PID CONTROL

- The basic configuration of the SISO control system is shown in the block diagram below (Fig. 1.). The block diagram gives the notation used for the signals in the control system.

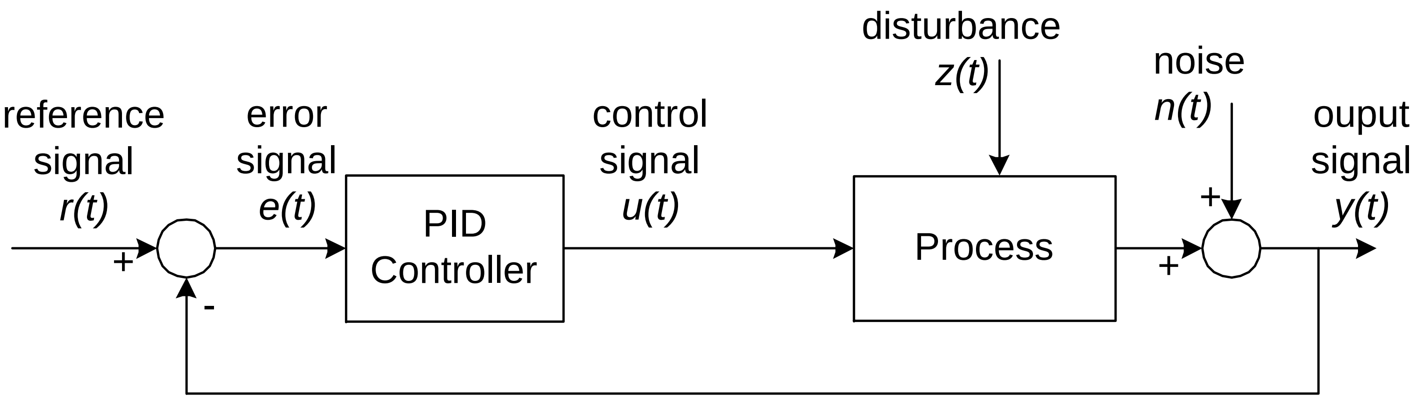


Fig. 1. Basic control system configuration.

- The block marked *»Process«* in Figure 1. includes all the elements of the control system which are considered as the parts of the process: actuator, plant, and sensor.
- The mathematical model of the process can be very complex, with complicated static and dynamic description.
- The identification of a complex model requires a lot of engineering effort.
- Since the performance obtained from control system with the PID controller is limited, many PID controller tuning methods use simple models which have similar complexity as the PID controller.
- These models require simple identification experiments and capture dominant dynamic properties. Usual representation of these models is low-order (first or second order) transfer function in Laplace domain.

- On the other hand, some tuning methods for PID controllers were developed for more complex process models (e.g. higher order models, models with non-linear characteristics) because of successful and widespread use of PID controllers in industry.
- A transfer function modeling the process is generally represented as follows:

$$G_P(s) = K\frac{1+b_1 s+b_2 s+b_2 s^2+\ldots+b_m s^m}{s^k\left(1+a_1 s+a_2 s+a_2 s^2+\ldots+a_n s^n\right)}e^{-T_t s}, \text{ with } m \le n, \qquad (1)$$

where is:

- o T_t - dead time,
- o m, n - degrees of complex variable polynomials, and
- o k - number of integrators present in the process.

- The type of the process model is determined by the exponent k, so that a process without an integrator (k=0) is called »type 0 process«, a process with an integrator (k=1) is called »type 1 process«, and so on.

- Two more common process models are used. The first model is the First Order with Dead Time (FODT) model, often used for the description of chemical processes, given by the transfer function:

$$G_P(s) = G_1(s) = \frac{K_1 e^{-T_{t1}s}}{1 + T_1 s} \ . \tag{2}$$

- The second model is frequently employed to describe electromechanical processes. It consists of an integration and a first order lag:

$$G_P(s) = G_2(s) = \frac{K_2}{s(1 + T_2 s)} \ . \tag{3}$$

- A PID controller consists of the three terms: proportional (P), integral (I), and derivative (D). Its behavior can be roughly interpreted as the sum of the three term actions:
  - o the P term gives a rapid control response and a possible steady state error;
  - o the I term eliminates the steady state error; and
  - o the D term improves the behavior of the control system during transients.

- A PID-type controller can be implemented variously. The following subsections describe how controllers apply the PID control law through the review of different PID controller forms and implementation aspects.
- The section proceeds with an outline of
    - different tuning rules and
    - explains the usage of PI controllers in dead-time compensating controllers.

1. Forms of the PID controller

- Different forms of PID controller reflect the development of the PID algorithm in different technologies and its use in diverse control systems.
- Besides, some PID forms ensure better performance and behavior of the control system than others.
- The textbook version of the PID control law in the time domain is:

$$u(t) = K_P e(t) + K_I \int_0^t e(\tau) d\tau + K_D \frac{de(t)}{dt} = K_P (e(t) + \frac{1}{T_I} \int_0^t e(\tau) d\tau + T_D \frac{de(t)}{dt}), \qquad (4)$$

where is:

- K_P - proportional gain,
- K_I - gain of the integral term,
- K_D - gain of the derivative term,
- T_I - integral time constant, and
- T_D - derivative time constant.

- The Laplace transformation of equation (4) gives the transfer function of the PID controller:

$$G_R(s) = \frac{U(s)}{E(s)} = K_P(1 + \frac{1}{T_I s} + T_D s) = K_P \frac{1 + T_I s + T_D T_I s^2}{s}. \quad (5)$$

- Since the numerator of the PID controller transfer function in (5) has a higher degree than the denominator, the transfer function is not causal and as such can not be realized.

- The form (5) of the PID controller is modified through the addition of a lag to the derivative term:

$$G_R(s) = \frac{U(s)}{E(s)} = K_P(1 + \frac{1}{T_I s} + \frac{T_D s}{1 + \frac{T_D}{N} s}), \quad (6)$$

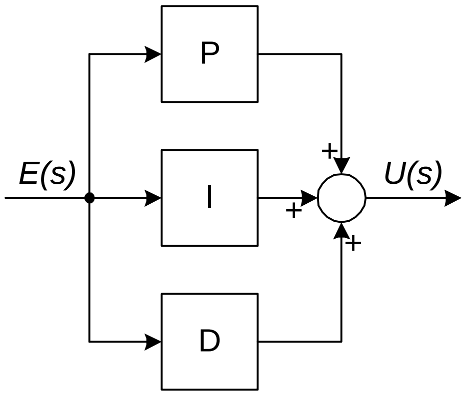
where is:

- T_D/N - time constant of the added lag.

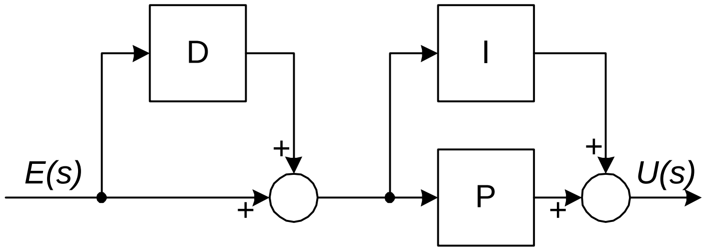
- Divisor N in (6) determines the gain K_{HF} of the PID controller in the high frequency range:

$$K_{HF} = \lim_{\omega \to \infty} G_R(j\omega) = K_P(1 + N) \, . \qquad (7)$$

- The gain K_{HF} must be limited because measurement noise signal *n(t)* often contains high frequency components and its amplification should be limited. Usually, the divisor N is chosen in the range $3 \div 10$.



a) Parallel form;



b) Series form;

Fig. 2. Forms of the PID controller.

- The form of the PID controller defined by (5) is called the **parallel** or **non-interactive** form (Figure 2.).
- Another form of the PID controller is the **series** or **interacting** form (Fig. 2.b.) with transfer function:

$$G_{RS}(s) = K_{PS}(1 + \frac{1}{T_{IS}s})(1 + T_{DS}s), \tag{8}$$

- The form (8) of the PID controller has a simple representation in the frequency domain, since all roots and zeros of $G_{RS}(s)$ are real and correspond to the inverses of the break frequencies.
- Based on equations (5) and (8), the relations for converting the parameters between the parallel and the series form of the PID controller are:

$$K_{PS} = \frac{K_P}{2}(1 + \sqrt{1 - 4T_D / T_I}), \tag{9}$$

$$T_{IS} = \frac{T_I}{2}(1 + \sqrt{1 - 4T_D / T_I}), \tag{10}$$

$$T_{DS} = \frac{T_I}{2}(1 - \sqrt{1 - 4T_D / T_I}), \tag{11}$$

- Equations (9) – (11) can be used only if $T_I \geq 4T_D$., i.e. when poles and zeros of the parallel form are real.
- As output of the above forms of the PID controller is the total value of the control signal *u(t)*, they are called positional (continuous version) or absolute (discrete versions) PID algorithms.
- Some actuators such as a motor may use the increment or derivative of the control signal as an input signal, because they have built-in integral action. PID controllers with such an output are termed velocity (continuous version) or incremental (discrete version) PID controllers.
- Standard PID controllers act on the error signal *e(t)* and give the control signal *u(t)* as the output. Such configuration of the controller uses the same parameters as in responding to set-point change (tracking) and to load disturbance (regulating). The two functions of the control system often impose contradictory demands on the value of the controller parameters. The contradiction is resolved by a trade-off in the controller's design.

- In order to avoid this trade-off a modification of the PID controller structure was devised (Figure 3.).

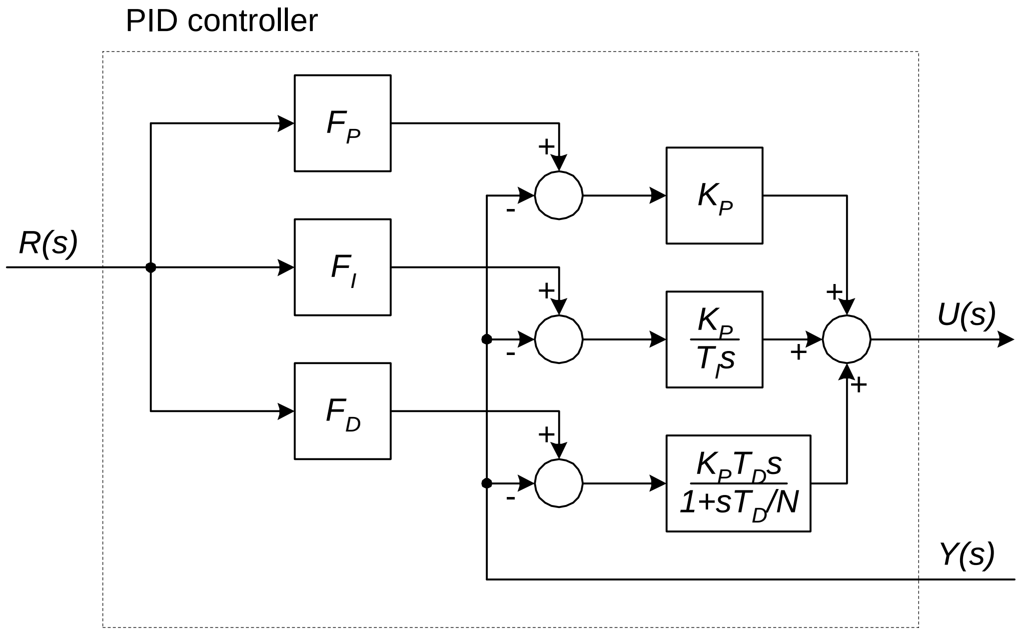


Fig. 3. Two-degrees of freedom PID controller

- Signal channels for reference signal and for measurement signal are separated, and a set of weights (F_P, F_I, F_D) in the channel of the reference signal is introduced.

- The PID controller with set-point weighting is tuned in two steps:
  - Controller parameters (K_P, T_I, T_D) are tuned for good regulation;
  - Weights F are adjusted in order to set zeros of the closed-loop transfer function and thus to improve the tracking behavior of the control system.
- Controllers that allow such separation of the design for regulating and for tracking are called two-degree-of-freedom controllers.
- The introduction of arbitrary weights in the reference channel of the PID controller gives design more freedom.
- In some PID controller implementations weights are set to $F_D = 0$ and $F_P = 0$ in order to avoid derivative and proportional bumps, which are present in response to step set-point change. Furthermore, the weight of integral term is set to $F_I = 1$.

- The standard form of the PID controller, similar to the structure depicted in Figure 3, and recommended by Instrument Society of America, is given below (K. J. Åstrom and T. Hägglund; PID Controllers: Theory, Design and Tuning; Instrument Society of America, North Carolina, 1995):

$$U(s) = K_R[(F_P R(s) - Y(s)) + \frac{1}{sT_I}(R(s) - Y(s)) + \frac{sT_D}{1 + sT_D / N}(F_D R(s) - Y(s))]. \quad (12)$$

- Frequently, only a part of a PID controller is used. Åstrom and Hägglund (1995) have noted that most control loops are of the PI type.

- As a rule, the PI controller is used for processes of the first order, or for processes not requiring tight control.

- A PD controller can be used for processes which contain integrators, and which do not have constant load disturbances, since PD controller can not compensate for it.

- The application of a P controller is limited to simple control tasks.

- Today, almost all control strategies are implemented as digital algorithms in various devices such as Programmable Logic Controllers (PLCs), Digital Signal Processors (DSPs), and in other microprocessor-based equipment.
- To become applicable in such equipment, the PID control algorithm has to be discretized. Using Euler integration method – rectangular integration, the discrete version of the positional algorithm (4) is calculated as:

$$u(k) = K_P \left[e(k) + \frac{T}{T_I} \sum_{i=0}^{k-1} e(i) + \frac{T_D}{T} \left(e(k) - e(k-1) \right) \right], \qquad (13)$$

where is:

- k - discrete time instant, and
- T - sampling time.

- Recursive equation describing the incremental version of the PID algorithm is obtained when the equation (13) for the time instant k-1 is subtracted from the same equation for the time instant k:

$$\Delta u(k) = u(k) - u(k-1) = q_0 e(k) + q_1 e(k-1) + q_2 e(k-2), \qquad (14)$$

where is:

$$q_0 = K_P \left(1 + T_D/T\right),$$

$$q_1 = K_P \left(1 + 2\,T_D/T - T/T_I\right) \text{ and}$$

$$q_2 = K_P\, T_D / T.$$

- Other relations for parameters q_i in (14) are obtained if a different integration method (e.g. trapezoidal method) is used.
- These discrete approximations of the continuous PID controller are valid only if the sampling time T is sufficiently short in comparison to the time constants of the controller.

- Otherwise, when sampling time T is not much shorter than the time constants of the controller, connection with continuous PID controllers is dropped and Z-transform form of the PID controller is used (Isermann, 1989).
- Discrete controller of the second order with an integrator has a transfer function in the Z-domain:

$$G_C(z) = \frac{q_0 + q_1 z^{-1} + q_2 z^{-2}}{1 - z^{-1}}. \qquad (15)$$

- Polynomial coefficients q_i in (15) have to satisfy the following relations (Isermann, 1989):

$$q_0 > 0,$$

$$q_1 < -q_0,$$

$$-(q_0 + q_1) < q_2 < q_0,$$

so that the obtained digital controller has a dynamic behavior of the continuos PID controller.

2. Practical issues in the application of PID control

- All controllers are designed to work with processes which have some physical constraint: valves have a limited operating range (0%-100%), pumps have limited power, motors have a maximum moment, and so on.
- These limitations can be regarded as non-linearities in the process and have to be considered in the application of the controller.
- Since many of these limitations appear at the input of the actuator (process), they are referred to as input limitations and are modeled with a non-linear element having a saturation characteristic, as shown in Figure 4.

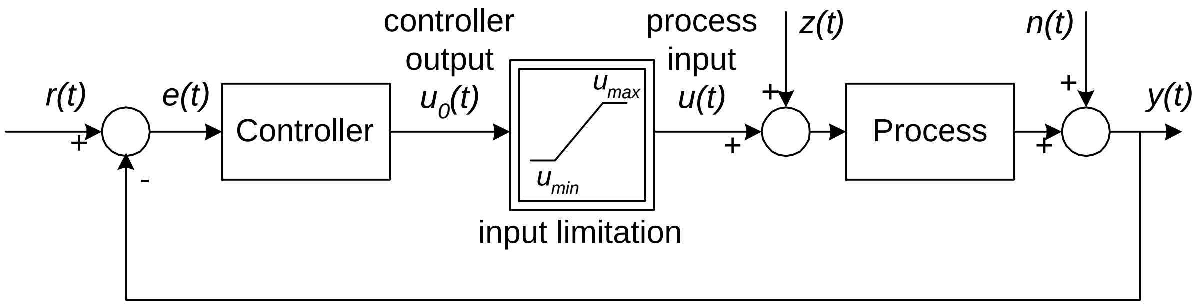


Fig. 4. Input limitation as a part of the control system.

- Beside input magnitude limitations, actuators often have defined rate limitations, or maximum rates at which the control signal $u(t)$ can be changed.
- When the controller output signal $u_0(t)$ exceeds the upper limit u_{max}, or when it falls below the lower limit u_{min} of the operating range, its value changes in the input limitation element so that the controller output signal $u_0(t)$ and the process input $u(t)$ do not coincide.
- Shouldn't the two coincide, the integrator in a controller with integral action would produce an inaccurate and highly excessive value which would cause oscillation and slowing down of the transient response. In other words, the effect would be a large overshoot and a long settling time.
- This behavior is called the **integrator windup**.
- Moreover, the feedback loop during the windup behaves as if it were broken.
- There are several **anti-windup** algorithms to avoid adverse effects of the integrator windup on the control system performance.

- Figure 5 shows the structure of the linear feedback anti-windup algorithm.
- A new signal $e_{aw}(t)$ is added as an additional input to the integrator of the controller. It is active when there is difference between the controller output $u_0(t)$ and the process input $u(t)$.
- It acts in direction opposite to the windup effect. The rate of anti-windup action is defined with the constant T_{AW} which can be explained as a time constant of this action. Åstrom and Hägglund (1995) calculated its value as follows: $T_{AW} = \sqrt{T_I T_D}$.

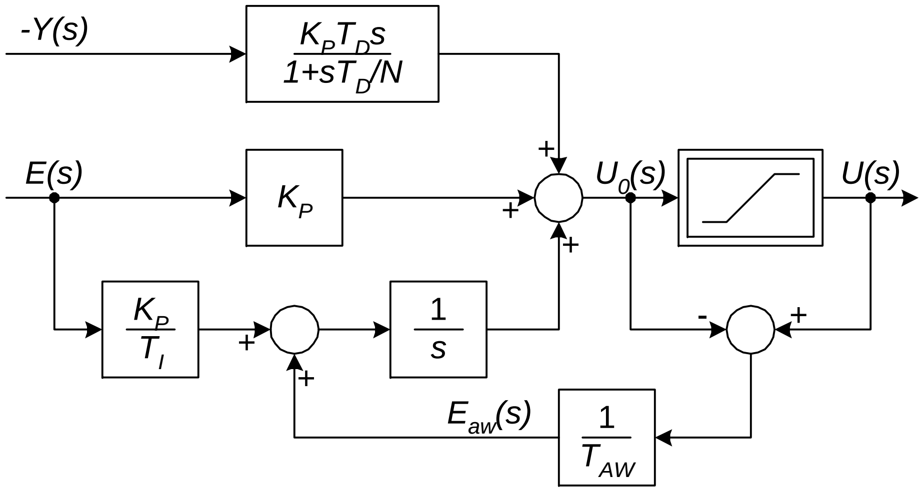


Fig. 5. Structure of the linear feedback anti-windup algorithm.

- Another anti-windup algorithm, suitable for discrete implementation, is the conditional integration algorithm. It allows integration of the error signal *e(t)* in the integrator element provided that some conditions imposed on the signals present in the control system are met. Otherwise, the integration is not permitted.
- Similar deterioration of the control system performance happens when the source of the control signal *u(t)* is changed, for example, when the controller is substituted with another, or when it is switched from manual to automatic mode.
- The switch between two different modes of control system operation is called plant-input substitution.

- A bump in the control signal *u(t)* reflects such plant-input substition unless the switching controllers are properly prepared.
- The effect can be avoided by using bumpless transfer techniques. These techniques calculate the states of the substitution controller before the switch happens, so that bump in the control signal does not occur.
- An important issue in the implementation of discrete control algorithms is the choice of sampling time.
- That choice depends on the control-loop dynamics and should follow the recommendation given in the Shannon's theorem.
- Since there are many signals and elements in the control loop with different dynamic properties, it is not always clear how to choose the sampling time.

- For the discrete PID controller, Isermann (R. Isermann: Digital Control Systems, Springer-Verlag, Berlin, 1989) relates the sampling time T to the settling time $T_{95\%}$ of the process (time required for the response to reach 95% of its final value):

$$\frac{T_{95\%}}{T} \approx 5 \div 15 . \tag{16}$$

- Some rules of thumb have been established for relating sampling time T to the parameters of PI and PID controllers (K. J. Åstrom i B. Wittenmart: Computer Controlled Systems – Theory and Design, Prentice-Hall, New Jersey, 1990):
  - PI controller:

$$\frac{T}{T_I} \approx 0.1 \div 0.3 ; \tag{17}$$

  - PID controller:

$$\frac{TN}{T_D} \approx 0.2 \div 0.6 , \tag{18}$$

where N is divisor constant from equation (6).

- Previous relations should serve as guidelines and, if necessary, should be adjusted for the particular use.
- Additionally, the discrete implementation of the PID controller raises several other issues which have to be addressed:
    - o Effects of finite word length;
    - o Signal quantization effects;
    - o Signal conditioning and prefiltering problems.
- Solutions and trade-offs concerning these issues have been addressed in many textbooks (e.g. Åstrom and Wittenmark, 1990; Isermann, 1989) and in specialized literature (e.g. Åstrom and Steingrimsson, 1991).

3. Tuning methods for PID controllers

- Controller tuning methods provide the controller parameters in the form of formulae or algorithms.
- They ensure that the obtained control system would be stable and would meet given objectives.
- These methods require certain knowledge about the controlled process. This knowledge, which depends on the applied method, usually translates into a transfer function.
- The objectives which should be achieved by the application of the control system are associated with the following control system features:
  - Regulating performance;
  - Tracking performance;
  - Robustness;
  - Noise attenuation.

- Often, the desired objectives put contradictory demands on the values of the controller parameters, so that various trade-offs have to be made. The objectives can be stated in many ways such as through:
  - Specifications within the time domain;
  - Specifications within the frequency domain;
  - Robustness specifications;
  - Other specifications.
- The specifications within the time domain give some values related to the shape of control system signals in the time domain.
- Figure 6. shows a typical output signal *y(t)*, a response to set-point change.
- Specification values within the time domain are marked on it: overshoot σ_m, undershoot σ_u, rise time t_r, time of first maximum t_m, settling time t_ε, and steady state error e_{ss}.

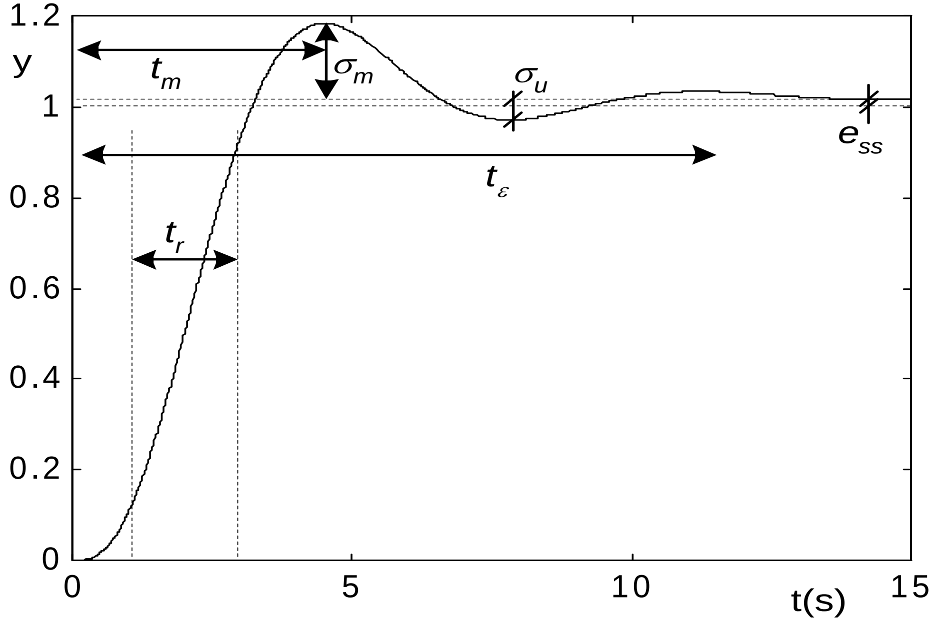


Fig. 6. Specifications in time domain.

- Similar specifications within the time domain are used to describe characteristics of the control system response to load disturbance: peak perturbation σ_{dm} and disturbance settling time $t_{d\varepsilon}$.

- The specifications within the frequency domain define some values related to the frequency characteristics of transfer functions of various elements in the control system.
- Bandwidth of a closed-loop control system with transfer function *G(s)* is the lowest frequency ω_b for which below relation holds:

$$\left|\frac{G(j\omega_b)}{G(0)}\right| = \frac{1}{\sqrt{2}} \tag{19}$$

- The gain margin A_r of the control system, described with the open-loop transfer function $G_O(s)$, is defined as the inverse of the open-loop gain at the phase crossover frequency ω_π:

$$A_r = \frac{1}{\left|G_O(j\omega_\pi)\right|}, \tag{20}$$

where the frequency ω_π is defined as the lowest frequency with

$$\arg\left[G_O(j\omega_\pi)\right] = -\pi\,. \tag{21}$$

- The phase crossover frequency ω_π is also called the ultimate frequency of the control system.
- The phase margin γ of the control system is defined as the phase of the open-loop transfer function $G_O(s)$ at the gain crossover frequency ω_c:

$$\gamma = \arg[G_O(j\omega_c)] + \pi, \tag{22}$$

where the frequency ω_c is defined as the lowest frequency with

$$|G_O(j\omega_c)| = 1. \tag{23}$$

- Maximum sensitivity M_s of the control system, also called modulus margin, is defined as:

$$M_s = \max_\omega \left| \frac{1}{1 + G_O(j\omega)} \right|, \tag{24}$$

- M_S can be interpreted as the inverse of the shortest distance between the critical point C (-1, i0) in the Nyquist plane and the Nyquist curve (Figure 7.).

- The definition (24) for the maximum sensitivity M_s makes it possible to relate M_s to gain and phase margin of the control system:

$$A_r \geq \frac{M_s}{M_s - 1}, \tag{25}$$

$$\gamma \geq 2\arcsin\left(\frac{1}{2M_s}\right). \tag{26}$$

- Robustness specifications define allowed deviation of the process parameters from nominal values.
- The control system should retain designed stability and performance in the range of these deviations.
- The parameter deviations from nominal values can be defined as multiplicative or additive parameter uncertainty characteristics, expressed in the frequency domain.

- Besides, the robustness of the control system can be specified in terms of gain margin A_r, phase margin γ, and as maximum allowed sensitivity M_s of the control system.

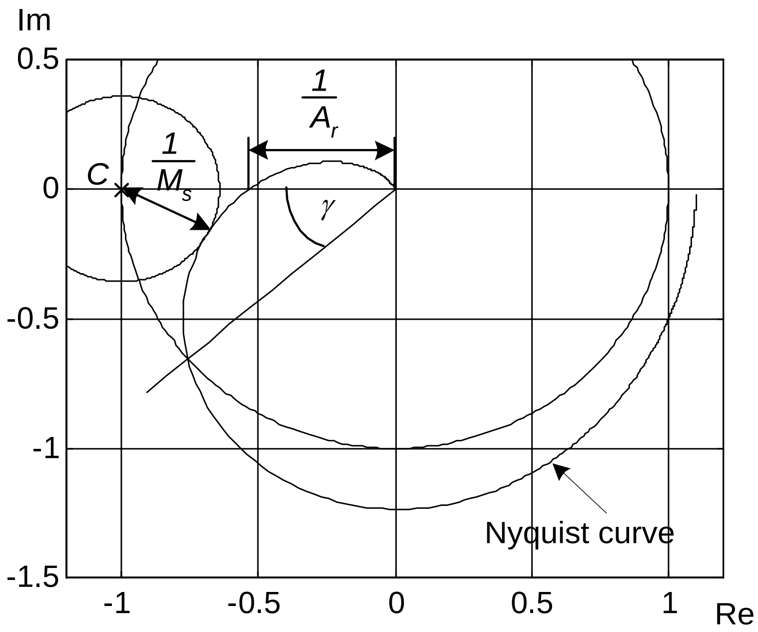


Fig. 7. Specifications in frequency domain.

- Some controller tuning methods include recommendations of a suitable controller structure and its parameters. Tuning methods for the fixed structure of PID controllers are presented here.
- P, PI, and PD controllers are considered as special cases of PID controller.
- The tuning methods for PID controllers can be grouped according to their nature and usage, as follows:
  - Heuristic methods evolved from practical experience in PID controller tuning;
  - Frequency methods employ frequency characteristics of the controlled process to tune PID controller parameters;
  - Analytical methods calculate PID controller parameters from analytical or algebraic relations that define control system by direct calculation;
  - Loop-shaping methods seek to shape the open-loop transfer function of the control system into a desirable form;

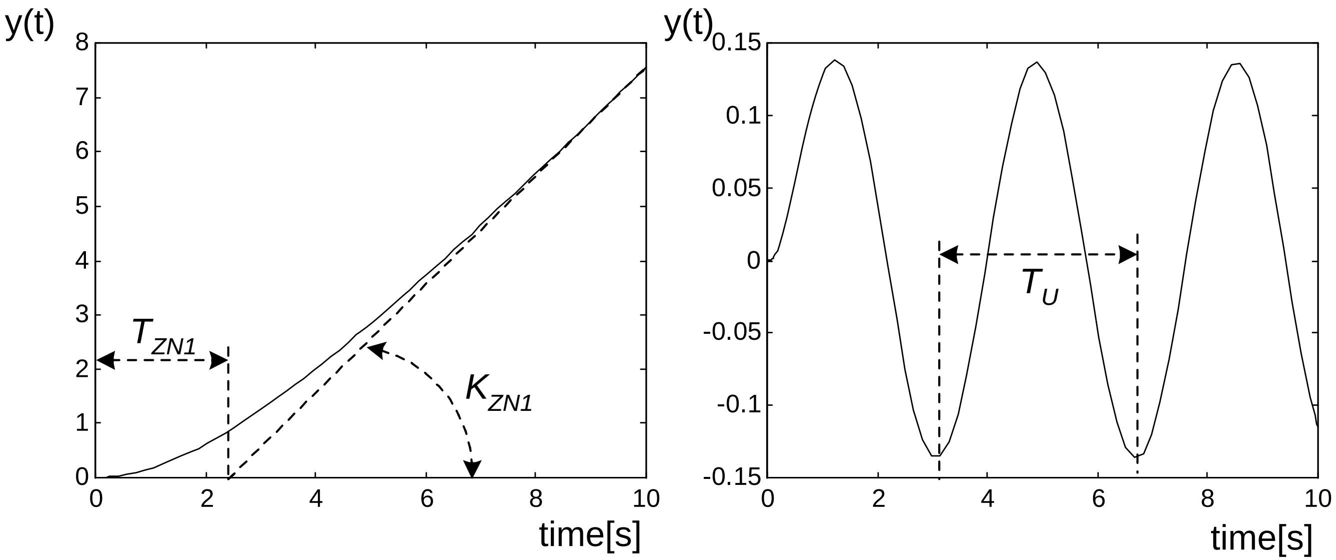
- o Optimization methods obtain PID controller parameters from different optimization algorithms;
- o Methods in which PID controller represents a restriction of possible controller structure (e.g. PID controller tuning in the framework of Internal Model Control);
- o Methods for tuning a PID controller which functions as a part of an advanced control strategy (e.g. usage of PI controller in dead-time compensating controllers).

- The above groups do not sharply distinguish and some methods may belong to more than one group.
- An important criterion in the evaluation of the presented tuning methods is the suitability of a particular method for the on-line usage.
- This especially refers to the possibility to use a particular method for autotuning.

3.1. Ziegler-Nichols tuning rules

- Ziegler and Nichols have introduced a useful methodology for controller tuning.
- It consists of a simple experiment with a controlled process and extracts some of its features.
- Once the experiment is completed, the method provides tables by which it is possible to calculate the controller parameters.
- The tuning tables were developed through numerous experiments which involved different processes.
- The goal of the design was to find a controller which gives the quarter amplitude damping (QAD) ratio of the control systems in response to load disturbance.
- This design specification arises from empirical observations and has been used traditionally, but gives too oscillatory control systems. Ziegler and Nichols considered P, PI, and PID controllers in their work.

- The first experiment consists of measuring apparent dead-time T_{ZN1} and the maximum slope of the response on the process reaction curve (response to step set-point change) K_{ZN1}.
- The measurements are shown in Figure 8.a, and relations for obtaining controller parameters in Table 1.a.



a) Experiment with the process reaction curve; b) Experiment with the process on the stability limit;

Fig. 8. Ziegler-Nichols experiments: output signal of the control system.

- In the second experiment, the process is controlled with a proportional controller.
- The gain of the controller is gradually increased until the control system reaches stable oscillations on the stability limit.
- The value of the controller gain K_U is called the ultimate gain, and the oscillation period T_U is called the ultimate period.
- The two values serve as the basis for the calculation of the controller parameters (see Table 1.b). Figure 8.b shows a typical process output *y(t)* during such an experiment.

Controller	K_P	T_I	T_D		Controller	K_P	T_I	T_D
P	$1/(T_{ZN1}\ K_{ZN1})$	-	-		P	$0.5\ K_U$	-	-
PI	$0.9/(T_{ZN1}\ K_{ZN1})$	$3\ T_{ZN1}$	-		PI	$0.45\ K_U$	$0.85\ T_U$	-
PID	$1.2/(T_{ZN1}\ K_{ZN1})$	$2\ T_{ZN1}$	$T_{ZN1}/2$		PID	$0.6\ K_U$	$0.5\ T_U$	$0.125\ T_U$

a) Experiment with process reaction curve (ZN1); b) Experiment with process on the stability limit (ZN2);

Tab. 1. Ziegler-Nichols relations for calculating controller parameters.

- Takahashi has developed similar relations (Table 2.) to calculate the discrete P, PI, and PID controller parameters.
- These relations are the form of the PID controller, as follows:

$$u(k) = u(k-1) + K_P\left[y(k-1) - y(k) + \frac{T}{T_I}e(k) + \frac{T_D}{T}\left[2y(k-1) - y(k-2) - y(k)\right]\right]. \quad (27)$$

- An additional parameter to the original relations in Table 2 is the sampling time *T*. In relation to other time constants, the sampling time *T* has to be small enough for the relations to produce useful results.
- Many authors assessed the performance of the control systems with controllers tuned according to Ziegler-Nichols (ZN) rules
- The comparison of the two ZN methods shows that the second can be regarded as better, since the first method fails to make it clear how to measure apparent dead time T_{ZN1}.

- Moreover, for many processes the controller gain obtained with the first method is 25% – 40% higher than the gain obtained with the second method, giving more oscillatory response to set-point change.

Controller	K_P	T/T_I	T_D/T
P	$\frac{1}{K_{ZN1}(T_{ZN1}+T)}$	-	-
PI	$\frac{0.9}{K_{ZN1}(T_{ZN1}+T)}-\frac{0.135T}{K_{ZN1}(T_{ZN1}+T/2)^2}$	$\frac{0.27T}{K_P K_{ZN1}(T_{ZN1}+T/2)^2}$	-
PID	$\frac{1.2}{K_{ZN1}(T_{ZN1}+T)}-\frac{0.3T}{K_{ZN1}(T_{ZN1}+T/2)^2}$	$\frac{0.6T}{K_P K_{ZN1}(T_{ZN1}+T/2)^2}$	$\frac{0.5}{K_P K_{ZN1} T}$

a) Experiment with process reaction curve;

Controller	K_P	T/T_I	T_D/T
P	$0.5\ K_U$	-	-
PI	$K_U\left(0.45-0.27\frac{T}{T_U}\right)$	$0.54\frac{K_U T}{K_P T_U}$	-
PID	$0.6K_U\left(1-\frac{T}{T_U}\right)$	$1.2\frac{K_U T}{K_P T_U}$	$\frac{3K_U T}{40K_P T_U}$

b) Experiment with process on the stability limit;

Tab. 2. Relations for controller tuning developed by Takahashi.

3.2. PID tuning based on integral criteria

- Methods based on integral criteria for tuning PID controller involve searching for the minimum of the cost function *I* in the general form:

$$I = \int_0^\infty t^n f[e(t)]dt, \tag{28}$$

where *e(t)* is the error signal.

- Optimum controller parameters and the minimum of the penalty function *I* is found when its partial derivatives, in respect to controller parameters, equal zero. Equations for the calculations of the PID controller parameters are:

$$\frac{\partial I}{\partial K_P} = 0, \quad \frac{\partial I}{\partial T_I} = 0, \quad \frac{\partial I}{\partial T_D} = 0. \tag{29}$$

- Generally, the set of equations (29) can not be solved analytically but numerically. Usually the choice of a particular function f and exponent n leads to formation of the following criteria (28):
  - Integral Error (IE): $f[e(t)]=e(t)$, $n=0$;
  - Integral Absolute Error (IAE): $f[e(t)]=|e(t)|$, $n=0$;
  - Integral Time multiplied Absolute Error (ITAE): $f[e(t)]=|e(t)|$, $n=1$;
  - Integral Squared Error (ISE): $f[e(t)]=e(t)^2$, $n=0$;
  - Integral Squared Time Error (ITSE): $f[e(t)]=e(t)^2$, $n=1$;
  - Integral Time square multiplied Squared Error (IT^2SE): $f[e(t)]=e(t)^2$, $n=2$.

- It is important to note that the error signal $e(t)$, used for optimization, can be a result of set-point change or of load disturbance. It is, therefore, possible to obtain two sets of parameters: one optimized for set-point change and the other for load disturbance.

3.3. Cohen-Coon tuning rules

- The Cohen-Coon tuning method is based on the FODT model (2) with main design specification for quarter amplitude decay (QAD) ratio in response to load disturbance.
- The design objectives (Åstrom and Hägglund, 1995) were to maximize the gain and minimize the steady-state error and QAD for P and PD controllers.
- The parameters of the PI controller were obtained through minimization of the IE criteria and demand for QAD response.
- The parameters for PID controller were calculated with same objectives as for the PI controller. The positioning of the additional closed-loop pole was on the negative real axis. It is placed at the same distance from the origin as the two complex poles of the closed-loop system.

- Relations for controller parameters in Table 3. are given in terms of parameters:

$$\alpha = \frac{K_1 T_{t1}}{T_1}, \tag{30}$$

$$\tau = \frac{T_{t1}}{T_{t1} + T_1}, \tag{31}$$

which are calculated from the parameters of the FODT model (2).

Controller	K_P	T_I	T_D
P	$\frac{1}{\alpha}\left(1+\frac{0.35\tau}{1-\tau}\right)$	-	-
PI	$\frac{0.9}{\alpha}\left(1+\frac{0.92\tau}{1-\tau}\right)$	$\frac{3.3-3\tau}{1-1.2\tau}T_{t1}$	-
PD	$\frac{1.24}{\alpha}\left(1+\frac{0.13\tau}{1-\tau}\right)$	-	$\frac{0.27-0.36\tau}{1-0.87\tau}T_{t1}$
PID	$\frac{1.35}{\alpha}\left(1+\frac{0.18\tau}{1-\tau}\right)$	$\frac{2.5-2\tau}{1-0.39\tau}T_{t1}$	$\frac{0.37-0.37\tau}{1-0.81\tau}T_{t1}$

Tab. 3. Cohen-Coon controller tuning rules.

- Åstrom and Hägglund (1995) observed that the Cohen-Coon tuning method suffers from a too small decay ratio, which results in low damping and high sensitivity of the closed-loop system.

3.4. PID tuning based on gain and phase margin specifications

- PID tuning methods based on gain and phase margin specifications (GPM methods) involve solving definition equations for gain and phase margins, given by (20)-(22). Generally, these equations are non-linear and complicated for solving.
- Therefore, usual design methods based on these specifications are solved numerically or graphically, using Bode diagrams.
- Let's analyze the control system consisting of a PI controller and a FODT process with transfer function (2). When transfer functions of these dynamic elements are put into (20)-(22), *arctan* function appears in relations determining gain and phase crossover frequency.

- It follows from (20) and (21):

$$\tfrac{1}{2}\pi + \arctan \omega_\pi T_I - \arctan \omega_\pi T_1 - \omega_\pi T_{t1} = 0, \qquad (32)$$

$$A_r K_P K_1 = \omega_\pi T_I \sqrt{\frac{\omega_\pi^2 T_1^2 + 1}{\omega_\pi^2 T_I^2 + 1}}, \qquad (33)$$

- It follows from (23) and (21):

$$K_P K_1 = \omega_c T_I \sqrt{\frac{\omega_c^2 T_1^2 + 1}{\omega_c^2 T_I^2 + 1}}, \qquad (34)$$

$$\gamma = \tfrac{1}{2}\pi + \arctan \omega_c T_I - \arctan \omega_c T_1 - \omega_c T_{t1}. \qquad (35)$$

- In order to simplify the procedure of solving these non-linear equations, the following approximation was introduced:

$$\arctan(x) \approx \begin{cases} \dfrac{\pi x}{4} & ;|x| \le 1 \\ \dfrac{\pi}{2} - \dfrac{\pi}{4x} & ;|x| > 1 \end{cases}. \qquad (36)$$

- The approximation and solving equations (32)-(35) for PI controller parameters gives:

$$K_P = \frac{\omega_\pi T_1}{A_r K_1}, \tag{37}$$

$$T_I = \left(2\omega_\pi - \frac{4\omega_\pi^{\ 2} T_{t1}}{\pi} + \frac{1}{T_1} \right)^{-1}, \tag{38}$$

where ω_π is calculated through:

$$\omega_\pi = \frac{A_r \gamma + \frac{1}{2}\pi A_r (A_r - 1)}{(A_r^{\ 2} - 1) T_{t1}}. \tag{39}$$

3.5. Approximate pole placement method: Dominant pole design

- Dominant-Pole Design (DPD) methods find controller parameters which place the dominant poles of the closed-loop system in specified locations. In other words, those methods can be viewed as a translation of the problem of finding controller parameters into the problem of placing dominant poles in desired locationsat the complex plain (*s* or *z*).
- The number of dominant poles to be placed depends on the number of free parameters, that is, on the number of controller parameters.
- A PI controller allows placement of two dominant poles, and a PID controller of three dominant poles. For these controllers, locations of the closed-loop dominant poles are parameterized with (Åstrom and Hägglund, 1995):
  - o PI controller:

$$p_{1,2} = \omega_n(-\zeta \pm j\sqrt{1-\zeta^2}),\ 0<\zeta<1; \qquad (40)$$

  - o An additional pole location for PID controller:

$$p_3 = -k_0\omega_n\,. \qquad (41)$$

- Pole locations (40) and (41) in s-plane are depicted in Figure 9., where angle α is determined through $\alpha = acos\,(\zeta)$.

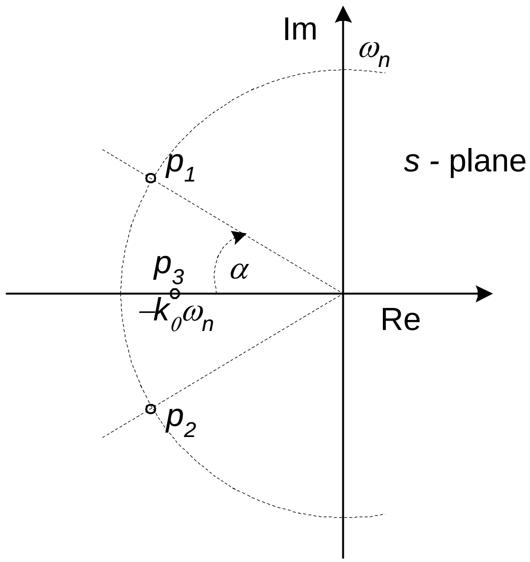


Fig. 9. Locations of dominant poles of closed-loop control system using PI and PID controllers.

- Calculation of the closed-loop poles involves solving characteristic equation of the control system:

$$1 + G_R(s)G_P(s) = 0. \tag{42}$$

- If the required pole location is p_1, and the controller $G_R(s)$ is a PI controller, the characteristic equation of the control system is:

$$1 + \left(K_P \frac{1 + T_I p_1}{T_I p_1} \right) G_P(p_1) = 0, \tag{43}$$

and it should be solved for the controller parameters K_P and T_I.

- To simplify calculation of the characteristic equation transfer function $G_P(s)$, modeling the process, is parameterized with:

$$G_P(s)\big|_{s=\omega_n e^{j(\pi-\alpha)}} = G_P(\omega_n e^{j(\pi-\alpha)}) = a(\omega_n, \alpha) e^{j\phi(\omega_n, \alpha)} \tag{44}$$

- Parameter functions $a(\omega_n, \alpha)$ and $\phi(\omega_n, \alpha)$ can described as 'frequency characteristics' of the process on the ray with angle α in s-plane, where a can be considered as 'gain characteristic' and ϕ can be considered as 'phase characteristic'.

- By putting parameterizations (40) and (44) in the characteristic equation (43), and solving it, following relations for PI controller parameters are obtained:

$$K_P = -\frac{\sin\left(\phi(\omega_n,\alpha)+\alpha\right)}{a(\omega_n,\alpha)\sin(\alpha)}, \tag{45}$$

$$T_I = \frac{\sin\left(\phi(\omega_n,\alpha)+\alpha\right)}{\omega_n \sin(\phi(\omega_n,\alpha))}. \tag{46}$$

- Dominant pole locations of the control system are set to p_1 and p_2 with suitable choice of ω_n and ζ, and through calculation of the PI controller parameters according to (45) and (46). Similar relations can be derived for parameters of the PID controller.

3.6. Magnitude optimum and symmetric optimum tuning methods

- Magnitude optimum (MO) and symmetric optimum (SO) are two loop-shaping tuning methods extensively employed by the German company Siemens. The first step in the application of these methods is to determine appropriate transfer function which models the process. Once the transfer function is determined, the controller is able to shape the open-loop transfer function in a desired.
- MO tuning method was devised with the objective to obtain a control system with a frequency characteristic as close to unity and as flat as possible for the maximum bandwidth. Its mathematical expression states the requirements posed on the closed-loop transfer function $G_C(s)$:

$$G_C(0)=1, \tag{47}$$

$$\lim_{\omega=0} \frac{d^n\left(G_C(j\omega)\right)}{d\omega^n} = 0, \tag{48}$$

for as many n as possible.

- Let it be desired open-loop transfer function is:

$$G_{O1}(s) = \frac{\omega_n^{\ 2}}{s(s+2\varsigma\omega_n)}, \tag{49}$$

where ζ is the damping of the closed-loop system and ω_n determines the closed-loop dynamics, that is, the speed of response.

- For example, the PI controller is employed when it is possible to approximate the model of the process with the transfer function:

$$G_P(s) = \frac{K}{(1+T_1 s)(1+T_2 s)}, \tag{50}$$

with $T_2 < T_1$,.

- By analyzing (47)-(50), and by setting ζ=0.707, PI controller parameters are calculated (Åstrom and Hägglund, 1995)

$$K_P = \frac{T_I}{2\ K\ T_2}, \tag{51}$$

$$T_I = T_1\ , \tag{52}$$

with ω_n=0.707/T_2.

- The dominant pole is cancelled by the PI controller zero, and the closed-loop dynamics are determined the smaller time constant T_2 of the process.
- MO design method optimizes the closed-loop transfer function $G_C(s)$ between the reference and the output signal. It often cancels the process poles by the controller zeros, which can lead to poor performance of the control system in response to load disturbance.
- The objective of the SO method, which was originally proposed by Kessler (1958), is to obtain an open-loop transfer function of the below formula:

$$G_{O2}(s) = \frac{a\omega_c^{\ 2}}{s^2}\frac{(s+\frac{\omega_c}{a})}{(s+a\omega_c)}, \tag{53}$$

where ω_c is the gain crossover frequency and a is related to the phase margin of the control system through:

$$\gamma = 2\ \text{atan}\left(\frac{a-1}{a+1}\right), \tag{54}$$

or conversely through:

$$a = \frac{1+\sin\gamma}{\cos\gamma}. \tag{55}$$

- The method maximizes the phase margin of the control system and leads to symmetrical phase and amplitude characteristics, as can be observed in Figure 10.
- The second multiplicand in (53) has the transfer function of a phase-lead network, which provides required phase uplifting at the frequency ω_c.
- Figure 10. shows the amplitude and gain characteristics of the control system with the open-loop transfer function equal to (53). In the example, the parameter a was set to 4.

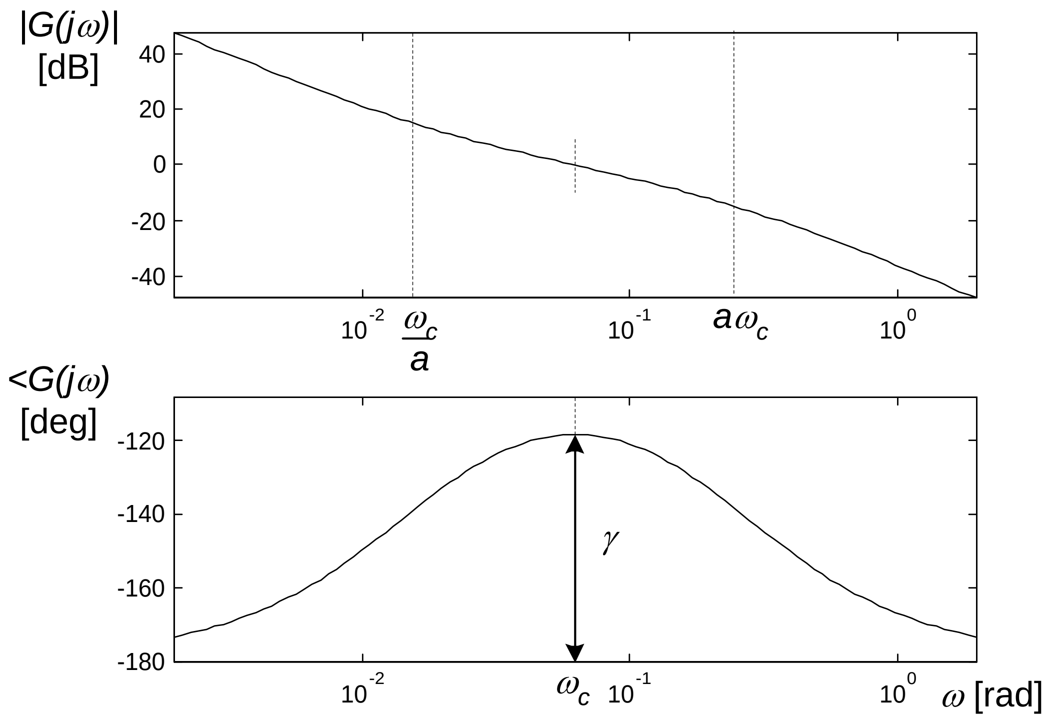


Fig. 10. Gain and phase characteristics of a control system tuned according to symmetrical optimum.

- For example, if the process can be modeled with the transfer function G_2 (3):

$$G_2(s) = \frac{K_2}{s(1+T_2 s)}, \tag{56}$$

it is suitable for the application of the SO tuning method.

- The procedure leads to a PI controller with the following settings (Perić, 1979):

$$K_P = \frac{1}{a\,K_2\,T_2}, \tag{57}$$

$$T_I = a^2 T_2, \tag{58}$$

and $\omega_c = \dfrac{1}{a\ T_2}$.

- The common choice for the parameter a is 2, which gives the phase margin of the control system $\gamma \approx 37^0$.

- The SO method is designed to give a good response to load disturbance, but the response of the control system to set-point change has large overshoot. The overshoot is commonly reduced through the usage of a two-degree-of-freedom controller or with a prefilter.
- The MO and SO tuning methods are widely used in the cascade control systems, especially to control motor drives (Perić, 1979,1989; Deur, 1999).

3.7. PID tuning in the framework of Internal Model Control

- Internal Model Control (IMC), thoroughly described by Morari and Zafiriou (1989), is a general design procedure for obtaining controllers that ‘optimally’ meet requirements for stability, performance, and robustness of the control system.
- The concept of IMC is based on the simulation of the process model $G_M(s)$ within the control structure.
- Figure 11. shows the arrangement of IMC.
- If the model of the process $G_M(s)$ perfectly matches the process $G_P(s)$, and load disturbance is not present, the output of the model cancels the output of the process annulling thus the feedback signal.
- In such case, the process is controlled in an open loop. The feedback signal and hence feedback control, exist only if there is the model mismatch or load disturbance $Z(s)$.

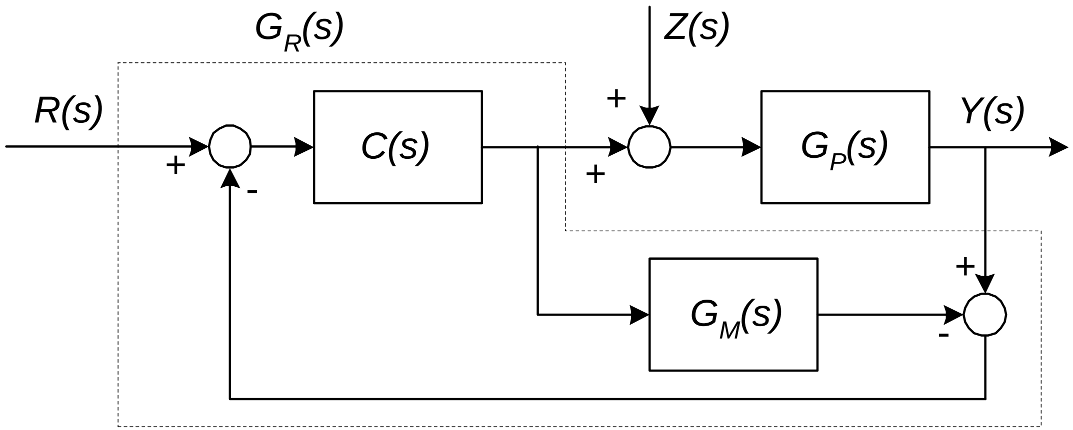


Fig. 11. Structure of Internal Model Control.

- IMC design is made through following steps. The first in IMC design is to factor the transfer function modeling the process:

$$G_M(s) = G_M^{+}(s)\, G_M^{-}(s), \tag{59}$$

where $G_M^{-}(s)$ contains only the left half plane poles and zeros, and $G_M^{+}(s)$ contains all time delays and the right half plane zeros.

- After that, the controller $C(s)$ is defined with:

$$C(s) = (G_M^-(s))^{-1} G_F(s), \tag{60}$$

where $G_F(s)$ is a filter which guarantees that the controller $C(s)$ is realizable.

- $G_F(s)$ also obtains the desired robustness and defines the closed-loop dynamics. The usual form of the filter is:

$$G_F(s) = \frac{1}{(1+T_F s)^n}, \tag{61}$$

- The IMC design procedure can be used to design conventional feedback controllers.
- Figure 11. shows the relation between a conventional feedback controller $G_R(s)$ and IMC controller $C(s)$, which may be expressed with the below formula:

$$G_R(s) = \frac{C(s)}{1 - C(s) G_M(s)}. \tag{62}$$

or inversely:

$$C(s) = \frac{G_R(s)}{1 + G_R(s) G_M(s)}, \tag{63}$$

- For the particular choice of model $G_M(s)$ and filter $G_F(s)$, IMC design procedure leads to PID controller (Morari and Zafiriou, 1989).
- Performance and robustness trade-off of the obtained control system is handled through the value of the adjustable parameter T_F, which determines the dominant time constant of the closed-loop system.
- For example, a low value of the parameter T_F produces fast response of the control system, but also results in low robustness margin.
- The FODT model (2), can be used within the frame of IMC, but the part of the transfer function modeling dead time e^{-sTt} has to be replaced with Pade approximations.
- Furthermore, the exponent n in the denominator of the filter transfer function (61) is set to 1.

- Pade approximation of the zero order is:

$$e^{-sTt} \approx 1, \qquad (64)$$

and leads to an IMC PI controller with the following parameters:

$$K_P = \frac{T_1}{K_1 T_F}, \qquad (65)$$

$$T_I = T_1, \qquad (66)$$

and the recommended value for the filter time constant $T_F > 1.7\ T_{t1}$.

- The first order Pade approximation:

$$e^{-sT_{t1}} \approx \frac{1 - sT_{t1}/2}{1 + sT_{t1}/2}, \qquad (67)$$

in the FODT model and IMC design lead to a PID controller with parameters:

$$K_P = \frac{2T_1 + T_{t1}}{K_1(2T_F + T_{t1})}, \qquad (68)$$

$$T_I = T_1 + \frac{T_{t1}}{2}, \qquad (69)$$

$$T_D = \frac{T_1 T_{t1}}{2T_1 + T_{t1}}, \qquad (70)$$

and the recommended value for the filter time constant $T_F > 0.8\ T_{t1}$.

- It can be concluded that the main advantages of IMC design are:
  - Model uncertainty is explicitly considered;
  - Trade-off between performance and robustness of the control system is clearly defined.
- The principal drawback of the method is that the process poles are cancelled with controller zeros according to (60), which results in sluggish response to load disturbance.
- IMC tuning rules are expressed in terms of process model parameters and can be applied after the identification of the process model. Such models can be obtained as a part of an autotuning procedure.

4. Usage of PI controller in dead-time compensating controllers

- One area of process control in which PID controllers fail to produce satisfactory results is the control of processes with large time delays.
- Time delay is considered large when its value exceeds the dominant time constant of the process.
- This type of dynamic behavior, termed time delay or dead time, is present in processes involving transport of materials such as rolling mills in metal industry and is a common result of composition analysis in chemical industry.
- The assertion that PID control is inadequate for the control of processes with large dead time is based on two arguments:
  - o the derivative action of the PID controller, needed for prediction, amplifies noise;
  - o the open-loop gain has to be small rendering the performance of the control system poor.

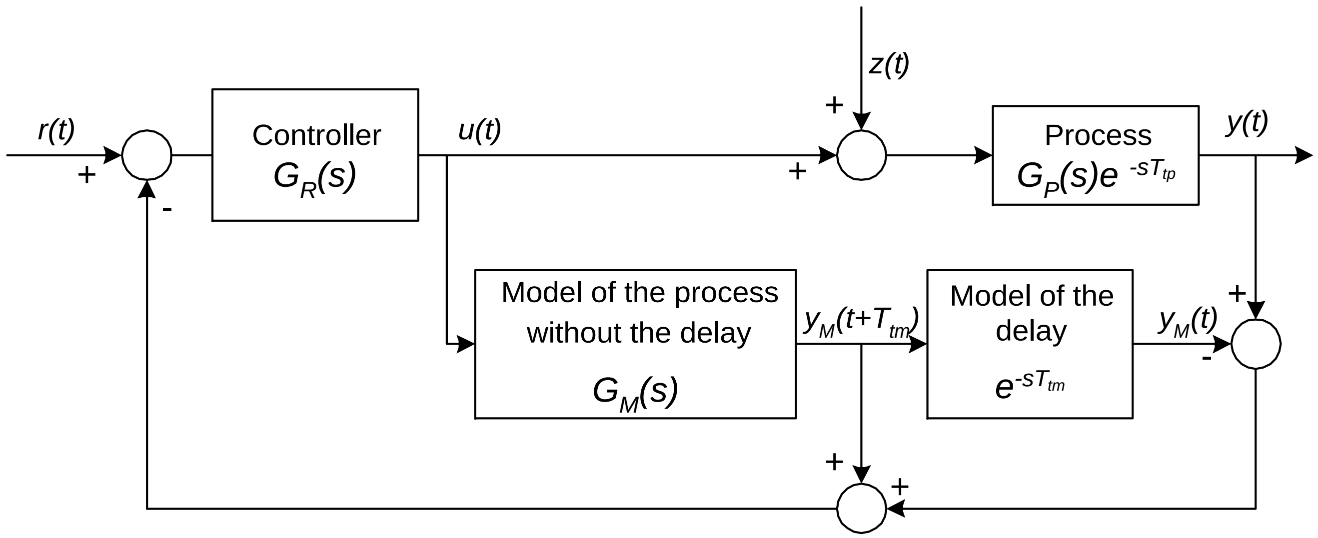


Fig. 12. Structure of the Smith predictor.

- Another approach to control of processes with large time delays is to incorporate the model of the system as in Internal Model Control
- The **Smith predictor**, shown in Figure 12, applies that approach.

- The model of the process is divided in two parts: one for modeling the dynamic behavior $G_M(s)$ of the process and the other for modeling time delay $e^{-sT_{tm}}$.
- The transfer function of the control system with the Smith predictor is:

$$G_X(s) = \frac{Y(s)}{R(s)} = \frac{G_R(s)G_P(s)e^{-sT_{tp}}}{1 + G_R(s)G_P(s) + G_R(s)G_P(s)e^{-sT_{tp}} - G_R(s)G_M(s)e^{-sT_{tm}}}. \qquad (71)$$

- The denominator of the transfer function (71) which is the characteristic equation of the closed-loop control system, plainly shows that when the modeling is exact ($G_M(s)=G_P(s)$) the two last terms are cancelled.
- In such case, the closed-loop transfer function becomes:

$$G_X(s) = \frac{Y(s)}{R(s)} = \frac{G_R(s)G_P(s)}{1 + G_R(s)G_P(s)} e^{-sT_{tp}}, \qquad (72)$$

- In that case the controller $G_R(s)$ can be designed as if the process did not contain dead time.

- In other words, the controller $G_R(s)$ can be designed just for the part of the process modeled by the transfer function $G_M(s)$.
- The cancellation of dead-time influence on the dynamic behavior of the control system is characteristic of dead-time compensating controllers.
- The main drawback of the Smith predictor is that the performance and the stability of the control system are very sensitive to inaccurate modeling of the process, especially to the inaccurate modeling of dead time.
- One important advantage of PID control over other control strategies, including the Smith predictor, is that operators are familiar with the tuning procedures for PID controllers.
- These involve finding of only three parameters.
- In comparison, the tuning of the Smith predictor involves identification of a suitable process model, and then tuning of the controller.

- The Smith predictor, which is based on the FODT model and on the PI controller, has five parameters and is very complicated to tune and operate.
- In order to simplify the tuning procedure of the Smith predictor, it has been proposed a restriction of the choice of the PI controller and FODT model parameters.
- This type of the Smith predictor is called the **predictive PI (PPI) controller**.
- Figure 13. shows the structure of the PPI controller.

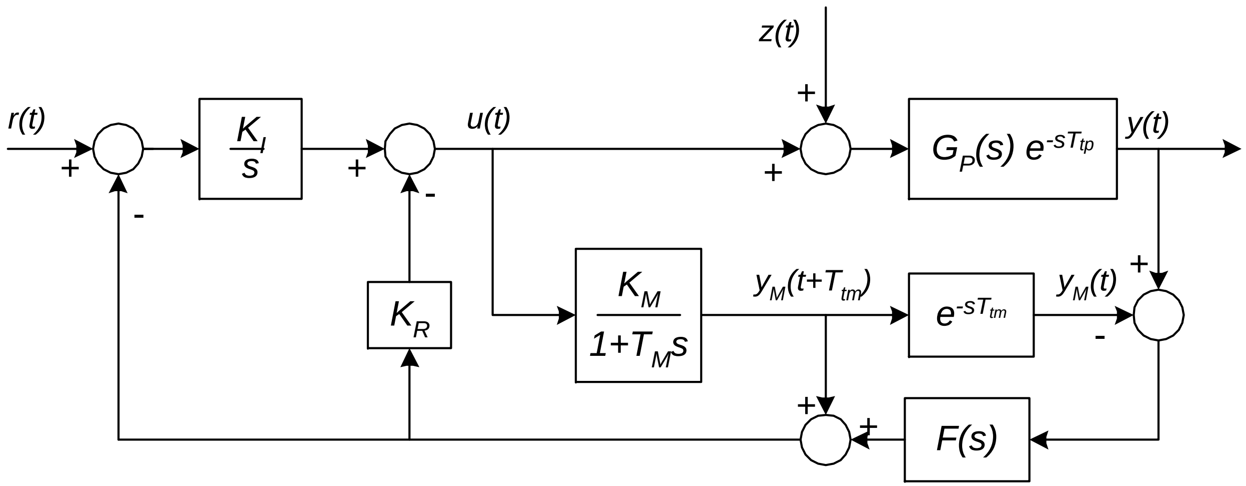


Fig. 13. Structure of the FPPI controller.

- In the PPI controller, the parameters of the PI controller are related to the FODT model parameters as follows:

$$K_R = \frac{\kappa}{K_M}, \tag{73}$$

$$T_I = \frac{1}{\tau} T_M, \tag{74}$$

where κ and τ are calculated from the desired performance of the closed-loop control system.

- Based on these parameters, the characteristic equation of the PPI is:

$$s^2 + \frac{1+\kappa}{T_M} s + \frac{\kappa\tau}{T_M{}^2} = 0. \tag{75}$$

- The PPI controller solves the problem of operational complexity of the Smith predictor since it decreases the number of controller parameters, but it does not solve the problem of high sensitivity of the Smith predictor due to the inaccurate modeling of the process.

- In order to increase the robustness of the PPI controller a filter in the PPI structure is introduced. Figure 13. shows such **filtered PPI (FPPI) controller**.
- In order to preserve the simple structure of the PPI filter, *F(s)* is chosen to be the first-order lag with static gain equal to one:

$$F(s) = \frac{1}{1 + T_{F_S} s} . \tag{76}$$

- It is recommended to choose:

$$T_{FS} = \frac{T_{tm}}{2} . \tag{77}$$

- The described PID controllers are basic components of many control systems operating in industry.
- It is worthwhile to automate some of the described tuning procedures (autotuning PID controller).