Chapter 4 Fundamentals of PID Control

4.1 Introduction

With its three-term functionality offering treatment of both transient and steady-state responses, proportional-integral-derivative (PID) control provides a generic and efficient solution to real world control problems. The PID controllers were the most popular controllers of the century. Many thousands of Instrument and Control Engineers world-wide are using such controllers in their daily work. According to a survey held in 1977, 34 out of 37 listed industrial analogue controllers were of the PID type. The same is true till today and well over ninety percent of existing control loops involve PID controllers. These controllers will well remain dominant in 21st century because of:

1. They are very simple to implement and easy to use
2. They are very effectiveness
3. They are robust, insensitive to changes to plant parameter and disturbance.
4. They have wide Application the application area cover all industry fields, such as, Petrochemical, Pharmaceuticals, Food, Chemical, Aerospace and Semiconductor, building HVAC systems, etc.

In PID control system design, controller parameters are usually tuned so that the closed-loop system meets the following five objectives:

1) stability and stability robustness, usually measured in the frequency domain
2) transient response, including rise time, overshoot, and settling time
3) steady-state accuracy
4) disturbance attenuation and robustness against environmental uncertainty, often at steady state
5) robustness against plant modeling uncertainty, usually measured in the frequency domain.

Early PID controllers were pneumatic and gained widespread industrial acceptance during the 1940s. Their electronic counterparts entered the market in the 1950s. Over the past fifty years, an enormous amount of effort has been expended in designing these controllers. Despite these advancements and the popularity of this approach, the design of PID controllers is still a challenge for engineers and researchers. Since the 1940s, many methods for tuning single-loop and multi-loop/multivariable PID controllers have been proposed but every method has some limitations. The purpose of this Chapter is to introduce fundamental concepts of PID control design, remedies for problems involving the integral and derivative terms. PID design objectives, methods, and implantations.

4.2 PID controller structure and functionality

4.2.1 PID controller structure

Consider the Block Diagram of a simple control system given as in Figure 4.2.1

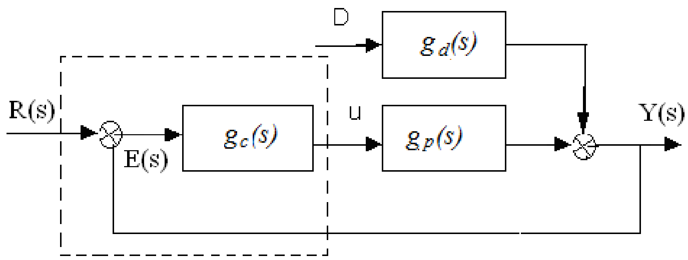


Fig 4.2.1 Block Diagram of Control Systems

We define

$$E(s) = R(s) - Y(s)$$

The transfer function of a PID controller is often expressed in the ideal form

$$g_c(s) = \frac{U(s)}{E(s)} = K_P\left(1 + \frac{1}{\tau_I s} + \tau_D s\right) \tag{4.1.1}$$

where *U(s)* is the control signal acting on the error signal *E(s)*, K_P is the proportional gain, τ_I is the integral time constant, τ_D is the derivative time constant, and *s* is the argument of the Laplace transform. The control signal can also be expressed in three terms as

$$\begin{aligned} U(s) &= K_P E(s) + K_I \frac{1}{s} E(s) + K_D s E(s) \\ &= U_P(s) + U_I(s) + U_D(s) \end{aligned} \tag{4.1.2}$$

where $K_I = K_P / \tau_I$ is the integral gain and $K_D = K_P \tau_D$ is the derivative gain.

If $\tau_I \geq 4\tau_D$, the PID controller can be realized in a series form

$$\begin{aligned} g_{PID}(s) &= K_P(\alpha + \tau_D s)\left(1 + \frac{1}{\alpha \tau_I s}\right) \\ &= g_{PD}(s) g_{PI}(s) \end{aligned} \tag{4.1.3}$$

where $g_{PD}(s)$ and $g_{PI}(s)$ are the factored PD and PI parts of the PID controller, respectively, and

$$\alpha = \frac{1 \pm \sqrt{1 - 4\tau_D / \tau_I}}{2} > 0 \tag{4.1.4}$$

The reason for using the series form is that the PID rules with derivative action are then much simpler. The corresponding settings for the ideal (parallel form) PID controller are easily obtained.

In comparison with the classical control theory, the PID control structure can be considered as:

1. PI controller structure is an extreme form of phase-lag compensators.
2. PD controller structure is an extreme form of phase-lead compensators.
3. PID controller structure is an extreme form of a phase lead-lag compensator with one

pole at the origin and the other at infinity;

The three-term functionalities include:

1) The proportional term provides an overall control action proportional to the error signal through the all pass gain factor.
2) The integral term reduces steady-state errors through low-frequency compensation.
3) The derivative term improves transient response through high-frequency compensation.

For optimum performance, K_P, K_I (or τ_I), and K_D (or τ_D) must be tuned jointly, although the individual effects of these three parameters on the closed-loop performance of stable plants are summarized in Table 1.

TABLE 1 Effects of independent *P*, *I*, and *D* tuning on closed-loop response.

	Rise Time	**Overshoot**	**Settling Time**	**Steady-State Error**	**Stability**
Increasing k_P	Decrease	Increase	Small Increase	Decrease	Degrade
Increasing k_I	Small Decrease	Increase	Increase	Large Decrease	Degrade
Increasing k_D	Small Decrease	Decrease	Decrease	Minor Change	Improve

Example 4.2.1: while K_I and K_D are fixed, increasing K_P alone can decrease rise time, increase overshoot, slightly increase settling time, decrease the steady-state error, and decrease stability margins.

4.2.2 Property of *P* Control

P control has steady state error. The controller output *u* is proportional to error signal *e*:

$$u = K_p e \tag{4.2.1}$$

Example 4.2.2: A temperature control for heating system as shown in Fig.4.2.1 where transducer θ_T obtains temperature of hot water θ and sends to temperature controller θ_c. It regulates the opening of steam valve to keep constant outlet temperature. The load of heater depends both on hot water flow rate *Q* and outlet temperature θ, using P control and opening of regulating valve μ as controller output.

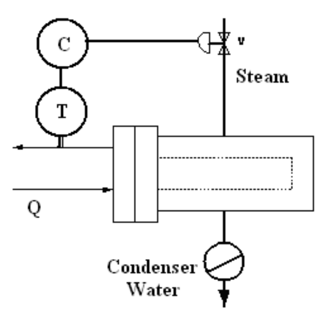


Fig.4.2.1 Temperature control for heating system

In Fig.4.2.2, Line 1 shows the static property of P controller, the higher of the hot water temperature, the smaller opening of regulating valve. Line 2 and 3 represent static property of heater at different flow rate, indicating the relation between hot water temperature and opening of regulating valve without control.

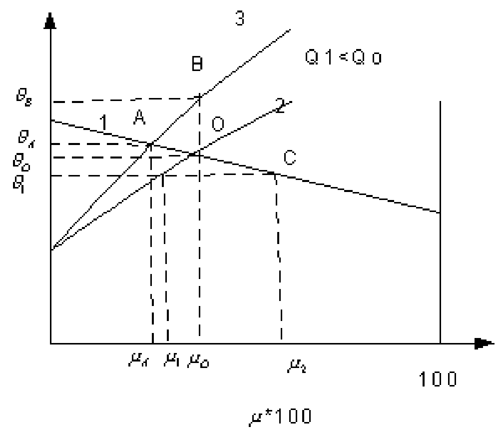


Fig. 4.2.2 Property of P Controller

Point *O* is initial steady state where flow rate is Q_o, temperature is θ_o and opening of regulating valve is μ_o. If the flow rate reduced to Q_1, new steady state will be at point *A* after regulation. There exists and error θ_A - θ_o because of P control.

Even with steady state error, it is still better than without control, which results the temperature end at θ_B.

4.2.3 Integral Function

In *I* regulation, the change rate of controller output is proportional to error *e*.

$$\frac{du}{dt} = K_I e \Rightarrow u = K_I \int_0^t e dt \tag{4.2.2}$$

However, *I* regulation alone is rarely used in industrial process control; It mostly combined with *P* function to form *PI* Control. By combining the advantages of the two; *P* to improve response time and reject disturbance, *I* to eliminate steady state error. Figure 4.2.3 shows the output response to each individual function of PI control.

$$u = K_p e + K_I \int_0^t e dt = K_p \left(e + \frac{1}{\tau_I} \int_0^t e dt \right) \tag{4.2.3}$$

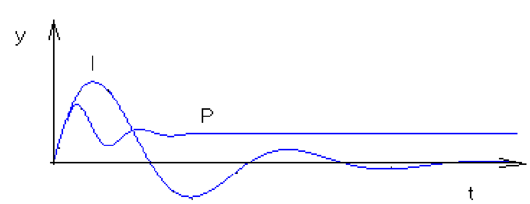


Figure 4.2.3 *PI* control

Example 4.2.3: Fig. 4.2.4 illustrates the regulating process when the hot water flow rate has a step reduction for the heating process.

Assume the hot water temperature θ is given as in Fig.4.2.1, μ_p is the proportional part of the *PI* controller (regulating valve) which is opposite to θ curve. μ_I is the integral part of the *PI* controller which is integration of θ curve. The total output of regulating valve μ_{pI} is the sum of μ_p and μ_I.

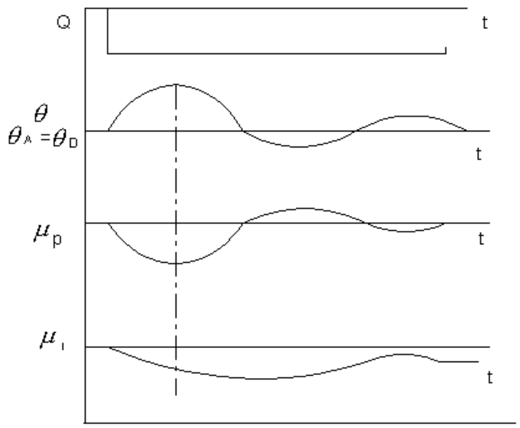


Figure 4.2.4 Step response of PI controller

Destabilizing Effect of the Integral Term

Referring to Equation (4.1.1) for $\tau_I \neq 0$ and $\tau_D = 0$, it can be seen that adding an integral term to a pure proportional term increases the gain by a factor of

$$\left|1+\frac{1}{j\omega\tau_I}\right| = \sqrt{1+\frac{1}{\omega^2\tau_I^2}} > 1 \text{ for all } \omega \qquad (4.2.4)$$

and simultaneously increases the phase-lag since

$$\angle\left(1+\frac{1}{j\omega\tau_I}\right) = \tan^{-1}\left(\frac{-1}{\omega\tau_I}\right) < 0 \text{ for all } \omega \qquad (4.2.5)$$

Considering

$$\log|g_c g| = \log|g_c| + \log|g|$$

$$\angle g_c g = \angle g_c + \angle g$$

Hence,

$$\log|g_c g| = \log|g_c| + \log|g| \leq \log|g|$$

$$\angle g_c g = \angle g_c + \angle g \leq \angle g$$

both gain margin (GM) and phase margin (PM) are reduced, and the closed-loop system becomes more oscillatory and potentially unstable.

In summary, the properties of *I* Regulation are:

1. The output of I control is constant only if $e=0$, no steady state error.

2. Reduce system stability. For the same process, I control is always slower than that of *P* control,
3. Open loop gain is proportional to K_I, increase K_I will reduce system stability till the system unstable.

4.2.4 Derivative Term

Derivative Function: The output of controller is proportional to the differential of error signal

$$u = K_D \frac{de}{dt} \tag{4.2.6}$$

The controller can adjust the output according to speed of error to perform the function of Prediction.

D control cannot work by itself. Any controller exists a dead zone; it takes no action when the rate of change is very small. This will accumulate the error. *D* function must be combined with other control functions to form PD or PID controllers.

PD Control

$$u = K_p e + K_D \frac{de}{dt} = K_p \left(e + \tau_D \frac{de}{dt} \right) \tag{4.2.7}$$

Derivative action is useful for providing phase lead, which offsets phase lag caused by integration. This action is also helpful in hastening loop recovery from disturbances. However, the derivative term is often misunderstood and misused. For example, it has been widely perceived in the control community that the derivative term improves transient performance and stability. But this perception is not always valid. τ_D must be carefully determined. Increase τ_D will increase system stability, however, the system becomes unstable when τ_D value over a certain limit.

To see this, note that adding a derivative term to a pure proportional term reduces the phase lag by

$$\angle(1 + j\omega\tau_D) = \tan^{-1}(\omega\tau_D) \in [0, 2\pi] \tag{4.2.8}$$

which tends to increase the PM.

However, the gain increases by a factor of

$$|1 + j\omega\tau_D| = \sqrt{1 + \omega^2\tau_D^2} > 1 \text{ for all } \omega \tag{4.2.9}$$

GM is uncertain and hence the overall stability may be improved or degraded.

To demonstrate that adding a differentiator can destabilize some systems, consider the typical first-order delayed plant

$$g(s) = \frac{Ke^{-Ls}}{1 + Ts} \tag{4.2.10}$$

where K is the process gain, T is the time constant, and L is the dead time or transport delay. Suppose that this plant is controlled by a proportional controller with gain K_P and that a derivative term is added. The resulting PD controller

$$g_{PD}(s) = K_P(1+\tau_D s) \tag{4.2.11}$$

leads to an open-loop feedforward path transfer function with frequency response

$$g(s)g_{PD}(s) = KK_P \frac{(1+\tau_D s)}{(1+Ts)} e^{-Ls} \tag{4.2.12}$$

For all ω, the gain satisfies

$$KK_P \sqrt{\frac{(1+\tau_D{}^2\omega^2)}{(1+T^2\omega^2)}} \geq KK_P \min\left(1, \frac{\tau_D}{T}\right) \tag{4.2.13}$$

where inequality (4.2.13) holds since

$$\sqrt{\frac{(1+\tau_D{}^2\omega^2)}{(1+T^2\omega^2)}}$$

is monotonic in ω. Hence, if $K_P > 1/K$ and $\tau_D > T/KK_p$, then, for all ω,

$$\left|g(j\omega)g_{PD}(j\omega)\right| > 1 \tag{4.2.14}$$

Inequality (4.2.14) implies that the *0-dB* gain crossover frequency is at infinity. Furthermore, due to the transport delay, the phase is

$$\angle g(j\omega)g_{PD}(j\omega) = \tan^{-1}\left(\omega\tau_D\right) - L\omega$$

Therefore, when ω approaches infinity,

$$\angle g(j\omega)g_{PD}(j\omega) < -180^0 \tag{4.2.15}$$

Hence, if $\tau_D > T/KK_P$ and $K_P > 1/K$, then by the stability criterion, the closed-loop system is unstable.

This analysis also confirms that some PID mapping formulas, as you will see later, such as the Ziegler-Nichol (Z-N) formula obtained from the step-response method, in which $K_P = (1.2(T/L))(1/K)$ and τ_D is proportional to L, are valid for only a limited range of values of the T/L ratio.

In summary, the properties of PD controllers are:

1. In steady state, *de/dt=0,* PD control has steady state error.
2. D function reduces oscillation, therefore, increases system stability.
3. Adding D function can increase open loop gain, therefore, increase response speed.
4. Sensitive to disturbance.
5. Care needs to be taken to tune and use the derivative term properly when the plant is

subject to delay (may cause system unstable).

Example 4.2.3: consider process (4.2.10) with K =10, T =1s, and L =0.1s. Control by means of a PI controller with K_P=0.644 and $\tau_I = 1.03$ s.

The PI controller yields reasonable stability margins and time-domain performance, as seen in Figures 4.2.5 and 4.2.6 red curve.

However, when a differentiator is added, gradually increasing τ_D from zero improves both GM and PM.

- The GM peaks when τ_D approaches 0.0303s; this value of τ_D maximizes the speed of the transient response without oscillation (Figures 4.2.5 and 4.2.6 blue curve).
- If τ_D is increased further to 0.1 s, the GM deteriorates and the transient exhibits oscillation (Figures 4.2.5 and 4.2.6 yellow curve).
- If τ_D increases to 0.2, the overall open-loop gain becomes greater than 2.2 dB for all ω. At ω = 30 rad/s, the phase decreases to −π while the gain remains above 2.2 dB. Hence, by the stability criterion, the closed-loop system is unstable (Figures 4.2.5 and 4.2.6 cyan curve).

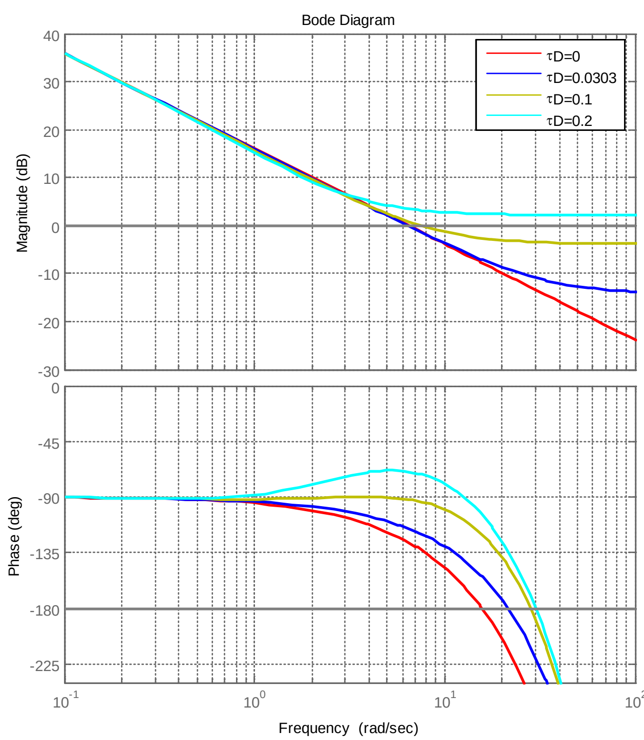
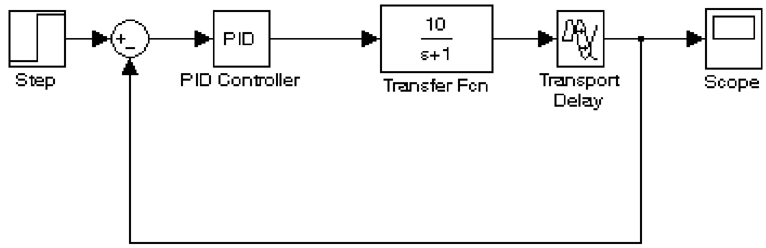


Fig. 4.2.5 Effect of the derivative term measured by GM and PM



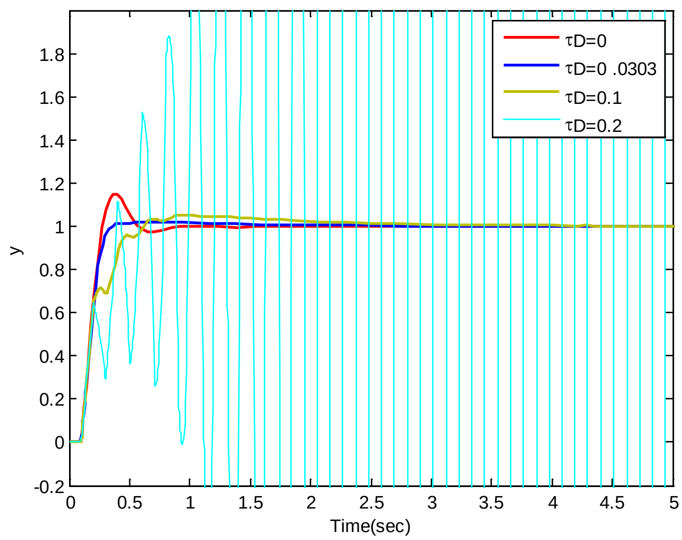


Fig. 4.2.6 Effect of the derivative term confirmed by the closed-loop step response.

4.2.5 PID Design Objectives and Methods

The design of the most appropriate controller for a given application depends on several factors, such as (1) the process characteristics, (2) how much variation from the set point is acceptable and under what conditions (such as start-up, running, idling), and (3) selecting the optimum controller type and tuning it properly. The process control characteristics are very important criteria, and experience shows that for easier controller tuning and lowest initial cost, the simplest controller that will meet requirements is usually the best choice. The design methods can be classified as:

Heuristic Methods: Heuristic methods evolve from empirical tuning (such as the ZN tuning rule), often with a tradeoff among design objectives. Heuristic search now involves expert systems, fuzzy logic, neural networks, and evolutionary computation.

Numerical Optimization Methods: Optimization-based methods can be regarded as a special type of optimal control. PID parameters are obtained by numerical optimization for a weighted objective in the time domain. Alternatively, a self-learning evolutionary algorithm (EA) can be used to search for both the parameters and their associated structure or to meet multiple design objectives in both the time and frequency domains.

Frequency Response Methods: Frequency-domain constraints, such as gain margin (GM), phase margin (PM), and sensitivities, are used to synthesize PID controllers offline. For real-time applications, frequency-domain measurements require time-frequency, localization-based methods such as wavelets.

Analytical Methods: Because of the simplicity of PID control, parameters can be derived analytically using algebraic relations between a plant model and a targeted closed-loop transfer function with an indirect performance objective, such as pole placement, internal model control (IMC), or lambda tuning. To derive a rational, closed-loop transfer function, this method requires that transport delays be replaced by Padé approximations.

4.3 Direct Synthesis Control

4.3.1 Basic Concepts

The purpose of a controller is to "shape" the response of the closed loop system. The response depends on the poles of the system.

From Fig. 4.3.1, it follows that,

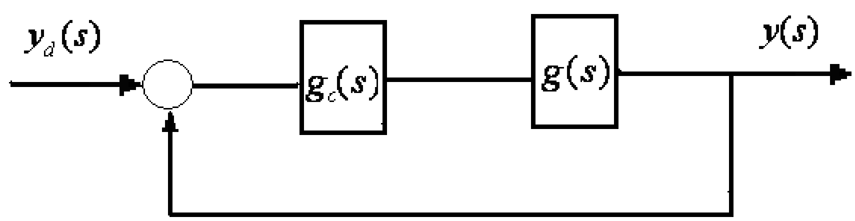


Fig. 4.3.1 Feedback Control system

$$y(s) = \frac{g(s)g_c(s)}{1+g(s)g_c(s)} y_d(s) \tag{4.3.1}$$

Suppose we require that

$$y(s) = q(s) y_d(s) \tag{4.3.2}$$

where *q(s)* is a specific pre-determined transfer function of our choice. The choice of the reference trajectory *q(s)* depends on the type of closed loop response desired. Then, the controller synthesis problem is now posed as follows:

> *What is the form of the controller* $g_c(s)$ *required to produce in the process the closed-loop behavior represented by the reference trajectory q(s)?*

Comparing (4.3.1) and (4.3.2),

$$q(s) = \frac{g(s)g_c(s)}{1+g(s)g_c(s)} \tag{4.3.3}$$

$$\Rightarrow g_c(s) = \frac{1}{g(s)}\left(\frac{q}{1-q}\right) \tag{4.3.4}$$

The controller parameter can be explicitly obtained from *q(s)* and *g(s).*

4.3.2 Synthesis of Minimum Phase Systems

Let the desired closed loop behaviour be

$$q(s)=\frac{1}{\tau_r s+1} \tag{4.3.5}$$

Substituting (4.3.5) into (4.3.4),

$$g_c(s)=\frac{1}{g(s)}\left(\frac{\frac{1}{\tau_r s+1}}{1-\frac{1}{\tau_r s+1}}\right)=\frac{1}{\tau_r s g(s)} \tag{4.3.6}$$

Once the process model is available, the controller can be easily obtained. It can also be shown that even when the actual plant dynamics are not matched exactly by the model transfer function *g(s)*, provided the overall closed-loop system is stable, this direct synthesis controller does not result in steady-state offset.

Before investigating the various types of controllers prescribed by Eq. (4.3.4) for various types of process models, it is important to note that this direct synthesis controller involves the inverse of the process model. This feature, common to all model-based controllers, will have significant implications later on, particularly when we consider the issue of synthesis for nonminimum phase systems.

Pure Gain Process: Let

$$g(s)=K$$

$$g_c(s)=\frac{1}{\tau_r s g(s)}=\frac{1}{K\tau_r s}$$

The controller is a pure Integral controller.

Pure Capacity Process: Let

$$g(s)=\frac{K}{s}$$

Then,

$$g_c(s)=\frac{1}{\tau_r s g(s)}=\frac{1}{K\tau_r}$$

The controller is a pure proportional controller.

First Order Process: Let

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

Then using (4.3.4) and (4.3.5),

$$g_c(s) = \frac{1}{\tau_r s g(s)} = \frac{\tau s + 1}{K\tau_r s} = \frac{\tau}{K\tau_r}\left(1 + \frac{1}{\tau s}\right)$$

The controller is a PI controller with

$$K_c = \tau / K\tau_r; \quad \tau_I = \tau.$$

Example: Direct Synthesis Controller for First Order System

Design a controller for the following first-order system

$$g(s) = \frac{0.66}{6.7s + 1}$$

using the direct synthesis approach, and given that the desirel closed loop behavior as in Eq. (4.3.5), with $\tau_r = 5$. Compare this controller with that $\tau_r = 1$.

Solution: The required direct synthesis controller (for $\tau_r = 5$) is a PI controller with K_c = 2.03 and $\tau_I = 6.7$.

$$g_c(s) = 2.03\left(1 + \frac{1}{6.7s}\right)$$

The controller obtained with the choice $\tau_r = 1$ (the faster closed-loop trajectory) is

$$g_c(s) = 10.15\left(1 + \frac{1}{6.7s}\right)$$

and the only difference between these controllers is that the latter has a proportional gain value which is *five times* that of the former; the integral times are identical. As we would expect, a controller with a higher gain value is required to provide the faster closed-loop response indicated by the smaller value chosen for τ_r. Further, it should be noted that a reduction in the value of τ_r by a factor of 5 resulted in a fivefold increase in the value of K_c .

Second order process: Let

$$g(s) = \frac{K}{\tau^2 s^2 + 2\varsigma s + 1}$$

The direct synthesis controller , using (4.3.4) and (4.3.5), is then given by

$$g_c(s) = \frac{1}{\tau_r s g(s)} = \frac{\tau^2 s^2 + 2\varsigma\tau s + 1}{K\tau_r s} \tag{4.3.7a}$$

or,

$$g_c(s) = \left(\frac{2\varsigma\tau}{K\tau_r}\right)\left(1 + \frac{1}{2\varsigma\tau s} + \frac{\tau s}{2\varsigma}\right) \tag{4.3.7b}$$

This is a PID controller.

Example: Suppose

$$g(s)=\frac{1}{(2s+1)(5s+1)}$$

and

$$q(s)=\frac{1}{5s+1}$$

Using (4.3.6)

$$g_c(s)=\frac{1}{5sg(s)}=\frac{(2s+1)(5s+1)}{5s}$$

or

$$g_c(s)=1.4(1+\frac{1}{7s}+1.43s)$$

PID controller with $K_c=1.4$, $\tau_I=7$ and $\tau_D=1.43$.

Now, suppose the rise time in the desired closed loop response is 5 times faster, i.e.,

$$q(s)=\frac{1}{s+1}$$

Then , using (4.3.6),

$$\begin{aligned} g_c(s)&=\frac{1}{sg(s)}=\frac{(2s+1)(5s+1)}{s} \\ &=7\left(1+\frac{1}{7s}+1.43s\right) \end{aligned}$$

a PID controller with $K_c=7, \tau_I=7$ and $\tau_D=1.43$. A reduction in the value of τ_r by a factor of 5 also resulted in a fivefold increase in the value of K_c

4.3.3 Synthesis of Time Delay Systems

Let the time delay process be represented by

$$y(s)=g(s)u(s)=\frac{Ke^{-\alpha s}}{\tau s+1}u(s)\,, \qquad \alpha>0 \tag{4.3.8}$$

and q is as in (4.3.5). However, direct using (4.3.4),

$$g_c(s)=\frac{1}{\tau_r sg(s)}=\frac{(\tau s+1)e^{\alpha s}}{K\tau_r s}$$

the controller has to predict α units of time ahead of response which is not realizable. The desired closed response is hence modified to take into account the time delay.

$$q(s)=\frac{e^{-\alpha_r s}}{\tau_r s+1} \tag{4.3.9}$$

Substituting (4.3.6) and (4.3.9) into (4.3.4), we get

$$g_c(s)=\frac{(\tau s+1)}{K}\left[\frac{e^{-(\alpha_r-\alpha)s}}{\tau_r s+1-e^{-\alpha_r s}}\right] \tag{4.3.10}$$

Putting $\alpha=\alpha_r$, we get ,

$$g_c(s)=\frac{(\tau s+1)}{K(\tau_r s+1-e^{-\alpha s})} \tag{4.3.11}$$

which results

$$u(s)=\frac{\tau s+1}{K(\tau_r s+1-e^{-\alpha s})}\varepsilon(s) \tag{4.3.12}$$

Rearranging (4.3.12), we obtain,

$$\frac{K\tau_r s}{\tau s+1}u(s)+\frac{K}{\tau s+1}u(s)=\frac{Ke^{-\alpha s}}{\tau s+1}u(s)+\varepsilon(s) \tag{4.3.13}$$

Let

$$y^*(s)=\frac{K}{\tau s+1}u(s)=g^*u(s) \tag{4.3.14}$$

where

$$g^*(s)=\frac{K}{\tau s+1}$$

Substitute (4.3.8) and (4.3.14) into (4.3.13) to get

$$\frac{K\tau_r s}{\tau s+1}u(s)+y^*(s)=y(s)+\varepsilon(s)$$

That is,

$$\begin{aligned}u(s)&=\frac{\tau s+1}{K\tau_r s}\left[\varepsilon(s)-(y^*-y)\right]\\&=K_c\left(1+\frac{1}{\tau_I s}\right)\left[\varepsilon(s)-\left(y^*(s)-y(s)\right)\right]\end{aligned}$$

Furthermore,

$$u(s)=\left(\frac{\tau}{K\tau_r}\right)\left(1+\frac{1}{\tau s}\right)\left[\varepsilon(s)-\left(y^*(s)-y(s)\right)\right] \tag{4.3.15}$$

where, $K_c=\tau/(K\tau_r)$ and $\tau_I=\tau$ are the parameters of a PI Controller.

Equation (4.3.16) corresponds to the Smith's Predictor Controller we will study later on in Chapter 7 in detail which is represented in block diagram form in Fig.4.3.2.

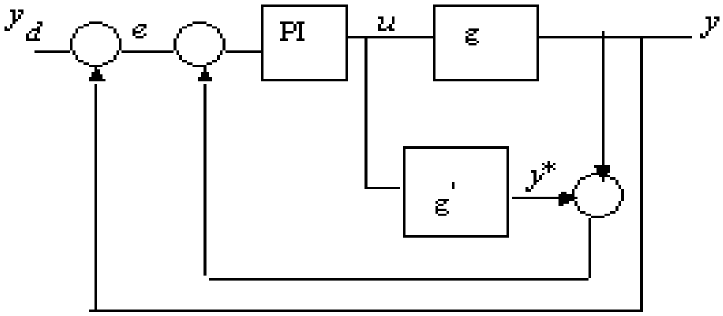


Fig. 4.3.2 Smith Predictor Controller

Approximation of Time Delay: Put

$$e^{-\alpha s} \cong 1-\alpha s$$

The controller given by (4.3.11) is expressed as

$$g_c(s) = \frac{(\tau s+1)}{K(\tau_r s+1-e^{-\alpha s})} = \frac{\tau}{K(\tau_r+\alpha)}\left(1+\frac{1}{\tau s}\right) \qquad (4.3.16)$$

a PI Controller.

Pade Approximation of Time Delay

Put

$$e^{-\alpha s} \cong \frac{1-\dfrac{\alpha s}{2}}{1+\dfrac{\alpha s}{2}}$$

Then, (10) becomes, after simplification,

$$g_c(s) = \frac{(\tau s+1)}{K(\tau_r s+1-e^{-\alpha s})} = \left(\frac{\tau}{K(\tau_r+\alpha)}\right)\left(1+\frac{1}{\tau s}\right)\left(\frac{1+\dfrac{\alpha s}{2}}{1+\tau^* s}\right) \qquad (4.3.17)$$

where

$$\tau^* = \frac{\alpha\tau_r}{2(\tau_r+\alpha)}$$

(4.3.17) corresponds to a PID controller. The proportional gain

$$K_c = \frac{\tau}{K(\tau_r+\alpha)}$$

and the Integral time $\tau_I = \tau$. The derivative part is the last product term in (4.3.17) which corresponds to a gain limited derivative controller.

Example: Let

$$g(s) = \frac{0.66e^{-2.6s}}{6.7s+1}$$

and

$$q(s) = \frac{e^{-2.6s}}{5s+1}$$

Then,

$$g_c(s) = \frac{q(s)}{g(s)\left[1-q(s)\right]} = \frac{6.7s+1}{0.66(5s+1-e^{-2.6s})}$$

Put

$$e^{-2.6s} \cong 1-2.6s$$

and simplify to get

$$g_c(s) = 1.34\left(1+\frac{1}{6.7s}\right)$$

a PI Controller with $K_c=1.34$ and $\tau_I=6.7$.

Alternatively, using Pade approximation,

$$e^{-2.6s} = \frac{1-1.3s}{1+1.3s}$$

Then, g_c can be written as

$$g_c(s) = \left(\frac{6.7}{0.66}\right)\left(1+\frac{1}{6.7s}\right)\frac{(1+1.3s)}{(7.6+6.5s)}$$

That is,

$$g_c(s) = (1.59)\left[1+\frac{1}{8s}+1.09s\right]\left(\frac{1}{1+0.86s}\right)$$

PID contoller with $K_c=1.59$, $\tau_I=8$ and $\tau_D=1.09$. The third product term corresponds to a low pass filter.

4.3.4 Synthesis of Inverse Response Systems:

Let

$$g(s) = \frac{K(1-\eta s)}{(1+\tau_1 s)(1+\tau_2 s)}$$

The desired closed loop response will be

$$q(s) = \frac{(1-\eta s)}{(1+\tau_{r1}s)(1+\tau_{r2}s)}$$

Then the direct synthesis controller is given by

$$g_c(s) = \frac{1}{g(s)}\left(\frac{q(s)}{1-q(s)}\right) = \frac{(1+\tau_1 s)}{K(\eta+\tau_{r1}+\tau_{r2})s}\frac{(1+\tau_2 s)}{(1+\tau^* s)}$$

where

$$\tau^* = \frac{\tau_{r1}\tau_{r2}}{\eta+\tau_{r1}+\tau_{r2}}$$

The controller can also be put as

$$g_c(s) = K_c\left[1+\frac{1}{\tau_I s}+\tau_D s\right]\frac{1}{(1+\tau^* s)} \tag{4.3.18}$$

where

$$K_c = \frac{\tau_1+\tau_2}{K(\eta+\tau_{r1}+\tau_{r2})}$$

$$\tau_I = \tau_1+\tau_2$$

and

$$\tau_D = \frac{\tau_1\tau_2}{\tau_1+\tau_2}$$

The third product term on the RHS of (4.3.18) coresponds to a first order filter term.

Example: Let

$$g(s) = \frac{1-3s}{(2s+1)(5s+1)}$$

and

$$q(s) = \frac{(1-3s)}{(1+s)(1+3s)}$$

Using (4), the controller is given, after simplification, as

$$g_c(s) = \left(1+\frac{1}{7s}+1.43s\right)\left(\frac{1}{1+0.43s}\right)$$

a PID controller in cascade with a filter.

4.3.5 Synthesis for Open-Loop Unstable Systems

Consider the first-order open-loop unstable system whose transfer function model is given as:

$$g(s) = \frac{K}{\tau s - 1} \tag{4.3.19}$$

the direct synthesis controller required for (4.3.6) , resulting

$$g_c(s) = \frac{1}{\tau_r s g(s)} = \frac{\tau}{K\tau_r}\left(1 - \frac{1}{\tau s}\right) \tag{4.3.20}$$

which looks like a PI controller with $K_c = \tau / K\tau_r$ but with a very important difference that its integral time is *negative;* i.e., $\tau_I = -\tau$.

The problem with this controller is that it will only function properly if all the process parameters are exactly equal to the model parameters; if this is not the case the overall closed-loop system will be unstable. The controller therefore has a severe robust stability problem.

Consider the situation in which the actual open-loop unstable process has the transfer function:

$$g_p(s) = \frac{K_p}{\tau_p s - 1} \tag{4.3.21}$$

for which (4.3.19) is now only an approximate model. When this system is operating in the closed loop with the controller in Eq. (4.3.20) derived on the basis of the approximate model of Eq. (4.3.19), the closed-loop characteristic equation:

$$1 + g_p g_c = 0$$

becomes

$$K\tau_r\tau_p s^2 + (K_p\tau - K\tau_r)s - K_p = 0$$

where $K_p \neq K$ and $\tau_p \neq \tau$. Because of negative coefficient of K_p , the system characteristic equation will always be unstable.

The direct synthesis philosophy, which requiring the closed-loop response to be as in Eq. (4.3.5), is responsible for this problem. According to Eq. (4.3.5), the single, closed-loop system pole is to be located at $s = -1/\tau_r$ to achieve this requires a controller that will do two things simultaneously:

- cancel the RHP pole;
- replace it with the desired LHP pole (the one located at $s = -1/\tau_r$).

Such perfect pole cancellation is possible only when the location of RHP pole is *exactly* known.

4.3.6 Observations on Direct Synthesis Method

The discussion above would allow us to derive the direct synthesis tuning parameters for any class of models and choice of reference trajectories. These formulas provide an excellent set of tuning parameters for a wide range of problems when an adequate linear model is available for the process. However

- PID parameters will be decided by a user-specified parameter: The desired closed-loop time constant (τ_r)
    - The shorter makes the action more aggressive. (larger K_c)
    - The longer makes the action more conservative. (smaller K_c)
- For limited cases, it results PID form.
    - 1st-order model without time delay: PI
    - FOPDT with 1st-order Taylor series approx.: PI
    - 2nd-order model without time delay: PID
    - SOPDT with 1st-order Taylor series approx.: PID
- The resulting controller from direct synthesis method could be quite complex and may not even be physically realizable.
- If there is RHP zero in the process, the resulting controller from direct synthesis method will be unstable.
- Unmeasured disturbance and modeling error are not considered in direct synthesis method.

These problems will be addressed in the next section by “Internal Model Control Strategy”.

4.4 Internal Model Control

4.4.1 Motivation

Consider the process whose dynamic behavior is represented by:

$$y(s) = g(s)u(s) + d(s) \tag{4.4.1}$$

with the block diagram shown in Figure 4.4.1 Here d represents the collective effect of *unmeasured* disturbances on the process output *y(s)*. If it is desired to have "perfect" control in which the output tracks the desired set-point $y_d(s)$ *perfectly, the control action required to achieve this objective is easily obtained by substituting* $y(s) = r(s)$ in Eq. (4.4.1):

$$r(s) = g(s)u(s) + d(s) \tag{4.4.2}$$

and then solving for *u(s)* to obtain:

$$u(s) = \frac{1}{g(s)}\left[r(s) - d(s)\right] \tag{4.4.3}$$

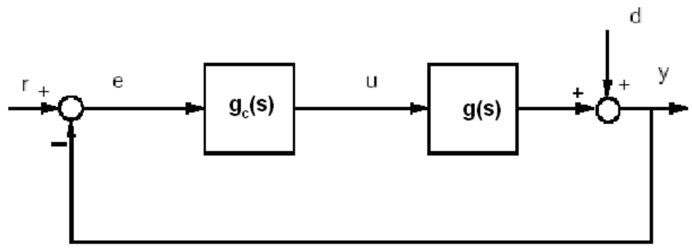


Figure 4.4.1 Conventional Feedback Control Structure

The implication is that if both *d(s)* and *g(s)* are known, then for any given $r(s)$, Eq. (4.4.3) provides the controller that will achieve perfect control. Since in reality, *d(s)* is unmeasured and *g(s)* be only modeled approximately by, say, $\tilde{g}(s)$, we may adopt the following strategy:

1. Assuming that $\tilde{g}(s)$ is our best estimate of the plant dynamics *g(s)*, then our best estimate of *d(s)* is obtained by subtracting the model prediction, $\tilde{g}(s)u(s)$, from the actual plant output *y(s)* to yield the estimate:

$$\hat{d}(s) = y(s) - \tilde{g}(s)u(s) \tag{4.4.4}$$

2. Let us choose the notation:

$$u(s) = \frac{1}{\tilde{g}(s)}\left[r(s) - \hat{d}(s)\right] \tag{4.4.5}$$

where $\hat{d}(s)$ is the estimate of *d* given by Eq. (4.4.4).

A block diagrammatic representation of Eqs. (4.4.4) and (4.4.5) takes the form shown in Figure 4.4.2 known as the "Internal Model Control" structure (IMC) which forms the basis for the systematic control system design methodology. Figure 4.4.2 is the “Internal Model Control” or “Q-parametrization” structure.

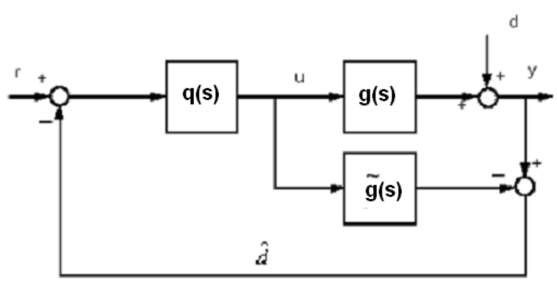


Figure 4.4.2: Internal Model Control Structure

The IMC structure and the classical feedback structure (Figure 4.4.1) are equivalent representations. Having designed $q(s)$, its equivalent classical feedback controller $g_c(s)$ can be readily obtained via algebraic transformations, and vice-versa

$$g_c(s) = \frac{q(s)}{1 - \tilde{g}(s)q(s)} \tag{4.4.6}$$

$$q(s)=\frac{g_c(s)}{1+\tilde{g}(s)g_c(s)} \tag{4.4.7}$$

The equivalence between the two are shown as in Fig 4.4.3.

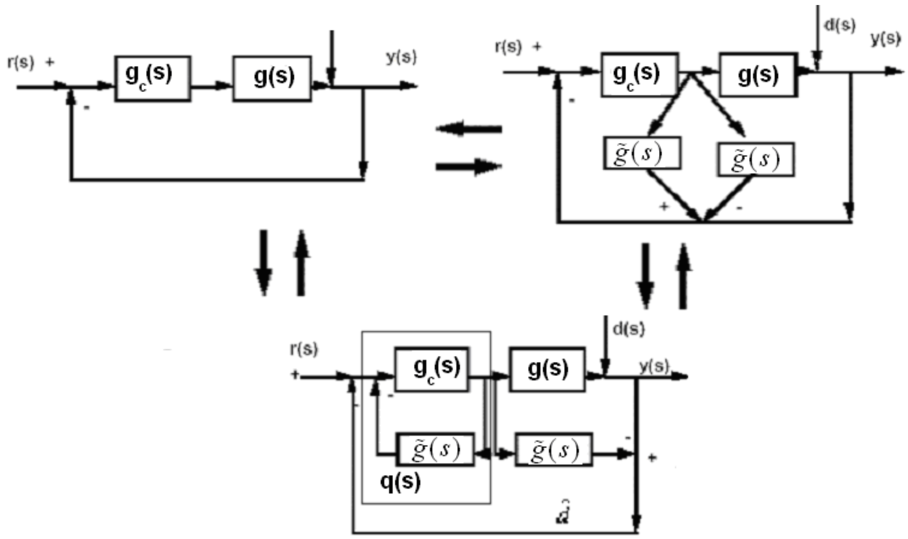


Figure 4.4.3 Evolution of the Internal Model Control

Closed-loop transfer functions

The closed loop transfer function of Fig. 4.4.2 is given by:

$$\begin{aligned} y(\mathrm{s}) &= \frac{g(s)q(s)}{1\ +\ q(s)(g(s)-\tilde{g}(s))}r(\mathrm{s})+\frac{1-\tilde{g}(s)q(s)}{1\ +\ q(s)(g(s)-\tilde{g}(s))}d(s) \\ &= \eta(s)r(s)+\varepsilon(s)d(s) \end{aligned} \tag{4.4.8}$$

where, in terms of the internal model $\tilde{g}(s)$ and controller $q(s)$, $\varepsilon(s)$ and *$\eta(s)$* are defined as sensitivity and complementary sensitivity, respectively. If no plant/model mismatch ($g(s)=\tilde{g}(s)$), they are simplified to

$$\begin{aligned} \tilde{\eta}(s) &= \tilde{g}(s)q(s) \Rightarrow q(s)=\tilde{g}^{-1}(s)\tilde{\eta}(s); \\ \tilde{\varepsilon}(s) &= 1-\tilde{\eta}(s)=1-\tilde{g}(s)q(s) \end{aligned} \tag{4.4.9}$$

Internal Stability

1. Assume a perfect model ($g(s)=\tilde{g}(s)$). The IMC system (Figure 4.4.2) is internally stable (IS) if and only if both $g(s)$ and *$q(s)$* are stable.
2. Assume that p is stable and $g(s)=\tilde{g}(s)$. Then the classical feedback system (Figure 4.4.1) with controller according to Equation (4.4.6) is IS if and only if *$q(s)$* is stable.

Asymptotic closed-loop behavior

We need to insure that the feedback control system leads to no offset for setpoint or disturbance changes; we thus need to define so-called Type 1 and Type 2 inputs:

Type 1 (Step Inputs): No offset to asymptotically step setpoint/disturbance changes is obtained if

$$\lim_{s\to 0}\tilde{g}(s)q(s)=\tilde{\eta}(0)=1 \Rightarrow q(0)=\frac{1}{\tilde{g}(0)} \tag{4.4.10}$$

Then regardless of the fact that $g(s)\neq\tilde{g}(s)$, (4.4.8) shows that $y(s)=r(s)$ at steady state. Thus, so long as the steady state gain of the controller is the same as the reciprocal of the gain of the process model.

Type 2 (Ramp Inputs): For no offset to ramp inputs, it is required that

$$\begin{aligned}&\lim_{s\to 0}\tilde{g}(s)q(s)=\tilde{\eta}(0)=1\\ &\lim_{s\to 0}\frac{d}{ds}(\tilde{g}(s)q(s))=\left.\frac{d\tilde{\eta}}{ds}\right|_{s=0}=0\end{aligned} \tag{4.4.11}$$

***Requirements for Physical Realizability on** q(s)*

In order for *q(s)*, the IMC controller, to result in physically realizable manipulated variable responses, it must satisfy the following criteria:

1. *Stability:* The controller must generate bounded responses to bounded inputs; therefore all poles of *q(s)* must lie in the open Left-Half Plane.
2. *Properness:* In order to avoid pure differentiation of signals, we must require that $q(s)$ be proper, which means that the quantity

   $$\lim_{|s|\to\infty} q(s) \tag{4.4.12a}$$

   must be finite. We say $q(s)$ is strictly proper if

   $$\lim_{|s|\to\infty} |q(s)|=0 \tag{4.4.12b}$$

   A strictly proper transfer function has a denominator order greater than the numerator order.

   $q(s)$ is semi-proper, that is,

   $$\lim_{|s|\to\infty} |q(s)|\geq 0 \tag{4.4.12c}$$

   if the denominator order is equal to the numerator order.

   A system that is neither strictly proper nor semiproper is called improper.
3. *Causality*: $q(s)$ must be causal, which means that the controller must not require prediction, i.e., it must rely on *current* and *previous* plant measurements. A simple example of a noncausal transfer function is the *inverse* of a time delay transfer function

   $$q(s) = \frac{u(s)}{e(s)} = K_c e^{+\theta s} \tag{4.4.13}$$

   The inverse transform of (4.4.13) relies on *future* inputs to generate a *current* output; it is clearly not realizable:

   $$u(t)=K_c e(t+\theta) \tag{4.4.14}$$

Advantages:

The IMC structure offers the following benefits with respect to classical feedback:

- No need to solve for roots of the characteristic polynomial $1+g_c(s)g(s)$; one simply examines the poles of *q(s)*;
- One can search for *q(s)* instead of $g_c(s)$ without any loss of generality.

4.4.2 Internal Model Control Design Procedure

The IMC design procedure is a two-step approach that, although sub-optimal in a general sense, provides a reasonable tradeoff between performance and robustness. The main benefit of the IMC approach is the ability to directly specify the complementary sensitivity and sensitivity functions *η(s)* and $\varepsilon(s)$, which as noted previously, directly specify the nature of the closed-loop response.

The IMC design procedure consists of two main steps. The frist step will insure that *q(s)* is stable and causal; the second step will require *q(s)* to be proper.

Step1: Factor the model $\tilde{g}(s)$ into two parts:

$$\tilde{g}(s)=\tilde{g}_+(s)\tilde{g}_-(s) \tag{4.4.15}$$

$\tilde{g}_+(s)$ contains all **Nonminimum Phase Elements** in the plant model, that is all Right-Half-Plane (RHP) zeros and time delays. The factor $\tilde{g}_-(s)$, meanwhile, is **Minimum Phase** and invertible; an IMC controller defined as

$$\tilde{q}(s)=\tilde{g}_-^{-1}(s) \tag{4.4.16}$$

is *stable* and *causal*.

The factorization of $\tilde{g}_+(s)$ from $\tilde{g}(s)$ is dependent upon the **objective function** chosen. For example,

$$\tilde{g}_+(s)= e^{-\theta s}\prod_i(-\beta_i s+1);\ \mathrm{Re}(\beta_i)\ >\ 0 \tag{4.4.17}$$

is Integral-Absolute-Error (IAE)-optimal for step setpoint and output disturbance changes. While, the factorization

$$\tilde{g}_+(s)= e^{-\theta s}\prod_i\frac{(-\beta_i s+1)}{(\beta_i s+1)};\ \mathrm{Re}(\beta_i)\ >\ 0 \tag{4.4.18}$$

is Integral-Square-Error (ISE)-optimal for step setpoint/output disturbance changes. Using ramp, exponential, or other inputs would imply different factorizations.

Step2: Augment *q(s)* with a filter *f*(*s*) such that the final IMC controller $q(s)=\tilde{q}(s)f(s)$ is, in addition to stable and causal, proper. With the inclusion of the filter transfer function, the final form for the closed-loop transfer functions characterizing the system is

$$\tilde{\eta}(s) = \tilde{g}(s)\tilde{q}(s)f(s)$$
$$\tilde{\varepsilon}(s) = 1 - \tilde{g}(s)\tilde{q}(s)f(s) \tag{4.4.19}$$

Step3: If necessary, the IMC controller may be converted to the conventional form for implementation by using (4.4.6)

The inclusion of the filter transfer function in here means that we no longer obtain "optimal control," as implied in Step 1. We wish to define filter forms that allow for no offset to Type 1 and Type 2 inputs; for no offset to step inputs (Type 1), we must require that $\tilde{\eta}(0) = 1$, which requires that $q(0) = \tilde{g}^{-1}(0)$ and forces

$$f(0) = 1 \tag{4.4.20}$$

A common filter choice that conforms to this requirement is

$$f(s) = \frac{1}{(\lambda s + 1)^n} \tag{4.4.21}$$

The filter order *n* is selected large enough to make *q(s)* proper, while *λ* is an *adjustable parameter* which determines the speed-of-response. Increasing *λ* increases the closed-loop time constant and slows the speed of response; decreasing *λ* does the opposite. *λ* can be be adjusted on-line to compensate for plant/model mismatch in the design of the control system; the higher the value of *λ*, the higher the robustness the control system.

The factorization of $\tilde{g}(s)$ is very important, recall that for classical feedback

$$y(s) = \eta(s)r(s) + \varepsilon(s)d(s)$$
$$\eta(s) = (1 + g(s)g_c(s))^{-1} g(s)g_c(s)$$
$$\varepsilon(s) = (1 + g(s)g_c(s))^{-1}$$

Using the IMC structure, for no plant/model mistmatch ($g(s) = \tilde{g}(s)$), we have

$$\tilde{\eta}(s) = \tilde{g}(s)q(s) \qquad \tilde{\varepsilon}(s) = 1 - \tilde{g}(s)q(s) \tag{4.4.22}$$

"Perfect" control (meaning *y* = *r* for all time) is achieved when $\tilde{\eta} = 1$ and $\tilde{\varepsilon} = 0$, which implies that

$$q(s) = \tilde{g}^{-1}(s) \tag{4.4.23}$$

However, in order for *u* = *q*(*r* - *d*), the manipulated variable response, to be physically realizable, *q* must be *stable, proper*, and *causal*. Nominimum phase behavior (deadtime and RHP zeros) will cause $q(s) = \tilde{g}^{-1}(s)$ to be noncausal and unstable, respectively; if $\tilde{g}(s)$ is strictly proper, then $q(s)$ will be improper as well.

Example: Consider the plant model

$$\tilde{g}(s) = \frac{K(-\beta s + 1)e^{-\theta s}}{\tau^2 s^2 + 2\zeta\tau s + 1}$$

where $\beta > 0$, which implies the presence of a Right-Half Plane zero. Nonminimum phase elemets for this plant are $(-$ 遱e$1)^{-\theta s}$. The "perfect" IMC controller for this system corrresponds to

$$q(s) = \tilde{g}^{-1}(s) = \frac{\tau^2 s^2 + 2\zeta\tau s + 1}{K(-\beta s + 1)} e^{+\theta s}$$

While $y = r$ using this controller, the manipulated variable response is physically unrealizable since 1) $q(s)$ is unstable as a result of a Right-Half Plane pole arising from $(-\beta s + 1)$; and 2) $q(s)$ is noncausal because of the presence of the time *lead* term $e^{+\theta s}$.

Applying an appropriate factorization to this model as described earlier results in stable, causal control action; a correctly chosen filter order will insure properness and a physically realizable response. One must keep in mind that the nonminimum phase elements $(-\beta s + 1)e^{-\theta s}$ will always form part of the closed-loop response!

4.4.4 Application of IMC Design to PID controller tuning

The IMC control design procedure, when applied to low-order models, will often result in PID and PID-like controllers. Developing these is the focus of this section:

PI Control for the First Order Model

A PI tuning rule arises from applying IMC to the first-order model:

$$\tilde{g}(s) = \frac{K}{\tau s \ + \ 1} \qquad \tau > 0$$

under the condition that d and r are step input changes.

Step 1: Factor and invert $\tilde{g}(s)$; since $\tilde{g}_+(s) = 1$, we obtain:

$$\tilde{q}(s) = \frac{\tau s + 1}{K}; \quad \text{no proper}$$

Step 2: Augment with a first-order filter

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

The final form for $q(s)$ is

$$q(s) = \frac{\tau s + 1}{K(\lambda s + 1)}; \quad \text{proper}$$

Step 3: We can now solve for the classical feedback controller equivalent $g_c(s)$ to obtain

$$g_c(s) = \frac{q(s)}{1 - g(s)q(s)} = \frac{\tau}{K\lambda}(1 + \frac{1}{\tau s})$$

which leads to the tuning rule for a PI controller

$$K_c = \frac{\tau}{K\lambda} \qquad \tau_I = \tau$$

The corresponding *nominal* closed-loop transfer functions for this control system are

$$\tilde{\eta}(s) = \frac{1}{\lambda s + 1}; \qquad q(s) = \tilde{g}^{-1}(s)\eta(s) = \frac{\tau s + 1}{k(\lambda s + 1)}; \qquad \tilde{\varepsilon}(s) = \frac{\lambda s}{(\lambda s + 1)}$$

Example: IMC Design for First Order Process

Design a controller for the first-order process whose transfer function is

$$\tilde{g}(s) = \frac{5}{8s + 1}$$

using the IMC strategy. Convert this controller to the convenentional feedback form

Solution: Observing that the transfer function is invertible, we obtain:

$$\frac{1}{\tilde{g}(s)} = \frac{8s + 1}{5}$$

which requires only a first-order filter *(n* = 1) in order that $g_c(s) = f(s)/\tilde{g}(s)$ be proper, Thus we have, in this case:

$$q(s) = \frac{1}{\tilde{g}(s)} f(s) = \frac{1}{5}\frac{8s + 1}{\lambda s + 1}$$

which can be implemented using a lead/lag element. Furthermore,

$$g_c(s) = \frac{8}{5\lambda}\left(1 + \frac{1}{8s}\right)$$

as the equivalent conventional feedback form a PI controller whose gain depends on the filter parameter λ.

PI Control Inverse Response Process

Consider now the first-order model with Right Half Plane (RHP) zero:

$$\tilde{g}(s) = \frac{K(-\beta s + 1)}{\tau s + 1} \qquad \beta, \tau > 0$$

again under the assumption that the inputs to *r* and *d* are steps.

Step 1: Use the IAE-optimal factorization for step inputs:

$$\tilde{g}_+(s) = (-\beta s + 1); \qquad \tilde{g}_-(s) = \frac{K}{(\tau s + 1)}; \qquad \tilde{q}(s) = \frac{(\tau s + 1)}{K}$$

Step 2: Use a first-order filter

$$f = \frac{1}{(\lambda s + 1)}; \qquad q(s) = \frac{\tau s + 1}{K(\lambda s + 1)}$$

Step 3: Solving for the classical feedback controller leads to another tuning rule for a PI controller:

$$g_c(s) = K_c(1 + \frac{1}{\tau s})$$

$$K_c = \frac{\tau}{K(\beta + \lambda)} \qquad \tau_{\rm I} = \tau$$

PI with filter control

Consider now the first-order model with Left Half-Plane (LHP) zero:

$$\tilde{g}(s) = \frac{K(\beta s + 1)}{\tau s \ + \ 1} \qquad \beta \ , \tau > 0$$

again under the assumption that the inputs to *r* and *d* are steps.

Step 1: No nonminimum phase behavior in $\tilde{g}(s)$; since $\tilde{g}_+(s) = 1$, we obtain:

$$\tilde{g}_-(s) = \frac{K(\beta s + 1)}{(\tau s + 1)}; \qquad \tilde{q}(s) = \frac{(\tau s + 1)}{K(\beta s + 1)}$$

Step 2: Use a first-order filter (*q(s)* is now *strictly proper*).

$$f(s) = \frac{1}{(\lambda s + 1)}; \qquad q(s) = \frac{\tau s + 1}{K(\beta s + 1)(\lambda s + 1)}$$

Step 3: Solving for the classical feedback controller,

$$g_c(s) = \frac{q(s)}{1 - \tilde{g}(s)q(s)} = K_c(1 + \frac{1}{\tau s})\frac{1}{(\tau_F s + 1)}$$

$$K_c = \frac{\tau}{K\lambda}; \qquad \tau_{\rm I} = \tau; \qquad \tau_{\rm F} = \beta$$

It is interesting to note that in IMC design, the presence of a Left-Half Plane zero in the model leads a low-pass filter element in the classical feedback controller!

PID Control

Consider now the second-order model with RHP zero:

$$\tilde{g}(s) = \frac{K(-\beta s + 1)}{(\tau_1 s + 1)(\tau_2 s + 1)} \qquad \beta, \tau_1, \tau_2 > 0$$

again under the assumption that the inputs to *r* and *d* are steps.

Step 1: Use the IAE-optimal factorization for step inputs:

$$\tilde{g}_+(s) = (-\beta s + 1) \qquad \tilde{g}_-(s) = \frac{K}{(\tau_1 s + 1)(\tau_2 s + 1)}$$

$$\tilde{q}(s) = \frac{(\tau_1 s + 1)(\tau_2 s + 1)}{K}$$

Step 2: Use a first-order filter (even though this means that *q(s)* will still be improper).

$$f(s) = \frac{1}{(\lambda s + 1)}; \qquad q(s) = \frac{(\tau_1 s + 1)(\tau_2 s + 1)}{K(\lambda s + 1)}$$

Step 3: Solving for the classical feedback controller

$$g_c(s) = \frac{q(s)}{1 - \tilde{g}(s)q(s)} = K_c(1 + \frac{1}{\tau_I s} + \tau_D s)$$

an ideal PID controller with

$$K_c = \frac{\tau_1 + \tau_2}{K(\beta + \lambda)}; \qquad \tau_\mathrm{I} = \tau_1 + \tau_2; \qquad \tau_\mathrm{D} = \frac{\tau_1 \tau_2}{\tau_1 + \tau_2}$$

PID with Filter Control

Consider the second-order model with RHP zero again and subject to step inputs to the closed-loop system. Applying the IMC design procedure gives:

Step 1: Use the ISE-optimal factorization

$$\tilde{g}_+(s) = \frac{(-\beta s + 1)}{(\beta s + 1)}; \qquad \tilde{g}_-(s) = \frac{K(\beta s + 1)}{(\tau_1 s + 1)(\tau_2 s + 1)}$$

Step 2: A first-order filter leads to q which is semiproper:

$$q(s) = \frac{(\tau_1 s + 1)(\tau_2 s + 1)}{K(\beta s + 1)(\lambda s + 1)}$$

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

Step 3: Solving for $c(s)$ as before results in a filtered ideal PID controller

$$g_c(s) = K_c(1 + \frac{1}{\tau_I s} + \tau_D s)\frac{1}{(\tau_F s + 1)}$$

with the associated tuning rule

$$K_c = \frac{\tau_1 + \tau_2}{K(2\beta + \lambda)}; \qquad \tau_\mathrm{I} = \tau_1 + \tau_2; \qquad \tau_\mathrm{D} = \frac{\tau_1 \tau_2}{\tau_1 + \tau_2}; \qquad \tau_F = \frac{\beta\lambda}{(2\beta + \lambda)}$$

Note the insight given by IMC design procedure regarding on-line adjustment (by changing the value for the IMC filter parameter λ).

Deadtime compensation (PI controller + Smith Predictor)

Consider the first-order with delay plant

$$\tilde{g}(s) = \frac{Ke^{-\theta s}}{\tau s + 1} \qquad \tau > 0$$

and step setpoint/output disturbance changes to the closed-loop system.

Step 1: The optimal factorization (IAE, ISE, or otherwise) is $\tilde{g}_+(s) = e^{-\theta s}$, resulting in:

$$\tilde{q}(s) = \tilde{g}_{-}^{-1}(s) = \frac{\tau s + 1}{K}$$

Step 2: A first-order filter makes *q* semiproper;

$$q(s) = \frac{\tau s + 1}{K(\lambda s + 1)} \qquad \tilde{\eta}(s) = \frac{e^{-\theta s}}{(\lambda s + 1)};$$

Step 3: The corresponding feedback controller is

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

which can be expressed as a PI controller using the Smith Predictor structure (same as in direct synthesis control).

PID Tuning Rules for 1st-order with Deadtime Plants

The PID tuning rule for plants with deadtime arises from using a first-order Padè approximation for the time delay.

$$\begin{aligned}\tilde{g}(s) &= \frac{Ke^{-\theta s}}{\tau s \ + \ 1} \\ &\approx \frac{K(-\theta s/2 + 1)}{(\tau s + 1)(\theta s/2 + 1)}\end{aligned}$$

The Padè-approximated plant is a second-order plant with RHP zero; using the analysis (4.4.39) leads to a PID tuning rule:

$$K_c = \frac{2\tau + \theta}{K(\theta + 2\lambda)}; \qquad \tau_{\mathrm{I}} = \tau + \frac{\theta}{2}; \qquad \tau_{\mathrm{D}} = \frac{\tau\theta}{2\tau + \theta}$$

The function of λ/θ independent of τ, is a measure related to robustness of the closed-loop system. Note that at $\lambda/\theta \approx 0.8$ the IMC-PID controller results in an ISE value that is only 10% greater than optimal. The controlled variable response of the IMC-PID controller for various settings of λ/θ is shown in Figure 4.4.4.

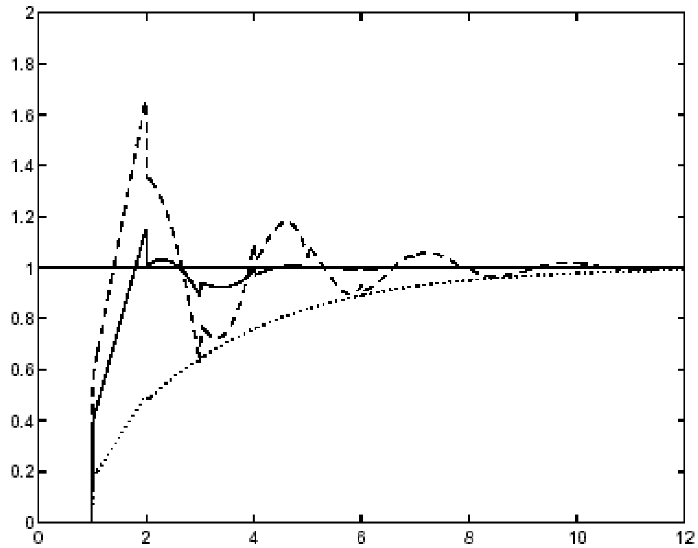


Figure 4.4.4: IMC-PID controlled variable responses for a step setpoint change, for various settings of λ/θ solid: $\lambda/\theta = 0.8$; dotted: $\lambda/\theta = 2.5$; dashed: $\lambda/\theta = 0.4$.

4.5 Gain and Phase Margin Method

4.5.1 General concepts

Let the process transfer function $G_p(s)$ be represented by

$$G_p(s) = \frac{b_{n-1}s^{n-1} + \cdots b_1 s + b_0}{a_n s^n + a_{n-1}s^{n-1} + \cdots a_1 s + 1} e^{-Ls} \tag{4.5.1}$$

In frequency domain, the Bode diagram is shown as in Figure 4.5.1, where ω_p is the phase crossover frequency at which the Bode diagram has a phase lag of $-\pi$, and ω_g is the gain crossover frequency at which the Bode diagram has a unity amplitude.

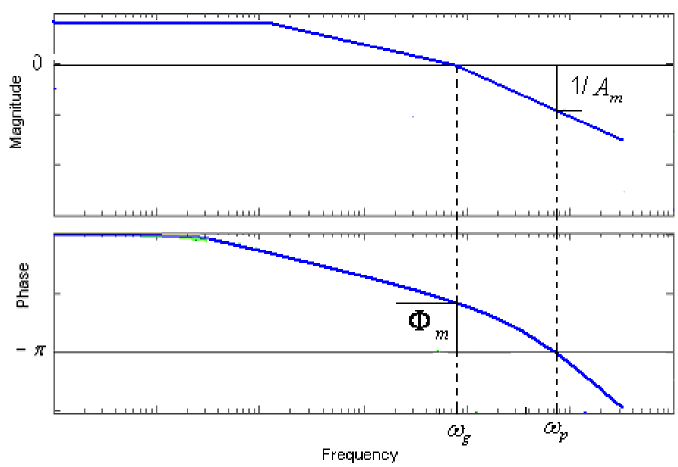


Figure 4.5.1 Bode Diagram

At these two frequency points, the gain and phases are defined as

$$\arg[G_c(j\omega_p)G_p(j\omega_p)] = -\pi \tag{4.5.2a}$$

$$A_m\left|G_c\left(j\omega_p\right)G_p\left(j\omega_p\right)\right| = 1 \tag{4.5.2b}$$

$$\left|G_c\left(j\omega_g\right)G_p\left(j\omega_g\right)\right| = 1 \tag{4.5.2c}$$

$$. \quad \Phi_m = \pi + \arg\left[G_c\left(j\omega_g\right)G_p\left(j\omega_g\right)\right] \tag{4.5.2d}$$

where A_m is the gain margin at phase crossover frequency ω_p and Φ_m is the phase margins at the gain crossover frequency ω_g. Now, let the system be controlled by a PID controller, $G_c(s)$, which is expressed by

$$G_c(s) = K_p + \frac{K_i}{s} + K_d s\,. \tag{4.5.3}$$

To ensure stability robustness of the feedback control system, the PID controller $G_c(s)$ is often so designed such that $G_c(s)G_p(s)$ meet the specified gain and phase margins, A_m and Φ_m at the ω_p and ω_g frequency points, respectively, which also shown in Figure 4.5.1. In control engineering practice, the gain and phase margins are normally specified to be between 2 to 5 and $\pi/6$ to $\pi/3\ rad$, respectively, for good system performance. Those margins are generally regarded as making the best compromise between performance and robustness in process control. For a given plant transfer function $G_p(s)$, A_m and Φ_m, the four designing equations of (4.5.2a) to (4.5.2d) contain five unknown parameters, i.e., K_p; K_i; K_d; ω_p, and ω_g. If one parameter is specified, the PID controller design problem is reduced to that of solving the four nonlinear algebraic equations for the remaining four unknown parameters. In other words, there is one degree of freedom among these parameters that must be specified in order to obtain a unique PID controller.

4.5.2 Simplified Design

To simplify the design procedure, we assume that the transfer function can be represented by a *second order plus dead time* (SOPDT) model, which can be used to describe most of the industrial processes:

$$G_p(s) = \frac{b_0}{a_2 s^2 + a_1 s + 1} e^{-Ls} \tag{4.5.4}$$

The PID controller transfer function in equation (4.5.3) can also be written as

$$G_c(s) = k\left(\frac{As^2 + Bs + C}{s}\right) \tag{4.5.5}$$

where $A = K_d/k$, $B = K_p/k$ and $C = K_i/k$.

By selecting $A = a_2$, $B = a_1$ and $C = 1$, *the open loop transfer function becomes*

$$G_p(s)G_c(s) = \frac{kb_0 e^{-Ls}}{s}, \tag{4.5.6}$$

where k is to be determined based on gain and phase-margin specifications. Insert Equation (4.5.6) into equations (4.5.3a) – (4.5.3d), resulting:

$$\arg[\frac{kb_0 e^{-jL\omega_p}}{j\omega_p}] = -\pi \quad \Rightarrow \quad \omega_p L = \frac{\pi}{2} \tag{4.5.7a}$$

$$A_m \left| \frac{kb_0 e^{-jL\omega_p}}{j\omega_p} \right| = 1 \quad \Rightarrow \quad A_m = \frac{\omega_p}{kb_0} \tag{4.5.7b}$$

$$\left| \frac{kb_0 e^{-jL\omega_g}}{j\omega_g} \right| = 1 \quad \Rightarrow \quad kb_0 = \omega_g \tag{4.5.7c}$$

$$. \quad \Phi_m = \pi + \arg\left[\frac{kb_0 e^{-jL\omega_g}}{j\omega_g} \right] \quad \Rightarrow \quad \Phi_m = \frac{\pi}{2} - \omega_g L \tag{4.5.7d}$$

Multiplied L on both sides of equation (4.5.7c), and substituting into (4.5.7d), we obtain

$$. \qquad \Phi_m = \frac{\pi}{2} - kb_0 L \tag{4.5.8}$$

By equation (4.5.7b)

$$\Phi_m = \frac{\pi}{2} - \frac{L\omega_p}{A_m}. \tag{4.5.9}$$

substitute (4.5.7a) into (4.5.9), we have

$$\Phi_m = \frac{\pi}{2}\left(1 - \frac{1}{A_m}\right). \tag{4.5.10}$$

And substitute (4.5.10) into (4.5.8), results

$$k = \frac{\pi}{2A_m L b_0} \tag{4.5.11}$$

Since from (4.5.4) and (4.5.5), we have

$$B = \frac{K_p}{k} = a_1 \Rightarrow K_p = \frac{\pi a_1}{2A_m L b_0} \tag{4.5.12a}$$

$$C = \frac{K_i}{k} = 1 \Rightarrow K_i = \frac{\pi}{2A_m L b_0} \tag{4.5.12b}$$

$$A = \frac{K_d}{k} = a_2 \Rightarrow K_d = \frac{\pi a_2}{2A_m L b_0} \tag{4.5.12c}$$

By this formulation, the gain and phase margin are interrelated to each other, some possible gain and phase margin selections are given in Table 4.5.1.

Table 4.5.1. Typical gain and phase margin values

Φ_m	$\pi/4$	$\pi/3$	$3\pi/8$	$2\pi/5$
A_m	2	3	4	5

The PID parameters are then given by

$$\begin{bmatrix} k_p \\ k_i \\ k_d \end{bmatrix} = \frac{\pi}{2A_m L b_0} \begin{bmatrix} a_1 \\ 1 \\ a_2 \end{bmatrix}. \tag{4.5.13}$$