

Chapter 8

Design of Conventional Regulators



The aim of controller design is to ensure that the closed-loop control system meets the quality specifications. A block diagram of a control system is shown in Fig. 6.1. This structure is called a series control structure, as the regulator (sometimes called controller) is serially connected to the plant. The regulator is designed considering the properties of the plant and the specifications.

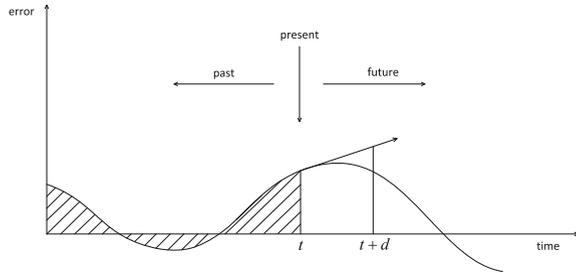
After the choice of the control structure, the parameters of the regulator have to be set. In the frequency domain with the appropriate choice of the regulator the frequency function of the open-loop is shaped according to the quality specifications (Fig. 6.2).

In Chap. 7 we could see, that for stable processes, precise theoretical methods are available to determine the optimal structure and the optimal parameters of the regulator for different cases. But already long before the elaboration of these theoretical methods, a well established class of control equipment was widely used in the control of industrial processes. This type of regulators has its determining significance even today. The development of the technology of electronic devices—which made possible the realization of more and more complicated transfer functions—played a significant role in the development of these conventional or classical regulators. Passive R-L-C circuits and precise operational amplifiers provided the technological basis for the development of the simple so-called *PID* regulator family. (In the case of regulators using mechanical, pneumatic, etc., signals only some restricted forms were realized, mainly because of realization constraints.)

PID regulators react proportionally to the current error value, take into consideration the past error signal history with the integral of the error signal, and count the future error signal trend by the differential of the error. Figure 8.1 shows [1] that the *PID* regulator calculates the manipulated variable (the control signal) with the *P* effect which is proportional to the error, with the *I* effect, the integral of the error, and the *D* effect, the differential of the error.

Application of *PID* regulators is quite general: more than 90% of the realized industrial control systems work with this type of regulator. In industrial process control the most frequently applied controllers are the *PID* regulators, as the quality

Fig. 8.1 The *PID* regulator calculates the manipulated variable from the current, the past and the future trend (the slope) of the error signal



specifications mostly can be fulfilled applying them, they have an easily realizable simple structure, and the effect of the parameter changes can easily be evaluated. The design of *PID* regulators can be discussed as a design method in the frequency domain or as a pole-zero cancellation technique.

8.1 The *PID* Regulator Family and Design Methods

The transfer function of the ideal *PID* regulator can be given in the following two forms:

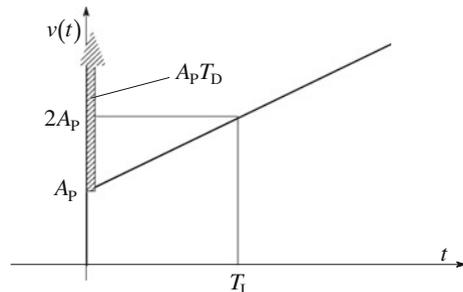
$$C_{PID} = A_P \left(1 + \frac{1}{sT_I} + sT_D \right) = A_P + k_I \frac{1}{s} + k_D s = \frac{A_P}{T_I} \frac{1 + sT_I + s^2 T_I T_D}{s} \quad (8.1)$$

Here the regulator parameters are A_P , the proportional transfer gain, T_I the integrating time constant, and T_D the differentiating time constant. The unit step response of the regulator is expressed as

$$v(t) = \mathcal{L}^{-1} \left[\frac{1}{s} C_{PID}(s) \right] = A_P + \frac{A_P}{T_I} t + A_P T_D \delta(t) \quad t \geq 0, \quad (8.2)$$

which is seen in Fig. 8.2.

Fig. 8.2 Unit step response of the ideal *PID* regulator



The *PID* regulator consists of parallel connected proportional (*P*), integrating (*I*) and differentiating (*D*) effects. The analytical expression (time function) of the ideal regulator, that is the operation executed on the error signal, $e(t)$ is given by

$$u(t) = A_P e(t) + k_I \int_0^t e(\tau) d\tau + k_D \frac{de(t)}{dt}. \tag{8.3}$$

This expression clearly shows the mentioned characteristics of the three regulator channels.

The regulator has a pole at the origin. It also has two zeros, which in the case of $T_I \geq 4T_D$, are located on the negative real axis. This condition requires a significant separation of the integrating and differentiating time constants, which requires a fourfold distance at least of the corresponding breakpoints in the amplitude-frequency diagram.

The location of the poles and zeros of the *PID* regulator in the complex plane is shown in Fig. 8.3. Its asymptotic BODE amplitude-frequency and phase-frequency diagram is seen in Fig. 8.4. The ideal *PID* regulator is non-realizable, it is used only for theoretical considerations and explanations. By ensuring the realizability of the *D* effect (combining it with a serially connected lag element with a small time constant) an approximate, but realizable *PID* regulator is obtained.

The transfer function of the approximate *PID* regulator can be given in the following form:

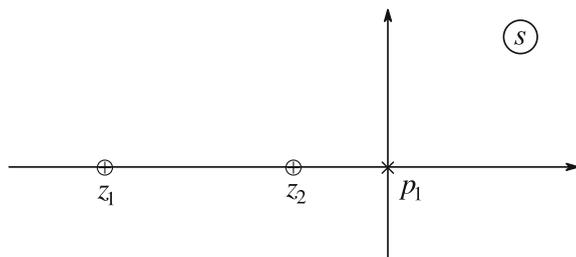
$$\widehat{C}_{PID} = A_P \left(1 + \frac{1}{sT_I} + \frac{sT_D}{1+sT} \right) = \frac{A_P}{T_I} \frac{1 + s(T_I + T) + s^2 T_I(T_D + T)}{s(1+sT)}. \tag{8.4}$$

The unit step response is expressed as

$$\widehat{v}(t) = \mathcal{L}^{-1} \left[\frac{1}{s} \widehat{C}_{PID}(s) \right] = A_P + \frac{A_P}{T_I} t + \frac{A_P T_D}{T} e^{-t/T} \quad t \geq 0 \tag{8.5}$$

and is shown in Fig. 8.5. The location of the poles and zeros of the approximate *PID* regulator is shown in Fig. 8.6. The asymptotic BODE amplitude-frequency and phase-frequency diagram is given in Fig. 8.7.

Fig. 8.3 Poles and zeros of the ideal *PID* regulator



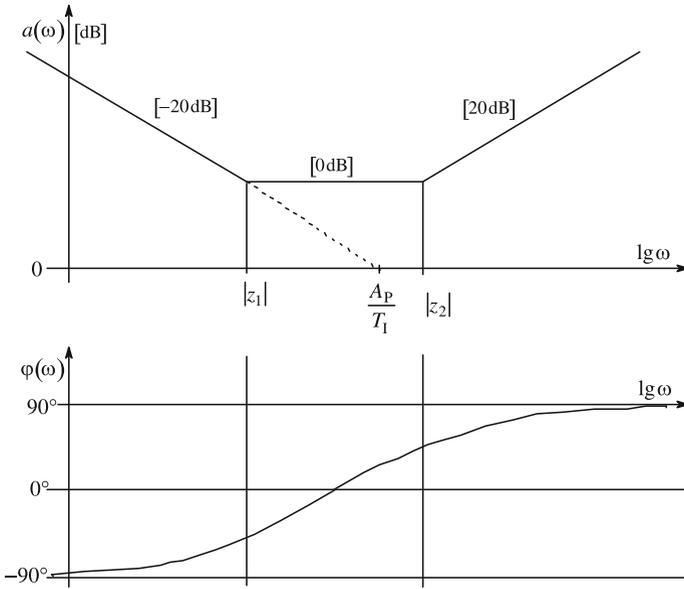


Fig. 8.4 Asymptotic Bode diagram of the ideal *PID* regulator

Fig. 8.5 Unit step response of the approximate *PID* regulator

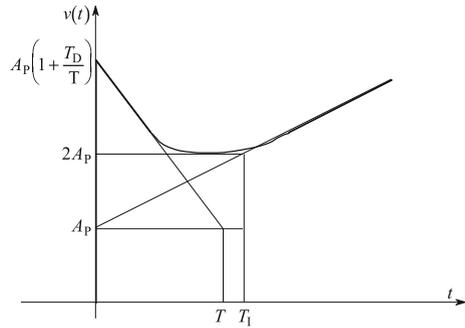
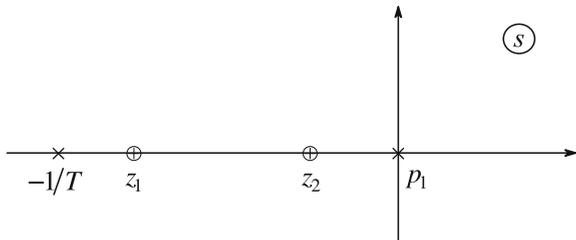


Fig. 8.6 Poles and zeros of the approximate *PID* regulator



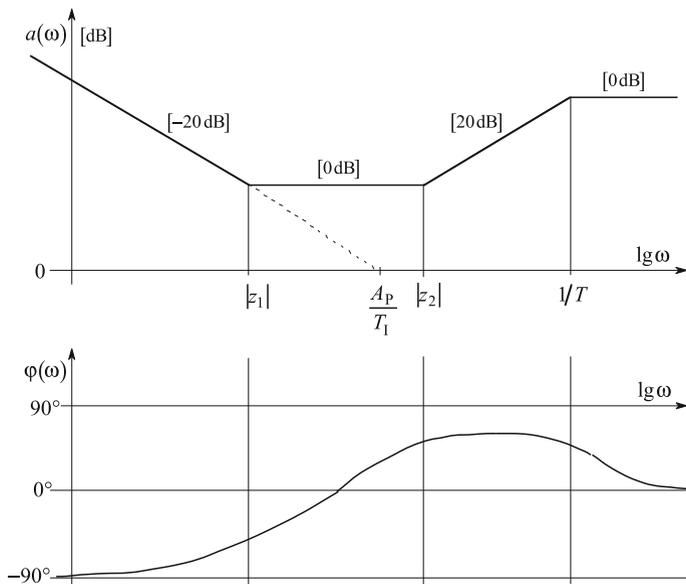


Fig. 8.7 Asymptotic Bode diagram of the approximate *PID* regulator

Now the regulator has two poles, one at the origin and the second at $-1/T$ in the complex plane, and its two zeros located on the negative real axis if $T_D \leq (T_I - T)^2/4T_I$. Again, this condition requires a significant separation of the integrating and the differentiating time constants, that is, a significant distance between the corresponding breakpoint frequencies.

The unit step response of the approximate *PID* regulator does not now start with the $\delta(t)$ (Dirac delta) function at $t = 0$, nevertheless the value of the initial jump expressed by $A_P(1 + T_D/T)$ may exceed the linearity range of the actuator and the equipment could be saturated. This situation is to be avoided, as on the one hand the normal operation mode of the actuator is ensured only in the linear range, and on the other hand the tuning of the regulator for stable performance, for the prescribed quality specifications, accuracy, etc., ensures the required behavior only in the case of linear models. The ratio T_D/T which determines the initial jump is called the overexcitation. Its value in real control systems should not exceed the limit of $T_D/T \leq 10$, and frequently should be even lower, with an allowable upper value of $4 \leq T_D/T \leq 6$. This realization limit determines first of all the fastest reachable transient (the cut-off frequency) for a control system of a given process.

Note that a modified version of *PID* regulators is often used, which significantly attenuates the effect of a possible abrupt change of the reference signal on the output signal in the closed-loop control system through a simple mechanism. In these regulators the manipulated variable—instead of the usual relationship (8.3)—is calculated by the expression $u(t) = A_P e_p(t) + k_I \int_0^t e(\tau) d\tau + k_D de_d(t)/dt$, where

$e_p(t) = br(t)$ and $e_d(t) = cr(t)$, setting parameters b and c whose values generally are between 0.2 and 0.8.

From the ideal *PID* regulator, the following different types of regulators can be obtained:

$$\begin{array}{ll}
 P & A_P \\
 I & \frac{1}{sT_I} = \frac{K_I}{s} \\
 PI & A_P \left(1 + \frac{1}{sT_I} \right) \\
 PD & A_P(1 + sT_D)
 \end{array} \quad (8.6)$$

In the approximate *PID* regulator, the *PD* effect is considered by the transfer function

$$PD \quad A_P \frac{1 + sT_D}{1 + sT}. \quad (8.7)$$

This element is also called a phase-lead or phase-lag element, as in the case of $T_D > T$ it realizes an approximate differentiating (*PD*), whereas if $T_D < T$, it provides an approximate *PI* effect.

In many practical cases the parameterization of the approximate *PID* regulator can be given in the following form:

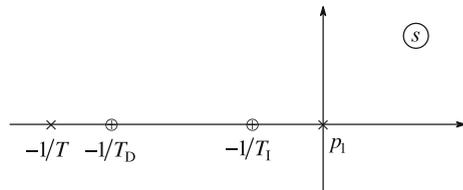
$$\widehat{C}_{PID} = A_P \frac{(1 + sT_I)(1 + sT_D)}{sT_I(1 + sT)}. \quad (8.8)$$

The advantage of this form is that it locates the breakpoint frequencies exactly at $1/T_I$, $1/T_D$ and $1/T$, which belong to the given integrating, differentiating time constants and to the time constant of the lag element. The location of the poles and the zeros is seen in Fig. 8.8.

Now the poles of the regulator are at the origin ($p_1 = 0$) and at $p_2 = -1/T$ and the zeros are at $-1/T_I$ and $-1/T_D$.

The *PID* regulator can also be considered as a general second order regulator, which has sufficient degrees of freedom to provide an appropriate solution for many simple control applications.

Fig. 8.8 Location of the poles and zeros of the approximate *PID* regulator according to (8.8)



8.1.1 Tuning of P Regulators

The P regulator ($C_P(s) = A_P$) generally results in a 0-type control system (except if the process itself contains an integrating effect). The 0-type control system has a finite steady error with a value of $1/(1+K)$, and the gain of the closed-loop transfer function is $K/(1+K)$, where $K = A_P P(0)$ is the gain of the open-loop (and not of the regulator!). It is also called the loop-gain and $P(0)$ is the gain of the process. In this case, when realizing the regulator, a static compensation factor (a calibration factor) with a value of $(1+K)/K$ is applied to ensure accurate reference signal tracking. The proportional regulator means the multiplication of the transfer function of the process by a constant factor, which influences only the amplitude of the frequency function, and has no effect on the phase characteristic: $|C_P(j\omega)| = A_P$; $\varphi(\omega) = \arg\{C_P(j\omega)\} = 0$.

By changing the value of the constant A_P the loop gain and thus the cut-off frequency can be set. The BODE amplitude diagram can be shifted parallel, thus the cut-off frequency can be changed to ensure the appropriate phase margin for the control system.

The control system can be stabilized by a P regulator, the overshoot of the unit step response can be kept within the permissible limit by setting the appropriate phase margin. As the cut-off frequency can not be increased significantly, the control system will work slowly. If the process does not contain integrating elements, the control system is of 0-type, which can track only the unit step reference signal. The control system works with a constant steady state error.

In industrial regulators, instead of the A_P gain of the proportional regulator, the so-called proportional band (PB) is set, which is defined as $PB = [1/A_P]100\%$. The proportional band is that range of the input signal in which the output signal runs through its entire range.

In the P regulator only one parameter, the A_P gain, can be set. This means that the prescribed static error and the prescribed phase margin can not always be ensured at the same time. These two prescriptions can be fulfilled together only in fortunate cases. In the usual design procedure first the maximal loop gain $K_{\max}(\varphi_{t0})$ is determined belonging to the prescribed phase margin φ_{t0} . The maximal reachable loop gain gives the minimal reachable static error

$$e_{\min} = \min\{e(s)|_{s=0}\} = 1/[1 + K_{\max}(\varphi_{t0})]. \quad (8.9)$$

If the task is given in a reverse way, that is, if K_{\min} has to be calculated from the maximal allowed error, then

$$K_{\min} = (1/e_{\max}) - 1; \quad e_{\max} < 1. \quad (8.10)$$

If the condition $K_{\min} < K_{\max}$ is fulfilled, then the twofold criterion has a solution, otherwise it does not. P compensation shifts the BODE amplitude-frequency diagram of the process parallel to the horizontal axis with the value of the regulator gain $A_P = K/P(0)$. The value of the shift initially is determined by K_{\max} or K_{\min} .

8.1.2 Tuning of I Regulators

The aim of applying an I regulator is to ensure a 1-type control system which provides zero steady state error in the case of a step reference signal. The design of the I regulator given by the transfer function

$$C_I = \frac{1}{sT_I} = \frac{K_I}{s} \quad (8.11)$$

is relatively simple, as it has only one parameter. Thus the maximal loop gain $K_{\max}(\varphi_{to})$ which provides the prescribed phase margin φ_{to} can be determined easily with the usual design methods, then $K_I = K_{\max}/P(0)$.

8.1.3 Tuning of PI Regulators

In the case of P regulators one could see that the allowable maximal loop gain $K_{\max}(\varphi_{to})$, whose value depends on the performance criterion generally is not high enough to ensure a small enough steady state error. To eliminate this problem PI regulators can be used, which ensure at least 1-type number for the control system, thus the steady state error will be zero in the case of a step-like reference or disturbance signal. The transfer function of the regulator is

$$C_{PI}(s) = A_P \left(1 + \frac{1}{sT_I} \right) = K_I \frac{1 + sT_I}{s}. \quad (8.12)$$

With the choice $T_I = \max\{T_i\} = T_1$, the largest time constant of the process can be cancelled. Then the maximal loop gain $K_{\max}(\varphi_{to})$ that provides the prescribed phase margin φ_{to} can be determined by the usual design methods, and then the integral gain of the regulator is calculated by $K_I = A_P/T_I = K_{\max}/P(0)$. The amplitude-frequency diagram of the open-loop $L(j\omega)$ starts with a slope of -20 dB/decade at low frequencies, as the PI regulator reallocates the smallest breakpoint frequency ($1/T_1$) belonging to the largest time constant of the process to the origin. Furthermore it shifts the amplitude-frequency diagram of the process parallel to the horizontal axis by K_I .

Based on the parallel form of the transfer function, the unit step response of the regulator can be drawn easily, whereas the product form provides the possibility of easily sketching the asymptotic BODE amplitude diagram. Figure 8.9 gives the unit step response and the BODE diagram of the PI regulator, the latter is plotted for $A_P = 1$. If the gain is different, the BODE amplitude diagram is shifted parallel.

PI regulators ensure a high gain in the low frequency domain. The integrating effect increases the type number of the control system by 1. In case of a proportional process, the static error will be zero for a step-like reference signal. By appropriately placing the cut-off frequency, a stable control system can be obtained with the required phase margin. As the cut-off frequency can not be put into the high frequency domain, the control system will be slow.

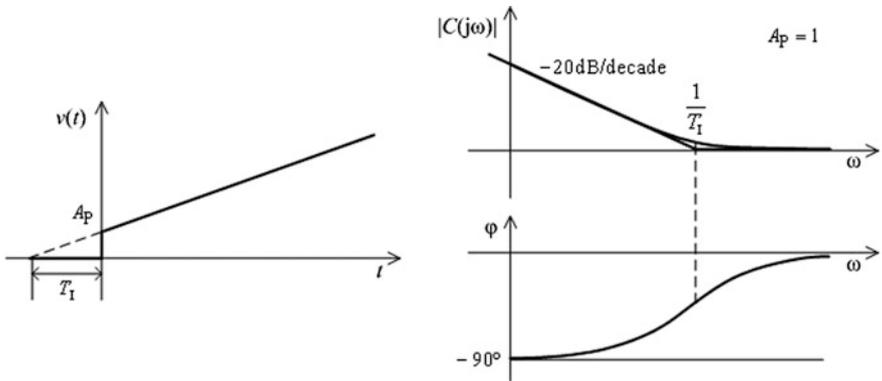


Fig. 8.9 The unit step response and the approximating BODE diagram of the PI regulator

An approximate realization of the PI regulator is the phase-lag element described by the transfer function

$$\tilde{C}_{PI}(s) = A_P \frac{1 + sT_I}{1 + sT} = C_{PL}(s), \tag{8.13}$$

where $T > T_I$. Now by the choice $T_I = \max\{T_i\} = T_I$ the smallest breakpoint frequency ($1/T_I$) is shifted left to the point $1/T$. The amplitude-frequency diagram of the loop transfer function $L(j\omega)$ starts now with 0 slope, thus the control system remains of 0 type. By applying \tilde{C}_{PI} the straight line section of slope -20 dB/decade becomes longer, thus a higher allowable $K_{\max}(\varphi_{10})$ gain can be ensured than with a simple P regulator. A single approximate PI regulator can reallocate a single breakpoint to a lower frequency.

Figure 8.10 gives the unit step response and the BODE diagram of the phase-lag element. It can be seen that this element approximates the characteristics of the PI element for small time instants in the time domain and in the high frequency range

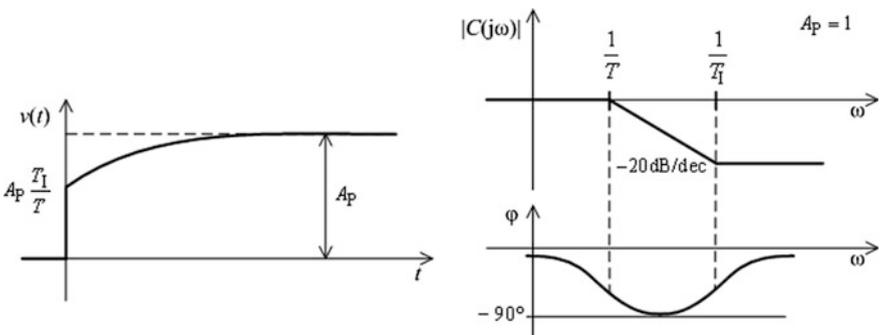


Fig. 8.10 Unit step response and the approximate BODE diagram of the phase-lag element

in the frequency domain. For sinusoidal inputs the output signal is phase-delayed related to the input signal. As this regulator does not contain an integrating effect, it does not ensure zero steady state error.

8.1.4 Tuning of PD Regulators

The ideal *PD* regulator given by the transfer function $C_{PD}(s) = 1 + sT_D$ can not be realized, therefore only the approximate realizable phase-lead form

$$\tilde{C}_{PD}(s) = A_P \frac{1 + sT_D}{1 + sT} = A_P \left(1 + \frac{s\tau}{1 + sT} \right); \quad T_D = T + \tau \quad (8.14)$$

has practical applications, which formally is the same as \tilde{C}_{PI} . The difference is that here $T_D > T$, and the time constant of the differentiating channel is τ . With the choice $T_D = T_2$, where T_2 is the second largest time constant of the process, the \tilde{C}_{PD} regulator lengthens the higher frequency part of the straight line of slope -20 dB/decade by shifting the breakpoint frequency $1/T_2$ to the right to point $1/T$. This improves the phase conditions, to reach a given phase margin the value of ω_c can be increased, which results in a faster settling process. There is a limit for the choice of the time constant T , as in the unit step response of the regulator at $t = 0$ there is a jump of value $A_P T_D/T$, which decreases asymptotically, its final value is A_P . So the value of the overexcitation is $\eta = T_D/T$, which is the same as the pole placement ratio. Not every actuator can execute this jump. The big mechanical actuators can tolerate 2-3-fold, whereas the more advanced electronic devices can bear at most a 10-fold jump. Therefore during regulator design an average 3-5-fold jump is allowed, then in practice it is tested whether the regulator is saturated (that is reaches the limit of its signal range). In some cases the choice of $T_D = T_3$ (where T_3 is the third largest time constant) is also possible, but the effect of this has smaller significance.

Figure 8.11 gives the unit step response and the BODE diagram of the *PD* regulator. The latter is drawn for $A_P = 1$. If the gain is different, the BODE amplitude diagram is shifted parallel. This element is also called a phase-lead element, as its phase angle is positive, that is in case of a sinusoidal input signal the output signal is accelerated in phase relative to the input signal.

The accelerating effect of the control system can be understood the most easily in case of a *PD* regulator. The inertial behavior of the processes can be overtaken only by conveying extra accelerating energy. This is ensured by the overexcitation (see Sect. 2.4).

The *PD* regulator is used if the system has to be accelerated. This acceleration is reached if the straight line section of slope $+20$ dB/decade of the *PD* regulator is placed in the frequency range where the slope of the BODE diagram of the process is -40 dB/decade. Thus the straight line section of the loop transfer function with slope -20 dB/decade is lengthened, and the cut-off frequency can be placed at a higher frequency. The higher the value of the parameter η , the bigger the acceleration.

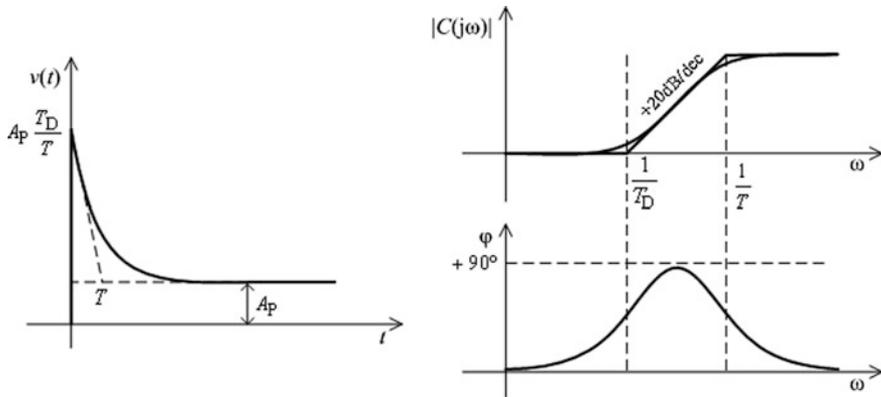


Fig. 8.11 Unit step response and the approximate BODE diagram of the *PD* regulator

Thus the *PD* regulator places one single breakpoint (generally the second one) of the BODE diagram at a higher frequency, but at the same time an overexcitation is produced which is equal to the ratio of the new and the old breakpoint. The amplitude-frequency diagram of the process is shifted parallel to the horizontal axis by A_P .

With the *PD* regulator the system can be stabilized. The system can be accelerated. Setting an appropriate phase margin the prescribed dynamic behavior can be reached. But as the regulator is of proportional type, with proportional processes the control system will have a static error for unit step reference signal.

The acceleration effect can be explained by the fact that at the beginning the error signal—exciting the *PD* regulator—produces a high signal at the regulator output, and the process temporarily starts to track this higher signal with its time constant. Thus the output signal starts with a big slope. In the control system then the error signal decreases, and the output signal settles at its steady value.

8.1.5 Tuning of *PID* Regulators

The simplest *PID* regulator parameterization is given by (8.8), that is

$$\tilde{C}_{PID} = A_P \frac{(1 + sT_I)(1 + sT_D)}{sT_I(1 + sT)}. \tag{8.15}$$

This regulator is the combination of the previous two (*PI* and *PD*) regulators, resulting in their series connection. Thus the design procedure shown previously has to be repeated here. The integrating time is set by the choice $T_I = \max\{T_i\} = T_1$, whereas $T_D = T_2$ is chosen for the differentiating time. By this design the straight line section of slope -20 dB/decade on the BODE diagram of the loop transfer function is lengthened by the maximal possible extent provided by the

structure of the regulator. In the control circuit after determining $K_{\max}(\varphi_{to})$, the maximal integrating gain belonging to the prescribed phase margin φ_{to} , the integrating gain factor of the regulator is obtained according to $K_I = A_P/T_I = K_{\max}/P(0)$. Regarding the choice of the time constant T of the approximate differentiating effect the considerations discussed in the design of PD regulators are also valid here, but the overexcitation in control circuits containing integrating effects is calculated in a different way. As the output of an integrator is changing until its input reaches the zero value, therefore the steady state of the control system is reached if the error signal has settled to zero. The initial jump of the PID regulator in the case of a unit step reference signal is $A_P T_D/T$. The steady state value of the process input is $1/P(0)$. The overexcitation now is calculated by $\eta = A_P P(0) T_D/T = K T_D/T$. Thus the overexcitation is obtained by the product of the loop gain and the pole placement ratio.

The form (8.15) of the regulator can be used straightforwardly for the analysis in the frequency domain. Figure 8.12 gives the unit step response and the BODE diagram of the PID regulator, the latter is drawn for $A_P = 1$. If the gain is different, the BODE amplitude diagram is shifted parallel.

PID regulators are used when the static accuracy of the control system has to be increased and the system also has to be accelerated. With the initial -20 dB/decade slope of the BODE diagram of the PID regulator, the low frequency range of the open-loop BODE diagram is modified, thus increasing the type number and the static accuracy. With the straight line section of slope $+20$ dB/decade of the BODE diagram of the regulator, the performance in the middle frequency range is modified. As by ensuring the appropriate phase margin, now the cut-off frequency can be placed to a higher value, thus a faster behavior of the control system can be reached.

The performance of the PID regulator can be approximated by the so-called phase-lag-lead element. Its transfer function is

$$C_{FKS}(s) = A_P \frac{1 + sT_1}{1 + sT_3} \frac{1 + sT_2}{1 + sT_4}, \quad \text{where } T_3 > T_1 > T_2 > T_4.$$

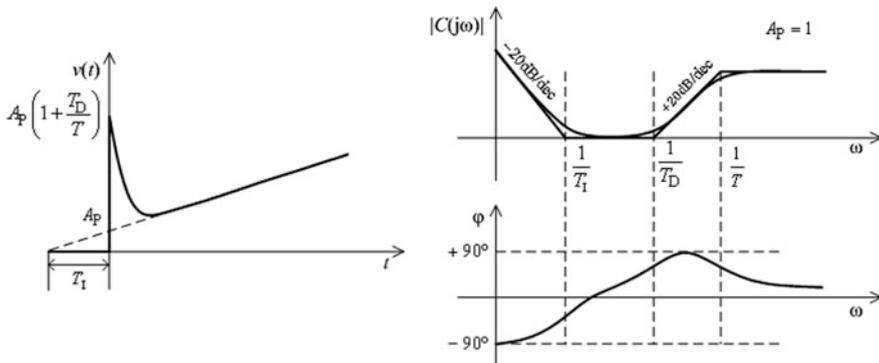


Fig. 8.12 Unit step response and the approximate BODE diagram of the PID regulator

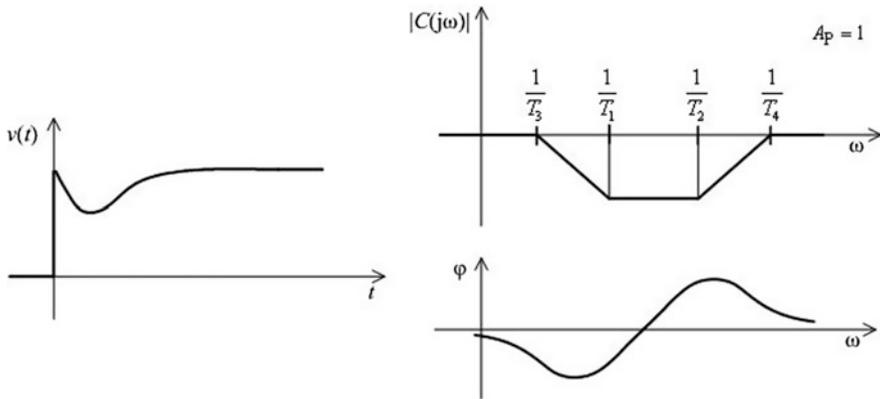


Fig. 8.13 Unit step response and the approximating BODE diagram of the phase-lag-lead element

The unit step response and the approximate BODE diagram of the phase-lag-lead element is given in Fig. 8.13. As this element does not contain an integrating effect, it will not ensure a zero steady state error in the control system.

The regulator is designed for the process (or for its model) to fulfill the quality specifications. First a decision has to be made about the regulator structure considering the given process and the prescriptions. Then the parameters of the regulator are chosen. For example in a *PID* regulator there are four free parameters, A_P , T_I , T_D and T . Regulator design means the choice of the parameters. The design can be executed in the time or in the frequency domain. Several procedures and methods have been elaborated to execute the design. The diversity of the design methods comes also from the fact that in the individual applications the design specifications may differ significantly. In many cases the prescriptions create contradictory requirements. In the general case, this corresponds to an optimization problem, when a proper compromise is formed to satisfy the contradictory requirements. In practice often some iteration procedure (intelligent guessing) is applied instead of executing the optimization (Table 8.1).

The presented *P*, *PI*, *PD* and *PID* regulators are also called compensators.

Let us summarize the practical rules for the design of the *PID*-like regulators using pole cancellation.

Table 8.1 Summarizing the regulator design

Regulator	T_I	T_D	A	K_I
<i>P</i>			$A_P = K/P(0)$	
<i>I</i>				$K_I = K_{max}/P(0)$
<i>PI</i>	T_I			$K_{PI} = K_{max}/P(0)$
<i>PD</i>		T_D		$K_{PD} = K_{max}/P(0)$
<i>PID</i>	T_I	T_D		$K_{PID} = K_{max}/P(0)$

The P regulator can be used if there are no high requirements for the static accuracy of the control system, and the control system can be slow. If the process contains an integrating effect, then the static accuracy will be appropriate also with the proportional regulator.

PI regulators are used if accurate tracking is required in steady state for a unit step reference signal. The integrating effect will ensure an accurate settling. With PI regulators, the control system will be slow.

PD regulators accelerate the control system. This effect is reached by the overexcitation provided by the regulator.

PID regulators are used if both the static accuracy and the speed of the control system have to be increased.

In the case of a proportional process, the parameters of the regulator applying pole cancellation technique are chosen as follows: the parameter T_I , the integrating time constant is chosen equal to the largest time constant of the process (this is the pole belonging to the smallest breakpoint frequency), and the parameter T_D is taken equal to the second largest time constant. Thus the zeros of the regulator cancel the poles of the process. The parameter T which appears in the denominator of the realizable differentiating element is given in by $T = T_D/\eta$, where η is the pole placement ratio, which specifies the frequency shift of the compensated pole realized by the PD element. If it is chosen to be a higher value, the control system will be faster at the price of a higher maximum value of the control signal. As A_p does not influence the phase-frequency course of the open-loop, it is used to set the prescribed phase margin.

If the process is not proportional, the type of the regulator can be decided on the basis of the approximate BODE diagram to fulfill the quality specifications. Pole cancellation can be applied expediently also in this case.

8.1.6 Influence of the Dead-Time

The effect of the dead-time can be considered relatively simply in the case of series compensation, as

$$H_H(s) = e^{-sT_d} \Rightarrow H_H(j\omega) = e^{-j\omega T_d} = e^{-j\varphi_d} \quad (8.16)$$

This means that $L(j\omega)$, the frequency characteristic of the loop transfer function, is modified by an element of unit amplitude and of phase angle $\varphi_d = -\omega T_d$, thus only the phase characteristic is changed. This can be taken into account by prescribing the required phase margin to be $\varphi_{to}^d = \varphi_{to} + \omega T_d$ instead of the original φ_{to} .

As the transfer function of the dead-time is a non-rational function, computer programs which are not prepared for handling such functions (e.g., MATLAB[®]) can not easily take its effect into account. In this case the transfer function of the dead-time can be approximated by a rational fraction.

Rational fractional approximations of the dead-time were discussed in Sect. 2.5.

8.1.7 Realization of PID Regulators

Analog regulators can be realized based on different physical conceptions (electronic, pneumatic, hydraulic, etc.). The electronic realization is built by a feedback operational amplifier (Fig. 8.14). The operational amplifier has a very high voltage amplification (10^5-10^7), its input resistance is also high. The relationship between its output and input is: $C(s) = -Z_v(s)/Z_b(s)$, where $Z_v(s)$ is the feedback impedance and $Z_b(s)$ is the input impedance. Figure 8.15 shows a realization of a pure integrator, and of a PI circuit.

Different versions of the circuits can be given. For example an aspect can be the realization of a circuit where changing the value of one element (generally a resistance) sets only one regulator parameter, and has no effect on the others.

Approximate phase-lag, phase-lead and phase-lag-lead regulators can be built from passive elements, the circuit does not contain operational amplifier. Figure 8.16 shows the circuits of these regulators with resistors and capacitors.

A compact regulator is produced by firms that manufacture automatic elements. In this regulator, the structure (*P*, *PI*, *PD* or *PID*) can be set by a switch, and the parameters generally can be tuned by setting potentiometers.

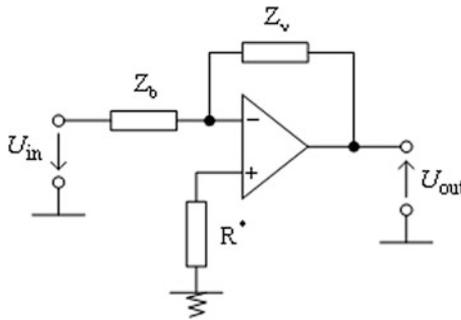


Fig. 8.14 Realization of a regulator with an operational amplifier

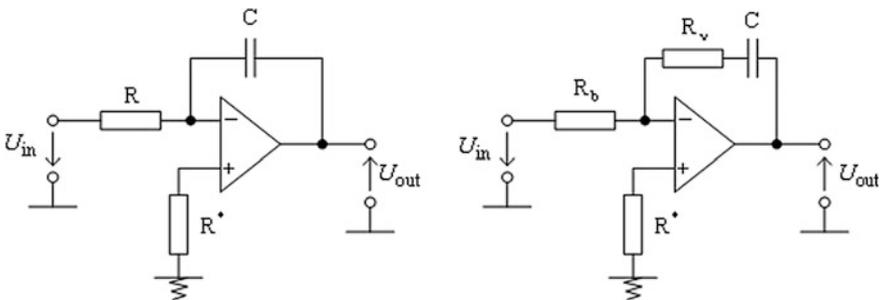


Fig. 8.15 Realization of an I and a PI regulator

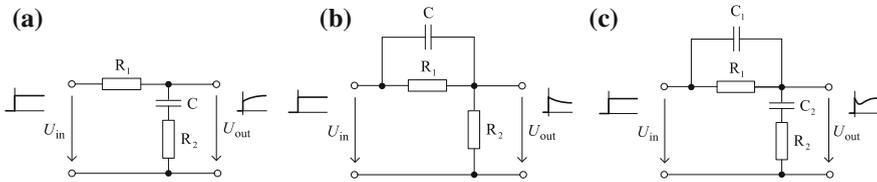


Fig. 8.16 Realization of phase-lag, phase-lead and phase-lag-lead regulator with resistors and capacitors

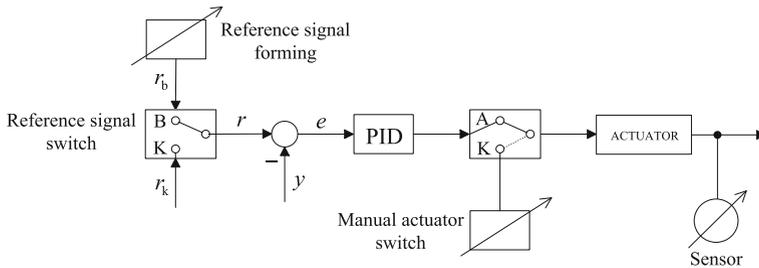
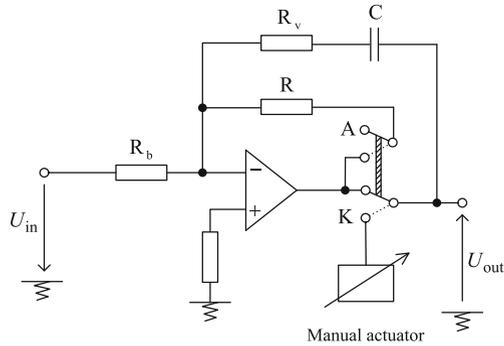


Fig. 8.17 Block-scheme of a *PID* regulator with manual-automatic switch-over

The regulators realize a bumpless transfer when switching from manual mode to automatic mode, and in the case of saturation, also handle integrator windup. Figure 8.17 gives a block diagram of a *PID* regulator, also showing the manual-automatic switch-over. The reference signal can be switched over between two signal generators. The system can be switched over between manual and automatic modes of operation. In the case of slow processes, when switching over from automatic operation mode to manual mode, the operator sets the value of the manipulated variable shown by the measuring instrument with the potentiometer to create the manual manipulated variable, and then executes the switch-over. Switching-over from manual mode to automatic mode is more critical, as during manual operation the signal on the integrating channel of the regulator is likely to “run away”. Figure 8.18 shows a possible solution for manual-automatic switch-over in the case of a *PI* regulator, avoiding windup. During the manual mode, the capacitor *C* is charged to the voltage of the manual manipulated signal, and after the switch-over the integration starts from this initial value. In manual mode the resistor *R* ensures the feedback of the operational amplifier, thus it will not be saturated. Handling of saturation will be discussed in Sect. 8.4.

Nowadays, instead of analog techniques, it is more and more the *programmable logic controllers (PLC-s)* or process control computers that realize the regulator. The process is connected to the computer via an *A/D* converter. The *PID* control algorithm is realized by a computer program. The regulator output is connected to the process input through a *D/A* converter. The program has to handle saturation

Fig. 8.18 *PI* regulator with manual-automatic switch-over



effects and also ensure bumpless transfer between the manual and the automatic operation modes. Digital *PID* regulators will be discussed in detail in Chap. 13.

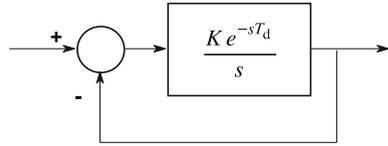
8.2 Design of Residual Systems

For some typical cases, regulator design can also be executed analytically. In the design of the usual regulators one can see that in the generally used simplest method, the breakpoints corresponding to the integrating and differentiating time constants of the *PID* regulator are fitted to the breakpoints belonging to the two largest time constants of the process. Then the only free parameter of the regulator is the gain, which has to be tuned appropriately. The gain is chosen to ensure a prescribed phase margin, gain margin, or NYQUIST stability margin. If after cancellation of the two dominant time constants the order of the so-called residual or reduced system is low, then the design can be executed easily, several times providing an explicit analytical result. If the residual system is a higher order, more complicated system, then only numerical methods (MATLAB[®]) can be applied. In the sequel, the design of the loop gain will be executed for some typical residual systems.

8.2.1 Simple Residual System with Dead-Time and Integrator

As seen previously, a pure dead-time element with negative feedback is at the borderline case of stability with unit loop gain. The control system can be stabilized and also its static accuracy can be improved applying an integrator instead of a proportional regulator. The considered residual system is shown in Fig. 8.19; its loop transfer function is

Fig. 8.19 Residual system containing an integrator and dead-time



$$L(s) = \frac{K e^{-sT_d}}{s} = \frac{e^{-sT_d}}{sT_1}; \quad L(j\omega) = \frac{K e^{-j\omega T_d}}{j\omega} = a(\omega) e^{j\varphi(\omega)}. \quad (8.17)$$

The amplitude-frequency curve of the open-loop is a straight line of slope -20 dB/decade, and the loop gain $K = 1/T_1$ gives the cut-off frequency of the open-loop. If φ_t is the prescribed phase margin, then both the phase condition

$$-\pi/2 - \omega T_d = -\pi + \varphi_t \quad (8.18)$$

and the absolute value condition

$$\frac{K}{\omega} = 1 \quad (8.19)$$

have to be fulfilled. Solving for the two equations the loop gain yields

$$K = \frac{\pi/2 - \varphi_t}{T_d} = \frac{\pi - 2\varphi_t}{2T_d} = K_{\varphi_t}. \quad (8.20)$$

If the gain of the regulator is A_c and the gain of the process is A_p , then the gain of the regulator is

$$A_c = \frac{K_{\varphi_t}}{A_p}. \quad (8.21)$$

The resulting formula (8.20) can be used also for the design of a pure integrating regulator if the process contains only a pure dead-time. Thus the integrating time constant of an integrating regulator used for compensation of a pure dead-time process is designed to be

$$T_1 = \frac{1}{K_1} = \frac{1}{K} = \frac{1}{K_{\varphi_t}}. \quad (8.22)$$

Relating it to the dead-time the following relationship is obtained:

$$\frac{T_1}{T_d} = \frac{1}{\frac{\pi}{2} - \varphi_t} = \frac{2}{\pi - 2\varphi_t}. \quad (8.23)$$

Some typical values: if $\varphi_t = 30^\circ = \pi/6$, then $T_1/T_d = 3/\pi \approx 1$; if $\varphi_t = 60^\circ = \pi/3$, then $T_1/T_d = 6/\pi \approx 2$. At the borderline of stability $\varphi_t = 0$, and $T_1/T_d = 2/\pi \approx 0.637$.

Thus in the case of a minimum phase system which also contains dead-time it is not sufficient to place the cut-off frequency on the straight line section of the BODE amplitude diagram of slope -20 dB/decade: to ensure a phase margin of about $\varphi_t = 60^\circ$ it has to be located around one half of the reciprocal of the dead-time.

Let κ be the prescribed gain margin (see Sect. 5.6). Now the phase condition is given by

$$-\frac{\pi}{2} - \omega T_d = -\pi, \quad (8.24)$$

whence the intersection frequency of the open-loop NYQUIST diagram with the negative real axis is expressed as

$$\omega_a = \frac{\pi}{2T_d} \quad (8.25)$$

and the absolute value of the loop frequency function at this point is

$$a(\omega_a) = \left. \frac{K}{\omega} \right|_{\omega=\omega_a} = \frac{K}{\omega_a} = 1 - \kappa. \quad (8.26)$$

The loop gain is obtained from the solution of the last two equations:

$$K = \frac{\pi(1 - \kappa)}{2T_d}. \quad (8.27)$$

A typical value: if $a_t = 0.5$, then $T_1/T_d = 4/\pi \approx 1.2$. The borderline of stability ($a_t = 0$) again is obtained at $T_1/T_d = 2/\pi \approx 0.637$.

The tuning relationship of the integrating time constant of an I integrating regulator used for compensating a pure dead-time process is given by

$$\frac{T_1}{T_d} = \frac{2}{\pi(1 - \kappa)}. \quad (8.28)$$

The NYQUIST stability margin $\rho_m = \rho_{\min}$ is defined as the smallest distance of the loop frequency function $L(j\omega)$ from the point $-1 + 0j$. Generally it is not easy to give this distance as an explicit algebraic expression, as it can be derived from the solution of an extremum seeking problem. Therefore generally its graphical representation is employed. In Fig. 8.20 ρ_{\min} is plotted versus the ratio T_1/T_d .

As seen in Chap. 5, a control system built of a pure dead-time element with unity negative feedback is stable only if its loop gain is less than 1. But then the static error is very high. Therefore the above considerations are frequently used not only when the pole cancellation regulator design technique is applied, but also in the

Fig. 8.20 NYQUIST stability margin ρ_{\min} versus T_I/T_d

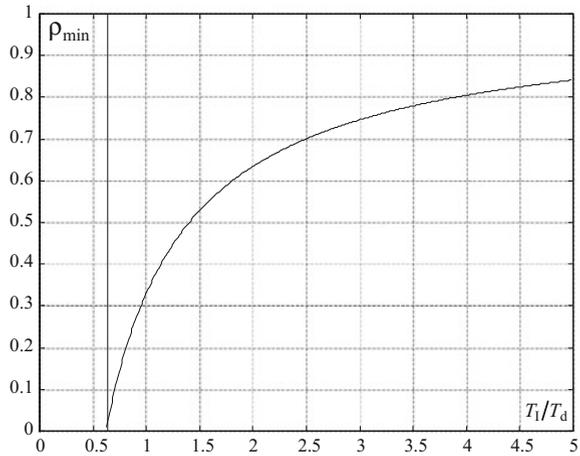
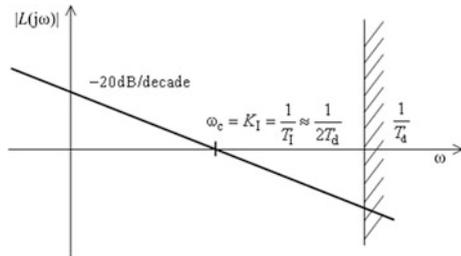


Fig. 8.21 BODE amplitude-frequency diagram of a compensated dead-time system



case of a pure I regulator. The BODE amplitude-frequency diagram of the open-loop for the case of $\varphi_t = 60^\circ = \pi/3$ is shown in Fig. 8.21. From (8.23) it can be seen, that when compensating a dead-time system, the cut-off frequency has to be placed at about on the half of the reciprocal of the dead-time on the long straight line section of slope -20 dB/decade. At the borderline of stability $\varphi_t = 0$ and then $K_I = \pi/2T_d \approx 1.57/T_d$.

An aperiodic process can be quite well approximated by a first-order (or second-order) lag element (see Sect. 8.3). To meet higher quality requirements, instead of an I regulator, a PI regulator can be used. With the zero of the PI element, the pole of the process can be cancelled, and instead an integrating effect is introduced. With this PI compensation the open-loop is given as an integrating element with dead-time:

$$L(s) = A_P \frac{1 + sT_1}{s} \frac{K_P e^{-sT_d}}{1 + sT_1} = \frac{A_P K_P e^{-sT_d}}{s} = \frac{K e^{-sT_d}}{s}. \tag{8.29}$$

The parameter K , which also gives the cut-off frequency of the open-loop, can be designed to ensure 60° for the phase margin according to Sect. 6.2:

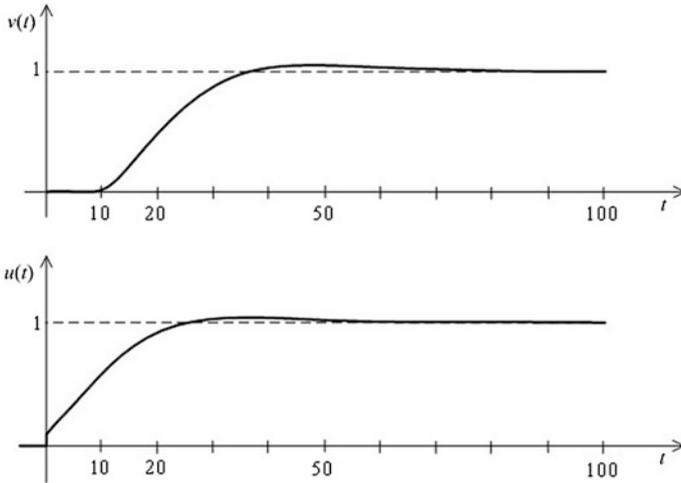


Fig. 8.22 Output and control signals of a dead-time system compensated by a *PI* regulator in the case of a unit step reference signal

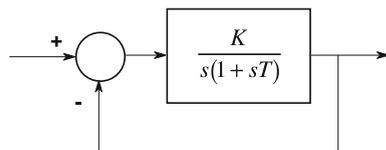
$K = \omega_c \approx 1/2T_d$. The output signal and the control signal of the control system for a unit step reference signal is shown in Fig. 8.22 in the case of $T_d = 10$ and $T_I = 2$. The dead-time can not be eliminated from the system; the output signal starts to change only after the dead-time.

(Note that in the case of a dead-time process with first order lag it makes no sense to add a *PD* compensation as well. In case of a dead-time process with two time lags, adding a *PD* effect will accelerate the system only if the time constants of the lag elements are the dominant ones. If the dead-time is intermediate or the largest time constant, then because of the significant phase shift of the dead-time, the phase margin can not be increased significantly by the effect of the *PD*, thus it makes no sense to apply it. As the increase of the cut-off frequency is limited by the dead-time, the behavior of a control system containing dead-time will be slow.)

8.2.2 Simple Residual System with Integrator and Time Lag

The residual system is shown in Fig. 8.23. Its loop transfer function is

Fig. 8.23 Residual system consisting of an integrator and a first order lag



$$L(s) = \frac{K}{s(1+sT)} = \frac{1}{sT_1(1+sT)}; \quad L(j\omega) = \frac{K}{j\omega(1+j\omega T)} = a(\omega)e^{j\varphi(\omega)}. \quad (8.30)$$

Using (8.30) the overall transfer function of the closed-loop system, that is the supplementary sensitivity function, is

$$T(s) = \frac{K}{K+s+Ts^2} = \frac{1}{1+\frac{1}{K}s+\frac{T}{K}s^2} = \frac{1}{1+2\xi\tau s+\tau^2s^2}, \quad (8.31)$$

which is a second order oscillating element whose damping factor can be set accurately by the loop gain. Comparing the coefficients the following relationships are obtained:

$$K = \frac{1}{4\xi^2T}; \quad \frac{T_1}{T} = 4\xi^2. \quad (8.32)$$

A damping factor, $\xi = \sqrt{2}/2 \cong 0.707$, which provides a “nice” transient response is obtained by a loop gain of $K = 0.5/T$. In some applications, an aperiodic transient is required. The limiting case of aperiodic response is $\xi \geq 1$, which corresponds to $K \leq 0.25/T$.

Let φ_t be the prescribed phase margin. Then based on the loop transfer function the following relationship can be written for the phase condition:

$$-\frac{\pi}{2} - \text{arctg}(\omega T) = -\pi + \varphi_t; \quad (8.33)$$

and for the absolute value condition,

$$\frac{K}{\omega\sqrt{1+\omega^2T^2}} = 1. \quad (8.34)$$

From the solution of these two equations, the loop gain is obtained as

$$K = \frac{1}{T} \text{tg}(90 - \varphi_t) \sqrt{1 + \text{tg}^2(90 - \varphi_t)} = \frac{1}{T} \frac{\sin(90 - \varphi_t)}{\cos^2(90 - \varphi_t)}, \quad (8.35)$$

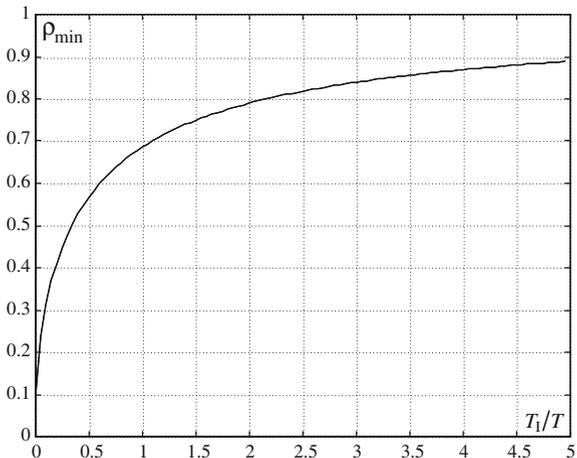
where the trigonometric identity

$$\text{tg}(x) \sqrt{1 + \text{tg}^2(x)} = \frac{\sin(x)}{1 - \sin^2(x)} = \frac{\sin(x)}{\cos^2(x)} \quad (8.36)$$

was taken into account.

Some typical values: if $\varphi_t = 45^\circ = \pi/4$, then $KT = \sqrt{2} \cong 1.414$; if $\varphi_t = 60^\circ = \pi/3$, then $KT = 2/3 \cong 0.667$. Also, the φ_t corresponding to a given ξ can be

Fig. 8.24 ρ_{\min} , the NYQUIST stability margin versus T_1/T



calculated. If $\xi = \sqrt{2}/2 \cong 0.707$ is required, then the corresponding phase margin is obtained as $\phi_t = 65.53^\circ$.

As the NYQUIST diagram of this residual system does not step across the third quadrant of the complex plane, therefore now the gain margin can not be designed ($\kappa_t \equiv 1.0$). This system is structurally stable.

Here ρ_{\min} , the NYQUIST stability margin, can be obtained only graphically. Its graph versus T_1/T is seen in Fig. 8.24.

A residual system containing an integrator and a first order lag is obtained for instance if a proportional second order lag element is compensated by a *PI* regulator using the pole cancellation technique. The loop transfer function is

$$L(s) = A_P \frac{1 + sT_1}{s} \frac{K_P}{(1 + sT_1)(1 + sT_2)} = \frac{A_P K_P}{s(1 + sT_2)} = \frac{K}{s(1 + sT)} = \frac{1}{sT_1(1 + sT)}, \tag{8.37}$$

which is of form (8.30), and the design formulae above can be applied. In the case of a prescribed damping factor ξ , the gain of the regulator can be calculated. If $\xi = 1$, then $K = 1/4T_2$, and the system has two coinciding real poles. The unit step response of the closed-loop just will not have any overshoot. If $\xi = \sqrt{2}/2 \approx 0.7$, then $K = 1/2T_2$, and the phase margin of the system will be $\phi_t = 65.53^\circ$. The overshoot of the unit step response of the closed-loop system will be about 5%. As $K = A_P K_P$ is the gain of the whole circuit, the gain of the regulator is obtained by $A_P = K/K_P = 1/2K_P T_2$.

8.3 Empirical Regulator Tuning Methods

Besides the model-based regulator design methods, which provide an outline of the foreseeable properties of the closed-loop control system, several experimental *PID* regulator tuning methods have been suggested in industrial process control, mainly for stable processes. These methods are often used during the installation of the regulator. The recommendations, the “recipes” for regulator parameter tuning, are based on some preliminary measurements executed on the process, or on simulations and practical observations.

Note that these methods are appropriate for fast regulator tuning when putting the control system in operation, but model based design methods provide a better basis for and overview of the behavior of the control system, or for modifying the tuning in case of changes. Usually, experimental methods are considered as the initial settings before introducing theoretically elaborated more extensive methods.

8.3.1 Methods of ZIEGLER and NICHOLS

Frequency response method

It is supposed that the technology of the process allows operating the closed-loop control system for a short time on the borderline of stability applying only a proportional regulator. During this experiment the integrating and the differentiating channels of the regulator are switched off ($T_I = \infty$ and $T_D = 0$), then by cautiously increasing A_P the borderline of stability is reached, when sinusoidal oscillations appear. After each change of A_P we have to wait for the new steady state to settle down, which can take a long time. Let $A_{P,cr}$ denote the critical gain and T_{cr} the time period of the constant sinusoidal oscillations. ZIEGLER and NICHOLS suggested the following Table 8.2 regulator tuning rules of thumb.

These tuning rules correspond to a damping factor of about $\xi = 0.25$ (which corresponds to a quite high overshoot of about 40%, thus in practice they can be used only for compensating slowly changing disturbances).

Tuning method based on the unit step response

The unit step response of several industrial processes shows an aperiodic characteristic with dead-time (Fig. 8.25). The straight line fitted to the inflection point of

Table 8.2 Regulator tuning according to ZIEGLER-NICHOLS (I)

Regulator	T_I	T_D	A_P
<i>P</i>			$0.5A_{P,cr}$
<i>PI</i>	$0.85T_{cr}$		$0.45A_{P,cr}$
<i>PID</i>	$0.5T_{cr}$	$0.125T_{cr}$	$0.6A_{P,cr}$

Fig. 8.25 Shape of the measured unit step response of the process

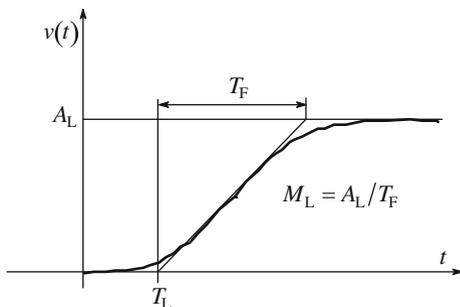


Table 8.3 Regulator tuning according to ZIEGLER-NICHOLS (II)

Regulator	T_I	T_D	A_P
<i>P</i>			$1/T_L M_L$
<i>PI</i>	$3T_L$		$0.9/T_L M_L$
<i>PID</i>	$2T_L$	$0.5T_L$	$1.2/T_L M_L$

the step response determines two quantities indicated in the figure, the so called latent dead-time (T_L) and the latent slope ($M_L = A_L/T_F$). Based on these quantities ZIEGLER and NICHOLS suggested tuning rules summarized in the Table 8.3.

From the above tables it can be seen, that the tuning of the three parameters $\{A_P, T_I, T_D\}$ in both cases is based on two observed values, then the *D*-channel is set as $T_D = T_I/4$. Of course this is a source of further design freedom.

8.3.2 Method of OPPELT

Several grapho-analytical methods can be used to fit an approximate first order lag element with dead-time given by (8.38) to the measured unit step response of the process.

$$\hat{P}(s) = \frac{A_L}{1 + sT_F} e^{-sT_L}; \quad A_L = \frac{y_\infty - y_0}{u_\infty - u_0}; \quad T_L = t_1 - t_0 \quad \text{and} \quad T_F = t_2 - t_1. \tag{8.38}$$

Fig. 8.26 Approximation of an aperiodic process by a first order lag with dead-time

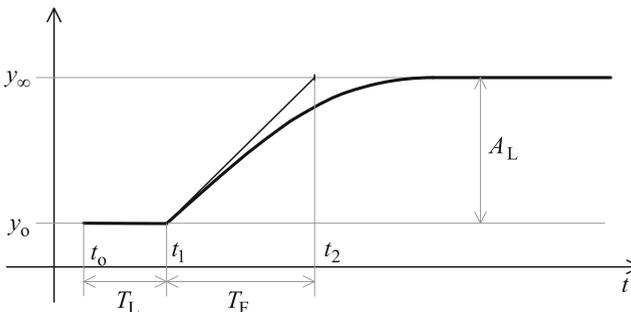


Table 8.4 Regulator tuning according to OPPELT

Regulator	$A_P M_L T_L$	T_I/T_L	T_D/T_L
<i>P</i>	1		
<i>PD</i>	1.2		0.25
<i>PI</i>	0.8	3	
<i>PID</i>	1.2	2	0.42

Let us set manually the nominal operating point, where for an input signal u_o the value of the output signal is y_o . At time t_o let us apply a step input, when the input signal u jumps from u_o to u_∞ . The output signal is seen in Fig. 8.26. (The definition of the time constant T_F in the two figures is different, as OPPELT does not suppose an inflection point.)

The tuning parameters according to OPPELT were determined for a damping factor of $\xi = 0.25$ considering the parameters of the approximate model given by (8.38). Therefore similarly to the ZIEGLER-NICHOLS method, in this case also a quite high overshoot can be expected. The proposed tuning values are summarized in the Table 8.4.

8.3.3 Method of CHIEN-HRONES-RESWICK

For the tuning of the regulator parameters, CHIEN, HRONES and RESWICK suggested the values summarized in the Table 8.5.

8.3.4 Method of STREJC

STREJC approximated the process by the model given by the transfer function

Table 8.5 Regulator tuning according to CHIEN-HRONES-RESWICK

Regulator	Fastest aperiodic transient	Fastest oscillating transient with 20% overshoot
<i>P</i>	$A_P = 0.3T_F/T_L$	$A_P = 0.7T_F/T_L$
<i>PI</i>	$A_P = 0.35T_F/T_L$ $T_I = 1.2T_F$	$A_P = 0.6T_F/T_L$ $T_I = 1.0T_F$
<i>PID</i>	$A_P = 0.6T_F/T_L$ $T_I = 1.0T_F$ $T_D = 0.5T_L$	$A_P = 0.95T_F/T_L$ $T_I = 1.35T_F$ $T_D = 0.47T_L$

Table 8.6 Regulator tuning according to STREJC

Regulator	A_P	T_I	T_D
<i>P</i>	$\frac{1}{A_L(n-1)}$		
<i>PI</i>	$\frac{n+2}{4A_L(n-1)}$	$\frac{T(n+2)}{3}$	
<i>PID</i>	$\frac{7n+16}{16A_L(n-2)}$	$\frac{T(7n+16)}{15}$	$\frac{T(n+1)(n+3)}{7n+16}$

$$\hat{P}(s) = \frac{A_L}{(1 + sT)^n}. \tag{8.39}$$

On the basis of the parameters of the approximate model, he gave tuning proposals according to the Table 8.6 to set the parameters of the *PID* regulator family.

8.3.5 Relay Method of ÅSTRÖM

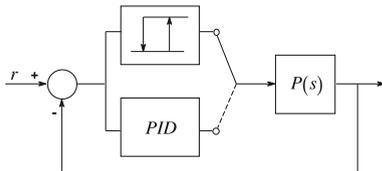
In the case of real industrial processes, it is rather difficult to set the critical loop gain corresponding to the borderline of stability of the closed-loop control system. Generally, approaching this range is not permitted because of safety requirements. If we want to use regulator tuning rules based on the critical loop gain, then it is expedient to determine its value with a different method. ÅSTRÖM suggested a method which can well be applied in practice. In a closed-loop system, the *PID* regulator has to be replaced by an amplifier realizing a relay characteristic with hysteresis (see Fig. 8.27).

For the stability analysis of special closed-loop nonlinear systems, the so called describing function method can be applied. The describing function $N(j\omega, a)$ is obtained by harmonic linearization, when the static nonlinear element is excited by a sinusoidal signal of amplitude a , and then the complex division of the basic harmonic of the output signal and the input sinusoidal signal is executed. Generally $N(j\omega, a)$ is a complex function with parameter a . In analyzing the stability, the role of the point $-1 + j0$ is replaced by the function $-1/N(j\omega, a)$. The system is at the borderline of stability at the point (ω_{cr}, a_{cr}) where the NYQUIST diagram of the loop frequency function $L(j\omega)$ intersects the inverse negative describing function $-1/N(j\omega, a)$, that is,

$$L(\omega_{cr})N(\omega_{cr}, a_{cr}) = -1, \quad \text{i.e.} \quad -1/L(\omega_{cr}) = N(\omega_{cr}, a_{cr}). \tag{8.40}$$

Here, a_{cr} is the approximate amplitude of the periodic oscillation in the borderline case of stability. (It is not an entirely accurate value, as the harmonic linearization considers only the first harmonic.) From the time period T_{cr} of the

Fig. 8.27 Tuning of a *PID* regulator with the relay method



periodic signal, a good approximation of the angular frequency belonging to the critical point can be calculated by $\omega_{cr} = 2\pi/T_{cr}$.

Let the deadband of the hysteresis be zero, so the regulator is a two-position relay. In this case the input of the process is a rectangular signal, and the process output in steady state is a periodic signal. In the linear case for the critical gain the characteristic equation is written as

$$1 + L(\omega_{cr}) = 1 + A_{P,cr}P(\omega_{cr}) = 0, \quad \text{i.e.} \quad A_{P,cr} = -1/P(\omega_{cr}). \quad (8.41)$$

From (8.40) and (8.41), a simple relationship can be obtained to estimate the critical gain:

$$A_{P,cr} = -1/P(\omega_{cr}) = N(\omega_{cr}, a). \quad (8.42)$$

If steady oscillations of amplitudes $\pm \Delta u$ and $\pm \Delta y$ are measured at the process input and at the process output, respectively, then the critical gain is

$$A_{P,cr} = N(a) = \frac{4\Delta u}{\pi\Delta y}, \quad (8.43)$$

where now $N(a)$, the describing function of the relay, depends only on the amplitude. The most important advantage of this method is that the oscillation of the process output can be gradually set to a still allowed value, namely to $\Delta u = h$, which is one half the height of the relay characteristic, and $\Delta y = a$. If an integrator is serially connected to the relay, then the loop gain belonging to the phase angle of -270° can be determined with this method.

Based on the describing function belonging to the hysteresis characteristic

$$-1/N(a) = -\frac{\pi}{4h}\sqrt{a^2 - g^2} - j\frac{\pi g}{4h} = -\frac{\pi}{4\Delta u}\sqrt{\Delta y^2 - g^2} - j\frac{\pi g}{4\Delta u}, \quad (8.44)$$

which is a straight line parallel to the negative real axis. Here g is the half-width of the hysteresis. Hence it can be easily checked that again the relationship

$$A_{P,cr} = |N(a)| = \frac{4\Delta u}{\pi\Delta y} \quad (8.45)$$

is obtained. Nowadays, in advanced electronic compact regulators, tuning with the relay method is already a built in possibility. (By changing the values h and g of the relay characteristic, further points of the NYQUIST diagram also could be analyzed.)

8.3.6 Method of ÅSTRÖM-HÄGGLUND

This method also uses the simple approximation of the process according to (8.38) as a starting point, but as a design parameter it also uses the maximum value M_{\max} of the complementary sensitivity function. The corresponding M curve characterizes the distance measured from the point $-1 + 0j$, thus to some extent the robustness of the regulator could be considered as well. The expression of the PID regulator used in this method is given by

$$u(t) = A_P \left[\beta r(t) - y(t) + \frac{1}{T_I} \int_0^t e(\tau) d\tau + T_D \frac{de(t)}{dt} \right] \tag{8.46}$$

where in forming the error signal the reference signal $r(t)$ and the output signal $y(t)$ are taken into account with different weights. (β is the weighting factor of the reference signal.) On the basis of the approximate form (8.38), introduce the following relative parameters:

$$\alpha = A_L \frac{T_L}{T_F} \quad \text{and} \quad \gamma = \frac{T_L}{T_L + T_F}. \tag{8.47}$$

ÅSTRÖM and HÄGGLUND suggested setting the regulator parameters according to the following function:

$$f(\gamma) = a_0 \exp\{a_1 \gamma + a_2 \gamma^2\}. \tag{8.48}$$

Table 8.7 Tuning of the parameters of the PI regulator

	$M_{\max} = 1.4$			$M_{\max} = 2$		
$f(\gamma)$	a_0	a_1	a_2	a_0	a_1	a_2
αA_P	0.29	-2.7	3.7	0.78	-4.1	5.7
T_I/T_L	8.9	-6.6	3.0	8.9	-6.6	3.0
T_I/T_F	0.79	-1.4	2.4	0.79	-1.4	2.4
β	0.81	0.73	1.9	0.44	0.78	-0.45

Table 8.8 Tuning of the parameters of the PID regulator

	$M_{\max} = 1.4$			$M_{\max} = 2$		
$f(\gamma)$	a_0	a_1	a_2	a_0	a_1	a_2
αA_P	3.8	-8.4	7.3	8.4	-9.6	9.8
T_I/T_L	5.2	-2.5	-1.4	3.2	-1.5	0.93
T_I/T_F	0.46	2.8	-2.1	0.28	3.8	-1.6
T_D/T_L	0.89	-0.37	-4.1	0.86	-1.9	-0.44
T_D/T_F	0.077	5.0	-4.8	0.076	3.4	-1.1
β	0.4	0.18	2.8	0.22	0.65	-0.051

The Tables 8.7 and 8.8 give the coefficients a_0 , a_1 and a_2 of (8.48) which determine the parameters of the *PI* and *PID* regulators, based on experimental considerations.

Similar tables have been elaborated for integrating processes.

8.4 Handling Amplitude Constraints: “*Anti-Reset Windup*”

Regulator design should take into account the limitations set for the control signal $u(t)$. These limitations may originate from several sources. The limitation may be the property of the actuator structure. Often the actuator is not able to provide an output value higher than a given maximum. For example, a valve in its fully open position provides a maximum flow rate. If it gets a command to transfer a higher value than its maximum, it is not able to execute it, it will be “saturated”. The role of a deliberately applied saturation at the process input is to prevent the process from a harmful level of overexcitation which would cause failure in the process.

Whether the restriction occurs because of the properties of the process, or is artificially introduced into the control loop, its effects have to be taken into account. It is expedient to consider the restrictions already in the phase of regulator design, and to design a regulator whose output signal will not reach the limit value. If this is not possible, then the additional phenomena appearing when the restriction occurs have to be handled.

Thus the linear range of the regulator or of the actuator operated by the regulator is finite. The relation of this amplitude limitation with the design goals has already been discussed in Sect. 7.4. During saturation, the closed-loop control system behaves similarly to the open-loop, as the output of the saturation is constant and thus the input of the process is also constant. The process output is settled according to its dynamics. But in the case of regulators which contain an integrating element, another problem may also occur: if the value of the error signal is high, the regulator output may reach the horizontal section of the saturation characteristic. If the integrator works further on, then the input of the saturation characteristic continuous to increase, and a long time has to pass till the sign of the error signal changes and the input signal gets back to the linear section of the characteristic, if this

Fig. 8.28 Regulator and actuator with saturation

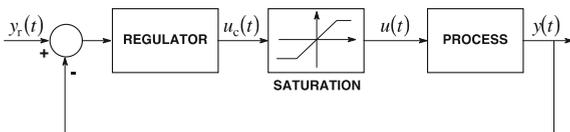
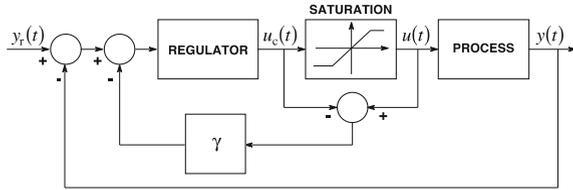


Fig. 8.29 Control system with an extended regulator ensuring the ARW effect



happens at all. Therefore the time of the transients will be increased unacceptably, and also steady oscillations may occur which are harmful to the process.

The problem can be solved by the technique of *Anti Reset Windup* (ARW) or “*antiwindup*”, for short. The main point of this technique is that it uses the model of the static saturation characteristic and with an appropriate feedback it guides the operating point to the crossing point of the linear and the saturated section. The usual saturation characteristic can be described by

$$u(t) = \begin{cases} U_{\max}; & \text{if } u_c(t) > U_{\max} \\ u(t); & \text{if } |u_c(t)| < U_{\max} \\ -U_{\max}; & \text{if } u_c(t) < -U_{\max} \end{cases}, \quad (8.49)$$

which is a more detailed form of (7.39). A closed-loop control system with the saturation is shown in Fig. 8.28.

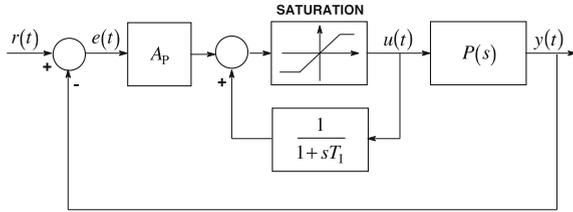
The ARW effect can be reached by the simple feedback illustrated in Fig. 8.29. The extra inner feedback works until the process input is on the horizontal saturation section of the characteristic. This ensures that the regulator output is set to $u(t) = u_c(t)$, belonging to the breakpoint of the characteristic.

In a continuous system this solution can be realized if the signal $u_c(t)$ is available. If the regulator output and the manipulated variable are distinct, then both $u_c(t)$ and $u(t)$ are measurable, and the feedback can be easily realized through the constant element γ . If this is not the case, then a model of the saturated process has to be built. (The realization of such algorithms is much easier in the case of sampled data systems, see Chap. 13.) It has to be ensured an appropriate choice of γ , that the inner feedback acts faster than the dynamics of the process itself.

Another possibility is also available: to reset the integrator component of the regulator when observing saturation. The disadvantage of resetting of the integrator in the case of saturation is that when the regulator comes out of the saturation, there is a mismatch between the state variables of the regulator and those of the process, which results in a deterioration of the control behavior. This can be compensated by a regulator structure where the input of the regulator and the input of the process are similarly restricted, that is, the regulator is put into the feedback path of the saturation. A typical example of this solution is the FOXBORO regulator (Fig. 8.30), which corresponds to a saturated PI regulator.

Now, if there is no saturation, then the overall transfer function of a proportional element fed back through a first order lag element with positive feedback results in the transfer function of a PI regulator:

Fig. 8.30 The FOXBORO regulator



$$C(s) = A_P \frac{1}{1 - \frac{1}{1+sT_I}} = A_P \frac{1+sT_I}{sT_I}. \tag{8.50}$$

In this structure, windup does not occur.

Most solutions set the integrator output to a given value after leaving the saturation. Several methods have been elaborated (some of them are rather complex) to calculate and set the “reset” value. There is no single procedure which ensures in all cases the appropriate behavior, but the above simple procedures in most cases provide satisfactory operation. Several other methods are known for taking the effect of saturation into account, but these are not discussed here.

8.5 Control of Special Plants

In the sequel, two examples will be shown of the regulator design of special plants, namely for a process containing two integrators, and for compensation of unstable processes. It will also be presented how in some cases the design can be executed analytically, approximating the plant by its dominant pole pair model.

8.5.1 Control of a Double Integrator

Let the transfer function of the process be $P(s) = K/s^2$. The process contains two integrators, in a closed-loop control system with feedback unity and with a

Fig. 8.31 NYQUIST diagram of a control system containing two integrators

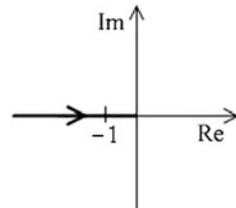
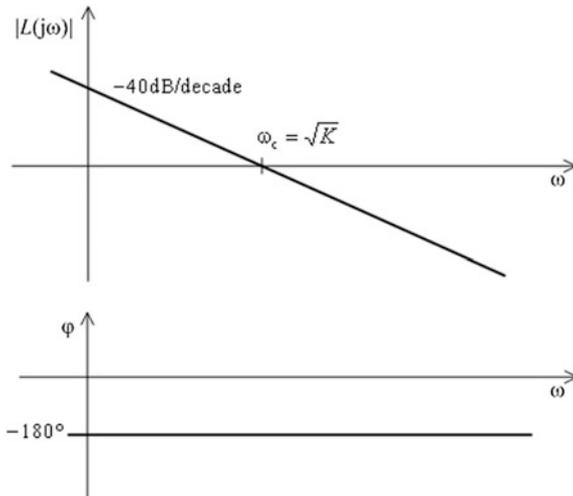


Fig. 8.32 BODE diagram of a control system containing two integrators



proportional regulator working at the borderline of stability. The characteristic equation is $1 + K/s^2 = 0$, or $s^2 + K = 0$. Its roots $s_{1,2} = \pm j\sqrt{K}$ are located on the imaginary axis. Instead of a good control system this system rather realizes a good oscillator. The NYQUIST diagram of the open-loop is shown in Fig. 8.31. The NYQUIST diagram crosses the point -1 , thus the system is at the borderline of stability. The BODE diagram is given in Fig. 8.32.

The quality specifications set for the control system are as follows: stability; the phase margin should be about 60° to ensure an appropriate dynamic response; and for step and ramp reference signals, the tracking error in steady state should be zero (that is, the type number of the control system after the compensation has to remain 2).

These requirements can be fulfilled applying a compensation element which is able to improve the phase conditions. The phase angle of the loop frequency function with the regulator is expressed as $\phi_L(\omega) = \phi_C(\omega) + \phi_P(\omega) = \phi_t(\omega) - 180^\circ$. The regulator has to provide a positive phase margin by $\phi_t(\omega) = 180^\circ + \phi_L(\omega_c) = \phi_C(\omega_c)$, because $\phi_P(\omega) = -180^\circ$. A phase-lead (approximate PD) regulator described by the transfer function

$$C(s) = A \frac{1 + s\tau}{1 + sT}; \quad \tau > T$$

guarantees the addition of a positive phase angle (see also Sect. 2.4), as $\phi_C(\omega) = \arctan(\omega\tau) - \arctan(\omega T) > 0$, if $\tau > T$. The higher the ratio τ/T , the higher the values that $\phi_C(\omega)$ can take. The phase lead element is also called an approximate PD element, as from the form

Fig. 8.33 BODE diagram of a system with two integrators compensated by a *PD* regulator

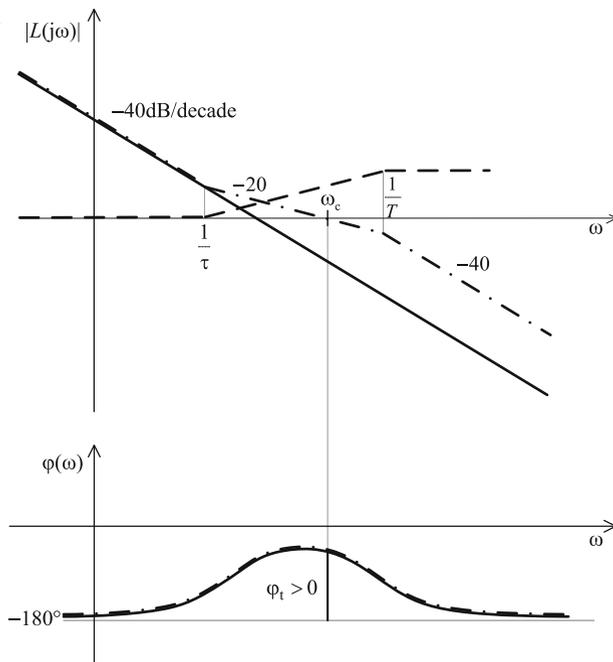
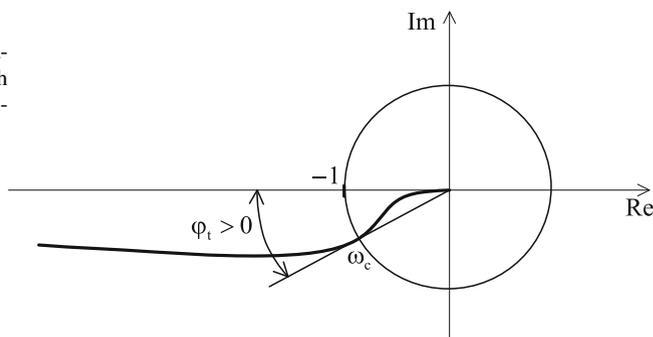


Fig. 8.34 NYQUIST diagram of a system with two integrators compensated by a *PD* regulator



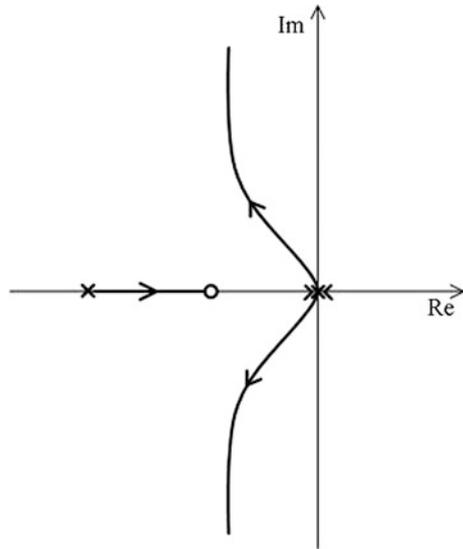
$$C(s) = A \frac{1 + s\tau}{1 + sT} = A \frac{1 + sT + s\tau - sT}{1 + sT} = A \left(1 + \frac{s(\tau - T)}{1 + sT} \right)$$

it can be seen, that $C(s)$ can be obtained as the parallel connection of a proportional and an approximate (that is realizable) differentiating channel. The transfer function of the regulator is often given in the form

$$\tilde{C}_{PD}(s) = \tilde{A}_{PD} \frac{1 + sT_D}{1 + sT}; \quad T_D > T, \tag{8.51}$$

where the notation $\tau = T_D$ is introduced for the differentiating time constant, see also (8.14). Instead of the correct denomination “approximate *PD*”, often the

Fig. 8.35 Root locus of a system with two integrators compensated by a *PD* regulator



slightly simpler denomination “*PD*” is used. Nevertheless this sloppiness is reasonable, since an accurate *PD* regulator without the pole is non-realizable (its high frequency gain would be infinity).

Figure 8.33 shows that with the *PD* regulator a straight line section of -20 dB/decade slope can be formed on the BODE amplitude-frequency curve. The cut-off frequency is placed on this section. Thus the system will have a positive phase margin; its performance will be fast, as ω_c is shifted toward the higher frequency domain. The gain of the regulator may be chosen to maximize the phase margin. The reachable maximum phase margin depends on the ratio T_D/T . This ratio also influences the maximum value of the control signal appearing at the input of the process at the time instant when the unit step reference signal is switched on. The maximum value of the control signal is $u_{\max} = A_p T_D/T$.

As seen from the BODE diagram, the control system is structurally stable: the phase margin is positive for any value of the loop gain.

The NYQUIST diagram of the compensated system is shown in Fig. 8.34. (Note that in the case of processes containing integrators, when there are poles at the origin of the complex plane, it is not necessary to apply the generalized NYQUIST criterion to evaluate the stability from the properties of the conformal mapping of the closed curve shown in Fig. 5.18 surrounding the origin by a circle of infinitesimal radius according to the loop frequency function. It is sufficient to plot the NYQUIST diagram only for the positive frequencies and to check whether when going through the curve from $\omega = 0$ to $\omega = \infty$, the point $-1 + 0j$ is to the left of the curve or not. If it is to the left, then the control system is stable, and the phase margin or the gain margin can be used to measure the distance from the borderline of stability.)

Figure 8.35 gives the root locus of the system. The points of the root locus, which are the roots of the characteristic equation, lie in the left half of the complex plane for all gain values, indicating that the control system is structurally stable.

8.5.2 Control of an Unstable Plant

In state space, the state variables are attached to the poles of the transfer functions of the process and the regulator. These state variables together form the state variables of the open-loop. When an undesired pole of the process is cancelled by a zero of the regulator, actually the corresponding state variable becomes inaccessible from the output or from the input side, namely the system becomes unobservable or uncontrollable (Sect. 3.4). But in spite of the fact that these variables do not appear in the overall transfer function of the loop, they remain parts of the system. To ensure the stability of the control system, not only do the poles of the transfer function have to be at left half side of the complex plane, but so do the unobservable and the uncontrollable poles.

Unstable poles of an unstable process must not be cancelled by the zeros of the regulator. This prohibition can be justified also by the fact that as seen in Chap. 4, the behavior of a closed-loop control system is characterized not only by the overall transfer function between the output and the reference signal, but also by the overall transfer functions between the output signal and the input and the output disturbances, and the overall transfer function between the control and the reference signal. The unstable pole does appear in the transfer function between the output signal and the input disturbance, thus in spite of the compensation the instability of the control system is maintained.

A further consideration is that in real systems, the values of the parameters are not accurate: generally they are obtained by measurements and lie within a range of their possible values. Therefore, an accurate cancellation of an unstable pole is not possible, and the instability is maintained in the control system. This phenomenon can be illustrated by the root locus. Let us consider as an example the control system in Fig. 8.36b. The loop transfer function is given by a proportional system with two lags, where one pole is unstable. From the root locus (Fig. 8.36a) it can be

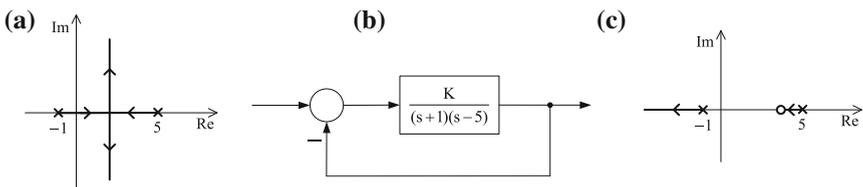


Fig. 8.36 With imperfect zero-pole cancellation the root locus has a branch on the right half of the complex plane

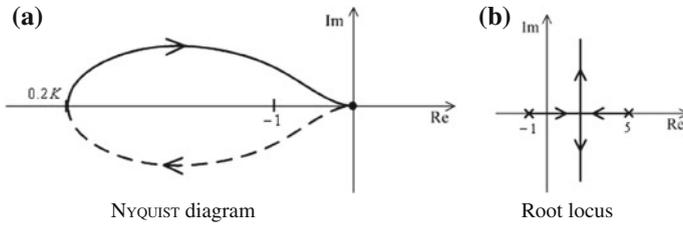


Fig. 8.37 Control of an unstable process with a proportional regulator: the system can not be stabilized!

seen that the closed-loop control system is unstable for all values of the gain. If the unstable pole could be accurately cancelled by the zero of the regulator, the control system would become structurally stable and the root locus would have a single branch on the left half of the complex plane. But as practically perfect cancellation can not be realized, there will remain a branch of the root locus on the right half of the complex plane, and thus the closed-loop control system remains unstable (Fig. 8.36c). (Note that in compensation, a zero by itself can not be realized, it always appears together with a pole.)

When compensating unstable processes, the generalized NYQUIST stability criterion has to be considered to ensure the stable behavior of the closed-loop control system. *PID*-like regulators can be used as compensating elements in such cases as well.

Example 8.1 Let us analyze whether the processes given by the transfer function

$$P_1(s) = \frac{1}{(s + 1)(s - 5)} = -\frac{0.2}{(1 + s)(1 - 0.2s)} \tag{8.52}$$

and

$$P_2(s) = \frac{1}{(s - 1)(s + 5)} = -\frac{0.2}{(1 - s)(1 + 0.2s)} \tag{8.53}$$

can be stabilized by the proportional regulator $C(s) = A_P$ or not.

As there is an unstable pole in the open-loop, stable behavior can be ensured if the NYQUIST diagram encircles the point $-1 + j0$ once anticlockwise.

For the first process, the NYQUIST diagram of the open-loop is shown in Fig. 8.37a. As seen, the diagram can encircle the point $-1 + 0j$ only clockwise, thus this process can not be stabilized by a proportional regulator. This property is demonstrated also with the root locus, which contains a pole in the right half-plane at any gain value. (Here, the stabilization of the system can be tried with a *PD*-like compensation, such as $C(s) = A_P(1 + 0.2s)/(1 + 0.02s)$.)

The second process can be stabilized by a proportional regulator, as the direction of encircling by the NYQUIST diagram is counterclockwise, thus choosing an

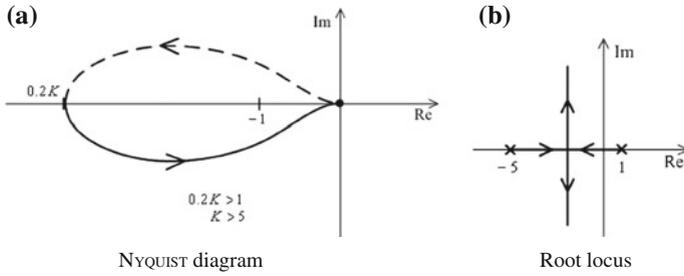


Fig. 8.38 Control of an unstable process with a proportional regulator: the system can be stabilized

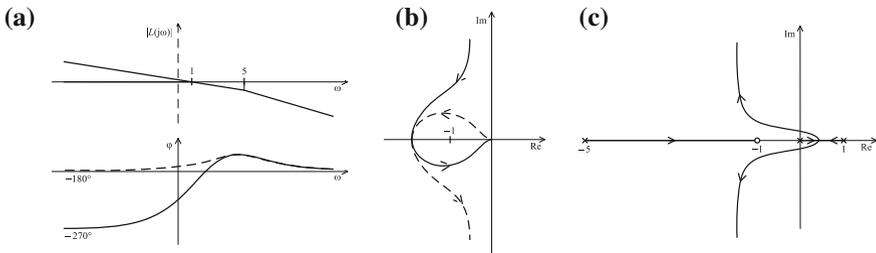
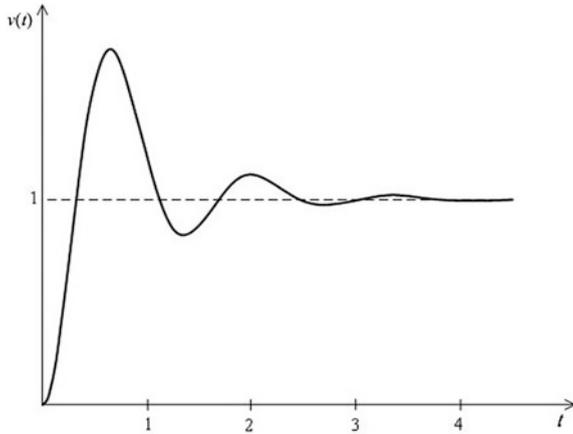


Fig. 8.39 BODE and NYQUIST diagram, and the root locus of the unstable process compensated by a PI regulator

Fig. 8.40 Unit step response of an unstable process compensated by a PI regulator



appropriate gain ($K > 5$) the NYQUIST diagram will encircle the point $-1 + 0j$ once (Fig. 8.38a). This is also demonstrated by the root locus shown in Fig. 8.38b. By increasing the gain, the root locus will get into the left half of the complex plane.

The type-number of the control system is zero, thus it has a static error. The static accuracy can be improved by a *PI* regulator which does not cancel the unstable pole. The transfer function of the regulator is: $C(s) = A_P(1 + s)/s$.

The loop transfer function is:

$$L(s) = C(s)P_2(s) = -A_P \frac{1+s}{s} \frac{0.2}{(1-s)(1+0.2s)}. \quad (8.54)$$

The BODE diagram of the original and of the compensated system is shown in Fig. 8.39a, the NYQUIST diagram of the compensated system (whose course can be derived from the BODE diagram) is given in Fig. 8.39b, and the shape of the root locus is provided in Fig. 8.39c. It can be seen that increasing the gain beyond a defined value the closed-loop control system will be stable, the generalized NYQUIST diagram encircles once the point $-1 + 0j$ in the positive (counterclockwise) direction. The parameter A_P can be designed for maximum phase margin. (The concept of phase margin can be used in this case as well.) Figure 8.40 shows the unit step response of the control system. ■

Example 8.2 The transfer function of an unstable process is:

$$P(s) = \frac{0.5}{(s - 0.1)(s + 1)(s + 5)} = -\frac{1}{(1 - 10s)(1 + s)(1 + 0.2s)} \quad (8.55)$$

Let us design a regulator which ensures stable behavior, tracking the unit step reference signal without static error, and the initial value of the control signal does not exceed the value of 50.

A qualitatively correct NYQUIST diagram of the open-loop with a proportional regulator is shown in Fig. 8.41. As the loop transfer function has one pole in the right half-plane, the closed-loop will be stable if $-1 + 0j$ lies within the left side curve in the figure. The condition for that is that the usually interpreted phase margin indicated in the figure be positive.

The asymptotic BODE amplitude-frequency and the phase-frequency diagrams are shown in Fig. 8.42. To reach a more favorable phase margin, *PD* compensation is applied. Thus the section of slope -20 dB/decade is elongated, and the cut-off frequency can be relocated to $\omega_c \approx 1$. To remove the static error, a further *PI* regulator is used. The transfer function of the entire *PID* compensator is

$$C(s) = 10 \frac{1 + 10s}{10s} \frac{1 + s}{1 + 0.2s}. \quad (8.56)$$

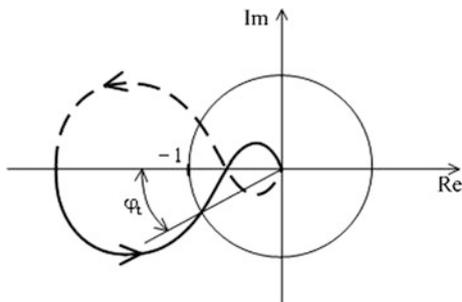


Fig. 8.41 NYQUIST diagram of an unstable process compensated by a proportional regulator

The loop transfer function is:

$$L(s) = -\frac{1 + 10s}{s(1 - 10s)(1 + 0.2s)^2}. \tag{8.57}$$

In the BODE amplitude diagram at frequency $\omega = 0.1$, the breakpoint disappears because of the contradictory effects of the zero and the unstable pole, but it remains as a corner-point, where the phase angle asymptotically changes from -270° to -90° .

The cut-off frequency and the phase margin of the open-loop are $\omega_c = 0.964$, and $\phi_t = 56.25$. The initial value of the control signal in the case of a unit step reference signal is just 50. The output and the control signals are shown in Fig. 8.43. ■

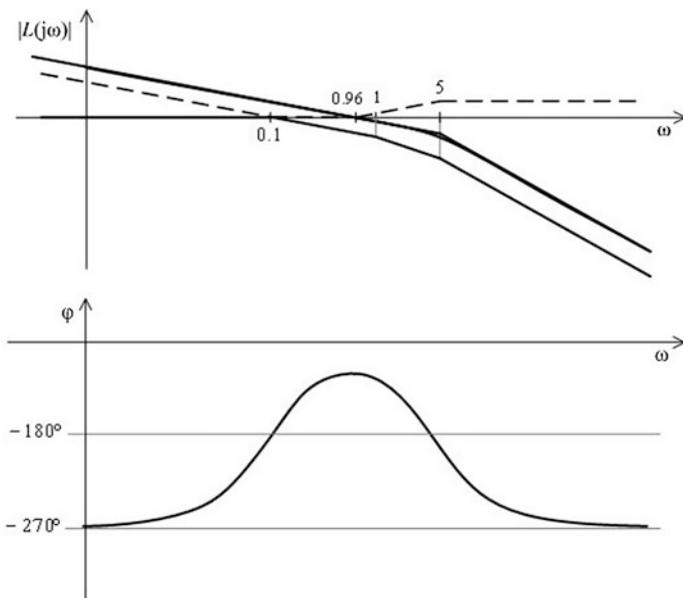


Fig. 8.42 BODE diagram of an unstable process compensated by a PID regulator

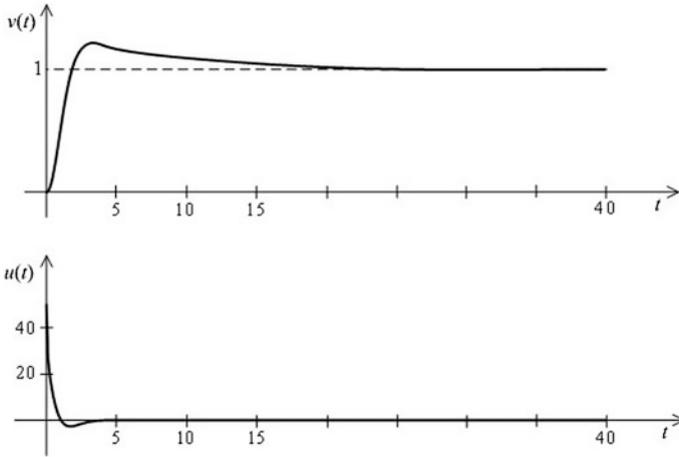


Fig. 8.43 Output and control signals of an unstable process compensated by a *PID* regulator in the case of a unit step reference signal

8.6 Regulator Design Providing a 60° Phase Margin by Pole Cancellation

A regulator is designed so that the process (or its model) meets the quality requirements set for the control system. A common compensation technique is pole cancellation, when the zeros of the transfer function of the regulator are chosen equal to the poles of the process: the unfavorable poles of the process are “cancelled”, and instead more favorable poles are introduced. As an example let us consider a proportional process with three time lags. The transfer function of the process is

$$P(s) = \frac{1}{(1 + sT_1)(1 + sT_2)(1 + sT_3)}; \quad T_1 > T_2 > T_3. \quad (8.58)$$

- (a) Suppose the prescription for the control system is stable behavior and an overshoot less than 10%. This latter requirement can be fulfilled in the frequency domain by ensuring a phase margin of about 60°.

The requirements can be met by applying a simple proportional regulator: $C(s) = A_P$. The approximate BODE diagram of the open-loop is shown in Fig. 8.44. To ensure stability the cut-off frequency ω_c has to be placed at a straight line section of slope -20 dB/decade. To reach the required phase margin, ω_c is located at the frequency where the phase angle is $\varphi = -120^\circ$. First the loop frequency function is analyzed supposing $A_P = 1$ (dotted line in the figure), then A_P is set to the reciprocal of the amplitude belonging to the phase angle $\varphi = -120^\circ$.

The given requirements can be fulfilled with a proportional regulator. The control system will be slow, as ω_c has to be placed at the straight line section of slope -20 dB/decade, which is in the low frequency domain. The control system is of 0-type, so it tracks the unit step reference signal with a static error, whose value depends on the loop gain.

- (b) Suppose the prescription for the control system is stable behavior and an overshoot less than 10%. Furthermore, that the static error be zero for step reference signal.

These prescriptions can be fulfilled by a *PI* regulator.

$$C_{PI}(s) = A_P \frac{1 + sT_I}{sT_I} \tag{8.59}$$

Let us choose the time constant T_I equal to the largest time constant of the process, $T_I = T_1$ (we “cancel” the largest time constant of the process, and “introduce” an integrating effect instead). According to Fig. 8.45 a long straight line section of -20 dB/decade slope is formed in the low frequency domain of the BODE amplitude diagram of the open-loop. Changing the gain A_P of the regulator, the BODE amplitude diagram is shifted parallel until the cut-off frequency is located to ensure the required $\sim 60^\circ$ phase margin.

With a *PI* regulator the type number will be 1, and besides meeting the prescriptions for stability and dynamic response, the control system also fulfills the static requirements. But as the cut-off frequency can be placed only in the low frequency range, the control system will be slow.

- (c) Let the prescription for the control system be stable behavior and an overshoot less than 10%, as well as that the operation of the control system has to be faster.

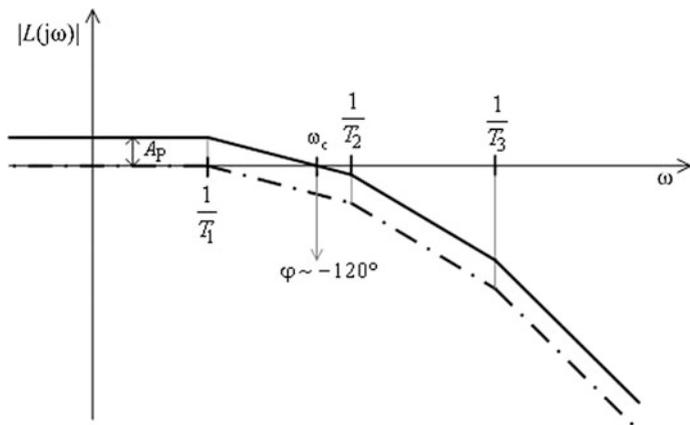


Fig. 8.44 Series compensation with proportional regulator

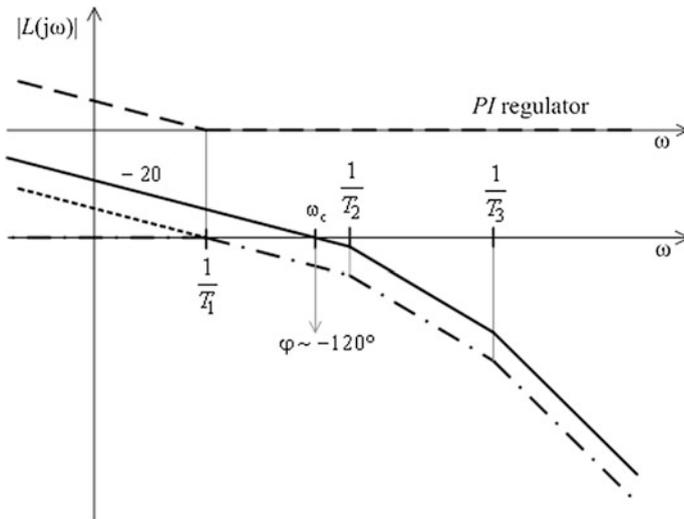


Fig. 8.45 Series compensation with a *PI* regulator

These specifications can be fulfilled by using a *PD* regulator.

$$C_{PD}(s) = A_P \frac{1 + sT_D}{1 + sT}; \quad T_D > T \tag{8.60}$$

Let us choose the time constant T_D equal to the second largest time constant of the process, $T_D = T_2$ (that is, equal to that time constant for which the slope changes from -20 dB/decade to -40 dB/decade at the corresponding breakpoint of the BODE amplitude diagram). The ratio $\eta = T_D/T$ is chosen according to the practical limit of the control signal. (We “cancel” the unfavorable time constant of the process and “introduce” a much smaller time constant instead.) Then changing the gain A_P of the regulator, the BODE amplitude diagram is shifted parallel until the cut-off frequency is located to ensure the required $\sim 60^\circ$ phase margin. The effect of the compensation on the BODE diagram of the open-loop is shown in Fig. 8.46.

The control system will be stable, it has a small overshoot, it will be fast, but as it remains of 0-type, it will have a static error, depending on the loop gain when tracking a unit step reference signal. The acceleration results from the high initial value $u(t = 0) = A_P \eta$ of the control signal.

- (d) Let the prescription for the control system be stable behavior, an overshoot less than 10%, fast operation and zero static error for tracking a step reference signal.

These prescriptions can be fulfilled by a *PID* regulator, combining the possibilities of the *PI* and the *PD* regulators.

$$C_{\text{PID}}(s) = A_P \frac{1 + sT_1}{sT_1} \frac{1 + sT_D}{1 + sT} \quad (8.61)$$

Two of the four free parameters are chosen considering the poles of the process. The parameter T_1 is chosen equal to the largest time constant of the process, and the parameter T_D is set equal to the second largest time constant. With this choice $T_1 = T_1$ and $T_D = T_2$ the loop transfer function can be simplified, that is the “introduced” zeros “cancel” poles of the process.

$$\begin{aligned} L(s) = C(s)P(s) &= A_P \frac{1 + sT_1}{sT_1} \frac{1 + sT_2}{1 + sT} \frac{1}{(1 + sT_1)(1 + sT_2)(1 + sT_3)} \\ &= \frac{A_P}{sT_1(1 + sT)(1 + sT_3)} \end{aligned} \quad (8.62)$$

It can be seen that the transfer function of the residual system became simpler, thus the further steps of the design become easier.

The remaining two parameters are chosen considering the prescriptions set for the acceleration and the overexcitation. The parameter T is chosen based on the pole placement ratio. The phase margin (and the overshoot of the step response) can be set with A_P . It can be seen in the BODE amplitude diagram of the open-loop that the section of slope -20 dB/decade will be longer because of the choice $T < T_D$. With the gain A_P of the regulator, the BODE amplitude diagram is shifted parallel until the cut-off frequency is located to ensure the required $\sim 60^\circ$ phase margin. This is done by checking the frequency where the phase angle is about -120° , and A_P is then set to the reciprocal of the amplitude corresponding to this

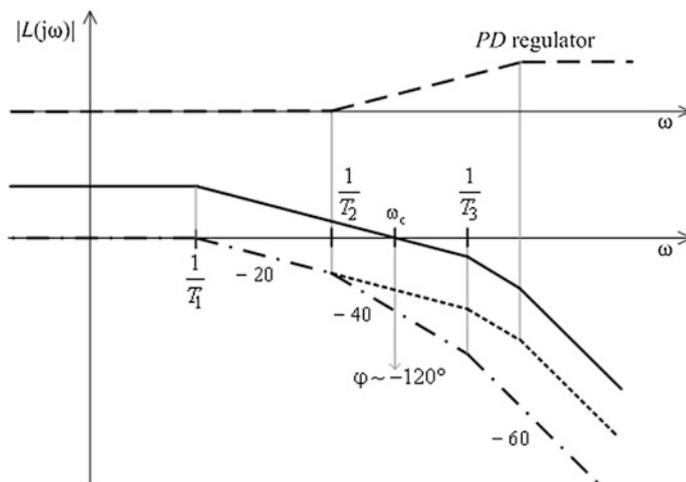


Fig. 8.46 Series compensation with a PD regulator

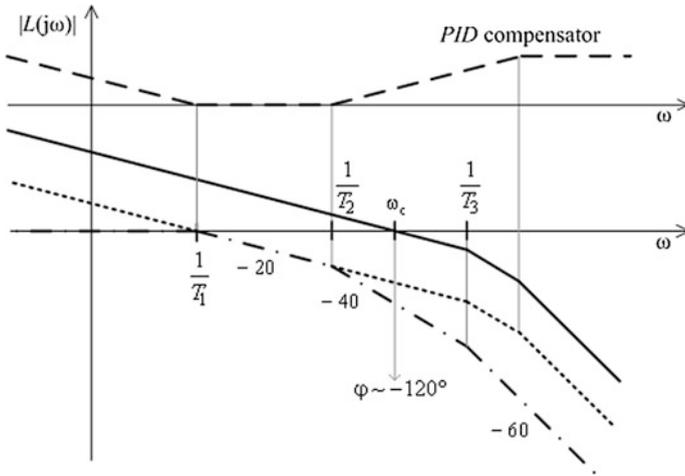
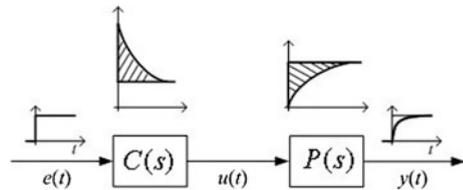


Fig. 8.47 Series compensation with PID regulator

Fig. 8.48 In compensation pole cancellation is virtual



frequency. This frequency will be ω_c , which is located now in the higher frequency domain, therefore the control system will be faster (Fig. 8.47).

But let us observe that the pole cancellation is just formal: in reality the poles do not disappear. The process can not be changed, its poles do exist. Actually the zeros and the poles do not cancel each other, only their effects compensate each other. Figure 8.48 demonstrates that in the case of zero-pole cancellation the overall transfer function of the serially connected regulator and process behaves as if a real pole cancellation has happened, but the effect of the zero of the regulator does appear in the signal $u(t)$. The overexcitation in the control signal depends on the ratio of the zero and the pole of the regulator. The so called *acceleration area* in the control signal decreases the so called *decelerating area* of the process which characterizes the settling time of its unit step response, thus yielding a faster response of the control system.

The main point of the pole cancellation method is that the unfavorable poles of the process are cancelled, and the poles of the regulator ensure a more favorable dynamics for the control system. As no real pole cancellation occurs, it is not necessary to set the zeros of the regulator quite accurately. The tuning can also be refined later, by moving the zeros a bit away from their pole cancellation location.

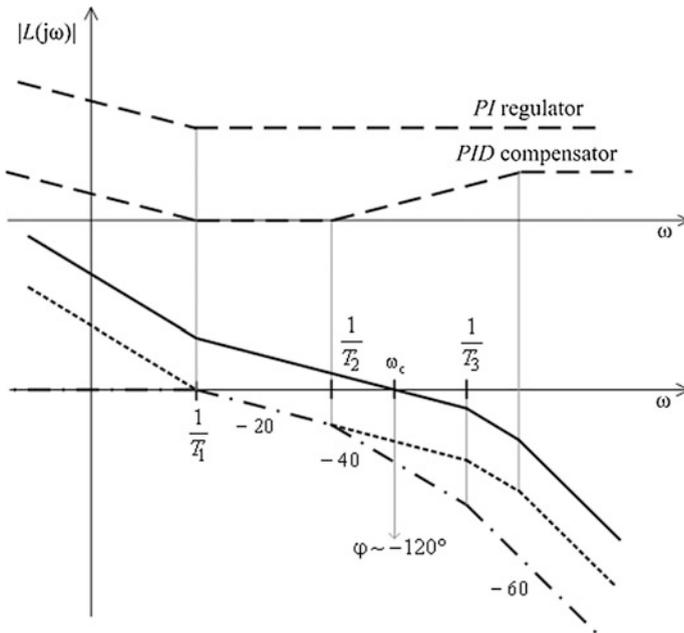


Fig. 8.49 Series PID compensation with additional PI regulator

(e) Let the prescription for the control system be its stable behavior, an overshoot less than 10%, fast operation, and zero static error for tracking both a step and also a ramp reference signal.

To meet the static requirements the control system has to be of 2-type containing two integrating effects. The low frequency part of the Bode diagram has to be of -40 dB/decade.

Table 8.9 PID-like regulators designed for a proportional system with three lags and the characteristic measures of the control systems

	$C(s)$	$L(s)$	e_∞	u_{\max}	ω_c
<i>P</i>	7.51	$\frac{7.51}{(1 + 10s)(1 + s)(1 + 0.2s)}$	0.1174	7.51	0.62
<i>PI</i>	$5.04 \frac{1 + 10s}{10s}$	$\frac{5.04}{10s(1 + s)(1 + 0.2s)}$	0	5.04	0.45
<i>PD</i>	$16.55 \frac{1 + s}{1 + 0.2s}$	$\frac{16.55}{(1 + 10s)(1 + 0.2s)^2}$	0.057	82.7	1.51
<i>PID</i>	$14.27 \frac{1 + 10s}{10s} \frac{1 + s}{1 + 0.2s}$	$\frac{14.27}{10s(1 + 0.2s)^2}$	0	71.36	1.33
<i>PI – PID</i>	$14.3 \frac{1 + 50s}{50s} \frac{1 + 10s}{10s} \frac{1 + s}{1 + 0.2s}$	$\frac{14.3(1 + 50s)}{500s^2(1 + 0.2s)^2}$	0	71.5	1.34

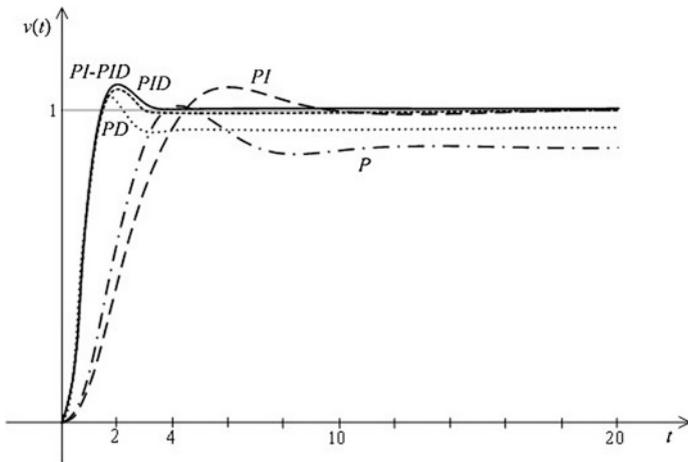


Fig. 8.50 Unit step responses of the compensated control systems

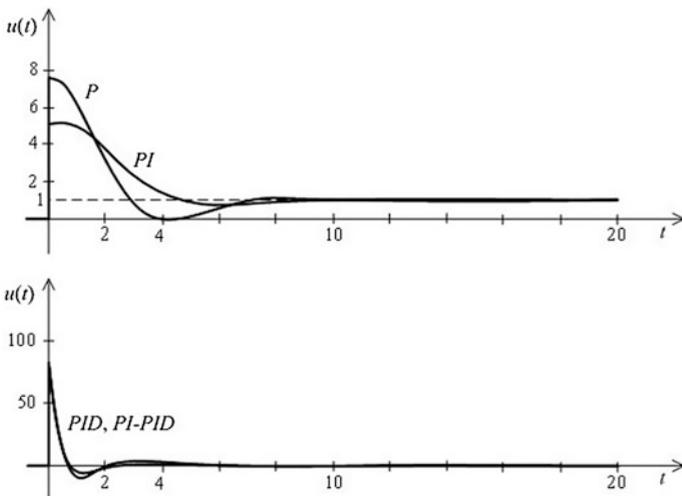


Fig. 8.51 Control signals of the compensated control systems

The previous regulator designed for pole cancellation is extended by a *PI* effect, which forms the BODE diagram of Fig. 8.49. The transfer function of the regulator is

$$C_{PI-PID}(s) = A_P \frac{1 + sT_{I1}}{sT_{I1}} \frac{1 + sT_{I2}}{sT_{I2}} \frac{1 + sT_D}{1 + sT}, \quad (8.63)$$

where $T_{I1} > T_{I2}$. This ratio is selected by the designer, its advisable value is $T_{I1} \approx 5T_{I2}$.

This regulator, considering Fig. 6.3, ensures not only a better reference signal tracking, but also a favorable disturbance rejection and less sensitivity to parameter changes.

Example 8.3 Let the time constants of the above process be $T_1 = 10$, $T_2 = 1$, $T_3 = 0.2$. The regulators designed for the above requirements and the characteristics of the control system are given in Table 8.9.

Figure 8.50 shows the unit step responses of the control circuits, whereas Fig. 8.51 presents the corresponding control signals. As can be seen, the static error is zero only in the cases when there is an integrating effect in the regulator. The control system with the P and the PI regulator is slow, whereas with the PD , PID and the $PI-PID$ regulators it is fast, at the cost of high overexcitation. ■