Chapter 9 Complex Control Strategies

Complex Control Strategies

9.1 The 2DOF Structure for Stable Plants

1DOF System

The effect of the reference on the error was the same as that of the output disturbance on the system output:

$$\frac{e(s)}{r(s)} = \frac{y(s)}{d(s)} = \frac{1}{1 + G(s)C(s)}$$

Such a system has merely one degree of freedom. When the reference and the disturbance have similar dynamic characteristics (for example, both of them are steps),a 1DOF controller can simultaneously satisfy the requirement on tracking response and disturbance response in many cases

Why do we Need a 2DOF System

Sometimes, the dynamic characteristics of the reference and the disturbance are different. For example, the reference is a step while the disturbance at the plant output is a ramp

If both good tracking response and good disturbance response are desired, the controller that achieves the two goals may not exist. In this case, an additional controller may have to be introduced so that the tracking response and the disturbance response can be adjusted independently. There are two loops in this system:

- One is the reference loop, which is from the reference to the system output
- The other is the disturbance loop, which is from the disturbance at the plant output to the system output

Such a system is of 2DOF

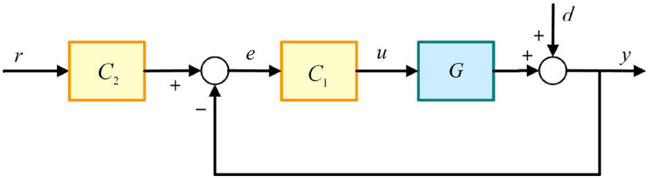


Figure: Typical 2DOF system

$C_1(s)$—The controller of the disturbance loop
$C_2(s)$—The controller of the reference loop. $C_2(s)$ is always stable

For convenience of presentation, the structure is named "**Structure I**". Structure I has many equivalents, as shown in Figures. It should be pointed out that the $C_1(s)$s in these Figures are not identical to each other

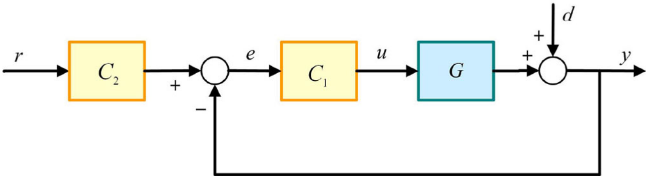


Figure: Typical 2DOF system

$C_1(s)$—The controller of the disturbance loop
$C_2(s)$—The controller of the reference loop. $C_2(s)$ is always stable

For convenience of presentation, the structure is named "**Structure I**". Structure I has many equivalents, as shown in Figures. It should be pointed out that the $C_1(s)$s in these Figures are not identical to each other

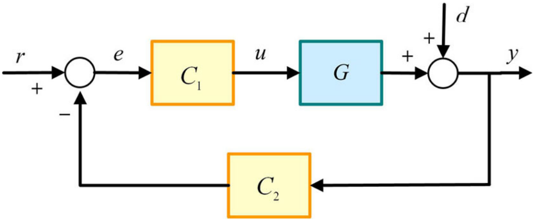


Figure: An equivalent of the typical 2DOF system

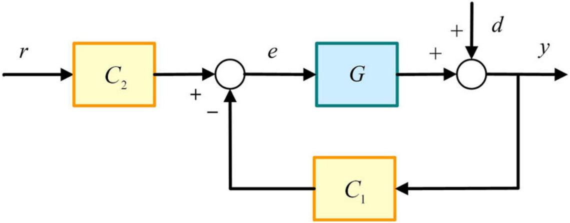


Figure: Another equivalent of the typical 2DOF system

Internal stability: Consider the typical 2DOF system. The input-output relationship is as follows:

$$\frac{e(s)}{r(s)} = \frac{C_2(s)}{1 + G(s)C_1(s)}$$

$$\frac{y(s)}{d(s)} = \frac{1}{1 + G(s)C_1(s)}$$

It can be seen that the internal stability of the closed-loop system is only determined by $C_1(s)$. The analysis for the internal stability is similar to that in a 1DOF system

Design: The design of a 2DOF system involves two steps:

1. Design $C_1(s)$ for good disturbance response
2. Design $C_2(s)$ for good tracking response

Internal stability: Consider the typical 2DOF system. The input-output relationship is as follows:

$$\frac{e(s)}{r(s)} = \frac{C_2(s)}{1 + G(s)C_1(s)}$$
$$\frac{y(s)}{d(s)} = \frac{1}{1 + G(s)C_1(s)}$$

It can be seen that the internal stability of the closed-loop system is only determined by $C_1(s)$. The analysis for the internal stability is similar to that in a 1DOF system

Design: The design of a 2DOF system involves two steps:

1. Design $C_1(s)$ for good disturbance response
2. Design $C_2(s)$ for good tracking response

$C_1(s)$: The design of $C_1(s)$ is the same as that for the controller in a 1DOF system

$C_2(s)$: After $C_1(s)$ is designed, the loop consisting of $C_1(s)$ and $G(s)$ is viewed as an augmented plant, of which the transfer function is denoted by $T(s)$. The system that consists of $C_2(s)$ and $T(s)$ forms the IMC structure with an exact model. Accordingly, $C_2(s)$ can be directly designed

To illustrate the design procedure, consider the following plant:

$$G(s) = \frac{K}{\tau s + 1} e^{-\theta s}$$

The system is required to track a step reference, and at the same time to reject the disturbance that is in the form of a ramp at the plant output

$C_1(s)$: The design of $C_1(s)$ is the same as that for the controller in a 1DOF system
$C_2(s)$: After $C_1(s)$ is designed, the loop consisting of $C_1(s)$ and $G(s)$ is viewed as an augmented plant, of which the transfer function is denoted by $T(s)$. The system that consists of $C_2(s)$ and $T(s)$ forms the IMC structure with an exact model. Accordingly, $C_2(s)$ can be directly designed

To illustrate the design procedure, consider the following plant:

$$G(s) = \frac{K}{\tau s + 1} e^{-\theta s}$$

The system is required to track a step reference, and at the same time to reject the disturbance that is in the form of a ramp at the plant output

First of all, design $C_1(s)$ for disturbance rejection. By utilizing the IMC controller $Q(s)$, $C_1(s)$ can be expressed as

$$C_1(s) = \frac{Q(s)}{1 - G(s)Q(s)}$$

The rational part of the plant is MP. Utilizing (**??**), we have

$$Q_{opt}(s) = \frac{\tau s + 1}{K}$$

Since the disturbance is a ramp, a Type 2 filter should be introduced. In light of the discussion in Section 5.7,

$$J(s) = \frac{2\lambda_1 s + 1}{(\lambda_1 s + 1)^2}$$

where λ_1 is the performance degree for disturbance rejection. Simple computations give

$$Q(s) = Q_{opt}(s)J(s) = \frac{(\tau s + 1)(2\lambda_1 s + 1)}{K(\lambda_1 s + 1)^2}$$

Next, design $C_2(s)$ for tracking. The loop that consists of $C_1(s)$ and $G(s)$ is regarded as an augmented plant, whose transfer function is

$$T(s) = G(s)Q(s) = \frac{2\lambda_1 s + 1}{(\lambda_1 s + 1)^2} e^{-\theta s}$$

Obviously, $T(s)$ is stable. According to the design procedure for the H_2 controller, the optimal $C_2(s)$ is the inverse of the rational part of $T(s)$. For a step reference the suboptimal controller is

$$C_2(s) = \frac{(\lambda_1 s + 1)^2}{(2\lambda_1 s + 1)(\lambda_2 s + 1)}$$

where λ_2 is the performance degree for disturbance rejection
Let $T_r(s)$ denote the transfer function from the reference to the system output, and $T_d(s)$ denote the transfer function from the disturbance at the plant input to the system output

It is easy to verify that the response of the reference loop is

$$T_r(s) = \frac{1}{\lambda_2 s + 1} e^{-\theta s}$$

of which the time domain response is

$$T_r(t) = \begin{cases} 0 & t < \theta \\ 1 - e^{-(t-\theta)/\lambda_2} & t \geq \theta \end{cases}$$

The reference response can be independently adjusted by the performance degree λ_2

The response of the disturbance loop can be written as

$$T_d(s) = \frac{K}{\tau s + 1} e^{-\theta s} \left[1 - \frac{2\lambda_1 s + 1}{(\lambda_1 s + 1)^2} e^{-\theta s}\right]$$

The corresponding time domain response is

$$T_d(t) = \begin{cases} 0 & t < \theta \\ K(1 - e^{-(t-\theta)/\tau}) & \theta \leq t < 2\theta \\ K\left[\begin{array}{l} \frac{\lambda_1}{\lambda_1-\tau} e^{-(t-2\theta)/\lambda_1} - \\ \frac{\tau}{\lambda_1-\tau} e^{-(t-2\theta)/\tau} - e^{-(t-\theta)/\tau} \end{array}\right] & t \geq 2\theta \end{cases}$$

The disturbance response can be independently adjusted by employing the performance degree λ_1

The disturbance loop of a 2DOF system **cannot** provide better disturbance rejection ability than a 1DOF system. Nevertheless, since the disturbance response and the reference response can be adjusted independently, better disturbance response and reference response can be reached **simultaneously** in a 2DOF system

Robustness: If there exists uncertainty, the reference response cannot be thoroughly isolated from the disturbance response. In this case, the robust stability and the disturbance response is only determined by λ_1, while the reference response is mainly determined by λ_2

For robustness tuning, one can monotonically increase the performance degrees until the required response is obtained.

Implementation:

If the plant is stable, $C_1(s)$ can be implemented in the IMC structure. When $C_1(s)$ is implemented as the unity feedback controller by using rational approximations, the augmented plant is

$$T(s) = \frac{G(s)C_1(s)}{1 + G(s)C_1(s)}$$

Robustness: If there exists uncertainty, the reference response cannot be thoroughly isolated from the disturbance response. In this case, the robust stability and the disturbance response is only determined by λ_1, while the reference response is mainly determined by λ_2

For robustness tuning, one can monotonically increase the performance degrees until the required response is obtained.

Implementation:

If the plant is stable, $C_1(s)$ can be implemented in the IMC structure. When $C_1(s)$ is implemented as the unity feedback controller by using rational approximations, the augmented plant is

$$T(s) = \frac{G(s)C_1(s)}{1 + G(s)C_1(s)}$$

Remove the time delay in the numerator of $T(s)$. The optimal $C_2(s)$ should be the inverse of the remainder of $T(s)$. Since the denominator of $T(s)$ contains a time delay, $C_2(s)$ contains a time delay. To implement $C_2(s)$, rational approximations have to be used. Design methods utilizing rational approximations have been studied well in foregoing chapters and thus are not repeated here

In Figure, a new 2DOF structure is given. To distinguish it from Structure I, it is named "**Structure II**". In Structure II, $C_3(s)$ is the controller for the disturbance loop, and $C_4(s)$ is the controller for the reference loop. If let

$$\begin{aligned} C_1(s) &= C_3(s) \\ C_2(s) &= \frac{C_4(s) + G(s)C_3(s)}{C_3(s)} \end{aligned}$$

then Structure I and Structure II are equivalent to each other

Remove the time delay in the numerator of $T(s)$. The optimal $C_2(s)$ should be the inverse of the remainder of $T(s)$. Since the denominator of $T(s)$ contains a time delay, $C_2(s)$ contains a time delay. To implement $C_2(s)$, rational approximations have to be used. Design methods utilizing rational approximations have been studied well in foregoing chapters and thus are not repeated here

In Figure, a new 2DOF structure is given. To distinguish it from Structure I, it is named "**Structure II**". In Structure II, $C_3(s)$ is the controller for the disturbance loop, and $C_4(s)$ is the controller for the reference loop. If let

$$\begin{aligned} C_1(s) &= C_3(s) \\ C_2(s) &= \frac{C_4(s) + G(s)C_3(s)}{C_3(s)} \end{aligned}$$

then Structure I and Structure II are equivalent to each other

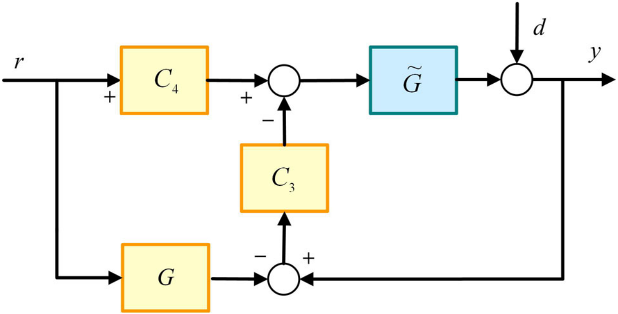


Figure: New 2DOF system

The feature of Structure II is that both of the two controllers can be directly designed. The design of $C_3(s)$ is similar to that for the unity feedback loop controller. Since the reference loop is an open one for the nominal plant, $C_4(s)$ can be designed as the inverse of the rational part of the plant

Explanation For RZN PID

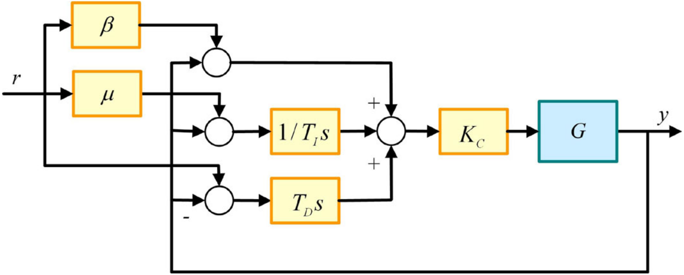


Figure: RZN PID control structure

In section 4.1, the RZN PID controller depicted in Figure **??** was introduced. By utilizing two parameters β and μ, an improved reference response is obtained

It is easy to verify that

$$T(s) = \frac{G(s)K_C\left(\beta + \frac{1}{\mu T_I s} + T_D s\right)}{1 + G(s)K_C\left(1 + \frac{1}{T_I s} + T_D s\right)}$$

$$= \frac{K_C\left(\beta + \frac{1}{\mu T_I s} + T_D s\right)}{K_C\left(1 + \frac{1}{T_I s} + T_D s\right)} \frac{G(s)K_C\left(1 + \frac{1}{T_I s} + T_D s\right)}{1 + G(s)K_C\left(1 + \frac{1}{T_I s} + T_D s\right)}$$

Let

$$C_1(s) = K_C\left(1 + \frac{1}{T_I s} + T_D s\right)$$

$$C_2(s) = \frac{K_C\left(\beta + \frac{1}{\mu T_I s} + T_D s\right)}{K_C\left(1 + \frac{1}{T_I s} + T_D s\right)}$$

Evidently, the introduction of β and μ is equivalent to providing an additional degree of freedom to the original system

The resulting system is in fact a 2DOF system

9.2 2DOF Structure for Unstable Plants

Problem: The reference response of the 1DOF system with an unstable plant usually exhibits excessive overshoot

Solution: The 2DOF system can be used to overcome this problem. This is because the 2DOF system can roll off high frequency signals in the reference independently

Design procedure: Similar to that for stable plants. To guarantee the internal stability, the controller for the disturbance loop can **only** be implemented in the unity feedback loop

The procedure will be illustrated by utilizing the first-order unstable plant. The design procedure for high-order plants is similar

9.2 2DOF Structure for Unstable Plants

Problem: The reference response of the 1DOF system with an unstable plant usually exhibits excessive overshoot

Solution: The 2DOF system can be used to overcome this problem. This is because the 2DOF system can roll off high frequency signals in the reference independently

Design procedure: Similar to that for stable plants. To guarantee the internal stability, the controller for the disturbance loop can **only** be implemented in the unity feedback loop

The procedure will be illustrated by utilizing the first-order unstable plant. The design procedure for high-order plants is similar

The first-order unstable plant can be expressed as

$$G(s) = \frac{K}{\tau s - 1} e^{-\theta s}$$

As the main goal of this section is to discuss how to depress the excessive overshoot, it is assumed that the system is required to track a step reference, and at the same time to reject the effect of a step disturbance at the plant output.

Consider Structure I. First, design $C_1(s)$. The rational part of the plant is MP. Based on the discussion in Section 8.4, we have

$$Q_{opt}(s) = \frac{\tau s - 1}{K}$$

When the disturbance at the plant output is a step and the plant has only one RHP pole, the filter can easily be determined

The H_2 suboptimal controller with the filter is

$$Q(s) = \frac{(\tau s - 1)\{\tau[(\lambda_1/\tau + 1)^2 e^{\theta/\tau} - 1]s + 1\}}{K(\lambda_1 s + 1)^2}$$

Then the controller for the disturbance loop is

$$\begin{aligned} C_1(s) &= \frac{Q(s)}{1 - G(s)Q(s)} \\ &= \frac{1}{K}\frac{(\tau s - 1)\{\tau[(\lambda_1/\tau + 1)^2 e^{\theta/\tau} - 1]s + 1\}}{(\lambda_1 s + 1)^2 - \{\tau[(\lambda_1/\tau + 1)^2 e^{\theta/\tau} - 1]s + 1\}e^{-\theta s}} \end{aligned}$$

Since

$$\lim_{s \to 1/\tau} \{(\lambda_1 s + 1)^2 - \{\tau[(\lambda_1/\tau + 1)^2 e^{\theta/\tau} - 1]s + 1\}e^{-\theta s}\} = 0$$

There exists a RHP zero-pole cancellation in $C_1(s)$. A rational approximation has to be used to remove it

This can be achieved in many ways. For example, the controller can be chosen as a PID controller in the form of

$$C_1(s) = K_C \left(1 + \frac{1}{T_I s} + T_D s\right)$$

With the Maclaurin series expansion, the following result was obtained in Section 8.5:

$$\begin{aligned}
T_I &= -\tau + \beta_1 - \frac{\lambda^2 + \beta_1\theta - \theta^2/2}{2\lambda + \theta - \beta_1} \\
K_C &= \frac{T_I}{-K(2\lambda + \theta - \beta_1)} \\
T_D &= \frac{-\tau\beta_1 - (\theta^3/6 - \beta_1\theta^2/2)/(2\lambda + \theta - \beta_1)}{T_I} - \\
&\quad \frac{\lambda^2 + \beta_1\theta - \theta^2/2}{2\lambda + \theta - \beta_1}
\end{aligned}$$

Second, design $C_2(s)$. Regard the feedback loop consisting of $C_1(s)$ and $G(s)$ as an augmented plant. The transfer function of the augmented plant is

$$T(s) = \frac{G(s)C_1(s)}{1 + G(s)C_1(s)}$$

The optimal $C_2(s)$ should be the inverse of $T(s)$ after the time delay in its numerator is removed. However, such a design procedure is tedious. Since $C_1(s)$ is an approximation of the ideal controller, $C_2(s)$ can be chosen as the inverse of the ideal $T(s)$ after the time delay in its numerator is removed. Then

$$C_2(s) = \frac{(\lambda_1 s + 1)^2}{\{\tau[(\lambda_1/\tau + 1)^2 e^{\theta/\tau} - 1]s + 1\}(\lambda_2 s + 1)}$$

Now consider Structure II. In Structure II, $C_3(s)$ equals $C_1(s)$ in Structure I. The optimal $C_4(s)$ can be chosen as the inverse of the rational part of the plant:

$$C_4(s) = \frac{\tau s - 1}{K(\lambda_2 s + 1)}$$

$C_3(s)$ is an approximation of the ideal controller. Hence, the reference response of the system with $C_3(s)$ approximates to the reference response of the system with the ideal controller:

$$T_r(s) = \frac{1}{\lambda_2 s + 1} e^{-\theta s}$$

The corresponding time domain response is

$$y(t) = \begin{cases} 0 & t < \theta \\ 1 - e^{-(t-\theta)/\lambda_2} & t \geq \theta \end{cases}$$

The reference response can be independently tuned by the performance degree λ_2. The disturbance response is approximately

$$G(s)T_d(s) = \frac{K}{\tau s - 1} e^{-\theta s} \left\{ 1 - \frac{\tau[(\lambda_1/\tau + 1)^2 e^{\theta/\tau} - 1]s + 1}{(\lambda_1 s + 1)^2} e^{-\theta s} \right\}$$

Let

$$a = \frac{1}{\lambda_1}, b = -\frac{1}{\tau}, c = \frac{1}{\tau[(\lambda_1/\tau + 1)^2 e^{\theta/\tau} - 1]}$$

The time-domain response is

$$y(t) = \begin{cases} 0 & t < \theta \\ K(1 - e^{-b(t-\theta)}) & \theta \le t < 2\theta \\ K\Bigg(- e^{-b(t-\theta)} - \frac{a^2(b-c)}{(a-b)^2 c} e^{-b(t-2\theta)} - & \\ \frac{ab(c-a)+bc(a-b)}{(a-b)^2 c} e^{-a(t-2\theta)} - & \\ \frac{ab(c-a)}{(a-b)c} t e^{-a(t-2\theta)} \Bigg) & t \ge 2\theta \end{cases}$$

The disturbance response can be independently tuned by λ_1.

Two classes of control methods are frequently used for unstable plants in literature:

- In the first method, a controller is directly used to control the plant, for example, the 1DOF control system and the 2DOF control system
- In the second method, an inner loop is introduced to stabilize the unstable plant, and then the controller is designed for the augmented stabilized plant

The control system with an inner stabilizing loop is shown in Figure, where $C_s(s)$ is the stabilizer. $C_s(s)$ and $G(s)$ construct an augmented plant. $C_s(s)$ should be chosen so that the augmented plant is stable. $C(s)$ is then designed for the augmented plant

The disturbance response can be independently tuned by λ_1.

Two classes of control methods are frequently used for unstable plants in literature:

- In the first method, a controller is directly used to control the plant, for example, the 1DOF control system and the 2DOF control system
- In the second method, an inner loop is introduced to stabilize the unstable plant, and then the controller is designed for the augmented stabilized plant

The control system with an inner stabilizing loop is shown in Figure, where $C_s(s)$ is the stabilizer. $C_s(s)$ and $G(s)$ construct an augmented plant. $C_s(s)$ should be chosen so that the augmented plant is stable. $C(s)$ is then designed for the augmented plant

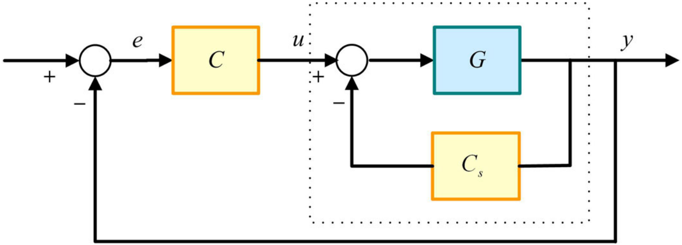


Figure: Control system with an inner stabilizing loop

Problem: The introduction of $C_s(s)$ makes the structure different from the unity feedback loop. As a result, the performance and robustness of the closed-loop system are difficult to analyze

Solution: It will be shown that the structure is in fact equivalent to the 2DOF structure

The closed-loop transfer function of the system is

$$T_r(s) = \frac{G(s)C(s)}{1 + G(s)C_s(s) + G(s)C(s)}$$

Rewrite it in the form of

$$T_r(s) = \frac{C(s)}{C_s(s) + C(s)} \frac{G(s)C(s) + G(s)C_s(s)}{1 + G(s)C_s(s) + G(s)C(s)}$$

which is equivalent to a 2DOF system with the following controllers:

$$\begin{aligned} C_1(s) &= C_s(s) + C(s) \\ C_2(s) &= \frac{C(s)}{C_s(s) + C(s)} \end{aligned}$$

The equivalent is illustrated in Figure

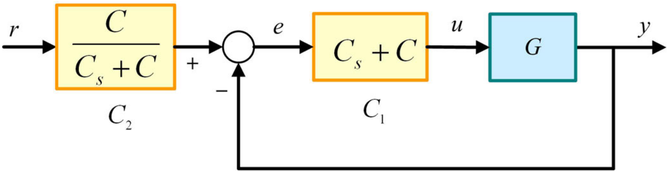


Figure: Equivalent of the control system with an inner stabilizing loop

For the reason of simplicity, the stabilizer $C_s(s)$ is usually chosen to be a proportional controller. Assume that the $C(s)$ in Figure **??** is a PID controller:

$$C(s) = K_c \left(1 + \frac{1}{T_I s} + T_D s \right)$$

Then the controller in Figure for the disturbance loop is also a PID controller

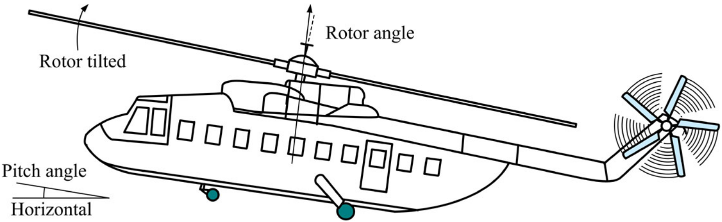
$$C_1(s) = C(s) + C_s(s) = K_c \frac{1 + (C_s/K_c + 1)T_I s + T_I T_D s^2}{T_I s}$$

Therefore,

1. The optimal method can be used to analytically design the controller in the system with an inner stabilizing loop
2. The closed-loop response can be quantitatively tuned
3. The robustness of the system can be analyzed by those methods developed for the unity feedback control system

Example

A plane with fixed wings has the feature of being self-regulating; i.e., it is stable. However, a helicopter is generally unstable and thus is involved to control. In a helicopter control system, the goal is to control the pitch angle by adjusting the rotor angle



Example (ctd.1)

The transfer function of a helicopter is

$$\boxed{G(s) = \frac{25(s+0.03)}{(s+0.4)(s^2-0.36s+0.16)}}$$

The plant is MP

Example (ctd.2)

Following the discussion in Section 8.4, we have

$$Q_{opt}(s) = \frac{(s+0.4)(s^2-0.36s+0.16)}{25(s+0.03)}$$

The plant has two RHP poles: $0.18 \pm 0.3572i$. Choose the following filter:

$$J(s) = \frac{\beta_2 s^2 + \beta_1 s + 1}{(\lambda_1 s + 1)^4}$$

If one takes $\lambda_1 = 0.5$, then $\beta_2 = 1.6781$ and $\beta_1 = 1.9164$. The controller of the disturbance loop in Structure I is

$$\boxed{C_1(s) = \frac{(s+0.4)(1.6781s^2+1.9164s+1)}{25s(s+0.03)(0.0625s+0.5225)}}$$

Example (ctd.3)

Since there is not any time delay in the plant, the augmented plant consisting of $C_1(s)$ and $T(s)$ is rational. Then, the controller of the reference loop is also a rational transfer function:

$$\boxed{C_2(s) = \frac{(\lambda_1 s + 1)^4}{(1.6781s^2 + 1.9164s + 1)(\lambda_2 s + 1)^2}}$$

Take $\lambda_2 = 0.5$. A unit step reference is added at $t = 0$ and a step load with the amplitude -0.1 is added at $t = 20$. The closed-loop system responses are shown in Figure. It can be seen that the disturbance responses are the same for the 1DOF control system and the 2DOF control system, while the reference response of the 2DOF control system is improved

Example (ctd.4)

The augmented plant consisting of $C_1(s)$ and $T(s)$ is rational. Then, the controller of the reference loop is rational:

$$\boxed{C_2(s) = \frac{(\lambda_1 s + 1)^4}{(1.6781s^2 + 1.9164s + 1)(\lambda_2 s + 1)^2}}$$

Take $\lambda_2 = 0.5$. A unit step reference is added at $t = 0$ and a step load with the amplitude -0.1 is added at $t = 20$. The closed-loop system responses are shown in Figure. It can be seen that the disturbance responses are the same for the 1DOF control system and the 2DOF control system, while the reference response of the 2DOF control system is improved

The inner stabilizer that corresponds to the result is

$$C_s(s) = \frac{(s + 0.4)[(1.6781s^2 + 1.9164s + 1) - (0.5s + 1)^2]}{25s(s + 0.03)(0.0625s + 0.5225)}$$

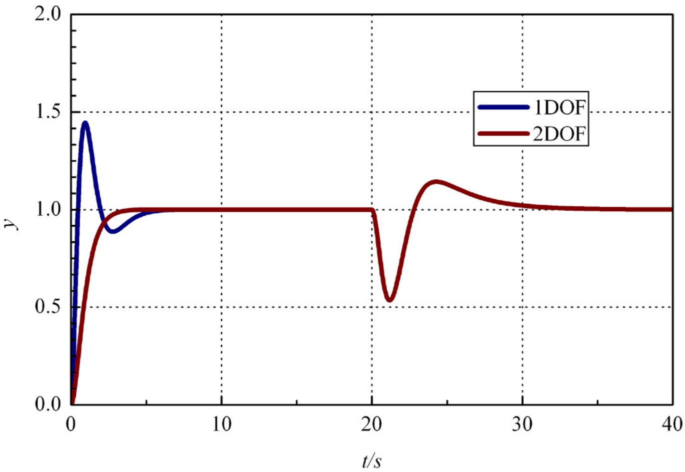


Figure: Responses of the 1DOF system and the 2DOF system

9.3 Cascade Control

Feature of the unity feedback configuration: It has only **one** loop for **one** system output

Feature of the cascade configuration: Uses **more than one** loop for **one** system output

To see how the cascade control system works, consider the distillation column shown in Figure. Assume that the single loop structure is used; that is, only the master controller is used. The temperature at the bottom of the distillation column is controlled by adjusting the steam flow rate to the reboiler. When the pressure for steam supply increases, the steam flow rate will increase. No correction will be made by the controller until the higher steam flow rate increases the vapor boilup and eventually raise the column temperature. Consequently, the whole system is disturbed by the change of the supply steam pressure

9.3 Cascade Control

Feature of the unity feedback configuration: It has only **one** loop for **one** system output

Feature of the cascade configuration: Uses **more than one** loop for **one** system output

To see how the cascade control system works, consider the distillation column shown in Figure. Assume that the single loop structure is used; that is, only the master controller is used. The temperature at the bottom of the distillation column is controlled by adjusting the steam flow rate to the reboiler. When the pressure for steam supply increases, the steam flow rate will increase. No correction will be made by the controller until the higher steam flow rate increases the vapor boilup and eventually raise the column temperature. Consequently, the whole system is disturbed by the change of the supply steam pressure

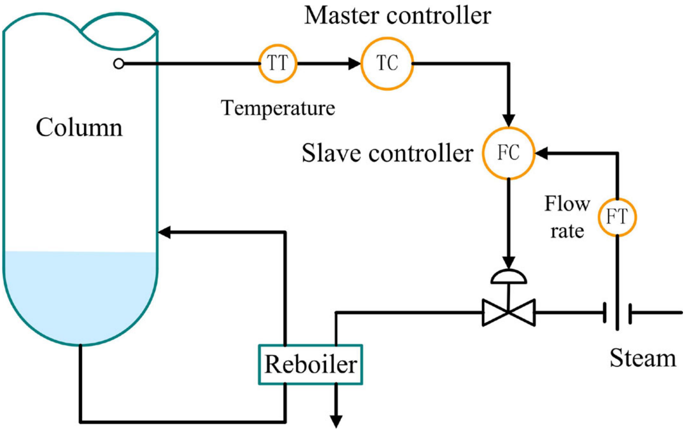


Figure: Temperature control system for distillation column

The response can be improved by installing a slave controller in between the temperature controller and the controlled steam flow rate. The controller is used to control the flow rate. Such an arrangement constitutes a **cascade control configuration**. In the cascade control system, the slave controller will immediately find the increase in the steam flow rate, and pinch the steam valve to pull the steam flow rate back to the desired value. As a result, the reboiler and the column are only slightly affected by the disturbance from the steam supply pressure.

The features of cascade control are as follows:

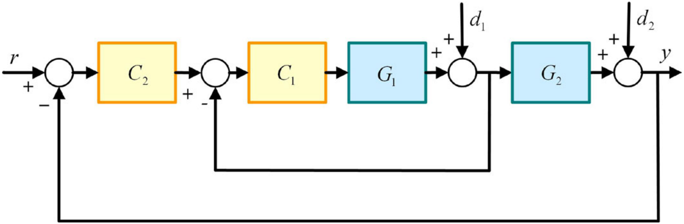
1. The controller, of which the reference is set up by the operator, is the **master** controller. The other is the **slave** controller. The output signal of the master controller serves as the reference of the slave controller.
2. The two feedback loops are nested, with the slave loop located inside the master loop.

The response can be improved by installing a slave controller in between the temperature controller and the controlled steam flow rate. The controller is used to control the flow rate. Such an arrangement constitutes a **cascade control configuration**. In the cascade control system, the slave controller will immediately find the increase in the steam flow rate, and pinch the steam valve to pull the steam flow rate back to the desired value. As a result, the reboiler and the column are only slightly affected by the disturbance from the steam supply pressure.

The features of cascade control are as follows:

1. The controller, of which the reference is set up by the operator, is the **master** controller. The other is the **slave** controller. The output signal of the master controller serves as the reference of the slave controller.
2. The two feedback loops are nested, with the slave loop located inside the master loop.

The diagram of a cascade control system is shown in Figure. In the example of distillation column, the input of $G_2(s)$ is the steam flow rate, and the output of $G_2(s)$ is the column temperature. $C_2(s)$ and $G_2(s)$ constitute the **master loop**. The output of $G_1(s)$ is the input of $G_2(s)$. $C_1(s)$ and $G_1(s)$ constitute the **slave loop**, whose reference is the output of $C_2(s)$. Normally, the slave controller is located close to a potential disturbance to improve the closed-loop response



Question: Whether can a cascade control system always provide superior performance to the single loop system?

Analysis: There are two disturbances in a cascade control system:

- One is the disturbance $d_2(s)$ at the master loop
- The other is the disturbance $d_1(s)$ at the slave loop

Consider $d_2(s)$ first. Assume that only a single loop (that is, the master loop) is applied. The system response to $d_2(s)$ is

$$y(s) = [1 - G_1(s)G_2(s)Q(s)]d_2(s)$$

where $Q(s)$ is the IMC controller corresponding to the plant $G_1(s)G_2(s)$. When the cascade control structure is employed, the effect can be written as

$$y(s) = [1 - G_1(s)Q_1(s)G_2(s)Q_2(s)]d_2(s)$$

where $Q_1(s)$ is the IMC controller corresponding to $G_1(s)$ and $Q_2(s)$ is the IMC controller corresponding to the plant $G_1(s)Q_1(s)G_2(s)$. **If $Q(s) = Q_1(s)Q_2(s)$ is taken, the two responses are the same**

Now consider $d_1(s)$. When only a single loop is used, the system response to $d_1(s)$ can be expressed as

$$y(s) = [1 - G_1(s)G_2(s)Q(s)]G_2(s)d_1(s)$$

If the cascade control structure is used, the system response to $d_1(s)$ is

$$y(s) = [1 - G_1(s)Q_1(s)]G_2(s)d_1(s)$$

The performance of **the cascade control is superior if $Q_1(s)$ can be designed such that**

$$\min_{Q_1} \|[1 - G_1(s)Q_1(s)]G_2(s)\| \le \min_{Q} \|[1 - G_1(s)G_2(s)Q(s)]G_2(s)\|$$

where $Q_1(s)$ is the IMC controller corresponding to $G_1(s)$ and $Q_2(s)$ is the IMC controller corresponding to the plant $G_1(s)Q_1(s)G_2(s)$. **If $Q(s) = Q_1(s)Q_2(s)$ is taken, the two responses are the same**

Now consider $d_1(s)$. When only a single loop is used, the system response to $d_1(s)$ can be expressed as

$$y(s) = [1 - G_1(s)G_2(s)Q(s)]G_2(s)d_1(s)$$

If the cascade control structure is used, the system response to $d_1(s)$ is

$$y(s) = [1 - G_1(s)Q_1(s)]G_2(s)d_1(s)$$

The performance of **the cascade control is superior if $Q_1(s)$ can be designed such that**

$$\min_{Q_1} \|[1 - G_1(s)Q_1(s)]G_2(s)\| \leq \min_{Q} \|[1 - G_1(s)G_2(s)Q(s)]G_2(s)\|$$

where $\|\cdot\|$ denotes some norm. The above inequility is an equality when $G_2(s)$ is MP and stable

Answer

The cascade control is only useful when $G_2(s)$ has RHP zeros or time delay

Design: Assume that $G_2(s)$ is NMP. The following two step design procedure can be used to design the controllers:

1. Assume that there is only the secondary loop. Design $C_1(s)$
2. Regard the secondary loop and $G_2(s)$ as an augmented plant. Design $C_2(s)$

The plants in the master loop and the slave loop are often described by the first-order model with time delay:

$$G_1(s) = \frac{K_1 e^{-\theta_1 s}}{\tau_1 s + 1}, G_2(s) = \frac{K_2 e^{-\theta_2 s}}{\tau_2 s + 1}$$

where $\|\cdot\|$ denotes some norm. The above inequility is an equality when $G_2(s)$ is MP and stable

Answer

The cascade control is only useful when $G_2(s)$ has RHP zeros or time delay

Design: Assume that $G_2(s)$ is NMP. The following two step design procedure can be used to design the controllers:

1. Assume that there is only the secondary loop. Design $C_1(s)$
2. Regard the secondary loop and $G_2(s)$ as an augmented plant. Design $C_2(s)$

The plants in the master loop and the slave loop are often described by the first-order model with time delay:

$$G_1(s) = \frac{K_1 e^{-\theta_1 s}}{\tau_1 s + 1}, \; G_2(s) = \frac{K_2 e^{-\theta_2 s}}{\tau_2 s + 1}$$

Use the two step design procedure. Design $C_1(s)$ first:

$$
\begin{aligned}
Q_1(s) &= \frac{\tau_1 s + 1}{K_1(\lambda_1 s + 1)} \\
C_1(s) &= \frac{Q_1(s)}{1 - G_1(s)Q_1(s)}
\end{aligned}
$$

The closed-loop transfer function is

$$
T_1(s) = \frac{1}{\lambda_1 s + 1} e^{-\theta_1 s}
$$

Regard $T_1(s)$ and $G_2(s)$ as an augmented plant. Design $C_2(s)$:

$$
\begin{aligned}
Q_2(s) &= \frac{(\tau_2 s + 1)(\lambda_1 s + 1)}{K_2(\lambda_2 s + 1)^2} \\
C_2(s) &= \frac{Q_2(s)}{1 - G_2(s)T_1(s)Q_2(s)}
\end{aligned}
$$

The overall closed-loop transfer function is

$$T_2(s) = \frac{1}{(\lambda_2 s + 1)^2} e^{-(\theta_1+\theta_2)s}$$

The following feature significantly reduces the complexity in tuning:

- The nominal performance of the overall system is only determined by the performance degree λ_2
- When there exists uncertainty, the robust performance relates to both λ_1 and λ_2. In general, the robustness is mainly determined by λ_2

Example

This example is used to illustrate how to design the cascade controller. Consider the temperature control system of the distillation column sketched in Figure

Example (ctd.1)

The dynamics of the steam flow rate can be expressed as

$$G_1(s) = \frac{0.68}{0.39s + 1}.$$

The dynamics of the column temperature is

$$G_2(s) = \frac{1.26e^{-0.5s}}{2.11s + 1}$$

Then

$$\begin{aligned} Q_1(s) &= \frac{0.39s + 1}{0.68(\lambda_1 s + 1)} \\ C_1(s) &= \frac{0.39s + 1}{0.68\lambda_1 s} \end{aligned}$$

Example (ctd.2)

The closed-loop transfer function of the slave loop is

$$T_1(s) = \frac{1}{\lambda_1 s + 1}$$

It is readily obtained that

$$\begin{aligned} Q_2(s) &= \frac{(2.11s+1)(\lambda_1 s+1)}{1.26(\lambda_2 s+1)^2} \\ C_2(s) &= \frac{(2.11s+1)(\lambda_1 s+1)}{1.26[(\lambda_2 s+1)^2 - e^{-0.5s}]} \end{aligned}$$

9.4 Anti-Windup Structure

Nonlinear in real systems: The manipulated variable in a real system may be constrained by a physical limit

An example: In a paper-making process the basis weight of the paper is controlled by adjusting the flow rate of stock. The flow rate from a valve has a maximum value, which is determined by the fully open valve

Windup: If the controller output exceeds the maximum value, the valve remains fully open despite that the controller output may continue to change. If the controller includes integral action, the persistent error is integrated. Thereby the output of the integrator becomes quite large

Anti-Windup: The windup puts the system in the state of saturation, which will affect the performance of the system. This issue has to be dealt with in an ad hoc fashion called anti-windup

Assume that the plant input is denoted by the saturation constraints of the controller output:

$$\hat{u}(t) = \mathrm{sat}[u(t)] = \begin{cases} u_{\min}(t) & u(t) < u_{\min}(t) \\ u(t) & u_{\min}(t) \le u(t) \le u_{\max}(t) \\ u_{\max}(t) & u(t) > u_{\ max}(t) \end{cases}$$

A simple method: Adjust the controller parameter, so that the controller output is kept within its physical limit. Recall the control problem for the first-order plant with time delay in Section 6.4. The controller output is

$$u(s) = Q(s)r(s) = \frac{\tau s + 1}{K(\lambda s + 1)} r(s)$$

If the reference is a unit step, its time domain response can be written as

$$u(t) = K\left(\frac{\tau}{\lambda}e^{-t/\lambda} + 1 - e^{-t/\lambda}\right)$$

The controller output can be restricted by properly increasing the performance degree

Advantage: It is simple and the control structure is not changed
Disadvantage: It does not sufficiently utilize $u(t)$ to improve the performance of the closed-loop system within the saturation scope

Optimization based method:
Consider the IMC structure shown in Figure. Define

$$\begin{aligned} y'(t) &:= G(t) * \hat{u}(t) \\ &= \int_0^{\infty} G(t-\tau)\hat{u}(\tau)d\tau \end{aligned}$$

If the reference is a unit step, its time domain response can be written as

$$u(t) = K\left(\frac{\tau}{\lambda}e^{-t/\lambda} + 1 - e^{-t/\lambda}\right)$$

The controller output can be restricted by properly increasing the performance degree

Advantage: It is simple and the control structure is not changed
Disadvantage: It does not sufficiently utilize $u(t)$ to improve the performance of the closed-loop system within the saturation scope

Optimization based method:
Consider the IMC structure shown in Figure. Define

$$\begin{aligned} y'(t) &:= G(t) * \hat{u}(t) \\ &= \int_0^{\infty} G(t-\tau)\hat{u}(\tau)d\tau \end{aligned}$$

where $y'(t)$ is the output of the constrained system. Because of the saturation constraints, $y'(t)$ necessarily differs from $y(t)$, the output of the unconstrained system

It is desirable to keep $y'(t)$ as close to $y(t)$ as possible. Mathematically, the problem is equivalent to solving the following optimization problem instantaneously at each time t:

$$\min_{\hat{u}(t)} |W(t) * y'(t) - W(t) * y(t)|$$

where $W(t)$ is a weighting function that makes $W(s)G(s)$ bi-proper. The introduction of the weighting function is based on the following fact: If $G(s)$ is strictly proper, $\hat{u}(t)$ does not affect $y'(t)$ instantaneously and thus the minimization is meaningless. For the IMC structure shown in Figure

$$\hat{u}(t) = \mathrm{sat}[u(t)] = \mathrm{sat} \int_0^{\infty} Q(t-\tau)e(\tau)d\tau$$

which is completely determined for a given $e(t)$. The conventional IMC implementation, in general, does not solve the problem

In the following figure, a modified IMC structure with the following controller is given:

$$Q(s) = \frac{Q_1(s)}{1 + Q_2(s)}$$

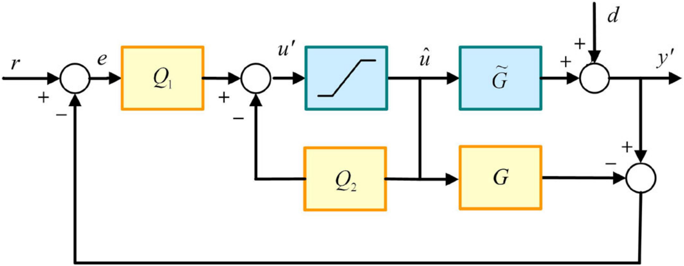


Figure: Modified IMC structure for anti-reset windup.

$Q(s)$ is bi-proper in general. Then

$$\begin{aligned} u(s) &= Q_1(s)e(s) - Q_2(s)\hat{u}(s) \\ &= Q_1(s)e(s) - [Q_1(s)Q^{-1}(s) - 1]\hat{u}(s) \end{aligned}$$

Let

$$Q_3(s) = Q_1(s)Q^{-1}(s)$$

In the time domain,

$$u(t) - \hat{u}(t) = Q_1(t) * e(t) - Q_3(t) * \hat{u}(t)$$

Theorem

Suppose that $Q(s)$ is bi-proper, the model is exact. If $W(s)G(s)$ is finite when $s \to \infty$, and $Q_1(s) = W(s)G(s)Q(s)$, then $\hat{u}(t)$ resulting from the modified IMC implementation is the solution of the optimization problem.

Proof.

Since $Q_1(s) = W(s)G(s)Q(s)$, from Figure we have

$$\begin{aligned} u(t) - \hat{u}(t) &= Q_1(t) * e(t) - Q_3(t) * \hat{u}(t) \\ &= W(t) * y(t) - W(t) * y'(t). \end{aligned}$$

When no saturation occurs,
$u(t) = \hat{u}(t), W(t) * y'(t) - W(t) * y(t) = 0$.Assume that saturation occurs. Since $\hat{u}(t)$ affects $W(t) * y'(t)$ linearly, $|W(t) * y'(t) - W(t) * y(t)|$ is a convex function of $\hat{u}(t)$. If $u(t) = \hat{u}(t)$ for which

$$|W(t) * y'(t) - W(t) * y(t)| = 0$$

is not feasible, the optimal solution must occur at the boundary, that is, $\hat{u}(t) = \text{sat}[u(t)]$ □

Stability:

$Q(s)$ is usually MP and always stable

If $Q(s)$ is MP and $Q_1(s)$ is NMP, then $[1 + Q_2(s)]^{-1}$ must be unstable

To guarantee the internal stability, $Q_1(s)$ must be MP and stable

$W(s)$ should be chosen so that $W(s)G(s)Q(s)$ is MP and stable

Choice of the weighting function: Different controller factorizations can be obtained by choosing different $W(s)$. Two special cases are discussed here.

1. $W(s) = G(s)^{-1}$. The optimization problem becomes

$$\min_{\hat{u}(t)} |u(t) - \hat{u}(t)|$$

The solution corresponds to the conventional IMC structure, which "chops off" the control input resulting in performance deterioration. The stability is guaranteed

2. $W(s)$ is chosen such that $Q_1(s)$ is a constant, for example, $Q_1(s) = Q(\infty)$. The optimization problem becomes

$$\min_{\hat{u}(t)} |Q_1(t)[e(t) - e'(t)]|$$

where

$$e'(t) = Q^{-1}(t) * \hat{u}(t)$$

The performance in this case is greatly improved, but the stability of the closed-loop system is not guaranteed

If the dynamics of $G(s)Q(s)$ are slow, minimizing the weighted error $e(t) - e'(t)$ may not be a good way to optimize the nonlinear performance. After the system comes out of the nonlinear region, the controller takes no action to compensate for the effect of the error introduced during the saturation

In Item 1, $W(s)$ is chosen to guarantee the stability, while in Item 2 $W(s)$ is chosen to enhance the performance. Therefore, $W(s)$ can be tuned to trade off the performance and the stability of the constrained system.

For stable plants, the IMC structure in forgoing figure and the unity feedback loop in the following figure are equivalent

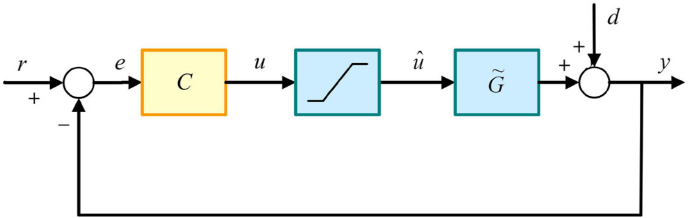


Figure: Unity feedback loop in the presence of actuator constraints

The result for the modified IMC structure shown in the forgoing figure can be extended directly to the unity feedback loop. The obtained anti-windup structure is shown in the following figure. The controllers are defined as follows:

$$
\begin{aligned}
C_1(s) &= Q_1(s) \\
C_2(s) &= Q_2(s) - Q_1(s)G(s)
\end{aligned}
$$

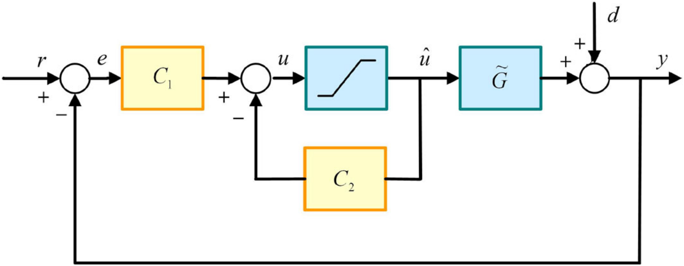


Figure: Modified unity feedback loop for anti-reset windup or

$$\begin{aligned} C_1(s) &= C(\infty) \\ C_2(s) &= \frac{C_1(s)}{C(s)} - 1 \end{aligned}$$

where

$$C(s) = \frac{Q(s)}{1 - G(s)Q(s)}$$

The latter controllers correspond to

$$W(s) = \frac{C_1(s)}{G(s)Q(s)}$$

which minimizes

$$\min_{\hat{u}(t)} \left| C_1(t)[e(t) - e'(t)] \right|$$

Example

Consider the following plant:

$$\boxed{G(s) = \frac{2}{100s + 1}}$$

It is easy to obtain the IMC controller:

$$Q(s) = \frac{100s + 1}{2(\lambda s + 1)}$$

It might as well take $\lambda = 20$.

Case 1 Choosing $W(s) = 2.5(20s + 1)$ results in

$$\boxed{Q_1(s) = 2.5, \;\; Q_2(s) = \frac{4}{100s + 1}}$$

Example (ctd.1)

Case 2 Choosing $W(s) = 50(s+1)$ results in

$$\boxed{Q_1(s) = \frac{50(s+1)}{20s+1}, \quad Q_2(s) = \frac{99}{100s+1}}$$

Here $W(\infty)$ is chosen such that $Q_2(s)$ is strictly proper.

$Q(s)$ in Case 1 corresponds to minimizing $|e(t) - e'(t)|$, while $Q(s)$ in Case 2 corresponds approximately to minimizing $|y(t) - y'(t)|$. Assume that the input is constrained between the saturation limits ± 1. The responses to a unit step disturbance with the conventional IMC implementation and the modified IMC implementation are shown in Figures. The control input in Case 1 stays saturated until $e(t) = e'(t)$, while the control input in Case 2 stays saturated until $y(t) \approx y'(t)$

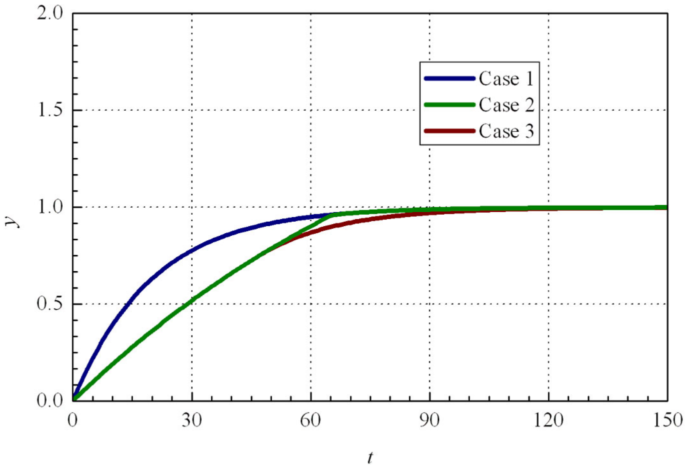


Figure: System output responses

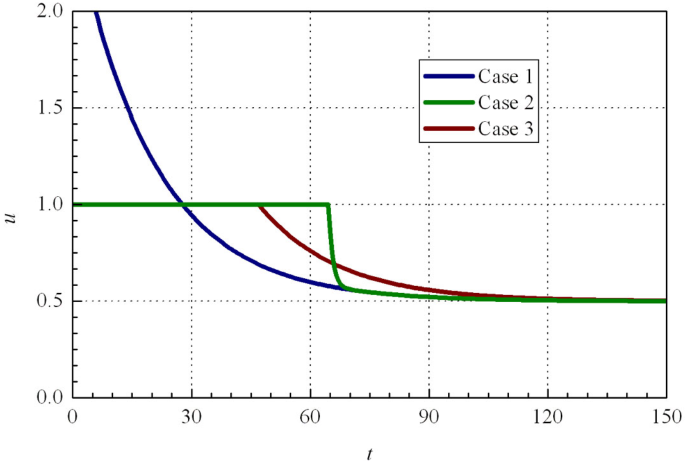


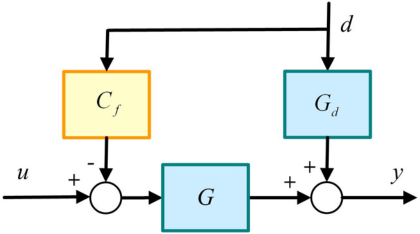
Figure: Plant input responses

9.5 Feedforward Control

Limitation of the feedback control: The disturbance can never be **completely** eliminated
Reason: The feedback controller takes action only after the error caused by the disturbance happens

Feedforward control: When the disturbance entering a system is known, it can **exactly** be compensated by introducing an additional loop

The feedforward structure is shown in Figure, where $C_f(s)$ is the feedforward controller, and $G_d(s)$ is the transfer function of the disturbance channel



The system output is

$$y(s) = G_d(s)d(s) - C_f(s)G(s)d(s)$$

Assume that the model is exact. To compensate the disturbance completely, the controller should be

$$C_f(s) = \frac{G_d(s)}{G(s)}$$

The exact compensation is achievable **only** when $C_f(s)$ is stable. This implies that $G(s)$ is MP or $G_d(s)$ has zeros wherever $G(s)$ has NMP zeros

It is possible that $G(s)$ is NMP, and has NMP zeros different from those of $G_d(s)$. In this case, $C_f(s)$ should be designed so that the effect of the disturbance on the system output is minimized

The transfer function from the disturbance to the system output is

$$\frac{y(s)}{d(s)} = G_d(s) - G(s)C_f(s)$$

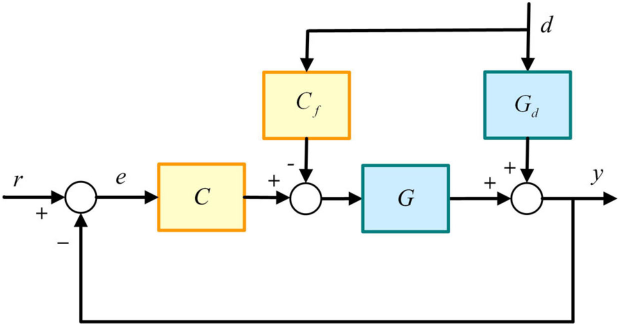
Then the optimization problem can be expressed as

$$\min \|W(s)[G_d(s) - G(s)C_f(s)]\|$$

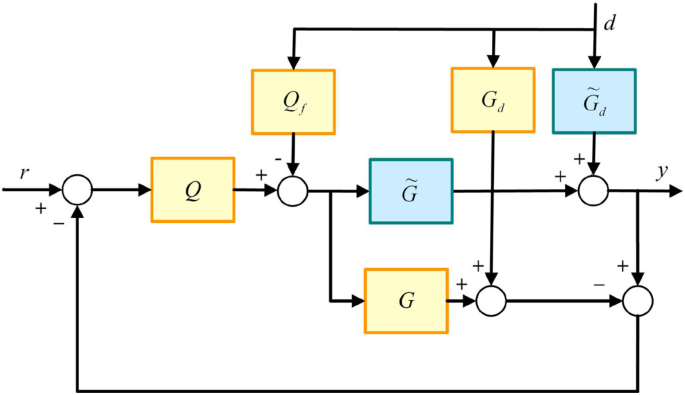
where $\|\cdot\|$ denotes some norm and $W(s)$ is a weighting function. This problem can be converted into the quasi-H_∞ design problem or the H_2 design problem.

Feedforward control is an open-loop control strategy. If the model is not exact, the effect of the disturbance on the system output will not vanish. The system output will show a deviation from the reference. Consequently, feedfoward control is seldom used alone. A frequently used scheme is the combined feedforward/feedback control, which is shown in Figure. In the structure, the effect of the disturbance on the system output is

$$y(s) = \frac{G_d(s)}{1 + G(s)C(s)}d(s) - \frac{C_f(s)G(s)}{1 + G(s)C(s)}d(s)$$



The feedforward controller is still $C_f(s) = G_d(s)/G(s)$. In the system, the feedforward controller is used to compensate the disturbance, whereas the feedback action attempts to eliminate the effects of uncertainty. Since the characteristic equation of the system is not changed, the introduction of the feedforward controller will not affect the stability



The feedforward controller can also be introduced to the IMC structure. In Figure, $Q(s)$ is the IMC controller and $Q_f(s)$ is the feedforward controller

They are related to the conventional controllers $C(s)$ and $C_f(s)$ through

$$\begin{aligned} C(s) &= \frac{Q(s)}{1 - G(s)Q(s)} \\ C_f(s) &= \frac{Q_f(s) - G_d(s)Q(s)}{1 - G(s)Q(s)} \end{aligned}$$

When there is no model error,

$$y(s) = [G_d(s) - G(s)Q_f(s)]d(s)$$

To compensate the disturbance completely the controller should be

$$Q_f(s) = \frac{G_d(s)}{G(s)}$$

Note that the feedforward controller in the IMC structure is in the same form as that in the unity feedback loop

Consider the design of a feedforward controller for a system in which the plant and the transfer function of the disturbance channel can be adequately described by the first-order model with time delay, that is,

$$\begin{aligned} G(s) &= \frac{Ke^{-\theta s}}{\tau s + 1} \\ G_d(s) &= \frac{K_d e^{-\theta_d s}}{\tau_d s + 1} \end{aligned}$$

The feedforward controller is as follows:

$$Q_f(s) = \frac{K_d}{K}\frac{\tau s + 1}{\tau_d s + 1} e^{(\theta - \theta_d)s}$$

A typical case is $\theta \approx \theta_d$. Then the controller is simplified to

$$Q_f(s) = \frac{K_d}{K}\frac{\tau s + 1}{\tau_d s + 1}$$

If the dynamic terms of the controller is removed, the steady-state feedforward controller Q_{fs} will be obtained:

$$Q_{fs} = \frac{K_d}{K}$$

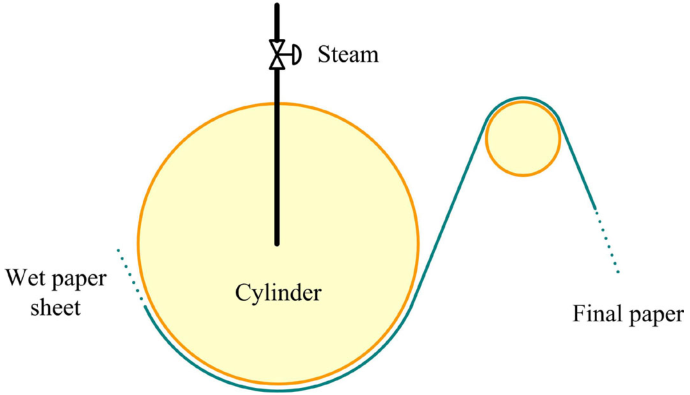
When the model is exact, the feedforward/feedback loop can be viewed as the unity feedback loop without disturbance

Example

An important step in the paper-making process is drying. The dryer of a paper-making machine consists of many cylinders that are 1.5m in diameter and 2.1m in length. Each cylinder is fully filled with vapor. When the wet paper goes through the surface of these cylinders, most of the water evaporates and the final paper with proper moisture content is obtained (Figure). To reduce the fluctuating magnitude and frequency of temperature change on the surface of the cylinder, the feedforward/feedback scheme is applied

Example (ctd.1)

By mechanism analysis, the temperature of the cylinder is chosen as the system output, the steam flow rate is chosen as the control variable, and the steam temperature is regarded as the disturbance



Example (ctd.2)

The models of a cylinder are as follows:

$$
\begin{aligned}
G(s) &= \frac{1.65 \times 0.48}{(48s+1)(10s+1)} \\
G_d(s) &= \frac{0.0636}{42.6s+1}
\end{aligned}
$$

Then the feedforward controller is

$$
C_f(s) = \frac{0.08(48s+1)(10s+1)}{42.6s+1}
$$

This is an improper transfer function. It can be physically realized by introducing a low-pass filter

9.6 Optimal Input Disturbance Rejection

Optimal output disturbance rejection: The goal is to minimize $\|W(s)S(s)\|_2$ or $\|W(s)S(s)\|_\infty$

Design methods of this kind have several merits:

1. The design procedure is simple and easy to understand
2. The resulting controller is usually of low-order and easy to implemented
3. Attention is paid on both reference tracking and disturbance rejection

Optimal input disturbance rejection: The design specification is to minimize $\|W(s)G(s)S(s)\|_2$

As introduced in Section 8.1, a general plant can be described by

$$G(s) = \frac{KN_+(s)N_-(s)}{M_+(s)M_-(s)}e^{-\theta s}$$

9.6 Optimal Input Disturbance Rejection

Optimal output disturbance rejection: The goal is to minimize $\|W(s)S(s)\|_2$ or $\|W(s)S(s)\|_\infty$

Design methods of this kind have several merits:

1. The design procedure is simple and easy to understand
2. The resulting controller is usually of low-order and easy to implemented
3. Attention is paid on both reference tracking and disturbance rejection

Optimal input disturbance rejection: The design specification is to minimize $\|W(s)G(s)S(s)\|_2$

As introduced in Section 8.1, a general plant can be described by

$$G(s) = \frac{KN_+(s)N_-(s)}{M_+(s)M_-(s)}e^{-\theta s}$$

The rational part of the plant is MP if $N_+(s) = 1$, and the plant is stable if $M_+(s) = 1$. For a step input, the $Q(s)$ that guarantees the internal stability and the asymptotic tracking property can be written as

$$Q(s) = \frac{[1 + sQ_2(s)]M_+(s)}{K}$$

where $Q_2(s)$ is any stable transfer function that makes $Q(s)$ proper and satisfies

$$\lim_{s \to p_j} \frac{d^k}{ds^k}\left\{1 - \frac{[1 + sQ_2(s)]N_+(s)N_-(s)e^{-\theta s}}{M_-(s)}\right\} = 0, k = 0, 1, ..., l_j - 1$$

Here $p_j (j = 1, 2, ..., r_p)$ are the l_j multiplicity unstable poles of $G(s)$. Take $W(s) = 1/s$. Then

$$\begin{aligned} & \|W(s)G(s)S(s)\|_2^2 \\ = \quad & \left\| W(s)G(s)\left\{1 - \frac{G(s)M_+(s)}{K}[1 + sQ_2(s)]\right\}\right\|_2^2 \end{aligned}$$

$$= \left\| \begin{array}{l} \frac{1}{s}\frac{KN_+(s)N_-(s)}{M_+(s)M_-(s)}e^{-\theta s} \\ \left[1 - \frac{N_+(s)N_-(s)}{M_-(s)}e^{-\theta s} - \frac{N_+(s)N_-(s)s}{M_-(s)}e^{-\theta s}Q_2(s)\right] \end{array} \right\|_2^2$$

$$= K^2 \left\| \begin{array}{l} \frac{N_-(s)N_+^2(-s)e^{\theta s} - N_+(s)}{sM_+(-s)M_-(s)N_+(s)} + \\ \frac{M_-(s) - N_-^2(s)N_+^2(-s)}{sM_+(-s)M_-^2(s)} - \frac{N_+^2(-s)N_-^2(s)}{M_+(-s)M_-^2(s)}Q_2(s) \end{array} \right\|_2^2$$

Suppose that the plant is MP, that is, $\theta = 0$ and $N_+(s) = 1$. Then

$$\|W(s)G(s)S(s)\|_2^2$$

$$= K^2 \left\| \frac{N_-(s) - 1}{sM_+(-s)M_-(s)} + \frac{M_-(s) - N_-^2(s)}{sM_+(-s)M_-^2(s)} - \frac{N_-^2(s)}{M_+(-s)M_-^2(s)}Q_2(s) \right\|_2^2$$

$$= K^2 \left\| \frac{M_-(s)N_-(s) - N_-^2(s)}{sM_+(-s)M_-^2(s)} - \frac{N_-^2(s)}{M_+(-s)M_-^2(s)}Q_2(s) \right\|_2^2$$

Since s is the factor of $M_-(s)N_-(s) - N_-^2(s)$, the optimal $Q_2(s)$ is

$$Q_{2opt}(s) = \frac{M_-(s)N_-(s) - N_-^2(s)}{sN_-^2(s)}$$

The optimal $Q(s)$ can readily be obtained:

$$Q_{opt}(s) = \frac{M_-(s)M_+(s)}{KN_-(s)}$$

Evidently, the controller is identical to the one designed for output disturbances

This implies that for MP plants the design for input disturbances cannot provide any improved disturbance response. Nevertheless, the disturbance response can be improved for NMP plants

To illustrate the problem, consider the simplest NMP plant described by the following transfer function:

$$G(s) = \frac{K(-z_r^{-1}s+1)}{(\tau_1 s+1)(\tau_2 s+1)}$$

where τ_1, τ_2, and z_r are positive numbers. We have

$$\begin{aligned}
&\|W(s)G(s)S(s)\|_2^2 \\
= \quad & K^2 \left\| \begin{array}{l} \frac{(z_r^{-1}s+1)^2-(-z_r^{-1}s+1)}{s(\tau_1 s+1)(\tau_2 s+1)(-z_r^{-1}s+1)} + \\ \frac{(\tau_1 s+1)(\tau_2 s+1)-(z_r^{-1}s+1)^2}{s(\tau_1 s+1)^2(\tau_2 s+1)^2} - \frac{(z_r^{-1}s+1)^2}{(\tau_1 s+1)^2(\tau_2 s+1)^2} Q_2(s) \end{array} \right\|_2^2 \\
= \quad & K^2 \left\| \begin{array}{l} \frac{z_r^{-2}s+3z_r^{-1}}{(\tau_1 s+1)(\tau_2 s+1)(-z_r^{-1}s+1)} + \\ \frac{(\tau_1\tau_2-z_r^{-2})s+(\tau_1+\tau_2-2z_r^{-1})}{(\tau_1 s+1)^2(\tau_2 s+1)^2} - \frac{(z_r^{-1}s+1)^2}{(\tau_1 s+1)^2(\tau_2 s+1)^2} Q_2(s) \end{array} \right\|_2^2
\end{aligned}$$

Let

$$
\begin{aligned}
a_0 &= \frac{4z_r^{-2}}{\tau_1+\tau_2+z_r\tau_1\tau_2+z_r^{-1}}\\
a_1 &= z_r^{-1}-a_0+\tau_1+\tau_2\\
a_2 &= (z_r\tau_1\tau_2-\tau_1-\tau_2)a_0+3z_r^{-1}(\tau_1+\tau_2)+\tau_1\tau_2-z_r^{-2}\\
a_3 &= [z_r\tau_1\tau_2(\tau_1+\tau_2)-\tau_1\tau_2]a_0+3z_r^{-1}\tau_1\tau_2\\
a_4 &= a_0 z_r\tau_1^2\tau_2^2
\end{aligned}
$$

It follows that

$$
\begin{aligned}
&\|W(s)G(s)S(s)\|_2^2\\
= \quad & K^2\left\|\frac{a_0}{-z_r^{-1}s+1}\right\|_2^2+\\
& K^2\left\|\frac{a_4s^3+a_3s^2+a_2s+a_1}{(\tau_1s+1)^2(\tau_2s+1)^2}-\frac{(z_r^{-1}s+1)^2}{(\tau_1s+1)^2(\tau_2s+1)^2}Q_2(s)\right\|_2^2
\end{aligned}
$$

Minimize the right-hand side of the equation. The optimal controller is

$$Q_{opt}(s) = \frac{1}{K} + \frac{(a_4 s^3 + a_3 s^2 + a_2 s + a_1)(\tau_1 s + 1)(\tau_2 s + 1)s}{K(z_r^{-1}s + 1)^2}$$

The corresponding optimal performance is

$$\begin{aligned} \text{ISE1} &= \min \|W(s)G(s)S(s)\|_2 \\ &= K \left\| \frac{a_0}{-z_r^{-1}s + 1} \right\|_2 \\ &= K a_0 \sqrt{\frac{z_r}{2}} \end{aligned}$$

If instead of the input disturbance, the controller is designed for the output disturbance. One can obtain the following optimal controller based on the discussion in Section 6.2:

$$Q_{opt}(s) = \frac{(\tau_1 s + 1)(\tau_2 s + 1)}{K(z_r^{-1}s + 1)}$$

The optimal performance corresponding to the input disturbance is

$$\begin{aligned}
\mathrm{ISE2} &= \min \|W(s)G(s)S(s)\|_2 \\
&= K \left\| \frac{2z_r^{-1}}{(\tau_1 s + 1)(\tau_2 s + 1)} \right\|_2 \\
&= K z_r^{-1} \sqrt{\frac{2}{\tau_1 + \tau_2}}
\end{aligned}$$

To compare the two performances, we calculate the ratio of them:

$$\begin{aligned}
\frac{\mathrm{ISE2}}{\mathrm{ISE1}} &= \frac{z_r(\tau_1 + \tau_2) + \tau_1\tau_2 z_r^2 + 1}{2\sqrt{z_r(\tau_1 + \tau_2)}} \\
&= \frac{\sqrt{z_r(\tau_1 + \tau_2)}}{2} + \frac{1}{2\sqrt{z_r(\tau_1 + \tau_2)}} + \frac{\tau_1\tau_2 z_r^2}{2\sqrt{z_r(\tau_1 + \tau_2)}}
\end{aligned}$$

It is concluded that the ratio is always greater than one

To see this, let $a = \sqrt{z_r(\tau_1 + \tau_2)}$. Since $(a-1)^2 \geq 0$, $a + 1/a \geq 2$. This implies that the sum of the first two terms in the right-hand side is greater than or equal to one. The third term is a positive number

For NMP plants the disturbance response can be improved by designing the controller for input disturbances

It is easy to verify that when the controller tends to be optimal in the system with an MP plant, $S(s)$ tends to be zero. Both of the two designs can reach the optimal rejection for input disturbances. Nevertheless, in the system with an NMP plant, when the controller tends to be optimal, instead of zero $S(s)$ tends to be a constant. In this case, the optimal performance with regard to input disturbances is affected by both $G(s)$ and $S(s)$

The effect of the **NMP zero** is evident. The larger the z_r, the larger the performance difference between the two designs

Consider the effect of the plant **poles**. When the pole of the plant is close to the imaginary axis (that is, τ_1 or τ_2 is large), the third term in the right-hand side is large. The performance of the design for output disturbances will be worse than that for input disturbances

In practice, the controller is seldom designed for input disturbances if the plant does not have poles close to the imaginary axis.

- On one hand, the result is complex
- On the other hand, the reference response is usually worsened

Example

Anesthesia can be administered automatically by a control system. For certain operations, such as brain and eye surgery, involuntary muscle movements can be disastrous. To ensure adequate operating conditions for the surgeon, muscle relaxant drugs, which block involuntary muscle movements, are administered.

A conventional method used by anesthesiologists for muscle relaxant administration is to inject a bolus dose and to inject supplements as required. However, an anesthesiologist may sometimes fail to maintain a steady level of relaxation, resulting in a large amount of drug consumption or unexpected side effect. Significant improvements may be achieved by introducing the concept of automatic control, which results in considerable reduction in the total drug consumed.

Example (ctd.1)

As the level of relaxation cannot be directly measured, the arterial blood pressure is chosen as its proxy. When the blood pressure increases, the level of relaxation decreases. Assume that the body dynamics is

$$\boxed{G(s) = \frac{Ke^{-\theta s}}{\tau s + 1}}$$

where $K = 2$, $\tau = 0.2$, and $\theta = 0.1$. If the controller is designed for output disturbances, one obtains

$$Q(s) = \frac{0.2s + 1}{2(\lambda s + 1)}$$

It might as well take $\lambda = 0.02$.

Example (ctd.2)

Consider the controller for input disturbances. Using the design method in this section, the following controller can be obtained:

$$Q(s) = \frac{(\tau s + 1)[1 + \tau s(1 - e^{-\theta/\tau})]}{K(\lambda s + 1)^2}$$

Substituting the plant parameters into the controller yields

$$\boxed{Q(s) = \frac{(0.2s + 1)(0.0787s + 1)}{2(\lambda s + 1)^2}}$$

To obtain the same rise time, take $\lambda = 0.04$. The closed-loop responses are shown in Figure. The design for the output disturbance has a better reference response, while the design for input disturbances has a better disturbance response. As the price, the controller for the input disturbance is complex and the the reference response is worsened

When the plant has poles close to the imaginary axis, three ways can be used to design the controller for input disturbances:

1. Introduce zeros by $S(s)$ to cancel the poles of $G(s)$ in the performance index $\|W(s)G(s)S(s)\|_2$
2. Design the system as a type 2 one
3. Introduce zeros by $W(s)$ to cancel the poles of $G(s)$ in the performance index $\|W(s)S(s)\|_2$

The first method has been discussed in the first half part of this section
The second method is, in fact, given in Chapter 7
The third method is mainly used in two cases:

- Simplify the design
- Tradeoff between the design for input disturbances and the design for output disturbances

In the plant whose order is greater than one, normally not all poles are close to the imaginary axis. Assume that there is only one pole, τ_1, close to the imaginary axis. The controller designed by the first method is complex. A simplified method is to choose a simple weighting function in the third method. For example, the following weighting function is taken:

$$W(s) = \frac{(\lambda s + 1)(\gamma s + 1)}{s(\tau_1 s + 1)}$$

where $\gamma \in [\lambda, \tau_1]$. The rest design is similar to the first method.

It can be seen that the choice implies that only the slow pole, τ_1, is considered in the weighting function. The main design objective is the input disturbance when $\gamma = \lambda$, while the main design objective is the output disturbance when $\gamma = \tau_1$. By choosing an appropriate γ in between λ and τ_1, one can easily tradeoff between the design for input disturbances and the design for output disturbances

9.7 Control of Plants with Multiple Time Delays

Frequently used model:

$$G(s) = \frac{KN_+(s)N_-(s)}{M_+(s)M_-(s)}e^{-\theta s}$$

More complex case: There are multiple time delays in the nominator or the denominator of the plant. This corresponds to the situation when the plant has multiple state time delays and output time delays

It is fairly difficult to rigorously treat such plants in control system design. Therefore, they are usually reduced to the form of rational transfer functions with time delays

9.7 Control of Plants with Multiple Time Delays

Frequently used model:

$$G(s) = \frac{KN_+(s)N_-(s)}{M_+(s)M_-(s)}e^{-\theta s}$$

More complex case: There are multiple time delays in the nominator or the denominator of the plant. This corresponds to the situation when the plant has multiple state time delays and output time delays

It is fairly difficult to rigorously treat such plants in control system design. Therefore, they are usually reduced to the form of rational transfer functions with time delays

To explain why this treatment is necessary, consider a plant with dual time delays. The plant is described by

$$G(s) = \frac{e^{-\theta_1 s}}{\tau s + 1} - \frac{Ke^{-\theta_2 s}}{\tau s + 1}$$

which consists of two parallel stable plants with the same time constant.

Without loss of generality, let $\theta_1 < \theta_2$ and $\theta = \theta_2 - \theta_1$. Then

$$G(s) = \frac{1 - Ke^{-\theta s}}{\tau s + 1} e^{-\theta_1 s}$$

If the design specification is to minimize $\|W(s)S(s)\|_2$, the optimal controller $Q(s)$ should be identical to that for the following plant:

$$G(s) = \frac{1 - Ke^{-\theta s}}{\tau s + 1}$$

For step inputs, the $Q(s)$ satisfying the internally stability and the asymptotic tracking property can be written as

$$Q(s) = \frac{1}{1-K} + sQ_1(s)$$

where $Q_1(s)$ is stable. Take $W(s) = 1/s$, then

$$\begin{aligned}
& \|W(s)S(s)\|_2^2 \\
= & \left\| W(s)\left\{1 - G(s)\left[\frac{1}{1-K} + sQ_1(s)\right]\right\}\right\|_2^2 \\
= & \left\| \frac{1}{s}\left\{1 - \frac{1-Ke^{-\theta s}}{\tau s+1}\left[\frac{1}{1-K} + sQ_1(s)\right]\right\}\right\|_2^2 \\
= & \left\| \frac{(\tau s+1)(1-K) - (1-Ke^{-\theta s})}{s(\tau s+1)(1-K)} - \frac{1-Ke^{-\theta s}}{\tau s+1}Q_1(s)\right\|_2^2
\end{aligned}$$

Let $1 - Ke^{-\theta s} = 0$. An infinity number of RHP zeros of $G(s)$ are obtained:

$$z_k = \frac{\ln K + 2k\pi j}{\theta}$$

where k is any integer. The all-pass function must have poles at the mirror images of the zeros:

$$p_k = \frac{-\ln K + 2k\pi j}{\theta}$$

which relates to $-K + e^{-\theta s} = 0$. Consequently,

$$\begin{aligned} & \|W(s)S(s)\|_2^2 \\ = & \left\| \begin{array}{l} \frac{[(\tau s+1)(1-K)-(1-Ke^{-\theta s})](-K+e^{-\theta s})}{s(\tau s+1)(1-K)(1-Ke^{-\theta s})} - \\ \frac{-K+e^{-\theta s}}{\tau s+1} Q_1(s) \end{array} \right\|_2^2 \end{aligned}$$

$$= \left\| \frac{\frac{-K+e^{-\theta s}-1+Ke^{-\theta s}}{s(1-Ke^{-\theta s})}+}{\frac{(\tau s+1)(1-K)-(-K+e^{-\theta s})}{s(\tau s+1)(1-K)}} - \frac{-K+e^{-\theta s}}{\tau s+1}Q_1(s) \right\|_2^2$$

$$= \left\| \frac{-K+e^{-\theta s}-1+Ke^{-\theta s}}{s(1-Ke^{-\theta s})} \right\|_2^2 +$$

$$\left\| \frac{(\tau s+1)(1-K)-(-K+e^{-\theta s})}{s(\tau s+1)(1-K)} - \frac{-K+e^{-\theta s}}{\tau s+1}Q_1(s) \right\|_2^2$$

Minimizing the right-hand side of the equality yields $Q_1(s)$. Then the optimal controller is

$$Q_{opt}(s) = \frac{\tau s+1}{-K+e^{-\theta s}}$$

Introduce the following filter:

$$J(s) = \frac{1}{\lambda s+1}$$

The unity feedback loop controller is

$$C(s) = \frac{\tau s + 1}{(\lambda s + 1)(-K + e^{-\theta s}) - (1 - Ke^{-\theta s})e^{-\theta_1 s}}$$

To implement the controller in the unity feedback loop, the reduction technique has to be used

For general plants, it is **impossible** to analytically design the controller when the time delay is rigorously treated. The difficulty is that the all-pass part of the plant cannot be constructed. For example, consider the following plant with dual time delays:

$$\begin{aligned} G(s) &= \frac{e^{-\theta_1 s}}{\tau_1 s + 1} + \frac{e^{-\theta_2 s}}{\tau_2 s - 1} \\ &= \frac{(\tau_2 s - 1)e^{-\theta_1 s} + (\tau_1 s + 1)e^{-\theta_2 s}}{(\tau_1 s + 1)(\tau_2 s - 1)} \end{aligned}$$

The unity feedback loop controller is

$$C(s) = \frac{\tau s + 1}{(\lambda s + 1)(-K + e^{-\theta s}) - (1 - Ke^{-\theta s})e^{-\theta_1 s}}$$

To implement the controller in the unity feedback loop, the reduction technique has to be used

For general plants, it is **impossible** to analytically design the controller when the time delay is rigorously treated. The difficulty is that the all-pass part of the plant cannot be constructed. For example, consider the following plant with dual time delays:

$$\begin{aligned} G(s) &= \frac{e^{-\theta_1 s}}{\tau_1 s + 1} + \frac{e^{-\theta_2 s}}{\tau_2 s - 1} \\ &= \frac{(\tau_2 s - 1)e^{-\theta_1 s} + (\tau_1 s + 1)e^{-\theta_2 s}}{(\tau_1 s + 1)(\tau_2 s - 1)} \end{aligned}$$

Let $(\tau_2 s - 1)e^{-\theta_1 s} + (\tau_1 s + 1)e^{-\theta_2 s} = 0$. The zeros of the plant are the solution of the following equation:

$$e^{-(\theta_1-\theta_2)s} = \frac{\tau_1 s + 1}{-\tau_2 s + 1}$$

It is a challenge to construct an all-pass transfer function with this equation

End of Chapter 9