Design of PID Controllers

4

4.1 INTRODUCTION

is chapter describes some methods for determining the parameters of a) controller. The properties of the design methods will be illustrated ng a fourth-order process model. The methods differ with respect to the owledge of the process dynamics they require. A PI controller is described two parameters (K and T_i) and a PID regulator by three or four param- rs (K, T_i, T_d and N). In the classical Ziegler-Nichols methods, the dynam- are characterized by two parameters. In the step response method, they taken from the step response. In the Ziegler-Nichols frequency response thod, the parameters are the frequency where the open-loop dynamics e a phase shift of 180° and the gain at that frequency. An obvious ension of the frequency response method is to develop methods that are ed on knowledge of the open-loop transfer function at two frequencies, , four parameters. Another way to obtain a characterization of process amics with few parameters is, of course, to use low-order dynamic mod- with few parameters. Design methods based on dynamic models of first l second order are discussed. A corresponding treatment of discrete time dels is also given. The discrete time models have the advantage that they describe time delays using finite dimensional models. Many of the ign methods described give good responses to load disturbances. The response to command signals will, however, often show a significant overshoot. The nature of this problem is discussed, and it is shown that the difficulty is due to a deficiency of the conventional PID structure. A simple way to alleviate this problem is suggested. The different design methods are compared, and some insight into the sensitivity problem and the differences between PI and PID control are also given.

4.2 ZIEGLER-NICHOLS METHODS

Two classical methods were presented by Ziegler and Nichols in 1942. These methods are still widely used, either in their original form or in some modification.

Ziegler-Nichols Step Response Method

The first design method presented by Ziegler and Nichols is based on a registration of the open-loop step response of the system, which is characterized by two parameters (see Figure 4.1). The point where the slope of the step response has its maximum is first determined, and the tangent at this point is drawn. The intersections between this tangent and the coordinate axes give the two parameters a and L. In Chapter 3, a model of the process to be controlled was derived from these parameters. Ziegler and Nichols have given PID parameters directly as functions of a and L. These are given in Table 4.1. An estimate of the period T_p of the dominant dynamics of the closed-loop system is also given in the table.

Example 4.1—The Ziegler-Nichols step response method will be applied to the process

$$G_p(s) = \frac{1}{(1+s)(1+0.2s)(1+0.05s)(1+0.01s)} \tag{4.1}$$

This process model is used as a test example throughout the chapter. Measurements on the step response give the parameters $a = 0.11$ and $L = 0.16$. The controller parameters can now be determined from Table 4.1. It follows that a PI controller should have the parameters $K = 8.2$ and $T_i = 0.48$. The parameters of a PID controller are $K = 10.9$, $T_i = 0.32$, and $T_d = 0.08$. Figure 4.2 shows the response of the closed-loop systems to a step command and a load disturbance.

Table 4.1
Recommended PID Parameters According to Ziegler-Nichols Step Response Method

Controller	K	T_i	T_d	T_p
P	1/a			4L
PI	0.9/a	3L		5.7L
PID	1.2/a	2L	L/2	3.4L

Notice that the response of the PI controller is poorly damped but that response of the PID controller is better. The overshoot in the response the command signal is, however, excessive even for the PID controller.

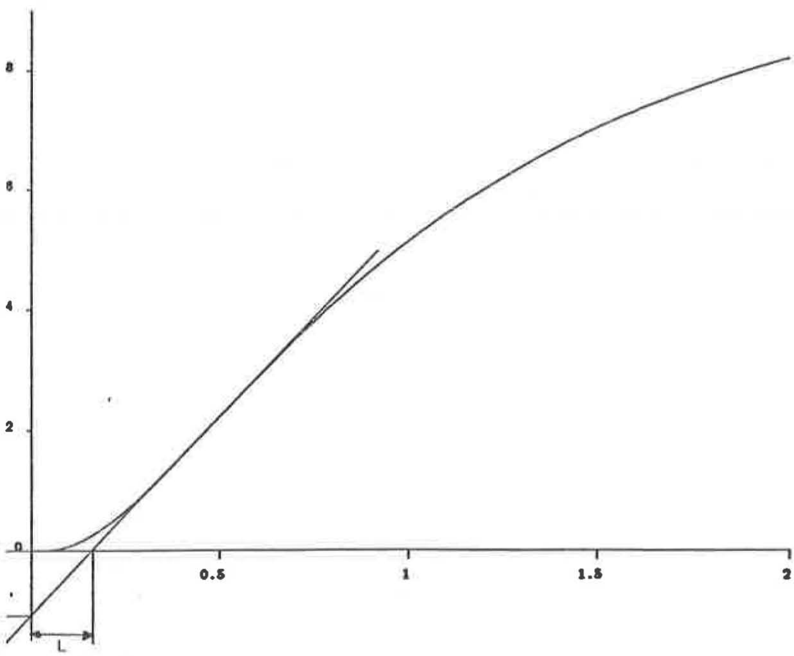

Figure 4.1
Characterization of a Step Response
Used in the Ziegler-Nichols Step Response Method

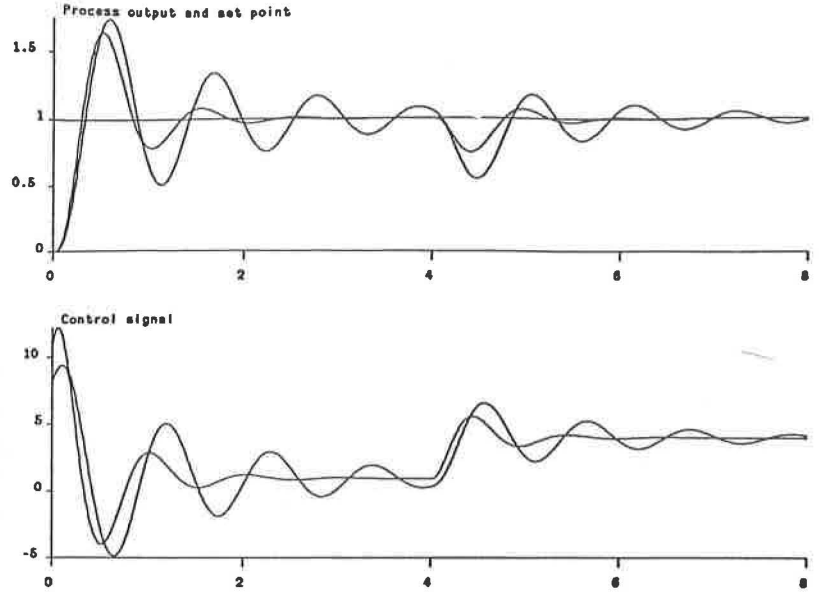

Figure 4.2
Step and Load Disturbance Response of the Process (Equation 4.1) Controlled by a PI Controller (thin lines) a PID Controller (thick lines) Tuned by the Ziegler-Nichols Step Response Method

Ziegler-Nichols Frequency Response Method

This method is also based on a very simple characterization of the process dynamics. The design is based on knowledge of the point on the Nyquist curve of the process transfer function G where the Nyquist curve intersects the negative real axis. For historical reasons this point is characterized by the parameters k_c and t_c, which are called the *ultimate gain* and the *ultimate period*. Section 3.2 described a method to obtain parameters k_c and t_c by increasing the gain in a proportional controller until the stability boundary is reached. The parameters can also be obtained using the relay feedback experiment presented in the same section. The Ziegler-Nichols design method gives simple formulas for the parameters of the controller in terms ıe ultimate gain and the ultimate period (see Table 4.2). An estimate of ɔeriod (T_p) of the dominant dynamics of the closed-loop system is also n in the table.

Table 4.2
Recommended PID Parameters According to Ziegler-Nichols Frequency Response Method

ɔntroller	K	T_i	T_d	T_p
	$0.5\ k_c$			t_c
	$0.4\ k_c$	$0.8\ t_c$		$1.4\ t_c$
ID	$0.6\ k_c$	$0.5\ t_c$	$0.12\ t_c$	$0.85\ t_c$

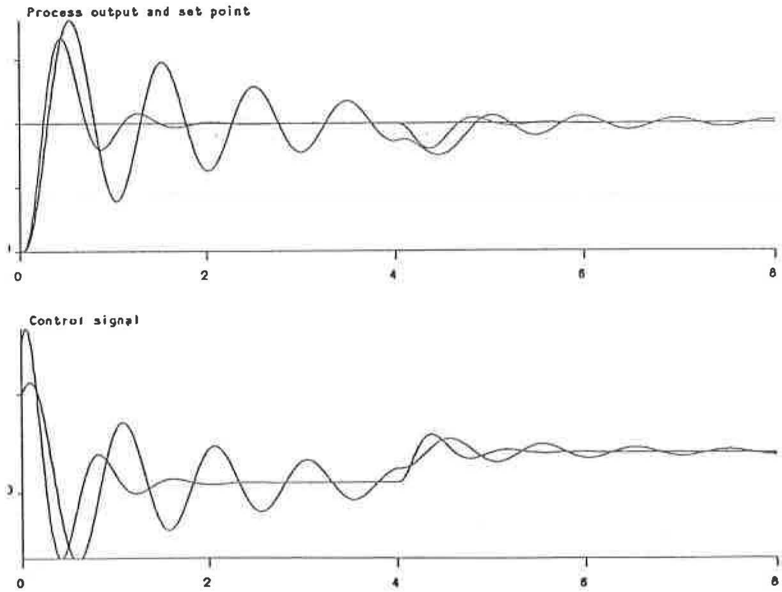

Figure 4.3
Step and Load Disturbance Response When the Process (Equation 4.1) is Controlled by a PI Controller (thin lines) and a PID troller (thick lines) Tuned by Ziegler-Nichols Frequency Response Method

Example 4.2—The process of Equation 4.1 has the ultimate gain $k_c \approx 25$ and the ultimate period $t_c \approx 0.63$. Table 4.2 gives the parameters $K = 10$ and $T_i = 0.50$ for a PI controller and $K = 15$, $T_i = 0.31$, and $T_d = 0.08$ for a PID controller. Figure 4.3 shows the closed-loop step and load disturbance responses when the controllers are applied to the Equation 4.1 process. The parameters and the performance of the controllers obtained with the frequency response method are quite close to those obtained by the step response method.

The Ziegler-Nichols tuning rules were originally designed to give systems with good responses to load disturbances. They were obtained by extensive simulations of many different systems. The design criterion was quarter amplitude damping. In Section 3.2, the relation between the damping (d) and the relative damping (ζ) is given as:

$$\zeta = \frac{1}{\sqrt{1 + (2\pi/\log(d))^2}}$$

Quarter amplitude damping, $d = 1/4$, gives the relative damping $\zeta = 0.22$, which is often considered too small. This is clearly seen in the examples above. The performance can be improved by the modification discussed below. In control loops where the major design objective is to quickly compensate for load disturbances, the high gain provided by the Ziegler-Nichols method is good. In these cases, large overshoots and oscillations during set point changes can be avoided by ramping the set point or performing the set point shift in several steps. In Section 2.4, another method to avoid large overshoots caused by set point changes was described.

Relations Between the Ziegler-Nichols Methods

The relations between the two methods can be seen by considering control of an integrator with a delay. Such a process has the transfer function

$$G(s) = \frac{b}{s} e^{-sT}$$

eq (3.1)

The step response parameters are $L = T$ and $a = bT$. The PID parameters obtained by the step response methods therefore become (Table 4.1)

$$K = \frac{1.2}{b \cdot T} \qquad T_i = 2T \qquad T_d = \frac{T}{2}$$

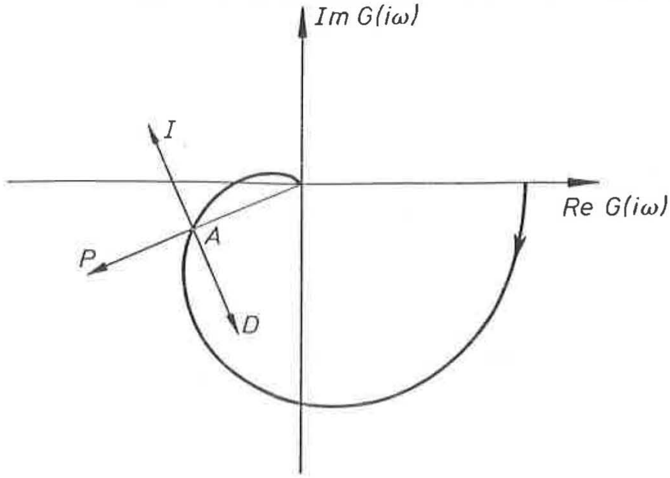

Figure 4.4
A Given Point on the Nyquist Curve May Be Moved to an Arbitrary Position in the G-plane by PI, PD, or PID Control. (Point A may be moved in the directions G(iω), G(iω)/iω, and iωG(iω) by changing the proportional, integral, and derivative gain, respectively)

e ultimate period of the system is $t_c = 4T$, and the ultimate gain is $k_c =$ T. The PID parameters obtained by the frequency response methods herefore,

$$\frac{0.6\pi}{2bT} \approx \frac{0.94}{bT} \qquad T_i = 2T \qquad T_d = \frac{T}{2}$$

An Interpretation of the Ziegler-Nichols Frequency Domain Method

e Ziegler-Nichols frequency domain method will be interpreted in of moving points in the Nyquist diagram. The method starts with nination of the point $(-1/k_c, 0)$ where the Nyquist curve of the open-system intersects the negative real axis. With PI or PID control, it is ɔle to move a given point on the Nyquist curve to an arbitrary position in the complex plane, as indicated in Figure 4.4. By changing the gain, it is possible to move the Nyquist curve in the direction of $G(i\omega)$, i.e., radially from the origin. Point A may be moved in the orthogonal direction by changing integral or derivative gain. It is thus possible to move a specified point to an arbitrary position, an idea that can be used to obtain design methods.

Let ω be the frequency that corresponds to A. The frequency response of the regulator at ω is

$$G_R(i\omega) = k\left[1 + \frac{1}{i\omega T_i} + i\omega T_d\right] = r_R e^{i\varphi_R}$$

with positive regulator parameters the angle φ_R is thus restricted to the range $-\pi/2 \leq \varphi_R \leq \pi/2$ where $\varphi_R = -\pi/2$ corresponds to pure integral control and $\varphi_R = \pi/2$ to pure derivative control.

Pure derivative control cannot be implemented (compare with Equation 2.7). The range of φ_R is therefore $-\pi/2 \leq \varphi_R \leq \varphi_0$ where φ_0 is about $\pi/3$ or 60°.

With the Ziegler-Nichols frequency response method it follows that

$$G_R(i\omega_c) = 0.6k_c\left[1 + i\left(\omega_c T_d - \frac{1}{\omega_c T_i}\right)\right]$$

$$= 0.6k_c\left[1 + i\left(\frac{2\pi}{8} - \frac{1}{\pi}\right)\right] = k_c(0.6 + 0.28i)$$

The Ziegler-Nichols frequency response method can thus be interpreted as finding regulator parameters so that the point where the Nyquist curve intersects the negative real axis is moved to $-0.6 - 0.28i$. This corresponds to a phase advance of 25° at ω_c.

A Modified Ziegler-Nichols Method

With the given interpretation, it is straightforward to generalize the Ziegler-Nichols frequency domain method. Other points of the Nyquist curve can be selected. They can also be moved to other positions. In this way it is possible to obtain design methods where the specifications are given in terms of amplitude margins or phase margins.

A general formulation is to start with a given point of the Nyquist curve of the process

$$G_p(i\omega) = r_p e^{i(\pi+\varphi_p)}$$

l a regulator so that this point is moved to

$\varphi_s)$

ılitude margin design corresponds to $\varphi_s = 0$ and $r_s = 1/A_m$ where mplitude margin; a phase margin design corresponds to $r_s = 1$ and here φ_m is the specified phase margin; and the Ziegler-Nichols domain method corresponds to $r_s = 0.66$ and $\varphi_s = 0.44$.
the frequency response of the controller as

$_R e^{i\varphi_R}$

$r_p r_R e^{i(\pi+\varphi_p+\varphi_R)}$

ıller should thus be chosen so that

$= \varphi_p$

:ulations give

$$\frac{(\varphi_s-\varphi_p)}{r_p}$$

$$= \tan(\varphi_s-\varphi_p)$$

ı k is uniquely given. However, only one equation determines the ; T_i and T_d. An additional condition must thus be introduced to these parameters uniquely. A common method is to specify a :lation between T_i and T_d, i.e.,

'ten is chosen as $\alpha = 0.25$. Straightforward calculations now give :ters T_i and T_d.

$$-\tan(\varphi_p-\varphi_p) + \sqrt{4\alpha+\tan^2(\varphi_p-\varphi_s)}\Big]$$

i

For systems where the amplitude and the phase of the transfer function decreases monotonously, the choice $r_s = 0.5$ and $\varphi_s = \pi/4$ guarantees an amplitude margin of at least 2 and a phase margin of at least 45°.

Assuming that a Ziegler-Nichols experiment is used to determine a suitable point, we have $r_p = 1/k_c$ and $\varphi_p = 0$. The controller parameters are then given by $k = 0.35\, k_c$, $T_i = 0.77\, T_c$, and $T_d = 0.19 T_c$. This can be compared with the values given by the Ziegler-Nichols frequency response method.

The Ziegler-Nichols frequency response method and the modified Ziegler-Nichols method are based on the idea of moving one point on the Nyquist curve to a desired position. The terms phase margin and amplitude margin

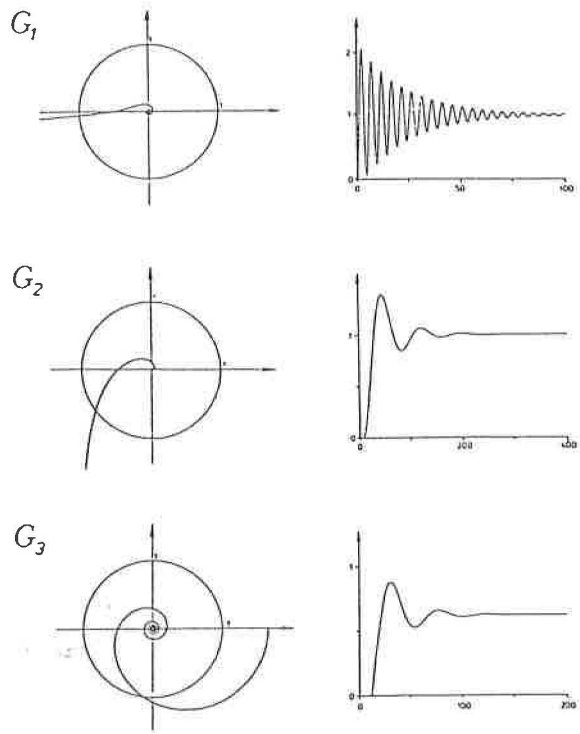

Figure 4.5
Nyquist Curves of Systems with Equal Amplitude Margin and Their Corresponding Closed-Loop Step Responses

o define one point on the Nyquist curve. In most cases these simple design les are sufficient, but there are exceptions. Figure 4.5 shows the Nyquist rves of three systems having the same amplitude margin, $A_m = 2$. This eans that all Nyquist curves pass through the point $z = -0.5$. Figure 4.6 ows the Nyquist curves of three systems having the same phase margin, = 45°. This means that all Nyquist curves pass through the point $z =$ .707 – 0.707i. The corresponding step responses clearly demonstrate that e transient behavior of the control loop is also influenced by other points the Nyquist curve. Design methods where several points on the Nyquist rve are determined are described below.

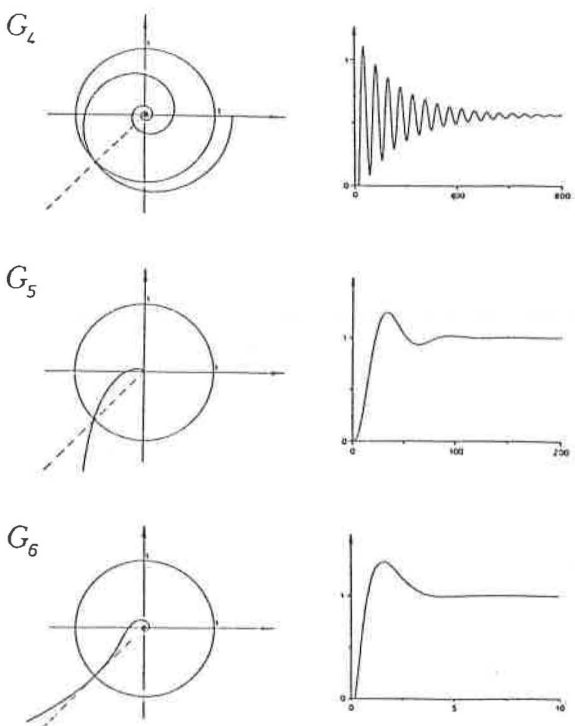

Figure 4.6
Nyquist Curves of Systems with Equal Phase Margin and Their Corresponding Closed-Loop Step Responses

4.3 DOMINANT POLE DESIGN

The Ziegler-Nichols methods discussed in the previous section were based on the knowledge of only one point on the Nyquist curve of the open-loop process dynamics. This section presents a design method that uses two points on the Nyquist curve. The method is based on a simple method of estimating the dominant poles of the closed-loop system from the open-loop transfer function. The notion of dominant poles is first discussed. The design method is then developed.

Dominant poles

Consider a closed-loop system obtained by negative feedback around a linear system with the transfer function $G(s)$ (see Figure 4.7). The transfer function of the closed-loop system from the command signal to the output is given by

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

Many properties of the closed-loop system can be deduced from the poles and the zeros of $G_c(s)$, which are the same as the zeros of $G(s)$ (i.e., the zeros of the plant and the controller). The closed-loop poles are the roots of the equation

$$1 + G(s) = 0$$

The pole-zero configurations of closed-loop systems may vary considerably. Many simple feedback loops will, however, have a configuration of the type shown in Figure 4.8 where the principal characteristics of the response are given by a complex pair of poles, p_1 and p_2, called the *dominant poles*. The response is also somewhat influenced by real poles and zeros p_3 and z_1, respectively. The position of z_1 and p_3 may be reserved. There may also be more poles and zeros far from the origin. Poles and zeros whose real parts are much smaller than the real part of the dominant poles have little influence on the transient response. Classical control was very much concerned with closed-loop systems having the pole-zero configuration shown in Figure 4.8.

Even if many closed-loop systems have a pole-zero configuration similar to the one shown in Figure 4.8, there are, however, exceptions. For instance, systems with mechanical resonances, which may have poles and zeros close to the imaginary axis, are generic examples of systems that do not fit the pole-zero pattern of the figure. Such systems are not treated in this section.

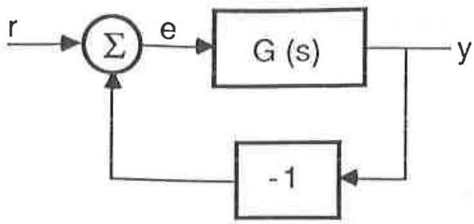

Figure 4.7
Block Diagram of a Simple Feedback System

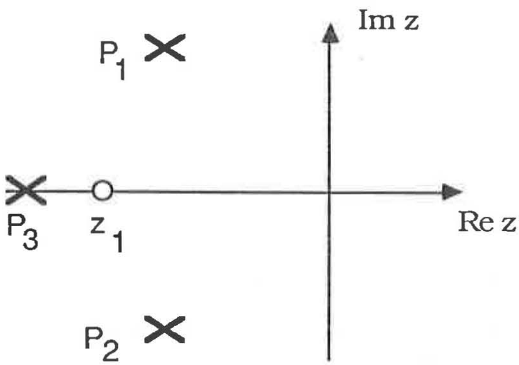

Figure 4.8
Pole-Zero Configuration of a Simple Feedback System

PI Control

ɜ dominant pole design method will first be applied to PI control. Two -loop poles can be specified, since a PI controller has two adjustable ɩeters. Let the PI controller be parameterized as

$$= k + \frac{k_i}{s}$$

where k is the proportional gain and k_i is the integral gain. The parameters k and k_i will be determined so that the closed-loop system has poles at $s = p_1$ and $s = p_2$, where

$$p_1 = -\zeta\omega_0 + i\omega_0\sqrt{1 - \zeta^2} = -\sigma + i\omega$$
$$p_2 = -\zeta\omega_0 - i\omega_0\sqrt{1 - \zeta^2} = -\sigma - i\omega \tag{4.2}$$

This implies

$$\begin{cases} 1 + \left[k + \dfrac{k_i}{p_1}\right] G_p(p_1) = 0 \\ 1 + \left[k + \dfrac{k_i}{p_2}\right] G_p(p_2) = 0 \end{cases}$$

where $G = G_R G_P$, and G_P is the transfer function of the process. The above equation is linear in k and k_i. It has a solution if $| G(p_1) | \neq 0$. The solution is

$$k(\omega_0) = -\frac{\sqrt{1-\zeta^2}\ A(\omega_0) + \zeta B(\omega_0)}{\sqrt{1-\zeta^2}\ [A(\omega_0)^2 + \zeta B(\omega_0)^2]}$$
$$k_i(\omega_0) = -\frac{\omega_0 B(\omega_0)}{\sqrt{1-\zeta^2}\ [A(\omega_0)^2 + B(\omega_0)^2]} \tag{4.3}$$

where $A(\omega_0) = Re\ G_P(p_1)$ and $B(\omega_0) = Im\ G_P(p_1)$.

The parameter ω_0 can be viewed as a design parameter that determines the response speed. Small values of ω_0 give a slow system, and large values give a fast system. If the process dynamics are of first order, the closed-loop system only has two poles: p_1 and p_2. The design parameter ω_0 can then be chosen arbitrarily. For higher-order dynamics, the closed-loop system will, however, have more poles. For stable systems with poles on the real axis, these poles will have real parts that are greater than $-\zeta\omega_0$ for large ω_0. The condition that the poles p_1 and p_2 are dominating will thus give an admissible range of the design parameter ω_0. The upper bound of ω_0 can be determined from the condition that the largest pole on the real axis is at $s = -\alpha\omega_0$. For stable processes G_P, the function $A(\omega_0)$ is positive and $B(\omega_0)$ is small for small ω_0. It then follows that the proportional gain $k(\omega_0)$ is negative for small ω_0. Since it is normally desirable to have positive controller gains, a lower bound for the design parameter is given by the condition $k(\omega_{0I}) = 0$. The value ω_{0I} corresponds to pure integral control. Analogously, the value ω_{0P} corresponds to pure proportional control. An alternative to choosing

,, based on pole dominance, is to select an ω_0 that gives the largest value of ə integral gain. This gives values that are very close to those obtained from e condition of pole dominance. A physical interpretation of the condition ll be given later in connection with the discussion of PID control.

PD Control

The dominant pole design can also be applied to PD control. Let the PD ntroller be parameterized as

$$_{R}(s) = k + k_d s$$

ıd require that the closed-loop system has poles at p_1 and p_2 given by]uation 4.2. Calculations analogous to those for the PI controller give

$$\omega_0) = \frac{-\sqrt{1-\zeta^2}\,A(\omega_0) + \zeta B(\omega_0)}{\sqrt{1-\zeta^2}\,[A(\omega_0)^2 + B(\omega_0)^2]}$$

$$(\omega_0) = \frac{B(\omega_0)}{\omega_0\sqrt{1-\zeta^2}\,[A(\omega_0)^2 + B(\omega_0)^2]} \qquad (4.4)$$

here $A(\omega_0) = Re\ G_P(p_1)$ and $B(\omega_0) = Im\ G_P(p_1)$. Notice that PI and PD ıntrol are complementary since Equation 4.4 gives $k_d(\omega_0) < 0$ for $\omega < \omega_{0p}$ ıd (4.3) gives $k_i(\omega_0) < 0$ for $\omega > \omega_{0p}$ The design parameter ω_0 is thus ways larger for PD control than for PI control as can be expected. An ›per bound for the design parameter for PD control is given by the condi- ›n $k(\omega_{0D}) = 0$, where parameter ω_{0D} corresponds to pure derivative con- ol. A reasonable choice of the design parameter is the value that corres- ›nds to the largest value of the proportional gain. Another alternative is to ıalyze the conditions for pole dominance.

PID Control

'ith PID control, it is possible to position three closed-loop poles. Let the ansfer function of the PID regulator be parameterized as

$$_R'(s) = k' + \frac{k_i'}{s} + k_d' s$$

here k' is the proportional gain, k_i' the integral gain, and k_d' the derivative ıin. Two closed-loop poles will first be positioned according to Equation 2, as was done for PI control. Assume that the PI design problem is already solved, i.e., that the functions $k(\omega_0)$ and $k_i(\omega_0)$ given by Equation 4.3 are known. The value of the regulator transfer function G'_R at $p_1 = -\sigma + i\omega$ is

$$G'_R(-\sigma+i\omega) = k' + \frac{k'_i}{-\sigma + i\omega} + k'_d(-\sigma + i\omega)$$

$$= k' - \frac{\sigma k'_i}{\omega_0^2} - \sigma k'_d + i\left[-\frac{\omega k'_i}{\omega_0^2} + \omega k'_d\right]$$

Requiring that this transfer function has the same value as the transfer function for PI control gives

$$k' - \frac{\sigma k'_i}{\omega_0^2} - \sigma k'_d = k - \frac{\sigma k_i}{\omega_0^2}$$

$$-\frac{\omega k'_i}{\omega_0^2} + \omega k'_d = -\frac{\omega k_i}{\omega_0^2}$$

Hence,

$$k'(\omega_0) = k(\omega_0) + 2\sigma k'_d = k(\omega_0) + 2\zeta\omega_0 k'_d$$

$$k'_i(\omega_0) = k_i(\omega_0) + \omega_0^2 k'_d \qquad (4.5)$$

Thus, there is a two-parameter (ω_0, k'_d) family of gains for a PID regulator, which gives a closed-loop system with poles at $s = p_1$ and $s = p_2$. The parameter k'_d will now be determined so that the closed-loop system also has a pole at $s = -\omega_0$.
Hence,

$$1 + \left[k' - \frac{k'_i}{\omega_0} - k'_d\omega_0\right] G_P(-\omega_0) = 0$$

Inserting the expressions in Equation 4.5 for k' and k'_i gives

$$1 + \left[k - \frac{k_i}{\omega_0} + 2\sigma k'_d - 2\omega_0 k'_d\right] G_P(-\omega_0) = 0$$

If $G_P(-\omega_0) \neq 0$, this equation can be solved with respect to k'_d. The solution is

$$k'_d(\omega_0) = \frac{1 + [k(\omega_0) - k_i(\omega_0)/\omega_0]G_P(-\omega_0)}{2\omega_0(1 - \zeta)G_P(-\omega_0)} \qquad (4.6)$$

ations 4.5 and 4.6 define a one-parameter (ω_0) family of controller s, which gives a closed-loop system with poles at $-\zeta\omega_0 \pm i\omega_0\sqrt{1-\zeta^2}$ $-\omega_0$. The parameter ω_0 may be viewed as a design parameter. Small es of ω_0 give a slow system, and large values a fast system. If there are onstraints on the signs of the regulator gains and the system dynamics of second order, arbitrary values of design parameter ω_0 can be speci- since the closed-loop system has only three poles. For systems with er-order dynamics, the condition that the chosen poles are dominating give constraints on the design parameter.

nple 4.3—The properties of the dominant pole design method will now lustrated. Consider a process with the transfer function given in Equa- 4.1. A PI regulator that gives closed-loop poles with relative damping .707 will first be designed. The smallest value of the design parameter gives a nonnegative proportional gain is $\omega_{0I} = 0.62$. This corresponds to integral control, i.e., $k = 0$ and $k_i = 0.394$. For $\omega_0 = 2.23$, the closed-loop m has poles at $-1.58 \pm 1.58i$, -2.24, -20.6, and -100. The controller s are $k = 1.62$ and $T_i = 0.70$. A comparison with Example 4.1 and nple 4.2 shows that the parameters obtained by the dominant pole n are significantly different from those obtained by the Ziegler-Nichols ods. For larger values of ω_0, the pole at -2.24 will move towards the and become dominating. For sufficiently large ω_0, gain k_i becomes tive. Controller gains for some different values of ω_0 are shown in Table The integral gain (k_i) has its largest value for $\omega_0 = 2.45$. The parameters $k = 1.73$ and $T_i = 0.74$, which are close to the values obtained for pole inance. The integral gain becomes zero for $\omega_0 = 3.72$. PD control can be for larger values of ω_0. The controller parameters for PD control are shown in Table 4.3. The proportional gain has its largest value for $\omega_0 =$. and becomes zero for $\omega_0 = 16.65$. With PID control, the design meter can be increased significantly compared with PI control. Table hows the parameters obtained for different ω_0. The closed-loop system have a double pole at $s = -\omega_0$ for $\omega_0 = 7.16$. The regulator parameters are $k = 11.9$, $T_i = 0.45$, and $T_d = 0.115$. To assess the different designs, first rve that the time to the peak of a step response is approximately $4.5/\omega_0$. value of ω_0 can thus be used to determine the response time. The value e integral gain (k_i) is also useful to assess the response to load distur- es. Consider a step in load disturbances. The control law is given by

$$= ke(t) + k_i \int^t e(s)ds + k_d \frac{de}{dt}$$

Table 4.3
PI and PD Regulators for Different Values of Design Parameter ω_0

ω_0	k	k_i	T_i	k_d	T_d
0.62	0	0.394			
1.00	0.51	0.887	0.574		
2.00	1.48	2.16	0.684		
2.20	1.60	2.29	0.700		
2.30	1.66	2.33	0.713		
2.40	1.71	2.35	0.728		
2.45	1.73	2.35	0.738		
2.50	1.76	2.35	0.748		
2.60	1.80	2.32	0.774		
3.00	1.90	1.96	0.972		
3.72	1.88	0	∞	0	0
4.00	2.24			0.080	0.036
6.00	5.09			0.598	0.118
8.00	7.73			1.02	0.132
10.0	9.43			1.33	0.141
11.0	9.72			1.45	0.150
12.0	9.52			1.55	0.163
14.0	7.36			1.67	0.226
16.0	2.33			1.68	0.322
16.65	0			1.66	∞

Table 4.4
PID Regulators for Different Values of Design Parameter ω_0

ω_0	k	k_i	T_i	k_d	T_d
2.00	1.15	1.70	-0.114		
2.24	1.62	2.31	0	0.704	0
3.00	3.32	4.98	0.336	0.668	0.101
4.00	5.78	10.0	0.705	0.587	0.122
5.00	8.20	16.1	0.996	0.510	0.121
6.00	10.3	22.1	1.21	0.466	0.118
7.00	11.8	26.2	1.35	0.500	0.115
7.16	11.9	26.5	1.37	0.452	0.115
7.50	12.2	26.8	1.40	0.454	0.115
8.00	12.3	26.1	1.42	0.473	0.115

\ssume that the system is initially at rest. With a controller having gral action, the error and its derivative are then zero. Let the system be ject to a load disturbance. For systems with constant static gain, the load urbance must be compensated with a change of the control signal Δu. s change is then given by

$$= k_i \int_0^{\infty} e(s)ds$$

'he error integral due to a load disturbance is then

$$e(s)ds = \frac{\Delta u}{k_i}$$

'or a given load disturbance, the error integral is thus inversely propor- ıal to k_i.

The properties of the different control laws can now be assessed. With e integral control, design parameter ω_0 is 0.62 and k_i is 0.39. The peak e is then approximately 7.2*s*. With PI control, the design parameter can :hosen in the range $0.62 < \omega_0 < 2.5$. This means that the response time be increased by a factor of 4 compared to pure integral control. The :gral gain k_i can also be increased from 0.39 to 2.35, which means that the)r integral for load disturbances can be reduced by a factor of 6 com- ed to pure integral control. Notice that the largest value of k_i is obtained $\omega_0 = 2.45$.

With PD control, the design parameter can be chosen in the range $3.7 \leq$ ≤ 11, with proportional gains in the range $1.9 \leq k \leq 9.7$. The largest ıe of the loop gain with PD control is 9.7, which means that PD control only be used if the largest steady-state error is less than 10%.

With PID control the design parameter ω_0 can be chosen in the range 0.6 $_0 \leq 7.5$. The value $\omega_0 = 7.5$ gives a threefold increase of response time ıpared to PI control. The integral gain k_i can be increased from 2.35 for control to 26.8, which corresponds to an error integral for load distur- ıces that is more than 11 times smaller.

Responses to step changes in the set point and load are shown in Figure The simulations support the results of the analysis.

The dominant pole design is useful since it gives predictable results. It , however, the drawback that the transfer function must be known in the ıplex plane. Approximate methods, which require only the values of the ıuist curve, are, therefore, developed below.

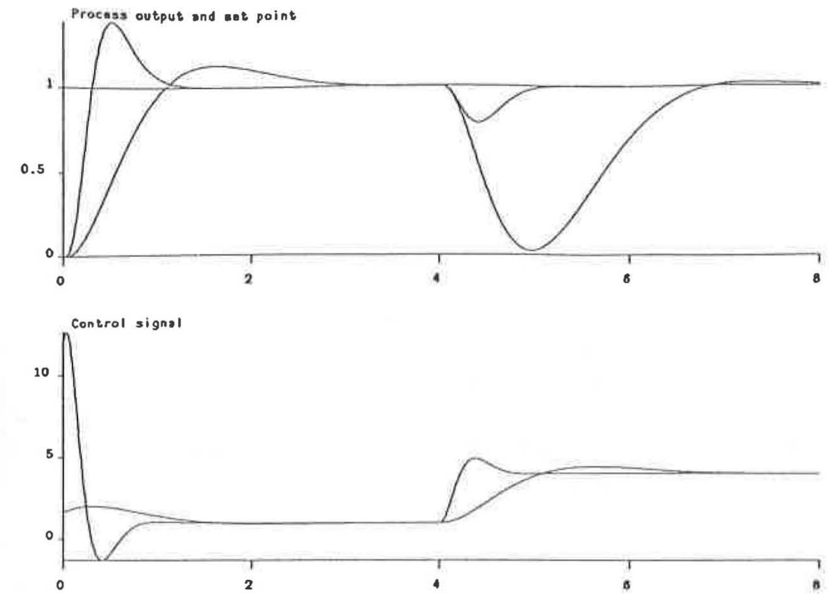

Figure 4.9
Step and Load Disturbance Responses of the Closed-Loop System in Example 4.3 Obtained with the Dominant Pole Design with $\omega_0 = 7.16$

Approximate Determination of the Dominant Poles

The following is a simple method for estimating the dominant poles from knowledge of the Nyquist curve of the open-loop system. The closed-loop poles are given by the characteristic equation

$$G(s) + 1 = 0$$

A Taylor series expansion around $s = i\omega$ gives

$$0 = 1 + G(-\sigma+i\omega) = 1 + G(i\omega) + i\sigma G'(i\omega) + \dots$$

where

$$G'(i\omega) = \frac{dG(i\omega)}{d\omega}$$

‹eglecting terms of second and higher orders in σ, we find

$$G(i\omega) + i\sigma G'(i\omega) = 0$$

ıce,

$$i\frac{1 + G(i\omega)}{G'(i\omega)} \tag{4.7}$$

both σ and ω of the dominant poles are determined. Notice that ω must :hosen so that σ becomes real. This analytic derivation shows that Equa- : 4.7 will give good results for small σ, i.e., when the dominant poles are e to the imaginary axis. The approximation will not hold if the function) has singularities inside a circle with the center in $i\omega$ and radius ω. This ıns that σ must be smaller than ω.

f the derivative is approximated by a difference between two close points the Nyquist curve, the following expression for determining σ is ɪined:

$$\frac{\omega_2) - G(i\omega_1)}{\omega_2 - \omega_1} = i\frac{1+G(i\omega_2)}{\sigma} \tag{4.8}$$

ɜy introducing a controller in the loop, the dominant poles may be ʼed to the desired new positions. The corresponding design problem may ı be expressed in terms of the frequency (ω) and the relative damping (ζ) he dominant poles.

ʼo perform the design, it is assumed that the values of the open-loop .sfer function at two neighboring frequencies, ω_1 and ω_2, are known, i.e.,

$$i\omega_1) = a_1 + ib_1$$

$$i\omega_2) = a_2 + ib_2$$

t is also assumed that frequencies ω_1 and ω_2 are close to the crossover uency. The design is not restricted to any particular controller structure, almost any controller with at least two adjustable parameters may be . A PID controller of the form

$$s) = K\left[1 + \frac{1}{sT_i} + sT_d\right]$$

ıosen as an illustration. Furthermore, it is assumed that there is a given ion between the integration time (T_i) and the derivative time (T_d).

$$: \alpha T_i \tag{4.9}$$

Hence,

$$G_R(s) = K\left[1 + \frac{1}{sT} + s\alpha T\right]$$

This regulator has two adjustable parameters: gain K, which moves the Nyquist curve radially from the origin, and time constant T, which twists the curve.

The design problem is then to determine a controller so that the transfer function of the compensated system has desired values at the two frequencies, i.e.,

$$\begin{aligned} G(i\omega_1) &= G_o(i\omega_1)\, G_R(i\omega_1) = c_1 + id_1 \\ G(i\omega_2) &= G_o(i\omega_2)\, G_R(i\omega_2) = c_2 + id_2 \end{aligned} \tag{4.10}$$

In the sequel, it is assumed that the desired frequency (ω) of the dominant poles is equal to ω_2. The following relation is then obtained from Equation 4.8:

$$\sigma = \frac{G(\omega_2) + 1}{G(i\omega_2) - G(i\omega_1)}\, i(\omega_2 - \omega_1)$$

The relative damping (ζ) is introduced by

$$\sigma = \frac{\zeta\omega_2}{\sqrt{1-\zeta^2}}$$

These two equations now give

$$\frac{G(i\omega_2) - G(i\omega_1)}{G(i\omega_2) + 1} = \frac{\sqrt{1-\zeta^2}}{\zeta} \cdot \frac{i(\omega_2 - \omega_1)}{\omega_2} \triangleq i\kappa$$

It follows from Equation 4.10 that

$$\frac{c_2 - c_I + i(d_2 - d_1)}{c_2 + 1 + id_2} = i\kappa$$

This gives

$$\begin{cases} c_2 - c_1 + \kappa d_2 = 0 & (4.11) \\ d_2 - d_1 - \kappa(c_2 + 1) = 0 & (4.12) \end{cases}$$

These conditions determine parameters K and T of the PID regulator. Equation 4.11 gives a second-order equation for T, from which T is solved. Gain K is then obtained from Equation 4.12.

:*ample 4.4*—Consider the system given by Equation 4.1. Two points on : Nyquist curve that are used for the design are given by

$(8 \cdot i) = -0.0593 - i \cdot 0.0135$

$(10 \cdot i) = -0.0396$

Using these two values of G(iω), the design method presented above can applied. The following set of PID parameters is obtained for $\alpha = 0.25$, ω_2 10, and $\zeta = 0.4$:

$= 14.2 \qquad T_i = 0.407 \qquad T_d = 0.102$

:p and load disturbance responses of the closed-loop system are given in gure 4.10.

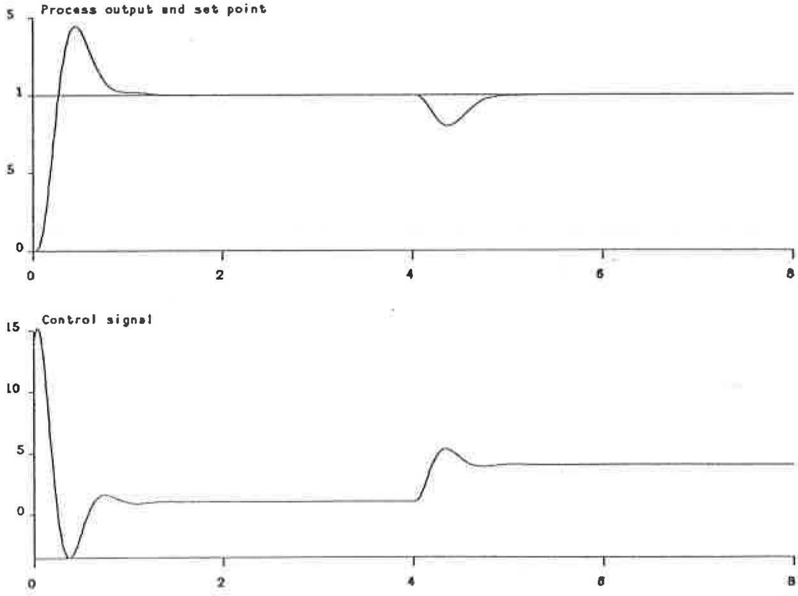

Figure 4.10
Step and Load Disturbance Responses of the Closed-Loop System in Example 4.4 Obtained with the Approximate Dominant Pole Design

A comparison of Figures 4.2 and 4.3 with Figure 4.10 shows that the responses obtained with the approximate dominant pole design are considerably better than those obtained by the Ziegler-Nichols methods. The price to be paid for the improved performance is that it is necessary to determine two points on the Nyquist curve of the open-loop system instead of one for the Ziegler-Nichols methods.

The parameters obtained by the approximate dominant pole design are quite similar to those obtained by the Ziegler-Nichols method. In the example, the gain is $K = 14.2$ versus 15 for the Ziegler-Nichols frequency response method. The other parameters are $T_i = 0.41$ (0.31) and $T_d = 0.10$ (0.08). The fact that the responses are different indicates that the parameter adjustment may be critical. This will be discussed further in Section 4.7.

Also notice that the design method is based on specification of only two parameters, σ and ω, the dominant poles. This implies that the gains of a PI or PD regulator are uniquely given. One extra condition has to be introduced to specify the three parameters of a PID controller, this condition being arbitrarily chosen as Equation 4.9.

4.4 FREQUENCY DOMAIN DESIGN

If several points on the Nyquist curve are known, many different design methods can be used. A common frequency domain approach attempts to find a compensator such that the magnitude of the closed-loop frequency response has unit gain at low frequencies and a resonance peak, M_P, which is less than a prescribed value. Such a design method is presented below.

M_P Values

Let $G = G_R G_P$ be the loop transfer function, i.e., the product of the transfer function of the controller and the process. The closed-loop transfer function is

$$G_s = \frac{G}{1 + G}$$

The curves in the G-plane where G_s has constant magnitude are given by

$$\left| \frac{G}{1 + G} \right| = M \tag{4.13}$$

ese are circles in the complex G-plane, called "M-circles". A few of the
:les are shown in Figure 4.11.

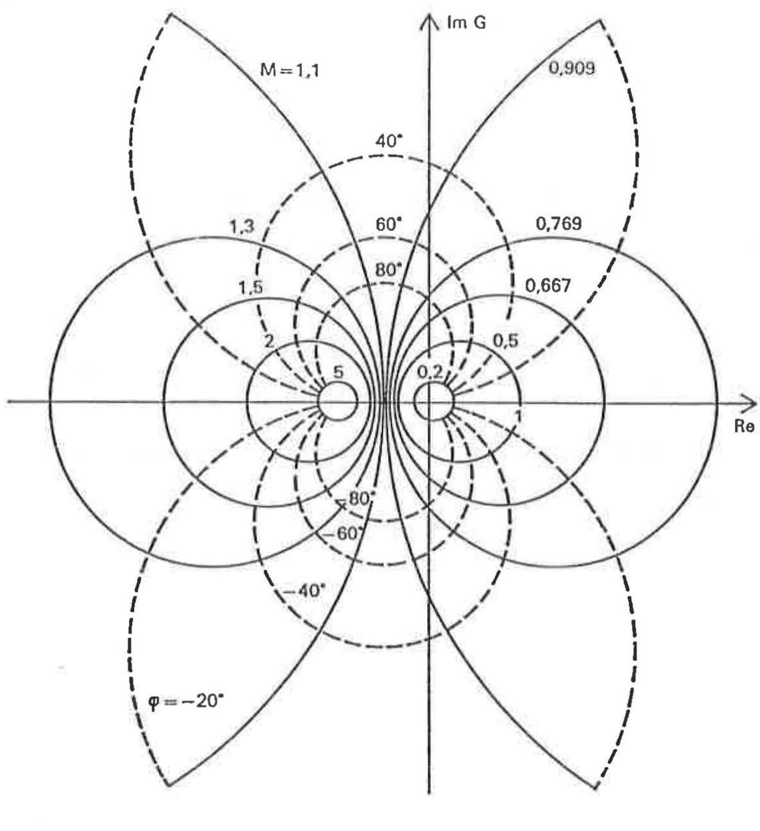

Figure 4.11
Complex G-plane with M-circles

The M_p value of a system is the largest value of M on its Nyquist curve. Notice that the Nyquist curve of a system is tangential to the M-circle, which corresponds to $M = M_p$. The M_p value can be related to other system characteristics and can be approximately computed from the relative damping (ζ) of the dominant poles in the following way:

$$M_P = \frac{1}{2\zeta\sqrt{(1-\zeta^2)}} \qquad \zeta \leq 1/\sqrt{2}$$

ζ is related to the full period damping as

$$d = e^{-\frac{2\pi\zeta}{\sqrt{1-\zeta^2}}}$$

Related values of M_P, ζ, and d are shown in Table 4.5, which also shows the radius (r) and the center (f) of the M-circles, given by

$$r = \frac{M}{M^2 - 1} \qquad f = \frac{M^2}{M^2 - 1}$$

Table 4.5.
Corresponding Values of M_P, Relative Damping (ζ), Absolute Damping (d), Radius (r) and Center (f) of the M-Circles.

M_P	ζ	d	r	f
1.1	0.54	0.018	5.24	5.76
1.2	0.47	0.034	2.72	3.27
1.3	0.42	0.052	1.88	2.45
1.4	0.39	0.071	1.46	2.04
1.5	0.36	0.091	1.20	1.80

Design Method

In the M-circle design method, the performance is specified by the M_p value, which is typically chosen in the range M_p = 1.1-1.5. The design rule is that the Nyquist curve of the compensated open-loop transfer function should avoid the interior of the circle associated with the specified M_p value and, instead, be tangent to it (see Figure 4.12).

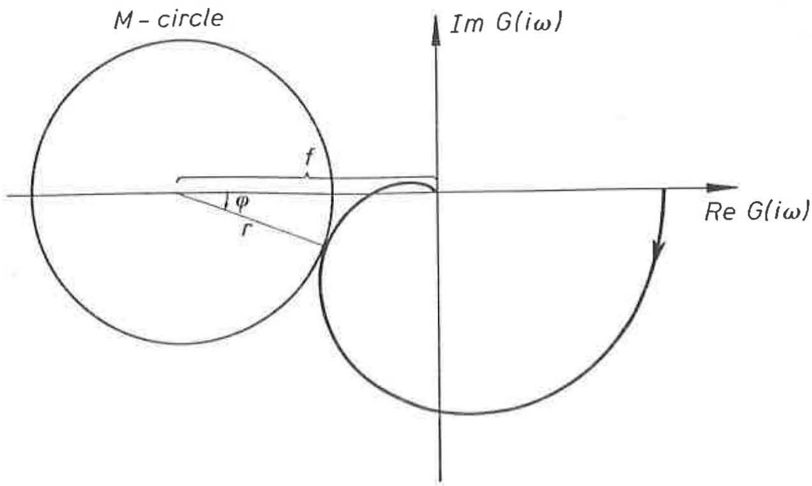

Figure 4.12
Graphical Illustration of the Design Procedure

The procedure can be described in some detail as follows. Let $G_P(i\omega)$ and $_R(i\omega)$ denote the transfer functions of the process and the controller. ssume that the open-loop frequency response of the process is measured at equency ω, i.e.,

$$_P(i\omega) = a + b$$

Also assume that the derivative of G_P is measured at the same frequency. his can be done by measuring G_P at two neighboring frequencies. Hence,

$$'_P(i\omega) = c + id$$

The transfer function of a PID controller is

$$_R(i\omega) = K\left[1 + i\left(\omega T_d - \frac{1}{\omega T_i}\right)\right]$$

ence,

$$''_R = iK\left[\; T_d \;\frac{1}{\omega^2 T_i}\right]$$

Let the point where the compensated Nyquist curve touches the M_P circle be specified by angle φ (see Figure 4.12). This point is then given by the complex number:

$$A = -f + r\cos\varphi - ir\sin\varphi$$

The open-loop transfer function of the compensated system is

$$G = G_P G_R$$

Requiring that the compensated Nyquist curve goes through A gives

$$G_P(i\omega)\, G_R(i\omega) = -f + r\cos(\varphi) - ir\sin(\varphi) \tag{4.14}$$

Separating the real and imaginary parts of this equation gives

$$K\left[a - b\left(\omega T_d - \frac{1}{\omega T_i}\right)\right] = -f + r\cos\varphi$$

$$K\left[b + a\left(\omega T_d - \frac{1}{\omega T_i}\right)\right] = -r\sin\varphi$$

The condition that the compensated Nyquist curve is a tangent to the M_P circle at A can be expressed as

$$\arg G' = \arg(G'_P G_R + G_P G'_R) = \frac{\pi}{2} - \varphi \tag{4.15}$$

This equation implies that

$$\tan\varphi = \frac{c - d\left(\omega T_d - \dfrac{1}{\omega T_i}\right) - b\left(T_d + \dfrac{1}{\omega^2 T_i}\right)}{d + c\left(\omega T_d - \dfrac{1}{\omega T_i}\right) + b\left(T_d + \dfrac{1}{\omega^2 T_i}\right)}$$

We thus obtain three conditions: two for positioning the point and one for the slope. Since point A can be positioned anywhere on the chosen M circle, one extra degree of freedom can be chosen as angle φ in Figure 4.12, thus obtaining three conditions to determine four parameters (K, T_i, T_d, and φ). An auxiliary condition is obtained from

$$\omega T_i = \alpha \tag{4.16}$$

where α is a number in the range 3-6. This requirement implies that the integral action acts at a time scale that is compatible with the bandwidth (ω). With this additional requirement, the design procedure gives unique values of the PID parameters.

Validation

It is important to test the validity of a design based on simplified assumptions. First notice that the given procedure is based on local properties of the Nyquist curve; hence, there is no guarantee that the Nyquist curve will remain outside the M_p disc globally.

Although it is not possible to guarantee the properties of a design without access to detailed models or experiments, several quantities can be computed to obtain indications of the validity of a design.

The dimensionless quantity ωT_d can be interpreted as the normalized prediction horizon. This quantity should be small for the prediction to be good. To obtain a number, we can observe that a straight line prediction of a sinusoid can be made with a precision of 10% if

$$\omega T_d < 0.8 \tag{4.17}$$

Another quantity of importance is the ratio T_i / T_d. The numerator of the regulator transfer function has zeros at

$$s = \frac{1}{2T_d}\left[-1 \pm \sqrt{1-4T_d/T_i}\right]$$

If T_i / T_d is too small, the zeros will have poor damping. Since the closed-loop poles will migrate towards the zeros, we will thus require that $T_i > T_d$. This condition is automatically guaranteed by Equations 4.16 and 4.17.

A third condition is that the quantity

$$\gamma = \arctan\left(\omega T_d - \frac{1}{\omega T_i}\right)$$

(which represents the phase shift in the controller) is of reasonable magnitude, say less than $\pi/3$. This condition is also guaranteed by Equations 4.16 and 4.17.

It can thus be concluded that it is practical to impose the conditions of Equations 4.16 and 4.17 since this will automatically guarantee that other important conditions hold.

Design Variables

The design variables are the frequency (ω) and the M_p value. Although M_p values close to one will give systems with good damping, there are several drawbacks in choosing too small a value, because the associated M circle will then have a large radius, and it is then a greater risk that the

Nyquist curve will enter it at some other frequency. With a large radius of the M circle, the design will also be more sensitive. Reasonable values are therefore in the range of 1.3 to 1.5. According to Table 4.5, this corresponds to a relative damping around 0.4.

The frequency (ω) is also a critical variable. Experience has indicated that it is sensible to choose a frequency where the Nyquist curve of the process is in the third quadrant.

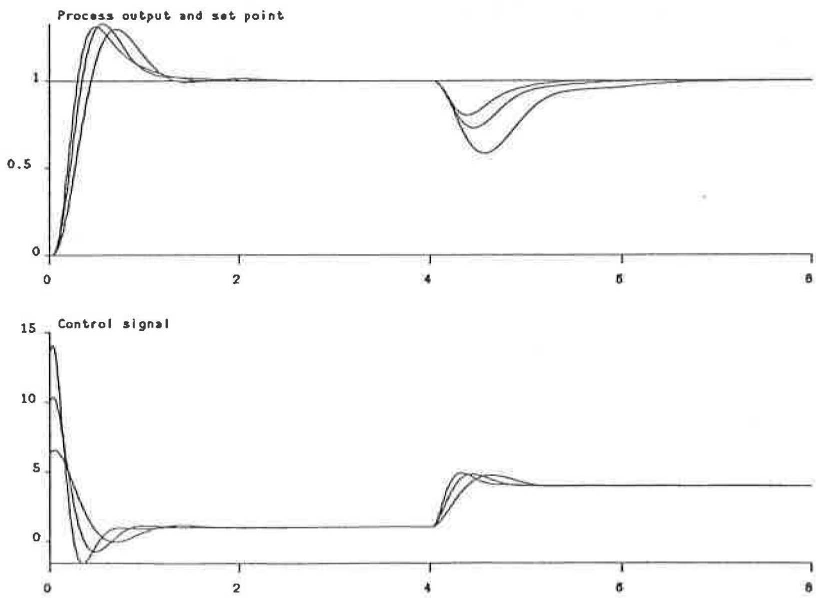

Figure 4.13
Step and Load Responses for the PID Controllers Obtained by the M circle Design Method. (The design parameters are M_p = 1.3 and ωT_i = 3. The responses for ω = 4, 5, and 6 rad/s are shown.)

xample 4.5—The design procedure described above can be illustrated ;ing the process model (Equation 4.1). The design parameters are chosen $M_p = 1.3$ and $\omega T_i = 3$. Solving the design equations, the following con-oller parameters for $\omega = 4$, 5, 6, and 7 are obtained:

ω	K	T_i	T_d	ωT_d	T_i/T_d
4	6.3	0.75	0.07	0.27	11.0
5	10.0	0.60	0.10	0.52	5.8
6	13.5	0.50	0.12	0.72	4.2
7	15.6	0.43	0.13	0.92	3.2

The PID controller obtained for $\omega = 4$ has phase lag. The value of ωT_d is little too high for $\omega = 7$. This indicates that PID control can be used for ındwidths up to 6 rad/s but not higher with the chosen M_p value. Figure 13 shows the responses of the regulators obtained for $\omega = 4$, 5, and 6 rad/s. the M_p value is increased to 1.5, a valid design can be obtained for $\omega = 7$ .d/s. The parameters are $K = 14.8$, $T_i = 0.43$ and $T_d = 0.105$. For this design, $T_d = 0.74$. A comparison with the previous results shows that the main fect of increasing M_p is that the derivation time decreases.

4.5 POLE PLACEMENT

ıe design methods presented previously in this chapter are all based on a nited knowledge of the process transfer function. Since the PID controller ıs only three design parameters, it cannot arbitrarily compensate more ımplicated process transfer functions. However, if the process is described / a low-order transfer function, a complete pole placement design can be :rformed, as described below.

PI Control of a First-Order System

Suppose that the process can be described by the following first-order odel:

$$_p = \frac{k_p}{1 + sT_1} \tag{4.18}$$

which has only two parameters, the process gain (k_p) and the time constant (T_1). By controlling this process with the PI controller,

$$G_R = K\left[1 + \frac{1}{sT_i}\right] \tag{4.19}$$

a second-order closed-loop system is obtained:

$$G_c = \frac{G_P G_R}{1 + G_P G_R} \tag{4.20}$$

The two closed-loop poles can be chosen arbitrarily by a suitable choice of the gain (K) and the integral time (T_i) of the controller. This is seen as follows. The poles are given by the characteristic equation, i.e., the equation

$$1 + G_P G_R = 0$$

The characteristic equation becomes

$$s^2 + s\left(\frac{1}{T_1} + \frac{k_p K}{T_1}\right) + \frac{k_p K}{T_1 T_i} = 0 \tag{4.21}$$

Now suppose that the desired closed-loop poles are characterized by their relative damping (ζ) and their frequency (ω). The desired characteristic equation then becomes

$$s^2 + 2\zeta\omega s + \omega^2 = 0 \tag{4.22}$$

Making the coefficients of these two characteristic equations equal gives two equations for determining K and T_i:

$$\begin{cases} \omega^2 = \dfrac{k_p K}{T_1 T_i} \\[2ex] 2\zeta\omega = \dfrac{1 + k_p K}{T_1} \end{cases} \tag{4.23}$$

Hence, the following PI parameters are obtained:

$$\begin{cases} K = \dfrac{2\zeta\omega T_1 - 1}{\mathrm{k_p}} \\[2ex] T_i = \dfrac{2\zeta\omega T_1 - 1}{\omega^2 T_l} \end{cases} \tag{4.24}$$

otice that in order to have positive controller gains it is necessary that hosen bandwidth (ω) be larger than $1/(2\zeta T_1)$. Also notice that if ω is the integration time T_i is given by

$$\frac{2\zeta}{\omega}$$

is thus independent of the process dynamics for large ω. There is no al upper bound to the bandwidth. However, a simplified model like ation 4.18 will not hold for large frequencies. The upper bound on the lwidth is therefore determined by the validity of the model.

PID Control of Second-Order Systems

uppose that the process is characterized by the second-order model

$$\frac{k_p}{(1 + sT_1)(1 + sT_2)} \tag{4.25}$$

his model has three parameters. By using a PID controller, which also hree parameters, it is possible to arbitrarily place the three poles of the d-loop system. The transfer function of the PID controller can be en as

$$= \frac{K(1 + sT_i + s^2 T_i T_d)}{sT_i}$$

he characteristic equation of the closed-loop system becomes

$$s^2 \left[\frac{1}{T_1} + \frac{1}{T_2} + \frac{k_p K T_d}{T_1 T_2} \right] + s \left[\frac{1}{T_1 T_2} + \frac{k_p K}{T_1 T_2} \right] + \frac{k_p K}{T_i T_1 T_2} = 0$$

suitable closed-loop characteristic equation of a third-order system is

$$\alpha\omega)(s^2 + 2\zeta\omega s + \omega^2) = 0 \tag{4.26}$$

h contains two dominant poles with relative damping (ζ) and frequency and a real pole located in $-\alpha\omega$. Identifying the coefficients in these two acteristic equations gives

$$\begin{cases} \dfrac{1}{T_1} + \dfrac{1}{T_2} + \dfrac{k_p K T_d}{T_1 T_2} = \omega(\alpha + 2\zeta) \\ \dfrac{1}{T_1 T_2} + \dfrac{k_p K}{T_1 T_2} = \omega^2(1 + 2\zeta\alpha) \\ \dfrac{k_p K}{T_i T_1 T_2} = \alpha\omega^3 \end{cases} \tag{4.27}$$

These three equations determine the PID parameters K, T_i, and T_d. The solution is

$$\begin{cases} K = \dfrac{T_1 T_2 \omega^2(1 + 2\zeta\alpha) - 1}{k_p} \\ T_i = \dfrac{T_1 T_2 \omega^2(1 + 2\zeta\alpha) - 1}{T_1 T_2 \alpha\omega^3} \\ T_d = \dfrac{T_1 T_2 \omega(\alpha + 2\zeta) - T_1 - T_2}{\omega^2 T_1 T_2(1 + 2\zeta\alpha) - 1} \end{cases} \tag{4.28}$$

Notice that pure PI control is obtained for

$$\omega_c = \frac{T_1 + T_2}{(\alpha + 2\zeta)T_1 T_2}$$

Notice also that the choice of ω may be critical. The derivation time is negative for $\omega < \omega_c$. The frequency (ω_c) thus gives a lower bound to the bandwidth. Also notice that the gain increases rapidly with ω. The upper bound to the bandwidth is given by the validity of the simplified model (Equation 4.25).

Example 4.6—In this example, the model (Equation 4.1) is approximated with the second-order model

$$G_P(s) = \frac{1}{(1 + s)(1 + 0.26s)}$$

Here, the longest time constant of the model is kept, and the three shortest time constants are approximated with their sum. If $\zeta = 0.5$ and $\alpha = 1$ are chosen, the design calculation gives the following PID parameters:

$$K = 0.52\omega^2 - 1$$

$$T_i = \frac{0.52\omega^2 - 1}{0.26\omega^3}$$

$$T_d = \frac{0.52\omega - 1.26}{0.52\omega^2 - 1}$$

In this case, pure PI control is obtained for $\omega = 2.4$. The derivative gain omes negative for lower bandwidths. The approximation neglects the d time constant 0.05. If the neglected dynamics are required to give a ase error of, at most, 0.3 rad (17 deg) at the bandwidth, $\omega < 6$ rad/s can obtained. In Figure 4.14, the behavior of the control is demonstrated for = 4, 5, and 6. It is straightforward to apply the direct design approach sed on the simplified process models. The specification of the desired

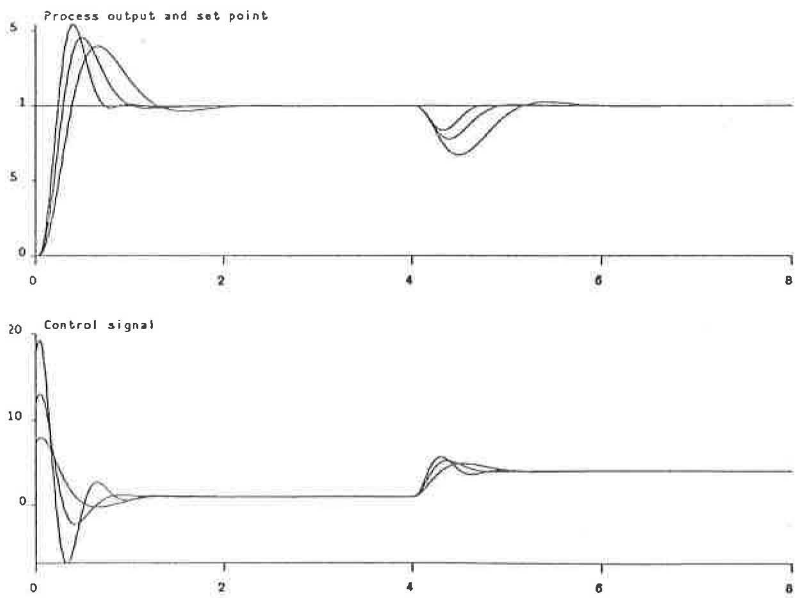

Figure 4.14
Step and Load Disturbance Responses of the Process (Equation 4.1) Controlled by a PID Controller Tuned According to Example 4.6 (The responses for ω = 4, 5, and 6 are shown.)

closed-loop bandwidth is, however, crucial since the controller gain will increase rapidly with the specified bandwidth. It is crucial to know the frequency range where the model is valid. Alternatively, an upper bound to the controller gain can be used to limit the bandwidth. Notice the effect of changing the design frequency (ω). The system with $\omega = 6$ responds faster and has a smaller error when subjected to load disturbances. Simulations indicate that the design will not work well when ω is increased above 8.

Cancellation of Process Poles

A particular class of design methods is based on the idea of choosing the parameters of the controller so that the dominant process poles are canceled. These methods are quite popular because they are very simple and give a good response to set point changes. They will, however, often give poor response to load disturbances.

To explain the methods, consider the transfer functions of a PI controller:

$$G_R(s) = k\left[1 + \frac{1}{sT_i}\right] = \frac{k(1 + sT_i)}{sT_i}$$

and an ideal PID controller with error feedback:

$$G_R(s) = k\left[1 + \frac{1}{sT_i} + sT_d\right] = \frac{k[1 + sT_i + s^2T_iT_d]}{sT_i}$$

One process pole can be canceled by a PI controller, and two process poles can be canceled by a PID controller. The response to load disturbances is poor for the designs based on cancellation because the dynamics corresponding to the canceled poles will appear in the response to the load disturbance. These modes will then recover in the same way as for the open-loop system. The same phenomena occur if the cancellation is not exact.

Example 4.7—PID design based on cancellation of process poles

Consider the system given by Equation 4.1. The system has the poles $p1 = -1$, $p2 = -1/0.2 = -5$, $p3 = -1/0.05 = -20$, and $p4 = -1/0.01 = -100$. Two of these poles can be canceled with a PID controller. Choosing the parameters T_i and T_d so that the slowest poles are canceled,

$$1 + sT_i + s^2T_iT_d = (1 + s)(1 + 0.2s) = 1 + 1.2s + 0.2s^2$$

This gives $T_i = 1.2$ and $T_d = 0.167$. To find a suitable value of the con-roller gain, proceed as in the direct pole placement method in Example 4.6. he compensated transfer function becomes

$$G_R(s)G_P(s) = \frac{k}{sT_i(1 + 0.05s)(1 + 0.01s)} \approx \frac{k}{sT_i(1 + 0.06s)}$$

The characteristic equation of the closed-loop system is, therefore,

$$(1 + 0.06s) + \frac{k}{T_i} = 0$$

r

$$s^2 + 16.7s + \frac{16.7k}{T_i} = 0$$

Identifying this with the characteristic equation

$$s^2 + 2\zeta\omega s + \omega^2 = 0$$

ives

$$k = \frac{4.2T_i}{\zeta^2} = \frac{5.0}{\zeta^2}$$

Choosing a relative damping $\zeta = 0.7$, then $k = 10$ and $\omega = 11.7$. Figure 4.15 hows the response of the closed-loop system obtained with these controller arameters. For comparison, the following results are obtained with a pole lacement controller without cancellation. This controller has the parame-ers $k = 12$, $T_i = 0.37$, $T_d = 0.11$, $N = 10$ and $b = 0.35$. Notice the fast response o command signals and the poor response to load disturbances. Also notice he "spike" in the control signal, which depends on the fact that the deriva-ive acts on the reference signal. The error due to a load disturbance decays vith a time constant of 1s, which corresponds to the cancelled mode $p1 = -1$ f the open-loop system. Because of the cancellation, the controller will not ttempt to control this mode. This is clearly seen in the fact that the control ignal settles much faster than the error at the load disturbance.

Although the designs based on cancellation of process poles are simple, hey will not be discussed further because of their poor performance when ubjected to load disturbances.

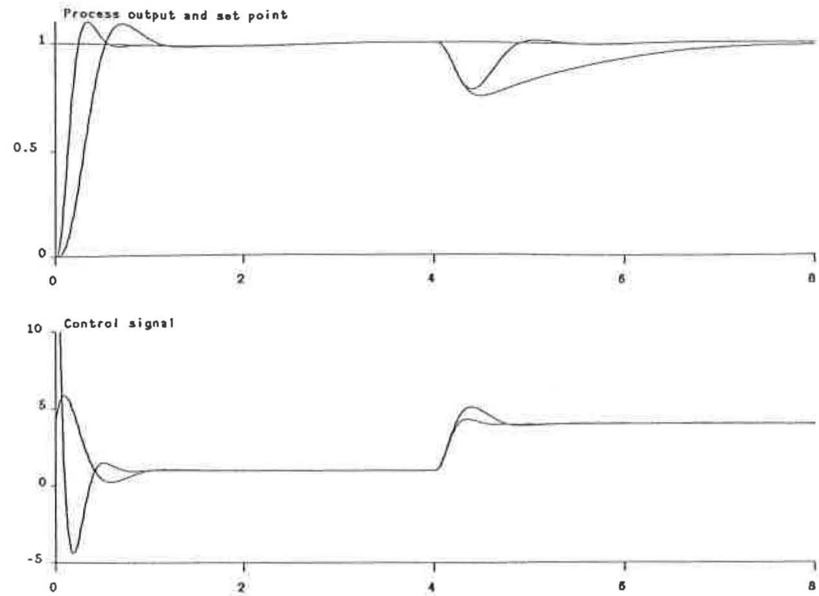

Figure 4.15
Simulation of PID Controller Based on Cancellation of Process Poles (For comparison, see the results of an equivalent design without cancellation shown by the thin lines.)

4.6 DISCRETE TIME POLE PLACEMENT

The examples have shown that PID controllers can be used for pole placement design when the process model is of low order. In the examples, continuous time models of the controller and the process have been used. It is also possible to use discrete time controllers for the pole placement design of discrete time process models, as shown below.

In Section 3.4, a discrete time process model was introduced using the z-transform instead of the Laplace transform used in continuous time models. Let the process be described by the transfer function

$$H_p(z) = \frac{B(z)}{A(z)} \tag{4.29}$$

Let $U(z)$ and $E(z)$ denote the z-transforms of the control signal, $u(t)$, and he error signal, $e(t)$. A general description of the controller is then

$$R(z)U(z) = S(z)E(z)$$

The transfer function of the controller can be written as

$$H_R(z) = \frac{U(z)}{E(z)} = \frac{S(z)}{R(z)} \tag{4.30}$$

The closed-loop transfer function is given by

$$H_C(z) = \frac{H_P H_R}{1 + H_P H_R} \tag{4.31}$$

and the characteristic equation therefore becomes

$$1 + H_P H_R = 0$$

Using Equations 4.29 and 4.30, the characteristic equation can also be written as

$$A(z)R(z) + B(z)S(z) = 0 \tag{4.32}$$

Now suppose that the process is of second order with the following transfer function polynomials:

$$A(z) = z^2 + a_1 z + a_2$$

$$B(z) = b_1 z + b_2$$

This structure of the process model captures many processes common in the process controller and is, for example, obtained by sampling the continuous time model (Equation 4.25) in the previous section. To ensure that the controller has integral action, the R-polynomial must be of the form

$$R(z) = (z-1)R_1(z)$$

The controller polynomials are given on the general forms

$$S(z) = s_0 z^2 + s_1 z + s_2$$

$$R(z) = (z - 1)(z + r_1)$$

Thus, the characteristic equation is obtained:

$$(z^2 + a_1 z + a_2)(z - 1)(z + r_1) + (b_1 z + b_2)(s_0 z^2 + s_1 z + s_2) = 0 \tag{4.33}$$

which is of fourth order. Assume that the desired closed-loop characteristic polynomial is given by

$$P(z) = (z - e^{-\alpha\omega h})^2(z^2 + p_1 z + p_2) \tag{4.34}$$

where

$$P_1 = -2\, e^{-\zeta\omega h} \cos\left(\omega h\sqrt{1 - \zeta^2}\right)$$

$$P_2 = e^{-2\zeta\omega h}$$

This corresponds to a fourth-order system having two dominant poles with relative damping (ζ) and frequency (ω), and two real poles located in $-\alpha\omega$.

The controller parameters can now be determined from the two descriptions of the characteristic equation, Equations 4.33 and 4.34. By comparing terms of equal power of z, parameters r_1, s_0, s_1, and s_2 can be determined, as illustrated in the following example. A detailed presentation of the discrete time design method is given in the book by Åström and Wittenmark (1984).

Example 4.8—In Example 4.6, the fourth-order model (Equation 4.1) was approximated by the second-order model:

$$G_P(s) = \frac{1}{(1 + s)(1 + 0.26s)}$$

If this model is sampled with the sampling period $h = 0.1s$, the following discrete time model is obtained:

$$H_p(z) = \frac{0.0164z + 0.0140}{z^2 - 1.583z + 0.616}$$

If the design parameters are $\zeta = 0.5$, $\omega = 4$, and $\alpha = 1$, the desired characteristic polynomial becomes

$$(z - 0.670)^2(z^2 - 1.54z + 0.670$$

Comparing this characteristic polynomial with the one obtained according to Equation 4.33, the following set of controller parameters is obtained:

$$\begin{cases} r_1 = -0.407 \\ s_0 = 6.74 \\ s_1 = -9.89 \\ s_2 = 3.61 \end{cases}$$

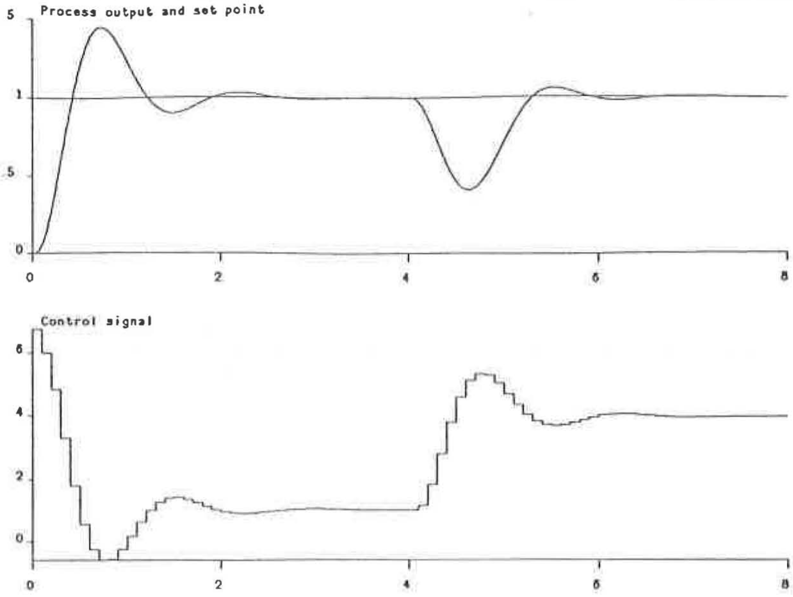

Figure 4.16
ɐp and Load Disturbance Responses of the Process (Equation 4.1) Controlled by a PID Controller Tuned According to Example 4.8

In Figure 4.16, the behavior of the control is demonstrated. Although the in is fairly high (see the control signal), the response to the load distur-ınce is quite slow because of the low value of ω.

A drawback with direct digital design is that it is normally difficult to anslate the controller to PID structure. The structure of the controller ed in this section is such a case. On the other hand, this general form is eful when trying to cope with problems that are hard to solve with the andard PID controller. Such an example is dead time compensation, here a suitable controller can be derived just by introducing the dead time the process model $H_P(z)$.

4.7 IMPROVEMENT OF SET POINT CONTROL

The controllers simulated in this chapter have responses to set point changes with excessive overshoot. Typical examples are given in Figures 4.2, 4.8 and 4.16. The reason for this is that the standard form of the PID controller with error feedback is used. The transfer function between the set point and the control signal of a PID controller is

$$G_{\text{PID}}(s) = K \frac{1 + sT_i}{sT_i}$$

The derivative part does not occur, since the derivation is performed on the process output only. The controller introduces a closed-loop zero at

$$s = -\frac{1}{T_i} \tag{4.35}$$

The influence of this zero was discussed in Section 2.4, where it was proposed to use a modified PID controller where only a fraction (b) of the reference signal is introduced in the proportional part. Such a controller is described by Equation 2.9, i.e.,

$$u = K\left[e_p + \frac{1}{T_i}\int_0^t e(s)ds + T_d \frac{de_d}{dt}\right] \tag{4.36}$$

where the error in the proportional part is

$$e_p = br - y$$

and the error in the derivative part is

$$e_d = -y$$

and the error in the integral part is

$$e = r - y$$

The modified controller has a zero at

$$s = -\frac{1}{bT_i}$$

which can be positioned properly by choosing the parameter b suitably. An estimate of the dominant closed-loop poles is necessary to do this. To avoid an excessive overshoot, parameter b should be chosen so that the zero is two to three times larger than the magnitude of the dominant poles. Estimates of the dominant poles are available for many of the design methods.

The Ziegler-Nichols Method

the Ziegler-Nichols method, estimates of the dominant poles are ned from the estimate of the closed-loop dominant period. This is listed in Tables 4.1 and 4.2. The design rule given above then gives $b = 0.2$ Figure 4.17 shows simulations with the modified controller with $b = 0$, 0.3, and 1.0. The figure shows clearly that the overshoot is reduced ically when the modified algorithm is used. It also indicates that the n rule gives a reasonable value of b.

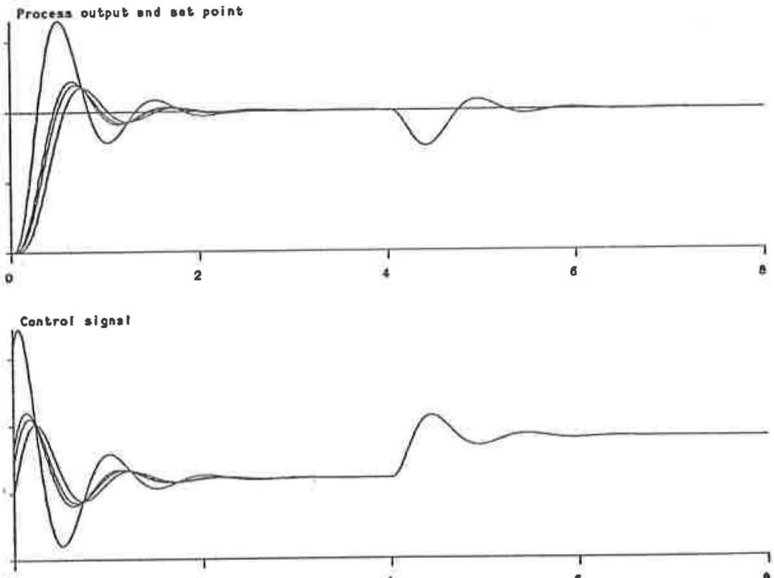

Figure 4.17
The Effect of Parameter b on the Step Response of a Closed-Loop System (The PID parameters are the same as in Figure 4.2.)

Direct and Dominant Pole Designs

In direct and dominant pole design methods, it is very easy to find good values for parameter b since these design methods deal directly with the dominant poles. Consider, for example, the direct design method used in Example 4.6 with $\omega = 6$ rad/s, which gives $T_i = 0.32$. The ordinary PID controller gives a zero at $s = 3.1$ rad/s, which is smaller than ω. To have the zero at $s = -12$, parameter b should be smaller than 0.26. To have the zero at $s = -18$, $b = 0.18$ should be chosen. Figure 4.18 shows a simulation of the modified PID controller. The figure shows clearly that the overshoot is reduced drastically when the modified algorithm is used. It also indicates that the rules for choosing parameter b are reasonable.

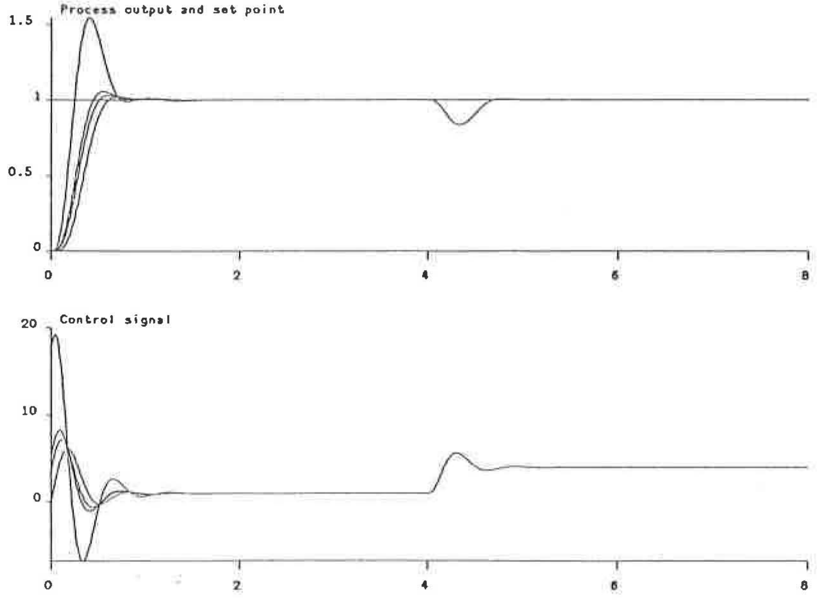

Figure 4.18
The Effect of Parameter b on the Step Response of a Closed-Loop System (The PID parameters are the same as in Figure 4.12 for ω = 6 rad/s)

Even in the case of direct digital design it is possible to improve the ponses to set point changes. In Section 4.6 the controller structure was en as

$$z)U(z) = S(z)E(z)$$

, error feedback was used. If the extended structure

$$z)U(z) = -S(z)Y(z) + T(z)Y_R(z)$$

used, the response to set point changes can be modeled by choosing lynomial $T(z)$ appropriately.

Conclusions

The results show conclusively that the responses to command signals are ıproved drastically by modifying the PID algorithm, as was discussed in :ction 2.4.

4.8 COMPARISONS OF DESIGN METHODS

though several methods have been given for designing PID controllers, all proaches have by no means been covered. There are many variations on : methods discussed herein, as well as a host of other techniques available the literature. Instead of going on to describe more methods, it is a good :a to provide some perspective on the different methods. Before going into : details of the design methods, it can first be observed that control system sign involves many different aspects, such as process dynamics, load dis-rbances, measurement noise, nonlinearities, and sensitivity. In this investi-tion, the focus has been on dynamics and set point changes, which is often equate for the design of simple controllers.

Overview of the Approaches

The Ziegler Nichols Methods

These are simple approaches based on information on two parameters ıly, either L and a (which characterize the step response) or K_c and T_c ·hich characterize the frequency response).

The analysis leading to the dominant pole design indicates that it is not possible to give estimates of the closed-loop dominant poles from the knowledge of one point on the frequency response only. It can thus be concluded that there will always be a large uncertainty with design methods like the Ziegler-Nichols, which are based only on this information. All the other design methods discussed in this chapter use more information.

The Dominant Pole Design

The method is based on positioning two or three dominant poles. The method is based on knowledge of the plant transfer function at the dominant poles. Approximate methods based on knowledge of the frequency curves are also given. The dominant pole design method has one design parameter, namely the distance of the poles from the origin.

An interesting feature of the dominant pole design is that it gives ranges of the design parameter that are achievable with different controller types. This can be used to choose P, PI, PD, or PID control. We illustrate this point by an example.

Example 4.9—PI and PD Control of $(s+1)^{-3}$.

Consider a plant with the transfer function

$$G_P = \frac{1}{(s+1)^3}$$

Since the plant is of third order, it is clear that exact pole placement cannot be obtained with PI, PD, or PID control. First, consider PI control. Using the equation for the approximate dominant pole design, the following parameters are obtained:

$$k = \frac{\sigma(-4\omega^4+20\omega^2) + 3\omega^4 + 2\omega^2 - 1}{\omega^2 + 12\sigma^2 + 6\sigma + 1}$$

$$k_i = \frac{-\omega^6 + 2\omega^4 + 3\omega^2 - 12\sigma(\omega^4-\omega^2)}{\omega^2 + 12\sigma^2 + 6\sigma + 1}$$

where $k_i = k/T_i$. PD control gives instead

$$k = \frac{-2\sigma\omega^4 + 3\omega^4 + 16\sigma\omega^2 + 2\omega^2 - 6\sigma - 1}{\omega^2 + 6\sigma^2 + 6\sigma + 1}$$

$$k_d = \frac{\omega^4 + 12\sigma\omega^2 - 2\omega^2 - 12\sigma - 3}{\omega^2 + 6\sigma^2 + 6\sigma + 1}$$

where $k_d = kT_d$. The controllers will have positive gains only if the specifications on the dominant poles are restricted to certain values. Figure 4.19 shows the combinations of σ and ω that give positive gains for the PI and the PD controllers, respectively. The border lines are given by the pure P, I, and D controllers. Notice that the approximative formulas are only valid if $\sigma < \omega$. From this figure it is seen that the bandwidth ω cannot be chosen too high if only a PI controller is used.

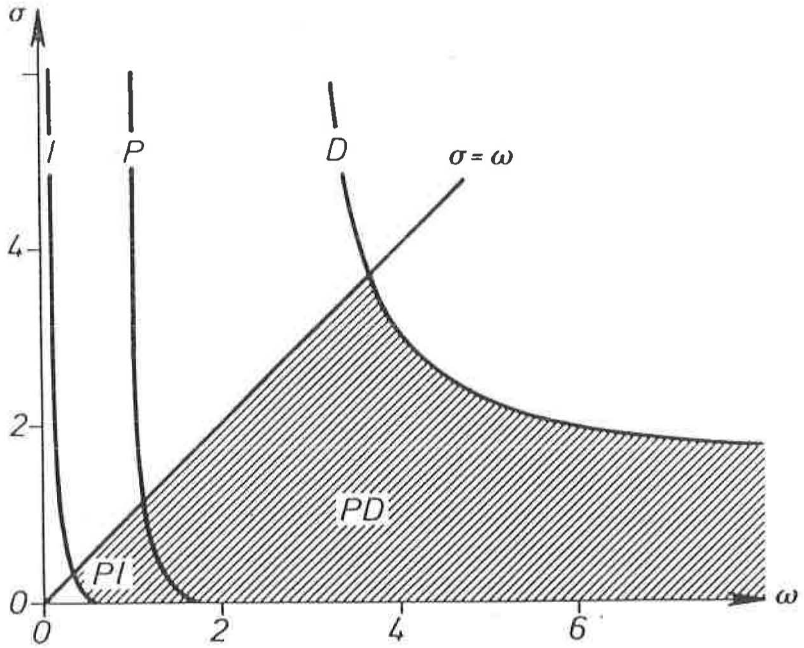

Figure 4.19
Regimes of Positive Gains for PI and PD Controllers

Simplified Frequency Domain Methods

This method is also based on knowledge of two *properly chosen* values of the open-loop frequency response of the system. The design method attempts to shape the closed-loop gain locally at the chosen frequencies.

When using the simplified frequency domain designs, it is clear that there are limitations on the shaping that can be done with a PID controller. It is thus necessary that the crossover frequency be chosen properly so that the loop can be shaped with a PID controller. It is also clear that the loop gain may behave badly at frequencies away from the chosen frequencies. This indicates that there will be problems with systems with resonances where the Nyquist curve twists and bends.

Pole Placement Methods

In the direct design methods, the dynamics are approximated by simplified models of first or second order, and the PID parameters are calculated from specifications on the desired closed-loop poles. The methods rely on making appropriate approximations and on the specifications being in harmony with these approximations.

Insight into the Problem

The direct design methods indicate superficially that any specification can be achieved. A closer inspection reveals, however, strict limitations. To obtain positive controller gains, it is necessary to choose the frequency (ω) sufficiently small (see Section 4.5). The formula for the controller gain also indicates that the gain will increase very rapidly with the chosen frequency. The frequency must also be chosen so low that the simplified model is valid well over ω. Experiments with continuous time and discrete time designs indicate that there is no large difference for small sampling periods. For longer sampling periods, the response to load disturbances will, however, be poorer for the discrete time algorithms because there will always be a time delay before the disturbance is captured.

Detailed Comparison

Table 4.6
Controller Parameters Obtained by the Different Design Methods

Method	K	T_i	T_d
Ziegler-Nichols step	10.9	0.32	0.08
Ziegler-Nichols frequency	15.0	0.31	0.08
Dominant pole design,			
$\omega = 5.3$	11.9	0.45	0.12
approximate method	14.2	0.41	0.10
M circle design			
$\omega = 4$	6.3	0.75	0.07
$\omega = 5$	10.0	0.60	0.10
$\omega = 6$	13.5	0.50	0.12
Direct pole placement			
$\omega = 4$	7.3	0.44	0.11
$\omega = 5$	12.0	0.37	0.11
$\omega = 6$	17.7	0.32	0.10
Direct pole placement with cancellation			
$\omega = 11.7$	10.0	1.20	0.17

Table 4.6 shows the parameters obtained when the different design thods are applied to the same problem. Several observations can be made m the table. First, with exception of the method based on cancellation, controller parameters obtained by the different methods are similar. For ample, the Ziegler-Nichols frequency domain method gives parameters t are quite close to the parameters obtained by the dominant pole design thod. The main difference is that the gain of the Ziegler-Nichols method oo high and the derivation time is too low. Another interesting fact is that Ziegler-Nichols method estimates the dominant frequency to be 12 d/s, which is much too high. Also notice that the dominant pole design es a value of the bandwidth ($\omega = \omega_0\sqrt{1 - \zeta^2}$), but that ω has to be chosen the designer for the direct pole placement.

Sensitivity

The comparison of the parameters obtained by the different methods indicates that the design may be quite sensitive to parameter variations. To investigate this, the parameters are perturbed in the Ziegler-Nichols frequency domain design. Figure 4.20 shows what happens when the derivation time is changed from T_d = 0.08 to 0.10 and 0.12. The figure indicates clearly that drastic improvements in the damping can be achieved by increasing the derivation time by 25%. Notice that the overshoot can be reduced drastically, as discussed in Section 4.7. This possibility was not used herein, because it is easier to see the improved damping with a large overshoot.

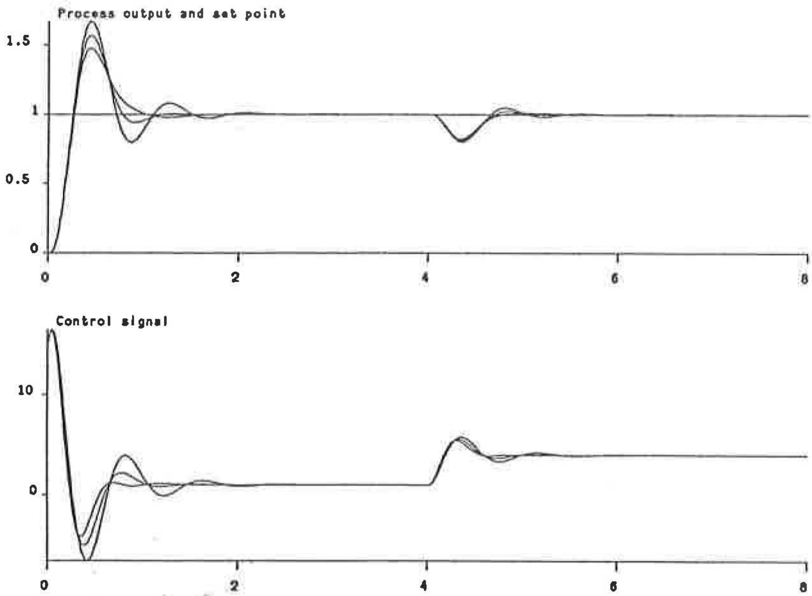

Figure 4.20
Effect of Changing the Derivation Time T_d in the Controller Obtained by the Ziegler-Nichols Frequency Domain Method

-igure 4.20 indicates that the system is sensitive, the reason being that the ;ed-loop bandwidth is quite high. It is a general rule that high bandwidth :ems are sensitive. The fact that the bandwidth is high can be concluded n the comparison with the direct pole placement method. The analysis formed in Section 4.5 indicated that $\omega = 6$ rad/s was on the high side and t a more reasonable value is $\omega = 4$ rad/s. This is illustrated in Figure 4.21 l Figure 4.22, which illustrate the sensitivity of the direct designs for $\omega = 4$ /s and $\omega = 6$ rad/s to changes in the controller parameters. The deriva- ı time (T_d) is changed by the same amount in both cases. Notice the stic influence in particular on the closed-loop period and damping in ure 4.22. The simulations strongly support reducing sensitivity by reduc- the bandwidth.

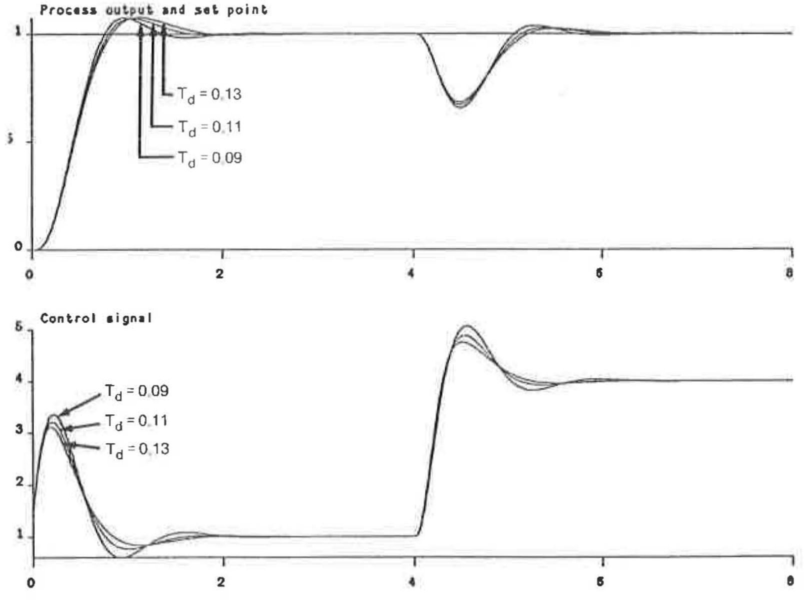

Figure 4.21
Effect of Changing Derivation Time T_d in the Controller Obtained by the Direct Pole Placement Design Method for $\omega = 4$ rad/s

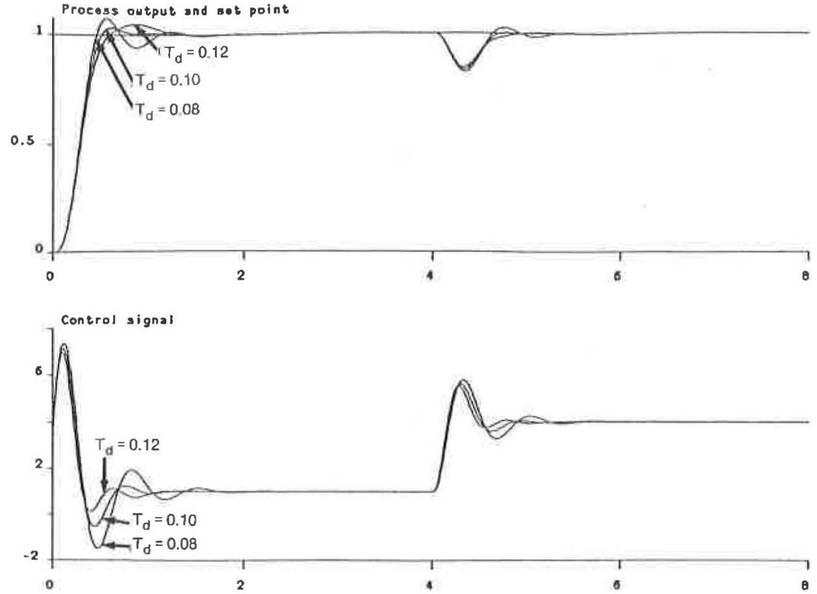

Figure 4.22
Effect of Changing Derivation Time T_d in the Controller Obtained by the Direct Pole Placement Design Method for ω = 6 rad/s

Conclusions

For processes with simple dynamics, it has been demonstrated that it is possible to find design methods that give good results. Some insight into the properties of different design methods have been developed. In particular, the desired closed-loop bandwidth has been found to be a crucial specification; too high a bandwidth gives excessive gain and a sensitive system. With a controller like the PID, which has restricted complexity, it may not be possible to achieve the desired bandwidth. The choice of the bandwidth thus emerges as a key issue.

It would be highly desirable to have a procedure that would allow determination of an appropriate bandwidth automatically. The ultimate frequency is a good starting value, but the analysis of the Ziegler-Nichols tuning procedures indicates that this frequency may be too high.

Since a PID controller has a limited complexity, it is clear that arbitrarily large values of ω cannot be chosen. This is also clearly illustrated in the examples. It is also clear that the approach will always work for open-loop stable systems if ω is chosen sufficiently low.

The dominant pole design gives a suitable value of ω directly. The following guidelines are useful for design methods where ω has to be chosen. The open-loop crossover frequency (ω_c) can serve as a first approximation. The phase lead generated by a PID controller depends on the ratio $\alpha = T_d / T_i$ and the maximum derivative gain. With $\alpha = 0.25$, the largest lead is approximately 40°. This means that a proper phase margin may be obtained with $\omega = \omega_c$. To obtain a good transient response it is, however, also necessary that the slope $d \log | G(i\omega) | / d \log(\omega)$ is close to −1 at the crossover. Evaluation of the slope at the open-loop crossover frequency indicates whether the crossover frequency can be chosen as ω. There is again some margin. A PID controller can, for example, increase the slope by at most 0.4 when $\alpha = 0.25$. If the slope conditions can not be satisfied, a lower value of ω must be chosen.

Evaluating how rapidly the phase and the amplitude change also indicates whether the system is minimum phase. For a system with pure time delay, for example, the slope of the amplitude curve at the crossover is zero. To obtain a proper slope of the amplitude curve, it is then necessary to introduce PI control. The integration time should be chosen so that the integral action dominates at crossover. This means that derivative action is useless and that the time delay should give a phase shift of about 90° at the crossover.

4.9 CONCLUSIONS

In this chapter, several approaches to design PID controllers have been presented. The design methods are based on the different process models given in the previous chapter. The derivation of the process model and the design calculation are closely related. All design methods reguire a model of the process to be controlled. As has been seen, different design procedures are based on different process characterizations. Using a complex design procedure such as the full pole placement design requires a transfer function description of the process with a high accuracy. The simple Ziegler-Nichols methods are based on limited process knowledge. If this kind of simple design procedure is used, there is no reason for making much effort in creating a detailed process model. Several design methods have been omitted in order to focus on those methods commonly used in the automatic tuning procedures. The process models described in Chapter 3 and the design methods presented in this chapter form the basis for the autotuning methods to be discussed in Chapter 5.