizing a nonlinear system, the parameters of the linearized model also depend on the operating conditions. It is straightforward to investigate the effects of parametric uncertainty simply by evaluating the performance criteria for a range of parameters. Such a calculation reveals the consequences of parameter variations. We illustrate by a simple example.

Example 12.1 Cruise control
The cruise control problem was described in Section 3.1, and a PI controller was designed in Example 10.3. To investigate the effect of parameter variations, we will choose a controller designed for a nominal operating condition corresponding to mass $m = 1600$ kg, fourth gear ($\alpha = 12$) and speed $v_e = 25$ m/s; the controller gains are $k_p = 0.72$ and $k_i = 0.18$. Figure 12.1a shows the velocity v and the throttle u when encountering a hill with a 3° slope with masses in the range $1600 < m < 2000$ kg, gear ratios 3–5 ($\alpha = 10$, 12 and 16) and velocity $10 \leq v \leq 40$ m/s.



(a) Disturbance response (b) Closed loop eigenvalues

Figure 12.1: Responses of the cruise control system to a slope increase of 3° (a) and the eigenvalues of the closed loop system (b). Model parameters are swept over a wide range.

The simulations were done using models that were linearized around the different operating conditions. The figure shows that there are variations in the response but that they are quite reasonable. The largest velocity error is in the range of 0.2–0.6 m/s, and the settling time is about 15 s. The control signal is marginally larger than 1 in some cases, which implies that the throttle is fully open. A full nonlinear simulation using a controller with windup protection is required if we want to explore these cases in more detail. Figure 12.1b shows the eigenvalues of the closed loop system for the different operating conditions. The figure shows that the closed loop system is well damped in all cases. ∇

This example indicates that at least as far as parametric variations are concerned, the design based on a simple nominal model will give satisfactory control. The example also indicates that a controller with fixed parameters can be used in all cases. Notice that we have not considered operating conditions in low gear and at low speed, but cruise controllers are not typically used in these cases.

Unmodeled Dynamics

It is generally easy to investigate the effects of parametric variations. However, there are other uncertainties that also are important, as discussed at the end of Section 2.3. The simple model of the cruise control system captures only the dynamics of the forward motion of the vehicle and the torque characteristics of the engine and transmission. It does not, for example, include a detailed model of the engine dynamics (whose combustion processes are extremely complex) or the slight delays that can occur in modern electronically controlled engines (as a result of the processing time of the embedded computers). These neglected mechanisms are called *unmodeled dynamics*.

Unmodeled dynamics can be accounted for by developing a more complex model. Such models are commonly used for controller development, but substantial effort is required to develop them. An alternative is to investigate if the closed loop system is sensitive to generic forms of unmodeled dynamics. The basic idea is to

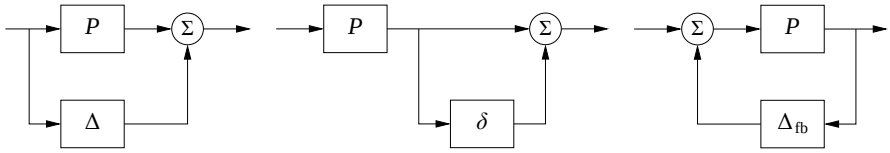


Figure 12.2: Unmodeled dynamics in linear systems. Uncertainty can be represented using additive perturbations (left), multiplicative perturbations (middle) or feedback perturbations (right). The nominal system is P, and Δ, $\delta = \Delta/P$ and Δ_{fb} represent unmodeled dynamics.

describe the unmodeled dynamics by including a transfer function in the system description whose frequency response is bounded but otherwise unspecified. For example, we might model the engine dynamics in the cruise control example as a system that quickly provides the torque that is requested through the throttle, giving a small deviation from the simplified model, which assumed the torque response was instantaneous. This technique can also be used in many instances to model parameter variations, allowing a quite general approach to uncertainty management.

In particular, we wish to explore if additional linear dynamics may cause difficulties. A simple way is to assume that the transfer function of the process is $P(s) + \Delta$, where $P(s)$ is the nominal simplified transfer function and Δ represents the unmodeled dynamics in terms of *additive uncertainty*. Different representations of uncertainty are shown in Figure 12.2.

When Are Two Systems Similar? The Vinnicombe Metric

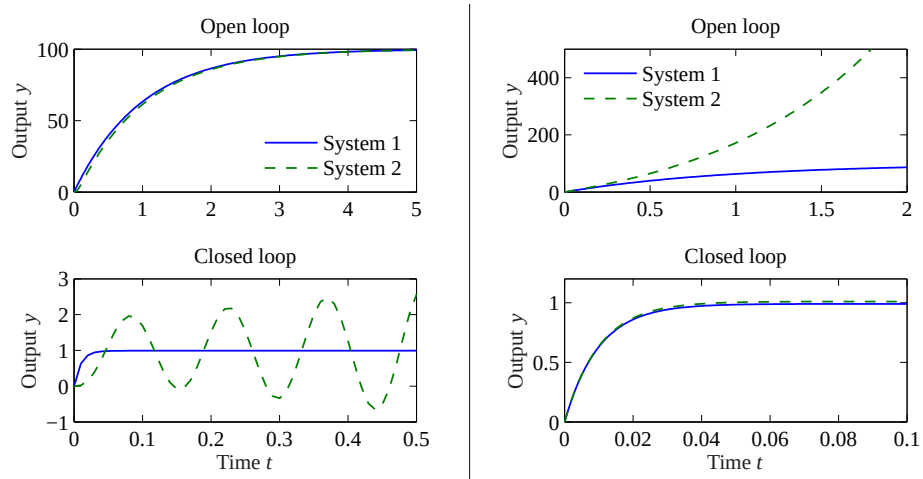
A fundamental issue in describing robustness is to determine when two systems are close. Given such a characterization, we can then attempt to describe robustness according to how close the actual system must be to the model in order to still achieve the desired levels of performance. This seemingly innocent problem is not as simple as it may appear. A naive approach is to say that two systems are close if their open loop responses are close. Even if this appears natural, there are complications, as illustrated by the following examples.

Example 12.2 Similar in open loop but large differences in closed loop
The systems with the transfer functions

$$P_1(s) = \frac{k}{s+1}, \qquad P_2(s) = \frac{k}{(s+1)(sT+1)^2} \tag{12.1}$$

have very similar open loop responses for small values of T, as illustrated in the top plot in Figure 12.3a, which is plotted for $T = 0.025$ and $k = 100$. The differences between the step responses are barely noticeable in the figure. The step responses with unit gain error feedback are shown in the bottom plot in Figure 12.3a. Notice that one closed loop system is stable and the other one is unstable. ∇

Example 12.3 Different in open loop but similar in closed loop



(a) Example 12.2 (b) Example 12.3

Figure 12.3: Determining when two systems are close. The plots in (a) show a situation when the open loop responses are almost identical, but the closed loop responses are very different. The processes are given by equation (12.1) with $k = 100$ and $T = 0.025$. The plots in (b) show the opposite situation: the systems are different in open loop but similar in closed loop. The processes are given by equation (12.2) with $k = 100$.

Consider the systems

$$P_1(s) = \frac{k}{s+1}, \qquad P_2(s) = \frac{k}{s-1}. \tag{12.2}$$

The open loop responses are very different because P_1 is stable and P_2 is unstable, as shown in the top plot in Figure 12.3b. Closing a feedback loop with unit gain around the systems, we find that the closed loop transfer functions are

$$T_1(s) = \frac{k}{s+k+1}, \qquad T_2(s) = \frac{k}{s+k-1},$$

which are very close for large k, as shown in Figure 12.3b. ∇

These examples show that if our goal is to close a feedback loop, it may be very misleading to compare the open loop responses of the system.

Inspired by these examples we introduce the *Vinnicombe metric*, which is a distance measure that is appropriate for closed loop systems. Consider two systems with the transfer functions P_1 and P_2, and define

$$d(P_1, P_2) = \sup_{\omega} \frac{|P_1(i\omega) - P_2(i\omega)|}{\sqrt{(1 + |P_1(i\omega)|^2)(1 + |P_2(i\omega)|^2)}}, \tag{12.3}$$

which is a metric with the property $0 \le d(P_1, P_2) \le 1$. The number $d(P_1, P_2)$ can be interpreted as the difference between the complementary sensitivity functions for the closed loop systems that are obtained with unit feedback around P_1 and P_2; see Exercise 12.3. The metric also has a nice geometric interpretation, as shown in

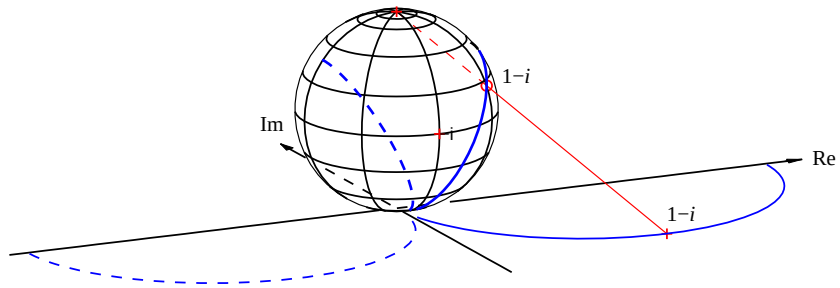


Figure 12.4: Geometric interpretation of $d(P_1, P_2)$. At each frequency, the points on the Nyquist curve for P_1 (solid) and P_2 (dashed) are projected onto a sphere of radius 1 sitting at the origin of the complex plane. The projection of the point $1 - i$ is shown. The distance between the two systems is defined as the maximum distance between the projections of $P_1(i\omega)$ and $P_2(i\omega)$ over all frequencies ω. The figure is plotted for the transfer functions $P_1(s) = 2/(s+1)$ and $P_2(s) = 2/(s-1)$. (Diagram courtesy G. Vinnicombe.)

Figure 12.4, where the Nyquist plots of P_1 and P_2 are projected onto a sphere with diameter 1 at the origin of the complex plane (called the *Riemann sphere*). Points in the complex plane are projected onto the sphere by a line through the point and the north pole (Figure 12.4). The distance $d(P_1, P_2)$ is the shortest chordal distance between the projections of $P_1(i\omega)$ and $P_2(i\omega)$. The distance is small when P_1 and P_2 are small or large, but it emphasizes the behavior around the gain crossover frequency.

The distance $d(P_1, P_2)$ has one drawback for the purpose of comparing the behavior of systems under feedback. If P_2 is perturbed continuously from P_1 to P_2, there can be intermediate transfer functions P where $d(P_1, P)$ is 1 even if $d(P_1, P_2)$ is small (see Exercise 12.4). To explore when this could happen, we observe that

$$1 - d^2(P_1, P) = \frac{(1 + P(i\omega)P_1(-i\omega))(1 + P(-i\omega)P_1(i\omega))}{(1 + |P_1(i\omega)|^2)(1 + |P(i\omega)|^2)}.$$

The right-hand side is zero, and hence $d(P_1, P) = 1$ if $1 + P(i\omega)P_1(-i\omega) = 0$ for some ω. To explore when this could occur, we investigate the behavior of the function $1+P(s)P_1(-s)$ when P is perturbed from P_1 to P_2. If the functions $f_1(s) = 1+P_1(s)P_1(-s)$ and $f_2(s) = 1+P_2(s)P_1(-s)$ do not have the same number of zeros in the right half-plane, there is an intermediate P such that $1 + P(i\omega)P_1(-i\omega) = 0$ for some ω. To exclude this case we introduce the set $\mathcal{C}$ as all pairs (P_1, P_2) such that the functions $f_1 = 1 + P_1(s)P_1(-s)$ and $f_2 = 1 + P_2(s)P_1(-s)$ have the same number of zeros in the right half-plane.

The *Vinnicombe metric* or *ν-gap metric* is defined as

$$\delta_\nu(P_1, P_2) = \begin{cases} d(P_1, P_2), & \text{if } (P_1, P_2) \in \mathcal{C} \\ 1, & \text{otherwise.} \end{cases} \tag{12.4}$$

Vinnicombe [Vin01] showed that $\delta_\nu(P_1, P_2)$ is a metric, he gave strong robustness results based on the metric and he developed the theory for systems with many

inputs and many outputs. We illustrate its use by computing the metric for the systems in the previous examples.

Example 12.4 Vinnicombe metric for Examples 12.2 and 12.3
For the systems in Example 12.2 we have

$$f_1(s) = 1 + P_1(s)P_1(-s) = \frac{1 + k^2 - s^2}{1 - s^2},$$
$$f_2(s) = 1 + P_2(s)P_1(-s) = \frac{1 + k^2 + 2sT + (T^2 - 1)s^2 - 2s^3T - s^4T^2}{(1 - s^2)(1 + 2sT + s^2T^2)}.$$

The function f_1 has one zero in the right half-plane. A numerical calculation for $k = 100$ and $T = 0.025$ shows that the function f_2 has the roots 46.3, -86.3, $-20.0 \pm 60.0i$. Both functions have one zero in the right half-plane, allowing us to compute the norm (12.4). For $T = 0.025$ this gives $\delta_\nu(P_1, P_2) = 0.98$, which is a quite large value. To have reasonable robustness Vinnicombe recommended values less than 1/3.

For the system in Example 12.3 we have

$$1 + P_1(s)P_1(-s) = \frac{1 + k^2 - s^2}{1 - s^2}, \qquad 1 + P_2(s)P_1(-s) = \frac{1 - k^2 - 2s + s^2}{(s + 1)^2}$$

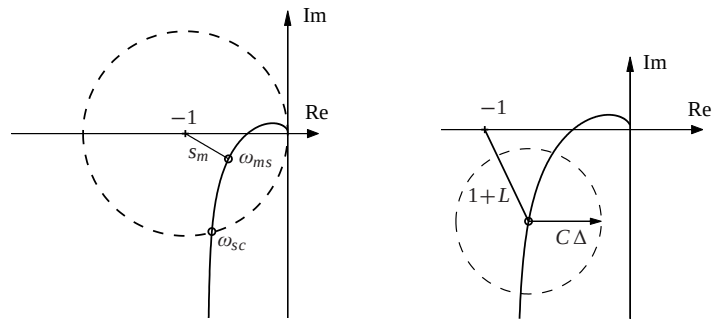
These functions have the same number of zeros in the right half-plane if $k > 1$. In this particular case the Vinnicombe metric is $d(P_1, P_2) = 2k/(1 + k^2)$ (Exercise 12.4) and with $k = 100$ we get $\delta_\nu(P_1, P_2) = 0.02$. Figure 12.4 shows the Nyquist curves and their projections for $k = 2$. Notice that $d(P_1, P_2)$ is very small for small k even though the closed loop systems are very different. It is therefore essential to consider the condition $(P_1, P_2) \in \mathcal{C}$, as discussed in Exercise 12.4. ∇

12.2 Stability in the Presence of Uncertainty

Having discussed how to describe uncertainty and the similarity between two systems, we now consider the problem of robust stability: When can we show that the stability of a system is robust with respect to process variations? This is an important question since the potential for instability is one of the main drawbacks of feedback. Hence we want to ensure that even if we have small inaccuracies in our model, we can still guarantee stability and performance.

Robust Stability Using Nyquist's Criterion

The Nyquist criterion provides a powerful and elegant way to study the effects of uncertainty for linear systems. A simple criterion is that the Nyquist curve be sufficiently far from the critical point -1. Recall that the shortest distance from the Nyquist curve to the critical point is $s_m = 1/M_s$, where M_s is the maximum of the sensitivity function and s_m is the stability margin introduced in Section 9.3.



(a) Nyquist plot (b) Additive uncertainty

Figure 12.5: Robust stability using the Nyquist criterion. (a) This plot shows that the shortest distance to the critical point s_m is a robustness measure. (b) This plot shows the Nyquist curve of a nominal loop transfer function and its uncertainty caused by additive process variations Δ.

The maximum sensitivity M_s or the stability margin s_m is thus a good robustness measure, as illustrated in Figure 12.5a.

We will now derive explicit conditions for permissible process uncertainties. Consider a stable feedback system with a process P and a controller C. If the process is changed from P to $P + \Delta$, the loop transfer function changes from PC to $PC + C\Delta$, as illustrated in Figure 12.5b. If we have a bound on the size of Δ (represented by the dashed circle in the figure), then the system remains stable as long as the process variations never overlap the -1 point, since this leaves the number of encirclements of -1 unchanged.

Some additional assumptions are required for the analysis to hold. Most importantly, we require that the process perturbations Δ be stable so that we do not introduce any new right half-plane poles that would require additional encirclements in the Nyquist criterion.

We will now compute an analytical bound on the allowable process disturbances. The distance from the critical point -1 to the loop transfer function L is $|1 + L|$. This means that the perturbed Nyquist curve will not reach the critical point -1 provided that $|C\Delta| < |1 + L|$, which implies

$$|\Delta| < \left|\frac{1 + PC}{C}\right| \qquad \text{or} \qquad |\delta| = \left|\frac{\Delta}{P}\right| < \frac{1}{|T|}. \tag{12.5}$$

This condition must be valid for all points on the Nyquist curve, i.e, pointwise for all frequencies. The condition for robust stability can thus be written as

$$|\delta(i\omega)| = \left|\frac{\Delta(i\omega)}{P(i\omega)}\right| < \frac{1}{|T(i\omega)|} \qquad \text{for all } \omega \geq 0. \tag{12.6}$$

Notice that the condition is conservative because it follows from Figure 12.5 that the critical perturbation is in the direction toward the critical point -1. Larger perturbations can be permitted in the other directions.

The condition in equation (12.6) allows us to reason about uncertainty without

exact knowledge of the process perturbations. Namely, we can verify stability for *any* uncertainty Δ that satisfies the given bound. From an analysis perspective, this gives us a measure of the robustness for a given design. Conversely, if we require robustness of a given level, we can attempt to choose our controller C such that the desired level of robustness is available (by asking that T be small) in the appropriate frequency bands.

Equation (12.6) is one of the reasons why feedback systems work so well in practice. The mathematical models used to design control systems are often simplified, and the properties of a process may change during operation. Equation (12.6) implies that the closed loop system will at least be stable for substantial variations in the process dynamics.

It follows from equation (12.6) that the variations can be large for those frequencies where T is small and that smaller variations are allowed for frequencies where T is large. A conservative estimate of permissible process variations that will not cause instability is given by

$$|\delta(i\omega)| = \left|\frac{\Delta(i\omega)}{P(i\omega)}\right| < \frac{1}{M_t},$$

where M_t is the largest value of the complementary sensitivity

$$M_t = \sup_{\omega} |T(i\omega)| = \left\|\frac{PC}{1+PC}\right\|_{\infty}. \tag{12.7}$$

The value of M_t is influenced by the design of the controller. For example, it is shown in Exercise 12.5 that if $M_t = 2$ then pure gain variations of 50% or pure phase variations of 30° are permitted without making the closed loop system unstable.

Example 12.5 Cruise control

Consider the cruise control system discussed in Section 3.1. The model of the car in fourth gear at speed 25 m/s is

$$P(s) = \frac{1.38}{s + 0.0142},$$

and the controller is a PI controller with gains $k_p = 0.72$ and $k_i = 0.18$. Figure 12.6 plots the allowable size of the process uncertainty using the bound in equation (12.6). At low frequencies, $T(0) = 1$ and so the perturbations can be as large as the original process ($|\delta| = |\Delta/P| < 1$). The complementary sensitivity has its maximum $M_t = 1.14$ at $\omega_{mt} = 0.35$, and hence this gives the minimum allowable process uncertainty, with $|\delta| < 0.87$ or $|\Delta| < 3.47$. Finally, at high frequencies, $T \to 0$ and hence the relative error can get very large. For example, at $\omega = 5$ we have $|T(i\omega)| = 0.195$, which means that the stability requirement is $|\delta| < 5.1$. The analysis clearly indicates that the system has good robustness and that the high-frequency properties of the transmission system are not important for the design of the cruise controller.

Another illustration of the robustness of the system is given in the right diagram

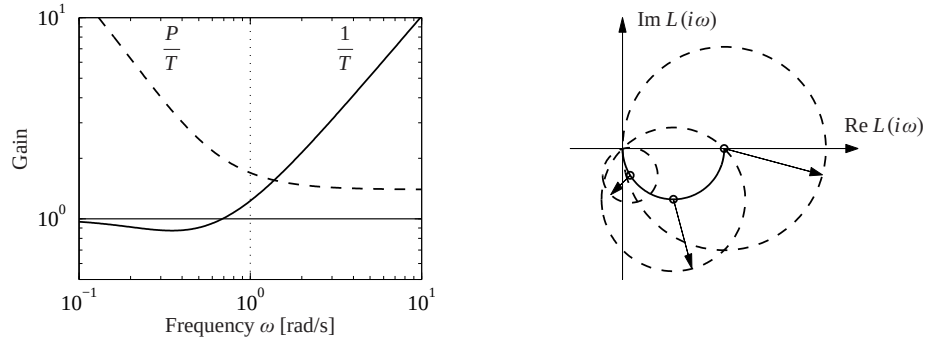


Figure 12.6: Robustness for a cruise controller. On the left the maximum relative error $1/|T|$ (solid) and the absolute error $|P|/|T|$ (dashed) for the process uncertainty Δ. The Nyquist curve is shown on the right as a solid line. The dashed circles show permissible perturbations in the process dynamics, $|\Delta| = |P|/|T|$, at the frequencies $\omega = 0$, 0.0142 and 0.05.

in Figure 12.6, which shows the Nyquist curve of the transfer function of the process and the uncertainty bounds $\Delta = |P|/|T|$ for a few frequencies. Note that the controller can tolerate large amounts of uncertainty and still maintain stability of the closed loop. ∇

The situation illustrated in the previous example is typical of many processes: moderately small uncertainties are required only around the gain crossover frequencies, but large uncertainties can be permitted at higher and lower frequencies. A consequence of this is that a simple model that describes the process dynamics well around the crossover frequency is often sufficient for design. Systems with many resonant peaks are an exception to this rule because the process transfer function for such systems may have large gains for higher frequencies also, as shown for instance in Example 9.9.

The robustness condition given by equation (12.6) can be given another interpretation by using the small gain theorem (Theorem 9.4). To apply the theorem we start with block diagrams of a closed loop system with a perturbed process and make a sequence of transformations of the block diagram that isolate the block representing the uncertainty, as shown in Figure 12.7. The result is the two-block interconnection shown in Figure 12.7c, which has the loop transfer function

$$L = \frac{PC}{1 + PC}\,\frac{\Delta}{P} = T\delta.$$

Equation (12.6) implies that the largest loop gain is less than 1 and hence the system is stable via the small gain theorem.

The small gain theorem can be used to check robust stability for uncertainty in a variety of other situations. Table 12.1 summarizes a few of the common cases; the proofs (all via the small gain theorem) are left as exercises.

The following example illustrates that it is possible to design systems that are robust to parameter variations.

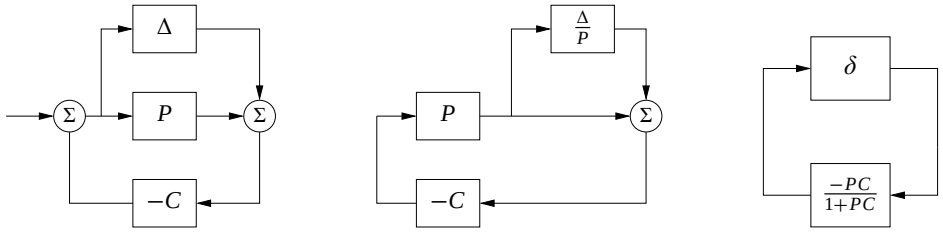


Figure 12.7: Illustration of robustness to process perturbations. A system with additive uncertainty (left) can be manipulated via block diagram algebra to one with multiplicative uncertainty $\delta = \Delta/P$ (center). Additional manipulations isolate the uncertainty in a manner that allows application of the small gain theorem (right)

Example 12.6 Bode's ideal loop transfer function
A major problem in the design of electronic amplifiers is to obtain a closed loop system that is insensitive to changes in the gain of the electronic components. Bode found that the loop transfer function $L(s) = ks^{-n}$, with $1 \leq n \leq 5/3$, was an ideal loop transfer function. The gain curve of the Bode plot is a straight line with slope $-n$ and the phase is constant $\arg L(i\omega) = -n\pi/2$. The phase margin is thus $\varphi_m = 90(2-n)^\circ$ for all values of the gain k and the stability margin is $s_m = \sin \pi(1 - n/2)$. This exact transfer function cannot be realized with physical components, but it can be approximated over a given frequency range with a rational function (Exercise 12.7). An operational amplifier circuit that has the approximate transfer function $G(s) = k/(s+a)$ is a realization of Bode's ideal transfer function with $n = 1$, as described in Example 8.3. Designers of operational amplifiers go to great efforts to make the approximation valid over a wide frequency range. ∇

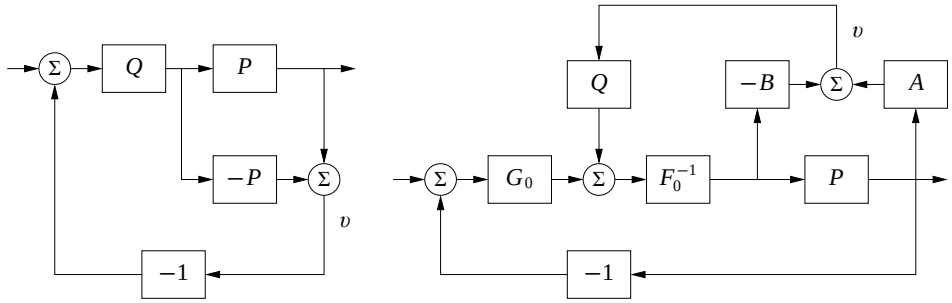
Youla Parameterization

Since stability is such an essential property, it is useful to characterize all controllers that stabilize a given process. Such a representation, which is called a *Youla parameterization*, is very useful when solving design problems because it makes it possible to search over all stabilizing controllers without the need to test stability explicitly.

We will first derive Youla's parameterization for a stable process with a rational transfer function P. A system with the complementary sensitivity function T can be obtained by feedforward control with the stable transfer function Q if $T = PQ$.

Table 12.1: Conditions for robust stability for different types of uncertainty

Process	Uncertainty Type	Robust Stability
$P + \Delta$	Additive	$\|CS\Delta\|_\infty < 1$
$P(1+\delta)$	Multiplicative	$\|T\delta\|_\infty < 1$
$P/(1+\Delta_{\text{fb}} \cdot P)$	Feedback	$\|PS\Delta_{\text{fb}}\|_\infty < 1$



(a) Stable process (b) Unstable process

Figure 12.8: Youla parameterization. Block diagrams of Youla parameterizations for a stable system (a) and an unstable system (b). Notice that the signal v is zero in steady state.

Notice that T must have the same right half-plane zeros as P since Q is stable. Now assume that we want to implement the complementary transfer function T by using unit feedback with the controller C. Since $T = PC/(1 + PC) = PQ$, it follows that the controller transfer function is

$$C = \frac{Q}{1 - PQ}. \tag{12.8}$$

A straightforward calculation gives

$$S = 1 - PQ, \qquad PS = P(1 - PQ), \qquad CS = Q, \qquad T = PQ.$$

These transfer functions are all stable if P and Q are stable and the controller given by equation (12.8) is thus stabilizing. Indeed, it can be shown that all stabilizing controllers are in the form given by equation (12.8) for some choice of Q. The parameterization is illustrated by the block diagrams in Figure 12.8a.

A similar characterization can be obtained for unstable systems. Consider a process with a rational transfer function $P(s) = a(s)/b(s)$, where $a(s)$ and $b(s)$ are polynomials. By introducing a stable polynomial $c(s)$, we can write

$$P(s) = \frac{a(s)}{b(s)} = \frac{A(s)}{B(s)},$$

where $A(s) = a(s)/c(s)$ and $B(s) = b(s)/c(s)$ are stable rational functions. Similarly we introduce the controller $C_0(s) = F_0(s)/G_0(s)$, where $F_0(s)$ and $G_0(s)$ are stable rational functions. We have

$$S_0 = \frac{AF_0}{AF_0 + BG_0}, \qquad PS_0 = \frac{BF_0}{AF_0 + BG_0},$$
$$C_0S_0 = \frac{AG_0}{AF_0 + BG_0}, \qquad T_0 = \frac{BG_0}{AF_0 + BG_0}.$$

The controller C_0 is stabilizing if and only if the rational function $AF_0 + BG_0$ does not have any zeros in the right half plane. Let Q be a stable rational function and

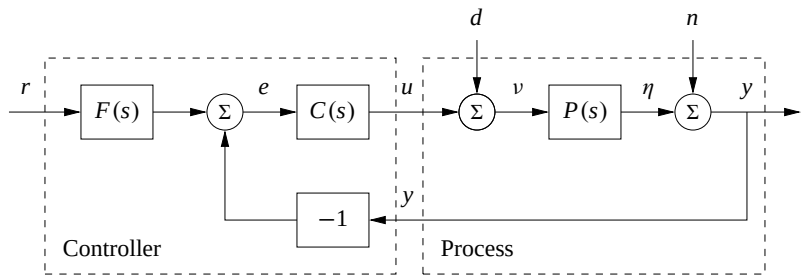


Figure 12.9: Block diagram of a basic feedback loop. The external signals are the reference signal r, the load disturbance d and the measurement noise n. The process output is y, and the control signal is u. The process P may include unmodeled dynamics, such as additive perturbations.

consider the controller

$$C = \frac{G_0 + QA}{F_0 - QB}. \tag{12.9}$$

The Gang of Four for P and C is

$$S = \frac{A(F_0 - QB)}{AF_0 + BG_0}, \qquad PS = \frac{B(F_0 - QB)}{AF_0 + BG_0},$$
$$CS = \frac{A(G_0 + QA)}{AF_0 + BG_0}, \qquad T = \frac{B(G_0 + QA)}{AF_0 + BG_0}.$$

All these transfer functions are stable if the rational function $AF_0 + BG_0$ does not have any zeros in the right half plane and the controller C given by (12.9) is therefore stabilizing for any stable Q. A block diagram of the closed loop system with the controller C is shown in Figure 12.8b. Notice that the transfer function Q appears affinely in the expressions for the Gang of Four, which is very useful if we want to determine the transfer function Q to obtain specific properties.

12.3 Performance in the Presence of Uncertainty

So far we have investigated the risk for instability and robustness to process uncertainty. We will now explore how responses to load disturbances, measurement noise and reference signals are influenced by process variations. To do this we will analyze the system in Figure 12.9, which is identical to the basic feedback loop analyzed in Chapter 11.

Disturbance Attenuation

The sensitivity function S gives a rough characterization of the effect of feedback on disturbances, as was discussed in Section 11.3. A more detailed characterization is given by the transfer function from load disturbances to process output:

$$G_{yd} = \frac{P}{1 + PC} = PS. \tag{12.10}$$

Load disturbances typically have low frequencies, and it is therefore important that the transfer function be small for low frequencies. For processes with constant low-frequency gain and a controller with integral action we have $G_{yd} \approx s/k_i$. The integral gain k_i is thus a simple measure of the attenuation of load disturbances.

To find out how the transfer function G_{yd} is influenced by small variations in the process transfer function we differentiate (12.10) with respect to P yielding

$$\frac{dG_{yd}}{dP} = \frac{1}{(1+PC)^2} = \frac{SP}{P(1+PC)} = S\frac{G_{yd}}{P},$$

and it follows that

$$\frac{dG_{yd}}{G_{yd}} = S\frac{dP}{P}. \tag{12.11}$$

The response to load disturbances is thus insensitive to process variations for frequencies where $|S(i\omega)|$ is small, i.e., for frequencies where load disturbances are important.

A drawback with feedback is that the controller feeds measurement noise into the system. In addition to the load disturbance rejection, it is thus also important that the control actions generated by measurement noise are not too large. It follows from Figure 12.9 that the transfer function G_{un} from measurement noise to controller output is given by

$$G_{un} = -\frac{C}{1+PC} = -\frac{T}{P}. \tag{12.12}$$

Since measurement noise typically has high frequencies, the transfer function G_{un} should not be too large for high frequencies. The loop transfer function PC is typically small for high frequencies, which implies that $G_{un} \approx C$ for large s. To avoid injecting too much measurement noise it is therefore important that $C(s)$ be small for large s. This property is called *high-frequency roll-off*. An example is filtering of the measured signal in a PID controller to reduce the injection of measurement noise; see Section 10.5.

To determine how the transfer function G_{un} is influenced by small variations in the process transfer, we differentiate equation (12.12):

$$\frac{dG_{un}}{dP} = \frac{d}{dP}\left(-\frac{C}{1+PC}\right) = \frac{C}{(1+PC)^2}C = T\frac{G_{un}}{P}.$$

Rearranging the terms gives

$$\frac{dG_{un}}{G_{un}} = T\frac{dP}{P}. \tag{12.13}$$

Since the complementary sensitivity function is also small for high frequencies, we find that process uncertainty has little influence on the transfer function G_{un} for frequencies where measurements are important.

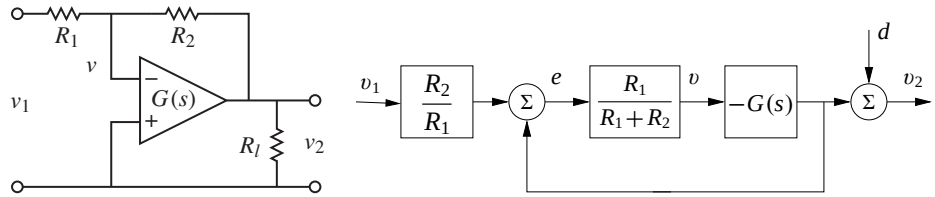


Figure 12.10: Operational amplifier with uncertain dynamics. The circuit on the left is modeled using the transfer function $G(s)$ to capture its dynamic properties and it has a load at the output. The block diagram on the right shows the input/output relationships. The load is represented as a disturbance d applied at the output of $G(s)$.

Reference Signal Tracking

The transfer function from reference to output is given by

$$G_{yr} = \frac{PCF}{1+PC} = TF, \tag{12.14}$$

which contains the complementary sensitivity function. To see how variations in P affect the performance of the system, we differentiate equation (12.14) with respect to the process transfer function:

$$\frac{dG_{yr}}{dP} = \frac{CF}{1+PC} - \frac{PCFC}{(1+PC)^2} = \frac{CF}{(1+PC)^2} = S\frac{G_{yr}}{P},$$

and it follows that

$$\frac{dG_{yr}}{G_{yr}} = S\frac{dP}{P}. \tag{12.15}$$

The relative error in the closed loop transfer function thus equals the product of the sensitivity function and the relative error in the process. In particular, it follows from equation (12.15) that the relative error in the closed loop transfer function is small when the sensitivity is small. This is one of the useful properties of feedback.

As in the last section, there are some mathematical assumptions that are required for the analysis presented here to hold. As already stated, we require that the perturbations Δ be small (as indicated by writing dP). Second, we require that the perturbations be stable, so that we do not introduce any new right half-plane poles that would require additional encirclements in the Nyquist criterion. Also, as before, this condition is conservative: it allows for any perturbation that satisfies the given bounds, while in practice the perturbations may be more restricted.

Example 12.7 Operational amplifier circuit

To illustrate the use of these tools, consider the performance of an op amp-based amplifier, as shown in Figure 12.10. We wish to analyze the performance of the amplifier in the presence of uncertainty in the dynamic response of the op amp and changes in the loading on the output. We model the system using the block diagram in Figure 12.10b, which is based on the derivation in Example 9.1.

Consider first the effect of unknown dynamics for the operational amplifier. If we model the dynamics of the op amp as $v_2 = -G(s)v$, then the transfer function

for the overall circuit is given by

$$G_{v_2 v_1} = -\frac{R_2}{R_1}\frac{G(s)}{G(s) + R_2/R_1 + 1}.$$

We see that if $G(s)$ is large over the desired frequency range, then the closed loop system is very close to the ideal response $\alpha = R_2/R_1$. Assuming $G(s) = b/(s+a)$, where b is the gain-bandwidth product of the amplifier, as discussed in Example 8.3, the sensitivity function and the complementary sensitivity function become

$$S = \frac{s+a}{s+a+\alpha b}, \qquad T = \frac{\alpha b}{s+a+\alpha b}.$$

The sensitivity function around the nominal values tells us how the tracking response response varies as a function of process perturbations:

$$\frac{dG_{yr}}{G_{yr}} = S\frac{dP}{P}.$$

We see that for low frequencies, where S is small, variations in the bandwidth a or the gain-bandwidth product b will have relatively little effect on the performance of the amplifier (under the assumption that b is sufficiently large.

To model the effects of an unknown load, we consider the addition of a disturbance at the output of the system, as shown in Figure 12.10b. This disturbance represents changes in the output voltage due to loading effects. The transfer function $G_{yd} = S$ gives the response of the output to the load disturbance, and we see that if S is small, then we are able to reject such disturbances. The sensitivity of G_{yd} to perturbations in the process dynamics can be computed by taking the derivative of G_{yd} with respect to P:

$$\frac{dG_{yd}}{dP} = \frac{-C}{(1+PC)^2} = -\frac{T}{P}G_{yd} \qquad \Longrightarrow \qquad \frac{dG_{yd}}{G_{yd}} = -T\frac{dP}{P}.$$

Thus we see that the relative changes in the disturbance rejection are roughly the same as the process perturbations at low frequency (when T is approximately 1) and drop off at higher frequencies. However, it is important to remember that G_{yd} itself is small at low frequency, and so these variations in relative performance may not be an issue in many applications. ∇

12.4 Robust Pole Placement

In Chapters 6 and 7 we saw how to design controllers by setting the locations of the eigenvalues of the closed loop system. If we analyze the resulting system in the frequency domain, the closed loop eigenvalues correspond to the poles of the closed loop transfer function and hence these methods are often referred to as design by *pole placement*.

State space design methods, like many methods developed for control system design, do not explicitly take robustness into account. In such cases it is essential to

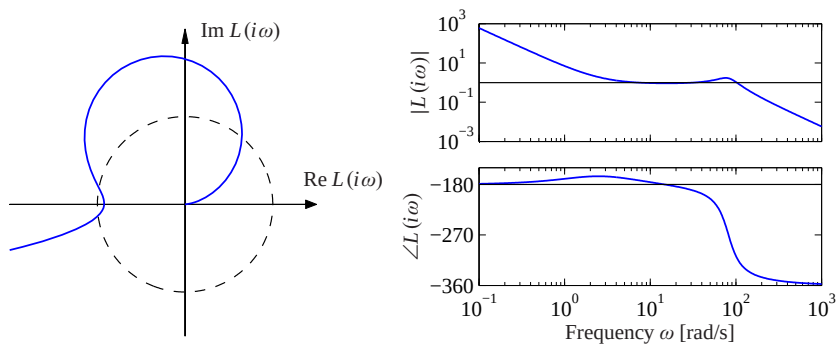


Figure 12.11: Observer-based control of steering. The Nyquist plot (left) and Bode plot (right) of the loop transfer function for vehicle steering with a controller based on state feedback and an observer. The controller provides stable operation, but with very low gain and phase margin.

always investigate the robustness because there are seemingly reasonable designs that give controllers with poor robustness. We illustrate this by analyzing controllers designed by state feedback and observers. The closed loop poles can be assigned to arbitrary locations if the system is observable and reachable. However, if we want to have a robust closed loop system, the poles and zeros of the process impose severe restrictions on the location of the closed loop poles. Some examples are first given; based on the analysis of these examples we then present design rules for robust pole (eigenvalue) placement.

Slow Stable Process Zeros

We will first explore the effects of slow stable zeros, and we begin with a simple example.

Example 12.8 Vehicle steering
Consider the linearized model for vehicle steering in Example 8.6, which has the transfer function

$$P(s) = \frac{0.5s + 1}{s^2}.$$

A controller based on state feedback was designed in Example 6.4, and state feedback was combined with an observer in Example 7.4. The system simulated in Figure 7.8 has closed loop poles specified by $\omega_c = 0.3$, $\zeta_c = 0.707$, $\omega_o = 7$ and $\zeta_o = 9$. Assume that we want a faster closed loop system and choose $\omega_c = 10$, $\zeta_c = 0.707$, $\omega_o = 20$ and $\zeta_o = 0.707$. Using the state representation in Example 7.3, a pole placement design gives state feedback gains $k_1 = 100$ and $k_2 = -35.86$ and observer gains $l_1 = 28.28$ and $l_2 = 400$. The controller transfer function is

$$C(s) = \frac{-11516s + 40000}{s^2 + 42.4s + 6657.9}.$$

Figure 12.11 shows Nyquist and Bode plots of the loop transfer function. The

Nyquist plot indicates that the robustness is poor since the loop transfer function is very close to the critical point -1. The phase margin is 7° and the stability margin is $s_m = 0.077$. The poor robustness shows up in the Bode plot, where the gain curve hovers around the value 1 and the phase curve is close to -180° for a wide frequency range. More insight is obtained by analyzing the sensitivity functions, shown by solid lines in Figure 12.12. The maximum sensitivities are $M_s = 13$ and $M_t = 12$, indicating that the system has poor robustness.

At first sight it is surprising that a controller where the nominal closed system has well damped poles and zeros is so sensitive to process variations. We have an indication that something is unusual because the controller has a zero at $s = 3.5$ in the right half-plane. To understand what happens, we will investigate the reason for the peaks of the sensitivity functions.

Let the transfer functions of the process and the controller be

$$P(s) = \frac{n_p(s)}{d_p(s)}, \qquad C(s) = \frac{n_c(s)}{d_c(s)},$$

where $n_p(s)$, $n_c(s)$, $d_p(s)$ and $d_c(s)$ are the numerator and denominator polynomials. The complementary sensitivity function is

$$T(s) = \frac{PC}{1 + PC} = \frac{n_p(s)n_c(s)}{d_p(s)d_c(s) + n_p(s)n_p(s)}.$$

The poles of $T(s)$ are the poles of the closed loop system and the zeros are given by the zeros of the process and controller. Sketching the gain curve of the complementary sensitivity function we find that $T(s) = 1$ for low frequencies and that $|T(i\omega)|$ starts to increase at its first zero, which is the process zero at $s = 2$. It increases further at the controller zero at $s = 3.4$, and it does not start to decrease until the closed loop poles appear at $\omega_c = 10$ and $\omega_o = 20$. We can thus conclude that there will be a peak in the complementary sensitivity function. The magnitude of the peak depends on the ratio of the zeros and the poles of the transfer function.

The peak of the complementary sensitivity function can be avoided by assigning a closed loop pole close to the slow process zero. We can achieve this by choosing $\omega_c = 10$ and $\zeta_c = 2.6$, which gives closed loop poles at $s = -2$ and $s = -50$. The controller transfer function then becomes

$$C(s) = \frac{3628s + 40000}{s^2 + 80.28s + 156.56} = 3628\,\frac{s + 11.02}{(s + 2)(s + 78.28)}.$$

The sensitivity functions are shown by dashed lines in Figure 12.12. The controller gives the maximum sensitivities $M_s = 1.34$ and $M_t = 1.41$, which give much better robustness. Notice that the controller has a pole at $s = -2$ that cancels the slow process zero. The design can also be done simply by canceling the slow stable process zero and designing the controller for the simplified system. ∇

One lesson from the example is that it is necessary to choose closed loop poles that are equal to or close to slow stable process zeros. Another lesson is that slow unstable process zeros impose limitations on the achievable bandwidth, as already

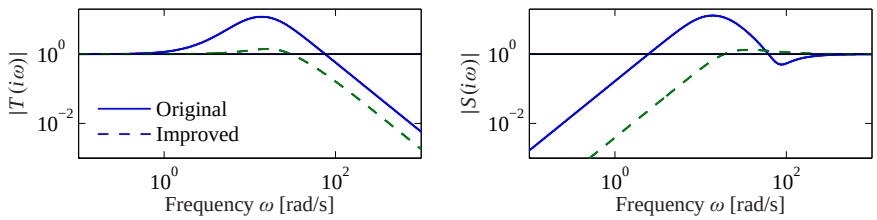


Figure 12.12: Sensitivity functions for observer-based control of vehicle steering. The complementary sensitivity function (left) and the sensitivity function (right) for the original controller with $\omega_c = 10$, $\zeta_c = 0.707$, $\omega_o = 20$, $\zeta_o = 0.707$ (solid) and the improved controller with $\omega_c = 10$, $\zeta_c = 2.6$ (dashed).

noted in Section 11.5.

Fast Stable Process Poles

The next example shows the effect of fast stable poles.

Example 12.9 Fast system poles
Consider a PI controller for a first-order system, where the process and the controller have the transfer functions $P(s) = b/(s+a)$ and $C(s) = k_p + k_i/s$. The loop transfer function is

$$L(s) = \frac{b(k_p s + k_i)}{s(s+a)},$$

and the closed loop characteristic polynomial is.

$$s(s+a) + b(k_p s + k_i) = s^2 + (a + bk_p)s + k_i b$$

If we specify the desired closed loop poles should be $-p_1$ and $-p_2$, we find that the controller parameters are given by

$$k_p = \frac{p_1 + p_2 - a}{b}, \qquad k_i = \frac{p_1 p_2}{b}.$$

The sensitivity functions are then

$$S(s) = \frac{s(s+a)}{(s+p_1)(s+p_2)}, \qquad T(s) = \frac{(p_1 + p_2 - a)s + p_1 p_2}{(s+p_1)(s+p_2)}.$$

Assume that the process pole a is much larger than the closed loop poles p_1 and p_2, say, $p_1 < p_2 \ll a$. Notice that the proportional gain is negative and that the controller has a zero in the right half-plane if $a > p_1 + p_2$, an indication that the system has bad properties.

Next consider the sensitivity function, which is 1 for high frequencies. Moving from high to low frequencies, we find that the sensitivity increases at the process pole $s = a$. The sensitivity does not decrease until the closed loop poles are reached, resulting in a large sensitivity peak that is approximately a/p_2. The magnitude of the sensitivity function is shown in Figure 12.13 for $a = b = 1$, $p_1 = 0.05$ and $p_2 = 0.2$. Notice the high-sensitivity peak. For comparison we also show the gain

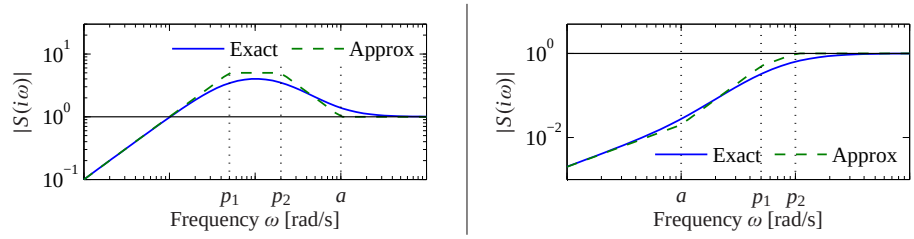


Figure 12.13: Gain curves for Bode plots of the sensitivity function S for designs with $p_1 < p_2 < a$ (left) and $a < p_1 < p_2$ (right). The solid lines are the true sensitivities, and the dashed lines are the asymptotes.

curve for the case when the closed loop poles ($p_1 = 5$, $p_2 = 20$) are faster than the process pole ($a = 1$).

The problem with poor robustness can be avoided by choosing one closed loop pole equal to the process pole, i.e., $p_2 = a$. The controller gains then become

$$k_p = \frac{p_1}{b}, \qquad k_i = \frac{ap_1}{l},$$

which means that the fast process pole is canceled by a controller zero. The loop transfer function and the sensitivity functions are

$$L(s) = \frac{bk_p}{s}, \qquad S(s) = \frac{s}{s + bk_p}, \qquad T(s) = \frac{bk_p}{s + bk_p}.$$

The maximum sensitivities are now less than 1 for all frequencies. Notice that this is possible because the process transfer function goes to zero as s^{-1}. ∇

Design Rules for Pole Placement

Based on the insight gained from the examples, it is now possible to obtain design rules that give designs with good robustness. Consider the expression (12.8) for maximum complementary sensitivity, repeated here:

$$M_t = \sup_{\omega} |T(i\omega)| = \left\| \frac{PC}{1 + PC} \right\|_{\infty}.$$

Let ω_{gc} be the desired gain crossover frequency. Assume that the process has zeros that are slower than ω_{gc}. The complementary sensitivity function is 1 for low frequencies, and it increases for frequencies close to the process zeros unless there is a closed loop pole in the neighborhood. To avoid large values of the complementary sensitivity function we find that the closed loop system should therefore have poles close to or equal to the slow stable zeros. This means that slow stable zeros should be canceled by controller poles. Since unstable zeros cannot be canceled, the presence of slow unstable zeros means that achievable gain crossover frequency must be smaller than the slowest unstable process zero.

Now consider process poles that are faster than the desired gain crossover fre-

quency. Consider the expression for the maximum of the sensitivity function:

$$M_s = \sup_{\omega} |S(i\omega)| = \left\| \frac{1}{1+PC} \right\|_{\infty}.$$

The sensitivity function is 1 for high frequencies. Moving from high to low frequencies, the sensitivity function increases at the fast process poles. Large peaks can result unless there are closed loop poles close to the fast process poles. To avoid large peaks in the sensitivity the closed loop system should therefore have poles that match the fast process poles. This means that the controller should cancel the fast process poles by controller zeros. Since unstable modes cannot be canceled, the presence of a fast unstable pole implies that the gain crossover frequency must be sufficiently large.

To summarize, we obtain the following simple rule for choosing closed loop poles: slow stable process zeros should be matched by slow closed loop poles, and fast stable process poles should be matched by fast closed loop poles. Slow unstable process zeros and fast unstable process poles impose severe limitations.

Example 12.10 Nanopositioning system for an atomic force microscope
A simple nanopositioner was explored in Example 9.9, where it was shown that the system could be controlled using an integral controller. The performance of the closed loop was poor because the gain crossover frequency was limited to $\omega_{gc} = 2\zeta\omega_0(1 - s_m)$. It can be shown that little improvement is obtained by using a PI controller. To achieve improved performance, we will therefore apply PID control. For a modest performance increase, we will use the design rule derived in Example 12.9 that fast stable process poles should be canceled by controller zeros. The controller transfer function should thus be chosen as

$$C(s) = \frac{k_d s^2 + k_p s + k_i}{s} = \frac{k_i}{s}\frac{s^2 + 2\zeta s + a^2}{a^2}, \tag{12.16}$$

which gives $k_p = 2\zeta k_i/a$ and $k_d = k_i/a^2$.

Figure 12.14 shows the gain curves for the Gang of Four for a system designed with $k_i = 0.5$. A comparison with Figure 9.12 shows that the bandwidth is increased significantly from $\omega_{gc} = 0.01$ to $\omega_{gc} = k_i = 0.5$. Since the process pole is canceled, the system will, however, still be very sensitive to load disturbances with frequencies close to the resonant frequency. The gain curve of CS has a dip or a notch at the resonant frequency, which implies that the controller gain is very low for frequencies around the resonance. The gain curve also shows that the system is very sensitive to high-frequency noise. The system will likely be unusable because the gain goes to infinity for high frequencies.

The sensitivity to high frequency noise can be remedied by modifying the controller to be

$$C(s) = \frac{k_i}{s}\frac{s^2 + 2\zeta a s + a^2}{a^2(1 + sT_f + (sT_f)^2/2)}, \tag{12.17}$$

which has high-frequency roll-off. Selection of the constant T_f for the filter is a compromise between attenuation of high-frequency measurement noise and ro-

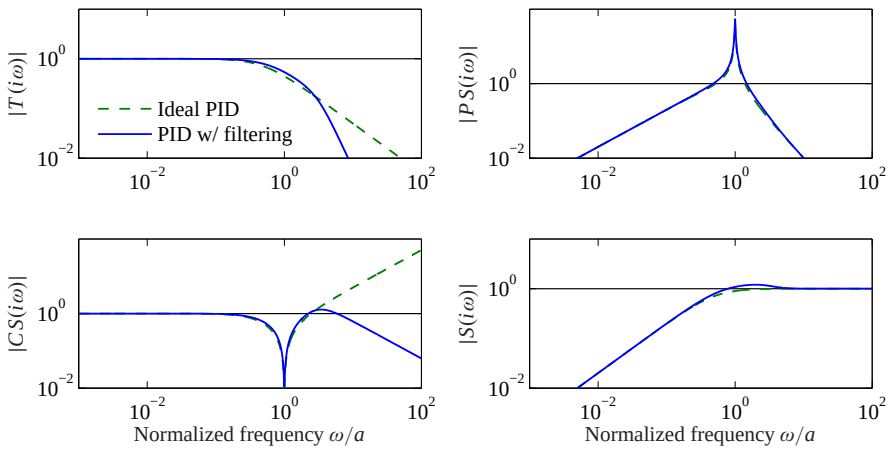


Figure 12.14: Nanopositioning system control via cancellation of the fast process pole. Gain plots for the Gang of Four for PID control with second-order filtering (12.17) are shown by solid lines, and the dashed lines show results for an ideal PID controller without filtering (12.16).

bustness. A large value of T_f reduces the effects of sensor noise significantly, but it also reduces the stability margin. Since the gain crossover frequency without filtering is k_i, a reasonable choice is $T_F = 0.2/T_f$, as shown by the solid curves in Figure 12.14. The plots of $|CS(i\omega)|$ and $|S(i\omega)|$ show that the sensitivity to high-frequency measurement noise is reduced dramatically at the cost of a marginal increase of sensitivity. Notice that the poor attenuation of disturbances with frequencies close to the resonance is not visible in the sensitivity function because of the exact cancellation of poles and zeros.

The designs thus far have the drawback that load disturbances with frequencies close to the resonance are not attenuated. We will now consider a design that actively attenuates the poorly damped modes. We start with an ideal PID controller where the design can be done analytically, and we add high-frequency roll-off. The loop transfer function obtained with this controller is

$$L(s) = \frac{k_d s^2 + k_p s + k_i}{s(s^2 + 2\zeta a s + a^2)}. \tag{12.18}$$

The closed loop system is of third order, and its characteristic polynomial is

$$s^3 + (k_d a^2 + 2\zeta a)s^2 + (k_p + 1)a^2 s + k_i a^2. \tag{12.19}$$

A general third-order polynomial can be parameterized as

$$s^3 + (\alpha_0 + 2\zeta)\omega_0 s^2 + (1 + 2\alpha_0\zeta)\omega_0^2 s + \alpha_0\omega_0^3. \tag{12.20}$$

The parameters α_0 and ζ give the relative configuration of the poles, and the parameter ω_0 gives their magnitudes, and therefore also the bandwidth of the system.

The identification of coefficients of equal powers of s in with equation (12.19)

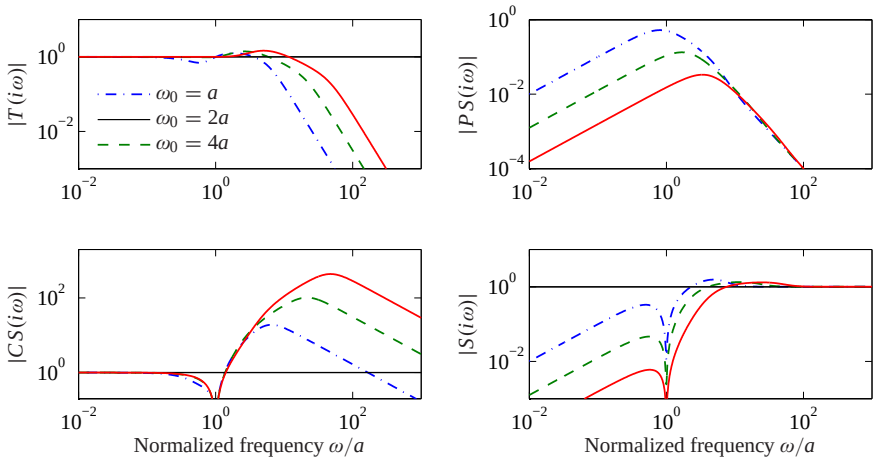


Figure 12.15: Nanopositioner control using active damping. Gain curves for the Gang of Four for PID control of the nanopositioner designed for $\omega_0 = a$ (dash-dotted), $2a$ (dashed), and $4a$ (solid). The controller has high-frequency roll-off and has been designed to give active damping of the oscillatory mode. The different curves correspond to different choices of magnitudes of the poles, parameterized by ω_0 in equation (12.19).

gives a linear equation for the controller parameters, which has the solution

$$k_p = \frac{(1 + 2\alpha_0\zeta)\omega_0^2}{a^2} - 1, \qquad k_i = \frac{\alpha_0\omega_0^3}{a^2}, \qquad k_d = \frac{(\alpha_0 + 2\zeta)\omega_0}{a^2} - 2\zeta a. \quad (12.21)$$

To obtain a design with active damping, it is necessary that the closed loop bandwidth be at least as fast as the oscillatory modes. Adding high-frequency roll-off, the controller becomes

$$C(s) = \frac{k_d s^2 + k_p s + k}{s(1 + sT_f + (sT_f)^2/2)}. \quad (12.22)$$

The value $T_f = T_d/10 = 0.1\,k_d/k$ is a good value for the filtering time constant.

Figure 12.15 shows the gain curves of the Gang of Four for designs with $\zeta = 0.707$, $\alpha_0 = 1$ and $\omega_0 = a$, $2a$ and $4a$. The figure shows that the largest values of the sensitivity function and the complementary sensitivity function are small. The gain curve for PS shows that the load disturbances are now well attenuated over the whole frequency range, and attenuation increases with increasing ω_0. The gain curve for CS shows that large control signals are required to provide active damping. The high gain of CS for high frequencies also shows that low-noise sensors and actuators with a wide range are required. The largest gains for CS are 19, 103 and 434 for $\omega_0 = a$, $2a$ and $4a$, respectively. There is clearly a trade-off between disturbance attenuation and controller gain. A comparison of Figures 12.14 and 12.15 illustrates the trade-offs between control action and disturbance attenuation for the designs with cancellation of the fast process pole and active damping. ∇

12.5 Design for Robust Performance

Control design is a rich problem where many factors have to be taken into account. Typical requirements are that load disturbances should be attenuated, the controller should inject only a moderate amount of measurement noise, the output should follow variations in the command signal well and the closed loop system should be insensitive to process variations. For the system in Figure 12.9 these requirements can be captured by specifications on the sensitivity functions S and T and the transfer functions G_{yd}, G_{un}, G_{yr} and G_{ur}. Notice that it is necessary to consider at least six transfer functions, as discussed Section 11.1. The requirements are mutually conflicting, and it is necessary to make trade-offs. The attenuation of load disturbances will be improved if the bandwidth is increased, but so will the noise injection.

It is highly desirable to have design methods that can guarantee robust performance. Such design methods did not appear until the late 1980s. Many of these design methods result in controllers having the same structure as the controller based on state feedback and an observer. In this section we provide a brief review of some of the techniques as a preview for those interested in more specialized study.

Quantitative Feedback Theory

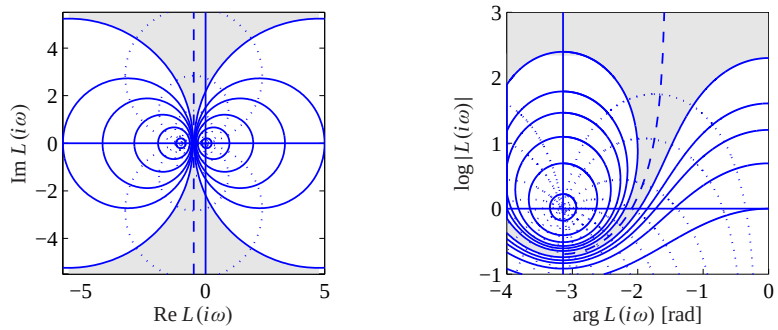
Quantitative feedback theory (QFT) is a graphical design method for robust loop shaping that was developed by I. M. Horowitz [Hor91]. The idea is to first determine a controller that gives a complementary sensitivity that is robust to process variations and then to shape the response to reference signals by feedforward. The idea is illustrated in Figure 12.16a, which shows the level curves of the complementary sensitivity function T on a Nyquist plot. The complementary sensitivity function has unit gain on the line $\operatorname{Re} L(i\omega) = -0.5$. In the neighborhood of this line, significant variations in process dynamics only give moderate changes in the complementary transfer function. The dashed part of the figure corresponds to the region $0.9 < |T(i\omega)| < 1.1$. To use the design method, we represent the uncertainty for each frequency by a region and attempt to shape the loop transfer function so that the variation in T is as small as possible. The design is often performed using the Nichols chart shown in Figure 12.16b.

Linear Quadratic Control

One way to make the trade-off between the attenuation of load disturbances and the injection of measurement noise is to design a controller that minimizes the loss function

$$J = \frac{1}{T}\int_0^T \left(y^2(t) + \rho u^2(t)\right) dt,$$

where ρ is a weighting parameter as discussed in Section 6.3. This loss function gives a compromise between load disturbance attenuation and disturbance injection because it balances control actions against deviations in the output. If all state



(a) Hall chart (b) Nichols chart

Figure 12.16: Hall and Nichols charts. The Hall chart is a Nyquist plot with curves for constant gain and phase of the complementary sensitivity function T. The Nichols chart is the conformal map of the Hall chart under the transformation $N = \log L$ (with the scale flipped). The dashed curve is the line where $|T(i\omega)| = 1$, and the shaded region corresponding to loop transfer functions whose complementary sensitivity changes by no more than $\pm 10\%$ is shaded.

variables are measured, the controller is a state feedback $u = -Kx$ and it has the same form as the controller obtained by eigenvalue assignment (pole placement) in Section 6.2. However, the controller gain is obtained by solving an optimization problem. It has been shown that this controller is very robust. It has a phase margin of at least 60° and an infinite gain margin. The controller is called a *linear quadratic control* or *LQ control* because the process model is linear and the criterion is quadratic.

When all state variables are not measured, the state can be reconstructed using an observer, as discussed in Section 7.3. It is also possible to introduce process disturbances and measurement noise explicitly in the model and to reconstruct the states using a Kalman filter, as discussed briefly in Section 7.4. The Kalman filter has the same structure as the observer designed by eigenvalue assignment in Section 7.3, but the observer gains L are now obtained by solving an optimization problem. The control law obtained by combining linear quadratic control with a Kalman filter is called *linear quadratic Gaussian control* or *LQG control*. The Kalman filter is optimal when the models for load disturbances and measurement noise are Gaussian.

It is interesting that the solution to the optimization problem leads to a controller having the structure of a state feedback and an observer. The state feedback gains depend on the parameter ρ, and the filter gains depend on the parameters in the model that characterize process noise and measurement noise (see Section 7.4). There are efficient programs to compute these feedback and observer gains.

The nice robustness properties of state feedback are unfortunately lost when the observer is added. It is possible to choose parameters that give closed loop systems with poor robustness, similar to Example 12.8. We can thus conclude that there is a

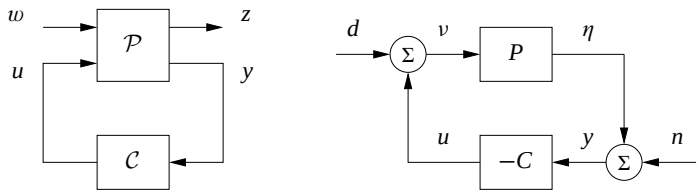


Figure 12.17: H_∞ robust control formulation. The left figure shows a general representation of a control problem used in robust control. The input u represents the control signal, the input w represents the external influences on the system, the output z is the generalized error and the output y is the measured signal. The right figure shows the special case of the basic feedback loop in Figure 12.9 where the reference signal is zero. In this case we have $w = (n, d)$ and $z = (y, -u)$.

fundamental difference between using sensors for all states and reconstructing the states using an observer.

H_∞ Control

Robust control design is often called H_∞ control for reasons that will be explained shortly. The basic ideas are simple, but the details are complicated and we will therefore just give the flavor of the results. A key idea is illustrated in Figure 12.17, where the closed loop system is represented by two blocks, the process $\mathcal{P}$ and the controller $\mathcal{C}$ as discussed in Section 11.1. The process $\mathcal{P}$ has two inputs, the control signal u, which can be manipulated by the controller, and the generalized disturbance w, which represents all external influences, e.g., command signals and disturbances. The process has two outputs, the generalized error z, which is a vector of error signals representing the deviation of signals from their desired values, and the measured signal y, which can be used by the controller to compute u. For a linear system and a linear controller the closed loop system can be represented by the linear system

$$z = H(P(s), C(s))w, \tag{12.23}$$

which tells how the generalized error w depends on the generalized disturbances w. The control design problem is to find a controller C such that the gain of the transfer function H is small even when the process has uncertainties. There are many different ways to specify uncertainty and gain, giving rise to different designs. The names H_2 and H_∞ control correspond to the norms $\|H\|_2$ and $\|H\|_\infty$.

To illustrate the ideas we will consider a regulation problem for a system where the reference signal is assumed to be zero and the external signals are the load disturbance d and the measurement noise n, as shown in Figure 12.17b. The generalized input is $w = (-n, d)$. (The negative sign of n is not essential but is chosen to obtain somewhat nicer equations.) The generalized error is chosen as $z = (\eta, \nu)$, where η is the process output and ν is the part of the load disturbance that is not

compensated by the controller. The closed loop system is thus modeled by

$$z = \begin{bmatrix} y \\ -u \end{bmatrix} = \begin{bmatrix} \dfrac{1}{1+PC} & \dfrac{P}{1+PC} \\ \dfrac{C}{1+PC} & \dfrac{PC}{1+PC} \end{bmatrix} \begin{bmatrix} n \\ d \end{bmatrix} = H(P,C) \begin{bmatrix} n \\ d \end{bmatrix}, \tag{12.24}$$

which is the same as equation (12.23). A straightforward calculation shows that

$$\|H(P,C))\|_\infty = \sup_\omega \frac{\sqrt{(1+|P(i\omega)|^2)(1+|C(i\omega)|^2)}}{|1+P(i\omega)C(i\omega)|}. \tag{12.25}$$

There are numerical methods for finding a controller such that $\|H(P,C)\|_\infty < \gamma$, if such a controller exists. The best controller can then be found by iterating on γ. The calculations can be made by solving *algebraic Riccati* equations, e.g., by using the command `hinfsyn` in MATLAB. The controller has the same order as the process and the same structure as the controller based on state feedback and an observer; see Figure 7.7 and Theorem 7.3.

Notice that if we minimize $\|H(P,C)\|_\infty$, we make sure that the transfer functions $G_{yd} = P/(1+PC)$, representing the transmission of load disturbances to the output, and $G_{un} = -C/(1+PC)$, representing how measurement noise is transmitted to the control signal, are small. Since the sensitivity and the complementary sensitivity functions are also elements of $H(P,C)$, we have also guaranteed that the sensitivities are less than γ. The design methods thus balance performance and robustness.

There are strong robustness results associated with the H_∞ controller. It follows from equations (12.4) and (12.25) that

$$\|H(P,C)\|_\infty = \frac{1}{\delta_\nu(P,-1/C)}. \tag{12.26}$$

The inverse of $\|H(P,C)\|_\infty$ is thus equal to the Vinnicombe distance between P and $-1/C$ and can therefore be interpreted as a *generalized stability margin*. Compare this with s_m, which we defined as the shortest distance between the Nyquist curve of the loop transfer function and the critical point -1. It also follows that if we find a controller C with $\|H(P,C)\|_\infty < \gamma$, then this controller will stabilize any process P_* such that $\delta_\nu(P,P_*) < 1/\gamma$.

Disturbance Weighting

Minimizing the gain $\|H(P,C)\|_\infty$ means that the gains of all individual signal transmissions from disturbances to outputs are less than γ for all frequencies of the input signals. The assumption that the disturbances are equally important and that all frequencies are also equally important is not very realistic; recall that load disturbances typically have low frequencies and measurement noise is typically dominated by high frequencies. It is straightforward to modify the problem so that disturbances of different frequencies are given different emphasis, by introducing

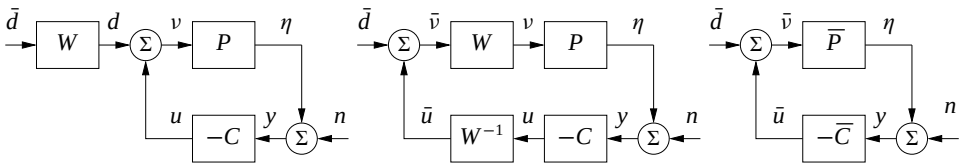


Figure 12.18: Block diagrams of a system with disturbance weighting. The left figure provides a frequency weight on processes disturbances. Through block diagram manipulation, this can be converted to the standard problem on the right.

a weighting filter on the load disturbance as shown in Figure 12.18. For example, low-frequency load disturbances will be enhanced by choosing W as a low-pass filter because the actual load disturbance is $W\bar{d}$.

By using block diagram manipulation as shown in Figure 12.18, we find that the system with frequency weighting is equivalent to the system with no frequency weighting in Figure 12.18 and the signals are related through

$$\bar{z} = \begin{bmatrix} y \\ \bar{u} \end{bmatrix} \begin{bmatrix} \dfrac{1}{1+\overline{P}\,\overline{C}} & \dfrac{\overline{P}}{1+\overline{P}\,\overline{C}} \\ \dfrac{\overline{C}}{1+\overline{P}\,\overline{C}} & \dfrac{\overline{P}\,\overline{C}}{1+\overline{P}\,\overline{C}} \end{bmatrix} \begin{bmatrix} n \\ \bar{d} \end{bmatrix} = H(\overline{P}, \overline{C})\bar{w}, \tag{12.27}$$

where $\overline{P} = PW$ and $\overline{C} = W^{-1}C$. The problem of finding a controller $\overline{C}$ that minimizes the gain of $H(\overline{P}, \overline{C})$ is thus equivalent to the problem without disturbance weighting; having obtained $\overline{C}$, the controller for the original system is then $C = W\overline{C}$. Notice that if we introduce the frequency weighting $W = k/s$, we will automatically get a controller with integral action.

Limits of Robust Design

There is a limit to what can be achieved by robust design. In spite of the nice properties of feedback, there are situations where the process variations are so large that it is not possible to find a linear controller that gives a robust system with good performance. It is then necessary to use other types of controllers. In some cases it is possible to measure a variable that is well correlated with the process variations. Controllers for different parameter values can then be designed and the corresponding controller can be chosen based on the measured signal. This type of control design is called *gain scheduling*. The cruise controller is a typical example where the measured signal could be gear position and velocity. Gain scheduling is the common solution for high-performance aircraft where scheduling is done based on Mach number and dynamic pressure. When using gain scheduling, it is important to make sure that switches between the controllers do not create undesirable transients (often referred to as *bumpless transfer*).

If it is not possible to measure variables related to the parameters, *automatic tuning* and *adaptive control* can be used. In automatic tuning the process dynamics are measured by perturbing the system, and a controller is then designed automat-

ically. Automatic tuning requires that parameters remain constant, and it has been widely applied for PID control. It is a reasonable guess that in the future many controllers will have features for automatic tuning. If parameters are changing, it is possible to use adaptive methods where process dynamics are measured online.

12.6 Further Reading

The topic of robust control is a large one, with many articles and textbooks devoted to the subject. Robustness was a central issue in classical control as described in Bode's classical book [Bod45]. Robustness was deemphasized in the euphoria of the development of design methods based on optimization. The strong robustness of controllers based on state feedback, shown by Anderson and Moore [AM90], contributed to the optimism. The poor robustness of output feedback was pointed out by Rosenbrock [RM71], Horowitz [Hor75] and Doyle [Doy78] and resulted in a renewed interest in robustness. A major step forward was the development of design methods where robustness was explicitly taken into account, such as the seminal work of Zames [Zam81]. Robust control was originally developed using powerful results from the theory of complex variables, which gave controllers of high order. A major breakthrough was made by Doyle, Glover, Khargonekar and Francis [DGKF89], who showed that the solution to the problem could be obtained using Riccati equations and that a controller of low order could be found. This paper led to an extensive treatment of H_∞ control, including books by Francis [Fra87], McFarlane and Glover [MG90], Doyle, Francis and Tannenbaum [DFT92], Green and Limebeer [GL95], Zhou, Doyle and Glover [ZDG96], Skogestand and Postlethwaite [SP05] and Vinnicombe [Vin01]. A major advantage of the theory is that it combines much of the intuition from servomechanism theory with sound numerical algorithms based on numerical linear algebra and optimization. The results have been extended to nonlinear systems by treating the design problem as a game where the disturbances are generated by an adversary, as described in the book by Basar and Bernhard [BB91]. Gain scheduling and adaptation are discussed in the book by Åström and Wittenmark [ÅW08].

Exercises

12.1 Consider systems with the transfer functions $P_1 = 1/(s+1)$ and $P_2 = 1/(s+a)$. Show that P_1 can be changed continuously to P_2 with bounded additive and multiplicative uncertainty if $a > 0$ but not if $a < 0$. Also show that no restriction on a is required for feedback uncertainty.

12.2 Consider systems with the transfer functions $P_1 = (s+1)/(s+1)^2$ and $P_2 = (s+a)/(s+1)^2$. Show that P_1 can be changed continuously to P_2 with bounded feedback uncertainty if $a > 0$ but not if $a < 0$. Also show that no restriction on a is required for additive and multiplicative uncertainties.

12.3 (Difference in sensitivity functions) Let $T(P, C)$ be the complementary sensitivity function for a system with process P and controller C. Show that

$$T(P_1, C) - T(P_2, C) = \frac{(P_1 - P_2)C}{(1 + P_1C)(1 + P_2C)},$$

and derive a similar formula for the sensitivity function.

12.4 (The Riemann sphere) Consider systems with the transfer functions $P_1 = k/(s + 1)$ and $P_2 = k/(s - 1)$. Show that

$$d(P_1, P_2) = \frac{2k}{1 + k^2}, \qquad \delta_\nu(P_1, P_2) = \begin{cases} 1, & \text{if } k < 1 \\ \dfrac{2k}{1 + k^2} & \text{otherwise.} \end{cases}$$

Use the Riemann sphere to show geometrically that $\delta_\nu(P_1, P_2) = 1$ if $k < 1$. (Hint: It is sufficient to evaluate the transfer function for $\omega = 0$.)

12.5 (Stability margins) Consider a feedback loop with a process and a controller having transfer functions P and C. Assume that the maximum sensitivity is $M_s = 2$. Show that the phase margin is at least 30° and that the closed loop system will be stable if the gain is changed by 50%.

12.6 (Bode's ideal loop transfer function) Make Bode and Nyquist plots of Bode's ideal loop transfer function. Show that the phase margin is $\varphi_m = 180^\circ – 90^\circ n$ and that the stability margin is $s_m = \arcsin \pi(1 - n/2)$.

12.7 Consider a process with the transfer function $P(s) = k/(s(s + 1))$, where the gain can vary between 0.1 and 10. A controller that is robust to these gain variations can be obtained by finding a controller that gives the loop transfer function $L(s) = 1/(s\sqrt{s})$. Suggest how the transfer function can be implemented by approximating it by a rational function.

12.8 (Smith predictor) The *Smith predictor*, a controller for systems with time delays, is a special version of Figure 12.8a with $P(s) = e^{-s\tau}P_0(s)$ and $C(s) = C_0(s)/(1 + C_0(s)P(s))$. The controller $C_0(s)$ is designed to give good performance for the process $P_0(s)$. Show that the sensitivity functions are

$$S(s) = \frac{1 + (1 - e^{-s\tau})P_0(s)C_0(s)}{1 + P_0(s)C_0(s)}, \qquad T(s) = \frac{P_0(s)C_0(s)}{1 + P_0(s)C_0(s)}e^{-s\tau}.$$

12.9 (Ideal delay compensator) Consider a process whose dynamics are a pure time delay with transfer function $P(s) = e^{-s}$. The ideal delay compensator is a controller with the transfer function $C(s) = 1/(1 - e^{-s})$. Show that the sensitivity functions are $T(s) = e^{-s}$ and $S(s) = 1 - e^{-s}$ and that the closed loop system will be unstable for arbitrarily small changes in the delay.

12.10 (Vehicle steering) Consider the Nyquist curve in Figure 12.11. Explain why part of the curve is approximately a circle. Derive a formula for the center and the radius and compare with the actual Nyquist curve.

12.11 Consider a process with the transfer function

$$P(s) = \frac{(3+3)(s+200)}{(s+1)(s^2+10s+40)(s+40)}.$$

Discuss suitable choices of closed loop poles for a design that gives dominant poles with undamped natural frequency 1 and 10.

12.12 (AFM nanopositioning system) Consider the design in Example 12.10 and explore the effects of changing parameters α_0 and ζ_0.

12.13 (H_∞ control) Consider the matrix $H(P, C)$ in equation (12.24). Show that it has the singular values

$$\sigma_1 = 0, \qquad \sigma_2 = \bar{\sigma} = \sup_\omega \frac{\sqrt{(1+|P(i\omega)|^2)(1+|C(i\omega)|^2)}}{|1+P(i\omega)C(i\omega)|} = \|H(P,C))\|_\infty.$$

Also show that $\bar{\sigma} = 1/d_v(P, -1/C)$, which implies that $1/\bar{\sigma}$ is a generalization of the closest distance of the Nyquist plot to the critical point.

12.14 Show that

$$\delta_v(P, -1/C) = \inf_\omega \frac{|P(i\omega) + 1/C(i\omega)|}{\sqrt{(1+|P(i\omega)|^2)(1+1/|C(i\omega)|^2)}} = \frac{1}{\|H(P,C))\|_\infty}.$$

12.15 Consider the system

$$\frac{dx}{dt} = Ax + Bu = \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix} x + \begin{bmatrix} a-1 \\ 1 \end{bmatrix} u, \qquad y = Cx = \begin{bmatrix} 0 & 1 \end{bmatrix} y.$$

Design a state feedback that gives $\det(sI - BK) = s^2 + 2\zeta_c\omega_c s + \omega_c^2$, and an observer with $\det(sI - LC) = s^2 + 2\zeta_o\omega_o s + \omega_o^2$ and combine them using the separation principle to get an output feedback. Choose the numerical values $a = 1.5$, $\omega_c = 5$, $\zeta_c = 0.6$ and $\omega_o = 10$, $\zeta_o = 0.6$. Compute the eigenvalues of the perturbed system when the process gain is increased by 2%. Also compute the loop transfer function and the sensitivity functions. Is there a way to know beforehand that the system will be highly sensitive?

12.16 (Robustness using the Nyquist criterion) Another view of robust performance can be obtained through appeal to the Nyquist criterion. Let $S_{\max}(i\omega)$ represent a desired upper bound on our sensitivity function. Show that the system provides this level of performance subject to additive uncertainty Δ if the following inequality is satisfied:

$$|1 + \tilde{L}| = |1 + L + C\Delta| > \frac{1}{|S_{\max}(i\omega)|} \qquad \text{for all } \omega \geq 0. \tag{12.28}$$

Describe how to check this condition using a Nyquist plot.