.

LECTURE NOTES

Control System Analysis and Design

Part 2: Analysis of Feedback Control Systems

Contents/concepts:

- Open-loop and closed-loop control:
  - plant model
  - static open-loop control
  - static closed-loop control
  - control and plant inversion
- Dynamic closed-loop control: sensitivity functions
- Stability and stability robustness
- Design goals and loop shaping
- Limitations of performance

Contents

1 Open-loop and Closed-loop Control **3**
1.1 Plant model . . . 3
1.2 Static open-loop control . . . 4
1.3 Static closed-loop control . . . 6
1.4 Control and plant inversion . . . 8

2 Dynamic Closed-loop Control: sensitivity functions **10**

3 Closed-loop Stability and Stability Robustness **14**
3.1 Internal stability . . . 14
3.2 Nyquist criterion . . . 15
3.3 Robustness . . . 16
3.3.1 Stability margins . . . 16

4 Design Goals **18**
4.1 Closed-loop stability . . . 18
4.2 Stability robustness . . . 19
4.3 Disturbance attenuation . . . 19
4.4 Command response . . . 19
4.5 Plant capacity . . . 20
4.6 Measurement noise . . . 20
4.7 Performance robustness . . . 20
4.8 Summary . . . 21

The purpose of these notes is to analyse the properties of feedback control systems. The focus is on the *analysis* of feedback systems, i.e. for a given plant and given compensator, what are the performance, stability, and robustness characteristics of the feedback loop? The process of determining the compensator to achieve certain feedback loop specifications is known as *control design* or *control synthesis*, and is treated in detail elsewhere in the lecture material and in [1, 2].

1 Open-loop and Closed-loop Control

We begin by considering possible open- and closed-loop control strategies, and we contrast their properties. A generic open-loop control structure is shown in figure 1. The part of the system to be controlled is referred to as the *plant*, and we aim to control the plant output y so that it follows the reference input r in some desirable way. The controller processes the reference signal to produce a control input u for the plant. The structure is called open loop because there is no feedback of the plant output y. A generic closed-loop control structure

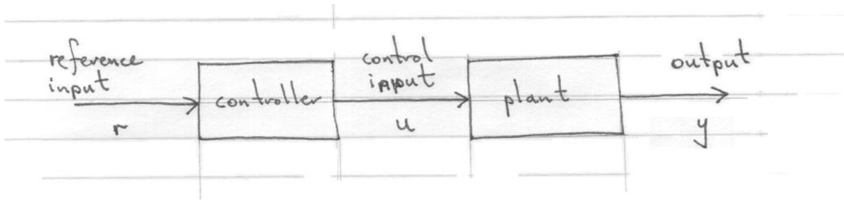

Figure 1: Generic open-loop control structure.

is shown in figure 2. The difference here is that a sensor is used (not shown) to provide a measure of the plant output (this usually introduces sensor noise). The measured output, in addition to the reference input, is processed by the controller.

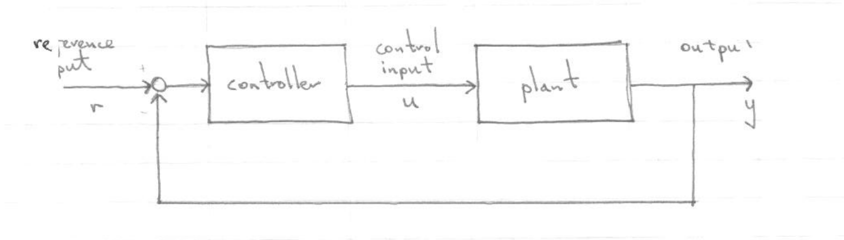

Figure 2: Generic closed-loop control structure.

Of particular importance is the task of trying to maintain the output close to the reference input, *in the face of disturbance signals and changes in the plant parameters.* In a first analysis, we consider the steady-state (i.e. static) properties of open-loop and closed-loop control with respect to these factors.

1.1 Plant model

For illustration consider the problems of knee-joint angle control, and posture control; the dynamic models for these problems were considered previously. The standard equations of motion for these two systems have exactly the same structure - only the definitions of the parameters ζ and ω_n are specific to the problem. In both cases we define the joint moment (i.e. knee joint or ankle joint) to have the form $\tau = \tau_e + \tau_i + \tau_d$. Here, τ_e, as before, is the equilibrium moment corresponding to the equilibrium position θ_e. The signal τ_i is a *control input* which is generated by the compensator to be chosen in the design process. τ_d, on the other hand, is a *disturbance signal* (a moment) over which the designer has no influence.

Referring back to our derivation of the equations of motion for the knee joint and ankle joint, we find readily that the standard equation of motion in both cases is given by

$$\ddot{\theta} + 2\zeta\omega_n\dot{\theta} + \omega_n^2\theta = \frac{1}{I}(\tau_i + \tau_d) \tag{1}$$

Taking Laplace transforms and ignoring initial conditions results in

$$\Theta(s) = \frac{1/I}{s^2 + 2\zeta\omega_n s + \omega_n^2}(\Gamma_i(s) + \Gamma_d(s)) \tag{2}$$

Here, $\Gamma_i(s)$ and $\Gamma_d(s)$ denote the Laplace transforms of $\tau_i(t)$ and $\tau_d(t)$. The plant transfer function is denoted as $P(s)$ and is given by

$$P(s) = \frac{1/I}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{3}$$

The structure of an open-loop controller for this plant is depicted in figure 3 and a closed-loop controller is shown in figure 4. In both of these figures the effect of the disturbance moment τ_d is represented at the output of the plant by the term $D(s) = P(s)\Gamma_d(s)$ (clearly, this is exactly equivalent to $\Gamma_d(s)$ acting instead at the plant input).

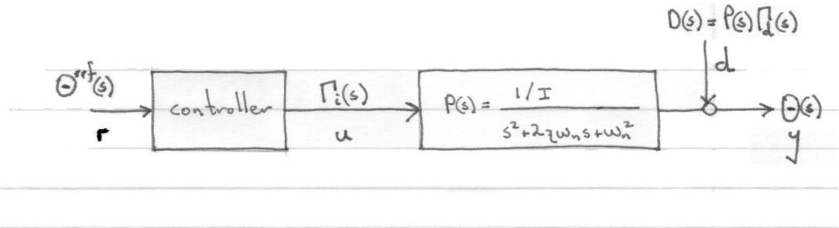

Figure 3: Open-loop control.

1.2 Static open-loop control

It is desirable to keep the output close to the reference, at least in steady-state conditions. We first consider a static (non-dynamic) open-loop controller described by $u = cr$. The static controller gain c is chosen to give $y = r$ in the absence of disturbances (i.e. with $d = 0$). We denote the static gain of the plant as P_{ss}. From equation (3) we have

$$P_{ss} = \frac{1}{I\omega_n^2}$$

which means that for constant signals u and d the steady-state output y_{ss} will be

$$y_{ss} = \frac{1}{I\omega_n^2}u + d$$

The appropriate value of c to achieve our goal is clearly given by the inverse of the plant's static gain, i.e.

$$c = P_{ss}^{-1} = I\omega_n^2 \Rightarrow u = I\omega_n^2 r \tag{4}$$

The steady-state output with this control strategy is then

$$y_{ss} = r + d \tag{5}$$

Thus, when $d = 0$ we achieve our goal of $y = r$.

Effect of disturbances: to analyse the effect of disturbances we define the steady-state *tracking error* e as $e = r - y_{ss}$. From equation (5) we get $e = -d$. For the specific problems of knee angle or posture control the error is $e = -1/(I\omega_n^2)\tau_d$. This shows that with open-loop control the error is directly proportional to the size of the disturbance. Importantly, the designer has no influence at all over the value $1/(I\omega_n^2)$ which determines the size of the error, since it is independent of the controller gain. This is a major disadvantage of open-loop control.

Effect of plant gain changes: changes in plant operating conditions or in the physical properties of the system lead to changes in the plant parameters. In the knee-angle or posture control problems, for example, the plant parameters are dependent on the equilibrium angle θ_e. In addition, the physical properties m, l or I may change (by the addition of weights at the ankle joint, for example), or they may be only approximately known. How do changes in the plant affect the overall system performance?

We assume a change in the static gain of the plant from P_{ss} to $P_{ss} + \delta P_{ss}$, but that the compensator is designed as before as $c = P_{ss}^{-1}$. Denoting the nominal overall open-loop static gain as T_{ol} we have

$$T_{ol} + \delta T_{ol} = P_{ss}^{-1}(P_{ss} + \delta P_{ss}) = 1 + \frac{\delta P_{ss}}{P_{ss}}$$

Since $T_{ol} = 1$ the overall gain changes by $\delta T_{ol} = \delta P_{ss}/P_{ss}$. In relative terms, the overall change is

$$\frac{\delta T_{ol}}{T_{ol}} = \frac{\delta P_{ss}}{P_{ss}} \tag{6}$$

Thus, a 10% change in the plant gain will result directly in a 10% change in the overall system gain T_{ol}. H.W. Bode defined the ratio of $\delta T_{ol}/T_{ol}$ and $\delta P_{ss}/P_{ss}$ as the **sensitivity**, $\mathcal{S}$, of the plant. Thus,

$$\frac{\delta T_{ol}}{T_{ol}} = \mathcal{S}\frac{\delta P_{ss}}{P_{ss}} \tag{7}$$

From equations (6) and (7) we see that for open-loop control $\mathcal{S} = 1$ which means that open-loop control is highly sensitive to plant variations, since there is a one-to-one relationship between relative plant variations and the resulting relative variation in the overall system. Sensitivity is a key concept to which we will return frequently.

The most important properties of open-loop control can be summarised as follows:

- Open-loop control provides no compensation for disturbances.
- The overall system is highly sensitive to possible changes in the plant and in operating conditions.
- A further key issue, not considered above, is that the plant may be unstable. Clearly, with open-loop control the overall system remains unstable: open-loop control cannot stabilise an unstable plant.

1.3 Static closed-loop control

We now analyse some of the basic properties of feedback control, in particular the effect of disturbances and plant variations. The feedback control structure we consider is shown in figure 4. In the first instance we consider again a static controller $C(s) = c$. The closed-loop

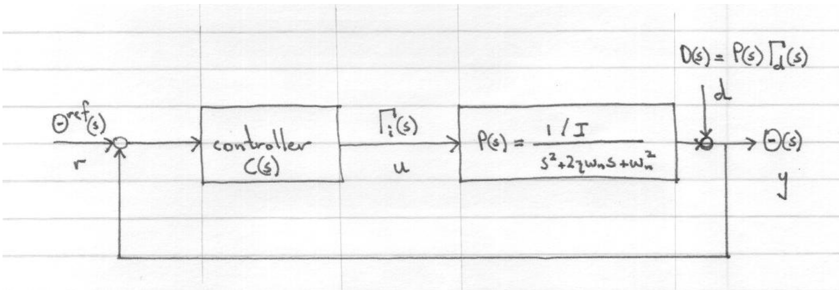

Figure 4: Closed-loop control.

equation for the plant output in steady-state is readily found to be

$$y_{ss} = \frac{cP_{ss}}{1 + cP_{ss}} r + \frac{1}{1 + cP_{ss}} d \tag{8}$$

for constant signals $r(t) = r$ and $d(t) = d$.

Effect of disturbances: if it is possible to choose the controller gain c to be large, such that $cP_{ss} \gg 1$, then clearly we have $y_{ss} \approx r$, regardless of the value of the disturbance d. In fact, comparison of equations (5) and (8) shows that as long as $cP_{ss} > 0$ then the tracking error caused by disturbances will be less than that for the open-loop case by a factor of $1 + cP_{ss}$. The first key property of feedback is therefore that disturbances can be significantly attenuated.

Effect of plant gain changes: the closed-loop static gain is defined as T_{cl}. The sensitivity $\mathcal{S}$ is defined as before as the ratio of relative change in the overall system gain resulting from a given relative change in the plant:

$$\frac{\delta T_{cl}}{T_{cl}} = \mathcal{S} \frac{\delta P_{ss}}{P_{ss}} \tag{9}$$

By rearrangement we see that $\mathcal{S}$ can be computed as

$$\mathcal{S} = \frac{P_{ss}}{T_{cl}} \left(\frac{dT_{cl}}{dP_{ss}} \right) \tag{10}$$

Now, from (8) the closed-loop gain from r to y is given by

$$T_{cl} = \frac{cP_{ss}}{1 + cP_{ss}} \tag{11}$$

By differentiation we get

$$\frac{dT_{cl}}{dP_{ss}} = \frac{c}{(1 + cP_{ss})^2} \tag{12}$$

Substitution of T_{cl} and dT_{cl}/dP_{ss} from (11)–(12) in (10) yields

$$\mathcal{S} = \frac{1}{1 + cP_{ss}} \tag{13}$$

Thus, if it is possible to choose the controller gain c to be large, such that $cP_{ss} \gg 1$, then $\mathcal{S} \approx 0$, which means that the closed-loop gain is largely unaffected by changes in the plant (see equation (9)). In fact, comparison of equations (6) and (13) shows that as long as $cP_{ss} > 0$ then the tracking error caused by changes in the plant will be less than that for the open-loop case by a factor of $1 + cP_{ss}$. This results illustrates a further key feature of feedback.

Comparison of equations (8) and (13) reveals that $\mathcal{S}$ is also equivalent to the transfer-function between the disturbance d and the output y. This establishes a direct link between the disturbance response and the closed-loop sensitivity to plant variations.

The most important properties of closed-loop control can be summarised as follows:

- With feedback control the system errors can be made insensitive to disturbances, i.e. disturbances can be significantly attenuated. The steady-state tracking error caused by disturbances will be less than that for the open-loop case by a factor of $\mathcal{S}^{-1} = 1 + cP_{ss}$.
- System errors are less sensitive to possible changes in the plant and in operating conditions. Compared to open-loop control, the steady-state error is less sensitive by a factor of $\mathcal{S}^{-1} = 1 + cP_{ss}$.
- A further key issue, not considered above, is that unstable systems can be stabilised by feedback. The dynamic closed-loop transfer function is $C(s)P(s)/(1 + C(s)P(s))$. Even if $P(s)$ is unstable, the closed-loop transfer function may be stable. Stability is considered in more detail later.

As we have now seen, feedback may result in very useful closed-loop properties. However, there are some potential drawbacks. These are summarised here, and studied in depth later:

- High gain feedback reduces the effect of disturbances and plant variations. However, high gain may result in large control inputs to the plant which exceed the plant capacity. Thus, the beneficial effects of feedback are limited by available plant capacity.

- Despite the fact that feedback may stabilise an unstable system, making the gain of the system large may in itself result in an unstable feedback system.

- Feedback requires using a sensor to measure the output. All sensors have limited accuracy, resulting in the injection of *measurement noise* into the loop. Measurement noise amplification may result, and this is undesirable.

1.4 Control and plant inversion

We have now considered the static properties of open and closed-loop control strategies. Open-loop control requires inversion of the plant's static gain (see equation (4)). However, this approach will only achieve the objectives if the plant is exactly known, if there are no disturbances, and if the plant is stable. With closed-loop control, the objectives can be achieved despite plant uncertainty (parameter variations and disturbances), provided that the controller gain can be made sufficiently high.

We will now see that one interesting property of feedback is that it generates an approximate inverse of the plant dynamics. This allows us to make a connection between open and closed-loop strategies, and provides insight into how feedback works. The idea of plant inversion is that if we know the effect of an input on the plant output (i.e. we have an accurate model of the plant), then to achieve a desired behaviour we need to invert the plant input-output relationship to determine the required plant input.

Consider the open-loop control structure shown in figure 5. Here, the plant dynamics are

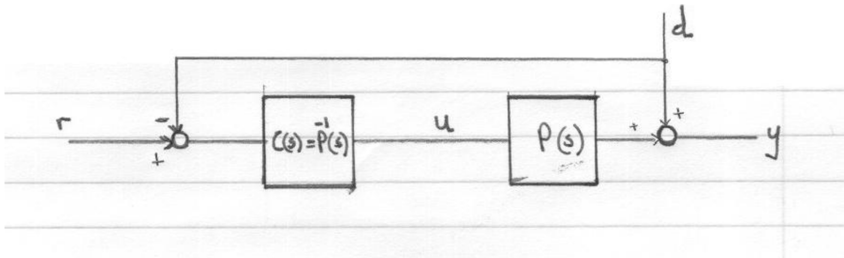

Figure 5: Open-loop control via explicit plant inversion.

given by $P(s)$ and the dynamic controller is given by the plant inverse $P^{-1}(s)$. The input to the controller includes the reference signal r and a measurement of the disturbance d, i.e. the signal $r - d$ is the controller input. With this structure it is clear that exact reference tracking, $y = r$, is achieved. However, the limitations of this approach are:

- the plant dynamics $P(s)$ must be known exactly;

- the plant dynamics must be stable;

- the disturbance d must be exactly measurable;
- the plant inverse dynamics $P^{-1}(s)$ must be stable and realisable (proper).

These are serious limitations. However, we will now see that high-gain feedback provides a mechanism for implicitly generating an approximate inverse of the plant. Consider the control structure shown in figure 6. Here, the plant model $P(s)$ is placed in a feedback configuration around the controller $C(s)$. Note that this realisation is still referred to as an open-loop configuration, as far as the plant is concerned, since the feedback provides no information about what is going on in the plant.

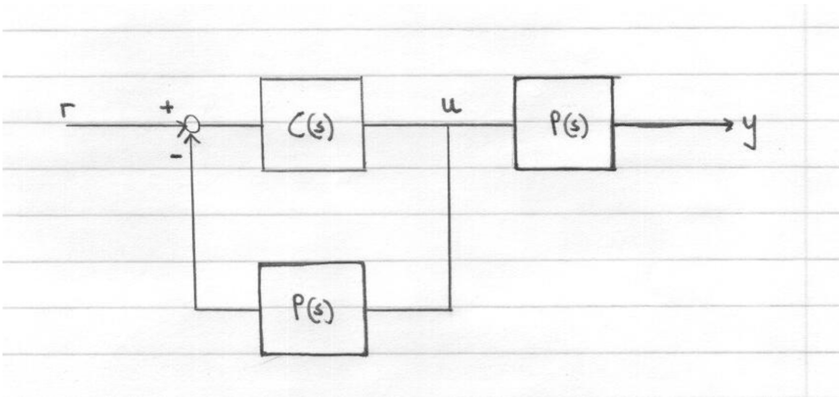

Figure 6: Open-loop control via implicit plant inversion.

In the first instance consider a static controller $C(s) = c$. The relationship between r and u is then given by $U(s) = [c/(1 + cP(s))]R(s)$. With very high controller gain we have $|cP| \gg 1$ and therefore $U(s)/R(s) \approx P^{-1}(s)$, which is an approximate inverse of the plant dynamics. In the absence of disturbances we have $y \approx r$ as desired. For reasons which will be discussed in detail below, it is unrealistic and undesirable to have very high controller gain for all possible frequencies of plant operation. Thus, in general we use dynamic controllers $C(s)$. Approximate plant inversion is then achieved *at those frequencies where* $|CP| \gg 1$, since at those frequencies we have

$$\frac{U(s)}{R(s)} = \frac{C(s)}{1 + C(s)P(s)} \approx P^{-1}(s)$$

and consequently $y \approx r$. These results show that high gain feedback allows us to achieve approximate reference tracking without the need for explicit inversion of the plant dynamics. However, since this is an open-loop strategy, we still have the problems listed above: there is no compensation for disturbances and plant variations; the plant dynamics must be accurately known and must be stable; the plant inverse dynamics must be stable. The solution is to move to a closed-loop control strategy, as shown in figure 7.

The first thing to notice about the schemes shown in figures 6 and 7 is that they are equivalent in the relationship from reference r to output y; in both cases we have $Y/R = CP/(1 + CP)$. Secondly, for those frequencies where the loop gain $|CP|$ is high, i.e. $|CP| \gg 1$, both structures achieve $y \approx r$. Since the two structures are equivalent in this respect, this means that the closed-loop structure can also be thought of as implicitly generating an approximate plant inverse (the inversion is clearly most accurate at those frequencies

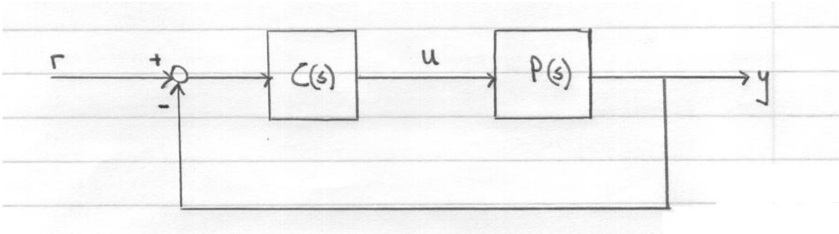

Figure 7: Closed-loop control.

where the loop gain is highest). The third and most important point is that despite this reference-output equivalence, the open-loop strategy has only an “internal” feedback, with no information about processes happening within the plant. In closed-loop control the plant output is measured for feedback, and therefore information about plant disturbances and variations is available to the controller and can be compensated for. The properties of dynamic closed-loop control are considered in depth below. For the moment we may summarise as follows:

- closed-loop control can compensate for disturbances acting on the plant;
- closed-loop control can compensate for variations in plant characteristics;
- unstable plants may be stabilised by feedback;
- closed-loop control implicitly performs approximate inversion of the plant dynamics, at least for those frequencies where the loop gain is high.

These observations echo the summary provided at the end of section 1.3. However, as noted there, the benefits of feedback do not come without a cost:

- high feedback gain may result in large inputs to the plant, possibly exceeding plant capacity;
- blindly making the feedback gain large may de-stabilise an otherwise stable system;
- measurement of the plant output is required. Measurement will be of limited accuracy and will introduce unwanted measurement noise into the system.

The task of feedback system design essentially amounts to finding the best trade-off between these different factors.

2 Dynamic Closed-loop Control: sensitivity functions

For the remainder of this chapter we consider the properties of dynamic closed-loop control, using the two-degree-of-freedom structure shown in figure 8. This generic structure has a control signal (plant input) u, an output disturbance d, and a measurement noise signal n. We also introduce a reference pre-filter F which allows the command response to be shaped independently of the feedback loop properties. At this point we introduce the concept of the *nominal plant*, denoted as P_o. This refers to a simplified model, often a linear model, which provides an approximation of the *true plant*. The nominal plant provides a basis for the design and analysis of the compensator C, which is later applied to the true plant. We initially analyse the nominal feedback loop, and return later to look at the errors which arise when the controller is applied to the true system.

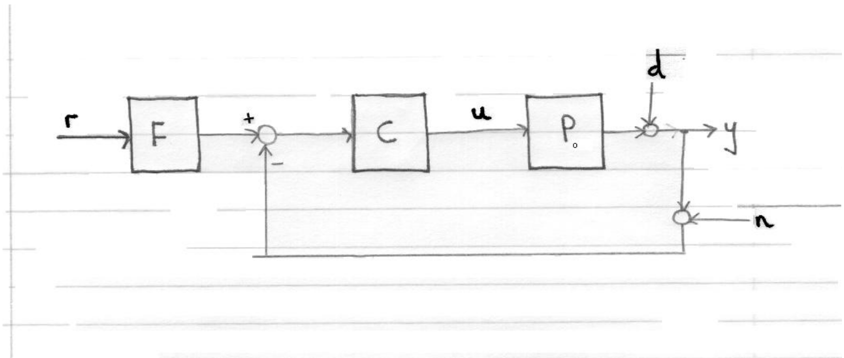

Figure 8: Two-degree-of-freedom control architecture.

For the two-degree-of-freedom architecture shown in figure 8 the nominal closed-loop equation for the output y is

$$y = \frac{CP_o}{1+CP_o}Fr + \frac{1}{1+CP_o}d - \frac{CP_o}{1+CP_o}n \tag{14}$$

The *nominal sensitivity function* S_o and the *nominal complementary sensitivity function* T_o are defined as

$$S_o = \frac{1}{1+CP_o}, \qquad T_o = \frac{CP_o}{1+CP_o} \tag{15}$$

and these definitions allow equation (14) to be written

$$y = T_oFr + S_od - T_on \tag{16}$$

We see that S_o determines the effect of disturbances on the output, since it is the transfer function from d to y. The function T_o describes the effect of measurement noise, since it is the transfer function from n to y. T_o additionally determines the response from the filtered reference signal Fr to y. For this reason T_o is commonly referred to as the *closed-loop transfer function*. The pre-filter F allows the overall command response to be shaped largely arbitrarily.

The closed-loop equation for the control signal u is found to be

$$u = \frac{C}{1+CP_o}Fr - \frac{C}{1+CP_o}(d+n) \tag{17}$$

It is clear that the transfer function $C/(1+CP_o)$ plays a key role in determining the effect on the control signal of the reference, disturbance and measurement noise signals. This transfer function is called the *nominal control sensitivity function* and is denoted as S_{uo}:

$$S_{uo} = \frac{C}{1+CP_o} \tag{18}$$

We note that S_{uo} is related to the other sensitivity functions by $S_{uo} = CS_o = T_o/P_o$. Using (18) the closed-loop relation (17) can be written

$$u = S_{uo}(Fr - d - n) \tag{19}$$

The various sensitivity functions are summarised as

Sensitivity:	$S_o = \dfrac{1}{1 + CP_o}$	(20)
Complementary sensitivity:	$T_o = \dfrac{CP_o}{1 + CP_o}$	(21)
Control sensitivity:	$S_{uo} = \dfrac{C}{1 + CP_o}$	(22)

The sensitivity function S_o is related to the sensitivity $\mathcal{S}$ discussed previously in relation to static open- and closed-loop control (see equations (7) and (13)). If the closed-loop transfer function (complementary sensitivity function) is T_o then its sensitivity with respect to plant variations, denoted $S_{P_o}^{T_o}$, is

$$S_{P_o}^{T_o} = \frac{P_o}{T_o}\left(\frac{\partial T_o}{\partial P_o}\right) \tag{23}$$

Using equation (21) and differentiating we obtain

$$S_{P_o}^{T_o} = \frac{1}{1 + CP_o} \tag{24}$$

which is equivalent to S_o in equation (20), i.e. $S_{P_o}^{T_o} = S_o$. This shows that in frequency ranges where $|S_o|$ can be made small, which means having large loop gain ($|CP_o| \gg 1$), T_o will be relatively insensitive to variations in the plant.

By a similar analysis we can investigate the sensitivity of S_o to plant variations. Recall that S_o determines the effect of disturbances on the plant output. Its sensitivity with respect to plant variations, denoted as $S_{P_o}^{S_o}$, is

$$S_{P_o}^{S_o} = \frac{P_o}{S_o}\left(\frac{\partial S_o}{\partial P_o}\right) \tag{25}$$

Using equation (20) and differentiating we obtain

$$S_{P_o}^{S_o} = \frac{-CP_o}{1 + CP_o} \tag{26}$$

which is equivalent to $-T_o$ from equation (21), i.e. $S_{P_o}^{S_o} = -T_o$. Thus, in frequency ranges where $|T_o|$ is small, which means having small loop gain ($|CP_o| \ll 1$), S_o is relatively insensitive to variations in the plant.

These results reveal a remarkable *symmetry* between S_o and T_o: the sensitivity of T_o with respect to plant variations is given by S_o, while the sensitivity of S_o is determined by T_o. This symmetry is a direct consequence of the analytical relationship between S_o and T_o. In fact, from equations (20) and (21) we see that the sensitivity functions are related by

$$S_o + T_o = 1 \tag{27}$$

This constraint underlies many of the trade-offs involved in feedback design, since clearly it is not possible to simultaneously make the magnitudes of S_o and T_o small. The solution is to make $|S_o|$ small in certain frequency ranges (usually at low frequencies), while allowing $|T_o|$ to be large (close to one) at those frequencies. Conversely, $|T_o|$ is made small at other frequencies (usually high frequencies), where $|S_o|$ is allowed to be large (close to one). These trade-offs, and further manifestations of the symmetry between S_o and T_o, will be studied in depth later.

3 Closed-loop Stability and Stability Robustness

We have seen that feedback has a range of beneficial effects, but an absolute requirement is that closed-loop stability must be maintained. Increasing the feedback gain improves disturbance rejection and makes the loop insensitive to changes in the plant, but for many systems continuing to increase the gain will eventually cause instability. For some systems, on the other hand, instability will result if the gain is too low (we shall see later that this is the case for plants which are open-loop unstable).

There are a number of formal definitions of stability: *asymptotic stability* means that for zero inputs and any initial conditions the system *state* will tend asymptotically to zero as time increases; *bounded-input bounded-output (BIBO) stability* means that a bounded input r results in bounded "outputs" u and y for any initial state (by design, we always assume that the pre-filter F is stable). Here, we focus on the concept of *internal stability*, which relates to interconnected systems such as the simple feedback loop shown in figure 8.

3.1 Internal stability

Internal stability means that the closed-loop systems resulting from all possible input injection points to all possible outputs are BIBO stable. To understand this, consider the feedback system shown in figure 9. Here, internal input signals v_i, $i = 1 \ldots 5$, have been injected at each "exposed interconnection" of the system, and additional internal output signals w_i, $i = 1 \ldots 5$, have been defined just after each injection point. The system is then said to be internally stable if and only if all possible input-output pairings are BIBO stable.

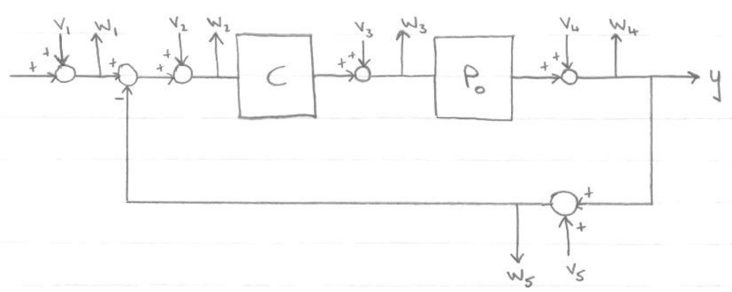

Figure 9: Internal stability: feedback system with added internal inputs and outputs.

Several important closed-loop transfer functions have a common denominator polynomial $A_oH + B_oG = A_{cl}$, which we call the closed-loop characteristic polynomial. In fact, A_{cl} is common to all possible closed-loop transfer functions considered in the definition of internal stability. Thus, closed-loop stability (internal stability) is achieved if and only if all the roots of the equation $A_{cl} = 0$ have strictly negative real parts. In other words, all closed-loop poles must lie in the left half of the complex plane.

It is important to note that internal stability implies more than just BIBO stability from the reference signal r to the output y. In fact, internal stability also excludes the possibility of cancellation of unstable poles between the compensator C and the nominal plant P_o (such cancellations are also prevented in more complex interconnections of subsystems). From the relation $A_{cl} = A_o H + B_o G$ it is clear that unstable pole-zero cancellations will result from unstable common factors in G and A_o, or in B_o and H, and A_{cl} will then be unstable (by assumption, we exclude *unstable hidden modes*, i.e. unstable cancellations *within* C or P_o). In such a scenario the system is not internally stable, despite the fact that the transfer function from r to y is stable.

Checking the roots of A_{cl} thus provides a direct test of internal stability.

The position of the closed-loop poles in the complex plane determines not only absolute nominal stability, but also gives a qualitative indication of the *relative stability* of the closed-loop, i.e. the likelihood that the loop will become unstable in the face of variations in the plant parameters. If the closed-loop poles are "close" to the imaginary axis then relatively small changes in the plant may cause the poles to migrate to the right half plane and the system will become unstable. However, these notions of relative stability and robustness are best studied using graphical methods in the frequency domain. We now consider such an approach - the Nyquist criterion.

3.2 Nyquist criterion

The Nyquist criterion is based on the frequency response of the nominal *loop gain* L_o, which is defined as the series open-loop connection of the compensator C and the plant P_o, i.e. $L_o = CP_o$. The Nyquist plot of L_o is the curve traced in the complex plane of $L_o(j\omega)$ for all frequencies ω (including negative frequencies). Because $L_o(-j\omega)$ is the complex conjugate of $L_o(j\omega)$ the entire plot of $L_o(j\omega)$ is obtained by reflecting the portion for $0 \leq \omega \leq +\infty$ about the real axis to get the portion for $-\infty \leq \omega \leq 0$.

The Nyquist criterion allows one to infer closed-loop stability from a plot of the open-loop function L_o. A proof of the Nyquist Criterion is not covered here (see [1] for details). However, the derivation is based on the *principle of the argument* and amounts to finding conditions on $L_o(j\omega)$ such that closed-loop poles in the right half plane are excluded. We summarise two versions of the criterion: the first is for open-loop stable systems, and the second is for open-loop unstable systems. Note that the results stated here are for functions L_o which are strictly proper, and which do not have any poles exactly on the imaginary axis (these conditions may be relaxed, but at the expense of a more complicated analysis, see [1, 2]).

1. **Nyquist Stability Criterion**: assume that the loop gain L_o has no right half plane poles (i.e. it is open loop stable). Then the closed-loop system is stable if and only if the Nyquist plot of L_o does not encircle the point -1.

2. **Generalised Nyquist Stability Criterion**: the generalised form of the Nyquist Criterion allows L_o to be open-loop unstable. We denote the number of open-loop right half plane poles as p, the number of closed-loop right half plane poles as z, and the number of clockwise encirclements of the point -1 by the Nyquist plot of L_o as n. The Generalised Nyquist Criterion then states that $z = n + p$. For closed-loop stability we require $z = 0$. Thus, the closed-loop system is stable if and only if the number of anti-clockwise encirclements of -1 is equal to p, the number of unstable open-loop poles.

It should be noted that to use the Nyquist Criteria to evaluate internal stability, we need additionally to exclude the possibility of cancellation of unstable poles between C and P_o, since the Criteria apply to the product $L_o = CP_o$, while internal stability requires also that no unstable pole-zero cancellations occur.

3.3 Robustness

The above discussion considered stability of the nominal closed-loop system. We previously introduced the distinction between the nominal plant P_o, which is often a simplified model of the "true" plant, and is used for compensator design and analysis. The compensator is then applied to the true plant P. We now examine how differences between P_o and P affect the stability and performance of the closed-loop. This is known as *robustness* analysis.

Consider the Nyquist plot of the nominal loop gain L_o of an open-loop stable system as shown in figure 10. The nominal closed loop is stable since $L_o(j\omega)$ does not encircle the point

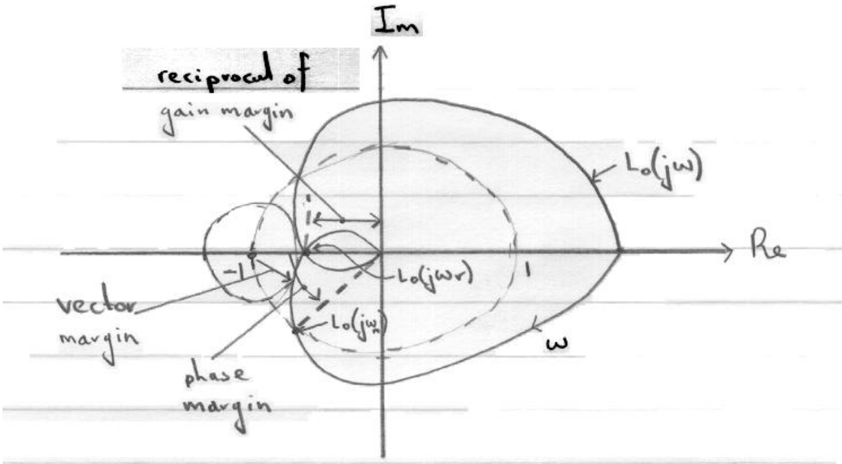

Figure 10: Stability margins.

-1. The actual closed-loop system will remain stable when P_o is perturbed to P (equivalently, when L_o is perturbed to L) as long as the Nyquist plot of the perturbed loop gain $L(j\omega)$ does not encircle -1.

3.3.1 Stability margins

Stability margins give a quantitative measure of how far from instability the loop is, i.e. they quantify the system's *relative stability.* There are three well-known measures which quantify the distance from the nominal loop gain to the critical stability point $(-1, 0)$.

1. **Gain Margin**: the gain margin k_m indicates the additional gain by which L_o may be multiplied so that it reaches the point -1. We denote the value of the loop gain at which the Nyquist plot intersects the negative real axis (i.e. the point where the phase is -180 deg) as $|L_o(j\omega_r)|$ (here, ω_r is the frequency at intersection). The gain margin is then
$$k_m = \frac{1}{|L_o(j\omega_r)|}$$

2. **Phase Margin**: the phase margin ϕ_m quantifies the additional phase delay which can be added to L_o to reach the critical condition. Thus, ϕ_m is the angle between the negative real axis and the point on L_o which has unity gain. This point is denoted $L_o(j\omega_m)$, and ω_m is the corresponding frequency (see figure 10).

3. **Vector Margin**: the gain and phase margins are the traditional indicators of relative stability. However, a much more informative measure is the vector margin, denoted s_m. As indicated in figure 10, this is the distance from the point -1 to the closest point on the plot of L_o. Equivalently, the vector margin is the radius of the smallest circle centred at -1 which is tangent to the Nyquist plot.

It is common for engineers to establish minimum acceptable values for the gain and phase margins. Common practical requirements are $k_m > 2$ and $\phi_m > 30\,\text{deg}$. However, the gain and phase margins do not always give an adequate characterisation of robustness. It is possible to have a plot of L_o which gives good gain and phase margins, but which nevertheless passes very close to the -1 point. In such a case, a relatively small *joint* perturbation of the gain and phase may suffice to destabilise the system. Thus the vector margin, which quantifies the nearest approach to -1, provides a more rigorous assessment of the actual stability margin.

It is straightforward to show that the vector margin s_m is related to the gain and phase margins by the inequalities

$$k_m \geq \frac{1}{1 - s_m}, \qquad \phi_m \geq 2 \arcsin \frac{s_m}{2}$$

Thus, if $s_m \geq 0.5$ then $k_m \geq 2$ and $\phi_m \geq 2 \arcsin 0.25 \approx 29\,\text{deg}$. (However, the converse is **not** true in general.) This shows that the traditional requirements for the gain and phase margins translate to the requirement for a vector margin of at least $1/2$, i.e. $s_m \geq 0.5$.

The vector margin s_m is directly related to the peak value of the nominal sensitivity function S_o. To see this, we denote the point on L_o which is closest to -1 as $L_o(j\omega_s)$, where ω_s is the corresponding frequency. It is easily seen that the vector from -1 to any point on $L_o(j\omega)$ is given by $1 + L_o(j\omega)$. We recognise that this is just the inverse of the sensitivity function, i.e. $1 + L_o(j\omega) = \frac{1}{S_o(j\omega)}$. Clearly, the minimal value of the magnitude of this function, which by definition is the vector margin, occurs at the maximal value of the nominal sensitivity, i.e.

$$s_m = |1 + L_o(j\omega_s)| = \frac{1}{|S_o(j\omega_s)|} = \frac{1}{\max\limits_{\omega} |S_o(j\omega)|}$$

This analysis shows that the nominal sensitivity peak is a key indicator of stability robustness, and implies that in feedback design peaking of the nominal sensitivity function S_o should be avoided.

By a similar argument, it can be shown that the peak value of the nominal complementary sensitivity function T_o is also important. If we plot the inverse Nyquist diagram, i.e. that of $1/L_o$, then loss of stability will occur as the plot comes close to -1. The minimal distance from $1/L_o$ to -1, denoted by r_m, and assumed to occur at frequency ω_r, is given by

$$r_m = \left| 1 + \frac{1}{L_o(j\omega_r)} \right| = \left| \frac{1 + L_o(j\omega_r)}{L_o(j\omega_r)} \right|$$

Since the complementary sensitivity function T_o is defined by $T_o = L_o/(1 + L_o)$ the above expression becomes

$$r_m = \frac{1}{|T_o(j\omega_r)|} = \frac{1}{\max\limits_{\omega} |T_o(j\omega)|}$$

This shows that peaking of T_o must also be avoided.

Finally, we note that adequate margins of the type discussed above are needed not only for stability robustness, but also to achieve a satisfactory time response of the closed-loop system. Small margins imply that the Nyquist plot comes close to -1 and that the stability boundary is approached. This is manifested by closed-loop poles which are close to the stability border (i.e. the imaginary axis). If the system is poorly damped then these closed-loop poles will cause an oscillatory response.

Thus, peaking of S_o and T_o must be avoided not only to achieve robust stability, but also to ensure acceptable closed-loop time-domain responses.

4 Design Goals

There are a number of goals which one attempts to achieve in the design of closed-loop feedback systems. The most important design goals include:

- closed-loop stability
- stability robustness (maintain stability in the face of plant variations)
- disturbance attenuation
- satisfactory closed-loop command response
- robustness of closed-loop performance (maintain performance in the face of plant variations)

There are a number of factors which limit the extent to which these goals can be achieved. The key limitations are:

- plant capacity
- measurement noise

These design goals, and the associated limitations, can be translated into requirements on the frequency responses of various closed-loop transfer functions, as discussed in the following sections. The discussion considers the two-degree-of-freedom control architecture of figure 8.

From the closed-loop equation (16) we recall that the output y is related to the command signal r, the disturbance d, and the measurement noise n by

$$y = T_o F r + S_o d - T_o n, \tag{28}$$

where S_o and T_o are the sensitivity and complementary sensitivity functions, respectively, and F is a reference pre-filter. This relationship will be used in the following analysis. We shall see that most of the design requirements and limitations can be translated into requirements on the shapes of S_o and T_o.

4.1 Closed-loop stability

By the Nyquist Criterion (see section 3.2), to achieve nominal closed-loop stability the loop gain L_o must be shaped such that the Nyquist plot of L_o does not encircle the point -1 (for a system which is open-loop stable).

4.2 Stability robustness

Robustness against loop gain perturbations requires the complementary sensitivity function T_o to be small. On the other hand, it was also shown that robustness against inverse loop gain perturbations requires the sensitivity function S_o to be small. Because of the constraint $S_o + T_o = 1$ (see equation (27)) it is not possible to make S_o and T_o small at the same time.

The solution is to make S_o and T_o small in different frequency ranges. Because of the low-pass nature of physical plants, S_o will naturally be small in the low frequency range, and will increase to 1 at high frequencies. Conversely, T_o will be close to 1 at low frequencies and will *roll off* to low values at high frequencies. Thus, T_o provides a "natural" robustness against high-frequency perturbations of the loop gain. This is important because fast, high-frequency dynamics are often neglected, and this is reflected in high-frequency loop gain perturbations. On the other hand, S_o provides a natural robustness against low-frequency variations in the inverse loop gain. These perturbations are often the result of slow variations in operating conditions and loads.

The most critical frequency range for robustness is the *crossover region*. This is the frequency region where the loop gain L_o crosses the value 1, and where neither S_o nor T_o can be small. In this region it is crucial to ensure that peaking in S_o and T_o is avoided so that the robustness inequalities are not violated. It was seen in section 3.3.1 that this is required in order to keep the Nyquist plot of L_o away from the -1 point.

4.3 Disturbance attenuation

The effect of disturbances d on the plant output y is determined by the sensitivity function $S_o = 1/(1+L_o)$ (see equation (28)), i.e. S_o is the nominal transfer function from d to y. Thus, the smaller is $|S_o(j\omega)|$, the greater is the attenuation of disturbances at frequency ω. $|S_o|$ is small if the magnitude of the loop gain L_o is large. Thus, it is necessary to shape the loop gain to make it large in frequency ranges where disturbance attenuation is needed.

Because of the natural shape of S_o (small at low frequencies, increasing to 1 at high frequency) it is normally only possible to make L_o large over a limited frequency band. The frequency up to which effective disturbance attenuation can be achieved is known as the *bandwidth* of the feedback loop.

Finally, peaking of S_o (to values significantly greater than 1, typically in the crossover region) must be avoided otherwise disturbance amplification will occur.

4.4 Command response

The closed-loop command response is governed by the complementary sensitivity function T_o, and by the reference pre-filter F. This can be seen from equation (28), since when $d = n = 0$ we have

$$y = T_o F r$$

Thus, without a pre-filter ($F = 1$) the closed-loop transfer function for command tracking is identical to T_o. Since T_o is typically close to 1 for low frequencies, and decreases beyond the crossover region, the output y will follow command signals r up to a frequency which is close to the bandwidth for disturbance rejection. If the closed-loop bandwidth is significantly different to the bandwidth required for command following then the pre-filter F can be used to directly compensate for this.

4.5 Plant capacity

It is important to ensure that inputs to the plant generated by feedback, i.e. the control signal u, are not too large for the plant to handle. Any physical plant will have a limited capacity to absorb high inputs, and at some point will enter a saturation state. From equation (19) the control signal u is governed by the *control sensitivity function* $S_{uo} = C/(1 + CP_o)$, i.e.

$$u = S_{uo}(Fr - d - n) \tag{29}$$

Thus, the control sensitivity function determines the way in which the plant input is affected by the command signal, disturbances, and measurement noise.

Analysis of the input sensitivity function illustrates the fact that open-loop plant zeros in the right half plane limit the achievable closed-loop bandwidth. We have seen previously that at certain frequencies (especially low frequencies) it is desirable to make the magnitude of the loop gain $L_o = CP_o$ large. If the loop gain is large then the input sensitivity function is approximately equal to the inverse of the nominal plant transfer function, i.e. $S_{uo} \approx 1/P_o$. If the open-loop plant has zeros in the right half plane then the transfer function $1/P_o$ is unstable, and for this reason open-loop plant zeros in the right half plane limit the closed-loop bandwidth, i.e. the frequency range over which L_o can be made large. In fact, the input sensitivity function S_{uo} may only be made equal to $1/P_o$ up to the frequency which equals the magnitude of the right half plane zero with the smallest magnitude.

As noted in section 2 the relationship between the complementary sensitivity function T_o and the input sensitivity function S_{uo} is

$$T_o = S_{uo}P_o$$

Thus, design requirements on the input sensitivity function may be translated into requirements on the complementary sensitivity function T_o. Equation (29) indicates that large values of S_{uo} may lead to plant input signals u which are too large for the plant to handle. This may conflict with the requirement at low frequencies to have a large loop gain L_o, and a correspondingly large S_{uo}, for robustness and disturbance rejection properties. If this conflict arises then the plant capacity may need to be increased (e.g. by using more powerful actuators) or the performance requirements may need to be relaxed. At frequencies above the bandwidth S_{uo} should decrease as fast as possible, which is in line with the requirement that T_o should roll off quickly above the bandwidth (for robustness and low noise sensitivity).

4.6 Measurement noise

Equation (28) shows that the closed-loop transfer function from measurement noise to output is identical with T_o, the complementary sensitivity function. Above the closed-loop bandwidth the magnitude of T_o will be less than 1, and noise will therefore be attenuated and will have only a small influence on the output. For low frequencies, however, $|T_o|$ is close to 1 and measurement noise fully affects the output. In the crossover region, $|T_o|$ may become greater than 1, in which case measurement noise will be amplified. These facts emphasise the need for low-noise sensors, and the need to avoid peaking of $|T_o|$ in the crossover region.

4.7 Performance robustness

We have seen that the various design goals for satisfactory closed-loop performance translate into requirements on the shapes of S_o and T_o. We now wish to investigate performance robustness, i.e. the extent to which changes in the plant affect these transfer functions. The answer to this question has in fact already been given in section 2. There, we derived the symmetry result that the sensitivity of T_o with respect to changes in the plant is given by S_o (see equation (24)), while the sensitivity of S_o with respect to plant variations is given by T_o (see equation (26)). Thus, the trade-off between S_o and T_o, due to the constraint $S_o + T_o = 1$, arises once more, and the solution is to address the different frequency ranges.

These results show that, because $|S_o|$ is naturally small at low frequencies, T_o will be robust at low frequencies, which is precisely the frequency range where T_o takes on significant values. At high frequencies the magnitude of T_o is small, which means that the theoretical loss of robustness here is not usually significant. Conversely, because $|T_o|$ is naturally small at high frequencies, S_o will be robust at high frequencies, which is precisely the frequency range where S_o has significant values. In the low frequency region the magnitude of S_o is small, and any changes in S_o resulting from plant variations are usually insignificant.

It is once more apparent that at intermediate frequencies (the crossover region) peaking of S_o and T_o needs to be avoided so that changes in the plant do not cause serious loss of performance.

4.8 Summary

The performance and robustness properties of the feedback loop are determined largely by the shapes of S_o and T_o. The requirements on S_o and T_o lead to general guidelines, which can be summarised as follows:

1. The sensitivity function S_o should be small at low frequencies. This will help to achieve:
    - good disturbance attenuation;
    - satisfactory command response;
    - robustness of closed-loop stability against perturbations in the inverse loop gain.
2. The complementary sensitivity function T_o should be small at high frequencies, in order to:
    - avoid exceeding plant capacity;
    - avoid unwanted effects of measurement noise;
    - achieve robustness of closed-loop stability against perturbations in the loop gain.
3. At intermediate frequencies (i.e. in the *crossover region*) peaking of both S_o and T_o should be avoided to prevent:
    - disturbance amplification;
    - amplification of measurement noise;
    - changes in the plant causing loss of stability or serious degradation of performance (i.e. loss of stability and performance robustness).

References

[1] G. F. Franklin, J. D. Powell, and A. Emami-Naeini, *Feedback Control of Dynamic Systems*, 6th ed. Prentice Hall, 2010.

[2] G. C. Goodwin, S. F. Graebe, and M. E. Salgado, *Control System Design*. Prentice Hall, 2001.