APROL Solutions
Advanced process control
SEM 890.3

Perfection in Automation
www.br-automation.com

APROL Solutions
SEM 890.3

Version: **V1.00 / Date: April 04, 2014**
Model number: **-**

Requirements:	
Training modules	Basic Training: Solutions Additional useful training modules: TM800, TM820, TM830, TM811, TM812, TM813,
Software	APROL Automation Runtime SuSE LINUX
Hardware	1 control computer, 2 controller

Table of contents

1 INTRODUCTION

You have chosen to take part in the APROL Solutions training module offered by B&R in order to become familiar with advanced process control methods.

Some of you may just want to learn about what types of configuration and expansion possibilities actually exist; others might actually be working with solutions projects and setting up your first installations after this seminar is over. We will do our very best to provide each and every one of you with as much information as possible to meet your requirements.

Nevertheless, the technical expertise of your seminar leader is not the only factor contributing to your personal success during this seminar. Success depends on cooperation and interaction between course members and the seminar leader, as well as the attitude of individual participants towards teamwork and the course in general.

Figure 1,APROL Solutions --> A perfect world

1.1 Learning objectives

The APROL Solutions seminar will provide you with the knowledge to identify control relevant optimization options of plants or rather processes and recognize their benefits.

Advanced Process Control (APC) is general term for controllers that go beyond conventional PID control or that improve them using additional function blocks. This seminar introduces the APC function blocks from the Process Automation Library (PAL). The function blocks have already been placed in CFC plans and connected. The goal here is to explore control methods that can be used to improve control quality. The basics that have been learned will then be reinforced by completing tasks for practical applications.

Different dynamical models will be used during the training to illustrate the control methods.

Figure 2,Overview of APROL solutions

The following control technology basics will be covered in the examples

Basic control technology terminology;

Dynamical systems (description, configuration of controlled systems);

The PID controller;

Adjustment of PID controller parameters, assessment of the control quality;

PID autotuning process;

Control performance monitoring (CPM);

Ratio control;

Linear interpolation of grid points (lookup table);

Split range control;

Override control;

Correction terms, dynamic disturbance variable disconnection;

Cascade control;

Gain scheduling;

Control of dead time dominant systems (Smith predictor);

Model predictive control (MPC);

Following exercises are gone through to stabilize theoretical information

- Second-order delay element with dead time.
  - Determination of the controlled system parameters
  - Manual tuning of a PID controller
    - Oscillation test according to the Ziegler/Nichols method
    - The Lambda tuning method
    - Trial on error – tuning
    - Short explanation of other tuning procedures (e.g. loop shaping, optimization procedures, status controller)
  - Automatic tuning of a PID controller
    - Oscillation methods
    - Step methods
- Positioning control for an actuating drive with friction (IT1 optional with non-linear friction model)
  - The deadband of a PID controller
  - The anti-windup function of a PID controller
  - Control performance monitoring
  - Identification of system dynamics (systems without inherent regulation)
- Mixture of two material flows
  - Ratio control
    - Mixture temperature control
    - Feed-forward through the ratio function block
    - Feed-forward with mixture temperature control
    - Enhanced feed-forward with mixture temperature control
  - Lookup table
    - Mixture temperature control, generation of external temperature-dependent setpoint
- Mixture of three material flows
  - Split range mixture temperature control (bipolar adjusting range)
- Heat exchangers
  - Override control with "external reset feedback" input
  - Dynamic disturbance variable compensation
    - The lead/lag link
  - Cascade control
- Horizontal cylindrical tank
  - Level control through PID gain scheduling
- Continuous stirred-tank reactor enclosed in a cooling jacket
  - PI control
  - Smith predictor
  - MPC single-input control
- Shell heavy oil fractionator
  - MPC multiple-input control

2 OVERVIEW OF ADVANCED PROCESS CONTROL (APC)

During the 3-day Advanced Process Control (APC) Solutions Training, basic and advanced knowledge of control technology will be shared with you in order to improve the control quality of standard control loops. In addition to discussions of fundamental control technology principles, the training will include ready-made examples in the distributed control system (DCS) APROL. Knowledge about project planning with APROL is not required in order to work with the examples. The tasks are already prepared to the point that they can be completed through operation of the graphic user interface.

Important:

The event is primarily concerned with control methods and their operation in APROL from the operator's view. Project implementation in APROL will not be discussed.

2.1 Why APC?

Improvements in process efficiency can very often be associated with an improvement in control quality. The degree of automation in processes is increasing. As a consequence, more process variables are available to users, which, in combination with APC methods, can lead to improvements in control quality. In the following figure, for example, the moisture content of the final product is controlled by a drying process. The red line shows the maximum moisture content, up to which point the product can still be sold. Basic automation (left) manages the task with relatively high variations in the control variable (moisture content). For this reason, the set point of the moisture content must be kept very low. Using APC methods, the variability of the control variable can be lowered. Subsequently, the moisture content setpoint can be increased. The decreased drying effort leads to lower energy consumption, which is converted into 1/1 profit.

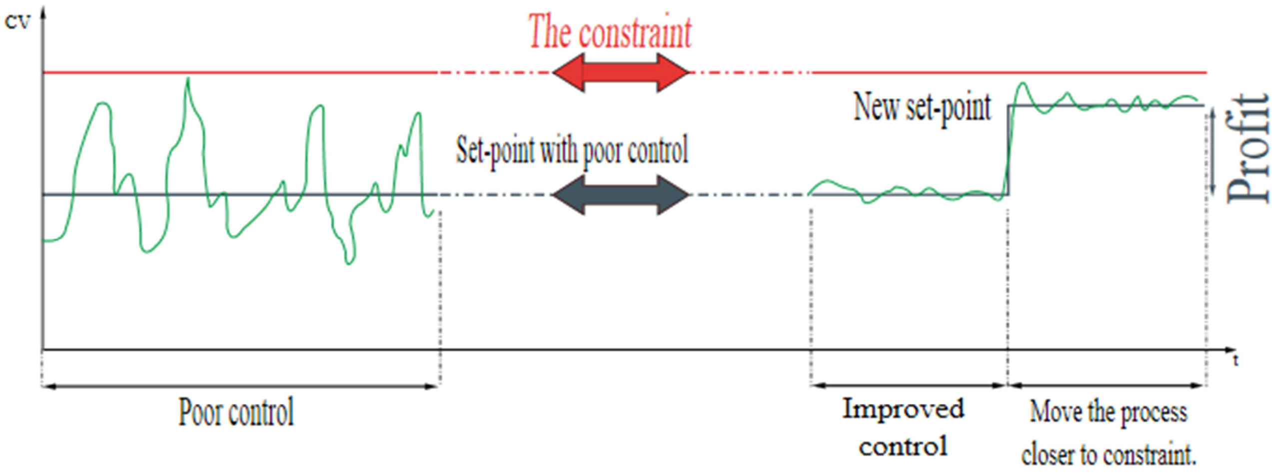


Figure 3, Advantage of APC

By reducing the variability of control variables, the product quality, for example, can be improved. If varying base materials are used, quality variations can be compensated for more effectively. Sensitivity to disturbances can typically be reduced through good control. Unintentional downtime of the controlled components can be reduced with APC and, in contrast, the makespan can be increased. With reduced failures, the burden on system operators in control rooms is reduced. In addition, improved control quality can have a positive effect on emissions, e.g. operation of incineration plants with less fuel. With a forward-looking control loop working method, a component can be operated in a way that is safer and which protects the materials.

2.1.1 Monitoring and improvement of control quality

All function blocks presented here are completely integrated into APROL. For this reason, it is mainly engineering costs that are incurred when commissioning additional function blocks. Costs for additional hardware, if new parts are needed, are kept very low.

If existing control loops are to be improved, the following may be carried out:

1. Identification of control loops which have a need for action. If, for example, several operator interventions are required, the manipulated variable is very often saturated and impairment to product quality occurs. Here, for example, the Control Performance Monitoring tool can be referred to.
2. Control loops that show potential for improvement or problems should be tested in relation to actuator and test signal sensor problems. Is there, for example, valve friction that is obstructing the current control loop? Are sensors impaired or damaged by residues or can any occurring signal disturbances be prevented? If an overriding control structure is incorporated, the basic automation that is closest to the system should be set first and then the overriding control structure.
3. If problems have been discovered and corrected, the control quality should be checked once more. Have improvements occurred?
4. All control loops that still exhibit poor control quality have the potential for improvement in the control quality using APC.
5. In the next step, we will differentiate between single-input control and multiple-input control.
6. Multiple input controls can, for example, be carried out with an MPC. If systems have non-linear parts, a subordinate control loop with gain scheduling control can be created, for example.
7. In single-input systems, linear and non-linear behaviors can be distinguished. In non-linear behavior, the use of gain scheduling can be tested. In linear behaviors it is possible to test whether disturbance signals are measurable or whether long periods of dead time occur and, for example, whether Smith predictor controls can be used.

Control qualities can impact on the efficiency of the system operating mode by increasing the flow rate, reducing the product changeover time during production or reduced energy consumption.

2.2 Agenda

Presentation
Exercise

	Day1	Day2	Day3
Block1	Overview of event and agenda Dynamic systems (description) Controlled systems Simulation of dynamical systems Parameter identification PT2 Features of dynamical systems Discrete-time observation	Split range control; Mixture (bipolar adjusting range) Override control Heat exchangers	MPC advantages MPC applications MPC operating principle Manipulated variable calculation SISO
Block2	The control loop, Stability, The basics of PID control technology Adjusting PID controllers Ziegler/Nichols tuning (stability margin) Lambda tuning Other tuning methods	Dynamic disturbance variable compensation, Lead/lag link Heat exchangers Cascade control; Quiz, Heat exchangers	Configuring the MPC Project development,
Block3	PID autotuning, Oscillation/step test Adjustment of PID parameters, PID expansions, Deadband (circuits with and without inherent regulation) Anti-windup	Gain scheduling, Level control for the horizontal cylindrical tank, Smith predictor, Temperature control for stirred-tank reactor	MPC multiple-input control
Block4	Control performance monitoring, CPM position control with static friction, Control/Feed-forward/Prediction Ratio control Temperature control, Feed-forward ratio control, Lookup table (1D, 2D) Feed-forward (lookup table)	MPC introduction, What is MPC? How does MPC work? Advantages of MPC, Target customers, Product status MPC single-input control	MPC application information / system topology, Recap of APC training, Summary

Table 1, Agenda

3 DYNAMIC SYSTEMS

A system is called a dynamic system if the output variables depend not only on the current input value of the input variables, but also on their historic values.

Figure 4, Dynamical system

The most important parameters of a dynamical system can change over time. For this reason, dynamic systems can be illustrated as functions of time t. In a dynamical system, a differentiation is made between the input variables that cause temporal changes within the system and the output variables.

The control technology uses models which describe the behavior of dynamical systems (control systems) through time-variable parameters (signals) and signal conversions. The models can be described in the form of differential equations. This document only discusses linear differential equations. The time-variable parameters and their derivatives depend on the physical parameters of the modeled process. The type of time-variable parameters (e.g. energy, mass or capital) does not play a role in terms of the system concept.

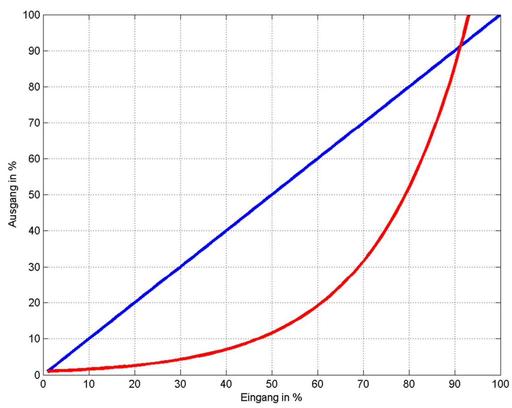
<table>
<tr>
<td>The stationary observation of systems assumes that the input variable does not change and the output variable is in a steady state ($t \to \infty$).
The stationary output is a function of the stationary input variable.</td>
<td>In the dynamic observation of systems, the output can be formulated as a function of the input signal and the time. The input signal may change over time.
The time-dependent output is a function of the time-dependent input signal.</td>
</tr>
<tr>
<td>

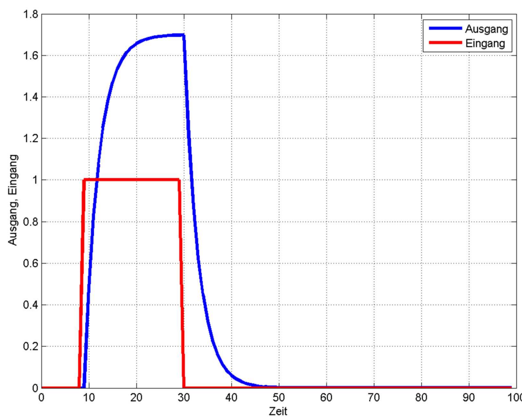
Figure 5, Stationary observation of dynamical systems</td>
<td>

Figure 6, Non-stationary observation of dynamical systems</td>
</tr>
</table>

Table 2, Stationary and non-stationary observation of dynamical systems

3.1 Systems, plants and processes

Plants or processes can be described through dynamical systems. Dynamical systems are also colloquially called systems, or controlled systems. In terms of control technology, dynamical systems (controlled systems) have different inputs and outputs.

- Input variables

- o The process can be actively influenced by manipulated variables.
- o Disturbance variables influence the dynamical system, but cannot be influenced. In addition, it is possible to differentiate between measurable and unmeasurable disturbance variables. Disturbance variables whose future course is already known (e.g. production plan) also exist.

- Output variables
  - o Control variables are process variables which should be specifically influenced by manipulated variables.

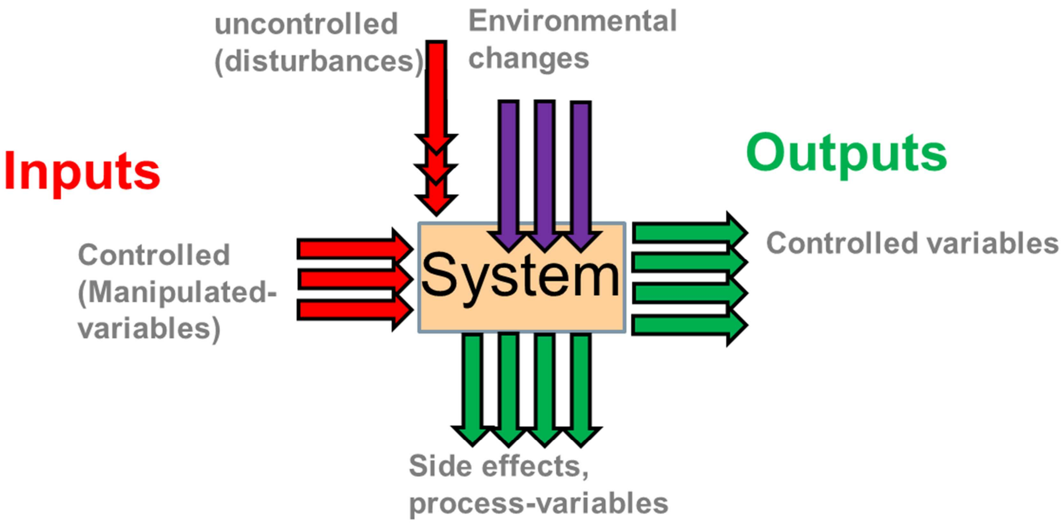


Figure 7, Dynamical system in detail

3.2 Description of dynamical systems

Dynamical systems can be described through various methods.

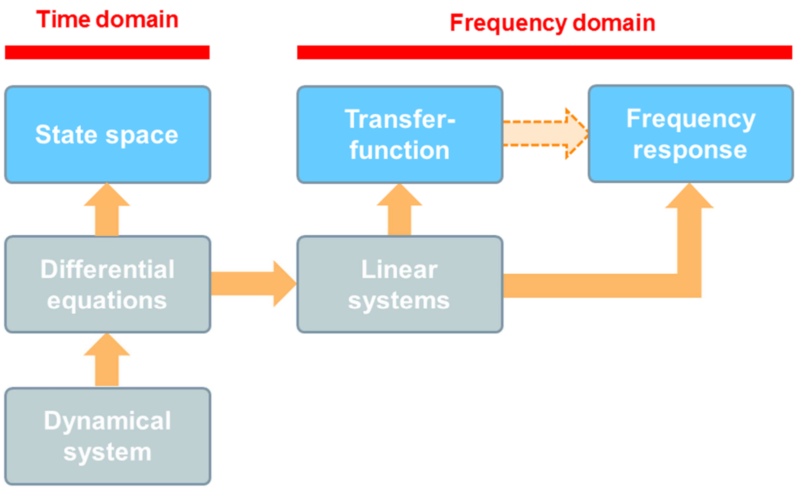


Figure 8, Description methods of dynamical systems

> **Important:**
>
> The following mathematical methods are only intended to provide a general overview of the description of dynamical systems. If you are interested in learning more, please refer to the relevant technical literature.

3.2.1 Differential equations

The conventional form is the differential equation, which can produce a very good description of the system behavior, often without restricting the behavior of dynamical systems. In differential equations, a basic distinction can be made between equations with lumped or distributed parameters.

The following figure illustrates a simple correlation between room temperature, external temperature and heat energy input in a house is made through a differential equation.

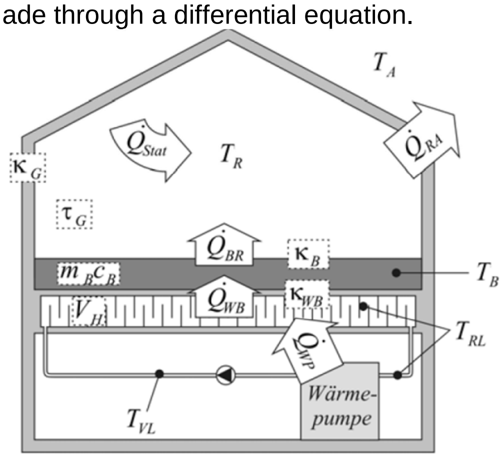


Figure 9, House model principle

As can be seen in the figure, the modeling can become complex very quickly. How many energy stores are taken into consideration? What forms of heat transfer are included in the model (solar radiation, heating conduction in the bricks, furnishing in the room)? Regardless of the level of abstraction, the differential equation should provide a correlation between room temperature and external temperature and heat output.

3.2.2 State space representation

Using state space representation, the differential equation can be mapped in an equivalent first-order differential equation system (linear hypothesis, lumped parameters). This form is very well suited to computer-aided processing. In state space representation, visible "states" of the dynamical system do not emerge at first glance. Based on the description of the house, in addition to the input and output variables, e.g. the return temperature of the floor heating (TRL) or the temperature of the concrete core of the covering emerge as "states" of the dynamical system.

$$\begin{bmatrix}\dot{T}_{RL}\\ \dot{T}_B\\ \dot{T}_R\end{bmatrix} = \begin{bmatrix} -\dfrac{k_{HB}}{\rho_w c_w V_w} & \dfrac{k_{HB}}{\rho_w c_w V_w} & 0 \\ \dfrac{k_{HB}}{m_B c_B} & -\dfrac{(k_{HB}+k_B)}{m_B c_B} & \dfrac{k_B}{m_B c_B} \\ 0 & \dfrac{k_{BR}}{\rho_g c_g V_g} & \dfrac{-(k_{BR}+k_{RA})}{\rho_g c_g V_g}\end{bmatrix}\begin{bmatrix}T_{RL}\\ T_B\\ T_R\end{bmatrix} + \begin{bmatrix} 0 & \dfrac{1}{\rho_w c_w V_w} \\ 0 & 0 \\ \dfrac{k_{RA}}{\rho_g c_g V_g} & 0\end{bmatrix}\begin{bmatrix}T_A(t)+\Delta T_{Stat}\\ \dot{Q}_{Hz}\end{bmatrix}$$

Equation 1, Thermal state space model of house

Important:

You do not need to fully understand the correlation as it is intended simply to give you a general overview of the various methods.

3.2.3 Frequency response

This is used to describe the dependency of the amplitude relationship and the phase shift dependent on the circuit frequency ω in steady state vibration.
BIBO stability is a prerequisite, i.e. the pole of the transfer function is on the left half plane without exception. The system is excited by a harmonic oscillation (u(t) = Uo*cos(ωo t)). On the output of the system, depending on the system dynamics, a harmonic oscillation will stop again (y(t)=Yo cos(ωo t + φo)). The amplitude changes depending on the factor |G(jωo)| (amplitude response) and the phase shifts to arg(G(jωo)) (phase response).

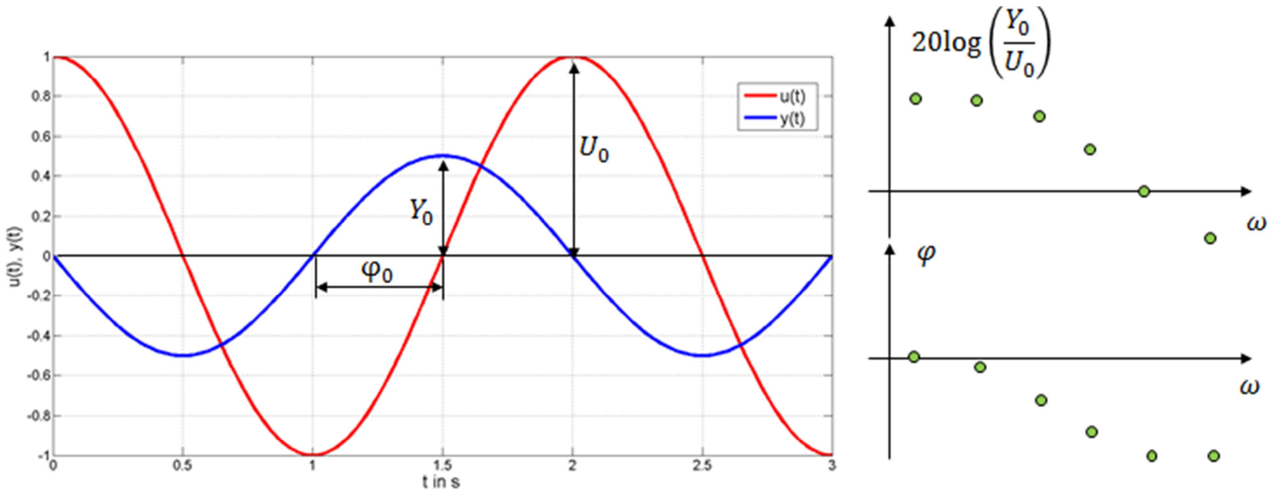


Figure 10, Correlation of frequency response / Bode plot

The relationship between the amplitudes of the I-/O- oscillations (Input variable u, Output variable y) presents the amplitude response |G(jωo)| = Yo/Uo. The time difference between the zero crossings φo leads to the phase response with arg(G(jωo)) = φo

Important:

The Bode plot can only be created if the frequency response has been recorded at several characteristic and various frequencies.

3.2.4 Calculating with dynamical systems

Once a dynamical system has been described by a differential equation, the output signal can be calculated with a given temporal progression of the input signal.
The initial states of the dynamical system can be set to 0 if just the transmission characteristic is of interest. For this reason, only the particular solution

$$x(t) = \int_0^t g(t-\tau) \cdot y(\tau)\mathrm{d}\tau,$$

Equation 2, Particular solution differential equation

must be calculated with the unit impulse response (weight function) (g). The weight function contains all characteristics of the dynamical system, but is not important for further considerations. This method is called convolution. *The dynamic transfer characteristic by the weight function is based on the idea that the system is excited by a very short impulse (Dirac impulse). The weight function is the system's response to the impulse-like excitation.*
The calculation of the integral becomes complex very quickly depending on the system characteristics and the input signal.

3.2.5 The Laplace transform

Using the Laplace transform, given functions of time t (f(t)) can be partitioned into elementary signals. The aim is the mapping of time functions t to complex variable functions s = σ+jω.
A unilateral Laplace transform can be carried out as follows

$$\hat{f}(s) = \int_0^\infty f(t) \cdot e^{-s \cdot t} \mathrm{d}t.$$

Equation 3, Unilateral Laplace transform

Dynamical systems that are described in the time domain can be transformed into the so-called (frequency) range. Initial conditions = 0 are taken as a basis for further observations. The transform does not have to happen by resolving the integral, there are correspondence tables for this purpose which make a calculation easier. It is important that the Laplace transform provides a considerable improvement when calculating linear dynamical systems. The convolution integral mentioned above represents a multiplication of the weight function and the input signal, according to the Laplace transform.

$$x(t) = \int_0^t g(t-\tau) \cdot y(\tau) \mathrm{d}\tau \underset{\ddot{a}quivalent}{\Longleftrightarrow} X(s) = G(s) \cdot Y(s)$$

Equation 4, Convolution integral (time domain / frequency range)

A differentiation is simplified by multiplication with (the Laplace operator) s and integration as a division with s.
The Laplace transform is very often used for the resolution, processing and testing of dynamical systems.

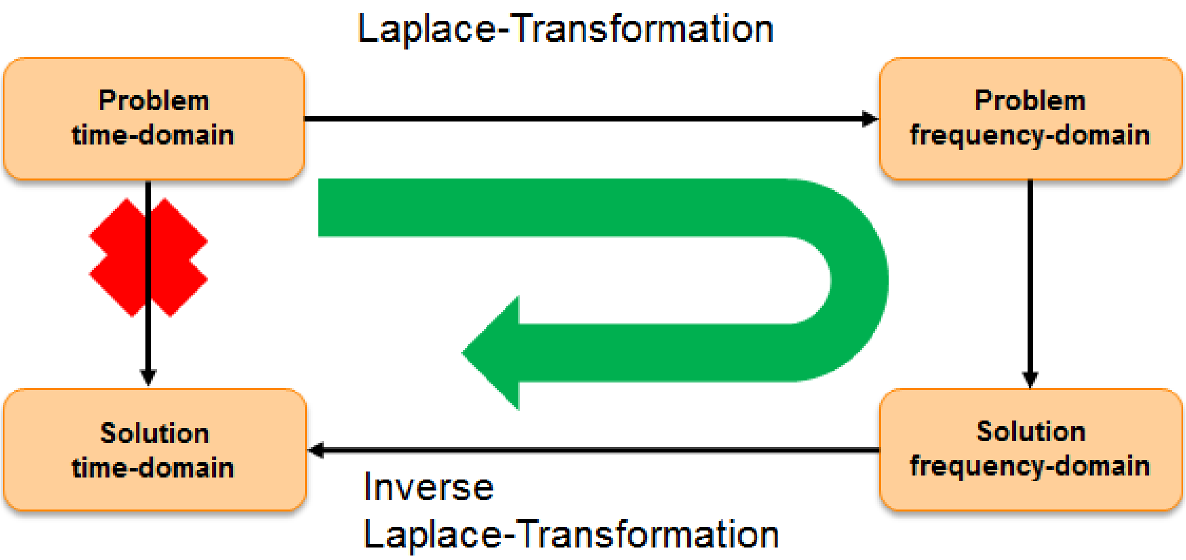


Figure 11, Resolution of dynamic problems using the Laplace transform

3.2.6 The transfer function

The transfer function is defined as the quotient of the Laplace transforms of the output variable and the input variable.

$$G(s) = \frac{X(s)}{Y(s)}.$$

Equation 5, The transfer function

The transfer function is used to describe the input/output behavior. For this reason, it is assumed that the system does not have an initial deflection (all initial conditions of the differential equation = 0).
Using the transfer functions, system characteristics can be analyzed more easily than through direct resolution of differential equations, for example. Using the transfer functions, structured, clear block diagrams can be created to describe and model dynamical systems.
The transfer function of a system can, for example, be obtained through the Laplace transform of a differential equation system. The differential equation system is used to mathematically describe a dynamical system. In the time domain, the use of differential equations is complex. The Laplace transform transfers the correlations from the time domain into the frequency range. By doing this, differential equations are reduced to algebraic equations (differentiation is, for example, only multiplication with s)
The Laplace transform can be regarded as a partition of a given time function f(t) into an infinitely large number of exponential signals. With the Fourier transform, there is the restriction that the signal must be absolutely integrable. Absolute integrability means that the system must be virtually 0 in finite time and, as a consequence, signals with unstable characteristics cannot be Fourier-transformed. In addition, a Fourier transform does not exist for the step signal used in control technology.

The dynamical system to describe the house can, for example, be structurally described using a disturbance variable transfer function (Ds = room temperature / external temperature) and a manipulated variable transfer function (room temperature / heating output).

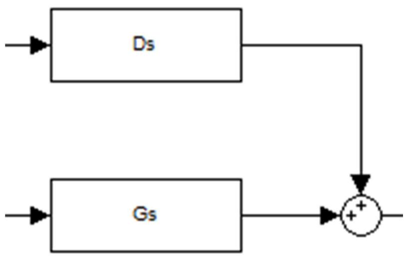


Figure 12, Thermal house model wiring diagram structure

3.2.6.1 Example of differential equation for transfer function

In the following example, a transfer function is created on a first-order low-pass filter.

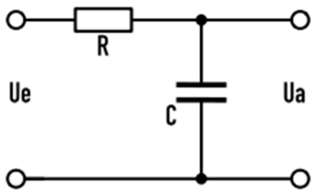


Figure 13, Electrical network

The following network equations can be constructed using Kirchhoff's law.

$$ir(t) = ic(t)$$

$$Ue(t) = ir(t) \cdot R + Ua(t)$$

Equation 6, Kirchhoff 1&2

Using the correlations between the current ic and the voltage uc on the capacitor

$$ic(t) = \frac{duc(t)}{dt} C,$$

Equation 7, Current in the capacitor

the following differential equation can ultimately be constructed

$$Ue(t) = \frac{dUa(t)}{dt} C \cdot R + Ua(t).$$

Equation 8, Differential equation of electrical network

The Laplace transform of a differentiation is the same as multiplication with the complex frequency parameter s = σ+ jω, under the assumption Ua(0)=0.

The Laplace transform of the differential equation can be written as follows:

$$Ue(s) \frac{1}{s \cdot R \cdot C + 1} = Ua(s),$$

Equation 9, Laplace transform of differential equation electrical network

which leads to the following transfer function (first-order delay element, PT1)

$$G(s) = \frac{Ua(s)}{Ue(s)} = \frac{1}{s \cdot R \cdot C + 1}.$$

Equation 10, Transfer function electrical network

3.2.6.2 Task: Standard transfer functions

Within the scope of a standard transfer function, various transfer functions are combined from the process automation.

As demonstrated in the following Figure 14, Standard APROL transfer functions, the characteristic of the transfer function can be selected and then configured using "Drag & Drop".

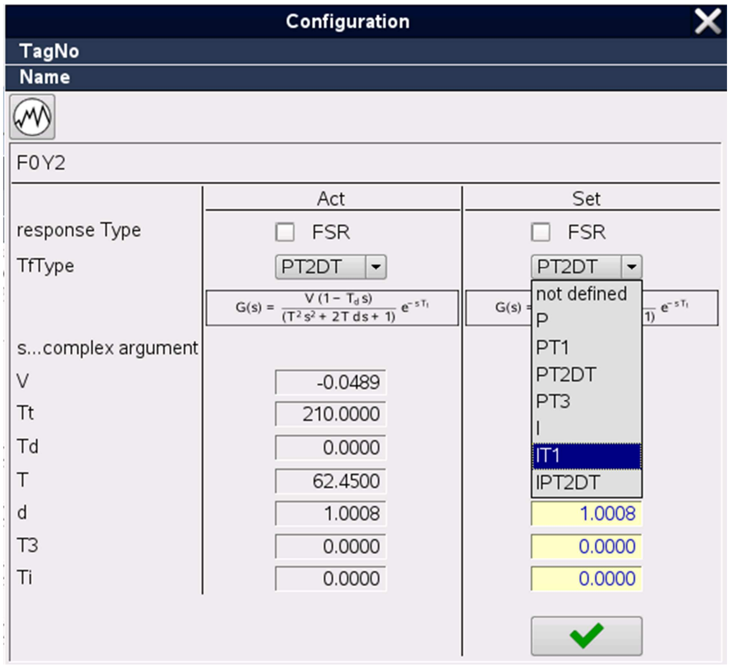


Figure 14, Standard APROL transfer functions

> **Important:**
>
> The relevant configuration is only applied when it has been confirmed by the green arrow.
>
> The step response of the configured transfer function can be seen by activating the trend button .

The following table lists the transfer functions from Figure 14, Standard APROL transfer functions in detail.

Proportional element (P)	$G(s) = e^{-sTt}\,V$
First-order delay element (PT1)	$G(s) = e^{-sTt}\,\frac{V}{T\,s+1}$
Second-order delay element (PT2DT) with differential portion	$G(s) = e^{-sTt}\,\frac{V\,(1-T_d\,s)}{(T^2\,s^2+2\,d\,T\,s+1)}$
Third-order delay element (PT3)	$G(s) = e^{-sTt}\,\frac{V}{(T^2\,s^2+2\,d\,T\,s+1)\,(T_{3f}\,s+1)}$
Integrator (I)	$G(s) = e^{-sTt}\,\frac{V}{s}$
Integrator with first-order integration time constant (IT1)	$G(s) = e^{-sTt}\,\frac{V}{T_i\,s\,(T\,s+1)}$
Integrator of second-order delay element with optional differential portion (IPT2DT)	$G(s) = e^{-sTt}\,\frac{V\,(1-T_d\,s)}{s\,(T^2\,s^2+2\,d\,T\,s+1)}$

Table 3, Standard APROL transfer functions

3.3 Bode plot

A Bode plot shows the magnitude |G(jω)| and the phase arg(G(jω)) of the frequency response as a function of the circuit frequency. Both characteristic curves are displayed separately with logarithmic axis scaling. In control technology, different frequencies are often layered by using correction terms, for example. In this context, the Bode plot has the advantage that an overall frequency response can be determined very quickly. The locus curve presentation (imaginary portion ->y axis and ordinate / real portion -> x axis or x-coordinate) would require the multiplication of two complex numbers in the laying of two frequency responses. In the Bode plot representation, it is simplified to two additions of real numbers. Individual terms can be easily layered using logarithmic displays (log(x)*log(y)=log(x+y)).

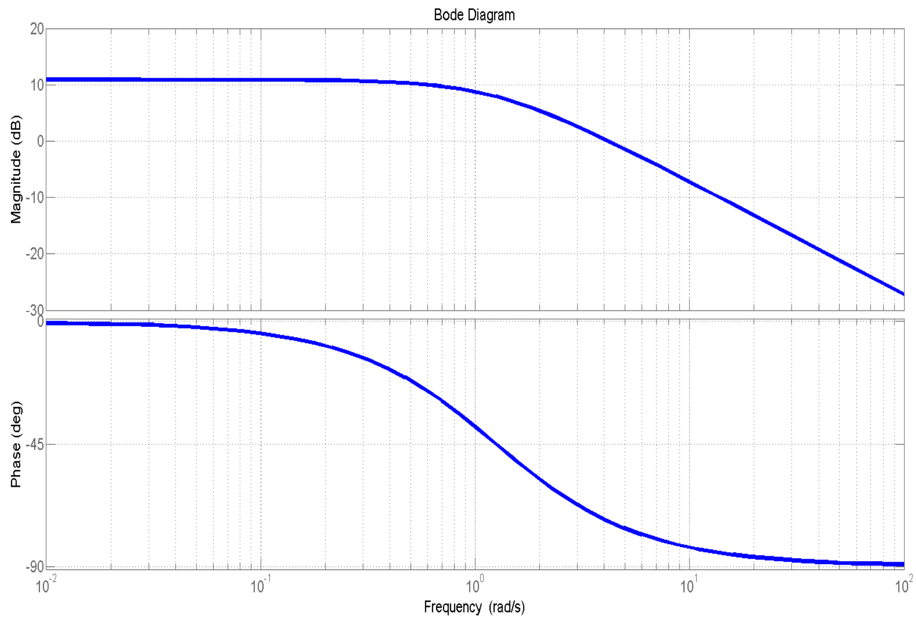


Figure 15, Bode plot of first-order delay element

3.4 The step response

The step response can be used to demonstrate the type of link between an input and an output of a system. The step response can be recorded by measuring the control variable after a step-like change in the manipulated variable.
The shape of the step response can be used to classify the system behavior.

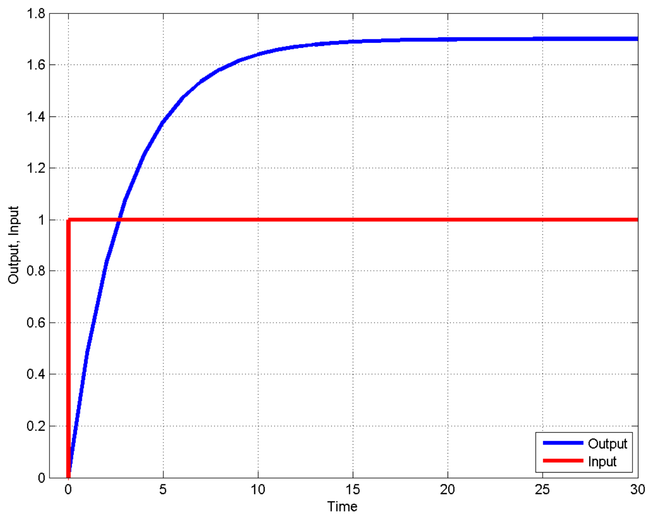


Figure 16, Step response

The derivation of the step response (unit step) results in the weight function and the impulse response.

4 CONTROLLED SYSTEMS

The following characteristics of linear, dynamical systems occur very frequently in the processing industry.

4.1 Processes with dead time

In processes with dead time, the control action only has an effect on the control variable after a certain period of time has expired (dead time). Dead times occur, for example, during transport problems. E.g. processes relating to upstream raw materials systems

$$G(s) = e^{-s\frac{Tt}{Ts}}$$

Equation 11, Transfer function dead time element

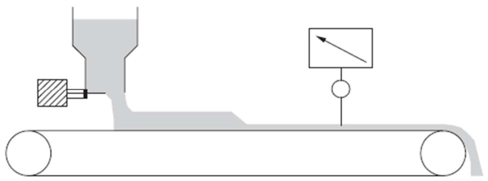


Figure 17, Dead time element principle

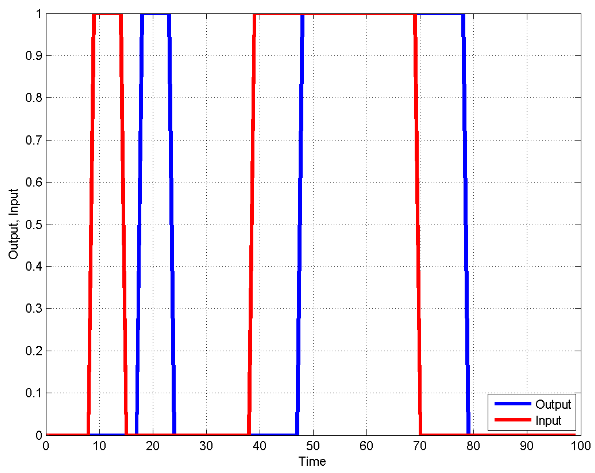


Figure 18, Trace representation of dead time

In control technology, dead times can lead to difficulties. The delay between the manipulated variable and control variable increases the system's oscillation tendency. For example, if the control variable periodically changes to the manipulated variable to shift the dead time (phase shifting increases in proportion to the frequency).

Important:

If possible, dead times should be avoided from the outset, in the planning phase, (short pipelines, sensor - actuator arrangement).

4.2 Proportional behavior

In systems with proportional behavior, the control variable x changes in proportion to the manipulated variable y. The characteristic of a P-transfer function is defined by the controlled system gain V. Proportional behavior is also spoken of if the time span between the controlling element and the control variables is negligibly low from a control technology perspective. Pure P-transfer functions would otherwise not exist in practice.

A valve that is used to impact on the flow rate in a pipe can be used approximately as a system with a proportional behavior. (Valve as actuator, connected flow rate measurement to generate the control variable).

$$G(s) = V$$

Equation 12, Transfer function proportional element

Examples:

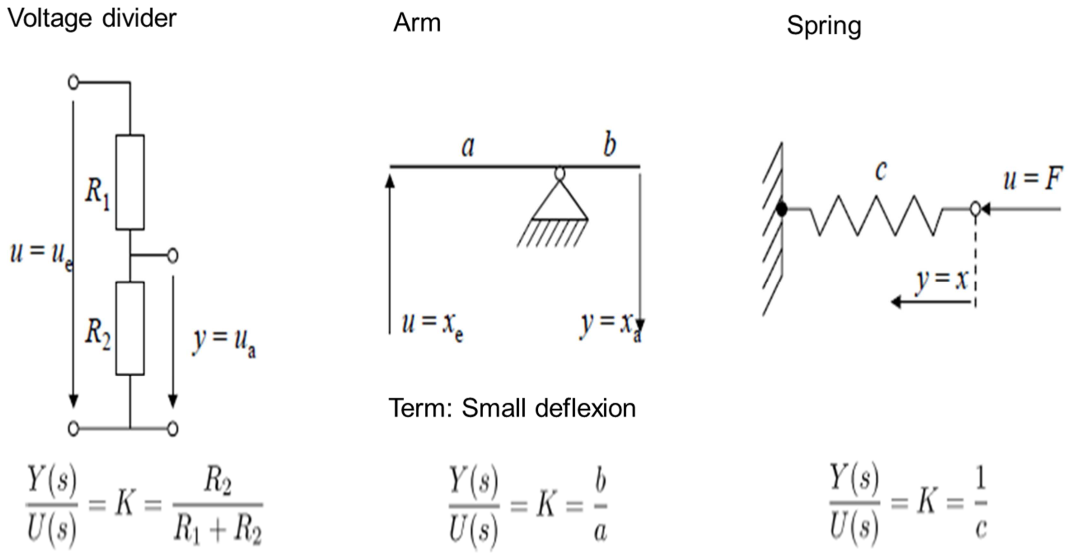


Figure 19, Examples of proportional elements

4.3 Proportional behavior with various-order delay elements

Processes in which energy is stored cause delays between manipulated variables and control variables, which do not change in a step-like way in contrast to dead time. Once a system with energy stores is brought into another state due to energy supply/output (released by manipulated or disturbance inputs), the transfer or conversion of the energy gradually takes place.

The control of room temperature with heating demonstrates a system with several energy stores. By changing the energy supply in the operating room, the room temperature changes depending on a certain time behavior. The time delays derive from e.g. the delivery [transport, heat conduction], heaters (diffusion, heat conduction), convective heat transfer to the air, heat loss in the room (external temperature, heat conduction and convection), solar heat production through heat radiation, etc. If, for example, the manipulated variable (heat quantity) of the heating is taken out, an equalization of the room temperature with the manipulated variables acting on the room takes place. For this reason, controlled systems with inherent regulation are referred to.
The characterization of the transfer behavior of a heating system can ultimately be approximated fairly well with a very simple model (PT1), in which the quick portions of the overall model are not considered at all initially (e.g. heat conduction in pipes). Depending on the system configuration and the requirements of the control design, smaller stores and quicker equalization processes can be disregarded in relation to larger storage elements which cause slower equalization processes.

Important:

In cascade control, the control task is divided into several partial interlaced tasks which have various time constants.

In these kinds of systems, the stationary final value is only achieved after a finite time. The speed at which the control variable changes is also not constant, but time-dependent. If the control variable is close to the steady state, this changes more and more slowly, asymptomatically to the stationary final value.

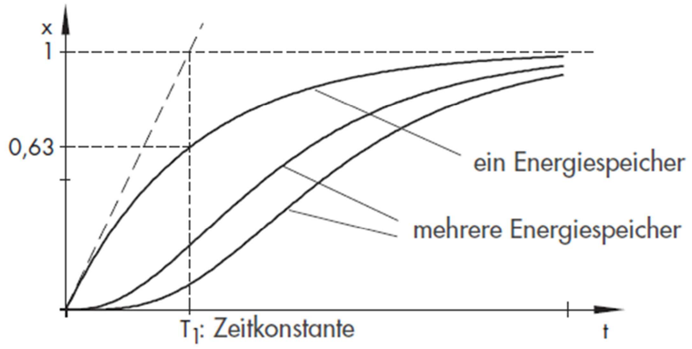


Figure 20, Proportional higher-order delay elements

A first-order system has a time-defining energy store (PT1), a second-order system (PT2) has two, etc.

Important:

The order of the system can be estimated by the number of energy stores. However, this advice does not always apply, as the control variables are also design variables in relation to the manipulated variables (e.g. mechanical systems), but they represent a good approximate value.

4.3.1 First-order delay element

The transfer function of a PT1 element:

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

Equation 13, First-order delay element

A tank is heated, for example, with heating pipes. The heating pipes are filled with a heating medium. The flow can be altered with a control valve.

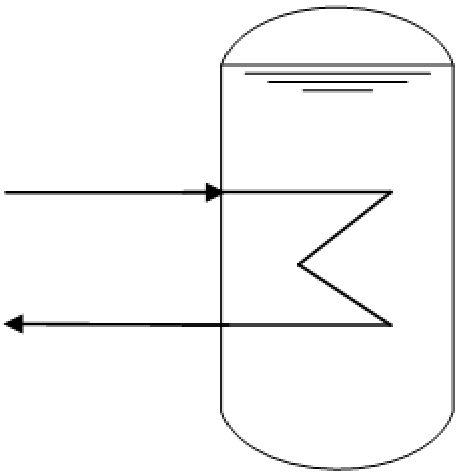


Figure 21, Heating process with PT1 behavior

The temperature of the tank is used as the control variable. When we let more heating medium through the valve to heat the tank, the temperature in the reservoir increases. Once the valve is closed, the temperature in the reservoir decreases gradually again and adjusts to the ambient temperature. In this case, equalization occurs.
More examples for transfer functions with PT1 character

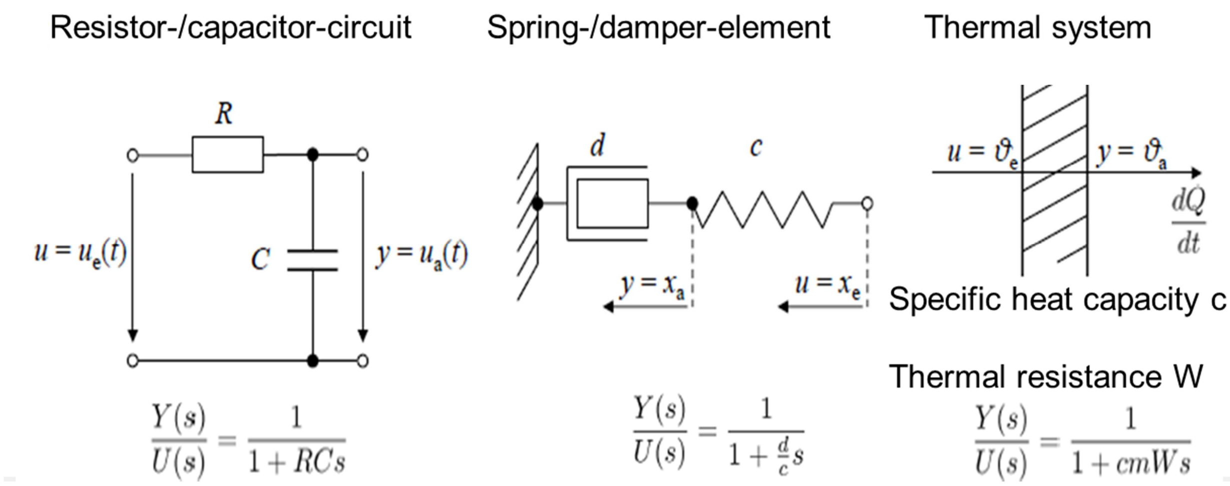


Figure 22, Examples of first-order delay elements

The speed behavior of a gasoline engine with step-like gas pedal operation.

4.3.1.1 PT1 element parameter identification

The parameters of a first-order delay element can be calculated with analytical modeling. However, this approach is only successful if the individual interacting processes occurring within the system and their parameters are known. The system identification provides the opportunity to obtain a model from measured input/output variables through different methods. In this chapter you will only be shown how parameters of PT1 transfer functions can be determined approximately from measured step responses.

> **Important:**
>
> In many cases, an approximately determined model that has been obtained without complex analytical methods or model identification procedures will lead to success in terms of control technology.

The next figure shows a step response of a first-order delay element. The static gain V is equal to the final value of the transfer function. The solution of the differential equation reads:

$$h(t) = V\left(1 - e^{-\frac{t}{T}}\right).$$

Equation 14, Step response PT1

The derivation of the solution in combination with a current tangent t = 0 presents the following correlation

$$\dot{h}(0)t = V\frac{t}{T} = Tangente.$$

Equation 15, Tangent on PT1 step response

The tangent reaches V (unit step hypothesis ΔIn = 1) with t = T. For this reason, the time delay T results from the point of intersection of the tangent with the output value (as parallel to the time axis).

The time constant T can, for example, also be determined by the fact that the time T has expired if the value of the measured step response shows 63% of the final value. The stationary final value is reached in approx. 5 time constants.

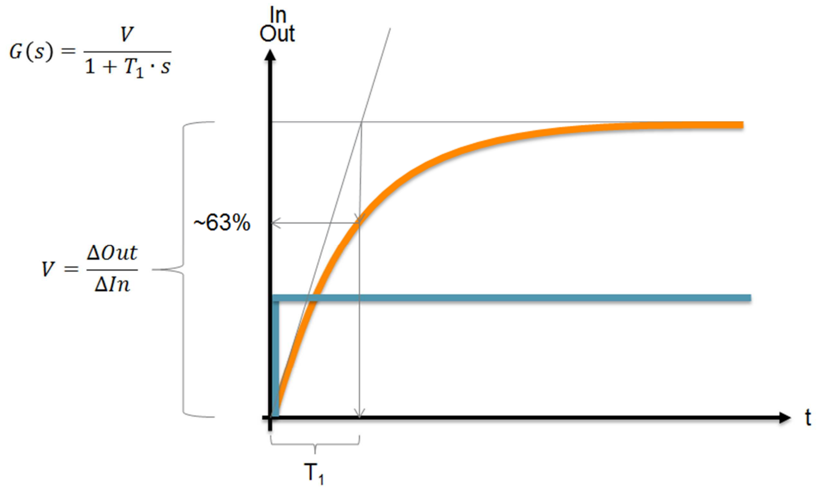


Figure 23, Step response PT1

4.3.2 Second-order delay element

A PT2 element has the differential equation

$$T^2\ddot{y}(t) + 2dT\dot{y}(t) + y(t) = Vu(t),$$

Equation 16, Differential equation PT2

where

$$V \ldots Verstärkungsfaktor,$$
$$T \ldots Zeitkonstante,$$
$$d \ldots Dämpfungsfaktor.$$

The differential equation can be demonstrated through the following transfer function

$$G(s) = \frac{V}{s^2T^2 + 2\,d\,T\,s + 1}$$

Equation 17, Transfer function PT2

For attenuations 0<d<1, the transfer function has a complex conjugated pole pair and can therefore oscillate. This behavior does not exist in first-order systems and cannot be reproduced through the series connection of two PT1 elements, as direct complex poles cannot be configured with the transfer functions available.

Typical processes that have a PT2 behavior (aperiodic oscillation) are heating circuits, DC motors with armature voltage and RC element series connections. Vibratory PT2 elements include mass-spring systems, RLC combinations, and as an approximation, the management behavior of closed control loops. Higher-order delay elements occur, for example, in heat exchange processes through dividing walls or physical mixing processes.

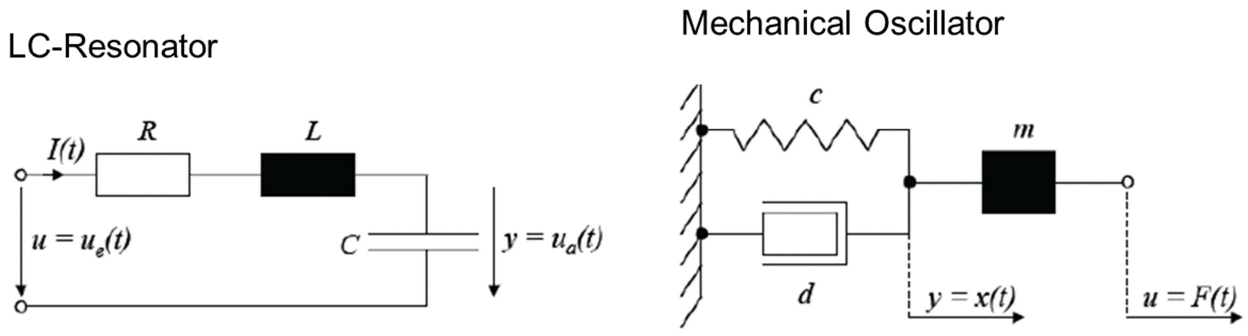


Figure 1, Examples of second order elements

4.3.2.1 PT2 element parameter identification

The dynamics of a PT2 element can be approximated through a PT 1 element, for example, if the delay time Tu is not too high in relation to the equalization time.

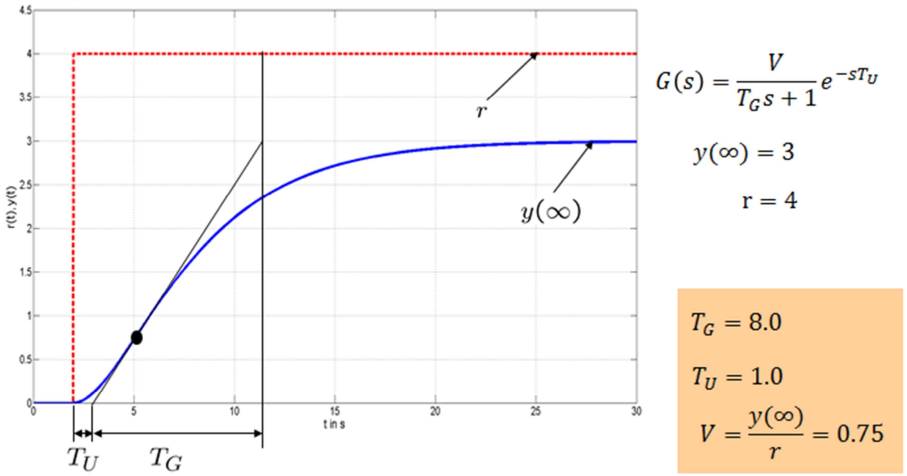


Figure 24, Approximation PT2 by PT1

In this case, the following errors would arise between the approximated PT1 element and the actual PT2 element.

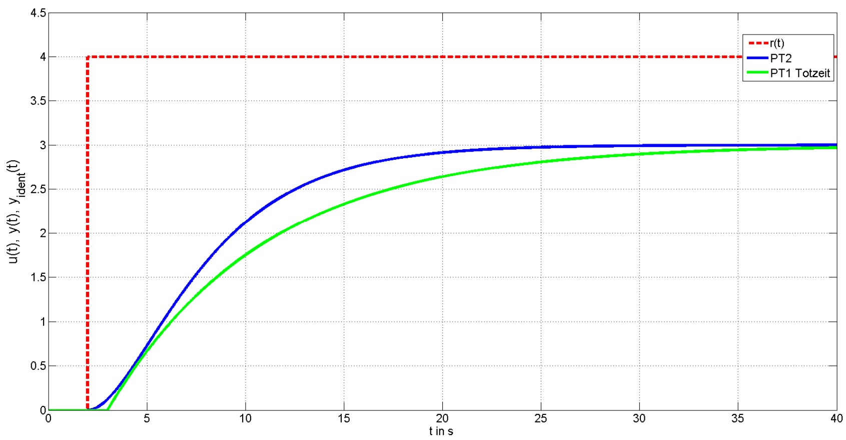


Figure 25, Comparison of approximated PT1 with PT2

The parameters for a vibratory PT2 element can be determined through the following procedure. The receipt of a step response is required, until the system is in steady state vibration.

PT2 (oscillating behavior)

Transfer function:

$$G(s) = \frac{V}{T^2 \cdot s^2 + 2 \cdot T \cdot d \cdot s + 1} \cdot e^{-s \cdot Tt}$$

Gain:

$$V = \frac{\Delta X}{\Delta Y}$$

Damping:

$$d = \frac{ln\left({}^{\Delta X}/_{\Delta Ov}\right)}{\sqrt{\pi^2 + \left(ln \frac{\Delta X}{\Delta Ov}\right)^2}}$$

Time constant:

$$T = \frac{tau \cdot \sqrt{1 - d^2}}{2 \cdot \pi}$$

Dead time

$$Tt \cong as\ per\ diagram$$

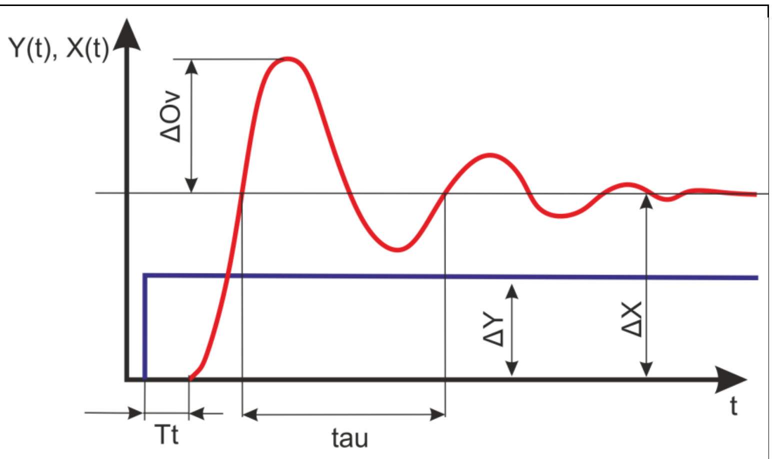


Figure 26, Parameter identification of vibratory PT2

Further identification procedures are not dealt with in detail here (see specialist literature on identification).

4.4 Integral behavior

In contrast to the systems with proportional behavior (controlled systems with inherent regulation), the systems with integral behavior (controlled systems without inherent regulation) do not take up a state of equilibrium if the manipulated variable is constant and not equal to 0.
Reservoirs, whose filling level should be controlled have I-behaviors, for example. The integration time Ti signifies the slew rate of the control variable. (See fig. G(s)= 1/Ti*1/s)

$$G(s) = \frac{V}{s} e^{sTt}$$

Equation 18, Transfer function I-element

The step response of a system with an IT-behavior (Tt = 1, Ti = 1, V = 0.3) can be illustrated as follows.

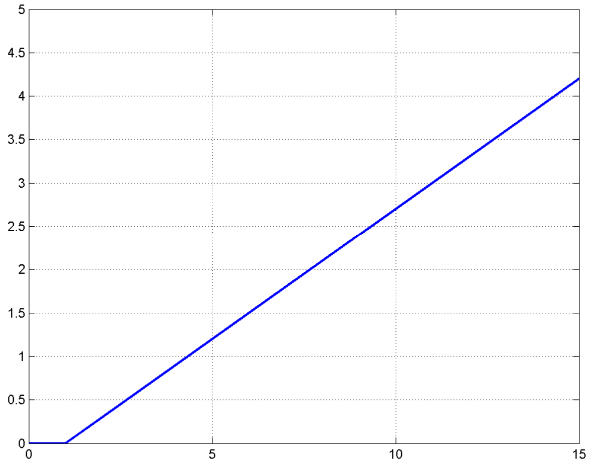


Figure 27, Step response I-element

A tank with a feed valve is filled through a controlling element (valve).

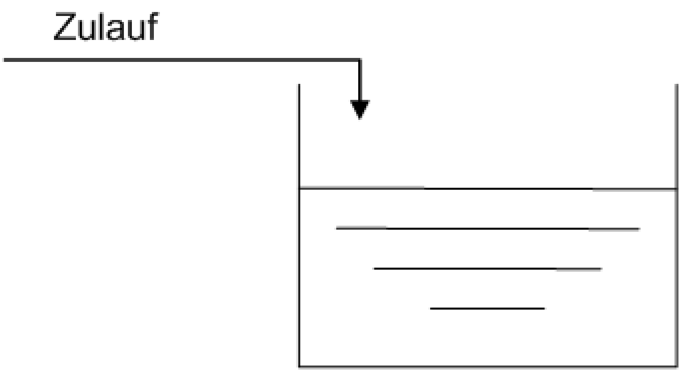


Figure 28, Tank model (I-behavior)

The control variable is the level of the tank. If the valve is opened, the container level rises. However, closing the valve does not result in a decrease in the container level. The filling level stays the same. In this case, equalization does not occur. An IT-transfer element represents a controlled system without inherent regulation.

Examples of processes with integral behaviors are:

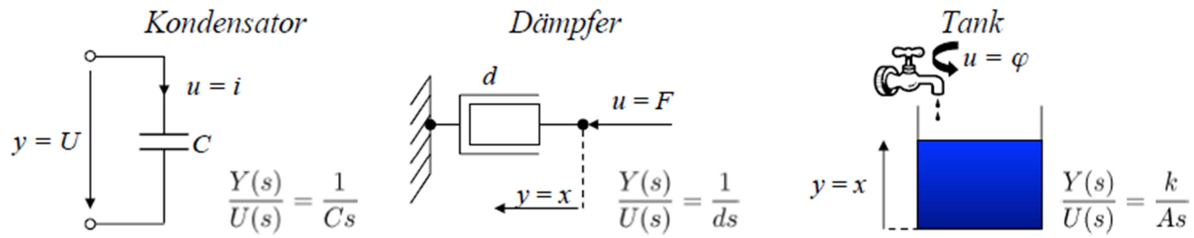


Figure 29, Examples of controlled systems with I-behaviors

4.4.1 Parameter determination I-element

The gain of I-elements can be determined through the step response in the following way.

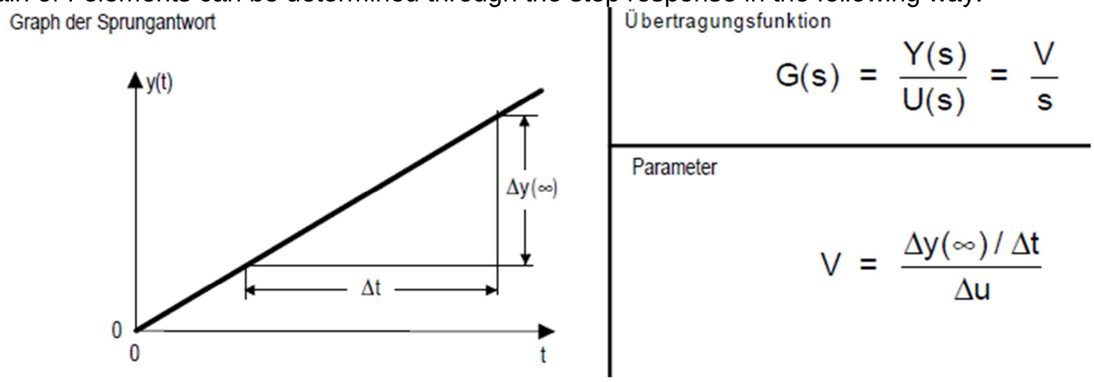


Figure 30, Parameter determination of circuits with I-behaviors

4.5 Integral behavior with various-order delay elements

An integral system with a first-order delay element arises as a result of a feed motor (PT1 behavior) with a downstream spindle (I-behavior), with which the turning movement of the motor (n(t)) is transformed into a position (x(t)).

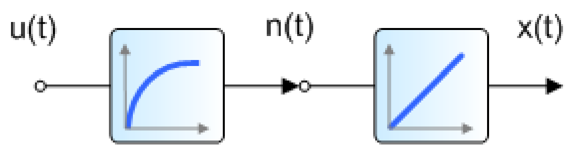


Figure 31, IT1 behavior principle

Important:

More extensive process models can be created by a simple series connection through the linear, time-invariant system class. However, in this case it must be ensured that the absence of feedback applies for each part transfer function.

The transfer function of an IT1 element can be written as follows:

$$G(s) = e^{-s\,Tt}\frac{V}{Ti\,s\,(T\,s+1)}.$$

Equation 19, Transfer function IT1

An IT1 element with the parameters V = 1; Tt = 10; Ti = 1; T = 3; Ts = 1 produces the following step response.

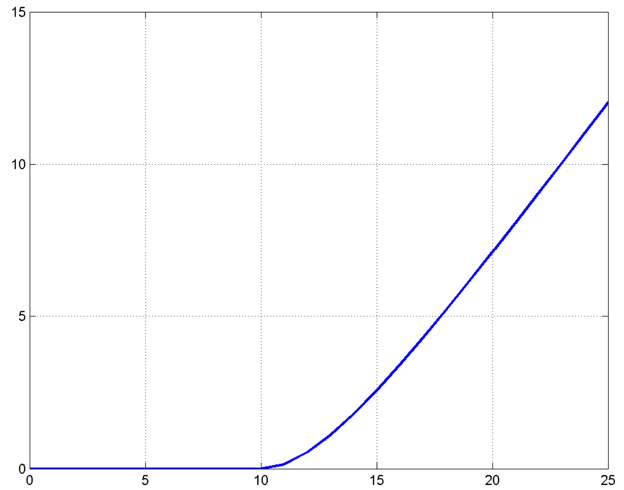


Figure 32, Step response IT1

4.5.1 IT1 element parameter identification

The specific values of an IT1 element can be determined as follows by the step response.

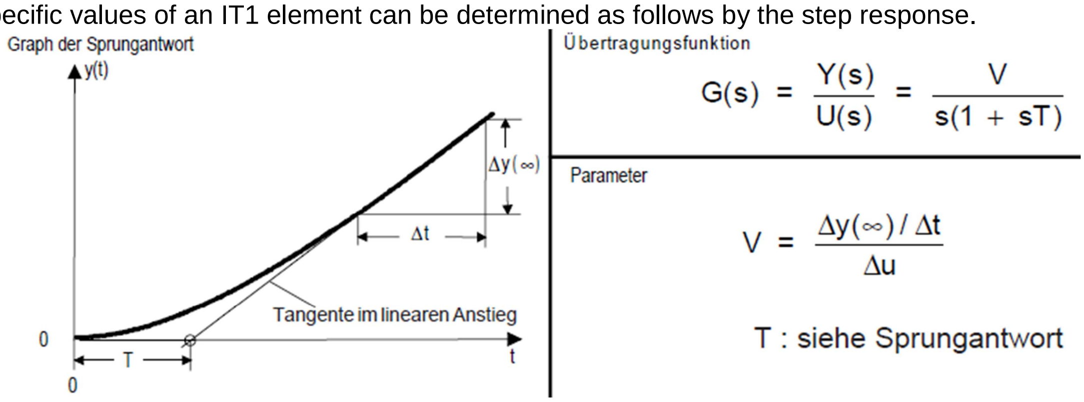


Figure 33, Parameter identification IT1

4.6 Inverse response behavior

Inverse response behavior is characterized by the fact that the control variable first responds in the opposite direction to the manipulated variable before the system passes into the stationary state in proportion to the manipulated variable. An input variable initially reaches the opposite of the intended effect.

> **Important:**
>
> Inverse response behavior is characterized by a non-minimum phase system from a system theory perspective. This characteristic is independent of whether a system has inherent regulation or not.

<u>Typical examples for processes with inverse response behaviors:</u>
Hydroelectric facilities where, for example, water is transported from a reservoir to a turbine through a pipe. The aim is to control the turbine performance (control variable), which can be affected by an inlet valve (manipulated variable) in the pipe. The turbine is already in operation and is in a stationary state. The power produced by the turbine depends on the kinetic energy delivered by the water, which is determined by the water mass and the water speed. A performance requirement causes a small, step-like opening of the inlet valve at the time t0. The flow resistance is disregarded in this observation. Observed long-term (stationary), increasing the mass flow leads to a proportional increase in the turbine power. However, the mass flow in the pipes cannot change in a step-like way at the time t0. Due to the increased valve cross section, the mass flow increases, but the flow velocity of the water decreases. For this reason, the kinetic energy reduces and the power initially changes to the inverted direction. The duration can be estimated by the velocity of propagation of the pressure wave in the pipe.

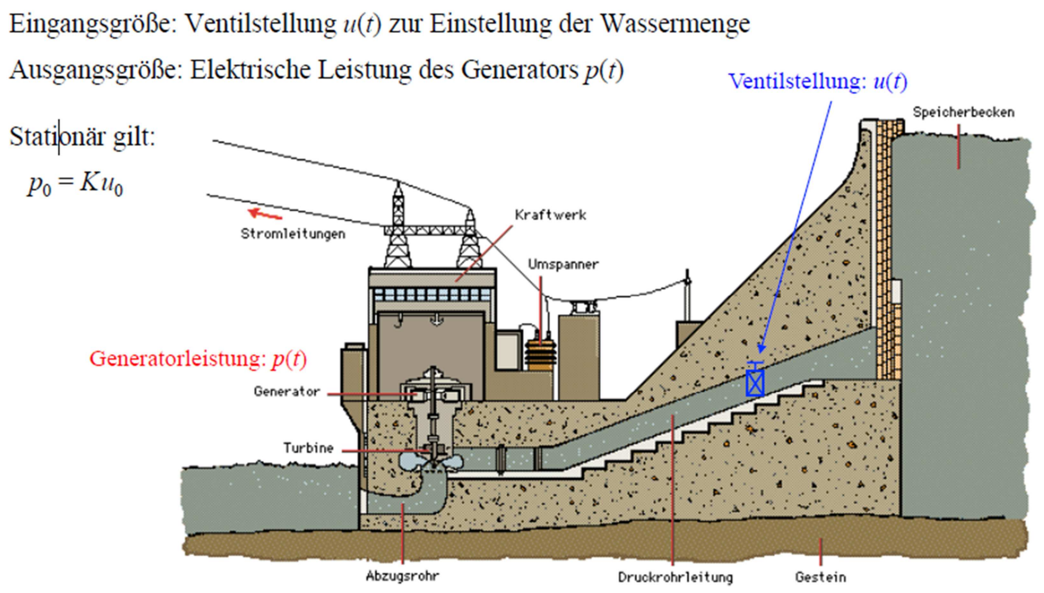


Figure 34, Hydroelectric facility principle

In order to counteract the non-minimum phase behavior, water turbines are not only controlled by adjusting the inlet valve, but also by beam guidance (example from Lunze Control Technology 1, page 320).

In **boiler drum water level control** the inflow to the drum (boiler drum feed pump) is the manipulated variable. The drum is positioned near the boiler roof and is connected with many pipes, which lead into the boiler as the feedline and out of the boiler as steam pipes. In the boiler drum, water and steam are separated.

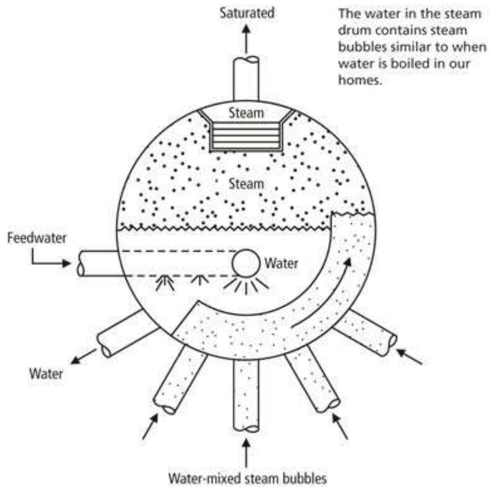


Figure 35, Boiler drum principle

The steam is transferred from the upper part of the boiler drum for overheating before the overheated steam is released into a steam turbine. If the feed water is considerably colder than the steam in the boiler drum and the feed water inlet is increased, the fluid level in the feed water tank may temporarily reduce. By increasing the "cold" inflow, the amount of steam decreases (the volume of the steam reduces) in the boiler pipes. Therefore, the boiler pipes can take in more feed water, whereby the head of liquid in the boiler drum reduces. Example from "Identification and Tuning or Integrating Processes with Deadtime and Inverse Response" by William L. Luyben

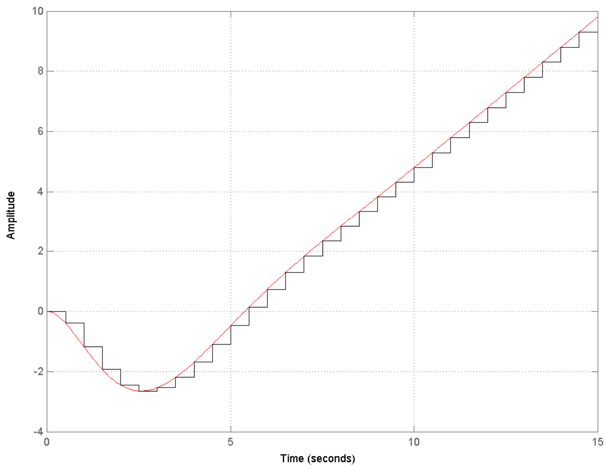


Figure36, Step response system with inverse response behavior (V=1; Ts=0.5; Td = 4; T=1; Tt = 0)

If the feed water inlet is considered as the input variable and the pressure of the produced steam is considered as the output variable, an increase in the water feed leads to a decrease in the steam volume. The pressure of the steam increases in the short-term. In the long-term, the increase in fresh water feed (constant fuel supply) leads to a fall in the steam temperature and therefore a reduced pressure.
(See Lunze 1, page 320).

$$G(s) = \frac{V(1 - Td\ s)}{s(T^2 s^2 + 2\ d\ T\ s + 1)}$$

Equation 20, Transfer function IPT2 non-minimum phase

5 FEATURES OF DYNAMICAL SYSTEMS

The control methods discussed in this training are based on the assumption that a system with a linear, time-variant behavior shall be controlled. In this chapter, we will look briefly at the characteristics of dynamical systems.

5.1 Linearity / nonlinearity

In principle, every process is nonlinear. The objective is to be able to describe the processes through linear models. This is because linearity results in many characteristics which simplify the treatment of dynamical systems considerably.
In many cases, nonlinearity can be disregarded. Furthermore, there is a possibility that nonlinear systems are linearized at an operating point.
A dynamical system is called linear if the effects of two linearly layered input signals are linearly layered in the same way at the system output. The superposition principle applies.

$$u(t) = k\ u1(t) + l\ u2(t) \rightarrow y(t) = k\ y1(t) + l\ y2(t)$$

Equation 21, Condition of linearity

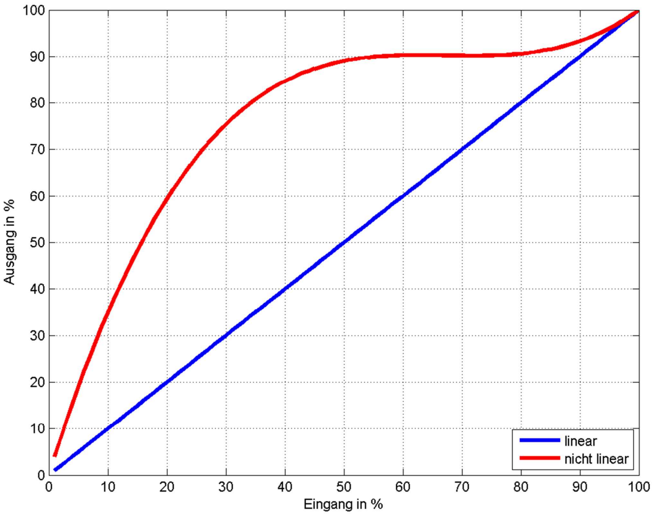


Figure 37, Comparison of linearity / nonlinearity

Very often, existing nonlinearities are dismissed from the outset solely through the formation of linear equations. This is possible especially if the nonlinearities are negligible, if the nonlinearities are compensated by external procedures, or if the control is only permitted for certain operating points. For small displacements to the operating point, nonlinear system equations can also be linearized at the operating point.
However, many controllers are only designed for one operating point of the process. If the process parameters change significantly subject to the operating point, the controller will no longer function as calculated in these ranges.

5.2 Causality

The output of a causal system at time t=t0, only depends on the present and past (t<t0). Conversely, an input variable at the time t=t0 can only affect the present and future progression of the output variable.
In physical systems, causality is a given. However, when designing controllers, there is a risk that there are control rules for which the causality characteristic does not apply.

$u1(t) = u2(t) \rightarrow y1(t) = y2(t)$ for $0 \le t \le T$

Equation 22, Condition of causality

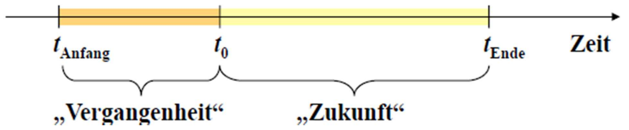


Figure 38, Causality principle

In differential equations, causality means that the input variables y(t) do not exist at a higher derivation than the output variables x(t).

5.3 Time invariance

A system is time-invariant if the system's reaction to a predefined input signal course, to which the system is excited by the signal, is independent of time. The shift of the input signal on the time axis to the right causes an equal shift in the output variable (shift principle)

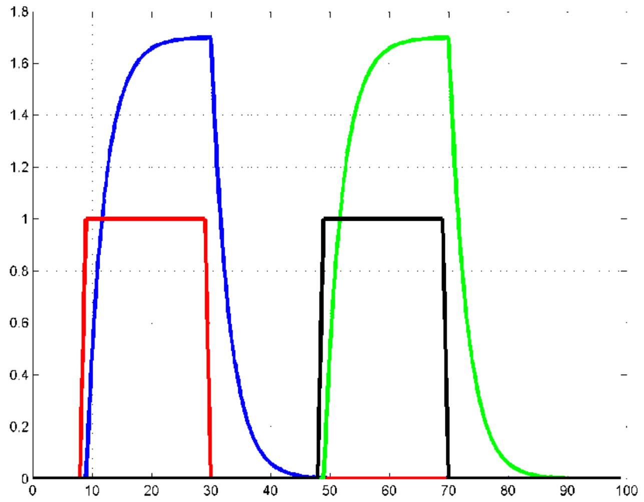


Figure 39, Time variance principle

It is a prerequisite that the initial conditions disappear. In control technology, systems are also regarded as time-variant, the system changes of which proceed very slowly and therefore have a negligible influence on the behavior of the control loop (e.g. minor parameter drift due to deposits, wear).

u1(t) = u2(t-Tt) -> y1(t) = y2(t-Tt)

Equation 23, Time variance prerequisite

A launching rocket shows typical time-variant behavior. The mass changes due to the consumed fuel. In a car, the mass changes (with the load). The change in mass is often < ~4% and can be disregarded.

5.4 LTI systems

(LTI) systems are a very important and frequently used system class in automation technology.
The script is restricted to systems that can be described through time-variant, linear, ordinary differential equations. The restriction is not as great as might first be suspected, since each real process is causal. In addition, in many processes, the effect of the time variance is so minor (e.g. slow) that it can be disregarded. Processes with shared parameters (partial differential equations) can very often be described through concentrated parameters (ordinary differential equations). Nonlinear processes (all processes are nonlinear) are broadly described through linear models. If this is not possible, several linear models can be used which are changed depending on the operating point (PID gain scheduling).

When models are discussed, the user must be aware that they can only be used to approximately describe a process anyway, since it is often the case that not all parameters are known or fluctuate over time. In addition, detailed process models for certain applications (controller design) are too extensive and require too much calculating time.

5.5 Signal-flow graphs

A block diagram is also called a flow diagram or schematic diagram and describes the structure of dynamical systems. Arrows represent time-variable variables and blocks represent dynamical systems. The signals have a clear effective direction. The blocks are reactionless, whereby processes within a block do not cause changes to the input signals (apart from returning the output signal). The blocks can also be identified as transfer elements, especially if the blocks only have one input and one output. The structural view of a system can significantly help to increase understanding of a system behavior, since this type of illustration can display signal courses that are coupled together or system feedbacks very clearly.

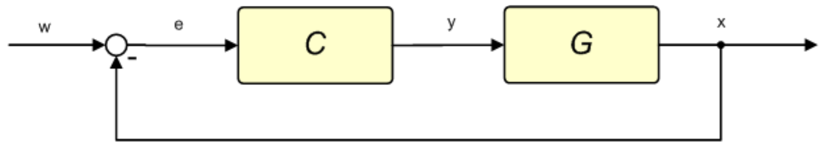


Figure 40, Signal-flow graphs, control technology block diagram

The block diagram shown in the last figure is equivalent to the following equations

- $e = w - x$,
- $y = C \cdot e$,
- $x = G \cdot y$.

According to DIN EN 60050-351 (international electronic vocabulary - control technology), action plans are symbolic representations of action processes in a system using blocks, addition and branch points, which are connected via lines of action. Individual elements of a system are condensed into assemblies or function units, which are called blocks. Lines of action show the direction and path of one or more physical factors. Signal-flow graphs are used to quickly understand the entirety of technical systems.

The following advantages come about as a result of the linear system class in the processing of partial systems.

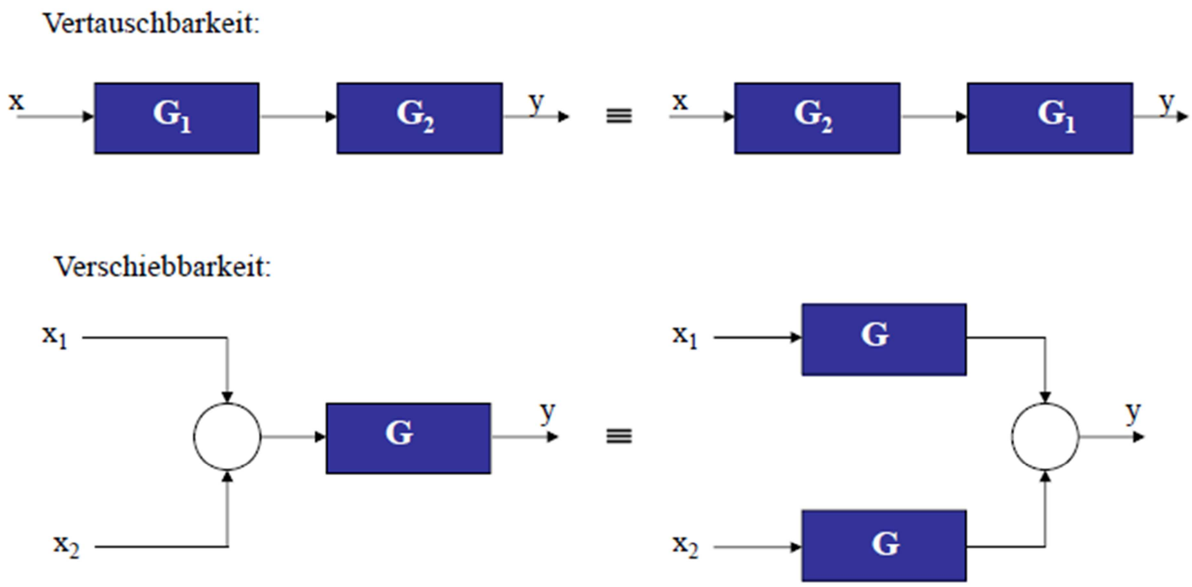


Figure 41, Equivalence of signal-flow graphs

5.6 Discrete-time systems

In contrast to analog technology, digital signal processing gives the control technician the advantage of easy re-configurability and an excellent long-term constancy of the parameters (except for the D/A conversion -> gain and offset errors). The time constants can be precisely set according to the system timer. By doing this, continuous design methods, including in the discrete realm, can be successfully transmitted. However, a few specific points must be observed. This is because the dynamic signal processing steps are implemented through concurrent discrete-time calculation programs.

5.6.1 Process sampling time

With sampling systems, the selection of the sampling time as well as the measurement and actuating frequency have a significant effect on the quality and stability of the control technology solution. For this reason, the amplitude quantization of a dynamical system must be selected carefully enough depending on the process time constants.

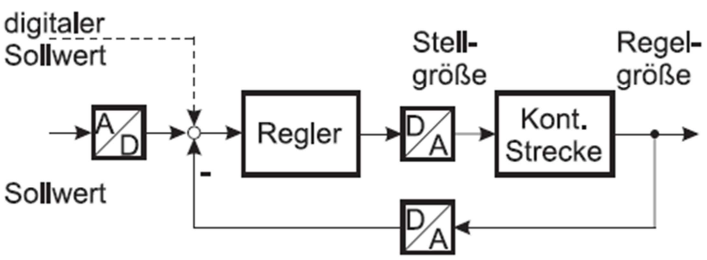


Figure 42, Control loop sampling

On the other hand, if sampling is done too often, there is a risk, depending on the signal quality, that the controller will react very severely to measurement noise, for example.

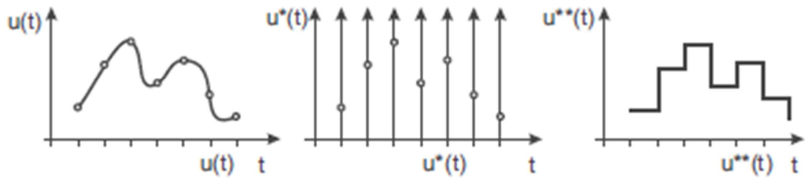


Figure 43, Signal sampling

For a large number of control loops, the sampling time can be determined broadly using the following equation

$$\frac{\min(Tmin)}{Ts} \approx 4 - 10.$$

Equation 24, Correlation of sampling time / system time constants

If the equation is satisfied, it can be assumed from a control technology point of view that the system dynamics are sampled to a sufficiently precise level depending on the dominant system time constant.

5.6.2 Sampling theorem, task class cycle time

You may have seen a car with aluminum-spoked wheels accelerating in a film before. First you see the wheel accelerate. After a certain speed, the wheel suddenly seems to change direction and turn backwards quickly (although the car is still accelerating). Then the wheel seems to become slower until it seemingly stops. The effect that leads to this contradiction is known as aliasing, which occurs in sampled systems. Problems obviously occur if signals are sampled whose frequency is in the range of the sampling frequency (e.g. in the film with a sampling frequency of 25 Hz).
In the following figure, a signal with a frequency of 1Hz is sampled with 1Hz.

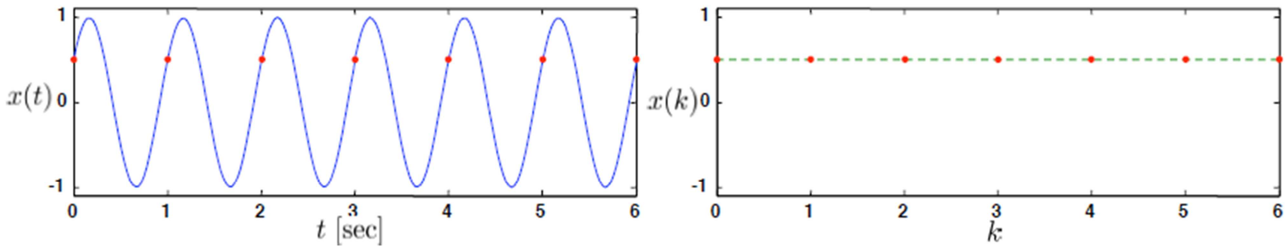


Figure 44, Violation of sampling theorem

In this case, the oscillation is completely lost and the sampled signal shows a frequency of 0 independent of the phasing. The equivalent value changes with the phasing.

It can be identified in the following examples that a minimum of double the signal frequency must be used for sampling, so the signal can be reconstructed after the sampling.

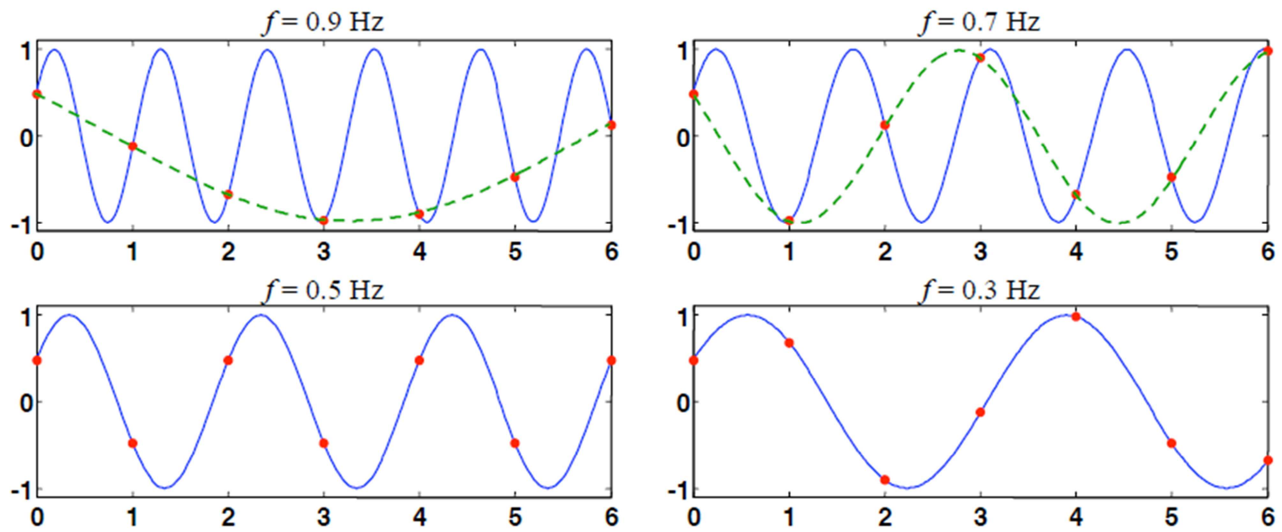


Figure 45, Sampling phenomena

Real signals consist of a mixture of many frequencies. Therefore, this requirement relates to the signal portion with the highest (relevant) frequency fmax.
Based on the Shannon sampling theorem, the sampling frequency f0 must be at least double the size of the highest frequency component fmax of the signal to be sampled

$$f0 > 2\ fmax.$$

Equation 25, Sampling theorem

If this condition cannot be met, aliasing occurs, whereby frequency components with f > ½ f0 are reflected in the lower frequency range. For this reason, high-frequency disturbance signals can cause a large amount of damage, as they are layered over lower frequency wanted signals.
Practice-related values are f0 = 5...10 fmax.

A signal sampled in compliance with the sampling theorem can be reconstructed. The sampling does not lead to loss of information. However, some signals are not bandlimited. In this case, there are no components with a maximum frequency (fmax = ∞). For example, in steps, rectangular signals or ramps, all frequencies are contained from 0 to ∞. These types of signals can be perfectly reconstructed from the sampled signal.

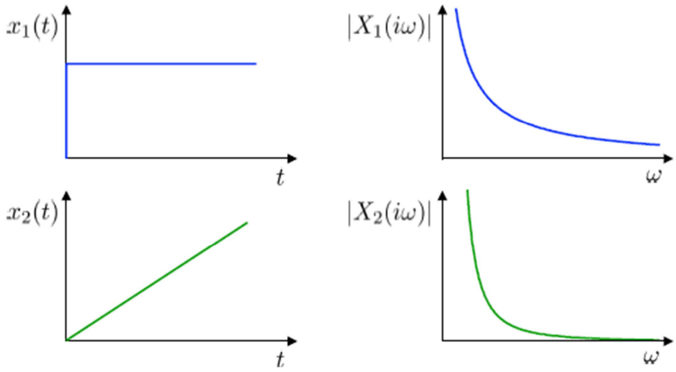


Figure 46, Reconstruction of a sampled signal

5.6.2.1 Anti-aliasing filter

Aliasing can be prevented by using an anti-aliasing filter. The signal is initially sampled at a much higher sampling rate (as required) (e.g. 10 f0). The frequency must be this high as no significant signal components exist in this range. As a result, the reflected frequencies (reflection is caused by sampling) are far outside of the range $|f| < ½ f0$. It can then be digitally filtered and the signal with the required sampling time can be downsampled.

Important:

If an analog Signal contains high frequency contents it is possible that signal sampling lead to eminent failures. Such failures could be prevented with an Anti-Aliasing-Filter. A lot of analog input modules offers per default appropriate input filters.

6 THE CONTROL LOOP

Controls can be differentiated according to various viewpoints. The information display (analog, digital), the automation degree (automatic, manual), the reference input (fixed value -> voltage control, follow-up control -> weather-compensated temperature control or time-oriented control- day/night lowering), the signal processing (time-continuous, multipoint, sample), the modality (electrical, mechanical, hydraulic, pneumatic).

Primarily, controls are set for the safe, economical operation of systems. It is important that controls have a stable behavior. Controllers should not be sensitive to parameter fluctuations or measurement noise.

6.1 Control objectives

Generally, you can distinguish between the following controls depending on the primary control objective. The characteristics that a controller has are defined by the structure of the controller (usually PID in this case) and the configuration.

6.1.1 Fixed value control

With fixed value control (disturbance variable control), a constant reference variable w is taken as a basis. A fixed value controller should primarily eliminate disturbances. For this reason, fixed value controllers are designed for a good disturbance behavior; the typical behavior of the disturbance variable should be known for this (steps, etc.)

6.1.2 Follow-up control

With follow-up control (after-running control), the reference variable w changes depending on the respective operating conditions. The follow-up controller should have a good reference variable behavior. A good reference variable behavior can be achieved by feed-forward signal generation. For this purpose, certain target trajectories can be planned. The additional control errors that arise due to inaccuracies in the generation of the feed forward signal or due to disturbances are removed by the feedback loop (as much control as possible, as little regulation as required).

6.2 Control

The definition of control according to DIN IEC 60050-351 is as follows:
a process whereby one or more variable quantities as input variables influence other variable quantities as output variables in accordance with the proper laws of the system (open-loop control).
A characteristic for control is the open action path or a closed action path where the fact that the output variables being influenced by the input variables are not continuously or sequentially influencing themselves and not by the same input variables (see Figure 47, Control chain according to DIN EN 60027-6*).*

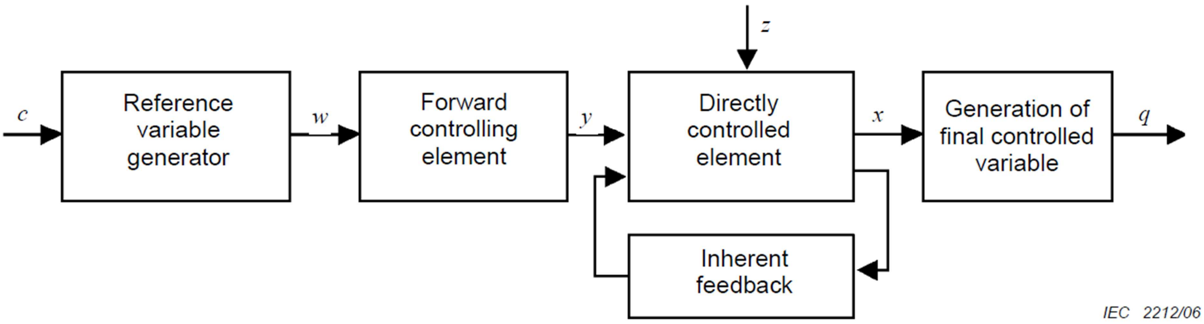


Figure 47, Control chain according to DIN EN 60027-6

6.3 Rules

Definition of rules according to DIN IEC 60050-351: *process whereby one variable quantity, namely the controlled variable is continuously measured, compared with another variable quantity, namely the*

reference variable, and influenced in such a manner as to adjust to the reference variable. A characteristic for the rule is the closed action in which the controlled variable continuously influences itself in the action path of the closed loop.
"Continuously" also refers to sufficiently frequent repetition of individual procedures, as occurs for sampling controllers.

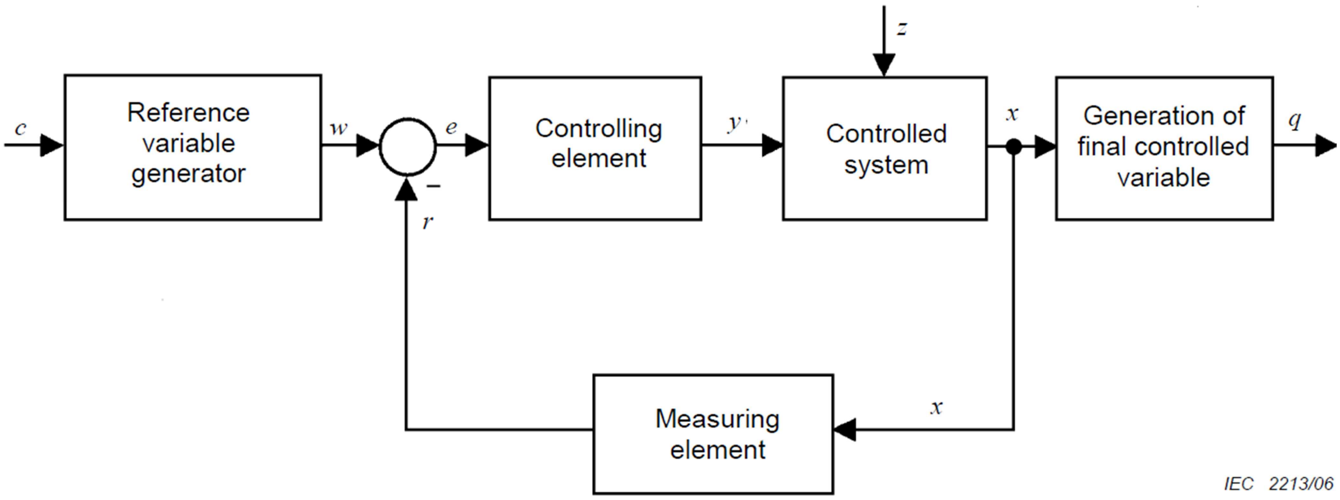


Figure 48, Control loop according to DIN EN 60027-6

A rule has the purpose of bringing the output variable of a controlled system (control variable x) to a certain value (reference value, setpoint w). The control variable must also be able to be maintained at the setpoint against the influence of disturbance variables z. The action flow is closed in contrast to control. The manipulated variables are formed internally with the available actuators from the measured control variables and not provided externally as for controls.

Controls are necessary in order to equalize manipulated variables, adjust the control variable according to the time progression of the reference variable, and to fulfill the control task in the event of changed characteristics of the controlled dynamical system.

A standard control loop is comprised of the following elements.

- The controlled system (plant), as the system that shall be controlled (system, process). 777
  - The control variable, which acts as the output variable of the controlled system (actual value).
- The reference variable, which defines the setpoint of the control variable.
- The measuring instrument (sensor, measuring device), which delivers the control variable measurement value to the controller.
- The control deviation (control error) as the difference between the reference and control variable.
- The controller, which forms a control signal from the control error, which impacts on the process.
- The actuator (controlling element), which acts as a connector between the controller and the system. The output variable of the actuator is the manipulated variable.
- The manipulated variable describes the influence of variables that affect the process, but cannot be affected. Manipulated variables can be unmeasurable, measurable and even predictable (data from the production planning).

The formula symbols and terms of control technology derive from DIN EN 60027-6 and can be specified as follows:

w	Reference variable	The reference variable is fed to the control loop externally and cannot be influenced by the control loop (∃ exception cascade, master tracking, manual slave). The reference variable sets the currently required value of the control variable x, which must be maintained by the control loop despite disturbances.
x	Control variable	The control variable is a physical variable, which can be recorded metrologically. The control variable should generally follow the reference variable. The control variable is the output variable of the controlled system and the input variable of the controller.
	Actual value	The actual value represents the control variable as a numerical value, i.e. the current value of the control variable measured by the sensor.
y	Manipulated variable	A physical variable to influence the controlled system (process). The manipulated variable is predefined by the controller and is the output variable of the controller and the input variable of the controlled system.
yR	Controller output variable	When dividing the controlling system into the controller and actuator, the variable yR stands for the output variable of the controller or the input variable of the actuator.
	Controlling element / actuator	Device that influences the manipulated variable through the control value. The controlling element is used to meter a power current or mass flow.
z	Disturbance variable	Cannot be influenced by externally-acting variables which have an undesired effect on the process to be controlled.
e	Control difference	Difference between the reference value and the control variable (e=w-x). The control difference has the same unit as the control variable.
G(s)	Transfer function	Mathematical relationship between output and input of a dynamical system in the frequency range.
System		A delimited configuration of interacting entities (e.g. panel that delimits the system from the surroundings). The connections between the system and the surroundings are created through signals.
Plant		This is a system, e.g. an automated part of a production process, in which at least one physical variable is to be controlled by a set input.

Table 4, Formula symbols for control technology

6.4 Control task solution process

Successful implementation of a control loop can be subdivided into the following points.

- Formulation of the control task:
  - Should the controller mainly be able to control reference variables or disturbance variables or both variables well? What characteristics should the closed loop have? What signal forms will be used to calculate the reference variables or disturbance variables?
- Determination of the control variables:
  - Which control variable(s) must be included in the control task?
- Determination of the manipulated variables:
  - Which manipulated variable(s) are necessary to be able to influence the control variable(s) (physical operating principles)?
- Depending on the control procedure, can or must a model of the controlled system by made?
  - What approach has been selected (analytical, system identification or is a step response sufficient)?
- Only now can the controller be designed, put in place and configured.

6.5 Stability criteria

The idle state of a dynamical system is called stable if the system goes back to its idle state by itself after being taken out of the idle state.

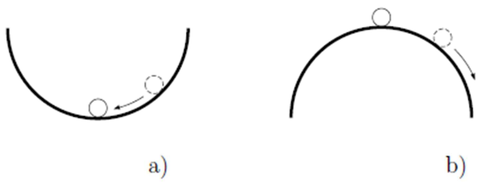


Bild 3.2: Stabilität von Ruhelagen
a) stabile Ruhelage b) instabile Ruhelage

Figure 49, Stability of the idle state

An LTI system is called asymptomatically stable if its weight function (impulse response) goes down to zero (absolute integrability of the impulse response).

$$\int_0^\infty |g(t)|\mathrm{d}t < \infty$$

Equation 26, Absolute integrability of the impulse response

Limit stability prevails in systems whose number of weighting functions with increasing t does not exceed a final value (system without inherent regulation).

6.5.1 BIBO stability

The stability of the idle state relates to autonomous systems. The transfer function of a system can be assessed by the BIBO stability criteria (Bounded Input Bounded Output) or input/output stability. An LTI system is properly BIBO stable if a limited output function can be generated for each limited input function (condition of infinitesimal initial deflection). A transfer function is precisely BIBO stable if all poles are situated on the right side of the plane.

Important:

If the idle state of a system is asymptomatically stable overall, the system is also BIBO stable. Inversion only applies if the order of the transfer function concurs with the order of the differential equation system

6.5.2 The phase margin / degeneration

The gain crossover frequency separates the frequency band which is amplified by the open loop from the frequency band which is weakened by the open loop. This gain crossover angular frequency ω_c is defined by the following correlation

$$|L(j\omega_c)| = 1.$$

Equation 27, Gain crossover frequency of an open loop

L(s) indicate a transfer function of an open loop. "A stable, open chain whose phase response only has one crossover frequency ωc leads specifically to an I/O stable control loop, if the phase margin is positive" (Control technology 1, (Lunze, 2010), p. 441)

The phase interval in place of the gain crossover frequency at -180°is called the phase margin.

$$\Phi = \arg\big(L(j\omega_c)\big) + 180°.$$

Equation 28, Phase margin

Φ can be used to assess the stability of control loops. In addition, the oscillation tendency of the closed loop can also be deduced from the phase margin (see frequency characteristic curve and loop shaping methods). To ensure that transfer behavior decays within an acceptable time period, the phase margin must not only be positive, but should range in the interval of 30°- 60°.

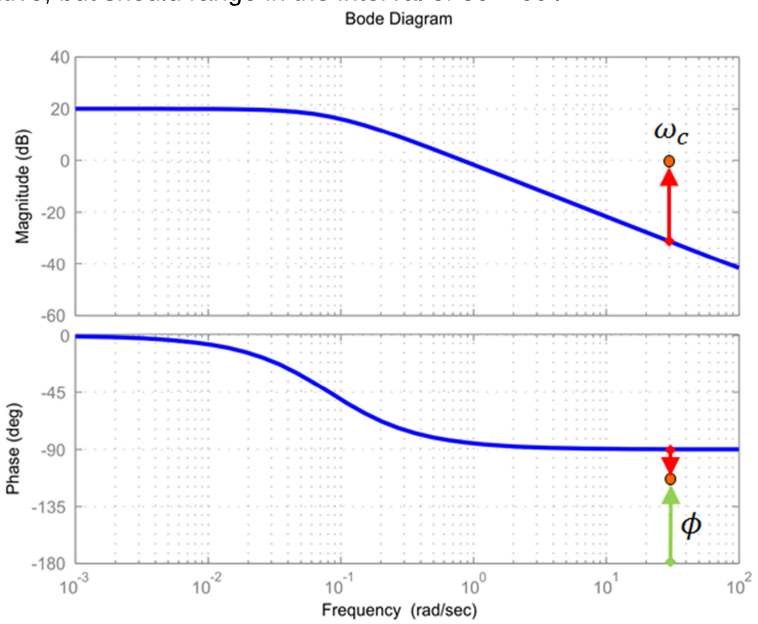


Figure 50, Bode plot with phase margin

A positive phase margin produces feedback. However, a negative phase margin leads to positive feedback. Reducing the phase margin causes an increase in the oscillation tendency and overshooting of closed loops. A phase margin of 0°is reached in a phase shift of 180°(open loop). For example, a sinusoidal oscillation of the control variable (steady state vibration) has a shift of exactly one half-wave compared to the control deviation e (it has an opposing algebraic sign). For this reason, the negative feedback (degeneration) becomes a positive feedback for this frequency. The system becomes unstable. Conversely, a stable behavior is reached when the gain of the open loop is smaller than one in the gain crossover frequency.

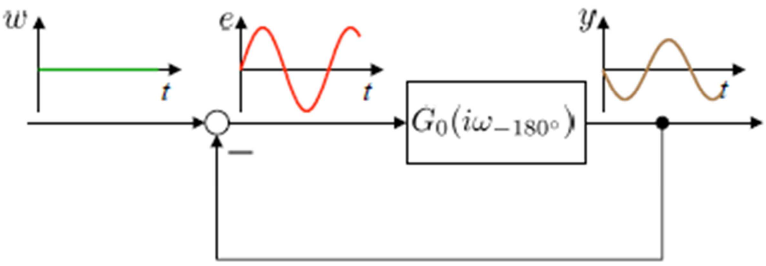


Figure 51, Phase rotation, positive feedback

Caution: dead time elements have a significant impact on the stability characteristic of feedback system, as dead time systems only influence the phase, not the amplitude. The phase shift increases in proportion to the frequency, without affecting the amplitude response. Dead time elements reduce the phase margin and lead to unstable control loops if the phase shift of the dead time element is greater than the phase margin of the system with no dead time.

7 DISCRETE CONTROLLER

Discrete controllers provide manipulated variables which can range between discrete values. Depending on the possible states of the controller and the manipulated variable, you can distinguish between two-/three- or multi-point controllers. The advantage of discrete control loops is their simplicity, as they can be operated with simple switching actuating devices. If discrete controllers are combined with energy stores, steady control variables are likely to arise as a result (heating control).

7.1.1 2-point controller

The two-point controller is the simplest controller and only has two states (On/Off). Two-point controllers are used in actuators that do not permit any other states or in systems which need to be controlled in the simplest way possible. Control quality with two-point control is secondary. In a conventional heating system, the water temperature in the boiler does not need to be controlled that precisely, as a few degrees of deviation does not have an impact on the room temperature (this is assumed by a continuous controller). However, the burner is often turned on and off with a two-point controller before a continuous dose of the fuel requirement is made depending on the air feed through a continuous control loop.

A two-point controller cannot reach a setpoint exactly (requirement of infinite quick on/off switch cycles). For this reason, hysteresis is present in a two-point controller, which is used to define the fluctuation margin of the actual value to the setpoint. Hysteresis is present through upper and lower switching thresholds. As soon as the setpoint is outside of the hysteresis, a switching operation is carried out in the two-point controller. In the case of burner control, for example, if the water temperature is below the set limit, the burner is turned on. If the water temperature limit is exceeded, the burner is switched back off.

The control variable of a two-point controller typically fluctuates around the setpoint. In the case of a switching hysteresis that is too small, the actuator is often turned on and off. In this case, the control variable fluctuates with little deviation and higher frequency around the setpoint. Due to the high frequency, the controlling element is burdened accordingly (wear, power-on load -> load peak).

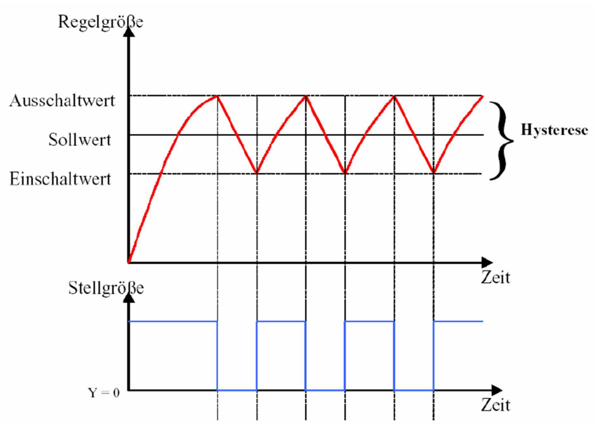


Figure 52, Two-point controller

Important:

Two-point and three-point controllers do not need any parameter tuning procedures and have almost no maintenance costs meaning they are very cost-effective to use.

<u>Two-point controller applications:</u>
For general fixed value control such as temperature control, e.g. using bimetallic strips, heating, filling level fixed value control e.g. toilet flushing.

7.1.2 3-point control

For three-point controllers, the manipulated variable can adopt three different values (e.g. positive signal/OFF/negative signal).

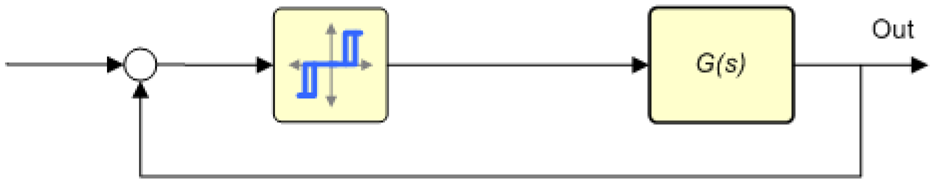


Figure 53, Three-point controller principle

Example: level control of a buffer tank. Removal is dependent on the downstream processes. The filling level is controlled by a motor valve, which can either be completely opened or completely closed. With three-point controllers, as with two-point controllers, the control variable is limited by the setpoint due to hysteresis. In this case, the controller would open the valve as soon as the value falls below the lower limit (OPEN -> OFF). The filling level of the tank increases again until the maxium set value of the switching hysteresis has been reached. Now, the controller would close the valve again (CLOSE -> OFF).

7.1.3 Comparison of switching controllers

Three-point controllers can, for example, be used to control actuating drives. The controller takes the actuator into account through the motor actuating time. The actuator and the control loop have a position feedback with which the discrete motor control can be directly influenced by an internal follow-up controller. The switching controller (e.g. two-point or three-point controller) is connected to a control loop without further dynamic wiring to the controlled system.

The amplitude and frequency of the switching oscillation on the control variable are only influenced by the system characteristics and the selection of the control output value. The application is limited to first-order controlled systems, as the probability of stability problems increases with increasing system orders.

In stead of the switching controller, a continuous controller can be used, for example, which can also provide discrete control signals by connecting a pulse width modulator.

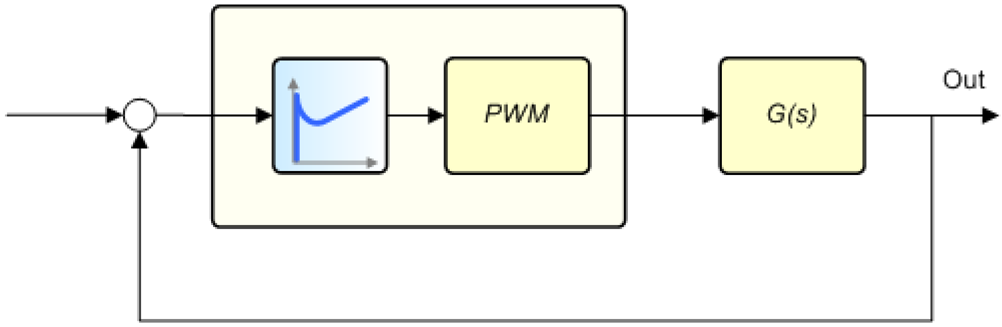


Figure 54, Continuous controller with PWM

The switching frequency, including a minimum pulse duration and a minimum pulse pause, can be predefined independent of the controller. As for standard control loops, the controller can be designed using a linear system model or a manual standard tuning process regardless of the switching frequency. As long as the switching frequency is considerably higher than the dominating system time constant (a small switching oscillation is to be expected for the control variable), the nonlinear effects of the PWM can be disregarded.

7.1.3.1 Step controller

A further improvement in controller behavior can be achieved by controlling the controlling elements, whose current position in the control loop can be utilized.

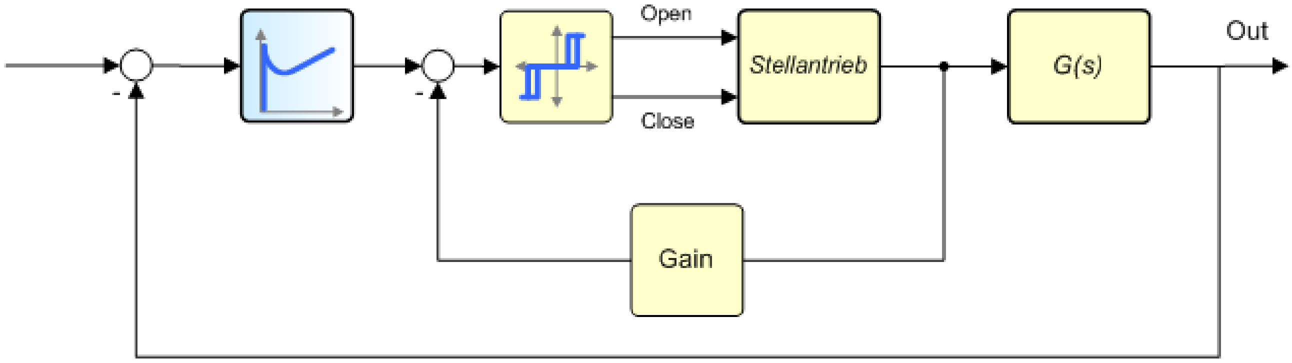


Figure 55, Step controller principle

8 CONTINUOUS CONTROLLER

Continuous controllers can assign the manipulated variable of any value within the adjusting range.

8.1.1 P controller

The higher the control deviation, the greater the manipulated variable must be (spring principle). The P-portion can reduce the control deviation quickly, but the control deviation cannot generally disappear. The manipulated variable is proportional to the control error. The P-controller only outputs the manipulated variable if the control error |e|>0 exists. The level of the manipulated variable depends on the level of the control error and the gain factor Kp. The higher the Kp, the quicker the controller reacts to control errors. The risk of overshooting and oscillation tendency increases with an increasing Kp.

Important:

The P-controller provides relatively quick reactions in a very simple control structure. The P-controller is not able to control permanent control deviation.

The input / output behavior of the P-controller can be written as follows:

$$y(t) = K_p \cdot e(t).$$

Equation 29, Input / output behavior of a P-controller

The P-controller is very simple and cheap and is often only implemented mechanically. Such as here, to control a fluid level.

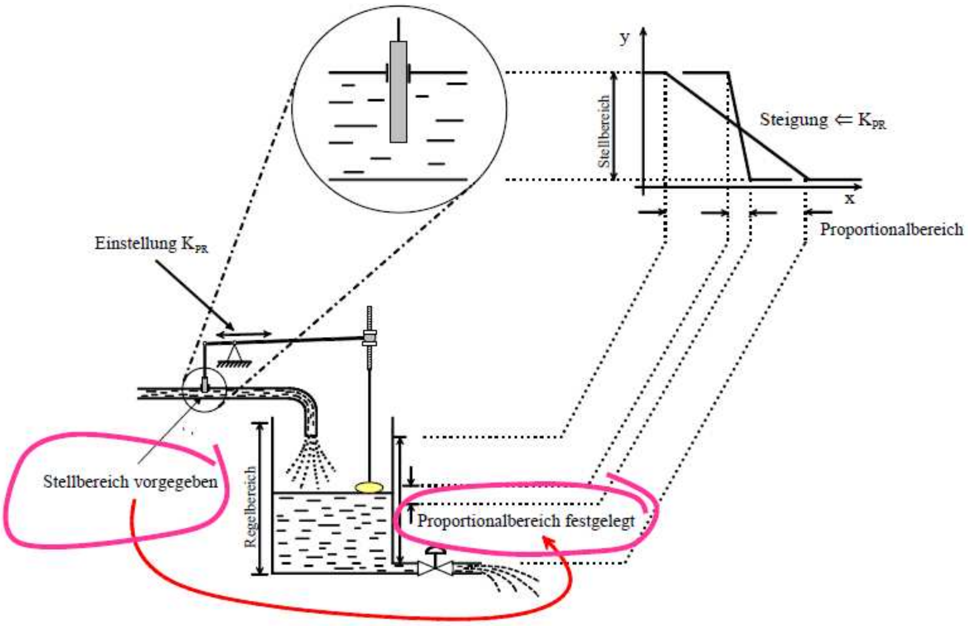


Figure 56, Mechanical water level control with a P-controller

The P-portion can be graphically represented as a spring. The spring constant is the gain factor Kp and the deflection of the spring symbolizes the control error. The higher the deflection, the stronger the force of the spring is, in order to return to the initial deflection (control error = 0).
The manipulated variable of the P-controller is limited by a maximum and a minimum manipulated variable. As soon as the control deviation is too great, the P-controller reaches saturation. The linear range between the limits is known as the "proportional range". The greater the proportional gain of the controller that is selected, the smaller the range becomes. The P-controller always needs an infinite control deviation to be able to react (except the operating point in the middle of the range randomly delivers the desired setpoint).
By increasing the Kp, the permanent control deviations could be kept small accordingly. The problem in this case is that, towards high frequencies, practically every controlled system reaches a phase shift of 180° due to its lowpass behavior. At this point, in stead of degeneration, positive feedback occurs, which, with a loop gain of >=1, leads to an unattenuated oscillation of the system. The P-controller gain should be kept low, so that before positive feedback is reached, an entire loop gain of the open loop <1 is reached (loop shaping methods).

<u>P-controller areas of application</u>
In stable processes (due to the high gain), in which a constant stationary control deviation is permitted.

Follow-up controller in the cascade. In cascade control, a remaining control deviation of the inner control loop is often not significant. The P-controller reacts very quickly. Remaining control deviations can then be managed by the master controller.

8.1.2 The I-portion

The manipulated variable changes for as long as a control deviation occurs. For as long as the control error is not equal to 0, the manipulated variable changes in proportion to the amount of time of the control difference. The dynamics of the I-controller can be indicated by the control deviation and the integral action time Tn.

$$y(t) = \frac{1}{T_n}\int e(t)\,dt$$

Equation 30, Manipulated variable through I-controller

With the I-portion, remaining control deviations can be eliminated. In contrast to the P-controller, the I-controller needs a longer integral action time for the control process. The I-controller regulates precisely but slowly and tends to oscillate. With the integral action time Tn, the speed of the I-controller can be changed.

<u>I-controller areas of application:</u>
For pulsating variables such as pressure and flow control, which have a relatively small equalization time in relation to delay time (Tg/Tu<3), pure dead times.

8.1.3 The D-portion

Through the D-portion the system has an addition phase margin of 90°, which is why greater gains are possible at higher frequencies. In this way, the controller reacts more quickly and can, for example, suppress oscillations faster during a transient response.

$$y(t) = T_v \frac{d}{dt} e(t)$$

Equation 31, Manipulated variable portion through D-portion

The greater the control variable change, the greater the control procedure (damper principle). Through the D-portion, the controller reacts with a greater manipulated variable if the control error increases greatly, even if the control deviation still has not adopted a great value. The pure D-portion is not technically feasible (s · Tv · Kp). If, for example, a step is operated on a D-portion, the step response must give an infinitely high, infinitely small impulse. However, this fact is of a purely theoretical nature. To prevent the D-portion from decaying immediately, the D-portion is implemented through a DT1 element. A current value of the filter time constant Tf in practice is Tf = Tv/4.

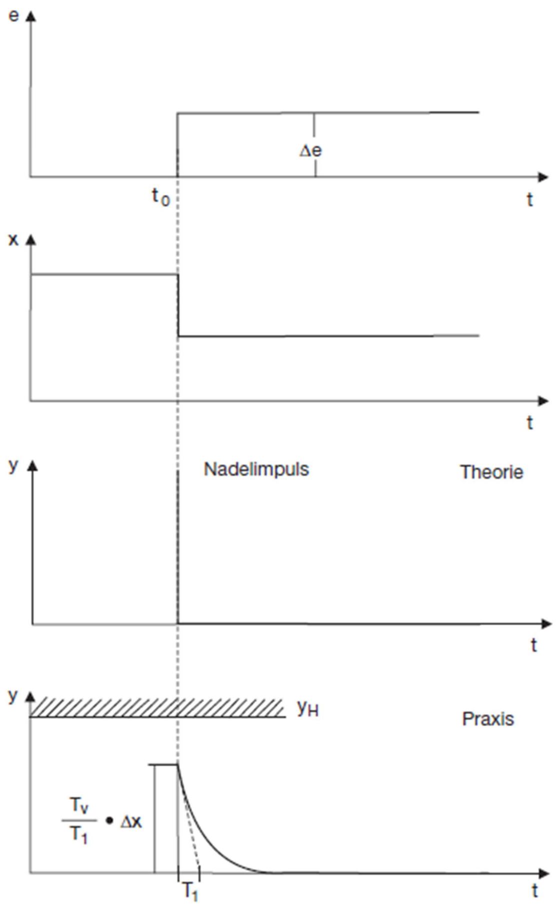


Figure 57, Reaction of D-portion

The correlation can be shown with the following differential equation

$$T_f \frac{d}{dt} y(t) + y(t) = T_v \frac{d}{dt} e(t)$$

Equation 32, Differebtial equation DT1 element

and transfer function

$$G(s) = \frac{y}{e} = \frac{T_v s}{T_f s + 1}$$

Equation 33, Transfer function DT1 element

.

The D-controller reacts to the time changes of the control error through a signal that is proportional to the change speed. With noisy signals, the D-portion should not be used in the PID controller.

Important:

As the D-portion only reacts to the changes in the control error and then returns to the initial value of the manipulated variable, a D-controller on its own is not suitable. The D-controller is not able to correct continuous control deviations.

The D-portion can be represented as an attenuator. The greater the change in position (speed and amplitude change) the greater the effort of the D-portion is to control this again. Remaining control deviations cannot be compensated through the D-portion. The D-portion is mainly used in conjunction with the P-portion as a controller.
The differentiator primarily works at high frequencies and causes a phase lead of 90°. This additional phase margin can be used so that a greater gain can be selected even at higher frequencies (PD controller). A controller with a D-portion (PD, PID) reacts more quickly and can, for example, suppress oscillations more effectively during a transient response. Time constant control errors are not corrected by the D-portion.
The D-portion reacts to changes in the control variable and counteracts this. For this reason, the D-portion is found in combination with the P- or the I-portion.

8.1.4 PD controller

The D-portion reacts very quickly to changes in the control error. The P-portion determines the long-term behavior.

$$R(s) = K_p\left(1 + \frac{T_v\, s}{\frac{T_v}{T_f}\, s + 1}\right).$$

Equation 34, Transfer function of PD controller

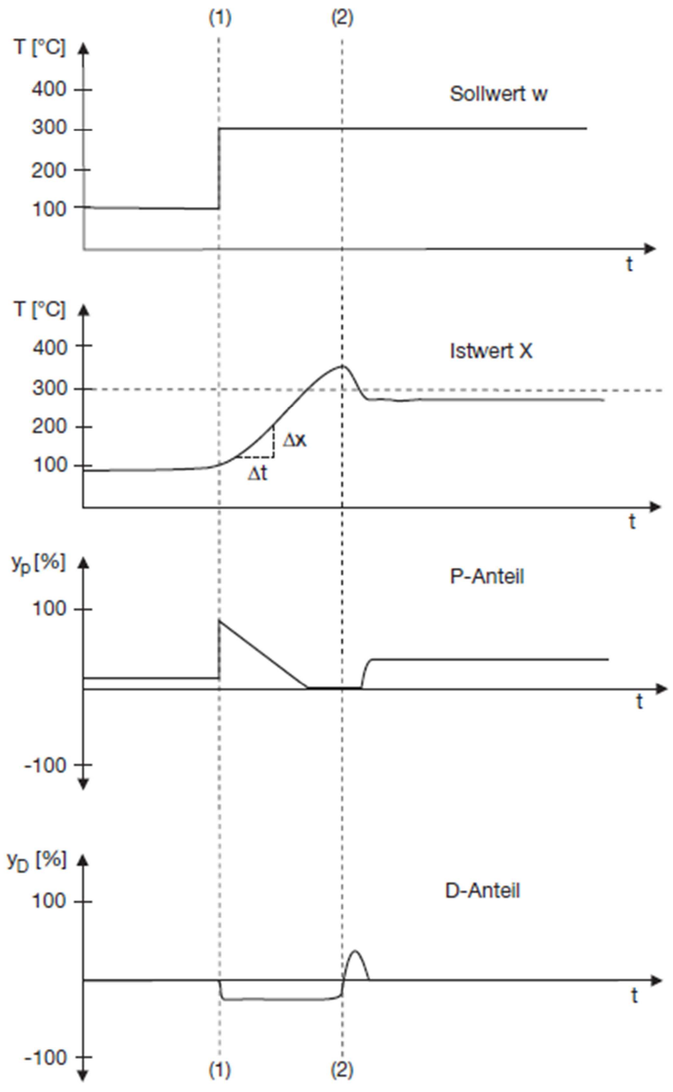


Figure 58, Reaction of PD controller

<u>PD-controller areas of application:</u>
Thermal processes with good insulation behave in a similar way to integrators. If, for example, power from an energy store was shifted to another, the systems can react very idly. The integrator would only slow down the process, and for this reason, a PD controller could be used here. It is only important that the measurements are free of measurement noise, since all step-like oscillations in the measurement signal by the PD controller act in an intensified way on the controlling element (the control gets worse with declining measurement signal quality).
PD controllers are often used in the control of moving objects, such as robots, airplanes, submarines, ships and rockets. The D-portion has a stabilizing effect for moving objects with regard to disturbances to the control variable. If it is necessary to control systems on which high setpoint steps are carried out, the setpoint attenuation (see Changing the PID controller structure, page 78) of the controller should be activated, so that the D-portion only reacts to fluctuations in the control variable.

8.1.5 PI controller characteristics

The PI controller combines the advantages of the P- and the I-portion (high dynamics and no remaining control deviation). The P-portion tries to quickly intercept the occurring control difference, without eliminating a remaining control deviation. Then, the I-portion removes the remaining control deviation. The P-portion reacts immediately depending on the gain factor Kp after the occurrence of a control error. The I-portion only comes into full effect after a while (integral action time Tn). The greater the gain Kp and the smaller the integral action time Ti is selected, the quicker the controller works.

The integral action time Tn is used to configure the time span that passes by until the I-portion creates the same level of the control output value in relation to the P-portion. The lower the integral action time, the higher the I-portion of the controller.

$$R(s) = K_p\left(1 + \frac{1}{sT_n}\right).$$

Equation 35, Transfer function of PI controller

As can be seen in the transfer function of the PI controller, the P-portion also affects the integrating behavior of the controller. If Kp is increased, the I-portion works even more quickly.

<u>PI-controller areas of application</u>
The PI controller is very often used in the technical processing industry, for applications such as pressure and temperature control. The D-portion is not used if a quick system reaction is not required immediately. In processes which can be affected by high disturbance variables or which struggle with very noisy signals, a D-portion is counter-productive too. The D-portion is also superfluous if only one energy store is present in the process to be controlled, such as in temperature or pressure control.
If systems with high dead times need to be controlled, a D-portion is also not needed.

8.1.6 PID control

The PID controller can be optimally adapted to various processes and is the controller that is most often used. The D-portion enables quick detection with regard to control differences. The P-portion is used to target and quickly adapt control variables to the setpoint. The I-portion enables the subsequent, precise control of the remaining control error. The PID controller is the ultimate universal controller and is the most frequently used.

8.1.6.1 Parallel form

A parallel form is also known as an independent PID controller.

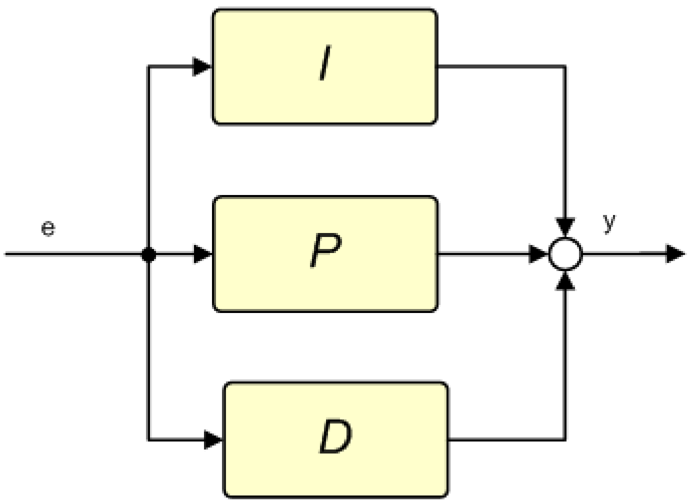


Figure 59, Parallel PID principle

The parameters of the controller structure cannot be ascribed to a clear physical correlation.

$$R_{parallel}(s) = K_p + \frac{K_n}{s} + \frac{K_v\, s}{\frac{K_v}{T_f}\, s + 1}$$

Equation 36, Transfer function of a parallel PID

8.1.6.2 Standard form

An uncoupled (T_n hat keinen Einfluss auf T_v) structure can be represented as follows.

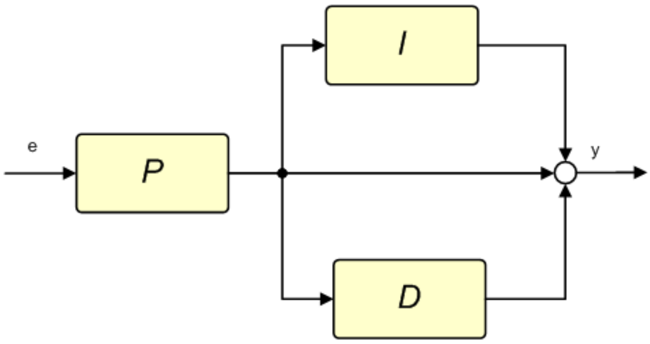


Figure 60, The PID standard form principle

A clear, physical effect can be ascribed to the parameters. The sum of the past errors is equalized over the time period of the integral action time (seconds). It is assumed that the control loop remains unchanged. The lead time Tv is used as an error prediction. The gain Kp is used for scaling.

$$R_{standard}(s) = K_p \left(1 + \frac{1}{sT_n} + \frac{T_v\ s}{\frac{T_v}{T_f}\ s + 1} \right)$$

Equation 37, Transfer function of PID standard form

This form is also known as the ideal or non-interacting form, see (Visioli, 2006).

8.1.6.3 Serial form

The coupled structure resulted from the construction of pneumatic controllers in the last century. According to (K. Aström, 1995), table 3.2, this structure is still widespread in commercial PID controllers.

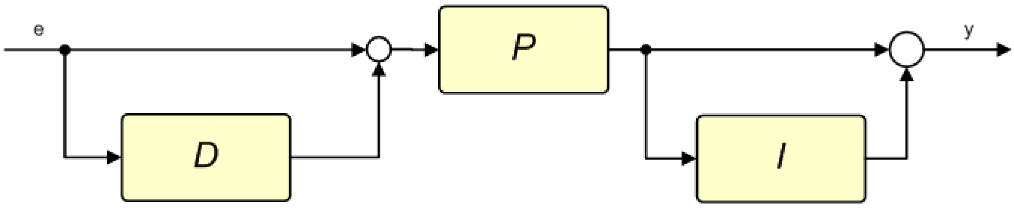


Figure 61, Serial PID principle

A change in the differential portion in the interacting form leads to a change in the integral portion. The form is the result of serial switching of a PD with a PI controller and derives from the area of the analog controller.

$$R_{Seriell}(s) = K_p' \left(\frac{T_n'\ s + 1}{T_n'\ s} \right) \left(\frac{T_v'\ s + 1}{\frac{T_v'}{T_f'}\ s + 1} \right)$$

Equation 38, Transfer function of a serial PID

When setting PID parameters which must be accepted by other controllers, you should always find out which equation has been used in the respective controller. If necessary, the parameters must be converted using the above equations.
The conversion of serial controller parameters into the parameters of the uncoupled representation is always possible.

$$Kp = \acute{K}_p \frac{T'_n + T'_v}{T'_n}$$

Equation 39, Conversion of Kp from serial to standard form

$$T_n = T'_n + T'_v$$

Equation 40, Conversion of Tn from serial to standard form

$$T_v = \frac{T'_i T'_v}{T'_n + \acute{T}_v}$$

Equation 41, Conversion of Td from serial to standard form

PID parameters of the uncoupled form can only be converted if

$$\acute{T}_n \geq 4T'_v$$

Equation 42, Condition of conversion from standard form to serial

applies.

$$\acute{K}_p = \frac{K_p}{2}\left(1 + \sqrt{1 - 4\frac{T_v}{T_n}}\right)$$

Equation 43, Conversion of Kp from standard to serial form

$$\acute{T}_n = \frac{T_n}{2}\left(1 + \sqrt{1 - 4\frac{T_v}{T_i}}\right)$$

Equation 44, Conversion of Tn from standard to serial form

$$\acute{T}_v = \frac{T_n}{2}\left(1 - \sqrt{1 - 4\frac{T_v}{T_i}}\right)$$

Equation 45, Conversion of Tv from standard to serial form

8.1.6.4 Industrial ISA standard form

The parallel and standard form enables a greater parameter setting range for the serial structure, as zeroing the counter when making a suitable selection for Ti and Td can also become complex. In this way, a band-stop filter can be created with a parallel structure. The parasitic Tv/Tf pole symbolizes the delay that occurs in the real difference formation. With the ideal PID controller, an impulse occurs through the ideal differentiation which is slurred by the real differentiation.
The following controller structure enables flexible programming of the controller. The controller structures presented in chapters 8.1.6.1, 8.1.6.2 and 8.1.6.3 can be completed by weighting the control error of the proportional portion with β and the differential part with γ to control loops with 2 degrees of freedom. If, for example, the standard form from Chapter 8.1.6.2 is completed with the setpoint weighting (P-portion with β and D-portion with γ), the following I/O behavior arises as a result

$$y = K_p\left(\beta \underbrace{w - x}_{e_P} + \frac{1}{sT_i}\underbrace{(w-x)}_{e_I} + \frac{T_d\, s}{\frac{T_d}{T_f}\, s + 1}\underbrace{(\gamma w - x)}_{e_D}\right).$$

Equation 46, Output- / Input behavior of ISA standard form

The block structure of the PID controller is summarized in the following figure.

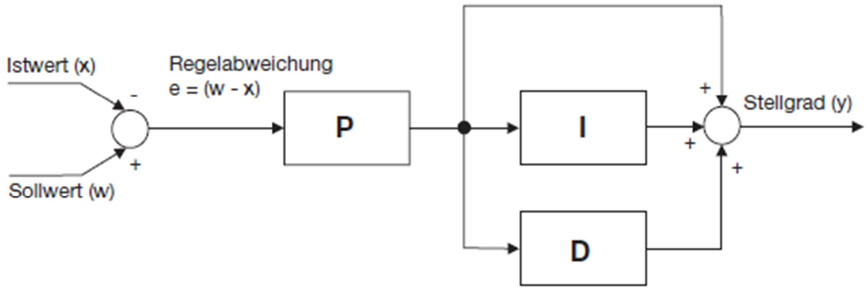


Figure 62, Standard PID in block diagram

8.1.7 PID controller structure

The ISA standard form is used as a controller structure for a B&R PID controller. Due to the variable setting option, different controller structures can be created through configuration.

Important:

A major advantage of this basic PID form is that, if required, the reference variable behavior can be configured independently of the disturbance variable behavior.

In the following figure, the PID's characteristics are presented in a structural block diagram.

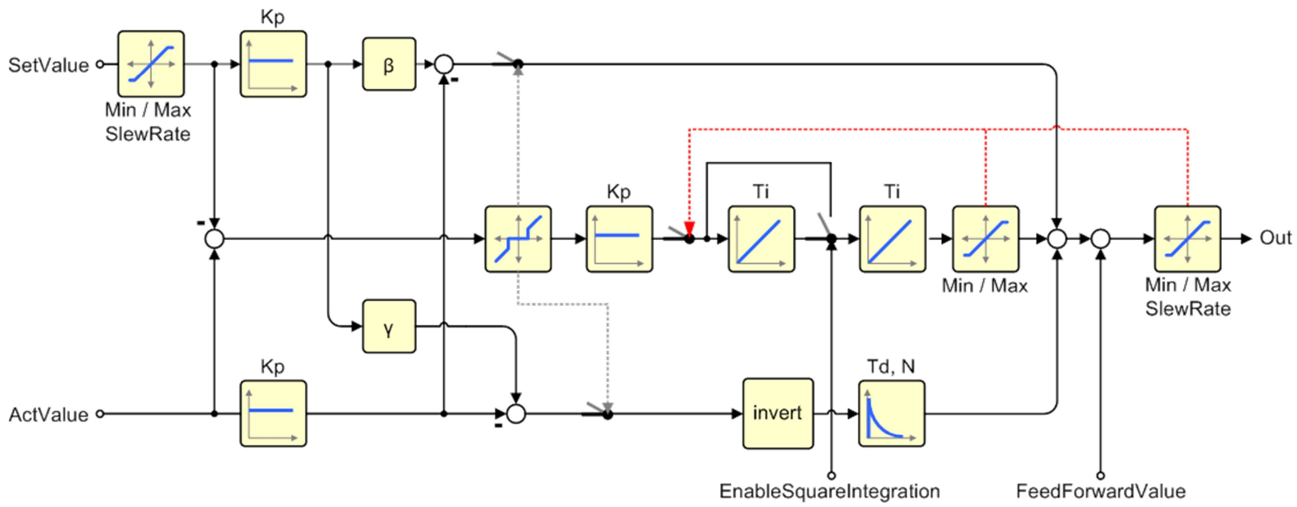


Figure 63, PID principle

The proportional gain K_p can be influenced by the configuration of the input and output ranges according to the correlation.

$$K_p = \text{Gain}\,\frac{\Delta\text{Range PID Output}}{\Delta\text{Range PID Input}}$$

Equation 47, Scaling options of PID

By doing this, the value ranges of the input and output variables can be standardized to a certain range. Standardization provides the advantage that the PID parameters can be easily adopted for the first step with controlled systems with similar dynamics.

By weighting the control error of the proportional portion of the reference variable with β and the differential portion of the reference variable with γ the following characteristics

- $\beta = \gamma = 1 \leftrightarrow$ PID controller in standard form (see chapter 8.1.6.2 Standard form, page 54).
- $\beta = \gamma = 0 \leftrightarrow$ I-PD controller, I to the control error, P and D to the control variable.
- $\beta = 1, \gamma = 0 \leftrightarrow$ PI-D controller, P and I to the control error, D to the control variable.

can be achieved by configuration. Generally, the parameters can be in the ranges $0 \leq \beta \leq 1$, $0 \leq \gamma \leq 1$, which in general terms (see Equation 46, Output- / Input behavior of ISA standard form) permit a splitting of the controller transfer function into the following 2 partial transfer functions.

$$V(s) = \frac{1 + \left(\beta T_i + \frac{T_d}{T_f}\right) s + T_i T_d \left(\gamma + \frac{\beta}{T_f}\right) s^2}{1 + \left(T_i + \frac{T_d}{T_f}\right) s + T_i T_d \left(1 + \frac{1}{T_f}\right) s^2},$$

Equation 48, PID pre-filter

$$R(s) = K_p \left(1 + \frac{1}{sT_i} + \frac{T_d\, s}{\frac{T_d}{T_f}\, s + 1}\right).$$

Equation 49, Transfer function of a PID

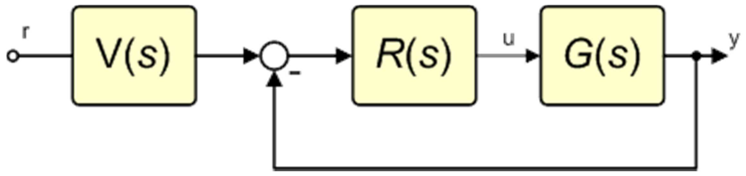


Figure 64, Principle of PID with pre-filter

See, for example, chapter 4 in (Visioli, 2006).

8.1.8 Basic information

- Standardization option of the manipulated variable output and reference variable input
- Deadband function (also asymmetrical => separate setting option of positive and negative deviation is supported).
- The controller can be configured as a PID or PI²D controller.
  The windup limits, the integral action time, the initial value specification and the bumpless switching option between the controller operation types are dependent on the respective active part (I or I²)
- P, D and I, or I² portion can be switched on/off separately
- Manipulated variable tracking option
- Cyclic tracking option of the integrator
- Initialization value of the integrator
- "Anti-windup" (windup limit I-portion == limit I²-portion), in addition the integrator can be capped above with "holdImax" and below with "holdImin".
- Once the manipulated variable output has reached an upper or lower absolute limit, a corresponding output is set (ymax_active, ymin_active). These outputs can then be used for the direction-dependent restriction of the integrator.

8.1.9 Controller operation types

The following operation types are built directly into the logic of the function block. The change from each operation type back to automatic operation can also be carried out bumplessly.

- Off
- Manual
- Open
- Closed
- Automatic
- Stop
- Autotune

8.1.10 Dead band

If the control error is in the dead band -NegDeadBand ≥ ControlError ≥ PosDeadBand, it will be set to 0 internally. The integral part is stopped and the proportional and derivative parts are set to = 0.

8.1.11 Windup

As soon as the controller output Out reaches the lower (MinOut) or upper (MaxOut) limit or the integral part reaches the limit (MinIntegrationPartValue, MaxIntegrationPartValue), the integral part will be frozen to the last value of IntegrationPart.

8.1.11.1 Guidelines for selecting the controller structure

The solution of the control task requires the selection of the controller structure as the first, very important, step. The controller structure selection depends on the controlled system characteristics, the type of reference variable signal and disturbance variables expected and on the quality requirements of the control loop. Below, basic guidelines for selecting the controller structure are provided.

If step-like reference or disturbance signals occur and are intended to prevent remaining control deviations, the open loop must have an integral behavior. In systems without inherent regulation, a P-controller is sufficient. In systems with inherent regulation, the controller must have an I-portion in order to prevent control deviations.
If an I-controller is used, it can be assumed that the manipulated variable will be changed very slowly. An I-controller cannot react quickly to control deviations.
With P- or D-portions, the transfer behavior of a control loop can be accelerated. Please note that if the differential time constants (Td) and Td*Kp factors are too high, the control loop tends to oscillate or becomes unstable. If rapidly changing disturbances such as measurement noises are expected to affect the control loop, the D-portion must be used with caution. This is because rapidly changing disturbances are amplified by the differentiation.

The following correlation can be used as a rough guideline for selecting the controller structure. The table is for guidance only as, in individual cases, there is a possibility that another correlation presents a better controller structure for particular reasons.

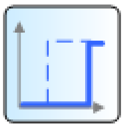
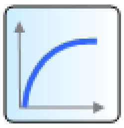
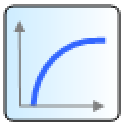
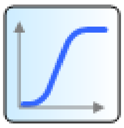
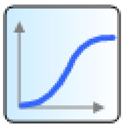
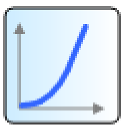
Plant	**Controller structure**				
	I	**P**	**PI**	**PD**	**PID**
Dead time 	Suitable (somewhat poorer than PI)	Unusable	Suitable, Reference + Disturbance	Unusable	Unusable
PT1 low dead time 	Not suitable	Reference	Well-suited, disturbance	Reference with dead time	Well-suited, Disturbance with dead time
PT1 with dead time 	Suitable (poorer than PI)	Unusable	Well-suited (poorer than PID)	Unusable	Reference + Disturbance
PT2 with dead time 	Suitable (but poor)	Not suitable	Disturbance	Reference	Well-suited, reference + Disturbance
PTn with dead time 	Suitable but poor	Not suitable	Suitable (poorer than PID)	Not suitable	Suitable, Reference + Disturbance
without inherent regulation 	Unusable, unstable	Reference (if no dead time)	Disturbance (if no dead time)	Well-suited, reference	Well-suited, disturbance

Table 5, Guidelines for selecting the controller structure depending on the controlled system

9 TUNING PID CONTROLLERS

Adjusting the controller can either be done in connection with a mathematical model of the process to be controlled (model-based adjustment) or without a model. However, before the tuning can be started, all measurements (including isolation amplifiers) and controlling elements of the controlled system should be checked (see control performance monitoring, from page 86). In valves, for example, friction problems or switching hysteresis can occur due to wear (lost motion). It may be that the valve packing just needs to be loosened.

9.1 Building up process understanding

Understanding the process and the aims of the control. The more you know about the process the better (operating range, starting and shut down scenarios, manipulated variables, process limits, actuator limits, measurement signal quality, delay times of the whole closed loop).
What effect does the control quality have on the overall process? In Figure 65, Tuning PID control loops, process understanding a process diagram is shown, where a reactor obtains raw materials through an upstream buffer tank. The buffer tank of the reactor serves as the supply for the reactor. The level control of the buffer tank must be set in connection with the flow control of the reactor feed. Otherwise, there is a risk that, in the event of a disturbance in the buffer feed pipe, the level controller will immediately reduce the reactor inlet immediately, so the buffer level can be retained.

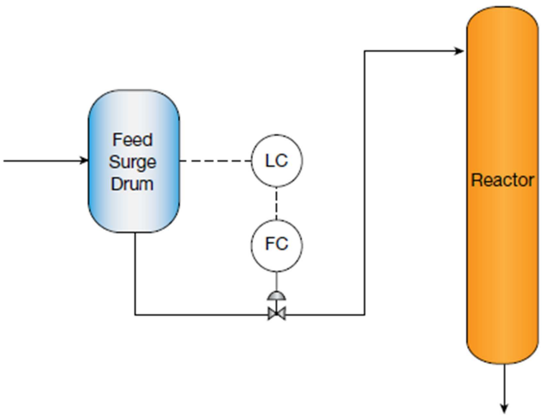


Figure 65, Tuning PID control loops, process understanding

With this understanding of the process, the control loops can be prioritized. Control loops with a strong coupling can work against each other in the event of equal prioritization.
In order for the economic objective can be optimally achieved, based on the reactor example of the feed (flow) control, a higher priority must be attributed to the reactor than to the level control of the buffer tank. In other words, the level of the buffer tank must be able to fluctuate more (e.g. 40% - 60%) (in order to equalize disturbances) than the setpoint of the reactor inflow.
The adjustment of the controller ultimately depends on the intended use. In a flow control loop that has a downstream store to smooth out oscillations, the control can be adjusted to be quicker and less attenuated as, for example, the position control of a milling machine, in which overshooting must be prevented without fail.

9.2 Selecting the tuning procedure

Selecting the tuning procedure is often dependent on the structural embedding of the control loop in the overall process. Can the process be excited by a test signal and, if yes, which? Does a process model already exist or what information is known about the controlled system? What demands are going to be placed on the control quality? Where are the measurements placed and what dynamics do they have? What characteristics are known about the actuator (linear I/O behavior)? Should the controller operate with standardized or absolute values? How many disturbance variables act on the controlled system and can these be measured? In what range is the intended operating point of the closed control loop? How does the commissioning and the shut down of the controlled system affect the overall system and vice versa?

An example of cascade control (see Cascade control page 120):

The dynamics of the internal controller should be between 5-10 times quicker than the external loop. This is hard to achieve using the Ziegler/Nichols or tune by feel method. Here, a tuning procedure is needed in which the speed of the control loop can be specified as a tuning parameter (e.g. Lambda tuning or possibly loop shaping). The setpoint of the inner loop is a function of the outer loop. For this reason, the controller will interact. In the event of equal priority (dynamics), there is also a risk of instability and oscillations, as the inner loop cannot follow the outer loop from the dynamics at all.

9.3 Carrying out the tuning

A clear distinction must be made here as to which processes the autotuning can be used in and in which processes the PID parameters can be better calculated with manual tuning. Limitations include whether the control loop can oscillate at all or how long a step can impact on a system without inherent regulation without causing damage.
How is it managed in nonlinear systems (operating points, gain scheduling, negligence)?

The operating points of the actuators should not be near the manipulated variable limits in the tests (no more margins). In this context, the behavior of the actuators must be taken into account (nonlinear valve characteristics curve).

9.3.1 Heuristic tuning rules

In this chapter, tuning rules will be discussed, which do not need a model of the controlled system. The parameters of PID controllers are determined by means of experiments with the controlled system and the control loop. The tuning procedures presented require processes which can be excited with certain test signals. A step, for example, can be used as a test signal. By using a step, it is possible to identify static friction problems affecting valves. In addition, knowledge of the process through the step response can be increased (simple test signals are easier to track). Moreover, any existing dead time is very difficult to identify. The stationary gain should be assumed for each step attempt and can be used as an indicator for nonlinearities in connection with different operating points.

The level of the step should be kept very low at first (0.25% - 1%). In this way, if necessary, the system is disturbed as little as possible. If a sufficiently good step response is not accomplished (e.g. including ramped response), the level of the step can be increased.

Manual tuning can be carried out as follows.

1. Establish or check whether the manual manipulated variable and the control variable are adopted into the APROL trend system.
2. Set the controller to MANUAL.
3. Wait until the process is in the stationary state of the manual manipulated variable.
4. During the test it should be ensured that no manipulated variable changes occur.
5. Enter desired manipulated variable change.
6. Wait until a stationary state has been reached (if necessary).
7. Turn the controller back to AUTO.
8. Analysis of the recorded data -> calculation of the controller parameters.

9.3.1.1 Stability margin method according to Ziegler and Nichols

The Ziegler and Nichols oscillation method (quarter decay ratio) can be used for systems with inherent regulation, which can be relatively quick (e.g. rotation speed control).
The stability margin method assumes that there is a technical possibility that the respective process can oscillate. If, for example, the speed controller of a cruise control (in a car) is set using an autotuning function, it must be assumed that the tuning procedure cannot be started at any random point in time. This is because an oscillation test in the gas tank / brakes will inevitably lead to problems.
During this behavior, no assumptions regarding the transfer behavior of the system are necessary. The control loop is closed using a P-controller (tuning for the closed loop). The controller gain Kp is increased until the closed loop achieves continuous oscillation after a change in the setpoint. The set controller gain in this case is called Kcrit, the cycle duration of the control variable oscillation is called Tcrit.

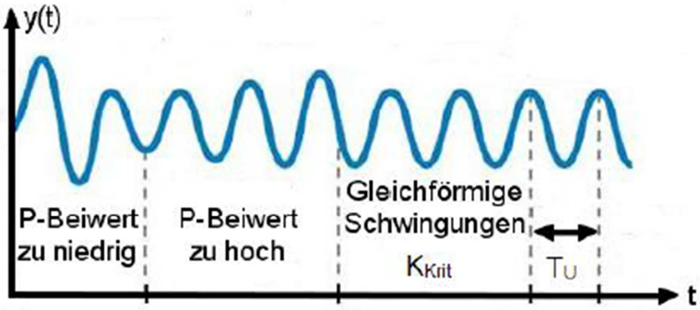


Figure 66, Stability margin method according to Ziegler/Nichols

With the critical P-portion and the critical cycle time, which emerge during the occurrence of the constant oscillation, the PID controller parameters can be determined according to the following table.

Controller structure	**Kp**	**Ti**	**Td**
P	0.5 Kcrit	-	-
PI	0.45 Kcrit	0.85 Tcrit	-
PD	0.55 Kcrit	-	0.15 Tcrit
PID	0.6 Kcrit	0.5 Tcrit	0.125 Tcrit

Table 6, Tuning rules of stability margin method according to Ziegler/Nichols

The stability margin and the inflection tangent method were predominantly designed to control disturbances. The reduction of the amplitude by ¼ each oscillation cycle can be identified as a criterion. These types of controllers can, for example, become unstable in a relationship of moderately poor attenuation and an oscillating control signal. Note that in the case of closed control loops in which overshooting is not allowed, neither the stability margin method nor the inflection tangent method are suitable. The methods should lead to a quarter wave decay behavior in a setpoint step (i.e. the decay behavior of the oscillation after a setpoint step is formed so that the amplitude decreases by ¼ in each cycle).

Important:

The cycle duration only changes slightly during the transition from a stable to an unstable oscillation. For this reason, the cycle duration can be concluded very quickly and this can be used to calculate the integral action time and the differentiation time constant. The gain can then still be adapted.

9.3.1.2 Inflection tangent method according to Ziegler and Nichols

The tuning procedure can be used in systems whose step response ends in a stationary state and therefore no overshooting occurs (for systems that cannot oscillate (e.g.: PT2 with attenuation <1)). An S-shaped output signal is a prerequisite.

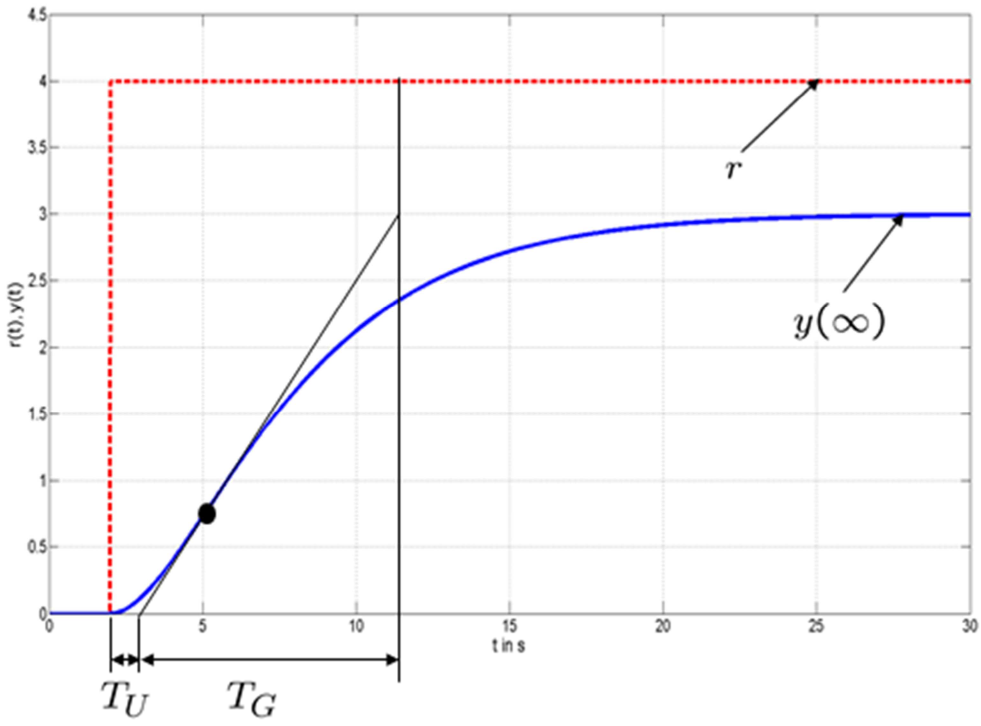


Figure 67, Inflection tangent method according to Ziegler/Nichols

The inflection tangents presents the process parameters of the delay time T_U and the equalization time T_G. The time constants of the controlled system should be determined from the process parameters using the inflection tangent process.
The dimensioning controller results due to the determined values T_U and T_G according to the following table.
The gain Ks can be calculated as follows

$$Ks = \frac{\Delta x}{\Delta y}.$$

Equation 50, System gain after step response

Controller type	Kp	Ti	Td
P	$1 / Ks \cdot T_G / T_U$		
PI	$0.9 / Ks \cdot T_G / T_U$	$3.33 \cdot T_U$	
PID	$1.2 / Ks \cdot T_G / T_U$	$2 \cdot T_U$	$0.5 \cdot T_U$

Table 7, Tuning rules of inflection tangent method according to Ziegler/Nichols

Important:

The following basic guidelines can be used to estimate how well the control will work.

Controllability of PTn systems	
Can be well controlled	$T_U / T_G < 1/10$
Controllable	$T_U / T_G \approx 1/6$
Difficult to control	$T_U / T_G > 1/3$

Table 8, Controllability of systems depending on delay time Tu and equalization time Tg

In principle, the controlled systems become more difficult to control with an increasing ratio of T_U / T_G. The advantage of this method is that very little information about the system is required. However, the reference variables and the manipulated variables are treated equally. In addition, in this procedure, overshooting must be taken into account.

9.3.1.3 T-total control

The T-total method is suitable for systems that are not able to oscillate with lowpass behaviors. The procedure is used in autotune procedures using a step test in APROL
The total time constant is defined as follows for a general transfer function with time constants in the numerator (T1...Tm), time constants in the denominator (Td1...Tdn) and a dead time of Tt

$$T_{\Sigma} = Td1 + Td2 + \cdots + Tdn - T1 - T2 - \ldots - Tm + Tt.$$

Equation 51, Total time constant, T-total tuning rule

If a model is available.

The total time constant can be calculated from the step response using the correlation A=Ks*TΣ.
The system must be steadied before the step test, i.e. the process is in a stationary state. The step test starts from a given value (which should be close to the current stationary value of the manipulated variable). The total time constant can also be determined by comparing the surfaces A1 and A2. When these are the same size, the total time constant has been found.

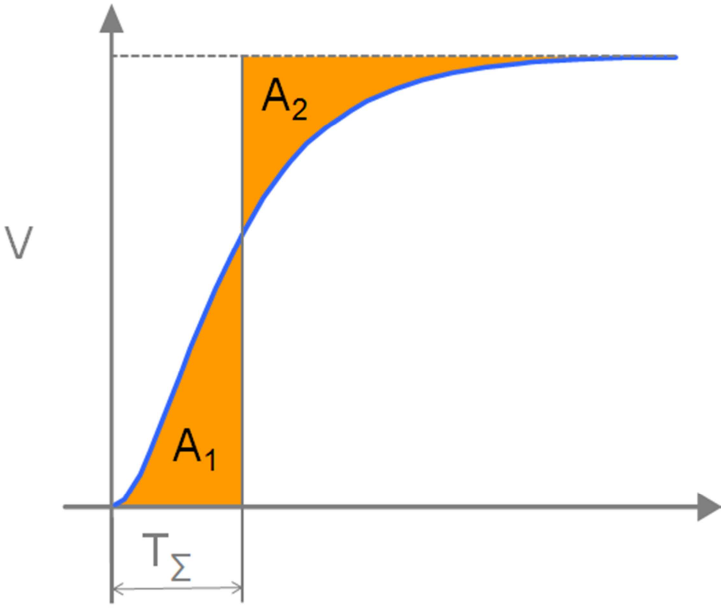


Figure 68, Total time constant principle

The gain V arises as a result of V= $\frac{Regelgrößenänderung}{Stellgrößenänderung}$.

	Controller type	**Kp**	**Ti**	**Tv**
Normal tuning	P	1/V	-	-
	PD	1/V	-	0.33 TΣ
	PI	0.5/V	0.5 TΣ	-
	PID	1/V	0.66 TΣ	0.167 TΣ
Quick tuning	PI	1/V	0.7 TΣ	-
	PID	2/V	0.8 TΣ	0.194 TΣ

Table 9, T-total tuning rule

9.3.1.4 Lambda tuning

The tuning parameter λ identifies the desired time constant of the controlled system depending on a reference variable step. The user is given the option to instantly adjust the controller to a less aggressive setting, depending on λ.
The speed of the closed loop can be set using λ, depending on the process requirements and physical limits. As a conservative starting point, triple the time constant T or the dead time Tt (depending of dominance) can be used for Lambda (λ = 3T(t)). With the tuning procedure, both systems with inherent regulation and systems without inherent regulation can be adjusted.

> **Important:**
>
> The procedure provides PID parameters which have a very calm, non-oscillating behavior in setpoint steps.

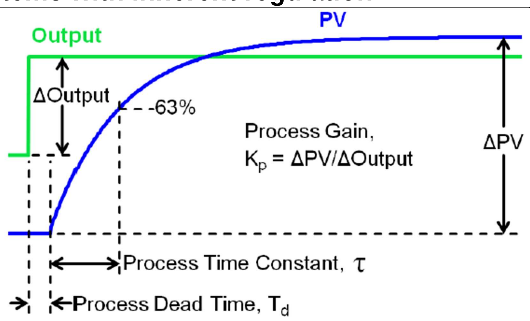
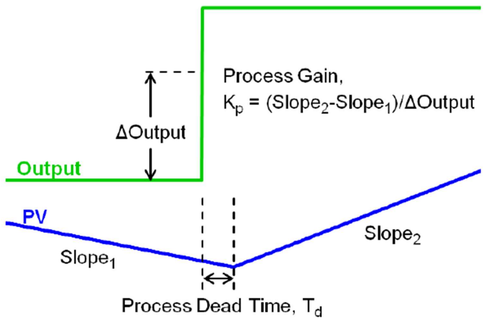
Systems with inherent regulation	Systems without inherent regulation
 Figure 69, Step response of system with inherent regulation	 Figure 70, Step response of system without inherent regulation
$Kp = \frac{\tau}{Ks(\lambda + Tt)}$ $Ti = \tau$ $Td = 0$ Equation 52, Lambda tuning of system with inherent regulation	$Kp = \frac{2\lambda + Tt}{Ks(\lambda + Tt)^2}$ $Ti = 2\lambda + Tt$ $Td = 0$ Equation 53, Lambda tuning of system without inherent regulation

Table 10, Lambda tuning rules

9.3.1.5 Chien Hrones Reswick

This method is considered a development of the inflection tangent method by Ziegler/Nichols. What makes this procedure interesting is that the control parameters can either be set for reference or manipulated variable behaviors.

For PID parameter calculation, the delay time, equalization time and system gain are required. The controlled system should be at least second-order. The manipulated variable step should be at least big enough to allow the control variable progression to be analyzed with sufficient precision. In addition, the control variable step must be in the setpoint range which is expected at a later stage.

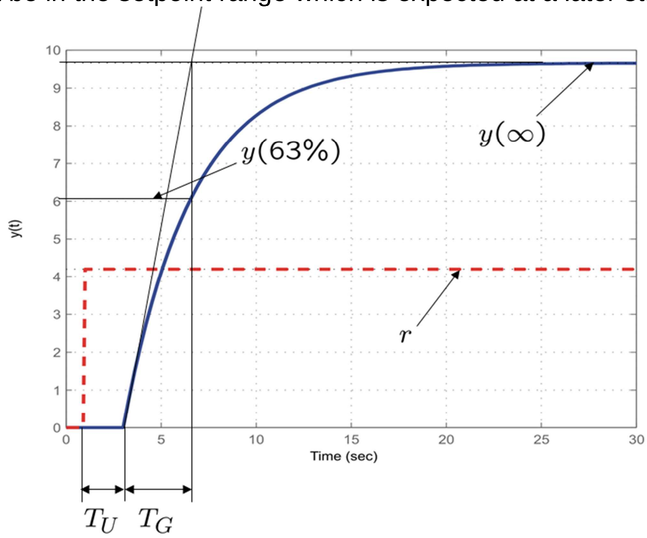


Figure 71, Step response PT1

The delay time Tu and the equalization time Tg can be calculated by determining the inflection tangent. The gain V arises as a result of V= $\frac{Regelgrößenänderung}{Stellgrößenänderung} = \frac{y(\infty)}{r}$.

The controller parameters for the acyclic case are as follows:

Controller structure	Reference	Disturbance
P	Kp = 0.3 Tg/(Tu*V)	Kp = 0.3 Tg/(Tu* V)
PI	Kp = 0.35 Tg/(Tu* V) Tn = 1.2 Tg	Kp = 0.6 Tg/(Tu* V) Tn = 4 Tu
PID	Kp = 0.6 Tg/(Tu* V) Tn = Tg Td = 0.5 Tu	Kp = 0.95 Tg/(Tu* V) Tn = 2.4 Tu Td = 0.42 Tu

Table 11, Chien Hrones Reswick tuning rules, acyclic case

The controller parameters for the acyclic case with 20% maximum overshoot are as follows:

Controller structure	Reference	Disturbance
P	Kp = 0.7 Tg/(Tu* V)	Kp = 0.7 Tg/(Tu* V)
PI	Kp = 0.6 Tg/(Tu* V) Tn = Tg	Kp = 0.7 Tg/(Tu* V) Tn = 2.3 Tu
PID	Kp = 0.95 Tg/(Tu* V) Tn = 1.35 Tg Td = 0.47 Tu	Kp = 1.2 Tg/(Tu* V) Tn = 2 Tu Td = 0.42 Tu

Table 12, Chien Hrones Reswick tuning rules, cyclic case

9.3.1.6 Trial on error tuning

Using this approach, consecutive favorable settings for the P-, I- and D-portions can be found. In this way, a typical setpoint, for example, can be gradually preset from a previously-defined initial state and the controller parameters are adjusted as a result.

The P-portion of the controller is set first (relatively low). An initial value is selected for this and then controlled with a setpoint step, which moves around an operating point. On the assumption that the initial value of the Kp was very low, the Kp can be increased until the control variable oscillates two or three cycles at the most, until a stable final value is reached. A remaining control deviation is ok in this case.

In the second step, the D-portion is activated but only if there is no dead time and the signals are noise-free. Starting from a very small lead time Tv, the setpoint is brought back into the operating point range and the lead time then increases gradually. The lead time Tv is favorably set if the control variable reaches the setpoint with as small an oscillation as possible. If the manipulated variable reaches 0 or 0% once or several times over the course of approaching the setpoint, the selected Td is too large.

In the last step the I-portion is activated. The integral action time Tn is set as the first approximation to Tn=4*Tv.

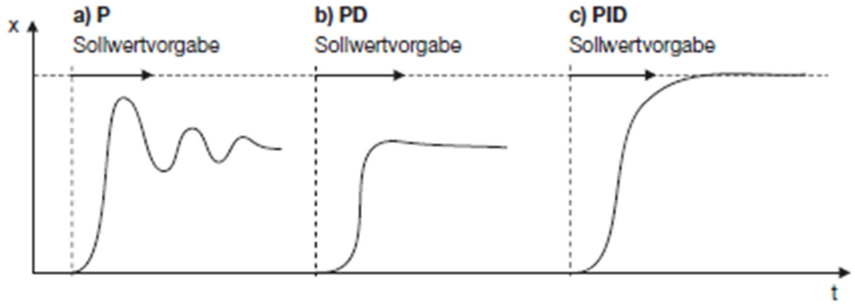


Figure 72, Trial on error tuning

Please note: not all parts of the PID controller can be activated in all systems. If, for example, it is established that a stable control cannot be achieved through the P-portion, a pure I-controller can be tested for use.
If the activation of the D-portion (PD controller) leads to instabilities, a PI controller should be reverted to.
In industrial practice, in many cases tuning is carried out on the controller by testing the PID parameters.

Based on the following step response of a closed loop, the following characteristics can be concluded.

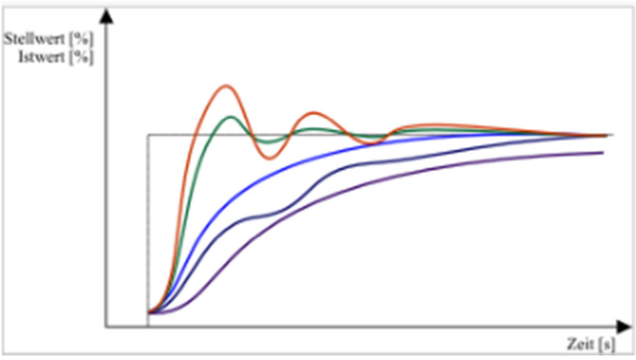


Figure 73, PID configuration after the step response of the closed loop

1.) Purple: The control variable is slowly approaching the setpoint. The proportional component could be increased in this case. If an improvement occurs, the integral action time can be reduced. The process can be repeated until a satisfactory result has been achieved.
2.) Dark blue: The control variable is approaching the setpoint slowly and with a slight oscillation. The proportional portion can also be increased in this case. If an improvement in the control quality

occurs, the lead time can then be reduced. The process can be repeated until the dynamics of the closed loop meets expectations (never forget the physical limits).

3.) Light blue: The control variable is approaching the setpoint without overshooting. If the closed loop still has a higher dynamic (overshooting is acceptable), proceed as described in point 1.

4.) Green: The control variable is approaching the setpoint with slightly attenuated overshooting. This tuning is particularly suitable for controlling disturbances (see oscillation method according to Ziegler/Nichols). If overshooting should be reduced, you can proceed as in point 5.

5.) Red: The control variable is oscillating greatly beyond the setpoint and is only approaching the setpoint weakly attenuated. In this case, the proportional portion should be reduced. In addition, the integral action time of the controller can be increased. The process can be repeated until a good transient response is reached.

9.3.2 Parameter optimization

The task of controller design can also be described and tackled as an optimization problem. If the control deviation needs to be as small as possible over time, but greater deviations are more strongly valued than small ones, this requirement can be formulated using the following quality function ("integral square error" criterion)

$$J = \int_0^\infty e^2(t)\mathrm{d}t$$

Equation 54, Quality function PID parameter optimization

. The quality function J describes the quadratic control area, which is dependent on the predefined controlled system, the control loop structure and the control parameters used. The goal is to minimize the quality function (see H" optimization). To do this, the type of input variable must be provided (constant, impulse-like, step-like). The optimization procedure requires a controlled system that is known exactly. However, the structure of the controller must not be predefined.

9.3.3 Analytical tuning

This procedure requires a state space model or a transfer function of the controlled system. By choosing suitable controller parameters, the eigenvalue of the system matrix or the poles of the transfer function of the closed loop should be specifically influenced. Output variables of linear systems depend on the eigenvalues of the system matrix and the poles of the transfer function. By choosing suitable poles, the dynamic characteristics of the control loop can be determined. Stability is achieved by placing the poles in the left complex half plane. If the pole is placed close to the real axis, the oscillations of the system are strongly attenuated. An analytical calculation of the controller tuning could be made as follows.

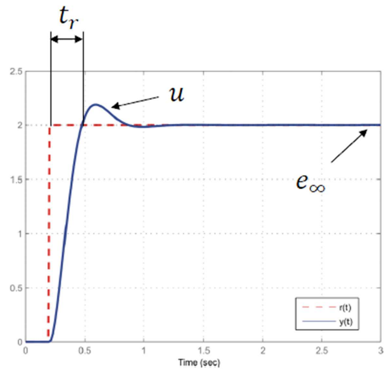
1. Determination / estimation of the structure of the controlled system (PT1, PT2, etc., IT1, IT2, etc., dead time
2. Identification of the parameters for the controlled system
3. Select a suitable controller structure
4. Algebraic calculation of the reference transfer function
5. Designing the controller parameters by specifying the desired pole distribution

With skilled selection of the controller structure, a coefficient comparison between the characteristic polynomial of a controller/system combination and the characteristic polynomial that puts the poles in place leads to pole placing and zeroing of the controller.

9.3.4 Loopshaping

This method is particularly suitable for frequently occurring problems of the same type. For a given process model, a type of controller is designed which meets the specifications of the closed control loop. In addition, the quality requirements of the closed loop translates into requirements for the open loop and, as a consequence, a controller is designed which meets the requirements of the open loop. The open loop is formed again (shaped) by this process. For this reason, the term "loop shaping" was formed. This method is also called a frequency response characteristics process.
The "loop shaping" process is carried out in the frequency range and refers to the following parameters

- Rise time,
- Overshoot,
- Remaining control deviation,



Equation 55, Loopshaping parameters of the closed loop

of the closed loop.

9.4 Adjusting the PID controller parameters

If a closed control loop is not tuned with sufficient precision after a tuning test, (there are many possibilities which allow a control loop to be poorly controlled, see control performance monitoring) the following adjustments can be made.

9.4.1 Stabilization of the PID control loops

A control loop tends towards instability if the Kp is too great or the Tn and Tv are too low (P-portion too great, I-portion integrated too quickly or the attenuation is too weak). In order to achieve more stability, decreasing the Kp can be started. If this does not stabilize the control variable, the Tn and Tv can be increased. As a practical rule of thumb, it can be checked, for example, whether Tv broadly corresponds to Tn/4. The correlation should also be constantly maintained when increasing the integral action time and lead time.

In the event of a very high number of applications, the PID structure has the best behavior. However, there are cases in which different portions of the PID structure must be deactivated.
In controlled systems with unsettled control variables, a D-portion, for example, may lead to instabilities. In this case, the P-portion also reinforces the disturbance of the system. Therefore, this could also be deactivated if necessary.
If the relationship of the delay time Tu to the equalization time Tg is relatively large (indicating a system that is difficult to control, see Table 8, Controllability of systems depending on delay time Tu and equalization time Tg), disconnecting the P and D-portion may be necessary in order to avoid instabilities.
Generally, a favorable controller structure cannot be universally provided depending on the system to be controlled. Nevertheless, a short overview of the system-dependent controller structures is provided here.

Discipline	Controller structure	Comment
Temperature	PID	The controlled systems always have inherent regulation. The equalization time Tg is very often greater than the delay time Tu, the PID structure is the most suitable
Pressure	I	This type of system usually has a very low ratio between equalization time/delay time (Tg/Tu<3). These types of systems can be dealt with in a similar way to systems influenced by dead time. In the event of unsettled actual values, the I-structure is the most stable solution.
pH value	PID flow path / sink P / PD (generally also gain scheduling)	If the PH value in a pipe system is controlled by flow path, a PID controller can be used. In the event of controls in

		a sink, the I-portion is not used, as it may lead to oscillations.
Speed	PI	Assuming that harmonics arise in the measurement signal due to the rotations or the control variable is very unsettled, the D-portion is not used.
Flow rate	I	In this case, the ratio between equalization time/delay time (Tg/Tu<3) can be very low. For this reason, the D- and P-portions are not used.
Level	PID	The control of systems without inherent regulation can be achieved with the PID structure. The I-portion, however, should not be too strong in this case (oscillation tendency). Note that due to the behavior of the system, these systems must not be used with a pure I-controller (phase margin, instabilities).
Delivery of bulk goods	I	Systems affected by dead time in combination with pure P-structures lead to oscillations. In many cases, the I-portion provides a good result.

Table 13, Selecting the controller structure according to the application

In the event of key processes where each improvement in the control has an impact on the profitability of a company (quicker product switching rates, higher flow rate, more efficient energy consumption), Advanced Process Control methods can further improve the control quality compared to PID. Using system models, model-based control methods (Smith predictor or MPC) can be used even more efficiently.

10 PID AUTOTUNE

The PID controller has a separate AUTOTUNE operating mode, which enables the tuning function block to have direct process access to the PID without additional wiring.

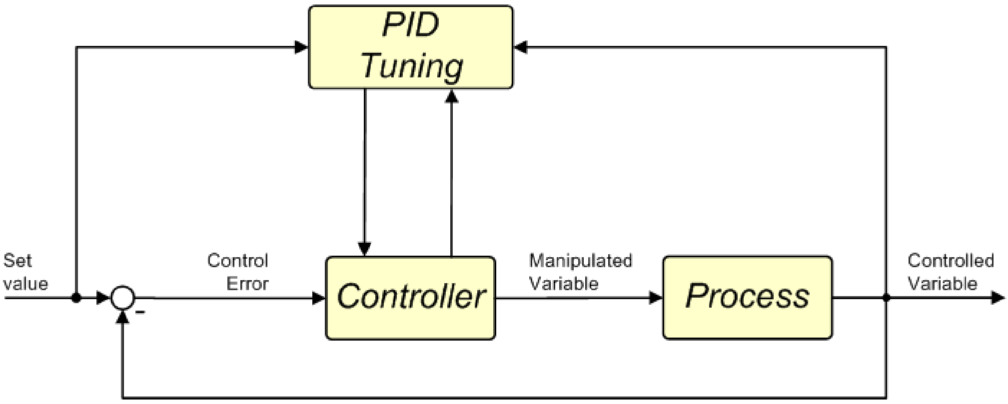


Figure 74, PID autotune principle

The autotune function block can optionally determine the required controlled system information to calculate the PID parameters based on an oscillation test or a step test.

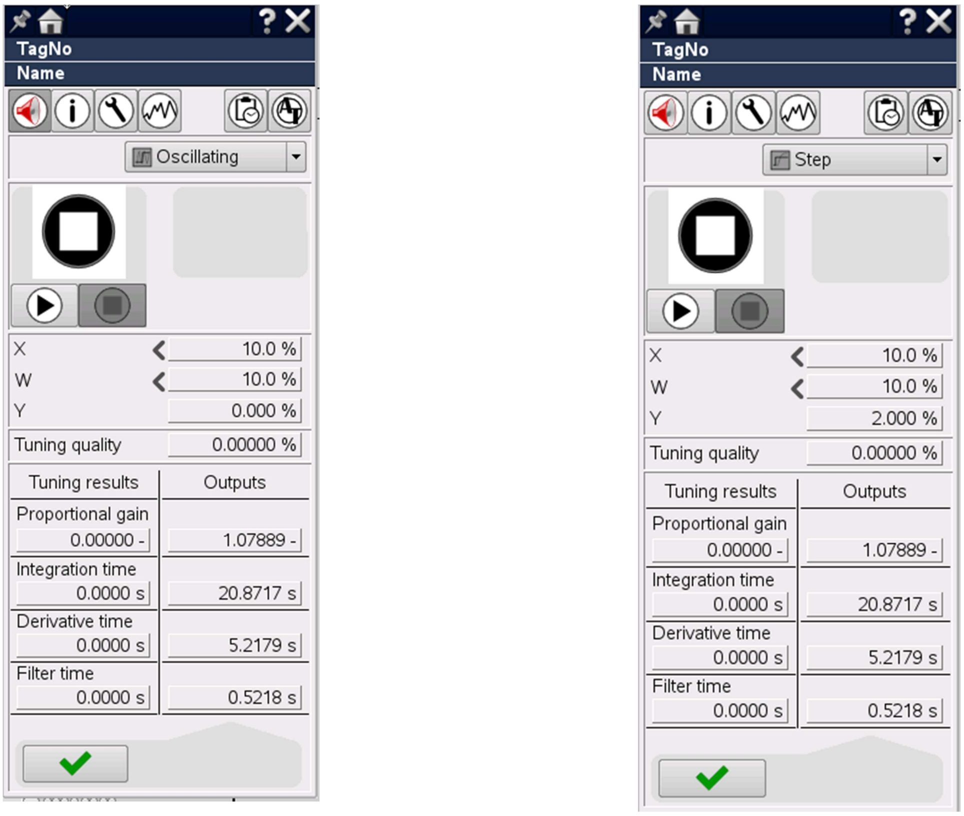


Figure 75, PID autotune, APROL operation

10.1 Oscillation test

The oscillation tuning procedure indirectly draws on the Ziegler/Nichols method and can be used for the closed loop. The system is stimulated to oscillate at the setpoint by cyclic manipulated variable oscillations.
The oscillation is an identification procedure in the closed control loop. In this procedure, the system is controlled so that the process variable oscillates periodically around the predefined set value. The cycle duration and amplitude ratio of the manipulated variable and the controlled variable are used to determine the control parameters of a PID disturbance variable controller in accordance with Ziegler/Nichols.
The goal of the oscillation test is to achieve the most symmetrical harmonic oscillation of the controlled variable possible around the set value. This goal can be achieved relatively easily using multiple consecutive oscillation tests, each with the best optimized parameters.
This function block can accurately determine PID control parameters for systems with approximated PTn-Tt characteristics (PT1-Tt, PT2, PT2-Tt, PT3, etc.) without requiring an extensive amount of work or special expert knowledge.
To achieve the highest degree of accuracy in the results, it is important to make sure that the sampling time (cycle time of the function block call) of the tuning function block is not larger than 1/10 of the smallest time constant in the entire controlled system (e.g. dead time, time constants of PT1 or PT2). Otherwise, the calculated control parameters may be imprecise.

10.1.1 Oscillation test 1

The oscillation stimulation occurs under the following conditions

- Control variable < setpoint => Controller output at maximum,
- Control variable > setpoint => Controller output at minimum.

This 2-step control results in a periodic oscillation of the controlled variable around the set value. The control parameters for disturbance variable controllers can be calculated according to Ziegler/Nichols from the amplitude ratio of the manipulated variable and the controlled variable as well as the duration of the oscillation.
Tests on ideal model systems (PT1-Tt, PT2, PT2-Tt) have shown that good controller parameters are calculated starting from *Quality* = 50%.

10.1.2 Oscillation test 2

The second oscillation test (stimulation of the system with four oscillations) is only performed if the first test does not deliver sufficient quality of the controller parameters. In this oscillation test, the threshold values of the 2-step controller are optimized automatically.
After the tuning has finished, the calculated PID parameters are shown in the graphical user interface of the autotune function block. With further confirmation, the parameters can be forwarded to the external parameter inputs of the PID controller. The new configuration can now be tested simply by switching between internal and external.

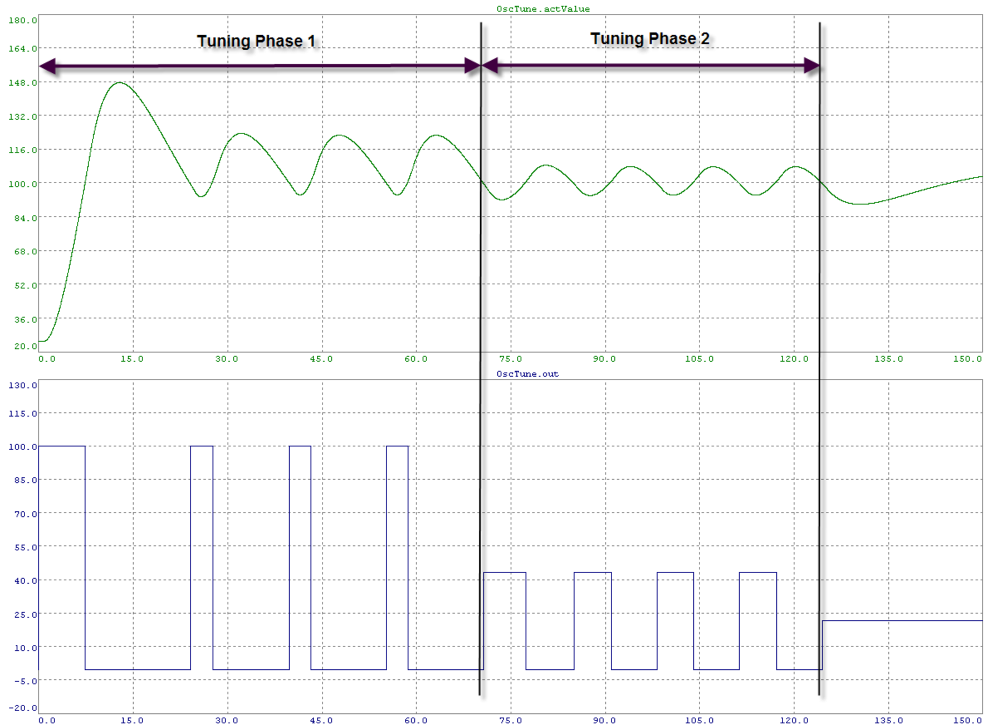


Figure 76, PID autotune, oscillation test

10.2 Step test

For processes that are not able to oscillate, a step test can be carried out assisted by the autotuning procedure.
During the step test, the PID parameters are calculated according to the T-total tuning rule (see page 66). This procedure is only suitable for systems that cannot oscillate. Before the step test can begin, the system must be in a stationary state. The initial value of the manipulated variable should be selected so that the current stationary state of the control variable can be approximately maintained.

Important:

This function block can accurately determine PID control parameters for non-oscillating systems with approximated PTn-Tt characteristics (PT1-Tt, PT2, PT2-Tt, PT3, etc.) without requiring an extensive amount of work or special expert knowledge.

To do this, the step response of the controlled system is recorded in the open control loop in order to calculate the control parameters. When doing this, an initial value of the manipulated variable it taken as a basis, which the control variable should keep approximately in the operating point. The system must be in a stationary state before starting the step test.

The manipulated variable range allowed during the tuning can be predetermined by configuring the limit values. The maximum duration of the self-optimization process is also defined. If one of these limits is violated during tuning, the tuning process is aborted. If it is not possible to calculate any parameters from the recorded step response, then an error is output.

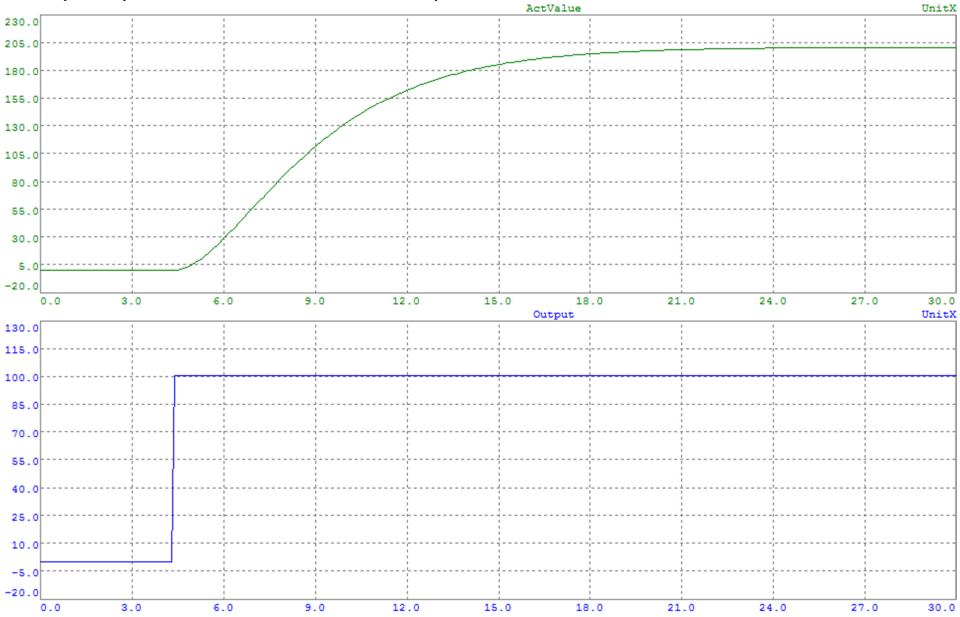


Figure 77, PID autotune, step test

11 ADDITIONAL FUNCTIONS OF THE PID CONTROLLER

The control quality of PID controllers can be improved in many cases by a very simple method.

11.1 Changing the PID controller structure

Depending on tuning, ordinary PID controllers can be designed so that either the reference variable is controlled very well or the controller has very good characteristics with regard to handling disturbances. For example, disturbance variables are controlled using a "high gain" controller, which tends to oscillate in the event of setpoint steps. This behavior can be counteracted by a PID controller with 2 degrees of freedom.

Using the different constellation options between β and γ, controllers with the following characteristics can be configured.

- $\beta = \gamma = 1$: Conventional PID controller in standard form
- $\beta = \gamma = 0$: (I-PD) I-portion acts on the control error, P-, D--portion act on the control variable.
- $\beta = 1$, $\gamma = 0$: (PI-D) P-, I-portion on control error, D-portion on control variable.
- $\beta = 0$, $\gamma = 1$: (ID-P) Reduction in overshooting during setpoint steps. Reference variable steps are only controlled through the I-portion.
- $0 \le \beta \le 1$, $0 \le \gamma \le 1$: Controller with 2 degrees of freedom.

See, for example, chapter 4 in (Visioli, 2006).

11.1.1 PI -> e / D -> x controller

When using a PID controller to control systems with dead time s portions or to control systems with high Tu/Tg values, differential impulses can weaken the control quality. If, for example, a reference variable step occurs, the D-portion (starting from a stationary state) is given a specific value. In the next sampling step (control error does not change due to the high dead time compared to the sampling time) the D-portion is set back to 0 (stays there until the control error changes). The reaction of the D-portion in this case will cause an unnecessary excitation of the process.

Important:

By using the PI-D controller (prerequisite of a controlled system affected by dead time) the D-portion will no longer excite the process by impulse-like D-portions ("derivative spikes") during reference variable steps.

In contrast, the transfer behavior with regard to disturbances is not influenced at all. The behavior could, of course, be improved by filtering; although in systems with high Tu/Tg values, the PI-D variant must be favored.

11.1.2 I -> e / PD -> x controller

In this type of PID control, the control variable is passed to both the P-portionand the D-portion instead of the control error. The I-component is, therefore, the only variable which reacts to step-like changes in the reference variable. This type of control is usually (possibly even incorrectly) used for controls in which the manipulated variable should be changed very slowly. However, this behavior can be also be achieved with a standard PID controller in combination with the relevant tuning.

With regard to the setpoint steps, the I-PD algorithm is almost equal in comparison to the PI-D algorithm, with low losses in the control quality (about 10%).

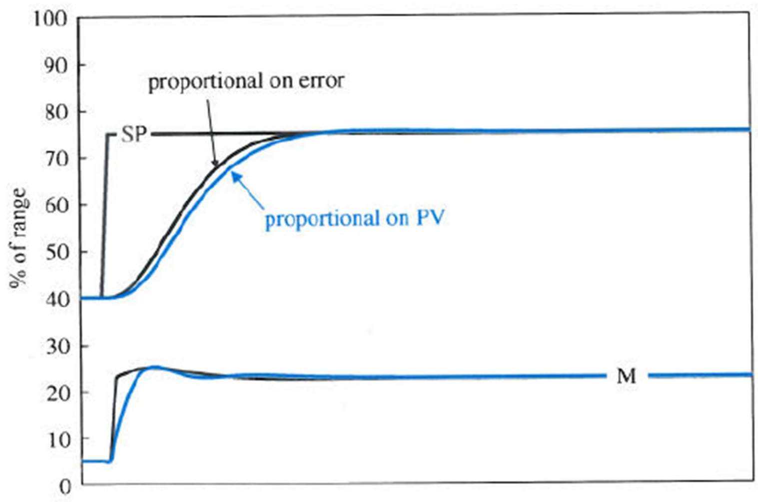


Figure 78, Step responses I-PD with PI-D structure

The missing proportional "kick" was compensated by a higher control gain in this case. As a result, the I-portion leads to a higher manipulated variable dynamic. In contrast to the PI-D, the I-PD does not provide any advantages for manipulated variables if both controllers have been tuned to compensate for manipulated variable steps.

In the following test, the control quality is accessed via the ITAE ("integral over time of absolute error") to assess the control quality. The higher the ITAE, the less control there is over the control error. Both controllers (I-PD and PI-D) have been set to handle setpoint steps. The next figure shows the development of the ITAE if an I-PD controller is changed to a PI-D controller.

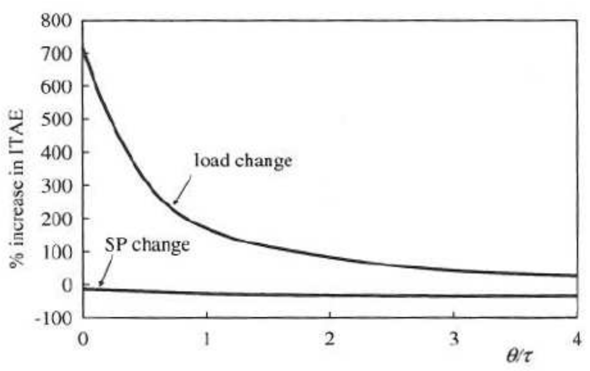


Figure 79, Comparison of PI-D with I-PD structure

The control quality compared to the setpoint steps is somewhat poorer with the I-PD than with the PI-D. In contrast, the control quality would deteriorate considerably for manipulated variables and with the use of a PI-D depending on the time constant.

Important:

In processes which have considerably fewer reference variable steps than manipulated variable changes, the I-PD variant is clearly the better solution.

For example, follow-up controllers in the cascade often have a low Tu/Tg ratio. The control quality can be increased by up to 600% with regard to the disturbance variables, with the disadvantage that a setpoint change causes a poorer control quality by 10%.

11.2 Operating point connection in the control loop

A reduction in the remaining control deviation can be achieved by an operating point shift. If a P-controller is used to control a particular setpoint (operating point), the energy that is needed to operate the system in the desired operating point can be converted into a manipulated variable offset y0 (operating point).

$$y = y0 + K_p \cdot e$$

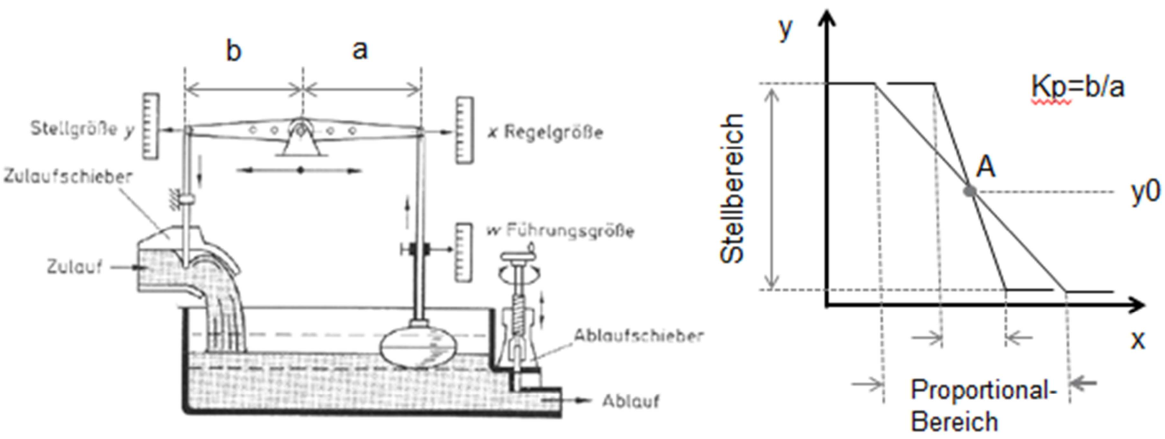


Figure 80, Operating point connection in mechanical water level controller

Important:

If no control deviation and no disturbance occurs and if y0 is optimally selected, the operating point of the control variable is achieved and the controller itself would not have to intervene at all in this case (for the nominal operating point).

In the case of mechanical control, the operating point can be adjusted by adjusting the mechanical bond between the vertical "float rod" and the horizontal transfer element.

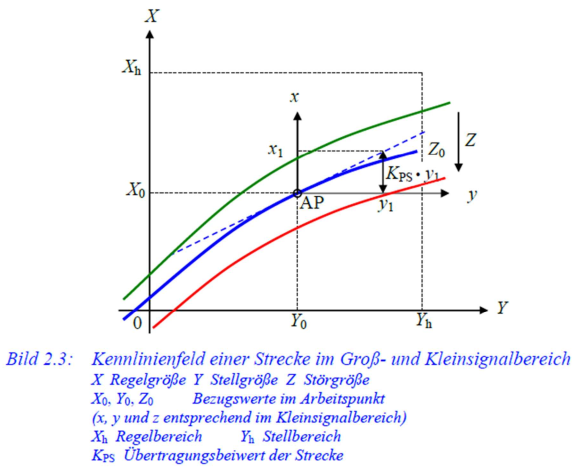


Figure 81, Operating point connection of a nonlinear system

Situation: From above, the tank for the mechanical level control is no longer cylindrical but spherical The relationship between the manipulated variable y and the control variable x is no longer linear (see Figure 81, Operating point connection of a nonlinear system). The controller can be tuned by linearization to the operating point. However, in this case it must be assumed that the deflection by the operating point is sufficiently small.

11.3 Reset windup

In reality, no single control loop is linear. The existing nonlinearities can be disregarded in many cases though, as the control loop can be controlled robustly, so that small deviations between the model and reality can be dealt with. A very significant nonlinearity is the limitation of the controlling element, which should always be taken into account when designing the control loops.

11.3.1 Manipulated variable limitations

Each controlling element has a lower and an upper limit (valve closed / open, motor at standstill / maximum revolutions). The manipulated variable can only range in this predefined interval. In the simplest case, this relationship is linear.

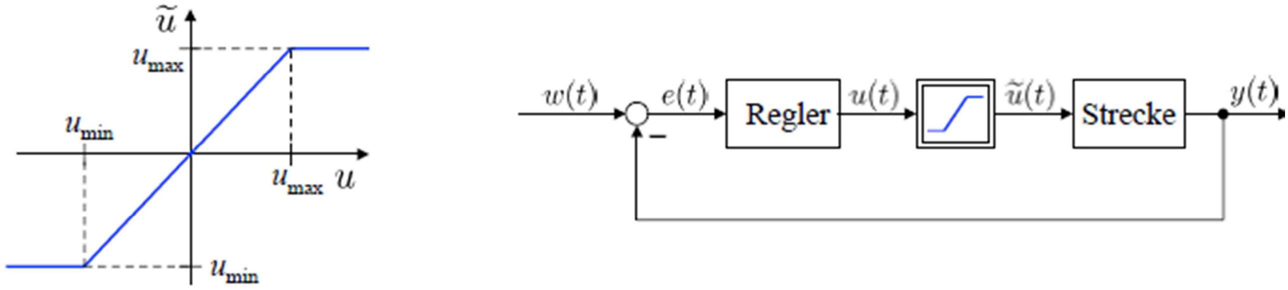


Figure 82, Manipulated variable limit

If the controller's manipulated variable is outside the range of the controlling element, a deviation between the desired manipulated variable and the actual manipulated variable produced by the controlling element occurs. As a result, the effect of the controller is not as strong as anticipated in the controller design. The controller design can be counteracted through the following options.

- The controller can be designed to be so inactive (gentle) that the desired manipulated variable never leaves the actual adjusting range. This leads to very slow controls with low loop gain, which is unacceptable.
- The controller is tuned according to a standard procedure and the manipulated variable limits are not considered. For this reason, windup effects may occur.
- The manipulated variable limits may be directly taken into account in the controller design (e.g. optimization procedure with auxiliary conditions). These types of methods can be implemented by using an MPC (model-based predictive controller), for example.
- Enhanced loop shaping methods “Schneider” offers the possibility to consider manipulated variable limits directly.

It is important that the manipulated variable limits are taken into account when configuring the controller. An actuating drive has a specific actuating speed. This actuating speed cannot be exceeded even with the best tuning. The control loop will only work as well as is maximally allowed by the process technology or machine construction (physical limits). If controlling elements make a noticeable impact on the dynamics of the overall system, this must be taken into account in the tuning. In addition, the limits can be embedded in the controller parameters, so that, for example, a change in the operating mode (automatic -> manual) leads to undesired manipulated variable changes. Most industrial controllers have the option to limit absolute manipulated variables and their maximum allowed rate of increase.

11.3.2 Anti-windup

Saturation of an integrator ("integral windup") occurs e.g. in a controller with an I-portion if a deviation arises between the desired manipulated variable (which is provided by the controller) and the actual manipulated variable from the controlling element. As a smaller control signal acts on the system than is desired by the controller, the control deviation slowly reduces, as assumed for the linear model without limits. The I-portion and therefore also the desired manipulated variable continue to increase, as the control deviation is integrated. Only when the control deviation disappears does the I-portion stop. During the period in which the manipulated variable adheres to its upper limit, the desired manipulated variable has continued to go upwards due to the increasing I-portion. If the control deviation becomes negative, the I-portion needs a specific period of time until it decreases to the level of the desired manipulated variable. Due to the windup effect, the desired manipulated variable stays above the maximum manipulated variable for a long time. The manipulated variable can no longer intervene as much and the control quality worsens. The stricter the limit, the poorer the control quality becomes.

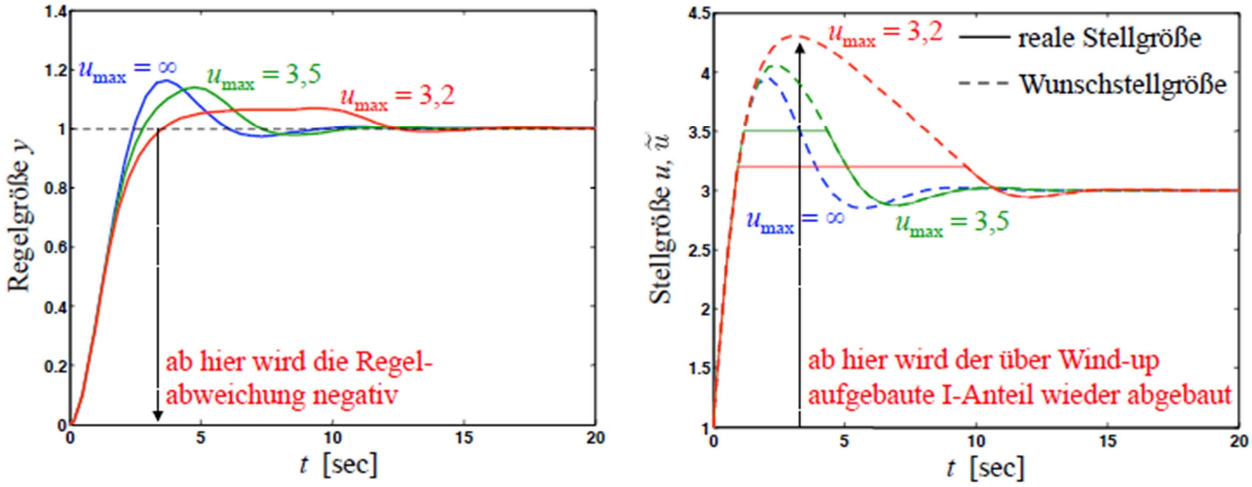


Figure 83, Anti-windup

If saturation of manipulated variables occurs, the control quality can be considerably improved using an anti-windup method. Industrial PID controllers usually have a minimum and a maximum manipulated variable parameter. If the manipulated variable is outside the allowed range, the integrator stopped.

11.4 Bumpless Transfer

Controllers usually have several operating modes. Bumpless transfer should be possible for switching between operating modes and also between a number of controllers. In this respect, standard controllers offer methods that are already integrated.

11.5 Control of an actuating drive with friction

The following section describes how to control a positioning drive using a direct current machine.

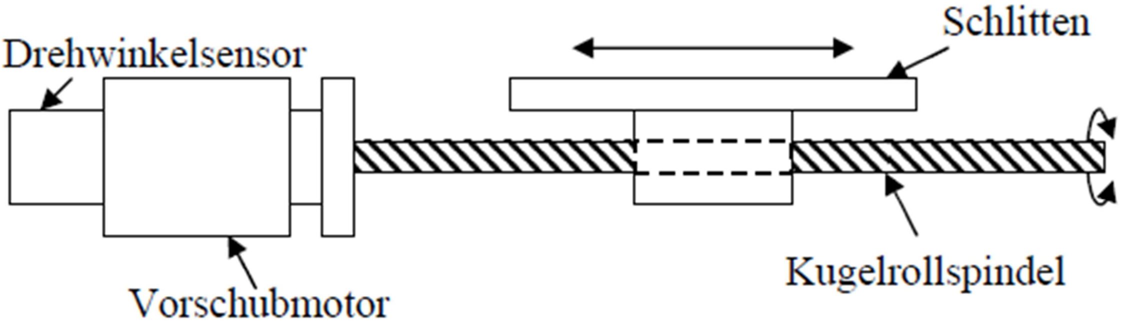


Figure 84, Actuating drive

The direct current has been modeled in a highly simplified manner as a PT1 element (normally PT2 due to armature circuit and mass inertia). The individual process, which is essentially determined by the armature circuit, has not been taken into account here. This is because the dynamic behavior of this direct current motor is predominantly determined by its mechanical properties. The mechanical model has also been expanded by a friction model.

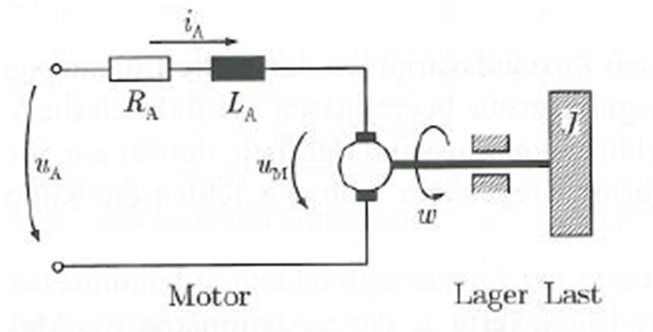


Figure 85, Modeling of direct current motor

The controlled system has an integrated character.
The friction model can optionally be switched on or off by a parameter. In addition, the *static friction portion* can be adjusted at any time. In addition to *static friction*, the mixed friction is also modeled. The mixed friction is caused by the growing velocity portion which causes the lubricant to sometimes be drawn into the friction gap. However, a load-bearing sliding film is not formed. With increasing velocity, the *mixed friction* or the friction portion of the dry friction is reduced (Stribeck effect). Eventually, the *mixed friction* becomes *viscous friction*, which increases linearly to the increasing velocity. With the *viscous friction,* the sliding layers are separated fully by a lubricating film.

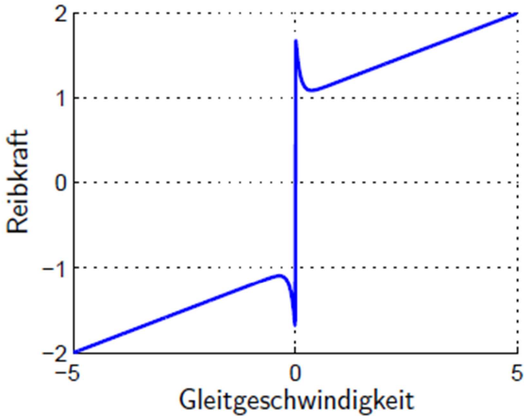


Figure 86, Friction model

The modeling was carried out in MATLAB and can be presented by the following Simulink structure.

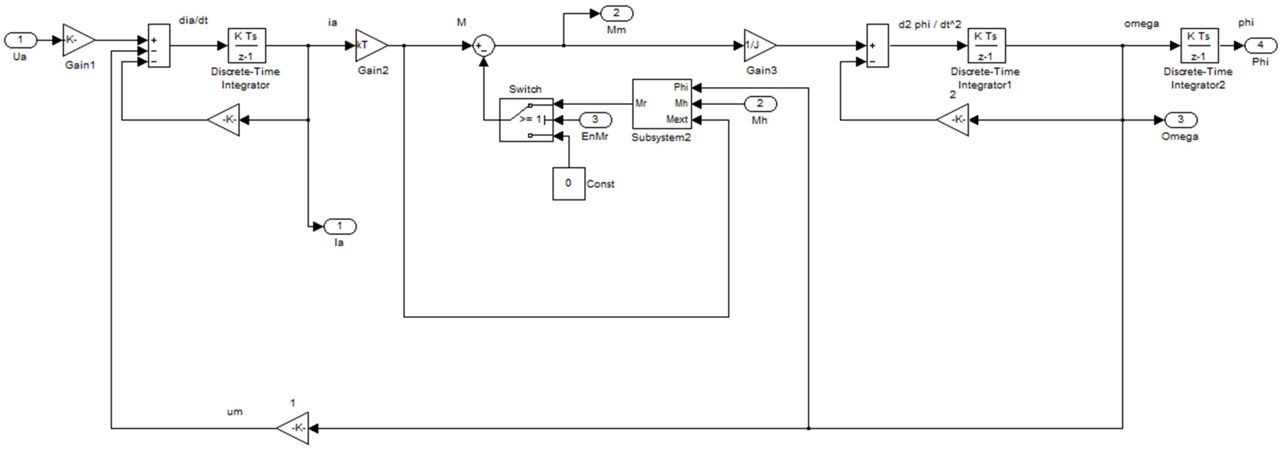


Figure 87, Principle of position control for actuating drive

Then, an Automation Studio program was created using the "B&R Automation Studio Target for Simulink".

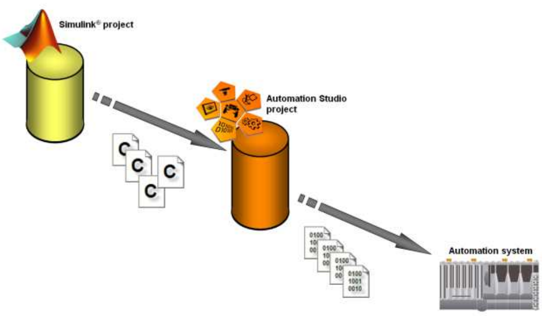


Figure 88, Code generation in MATLAB/Simulink

11.5.1 Deadband for systems without self regulation

The controlled variable of the controller is disturbed by a proportionately high level of measuring noise. The following section shows how to use the deadband in the controller to keep the manipulated variable constant below a defined limit so that the actuating drive is not continually controlled.

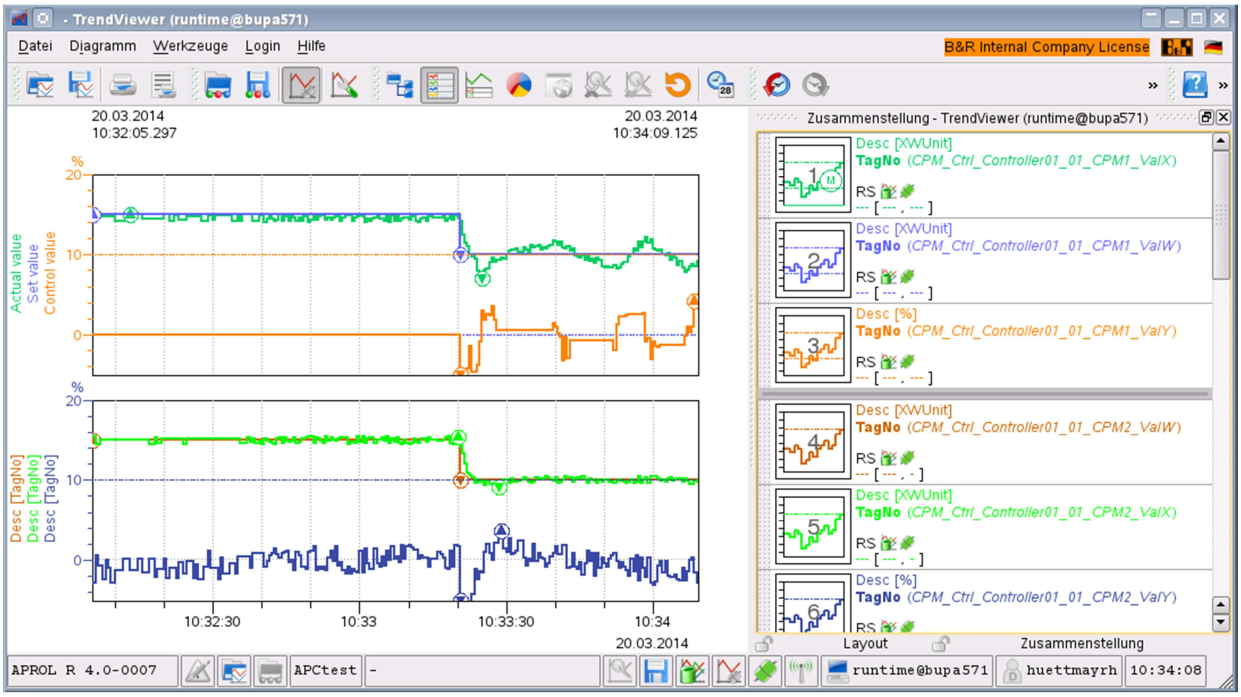


Figure 89, PID control with and without deadband (system without inherent regulation)

The following settings can be defined for the controller:
Kp =3.65, Tn = 1.93, deadband = 2

Conclusion: The deadband is easier on the actuator. Furthermore, the user saves money because less process air or power is used. What must be taken into account is that the control performance is reduced to the envisaged extent by the use of a deadband.

Important:

With systems without inherent regulation, the use of the deadband function can cause problems with the control performance.

11.5.1.1 Manual tuning of controlled systems without inherent regulation

In addition, the following tests can be carried out.
Test 2:
The following PID controller (Kp = 6, Tn = 2, Td = 15, Tf = 5) should be used for controlling the system.

Test 3:
A P controller is used for control, which is operated with an amplification factor of Kp = 20.

The following conclusions can be drawn from this.
1.) If the controller structure is unclear, always start with a P controller.
2.) The anti-windup improves the control performance.

12 CONTROL PERFORMANCE MONITORING (CPM)

This function block is used for the permanent, non-invasive monitoring of control performance. The control performance is represented by various dimensionless parameters, with which the calculated indices can be compared with values from similar control loops. The user can also use the configuration of alarm threshold values or notifications in case the control performance deteriorates. The function block can be easily attached to existing controller configurations.

12.1 Overview of CPM

The control performance of control loops can often be increased by improving efficiency in process control in the following ways:

- Reducing energy consumption
- Reducing production downtime
- Reducing the amount of mechanical wear
- Achieving faster startup times
- Achieving faster changeover times in production
- Increasing tolerance consistency in the end product
- Increasing throughput

Proportional effect on the equipment's profitability

Process systems often contain a vast amount of fundamental rules involved in stabilizing the various process parameters. The basic automation items must be setup up properly and in accordance with the current situation in order to ensure both safe and efficient operation. Furthermore, "Advanced Process Control" (APC) solutions are, first and foremost, dependent on the quality of the subordinate basic automation. Insufficient control performance can be caused by states or status changes such as these:

- Controller not in automatic mode
  - Operating personnel are not satisfied with controller performance and switch to manual mode.
  - Controller has not been activated since being serviced.
  - Analog values are still in "Force" state following the most recent maintenance work.
- Problems with the actuator
  - Mechanical wear can change characteristics (cavitation, friction, chipping).
  - Actuators are over- or under-dimensioned for the current requirements.
  - Residue buildup or leakage changes operating behavior.
  - Operation outside the permissible temperature range (with repercussions for the actuator performance).
- Measurement problems
  - Orifice plates have become offset or worn out of their original shape.
  - Sensors calibrated incorrectly or not at all.
  - Measuring ranges defined incorrectly or not at all.
  - Insufficient measuring resolution.
  - Measurement signal disrupted by external influences (poor shielding).
  - Measuring points physically moved due to subsequent modifications to the system.
  - Hardware or software error
- Changes to the process
  - Parts of the system have been expanded, modified, serviced, optimized, etc.
  - Standard production quantity has been increased or decreased (operating point shifted).
  - Changes to the properties of materials or components used for production.
  - Residue and/or wear can change the process performance.
  - Changes to the environmental conditions surrounding the process (temperatures, seasonal components, etc.).
  - Changes to recipes or operating methods.
- General causes of degraded control performance
  - Control loops not accurately setup (no accurate adjustment to the actual environmental conditions).

- No personnel available to check and/or maintain control loops during operation.

- Control-related failures
  - Failure to account for process-relevant inputs/outputs and linkages.
  - Incorrect composition of measurement and process parameters.
  - Insufficient degree of freedom for controller. For example, a PID that follows reference variables should also adjust for disturbance variables.
  - Failure to account for non-linearity in certain areas of operation.
  - Failure to account for, or incorrectly accounting for, dead times.
  - Disturbance variables not taken into consideration in the design of the control loop.

Control performance is affected by many external influences. When problems arise in the control loop, an analysis of the data from the CPM block can help identify the source of the problem. The process control system APROL can be easily used to display and analyze the data necessary for evaluating the control performance via the trend system.
The fundamental idea behind the CPM solution is to provide and display data that is relevant to the control loop. This function block was mainly designed for a classical PID controller.

12.2 Control performance monitoring for control loops

Manual operating mode ensures stable operation.
It is a well-known fact that PID controllers are generally not configured optimally and must therefore be operated manually at least some of the time. Plant operators are also not able to monitor the control loops allocated to them all of the time. This can only be achieved with an integrated functionality of the process control system.

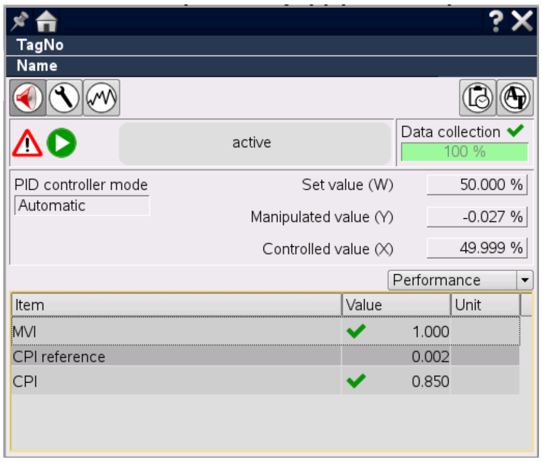


Figure 90, User interface for APROL Control Performance Monitoring

Creeping degradation

Monitoring control performance helps detect creeping degradations of control loop performance so that appropriate maintenance measures can be taken or control parameters optimized.

Statistical data identifies tendencies
The "ControlPerformanceMonitor01_01“ control module of the "Process Automation Library“ (PAL) provides statistics that can be used to assess the quality of control loops. These figures are crucial for achieving the maximum possible increase in process control efficiency.

Even Advanced Process Control solutions are quite fundamentally dependent on the quality of the subordinate basic automation.

The fundamental idea is to provide and display data that is relevant to the control loop. The figures calculated must be interpreted by the user. The Minimum Variance Index is obtained by comparing the variance of the controlled variable on the controller being used and the variance that would arise by using a minimum variance controller.

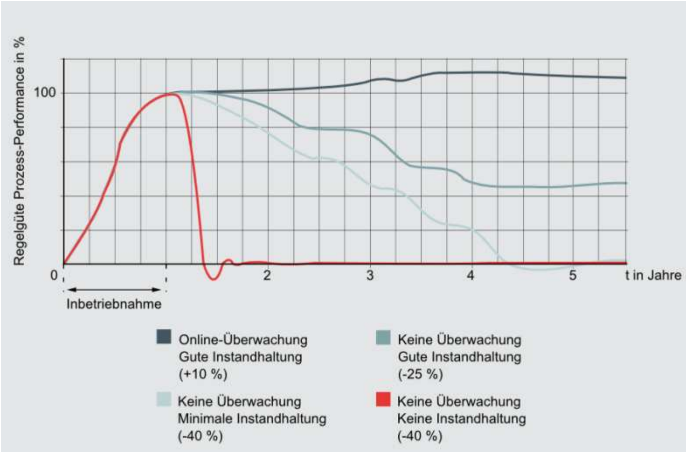


Figure 91, Control performance trend

If a drift in the MV index is identified over a given period of time, a change in control performance can be inferred from this. This means that it is not the absolute value which is relevant for an evaluation, but rather the relative value or drift.

The graph shows a possible curve of the control performance of control loops or involved assets over the life cycle of the system.

12.3 CPM figures

The CPM solution can be connected to an existing control loop at any time and it does not have any impact on ongoing processes or modules.

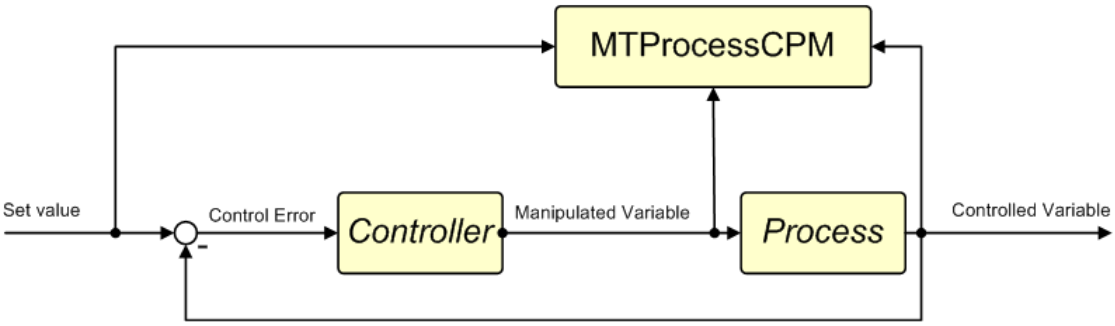


Figure 92, The CPM principle

As soon as the software function block is connected to the control loop and initialized, the following figures are calculated depending on the parameter settings. For the connection, only the setpoint, manipulated variable, controlled variable and operating mode of the controller need to be entered into the CPM function block.

12.3.1 Permanently calculated figures

The following specific values are provided by the module.

- Standard deviation of the set value
- Control average error
- Sum of the absolute values of the control error ("Integral Absolute Error", IAE).

$$IAE = \int |\, e(t) \,|\, dt \,)$$

Equation 56, Integral Absolute Error (IAE)

- Sum of the square of control errors, ("Integral Squared Error", ISE).
- Standard deviation for the control error
- Manipulated variable average
- Standard deviation for the manipulated variable
- Standard deviation for the process variable

- Numerator of the controlled variable setpoint crossings
- Presentation of process limits. E.g. How often or for how long is the actuator operated with limited manipulated variables (saturation).

12.3.2 Figures calculated depending on status

- Proportion of time in % in which the controller is operated in automatic mode ("loop uptime").
- System dead time
- Minimum Variance Index: The index is determined by the ratio of the controlled variable variance that would result from a minimum variance controller to the variance of the controlled variable of the controller being used. The minimum variance controller is a fictitious controller which very quickly compensates for the control error. The response of the manipulated variable can therefore be very dramatic. In controlled systems that permit small manipulated variable dynamics which are limited by the actuator, the value of the indices does not permit immediate information on the control quality. However, changes to the index over longer periods of time may well indicate a dropping control quality.
- "Control Performance Indices“ are generally described as differences between optimal/desired and current evaluation criteria.

$$\eta = \frac{J_{ref}}{J_{akt}} .$$

Equation 57, Dimensionless Control Performance Index

In this case, the variance of the controlled variable is used as a criterion for assessing the control performance. During startup, the function block is given a reference variable for the standard deviation, and this can be calculated based on the current measured values of the closed control loop. Assuming the control loop has been well configured and provided that the index is normalized to 1, the index will be in the range of 1 for very well configured controllers. Controlled processes or the dynamic properties of the processes may change over time. The normalized "Control Performance Index" also changes according to the predetermined reference variance. If the index has worsened accordingly, it can be concluded that the control loop must be reconfigured. As a function, the index is subject to a self-defined reference variable for certain deviations. For this reason, it is also possible that the index can, for a short time, also be greater than 1. If the index has worsened considerably, reconfiguring the controller is not always the solution to the problem, as defective actuators (e.g. valve friction) can also be a cause of the deterioration. The CPM solution can be based on the following reference variables.

- Standard deviation of controlled variable as reference
- IAE as reference variable
- ISE as reference variable

12.4 Evaluation of CPM

The calculated figures are summarized and displayed in the graphical user interface of the module.

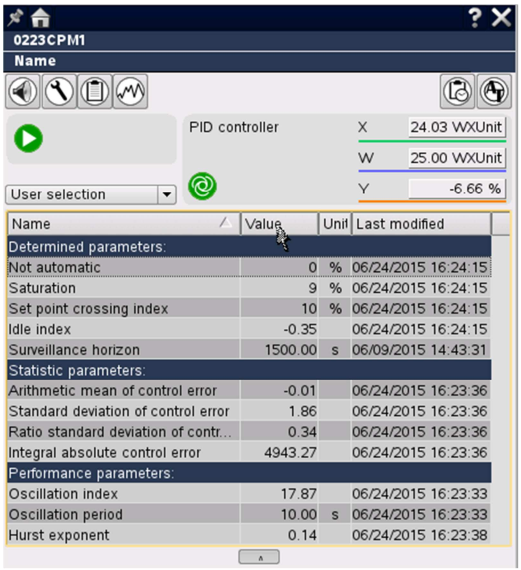


Figure 93, Statistical figures CPM/APROL

The figures are categorized into three different areas in the user interface.

- Performance: This area presents general dimensionless figures for evaluating control performance.
- Statistic: This area summarizes all figures with a statistical background.
- Determined: This area summarizes all figures that can only be calculated at a certain point in time (e.g. after a manipulated variable step).

In addition, the figures of the respective control loop are saved in the trend system for long-term storage. The data can be displayed either as a function of time or in x-y presentation over any periods. The high display performance of the APROL trend system also means large quantities of data (longer time ranges) are very quick to transfer.

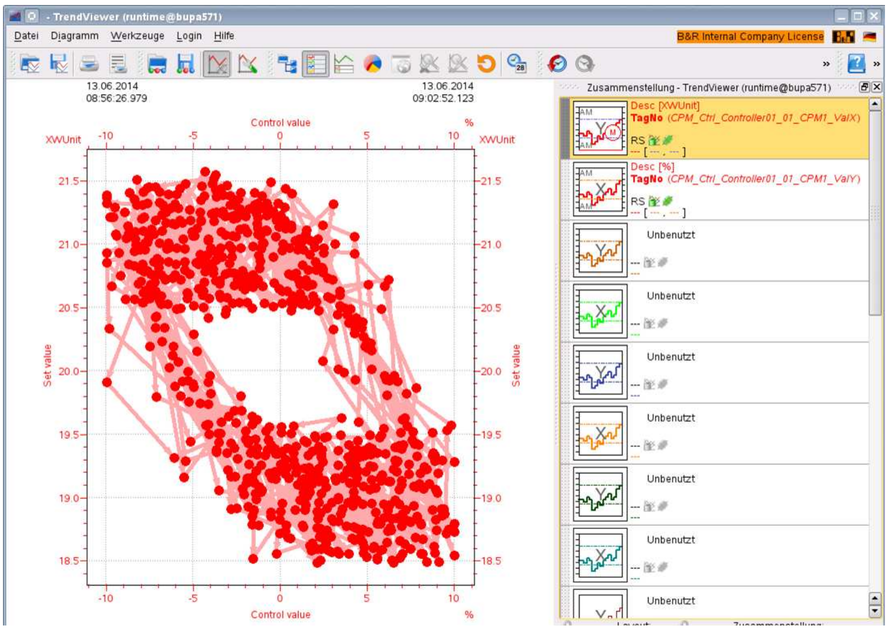


Figure 94, X-Y Presentation (“sticking valve”) of Trendviewer

The presence of the CPM solution can be configured by alarm or warning thresholds depending on the user. Furthermore, the user also has the option to evaluate the performance using APROL Reports. The "iReport Designer" function enables pre-prepared reports to be modified or adapted to the respective customer requirements.

Furthermore with the optional „Business Intelligence-Suite“ (Jaspersoft-Studio by TIBCO) the user is able to create web-based dashboards or rather reports.

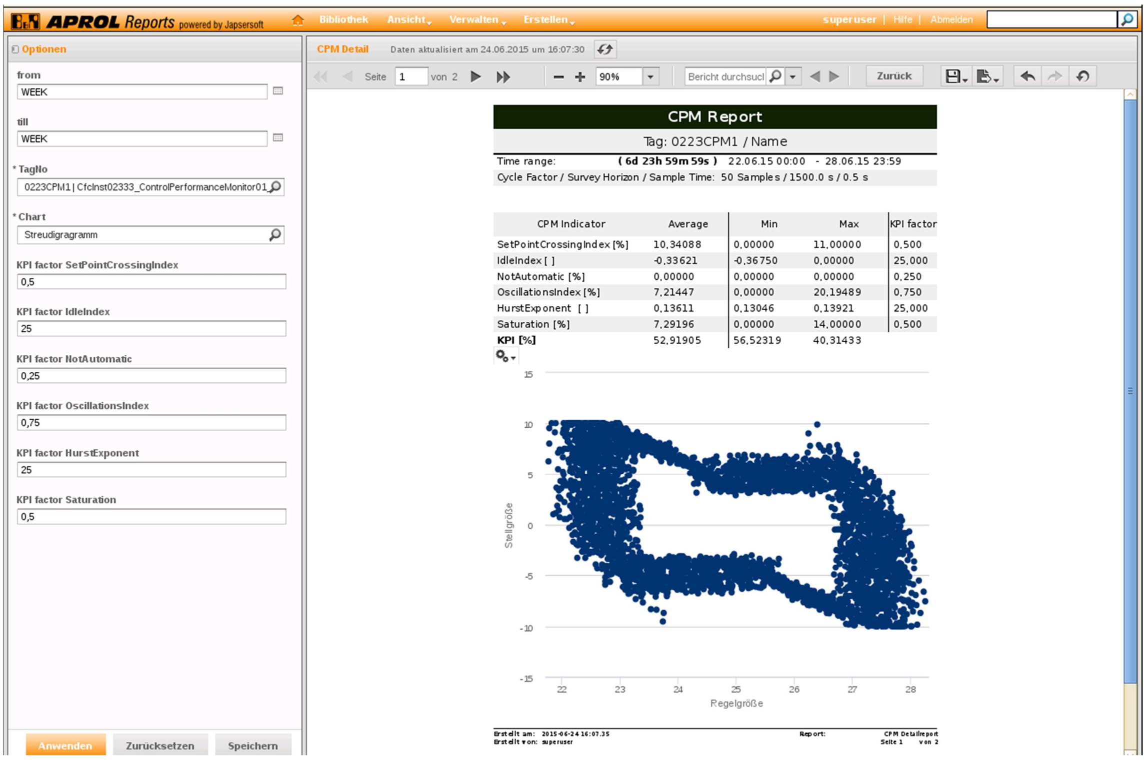


Figure 206, CPM report

13 RATIO CONTROL

The feed-forward control provided by the ratio function block is used to set a defined relation between the process variables of a system. The ratio function block can for example also be used to adjust the reference variable specification in a cascade control or for the direct feed-forward control of a second actuating element (synchronization control). A benefit of ratio control is the increased dynamics in relation to a further manipulated variable. A feed-forward signal can be generated immediately in relation to a specific ratio and a process variable (see Inertia through a further controller).

The output of the "Ratio Function Block" can be calculated as follows x = y * yRatio + yOffset. The setpoint of the secondary process variable is set as a specific ratio to the primary process variable in the ratio control.

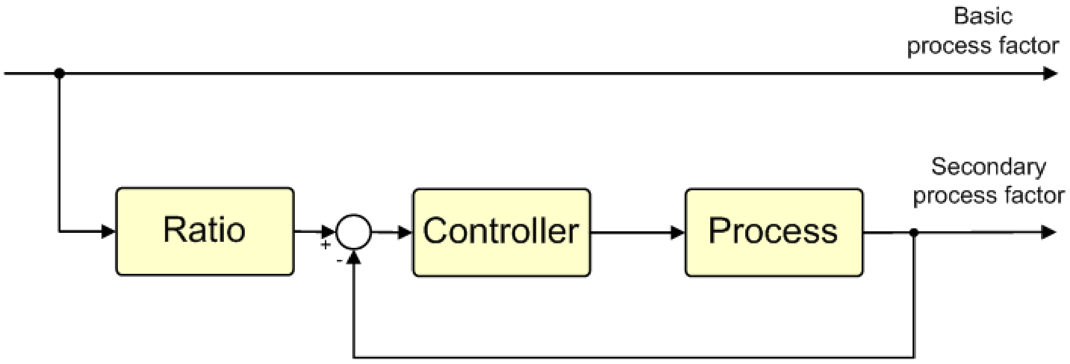


Figure 95, Ratio control principle

Ratio controllers can for example be used in analysis measurement engineering for the mixing of reactants or in process engineering for the manufacture of mixtures. A ratio function block can also be used to generate setpoints for controlling burners. For this, the combustion air volume is measured for example. The measuring signal is used to generate the setpoint for the gas volume controller, for example, in relation to a specific ratio. For the rate control a further controller is used in this case to control the combustion air volume or the furnace temperature for example.
First the ratio controller is configured and the master controller is in manual mode. Caution: the threshold for the gas/combustion air mixture after which no more combustion can take place must be clarified in the field for example.

Typical uses for ratio control:

- All types of mixture control (e.g. gas and air mixtures in combustion processes or acid/alkaline neutralization)
- Specification of a speed ratio for rolling processes
- Chemical reactions such as neutralization systems
- Generation of follow-up control signal
- The lighting affects the temperature of a climatic chamber due to the radiation of heat. As soon as the lighting is switched on, the manipulated variable is changed in relation to the lighting current.

- The following figure shows a ratio-dependent mixture control of two starting substances. The controlled flow is the precursor to the medium to be mixed. The ratio function block generates the setpoint for the delivery of additive. The output value of the ratio function block can either be delivered immediately as a setpoint for the actuating element (feed-forward only) or be used as a setpoint for a subordinate flow control.

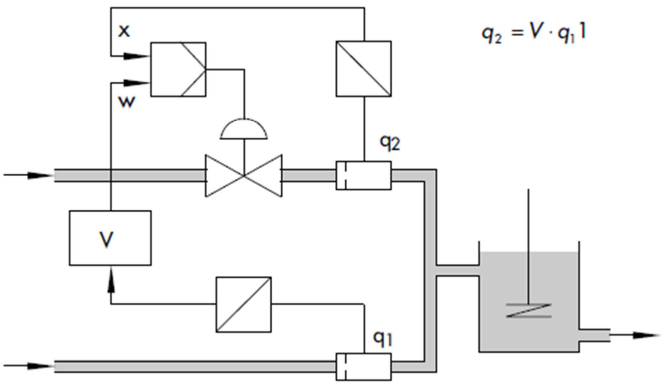


Figure 96, Ratio-dependent mixture control

13.1 Typical ratio control application

Many industrial processes require two 2 controlled variables with a fixed ratio to each other.
The ratio function block can keep the ratio between two specified variables constant.
The ratio function block is provided through the measurement of a process variable. The output of the ratio function block can be used to generate a setpoint for the control loop to control the second process variable in such a way that the set ratio is achieved between the process variables.

13.1.1 Temperature control

The following section shows how the temperature which arises as a result of mixing the two substance throughputs (A and B) is controlled by affecting substance throughput B. The substance throughput of flow A cannot be affected. Flow A or changes to the flow can be considered as disturbance variables in this respect.

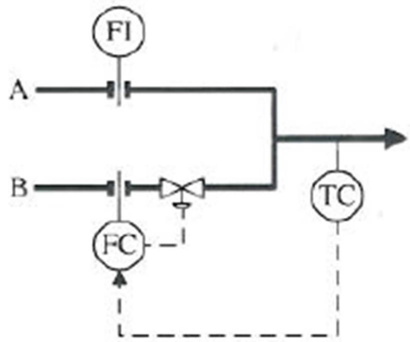


Figure 97, Mixing temperature control diagram

The structure of the control loop can be presented as follows.

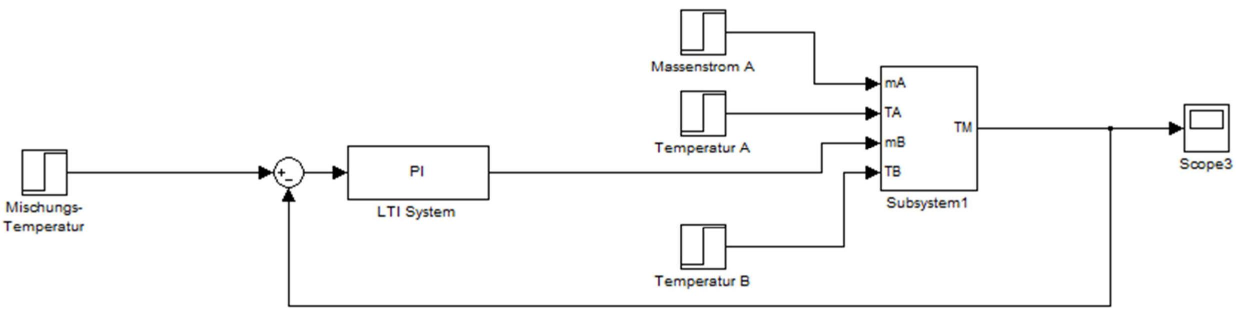


Figure 98, Mixing temperature control principle

The flow control can be put into operation with the following parameters Kp = 0.02, Tn =0.3. The temperature in A can be set to 10℃, the temperatur e in B to 35 ℃ and the mass throughput in B to 1 kg/s. In addition, after starting up the control loop (mixing temperature setpoint 20℃), a step-like disturbance of the mass throughput in A should be simulated (e.g. from 1 kg/s to 3 kg/s).

The control loop can respond to changes in A only as a disturbance controller. The following problems can occur with mass throughput changes in A.

- The temperature controller can (if a controller with 2 degrees of freedom is not used) either be configured as a master controller or as a disturbance variable controller. Under these conditions, an optimum result cannot be achieved.
- If there is a high system time constant between the controlled variable (temperature) and the manipulated variable (flow B), or even simply just a portion of dead time (if for example the temperature measurement equipment is not placed close enough to the mixer from a local perspective), the control error will also reach a certain value (step-like change in A -> possibly calculate max. error) even with an optimally configured disturbance variable controller (from a dynamic perspective).

13.1.2 Feed-forward through the ratio function block

The following control concept uses the flow measurement of A (disturbance variable) to generate a feed-forward signal for the flow control of B. The flow measurement of A enables a response to be made even before the disturbance of or change in mixing temperature (predictive components) due to a ratio-dependent change in the feed-forward signal (control loop B).

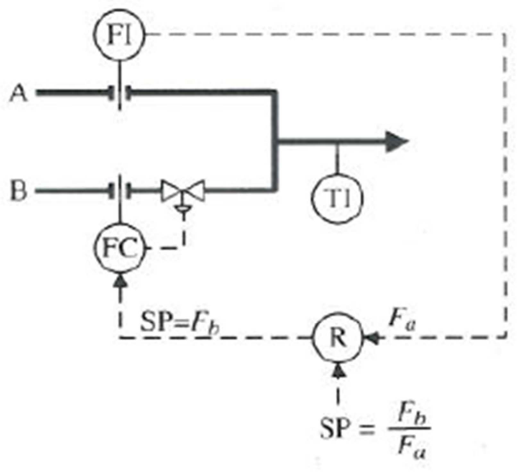


Figure 99, Mixing temperature feed-forward control diagram

The setpoint for B is determined by multiplying the measured disturbance variable A and the adjustable (plant operator) ratio (Flow B to Flow A). If the flow in A changes, the ratio function block will adjust the flow in B so that the temperature remains constant. The ratio can be calculated with various specific heat capacities (cp) as follows

$$R = \frac{cp_A(T - T_A)}{cp_B(T_B - T)},$$

Equation 58, Feed-forward ratio of mixing temperature

so that the desired mixing temperature T can be achieved. The disadvantage is that at least 2 further temperature measurements are needed for this control method.

Strictly speaking, this example is not a ratio control (but an open control) because the actual value of the ratio is not calculated or measured and compared with the actual setpoint of the ratio. However, in practical terms and in everyday language the term ratio control is used in both cases (control of the ratio and open control).
Assuming that the specific heat capacities of the two liquids are the same, the following correlation results

$$\dot{m}_A T_A + \dot{m}_B T_B = (\dot{m}_A + \dot{m}_B) T.$$

Equation 59, Energy balance for mixture of substance throughputs

With $R = \frac{\dot{m}_B}{\dot{m}_A}$ the ratio in relation to the energy balance can be formulated as follows

$$R = \frac{T - T_A}{T_B - T}$$

Equation 60, Feed-forward ratio of mixing temperature with equal heat capacities

.
The concept of feed-forward control is as follows.

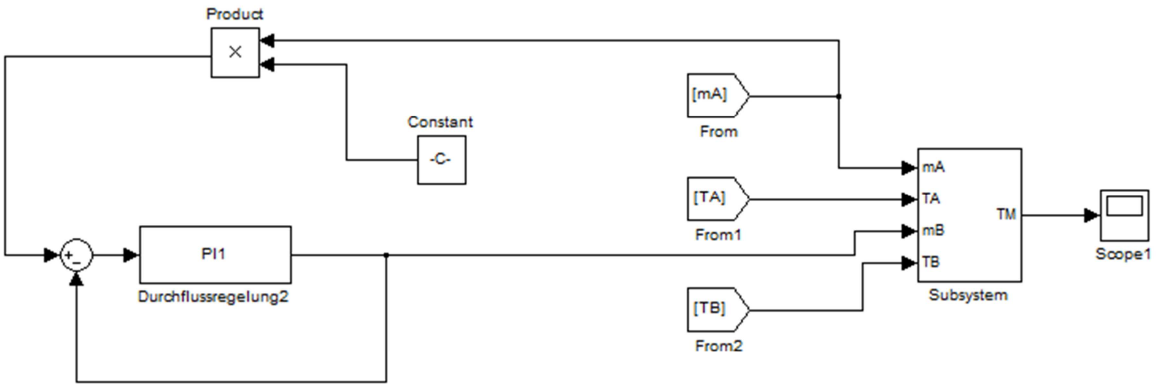


Figure 100, Principle of feed-forward control of mixing temperature through the ratio function block

In this task the ratio (according to Equation 60, Feed-forward ratio of mixing temperature with equal heat capacities) can be calculated with the numerical values of the previous task.

13.1.3 Feed-forward components with temperature control

As soon as the process is disturbed by other variables, the feed-forward control cannot actively act against the disturbance variable. The feed-forward control assumes that the flow measurement instrumentation is correct. If deviations occur here, the mixing temperature will no longer correspond with the setpoint.
For this reason, a simple solution is needed which includes the actual mixing temperature in the control loop. The feed-forward signal should nevertheless be used to improve the temperature control. Under these conditions, the control diagram can be presented as follows.

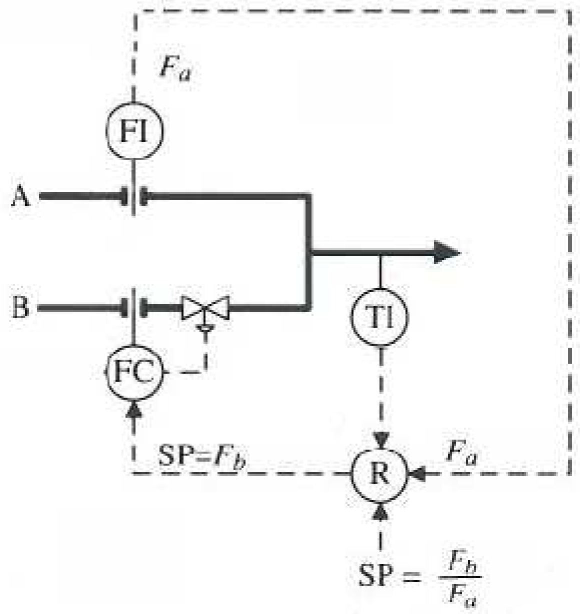


Figure 101, Diagram of mixing temperature control with feed-forward control

The measured mixing temperature is used in this case for influencing the ratio setpoint.

The temperature control in combination with the feed-forward control can be presented as follows.

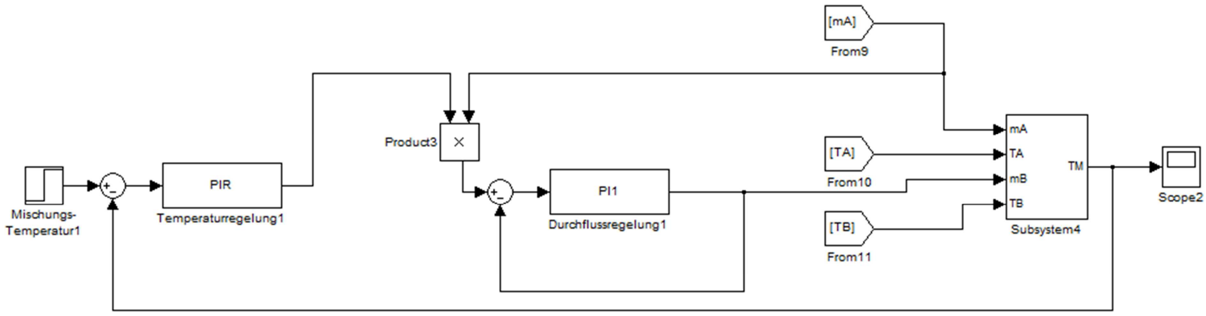


Figure 102, Principle of mixing temperature control with feed-forward control

In this case, the ratio controller can be operated with Kp = 0.1 and Tn = 0.5. After starting up the controller

1. a step in the mass throughput in B can be simulated (from 1kg/s to 3kg/s),
2. a step in the temperature in A can be simulated (from 10℃ to 4℃).

The control performance should then be compared with the two previous control methods carried out.

13.1.4 Feed-forward components with temperature control and new reference variable

Up to now, the ratio is calculated as follows $R = \frac{\dot{m}_B}{\dot{m}_A}$. Using the energy balance, the mixing temperature can be formulated as a function of the ratio as follows

$$T = \frac{T_A + RT_B}{R + 1}.$$

Equation 61, Mixing temperature as a function of the mixing ratio

The process gain $V = \frac{dT}{dR} = \frac{T_A+T_B}{(R+1)^2}$ in this case is directly dependent on the ratio. In the event that the ratio is subject to various changes in operating conditions, this dependency can lead to problems with the higher-level temperature control loop (non-linear behavior). As the process gain changes according to the operating points, the parameter settings of the temperature must be adjusted in the worst case scenario.

This phenomenon can be prevented simply by changing the reference variable of the ratio $R = \frac{\dot{m}_B}{\dot{m}_A+\dot{m}_B}$ to the overall mass throughput. The mixing temperature, as a function of the ratio, is in this case

$$T = (1-R)T_A + RT_B.$$

Equation 62, Change of reference variable for mixing temperature as a function of the mixing ratio

In this case, the process gain is not a function of the ratio R. The structure of the control loop would change as follows.

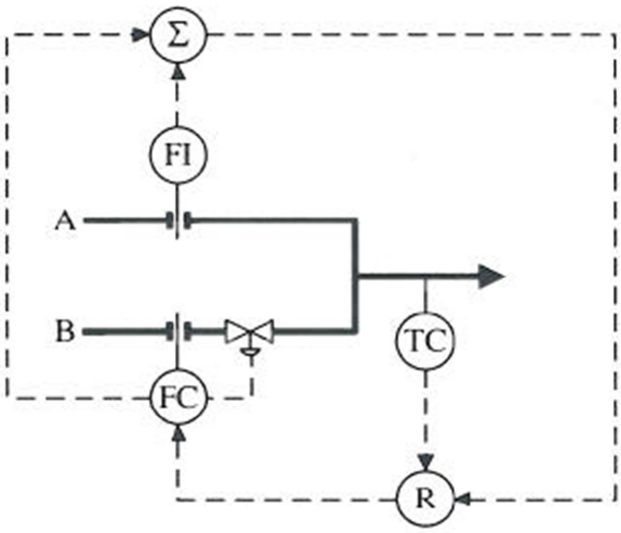


Figure 103, Diagram of mixing temperature control with feed-forward control and change of reference variable

The temperature control with the modified feed-forward control can be presented as follows.

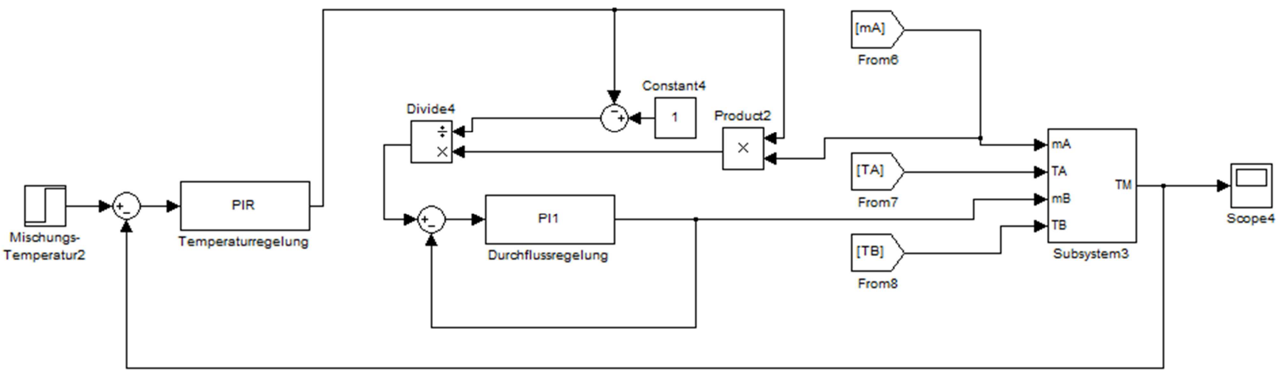


Figure 104, Principle of mixing temperature control with feed-forward control and change of reference variable

13.1.5 Summary

The four different methods for controlling the mixing temperature can be summarized as follows.

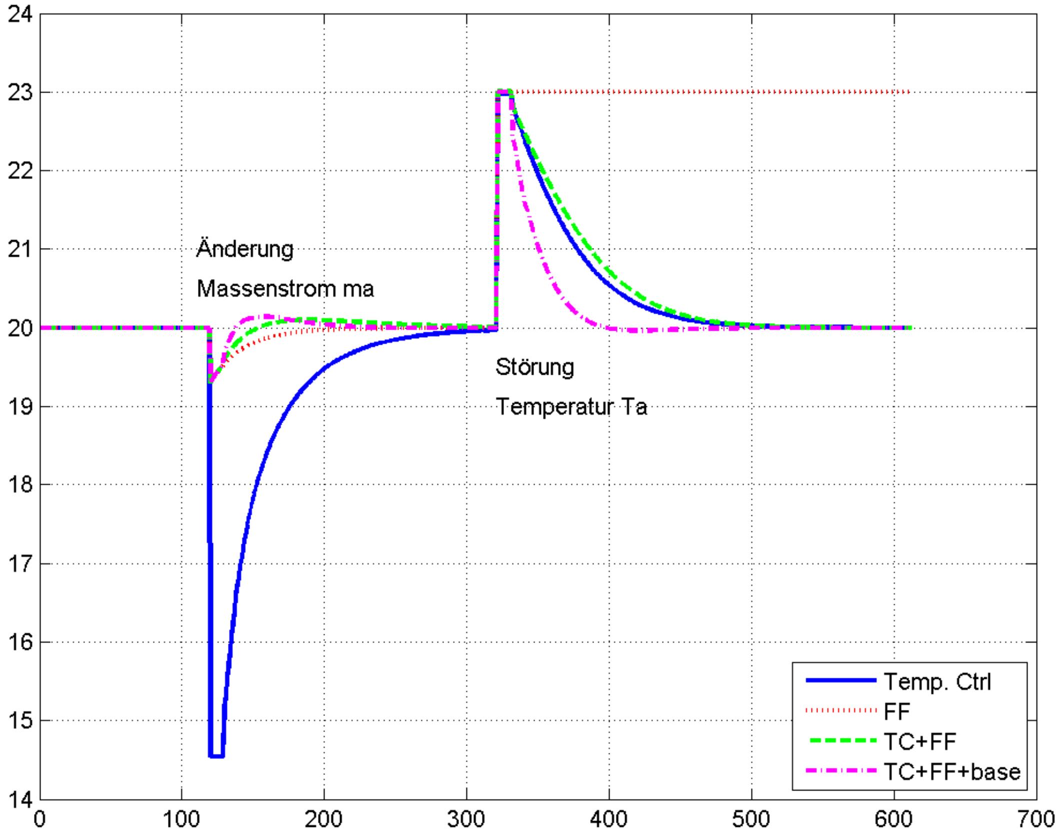


Figure 105, Comparison of mixing temperature control methods

The setpoint of the mixing temperature was kept constant over the simulation period. In this example, the temperature measurement of the mixing temperature shows a dead time of 1 second. As shown in the figure, this feed-forward control (FF) cannot compensate for any unplanned disturbances (e.g. a temperature fluctuation in loop A). The further simulation data is as follows.
Ta = 10℃ (disturbance +5℃), ma = 1kg/s (change to 3 kg/s), Tb = 35℃, cp = const. = 4.18 kJ/(kg K).
The tasks were performed with a PI controller with a 100ms process sampling time.

Note:
The control of the control loop with the ratio function block can, for example, also be moderated if the primary setpoint is used instead of the primary controlled variable. In this case, it must be assumed that the controller for controlling the setpoint of the reference variable is very well configured.

14 LINEAR INTERPOLATION OF GRID POINTS (LOOKUP TABLE)

In order to generate piecewise linear functions, a one-dimensional or two-dimensional grid point interpolation can be used as needed.

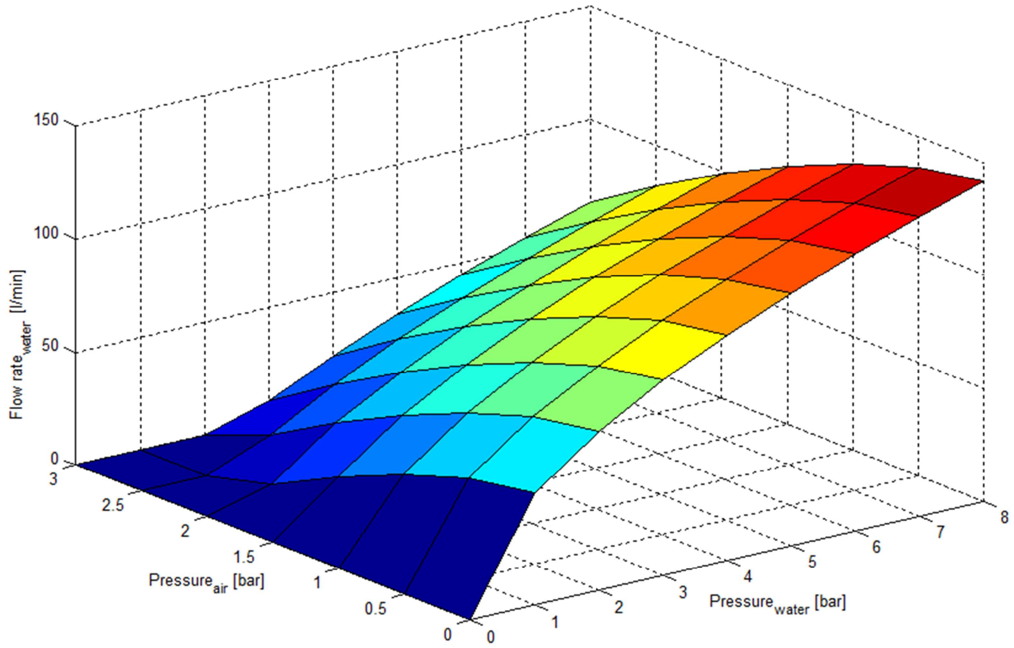


Figure 106, Cooling a continuous casting plant

Air/water cooling injectors are used for cooling continuous casting plants. The image above shows the cooling action arising from the mixture of water and air pressure.

14.1 Lookup 1D

This module can be used to generate a one-dimensional function using grid points. The "LookUpTable01_01" module in APROL can be used to process up to 50 grid points.

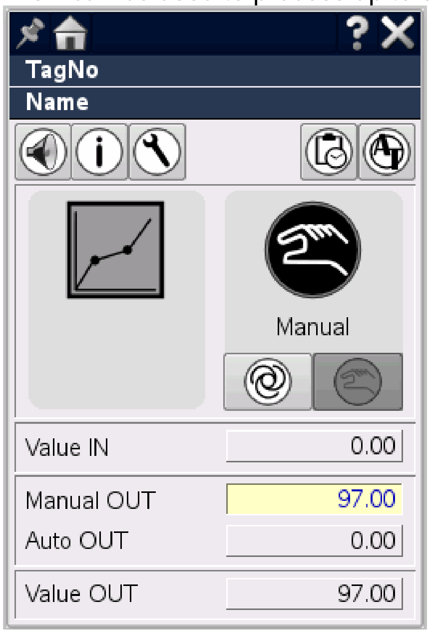


Figure 107, User interface for 1D grid point interpolation

14.1.1 Operating principle

Figure 108, Characteristic curve for LookUP Table 1D shows a characteristic curve which was generated with the following parameters.

Input	Output
0.5	0.5
1.0	1.0
1.5	1.3
2.0	1.5
2.5	2.0
3.0	2.2
3.5	3.5

Table 14, Parameter settings for LookUP Table 1D

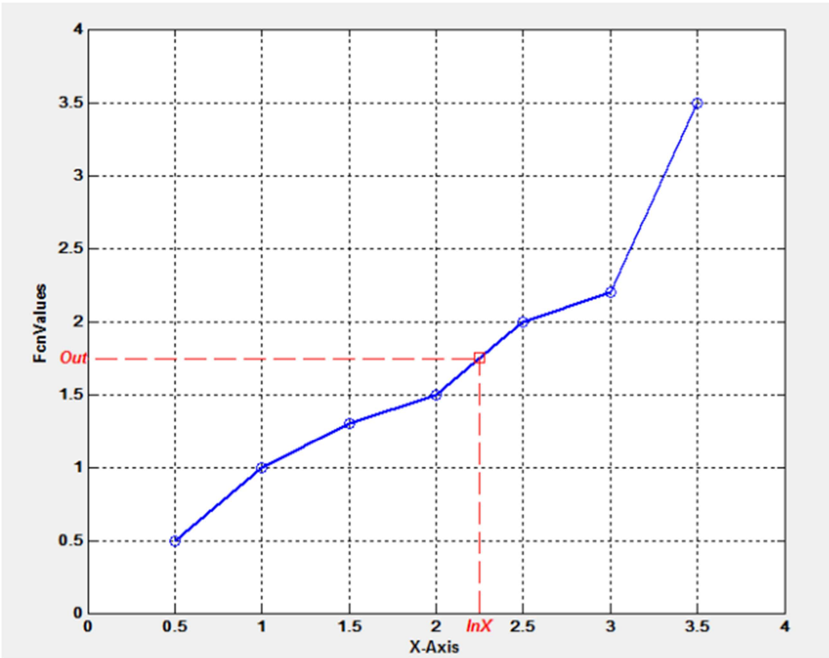


Figure 108, Characteristic curve for LookUP Table 1D

The input value InX = 2.25 gives an output value of Out = 1.75 through the linear interpolation between the grid points.

If InX is outside the node range, Out cannot be calculated using interpolation. In this case, there are two different methods available.

- "Calculation mode out of range" = "Constant end value". In this case, the boundary values (outside the node range) are kept constant.
- "Calculation mode out of range" = "Linear extrapolation". This configuration leads to a linear extrapolation, which is carried out based on the two left or right boundary values.

The next figure demonstrates the mode of action of the two methods.

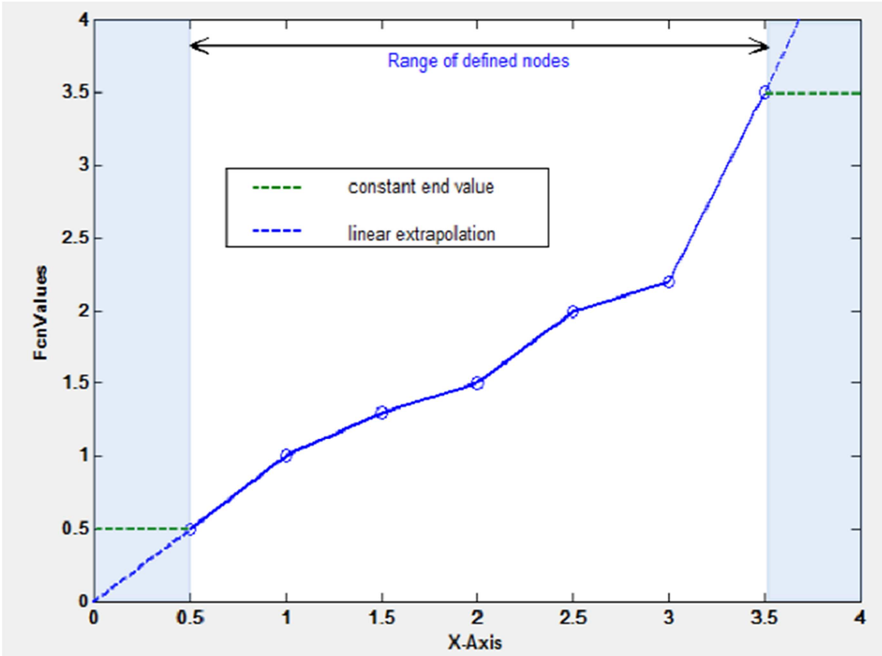
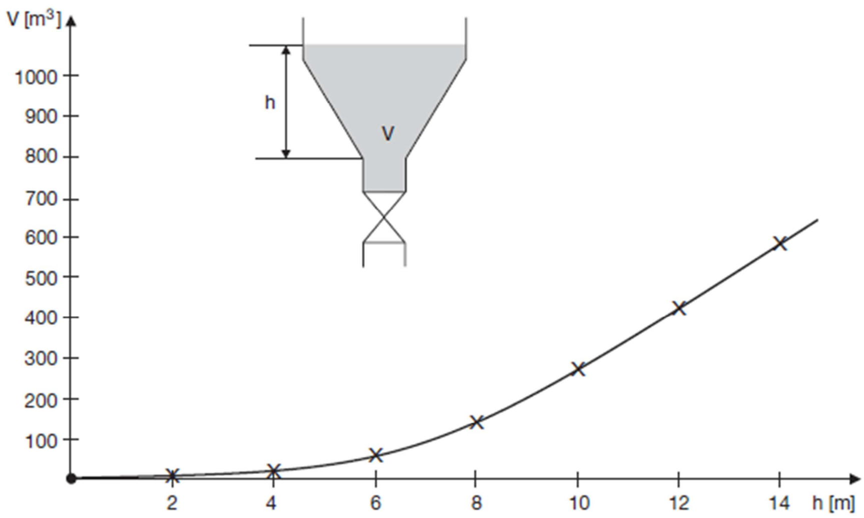


Figure 109, LookUp Table 1D, Boundary points methods

Applications:

The LookUp Table can for example be used for customer-specific linearization (taking into account the non-linearity of a PT100 measuring sensor). The following example uses the LoopUp Table for calculating a container fill volume depending on the filling level.



Equation 63, Use of 1D linear grid point interpolation

Furthermore, the linear grid point interpolation can be used for the following applications, among other things.
Generating ratio-dependent feed-forward signals (flow temperature of heating, formulas), adjustment of signals, linearization.

14.2 Lookup 2D

A typical curve family comes from the control of combustion engines. In this case, the engine torque can be specified as a function of the input variables of engine speed and throttle angle (gasoline engines) or the injection volume (diesel engines).

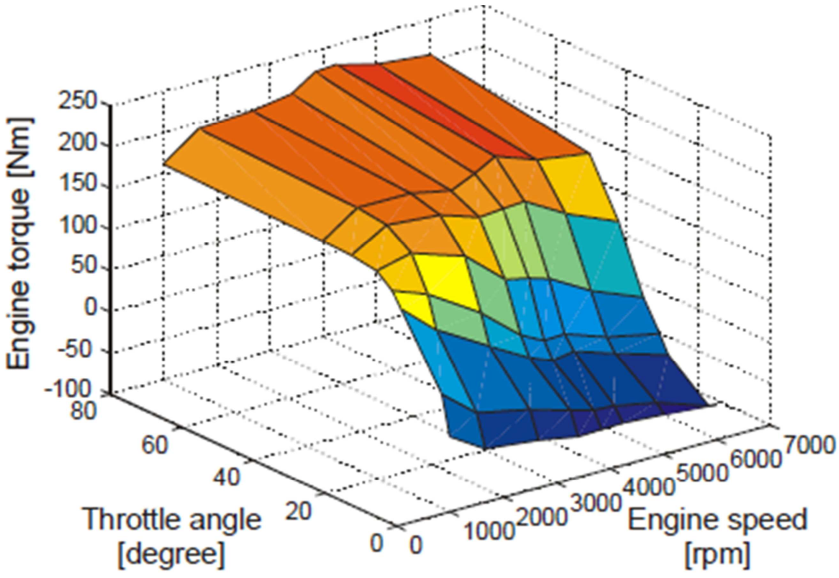


Figure 110, Calculating the engine torque

Application:
Non-linear control, interpretation and evaluation of multi-dimensional correlations. Connection between pressure and temperature to evaluate density or dynamic viscosity of certain media.

15 SPLIT RANGE CONTROL

The "Split Range" control is available to several actuators for controlling a controlled variable. This principle enables the manipulated variable of a controller to be divided over various actuating elements. This division is necessary for example when an actuator cannot provide the necessary output, or the adjusting range of an actuator is not sufficient to control the process. The figure below shows a split range control with a case of poor cooling capacity of a cooling tower; if the cooling is no longer sufficient the additional cooling capacity is drawn from a chiller.

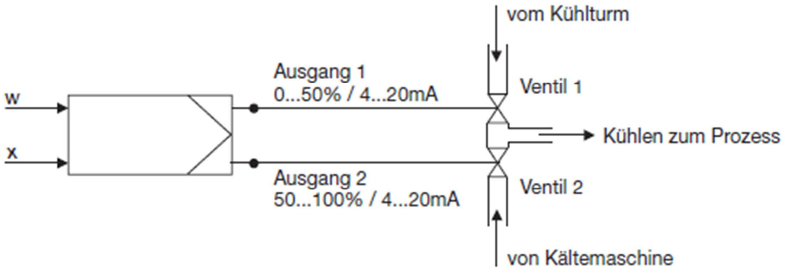


Figure 111, Cooling with split range control

The division is made using specific values for the input variable. Two independent actuator characteristics can be configured for the split range function block. This gives the user the best possible flexibility when using the function block.

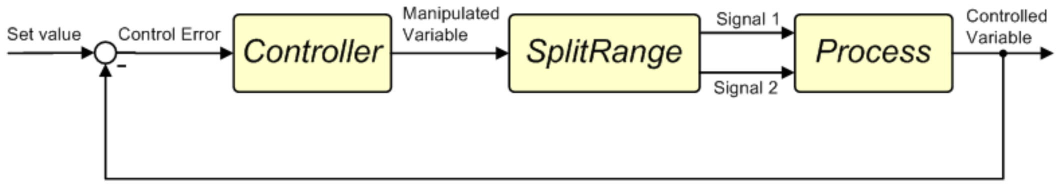


Figure 112, Principle of split range control

The split range control can for example be used in processes with bidirectional manipulated variable requirements and actuating elements with a unilateral action. For example, temperature control systems which use one actuator for heating and one actuator for cooling. In this case, the adjusting range of 0..100% needs to be divided between both complementary elements using a deadband.

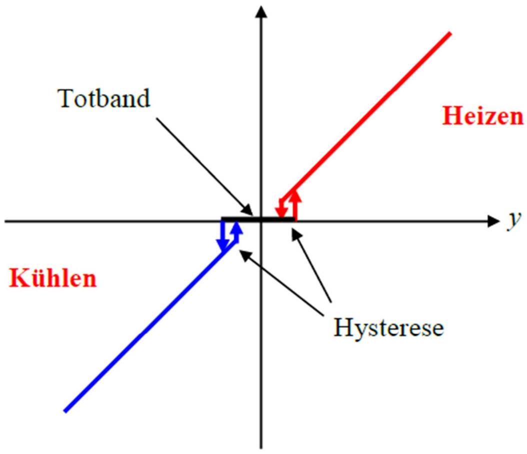
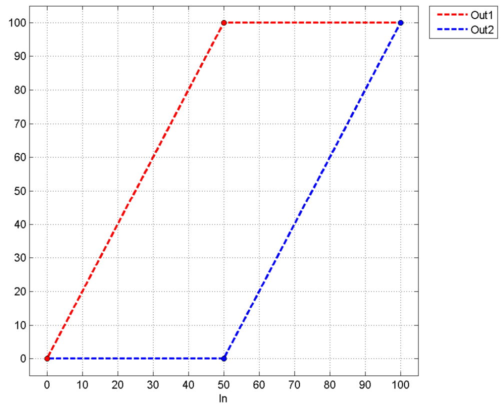


Figure 113, Split range control with deadband

The parameter settings for the "SplitRange01_01" module can be as follows. Two reference points are defined for each output signal. If the input signal is between the reference points, then the respective output signal is calculated by linear interpolation. If the input signal is outside the defined coordinates,

then the respective output signal is driven constantly with the function value of the coordinate closest to the reference input.
Definition of the individual coordinates makes it possible to also individually configure characteristics for the outputs with overlapping or deadband without requiring extra input parameters. The following graphic shows different Split-Range configuration options.

<table>
<tr><th>I/O behavior</th><th>Configuration</th></tr>
<tr><td>


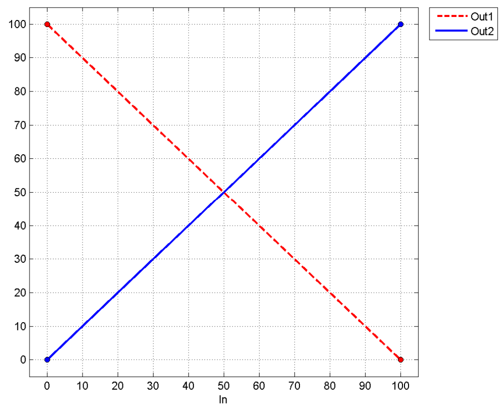
</td><td>

Output 1

<table>
<tr><td>Parameter</td><td>Input</td><td>Output</td></tr>
<tr><td>Start</td><td>0.0</td><td>0.0</td></tr>
<tr><td>End</td><td>50.0</td><td>100.0</td></tr>
</table>

Output 2

<table>
<tr><td>Parameter</td><td>Input</td><td>Output</td></tr>
<tr><td>Start</td><td>50.0</td><td>0.0</td></tr>
<tr><td>End</td><td>100.0</td><td>100.0</td></tr>
</table>

</td></tr>
<tr><td>


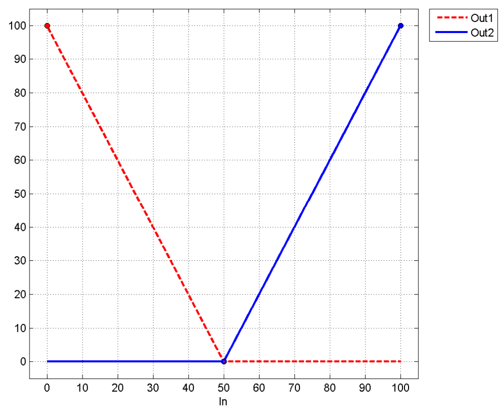
</td><td>

Output 1

<table>
<tr><td>Parameter</td><td>Input</td><td>Output</td></tr>
<tr><td>Start</td><td>0.0</td><td>100.0</td></tr>
<tr><td>End</td><td>100.0</td><td>0.0</td></tr>
</table>

Output 2

<table>
<tr><td>Parameter</td><td>Input</td><td>Output</td></tr>
<tr><td>Start</td><td>0.0</td><td>0.0</td></tr>
<tr><td>End</td><td>100.0</td><td>100.0</td></tr>
</table>

</td></tr>
<tr><td>


</td><td>

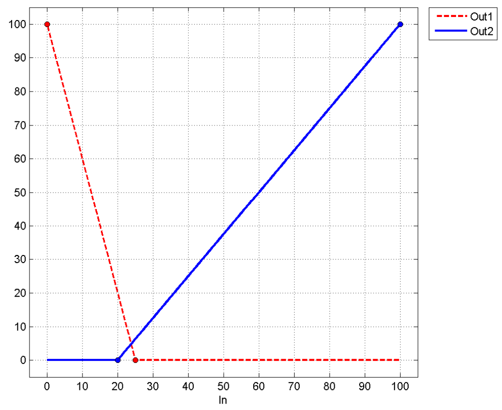
Output 1

<table>
<tr><td>Parameter</td><td>Input</td><td>Output</td></tr>
<tr><td>Start</td><td>0.0</td><td>100.0</td></tr>
<tr><td>End</td><td>50.0</td><td>0.0</td></tr>
</table>

Output 2

<table>
<tr><td>Parameter</td><td>Input</td><td>Output</td></tr>
<tr><td>Start</td><td>50.0</td><td>0.0</td></tr>
<tr><td>End</td><td>100.0</td><td>100.0</td></tr>
</table>

</td></tr>
</table>

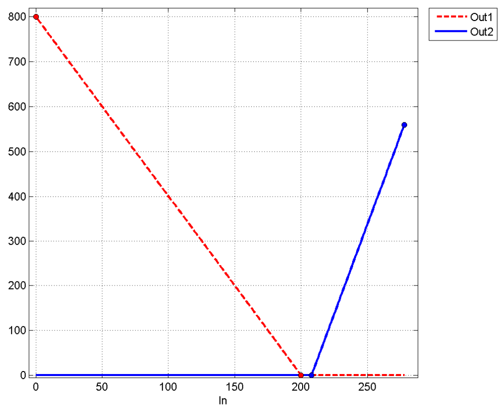


Output 1

Parameter	Input	Output
Start	0.0	100.0
End	25.0	0.0

Output 2

Parameter	Input	Output
Start	20.0	0.0
End	100.0	100.0



Output 1

Parameter	Input	Output
Start	0.0	800.0
End	200.0	0.0

Output 2

Parameter	Input	Output
Start	208.0	0.0
End	278.0	560.0

Table 15, Examples of configurations for the split range control

Typical applications for the split range control:

- With temperature control, the controlled variable is increased by reducing the cooling capacity. If the setpoint is within a range that cannot be achieved by cooling (actuator 1) (definition:"Split" point), the second actuator for heating capacity is brought into the system, and the setpoint can be achieved with this.
- Pressure control of steam collection bars supplied by various steam generators.
- Flow control for large pipes. With a single actuating element, the limits for controllability are soon reached, as the necessary precision cannot be provided over the entire flow range. The adjusting range can be divided between two smaller valves. With low quantities, only the smaller valve is controlled by the split range block. With large quantities, both valves are controlled. In addition, the split range block can be used to normalize the actual control range of the fitting to the input signal of the valve.
- Thanks to the individual adjustability of the split range function block, this can also be used for the simultaneous, mirror-inverted control of valves. Thus the function of a three-way valve can be emulated using two standard valves. The following figure shows the cooling down of a process through an air-cooled condenser. As soon as the pressure increases, the controller closes the bypass valve, and at the same time the condenser valve is opened.

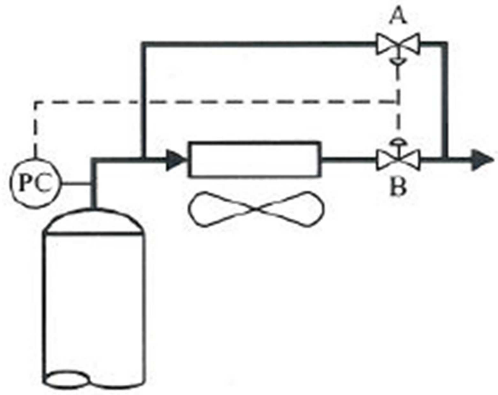


Figure 114, Simultaneous valve control with split range function block

- Split range valve controls can for example be implemented if the controller gives a read out from 0-100% that the valve is showing a fail-open behavior (at 0% it is fully open), or if adjustments need to be made to the gain between the control signal of the controller and the control signal sent to the actuator.

16 OVERRIDE CONTROL

Multiple controllers affect the same actuator when release control is involved. The controllers are wired so that only one controller is active at a time. A controller is given access to the actuator based on the current process state. The controllers are capable of overriding each other. The decision about which controller should be active at a given time can be made either by comparing the two controllers' control values (minimum or maximum manipulated variable is given priority no. 1) or by using definable threshold values related to the controlled variable or an external reference variable.
override criterion based on the manipulated variable of both controllers. The controller with either the largest or smallest output signal in terms of magnitude is given access to the actuator. The occurrence of windup effects needs to be prevented for the passive controller by tracking the I portion. With a low measuring signal quality, there is the risk that one override process will follow another.
Problems of this kind can be prevented by using an "External Reset Feedback" tracking input.

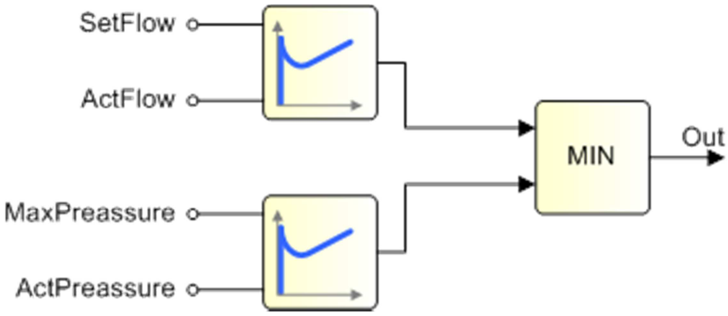


Figure 115, Override control principle

An APROL template is available on this subject, in which two PI controllers are controlled via a minimum selector on an actuating element. The current value of the actuating element is delivered to the two controllers via an "external reset" input.

Selective control can be viewed as a form of multiple variable control, in which the selective element enables the online modification of the control strategy. The following tasks can be dealt with using override control.

- Automatic startup and shutdown processes
- Protection against measuring equipment errors, e.g. by calculating the median from several measurements.

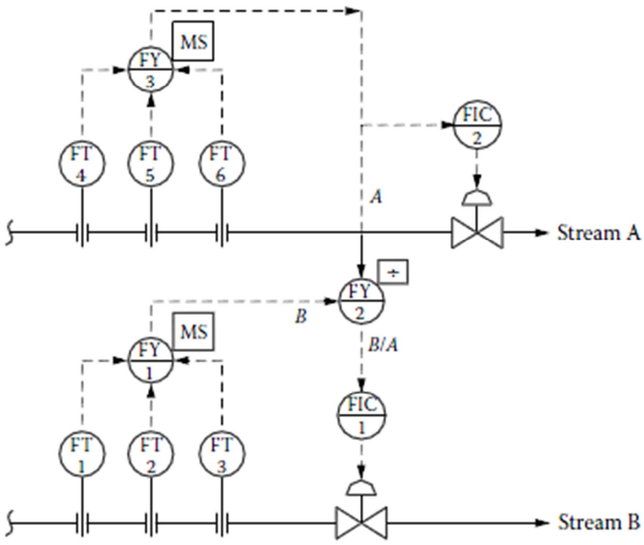


Figure 116, 2 out of 3 measurements

- Protection of plant parts. Override controls are an advantage with regard to preserving the availability of plant parts, as the plant parts are not shut down as a result of suboptimal operation, unlike with interlocking mechanisms for example. With override control, the plant parts can continue to be operated at least with reduced productivity. As shown in the following figure, a further limit can be used to change the control strategy during operation before reaching the interlocking point ("Hard Constraint").

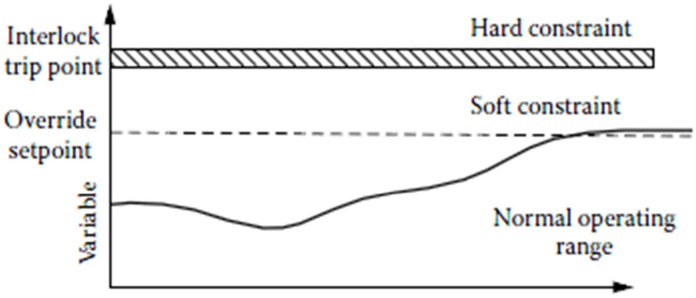


Figure 117, Override control for protecting plant parts

An inactive controller of an override configuration in standard operation remains inactive until a limit is reached or exceeded. From this point, the previously inactive controller takes over the calculation of the manipulated variable through a selection element and prevents the limit from being overshot any further.

16.1 Tuning

Both control loops can for example be configured using a manual tuning process or the auto-tune function. The parameters of each individual controller are calculated separately. For this purpose, the system must be operated in the respective operating range of the controller, which, for tuning, should be an appropriate distance from the changeover point with the other controller.
In override controls in which a controller takes on a safety function, the operation or tuning process is not always possible in this range. In this case, as much data as possible on the system behavior should be collected in this range, which can then be used to calculate the controller parameters

16.2 Typical uses for override control

- Flow control of a steam pipe with pressure limitation. In this case, the two output signals of the controller are fed to a minimum selector, which forwards the lower manipulated variable of the two signals to the control valve. The setpoint of the pressure controller is set to the maximum allowable steam pressure that can still be coped with by the controlling users but which does not yet reach the critical safety limit (overpressure trip, emergency off switch). As long as the nominal value of the pressure controller is not exceeded, the manipulated variable of the pressure controller is blocked. If the allowable pressure increases at the controlling users, the manipulated variable of the pressure controller decreases. As of a certain limit, the pressure controller, triggered by the minimum selector, takes over and results in a restriction of the control valve. The pressure controller therefore triggers the flow control and prevents a critical increase in the steam pressure. An "external reset feedback" input (see Override control with “back calculation” input

An industrial PID-controller usually exhibit an integral-tracking input which could be used to do override control applications. Integral tracking could be achieved via the so called back-calculation principle pictured in following figure.

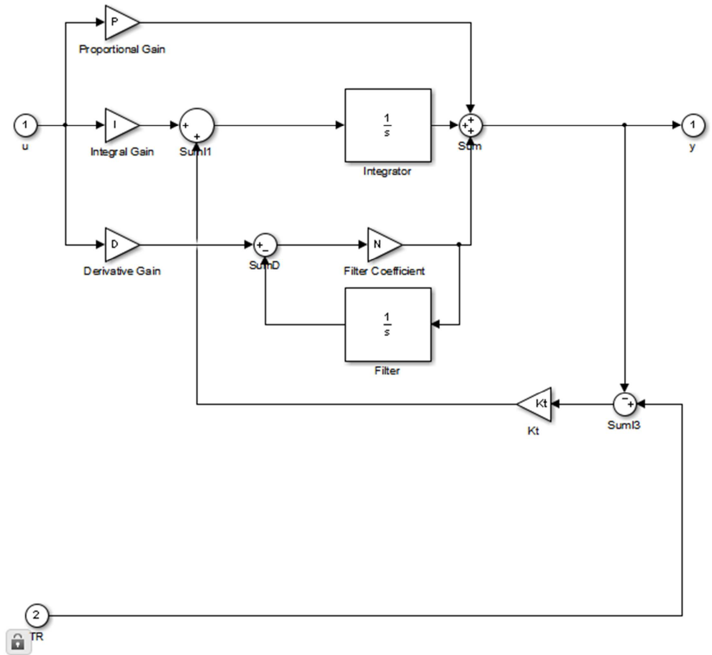


Figure 120, Integral tracking via back-calculation principle

A typical override control configuration could be carried out as follows. The “Feedback Value” labeling the integral tracking input. The controller calculating the lower manipulated variable overrides the other controller. Based on the back-calculation principle the controller which is active is not influenced by the back-calculation branch (TR-y=0). The non-active controller operates in passive mode. The integrator of the passive controller is tracked permanently. So an override process between the controllers is executed inherently bump less.

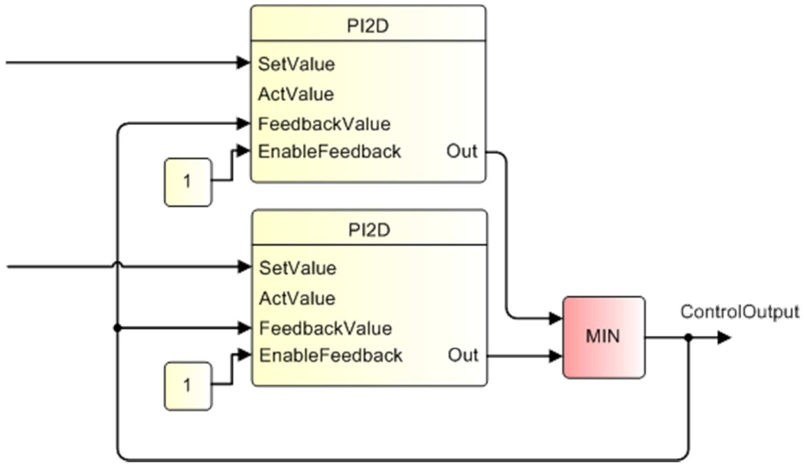


Figure 121, override control via back-calculation principle

With override control, there is only ever one PID controller connected to the actuating element, and the second controller will not be able to reach the setpoint due to the broken connection with the actuating element. For this reason, the I portion will load until the saturation limit is reached (if the windup protection is enabled).

Important:

Back-calculation principle delivers anti-windup inherently.

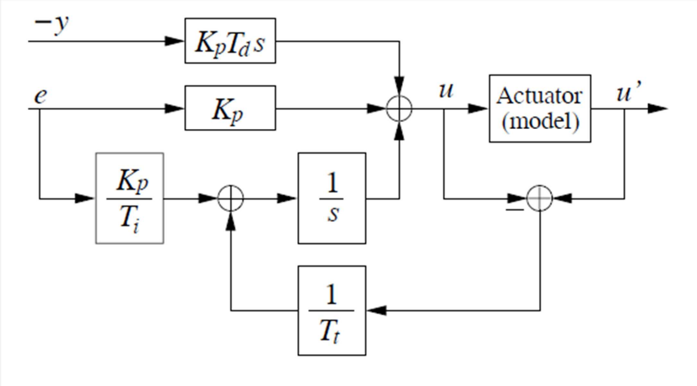


Figure 122, Anti windup scheme via back-calculation principle

If the actuator, actuator-model or saturation-block differs from original manipulated variable u the integral part is debilitated by back-calculation branch. The correction of integral part could be done dynamically via time constant Tt.

- Override controls with "external reset feedback" input, Page 112) on both controllers can enable a step-free changeover to be achieved, as the error cannot be integrated with an inactive controller.

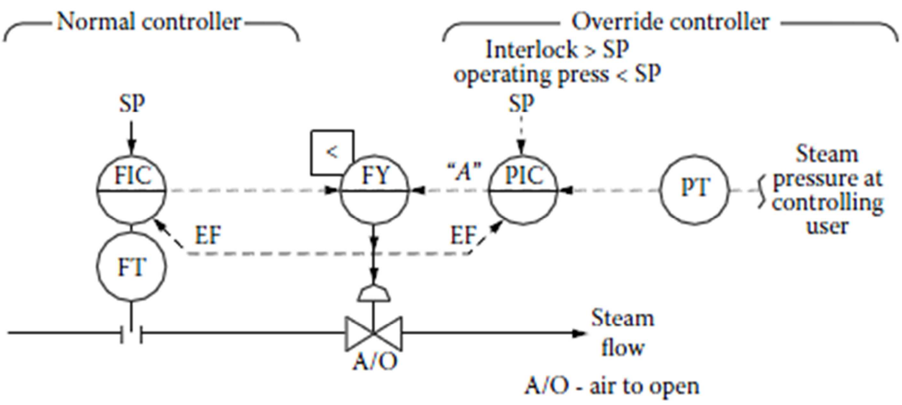


Figure 118, Flow control of a steam pipe with pressure limitation

- Heat exchanger for the condensation of steam using river water. Normally, the condensate temperature is controlled by restricting the river water supply. River water in the primary circuit runs the risk that so-called fouling will occur in the heat exchanger as of a certain outlet temperature threshold, which would reduce the performance of the heat exchanger or put it out of operation for a certain time. For this reason, the condensate temperature control is triggered by the flow temperature control as of a certain temperature limit. Now the primary circuit is controlled instead of the flow quantity of the secondary circuit. This means the outflow temperature of the river water can be reduced. The condensate temperature will however fall below the setpoint of the triggered condensate temperature controller, but fouling will be prevented in the heat exchanger.

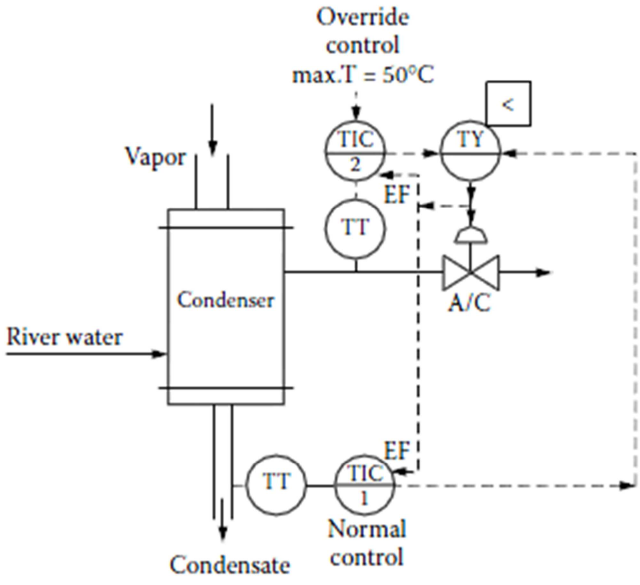


Figure 119, Anti-fouling through override control at the heat exchanger

- Compression systems are primarily controlled according to throughput, but the pressure is monitored as well. As of a certain pressure, the throughput control of the pressure control is triggered, and this also has an effect on the motor speed of the unit.
- Plant steam networks with high and medium pressure levels are primarily stabilized by the medium pressure controller in order to correct fluctuations in use. If, for example, a certain pressure is exceeded in the high pressure unit, the medium pressure controller will be triggered by the high pressure controller and the pressure in the high pressure network will be reduced or limited. The pressure in the medium pressure unit will increase analogously.

16.3 Override control with “back calculation” input

An industrial PID-controller usually exhibit an integral-tracking input which could be used to do override control applications. Integral tracking could be achieved via the so called back-calculation principle pictured in following figure.

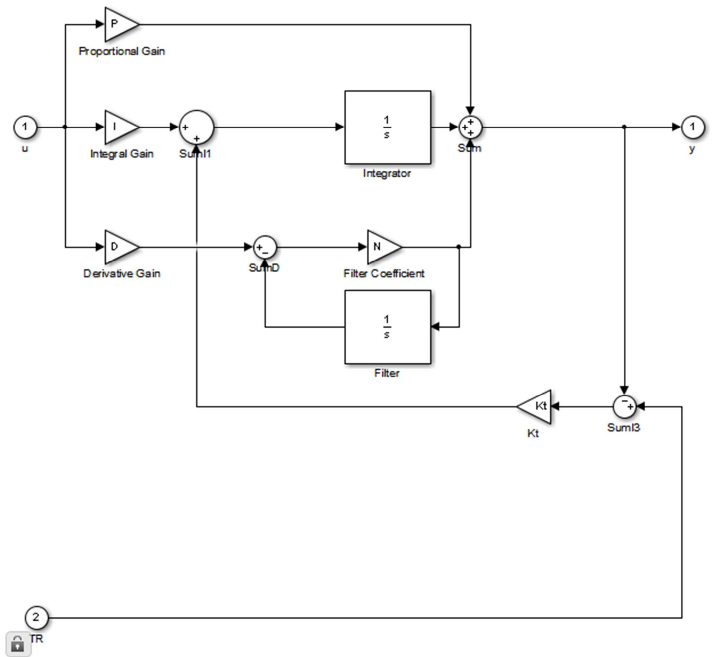


Figure 120, Integral tracking via back-calculation principle

A typical override control configuration could be carried out as follows. The “Feedback Value” labeling the integral tracking input. The controller calculating the lower manipulated variable overrides the other controller. Based on the back-calculation principle the controller which is active is not influenced by the back-calculation branch (TR-y=0). The non-active controller operates in passive mode. The integrator of the passive controller is tracked permanently. So an override process between the controllers is executed inherently bump less.

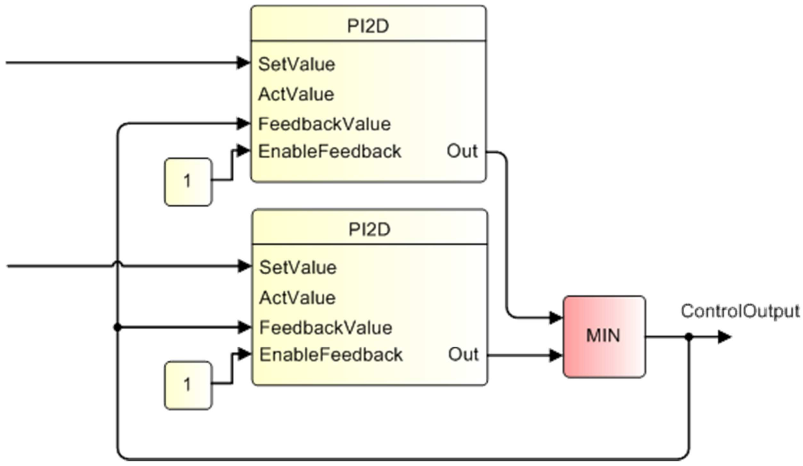


Figure 121, override control via back-calculation principle

With override control, there is only ever one PID controller connected to the actuating element, and the second controller will not be able to reach the setpoint due to the broken connection with the actuating element. For this reason, the I portion will load until the saturation limit is reached (if the windup protection is enabled).

Important:

Back-calculation principle delivers anti-windup inherently.

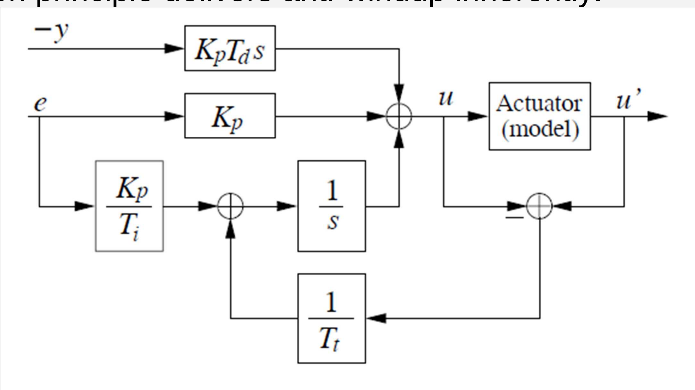


Figure 122, Anti windup scheme via back-calculation principle

If the actuator, actuator-model or saturation-block differs from original manipulated variable u the integral part is debilitated by back-calculation branch. The correction of integral part could be done dynamically via time constant Tt.

16.4 Override controls with "external reset feedback" input

Override control could be also done with an so called “external reset feedback” input which often could be found in US-regions.
The following assumes a PI controller, as the D portion does not contribute to the integrator saturation. In order to deal with this problem and design a bumpless changeover process, a so-called “external reset feedback“ input or external tracking of the I portion can be used, which tracks the I portion with a delay. This characteristic can be achieved through the following structural diagram.

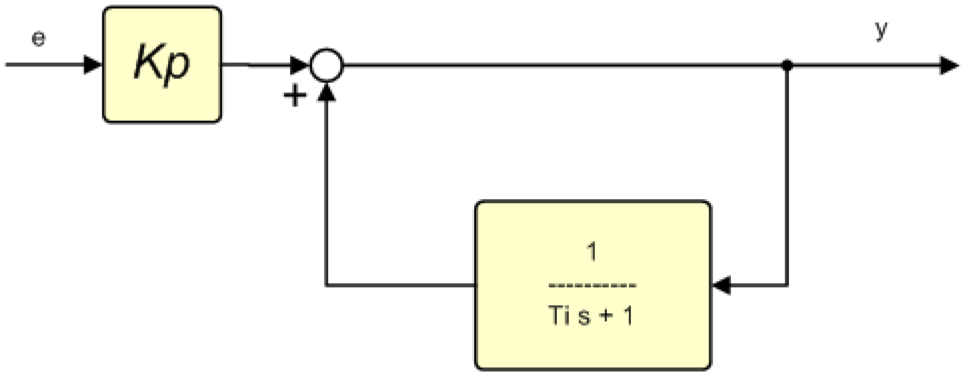


Figure 123, Alternative PI controller structure

The block circuit diagram represents for now a standard PI controller, which is characterized by the following transfer function

$$PI(s) = \frac{y}{e} = K_p\left(1 + \frac{1}{T_n s}\right).$$

The advantage of this structure is the possibility of connecting the input signal of the PT1 system (time-constant integral action time Tn) with the output Ext. y of the selector of the override control rather than with the manipulated variable of controller y.

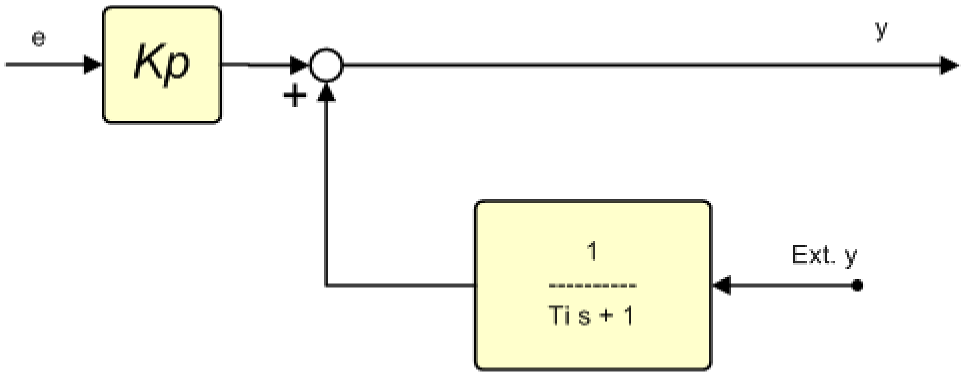


Figure 124, External reset feedback input

As a result, the integral portion will not go into saturation when the controller is inactive. This is because if the controller was selected by selection, the feedback signal is the only signal and otherwise the feedback variable is the output of the controller which is already active.
The blocked controller will therefore be tracked with the output of the active controller. The tracking speed depends on the integral action time of the integrator. The greater the integral action time Tn, the slower the tracking. Too slow tracking can cause the problem of the integrators of the two controllers that lead to the selector showing major differences in their loading. This means that the slower controller can be selected incorrectly as the output of the selector cannot be tracked quickly enough. For this reason, it may be necessary to temporarily reduce the integral action time of the inactive controller so that the active manipulated variable can be tracked at the appropriate speed.

Important:

Compared with the direct tracking of the I portion (without delay=> not back-calculation principle), the delayed integral action time will result in a less abrupt response by the controller. Therefore, noisy measuring signals for example will be lowpass filtered by the integral action time. This lowers the risk of the controllers being repeatedly switched between in case of override control.

17 DYNAMIC DISTURBANCE COMPENSATION CONTROL

If disturbance variables are to be corrected by the controller, there must always first be a control difference so that the controller is able to make the correction. If a controller is also to be configured so that it is faithful to the reference variable, the characteristics relating to the disturbance variable control behavior also need to be valued lower. If the disturbance variable is unknown or cannot generally be measured, the user often doesn't have another choice when it comes to basic control technology methods.
If impacting disturbance variables can be measured, it is possible to compensate for these directly via a correction element by generating a feed forward signal. The control signal is changed by the feed forward control to the effect that the measured disturbance impacting the process is prevented from the outset from having a negative impact on the system.

The concept of disturbance variable compensation assumes there is a measurable disturbance variable that can be directly compensated for by known system dynamics and feed forwards controls.

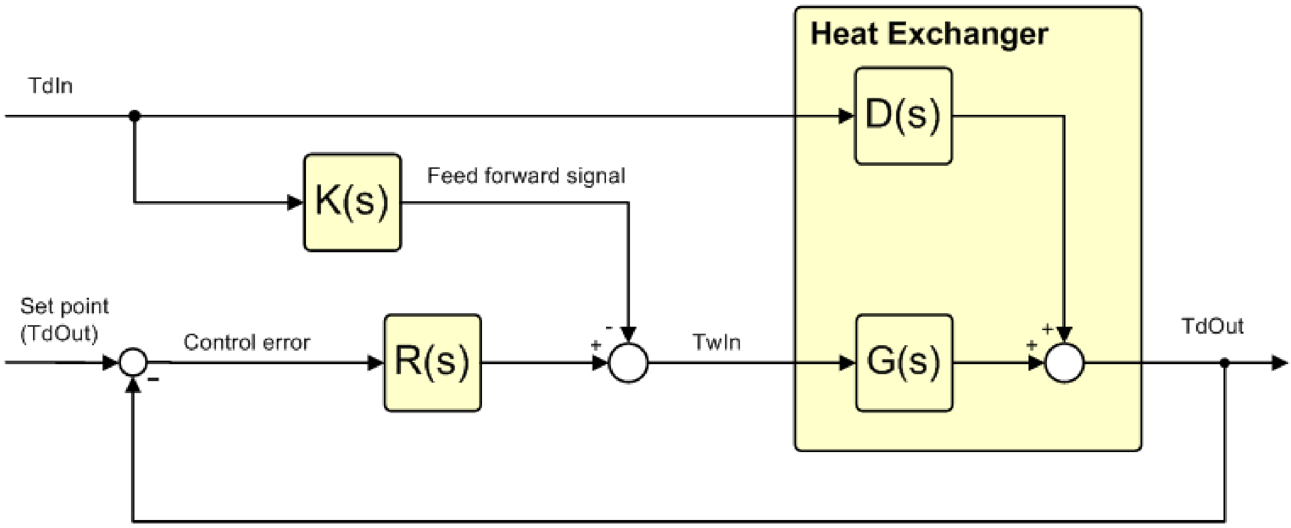


Figure 125, Principle of dynamic disturbance variable compensation

The disturbance variable transfer function with associated disturbance variable compensation can be formulated as follows

$$\frac{\hat{x}}{\hat{z}} = \frac{D(s) - K(s) \cdot G(s)}{1 + R(s) \cdot G(s)}$$

Equation 64, Disturbance variable transfer function

. Assuming that the disturbance should not have any impact on the controlled variable x, the following correlation results

$$D(s) - K(s) \cdot G(s) = 0.$$

Equation 65, Simplification of disturbance variable transfer function to calculate the compensation dynamics

The compensation dynamics K(s) can be calculated as follows using the disturbance variable and control system transfer function

$$K(s) = \frac{D(s)}{G(s)}$$

Equation 66, Compensation dynamics of dynamic disturbance variable compensation

.

Typical uses for disturbance variable feed forward:

- Temperature control of an industrial boiler. The flow in the furnace is the measured disturbance variable which is corrected to the output of the temperature controller via the compensation element.
- Level controls. The extraction can be fed back as a disturbance variable to the output of the tank load via the compensation element.

18 LEAD / LAG ELEMENT

The following section deals with the correction elements (lead/lag element), which can be used to deal with control tasks like the controllers outlined above. Correction elements differ from controllers in terms of process design. Correction elements can be used above all to correct the behavior of the open chain, which can also enable the behavior of the closed loop to be influenced. Correction elements very often show a static gain of 1 and can be used to influence the phase shift of systems. Controllers work with fed-back signals and are used for influencing the closed loop statically and dynamically.

18.1.1 Lag element

The phase-lag correction element has the following transfer function

$$G(s) = \frac{T_d s + 1}{Ts + 1}, mit\ T > T_d$$

Equation 67, Transfer function of lag element

The transfer function works between the frequencies of ω1=1/T and ω2 = 1/Td and there it generates a phase-lag (integrating character). The amplitude response is reduced in the high-frequency range.

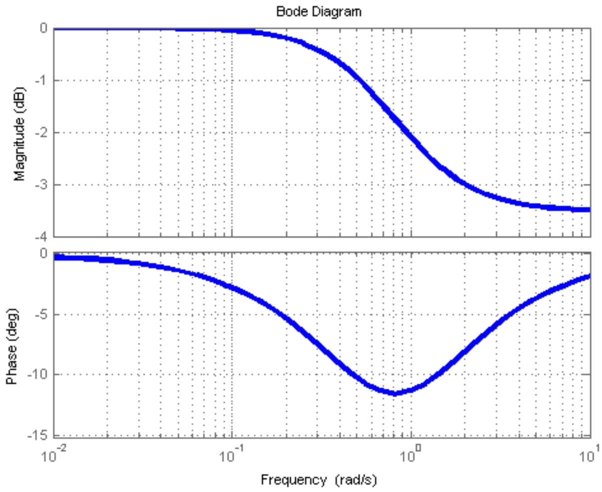


Figure 126, Bode plot of lag element

The lag element can be used to increase the stationary precision. Unlike with the I element, this does not have a destabilizing effect. It can be used to reduce a residual control deviation. The phase response is only subject to a slight change by this (unlike with the I element) and consequently stability problems can be prevented. However, this type of correction element cannot prevent the residual control deviation.

18.1.2 Lead element

The phase-lead correction element has the following transfer function

$$G(s) = \frac{T_d s + 1}{Ts + 1}, mit\ T < T_d$$

Equation 68, Transfer function of lead element

The correction element boosts the phase between ω1 = 1/Td and ω2 = 1/T (differentiating character). The amplitude response is reduced in the low-frequency range.

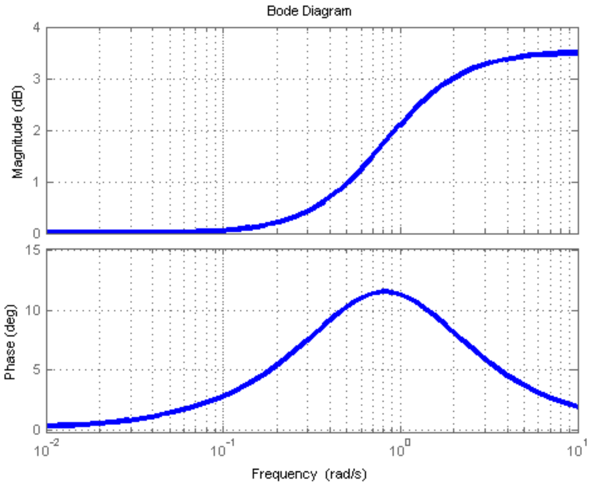


Figure 127, Bode plot of lead element

The lead element has a stabilizing effect (increases the phase margin) similar to a D element, but with a lower noise amplification at high frequencies. Lead elements can be used to increase the attenuation and reduce the overshoot.

18.1.3 Characteristics of lead/lag elements

A lead element has similar characteristics to a PD controller. Unlike the PD controller, measuring noises for example are not amplified so strongly by the additional pole in the lead element. A lead element is mainly used to increase the attenuation behavior of control loops.
A lag element is comparable to the behavior of a PI controller. Unlike the PI controller, the lag element can also be integrated as a passive element in the control loop. The lag element is used to reduce the stationary control deviation so that the system dynamics are not so strongly influenced. The combination of both elements is very similar to a PID controller (stationary and transient control errors are reduced).

19 CASCADE CONTROL

A control structure in which several control loops are layered on top of one another is called a cascade control. Cascade control utilizes a master controller, which sends the manipulated variable as a setpoint to the downstream slave controllers. This type of connection results in control loops encapsulated in one another that are able to more quickly compensate for any disturbances (z1, z2) that occur in the respective control loop. Connecting control loops in series requires that the lower-level control loops feature successively higher dynamics (e.g. flow rate control with higher-level temperature regulation or MPC as master controller supplies setpoints for lower-level PID control loop).

- With a cascade control it must be ensured that, if there are manipulated variable limitations of the slave controller, the integral part of the master controller is shutdown.
- In addition, the adjusting range of the master controller must be coordinated with the setpoint range of the slave controller.
- If the slave controller is not in automatic operation (with an external setpoint), the master controller must be tracked so that the I portion cannot be integrated in the master controller. In order to ensure a bumpless switching back again at a later time, the manipulated value of the master controller must be tracked with the current setpoint of the slave controller.

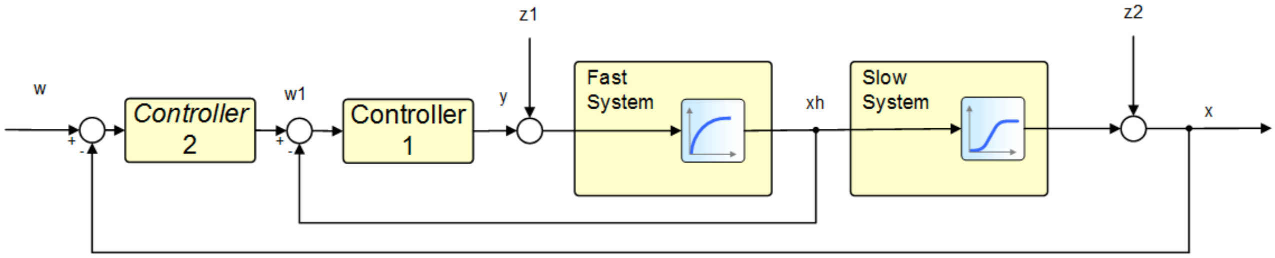


Figure 128, Cascade control principle

An oven should maintain a certain temperature and for technical safety reasons the temperature of the electric heating element should not exceed a certain setpoint.

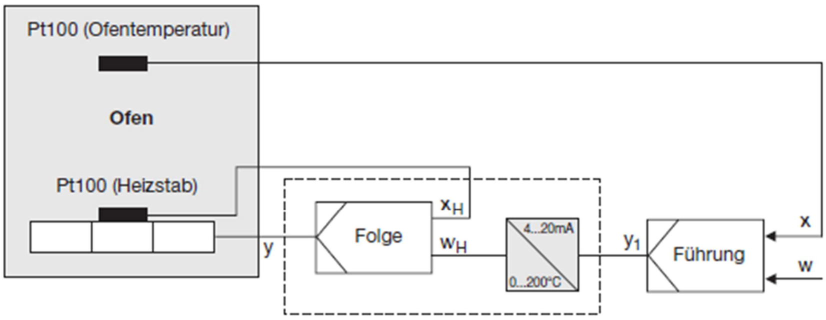


Figure 129, Limitation of heating element temperature with higher-level oven temperature control

The dynamics of the heating element are considerably quicker to assess than the dynamics of the entire system. For this reason, the temperature limitation of the heating element can be assumed by the internal controller (slave controller). The output signal of the master controller provides the setpoint of the slave controller. 100% control signal of the master controller means for example a 200℃ setpoint for the heating element temperature, which is controlled as a measurement variable in the slave controller. By adjusting the control range of the slave controller, the temperature limit of the heating element can be set.

> **Important:**
>
> The advantage of a cascade control (in this case) is that the temperature of the heating element can stay under control. The internal loop, which has higher dynamics, can be used to prevent the controlled variable from exceeding the setpoint.

The control loop is configured in an inwards to outwards direction. The master controller is switched on manually and the slave controller can for example be set by an auto-tuning process. The auto-tuning process can also be used in the outer control loop. The slave controller remains in automatic mode in this case.
The slave controller should respond more quickly than the master controller. For this reason, a pure P or PD controller is very often used as a slave controller. Whether or not the slave controller can compensate for the controlled variable exactly is of secondary importance in this case, because this task can be performed by the master controller.

Typical uses for cascade control:

- Master controller for fill level control with slave controller for controlling the intake or outlet quantities.
- Loading process of energy stores (limited capacity).
- Position control with lower-level speed control.
- Temperature control with lower-level flow control.
    - Temperature control of a container with subordinate flow control of the heating medium.
    - Master controller for temperature control of a furnace with slave controller for flow control of the fuel quantity.
    - Outlet temperature control of the primary circuit of a heat exchanger as a master controller, inlet quantity control of the secondary circuit as a slave controller. The advantage lies in the compensation of pressure fluctuations in the inlet area of the secondary circuit. With a cascade control, the flow controller can respond quickly. With a conventional control, the temperature controller with the delay in system dynamics responds between the secondary inlet and primary outlet.

19.1 Tuning

The controller can be configured in the following order:

- The master controller must be switched to manual mode.
- The slave controller can then be configured.
- Then the master controller is switched to automatic mode. It should also be ensured that the slave controller receives the setpoint of the master controller.
- The master controller can then be configured.

Cascade controllers can be configured using conventional methods (Ziegler/Nichols). The slave controller should be configured so that no or only few oscillations occur in the presence of disturbances, otherwise the reference variable controller will also be affected by the fluctuations.

19.2 Identification of applications

Cascade controls can be used if the control performance from a standard control loop is not acceptable or if an additional measurement is available, which can be used for cascade control.
The starting point is a primary measurement of the controlled variable, which can be included in a control loop through a controller with an actuating element.

> **Important:**
>
> The system part is an ideal candidate for cascade regulation if, for example, the process is subject to disturbances that can be recorded by an additional (intermediate) measurement. It enables an internal loop to be included in the control structure. The dynamics of the internal subprocess must always be faster than the dynamics of the outer subprocess.

The next example assumes a process with one manipulated variable and 2 disturbance variables.

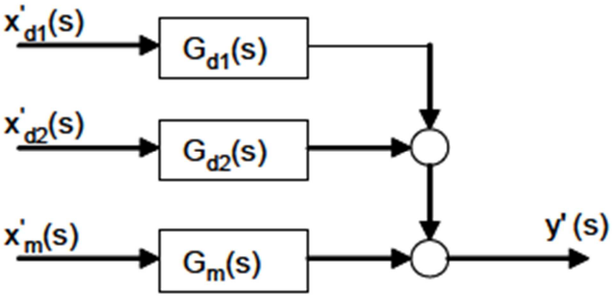


Figure 130, Identification of cascade controls

In very general terms, the output variable y‘(s) is dependent on all three inputs.

The following step tests whether the process can be divided into two parts. The connection between the subsystems can be made through the measurable process variable xi‘(s). The process description using the transfer function was only divided in this case (Gm(s) = Gm1(s) Gm2(s)) because the process had not changed.

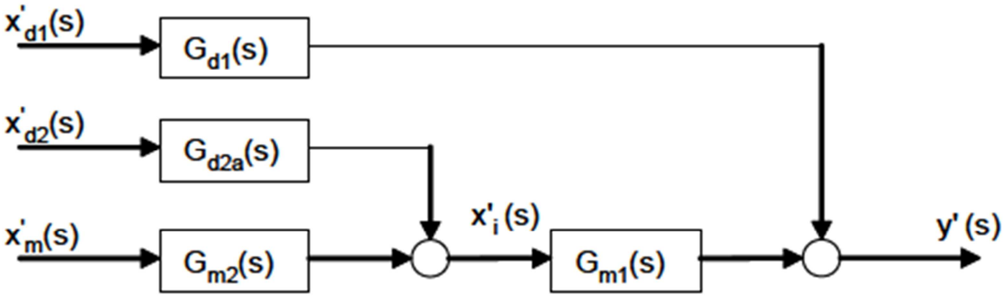


Figure 131, Introduction of auxiliary controlled variable in cascade control

The auxiliary controlled variable x‘(s) of the process responds to the disturbance variable xd2‘(s) before the controlled variable y’(s) can change, and it is also influenced by the manipulated variable x’m(s). This structural design enables an additional control loop, which can compensate for the auxiliary controlled variable quicker. In general, the auxiliary controlled variable must meet the following requirements.

- Dependency on disturbance variables which significantly affect the process.
- There must be a link with the main controlled variable.
- There must be a link with the manipulated variable.
- The time constants of the primary controlled systems should be greater than the time constants of the secondary controlled systems by a factor of at least 4.

19.3 Advantages of cascade control

Non-linearities can be compensated for by the slave controller, using additional function blocks which can be added into the closed loop of the slave controller. Thus, for example, an MPC with linear process models can be operated as a master controller in the cascade with non-linear actuator behavior. Non-linear valve curves can, in the same way, be compensated for by function blocks in the slave control loop. The master controller can then be configured without taking into account non-linearities.

A cascading approach in case of level controls (master controller for level, slave controller for flow) can only achieve its aim if the dynamics requirements for the two controlled variables are both different. With buffer tanks, downstream pressure fluctuations are possible for example (e.g. the outlet pressure fluctuates). These pressure fluctuations cause pressure drops in the control valve which also cause the flow to fluctuate. If the flow of the valve is controlled, these disturbances, which are identified by minor changes in the buffer level, are quickly suppressed. If there is no flow control, the level controller will compensate for the disturbance. In case of disturbances in the flow, the parameter settings of the level controller should be as strong as possible. However, if the in-flow quantity of the buffer tank changes, the parameter settings of the level controller should be as weak as possible. A cascade control enables both tasks to be carried out as required and without making any compromise.

In addition, cascade control offers these additional advantages:

- Use of a simple, standard PID controller.
    - No process models required (no identification required)
    - PID auto-tune can be used (no great time or effort required)
- Clear, transparent control method (open the cascade and all superordinate controllers are put out of operation).
- Increases in control performance.
- Improved response speed of the controlled system (in relation to disturbances).
- Non-linearities can be compensated for in the internal loop.
- Masses or energy (flow control) can be controlled better.

19.4 Slave controller without automatic mode

The wiring between master and slave controller is shown in Figure 132, Principle of PID connection for cascade control. The slave controller signals to the master controller through the binary output YTrkReq that it is not in automatic operating mode. The master controller receives this information via the input yTrkEnable and tracks its manipulated value against the input yTrkValue.
yTrkValue is linked to the output yTrkRequestValue of the slave controller, which issues the actual value of the slave controller. If a separate processing of the setpoint is required in internal and external mode, the tracked values must be manipulated by external wiring outside of the controller core. This wiring should enable the bumpless switching back of the slave controller to automatic mode.

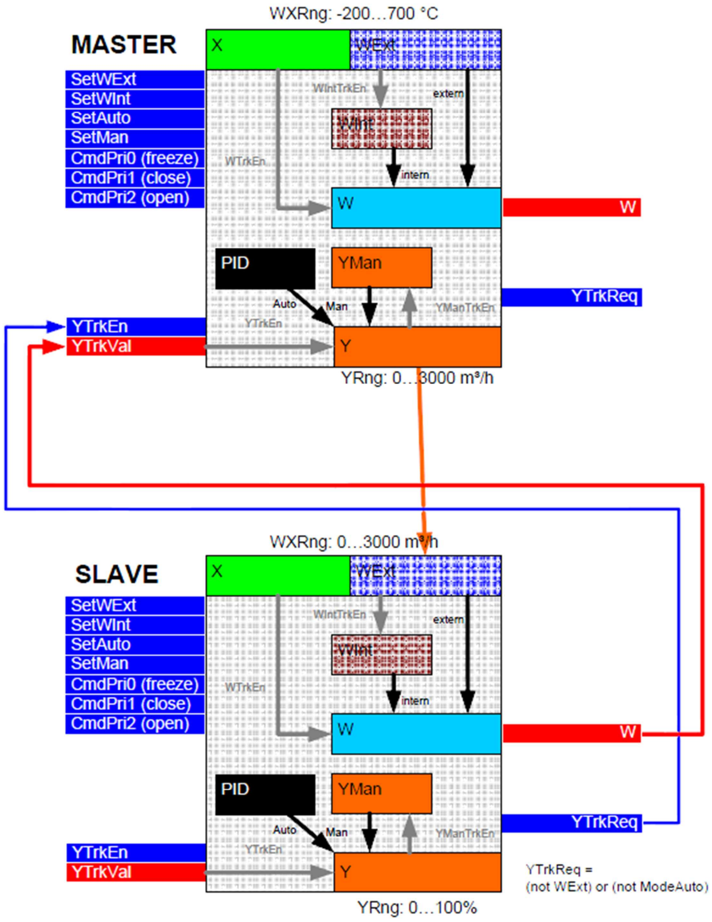


Figure 132, Principle of PID connection for cascade control

19.5 Windup protection in the cascade

Implementing windup protection in cascade control requires the BOOL outputs IntegrationPartStoppedMin and IntegrationPartStoppedMax on the slave controller to the connected with the inputs HoldIntegrationPartMin and HoldIntegrationPartMax on the higher-level controller. This input can be fed back to the master controller, which blocks the integrator of the master in an upwards (to downwards) direction starting from the proportional (inverting) transit behavior of the cascade.

The image below shows the connections needed to ensure the bumpless startup of the cascade control with windup protection.

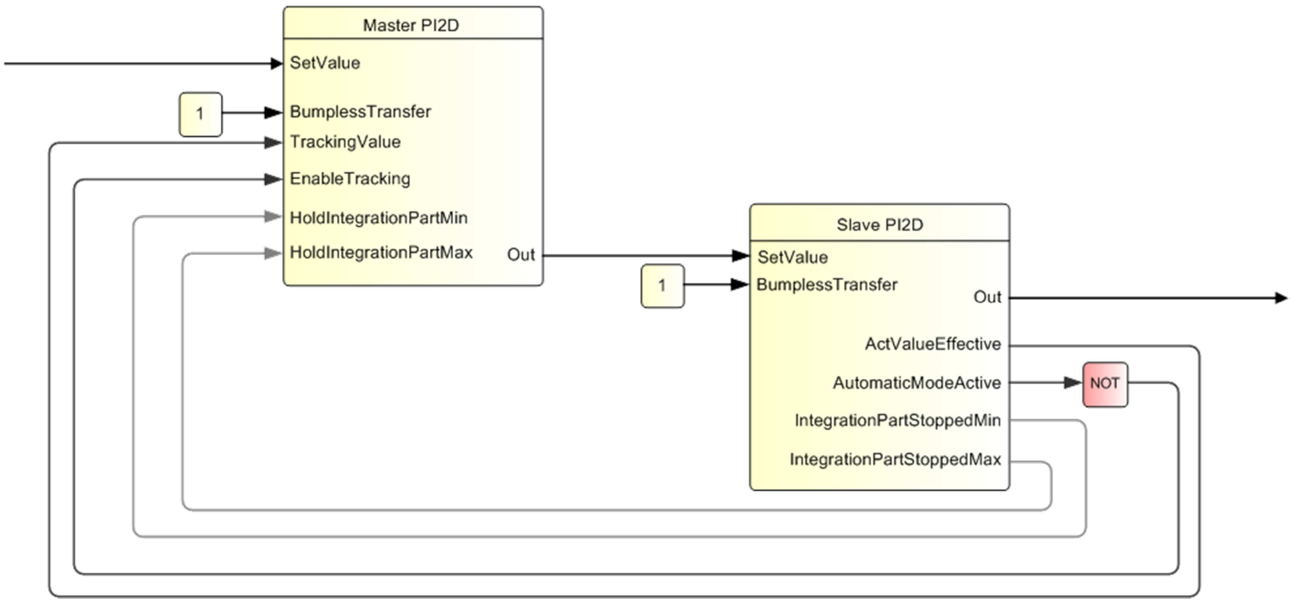


Figure 133, Principle of bumpless start up of cascade controller

The connections ("CFC") between the PID controllers are already prepared and therefore they do not need to be implemented manually.

19.6 Comparison of dynamic disturbance variable compensation with cascade control

The two control methods show the following similarities.

1. The control strategy requires an additional measurement and an additional dynamics function block.
2. The setpoint of the primary control loop remains nearly unchanged in the event of a disturbance.
3. The primary control loop serves mainly as a setpoint control.
4. The manipulated variable of the control loop (cascade master controller) is not affected.

However, the control processes differ in the following ways.

1. The cascade control reduces the extent of disturbances to the controlled variable. In case of disturbance variable compensation, only the effect of a certain disturbance variable is reduced.
2. Cascade controls make do with an additional standard PID controller. Compensation dynamics for disturbance variable feed forward controls are adjusted specifically for the requirement.
3. In cascade control, the additional measurement relates to the controlled variable of the slave controller, which is influenced by the disturbance variable and manipulated variable. Disturbance variable feed forward controls require the measured value of the disturbance variable occurring, which cannot be influenced by a manipulated variable.

20 PID CONTROL WITH OPERATING POINT-DEPENDENT PARAMETER CONTROL (GAIN SCHEDULING)

Gain scheduling is used in control loops in which process parameters change so severely over the target control range that the parameter settings of the controller need to be adjusted in relation to the current operating point.

Compared with the mechanical example of level control, see (Figure 56, Mechanical water level control with a P-controller), the cylindrical tank is no longer vertical but in a horizontal position. The correlation between the manipulated variable (in-flow) and the controlled variable (level) is no longer linear due to the curvature of the tank. The non-linearity of the controlled system can be taken into account using "Gain Scheduling" (operating point-dependent adjustment of the PID parameter). Beforehand, several distinctive operating points of the process are defined. The controller will then be configured for each operating point. The PID parameter sets can be supplied to the "Gain Scheduling" function block in connection with the operating point. Depending on the current operating point, the "Gain Scheduling" module can calculate the corresponding PID parameters through linear interpolation between the operating points dependent on the current operating point and forward this directly to the PID controller.

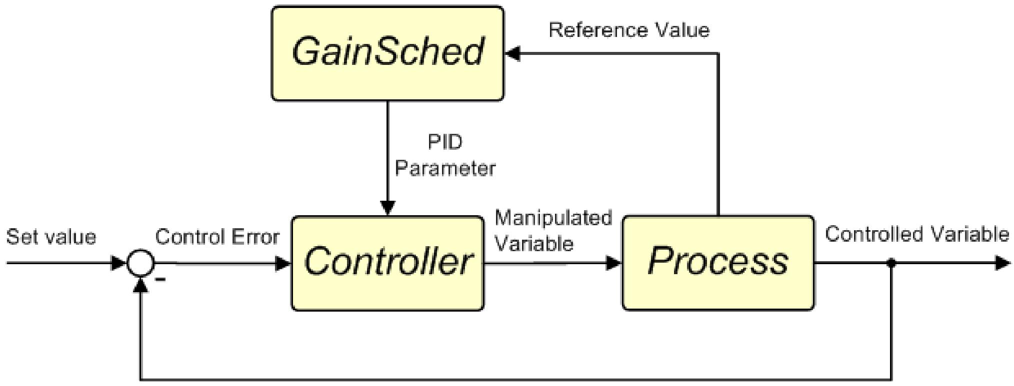


Figure 134, Gain scheduling principle

Important:

Gain scheduling can only be used if the PID controller can take on the new PID parameter in every sampling sequence.

If the controlled variable or external reference values are below the first working point or above the last working point then the parameters of the next working point will be transferred to the controller unchanged.
If a signal on the *In* reference input is between between two configured reference points, then the PID parameters (*Gain*, *IntegrationTime*, *DerivativeTime*) will be linearly interpolated based on the configured reference values.

Typical uses of gain scheduling:

- Ph value control. The gain factor of the PID controller is constantly adjusted piecewise in relation to the PH value according to the acid-base titration curve.
- Control of non-linear integrating processes (cylindrical tank in horizontal position, spherical tank)
- Control of heat exchangers operated in several operating points (change in load results in non-linear behavior).
- Control of tanks, the discharge of which can only be controlled by gravity (the process gain changes with the fill level).
- Control of reactors in which chemical reactions occur and which show non-linear behavior.
- Temperature control of pressure boilers.

Important:

In many cases, it is enough to linearize the measurements to improve the behavior of a control loop without requiring gain scheduling.

21 SMITH PREDICTOR

A Smith predictor is a control structure which accesses additional information supplied by a stored model of the system.

Important:

The Smith predictor controller configuration is suitable for processes with high dead times.

The additional supply of information enables the controller to be configured offensively, which increases the control performance in comparison with the conventional controller configuration.

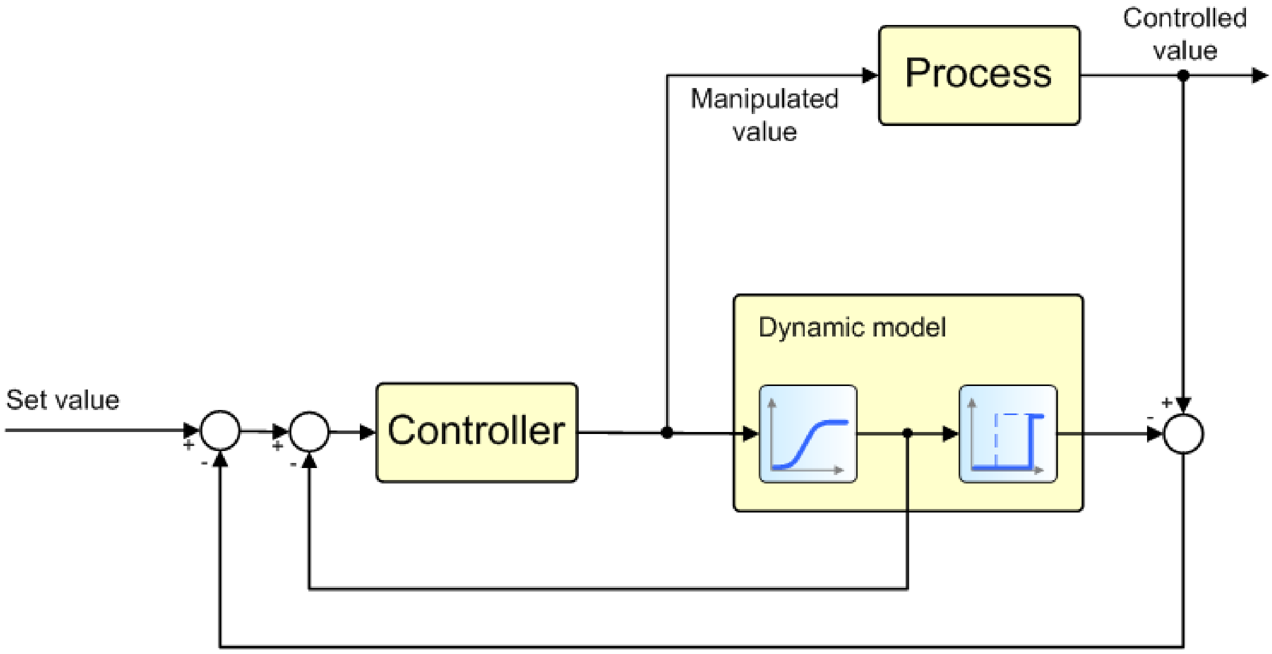


Figure 135, Smith predictor principle

The following process models are available for configuring the dynamic process model (P, PT1, PTDT1, PT2, PTDT2, I, IT1, IPT2DT, PT3, DT1, dead time).

21.1.1 Taking into account dead time

If the control of the integral portion is delayed by the dead time, the control performance can be increased. This is because the controller can be changed by the tuning using the model information on the dead time.

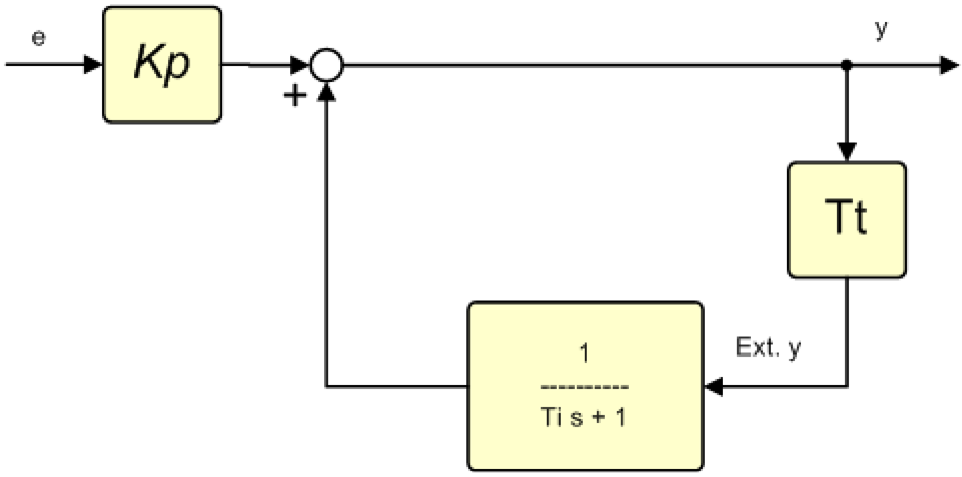


Figure 136, External reset feedback input dead time components

22 MODEL-BASED PREDICTIVE CONTROL

Model-based predictive control is based on an optimization method that cyclically minimizes predicted control errors. Predicting control errors is backed up by the use of a model. To do so, pulse and/or step responses are saved in the control function block. The calculation is based on the assumption that the model is linear and time-invariant. Through the use of the dynamic model, the future course of the controlled variables can be expressed as a function of future changes to the manipulated variables.

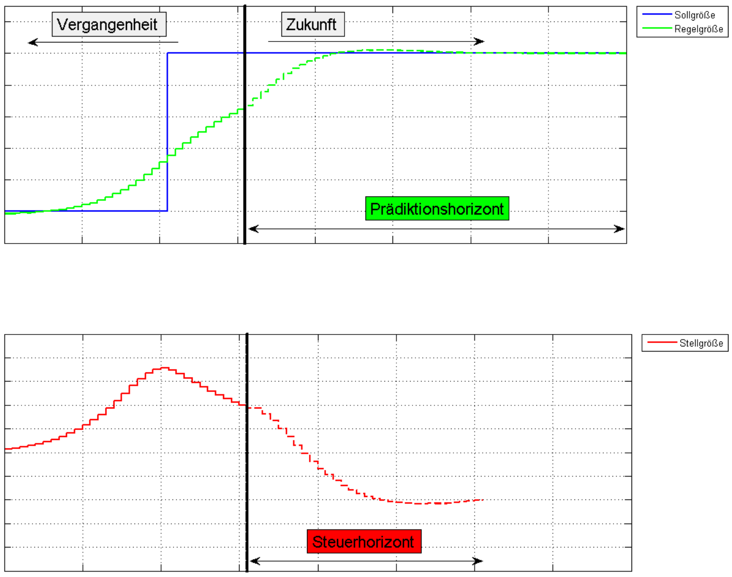


Figure 137, Principle of model-based predictive control

Together with the desired setpoints of the controlled variables, the prognosticated quadratic control error can be formulated as a function of the future changes to the manipulated variables. This formulation is then provided to the optimization procedure and minimized. While determining the solution, it is possible for constraints that affect the process such as manipulated variable limits of actuators to be taken into consideration directly. The minimum, future quadratic control error then results from a clearly determinable sequence of future changes to the manipulated variables. The number of sampled values used to calculate the sum of the predicted quadratic control error can be configured with the prediction horizon. Manipulated variable changes are usually calculated using several sampling sequences over time (control horizon). Only the first value is used to output the manipulated variable before the optimization process restarts, taking the latest measured values into consideration.

22.1 Advantages of the use of MPC controls

- Highly suited to MIMO systems. Global optimization and prioritization of controlled variables, meaning that complex decoupling networks are no longer required.
- Manipulated and control variable restrictions are taken into account directly in the algorithm. There is also the possibility of weighting the manipulated variable dynamics individually.
- Ideal for systems with high dead times (significantly greater than system time constants) or for processes whose step response first responds in the "wrong" direction ("inverse-response" behavior).
- No special design required for disturbance variable feed forward control, as this is taken account of directly in the algorithm (considerably less complicated than with PID controllers).
- Economic advantages due to improved product quality or increased throughput. The prediction of the controller or operating mode, which can be based more closely on previously defined process limits

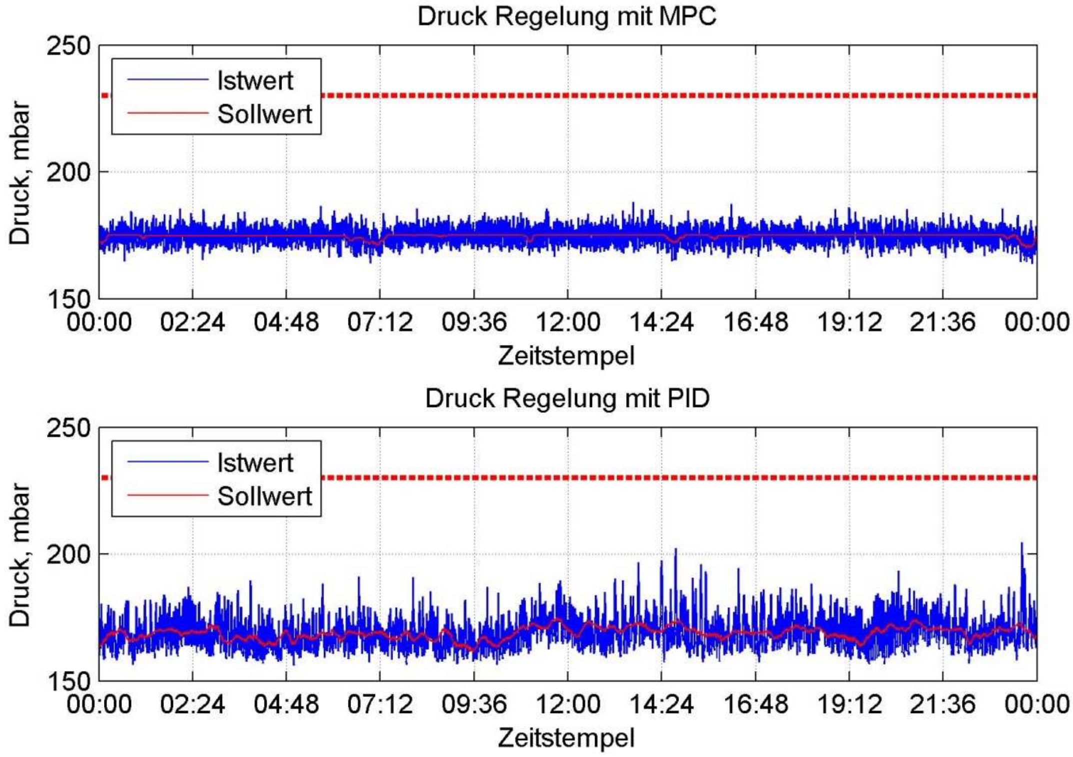


Figure 138, More efficient process control by MPC

Figure 138, More efficient process control by MPC compares a PID control with an MPC. Unlike with a PID, the pressure control of the MPC considerably reduces the variance of the controlled variable. This means the setpoint of the process controlled by the MPC is already closer to the red safety shut-off value. Consequently, the process is a lot more efficient to operate.

22.2 Overview of MPC control offered by B&R

- Fully integrated in APROL.
    - Full scope of the process control system can be used at any time (alarm system, reports, etc.).
    - As a data archive, the PCS trend system enables later analyses (process model drift, data export for re-identification).
    - Online performance monitoring. Various indices for evaluating the control performance and classifying model/system deviations are calculated directly in the MPC and can be filed in the APROL trend system (various parameters, prediction errors).
    - Use as a stand-alone solution or use in combination with external process control systems possible at all times (e.g. OPC connection).
- Very low response times between the process variables and controller, as the controller algorithm is operated directly on the real-time hardware (controller level) (process variable update with the most up-to-date measured values).
- Reduced commissioning time
    - Standardized operating images and hardware configurations.
    - Simple configuration of the controller through standard selectable process dynamics.

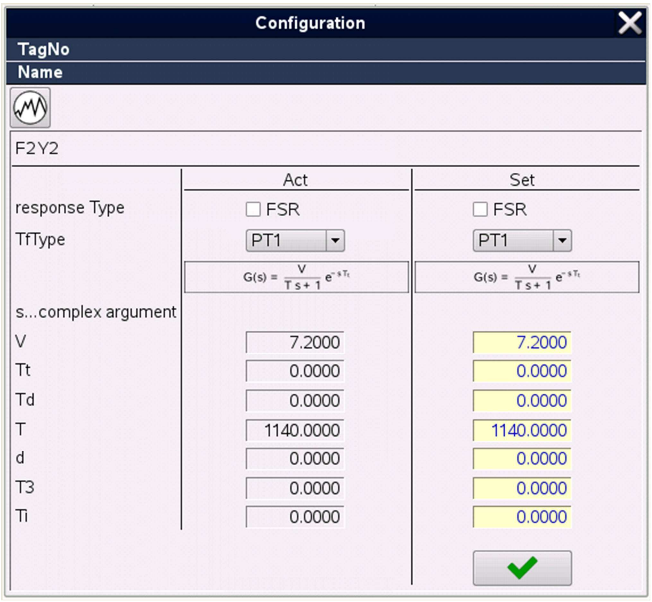


Figure 139, Parameter settings for MPC process model

  - Complete "Hardware in the Loop“ (HIL) simulation option (controller/master computer level) using Matlab/Simulink code generation.
    - Testing of different controller configurations or pre-configuration in the simulation.
    - Training of operating personnel.
- Single/multiple variable control.
  - Control of a single variable system through to MIMO systems with 10 control variables and 10 manipulated variables.

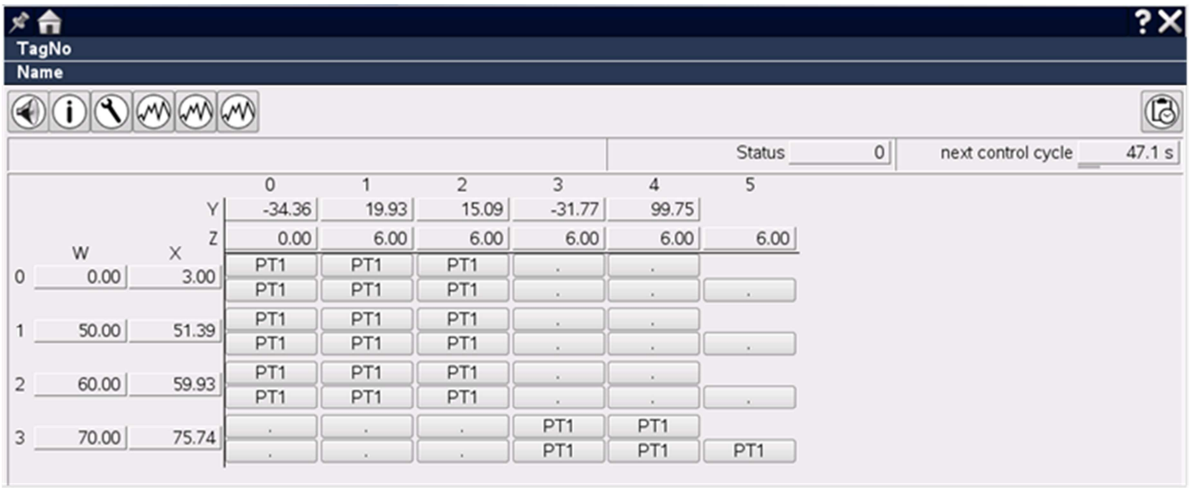


Figure 140, APROL process coupling

  - Suitable as standard either for stable processes with step/impulse response or for marginally stable processes with integrating behavior with an impulse response (e.g. fill level controls). With MIMO systems there is also the option, depending on the controlled variable, to combine the characteristics (stable, marginally stable or proportional, integral) into one controller configuration.
  - Integrated disturbance variable feed forward.
    - Up to 10 measurable disturbance inputs configurable directly in the controller.
    - With a known disturbance variable curve, disturbances occurring in future can be taken into account directly in the algorithm (data from production/maintenance planning, e.g. extraction plans).

- Limitations of the absolute manipulated variables, the manipulated variable changes and "soft constraints" of the controlled variables.
- Structure changes possible in the current controlled operation (update of process dynamics and re-configuration).

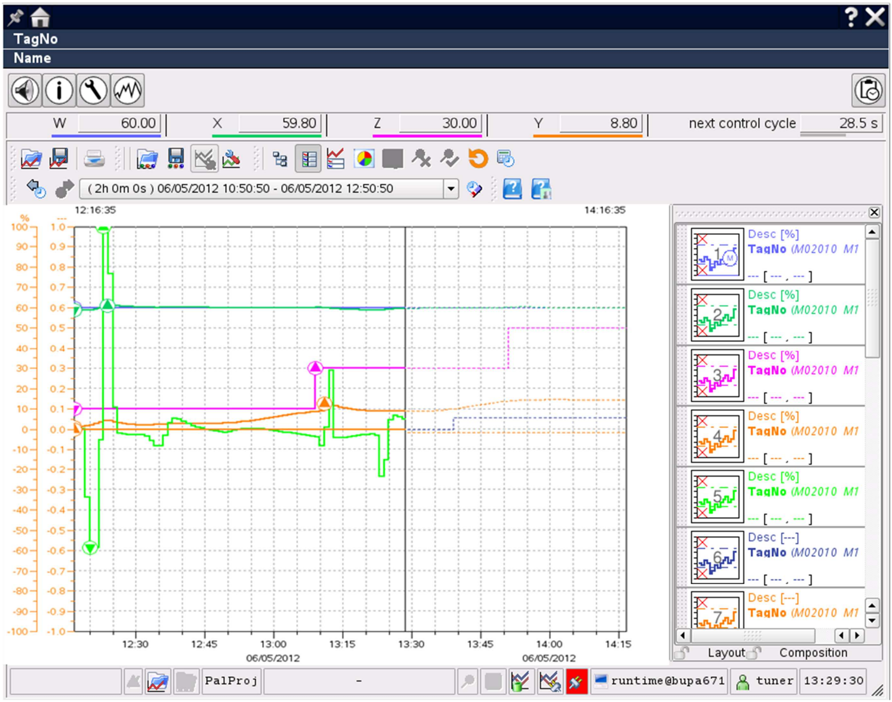


Figure 141, Trend + predictive components

22.3 MPC operating principle

The following information assumes a system with one controlled variable, one manipulated variable and one disturbance input.
The open loop controlled variable, calculated according to the model, results from superimposing the effect of the manipulated variable on the controlled variable onto the effect of the disturbance variable on the controlled variable. Model/System deviations or unknown disturbances (at undefined places in the process) will cause the calculated controlled variable of the open loop to deviate from the measurement. This difference is taken into consideration by introducing an additive disturbance input to the model.
The prediction of the controlled variable can ultimately be illustrated as follows:

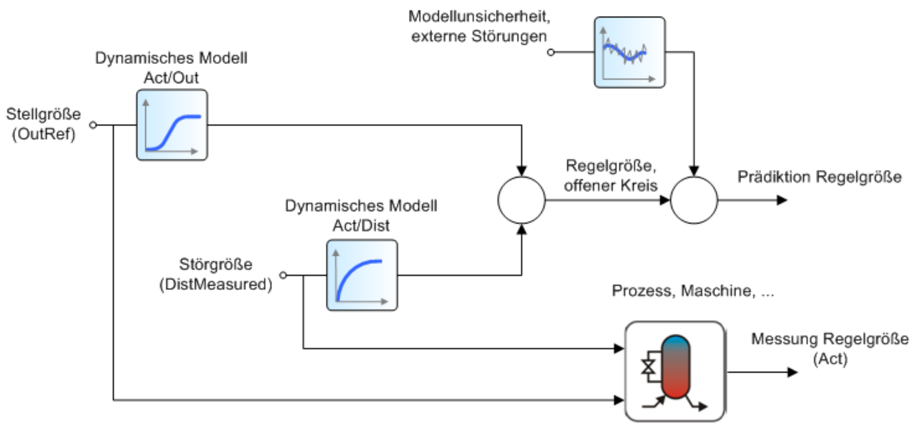


Figure 142, The MPC principle

22.3.1 Optimizing and calculating the manipulated variables

The mathematical prediction of the future course of the controlled variable is a function of how the input variables have behaved in the system in the past as well as the future manipulated variables **y**. The predicted control error can be formulated as follows in conjunction with the desired setpoint curve **w**

$$\boldsymbol{e} = \boldsymbol{x}(\boldsymbol{y}) - \boldsymbol{w}.$$

Equation 69, Predictive control error

The goal is to minimize the quadratic control error. For this reason, the control loop quality factor *J* is introduced, which should be minimized.

$$J = \int_0^{N_x\, T_s} \left([\kappa\, [\mathbf{w} - \mathbf{x}]]^2 + [\lambda\, \Delta \mathbf{y}]^2 \right) \mathrm{dt} \rightarrow \min$$

Equation 70, Quality function of MPC

Parameter κ ("weight error") is used for weighting the quadratic control error $(\mathbf{w}\text{-}\mathbf{x})^2$. The "Y Move Suppression" parameter is also introduced in order to weight the manipulated variable dynamics. Our goal is to keep the control error as minimal as possible while still honoring all constraints. The only unknown influencing variable is the future curve of the manipulated variable $\Delta\mathbf{y}$. Optimization changes the sequence until the control error is minimal.
The control concept can be shown as follows:

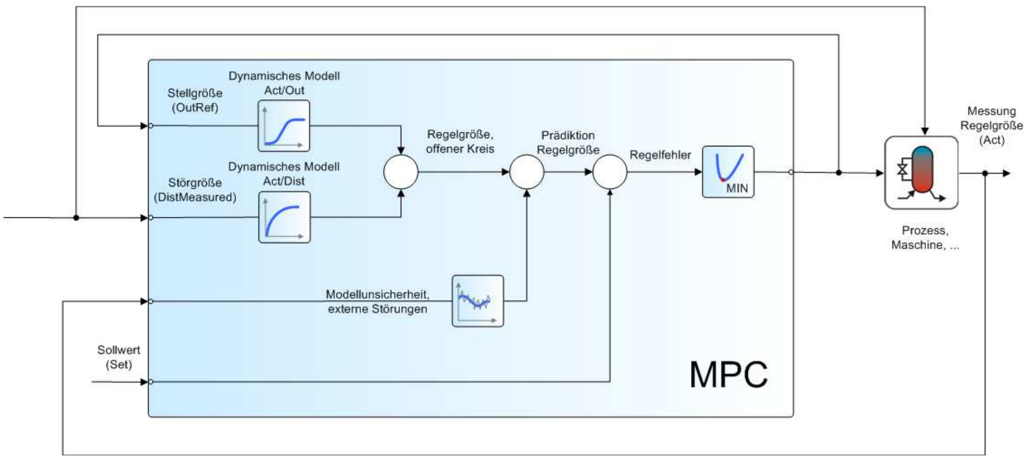


Figure 143, The MPC control concept

The optimal solution delivers a sequence of future changes to the manipulated variable. For the control loop, however, only the most current, first part of the optimal solution is used as the manipulated variable. After the manipulated variable has been output, then the measured values of the process being controlled are reread cyclically (based on the process sampling time) and used to calculate the new prediction (i.e. the algorithm starts back at the beginning).

22.4 System classes

The implemented algorithm can be used for controlling single and multiple variable systems which can have up to 10 manipulated variables with 10 manipulated inputs and 10 disturbance inputs. The use of disturbance inputs gives the option of taking future disturbances into account directly in the algorithm or compensating for them. In general, all systems that are linear/have been linearized and are stable can be controlled.
The nature of manipulated variable calculation can take large amounts of calculation time even in MIMO systems. For this reason, in systems with a high level of dynamics, the minimum possible cycle time

should be tested through a passive use of the controller. The controller is already linked to the system measurements in passive operating mode. However, the manipulated variables are not forwarded to the process. Passive operation is only possible if an existing control concept is already in operation. If the system or system part has not yet had its first operation, a complete "hardware in the loop" simulation can be created. To do this, a corresponding model of the process to be controlled is installed in the process control system (or externally) which is connected to the APROL MPC. This simulation can also test various setting options of the controller to speed up the commissioning of the actual system. Furthermore, the design can also be used for training operating personnel.
The system characteristics of processes that can be controlled with the MPC library from B&R must meet the following requirements.

a. Linear systems not capable of producing steps with a proportional character (PT), with and without dead time. The system dynamics are transferred to the controller by the step response ("Response type" setting = FSR) or impulse response ("Response type" setting = FIR).
b. Linear systems with an integrating character (IT) with and without dead time, with and without a disturbance input. The system dynamics are transferred to the controller by the impulse response (FIR).

The algorithm for the manipulated variable calculation works with a model of the system to be controlled. For this reason, the control performance is very highly linked to the quality of the model. The robustness of the controller is inversely proportional to the level of deviation between the model and the actual process (model/system deviation). The model dynamics are transferred to the controller via a sampled sequence of step or impulse responses. It is also assumed that the necessary step/impulse responses are either obtained through external identification processes or are already present, as no identification algorithm is integrated in the initialization part of the "Model-based Predictive Control" process.

Important:

All disturbance transfer functions must show the same form of discretized system response (FSR, FIR) as the manipulated transfer functions (per controlled variable)

If, for example, a system with integrating character has a disturbance input with PT behavior, the system dynamics must be implemented as FIR in the controller function block in relation to the disturbance. In principle, all linear systems (PT, IT) can be controlled via a stored "FIR" model; however, for this, the setting "FIR" must always be applied in the parameter settings of the transfer functions. The control of mixed MIMO systems (PT, IT) is also possible (in terms of the controlled variables).

22.5 Summary of the calculation process

A quality function is used as a basis for calculating the manipulated variable y, in which the integral squared error

$$\min_{\Delta y}\left[\sum\left(\vec{w}-\vec{x}(\Delta y)\right)^2\right]$$

Equation 71, Minimization of the integral squared error

is minimized. The control error results from the setpoint w and the controlled variable x, which is described as a function of the manipulated variable y. The integral squared error is then calculated using the prediction horizon (X). The "Weight Error" factor can be used to change the priority for each controlled variable, which should minimize the errors through the manipulated variable changes. There is also the option of using λ with the control horizon to influence the dynamics of the predicted controlled variable changes (for each manipulated input).
Future setpoint curves need to be supplied to the respective block for each system output as a data sequence with a length of 200 sample values per controlled variable. The future disturbance variable curves must also be supplied to the function block as data sequences with a length of 200 sample values per disturbance variable. If there is no future data for the respective disturbance input, a data sequence must nevertheless be linked, and in this case this is initialized with 0. In this case, the function to take into account future disturbance variables should be disabled ("Enable Z trajectory" = Off). The calculation of

the current manipulated variables takes place from a time perspective on the assumption that the current disturbance variables of the system are already available (through measurement). This means the disturbance variable which is read before the start of the calculation acts on the process over the period of the sampling time. To input the current disturbance variable, there is a separate input in the function block interface.
The calculation time and use of the dynamic memory increases with the length of the control horizon (see graphical configuration level APROL: Control horizon (Y)).
The "Sample time factor" parameter in the APROL user interface can be used to change the number of function block call-ups of the sample time. The sample time (ST) of the process is set as follows in the controller, depending on the task class cycle time of the MPC function block (Sample time (ST) = Controller task cycle time x Sample time factor).
The "Y Operating point" and "X Operating point" parameters give the option of operating the used system dynamics in the relevant applicable operating points.

22.6 Dependency on other B&R libraries

In order to ensure the functionality of the MPC library, the following B&R libraries must be uploaded into the relevant AS project.

- Runtime
  a. Non-cycle-time-monitored task class.
- AsTime
  a. Internal calculation of the sample time of the process and the duration of an initialization process and an equation solving process.
- Sys_lib
  a. Memory management with TMP_alloc, TMP_free.
- Brsystem
  a. Monitoring of the free DRAMs (MEMInfo) of the hardware used.

22.7 Allocating hardware to the system scope

With MPC control, DRAM is allocated depending on the problem in question. The size of the memory depends on the number of controlled variables, manipulated variables, disturbance variables to be taken into account and the horizon lengths (prediction horizon, control horizon).
The following figure lists all input/output configurations of processes that can be controlled with a X20CP3586 / 512MB DRAM.

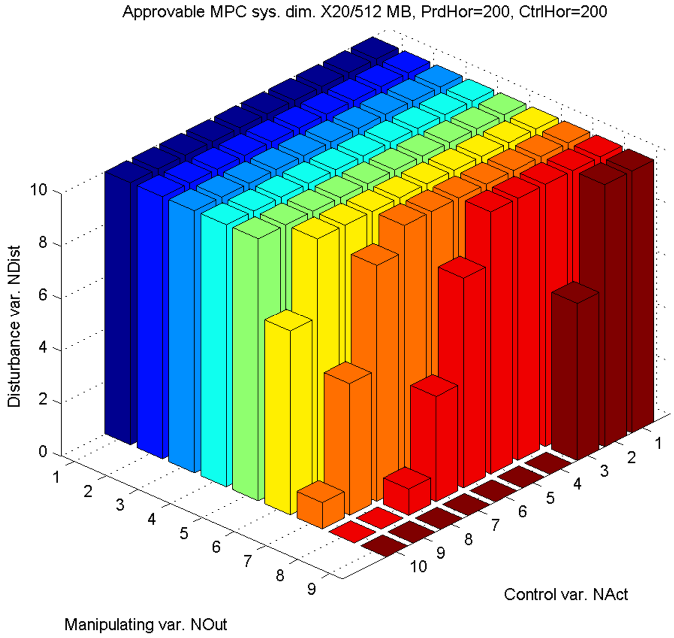


Figure 144, Possible MIMO system scope on x20 CP3586

Systems which are too large to be controlled with a x20 system in view of the maximum allowable prediction horizon Nx and control horizon Ny, must use an Automation PC (approved machine number: 5P62:220198.047-00).

22.8 APROL/MPC structure

The model-based predictive controller used is fully integrated in the APROL process control system. In its library, there is a pre-prepared controller module available, which contains the relevant image function blocks, the integration in the trend system, the alarm functionality and the system logics. For implementation, the function block simply needs to be put in the function plan, connected and uploaded to the relevant controller. The parameter settings for the controller are then applied via the graphical user interface in APROL.
This implementation method enables maximum process proximity between the measurement signals, the technical control algorithms and the actuators, as a result of which the controller can quickly influence the process or respond to disturbance variables. Further advantages include increased availability and time and cost savings when implementing and creating interfaces. For this reason, an MPC solution can be implemented very quickly and cost-effectively.
Data on the manipulated and controlled variables is shown in the APROL trend system as a supplement to the system history. An overview of the typical architecture and graphical user interface of the APROL/MPC is shown in Figure 145, Architecture of APROL-MPC.

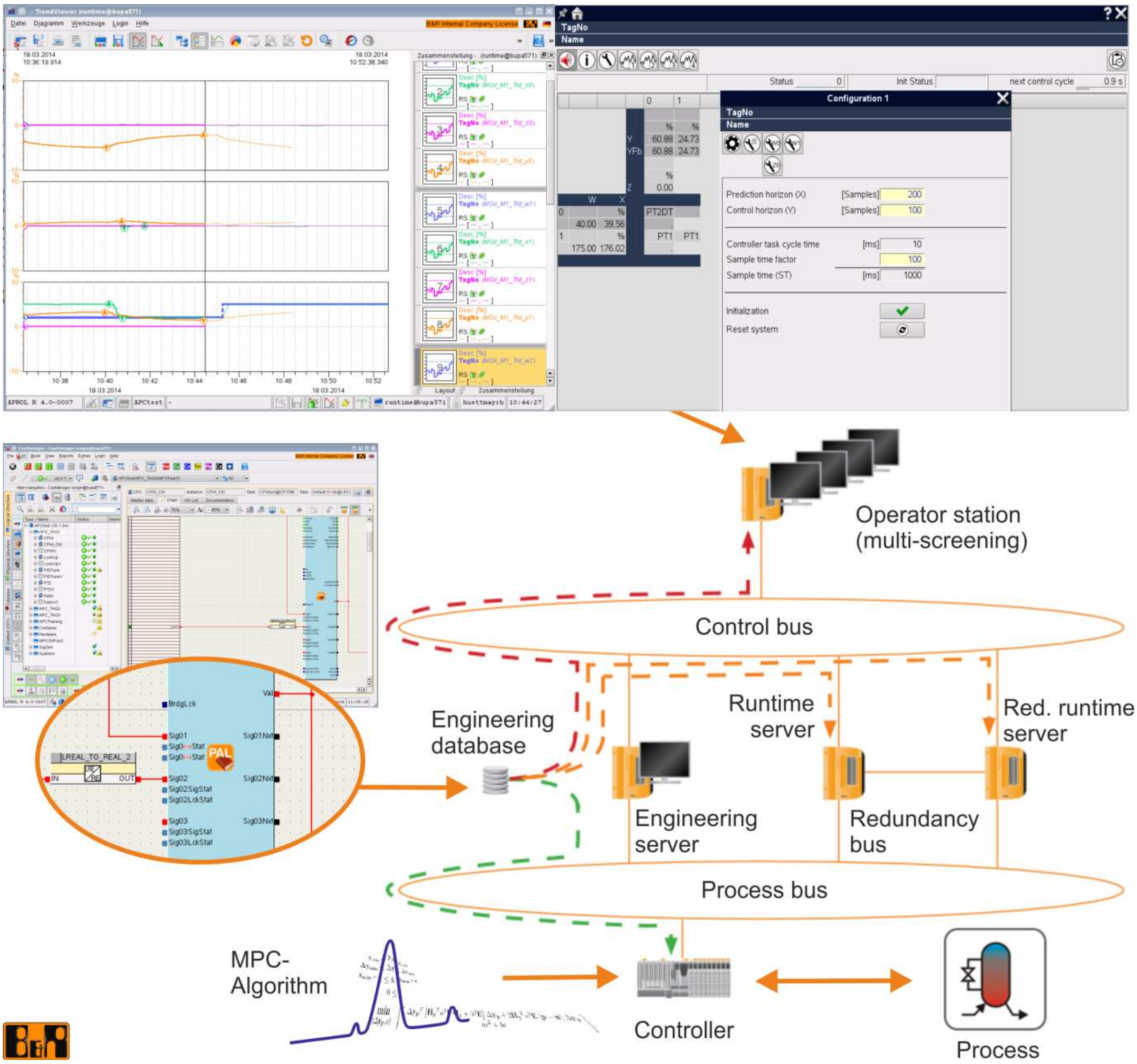


Figure 145, Architecture of APROL-MPC

22.9 Scenario for MPC project development

When using an MPC, project development can be broken down into the following detailed steps.

22.9.1 Analysis of relationships in the process

Successful implementation of an MPC requires a thorough analysis of the process being controlled. This requires exact understanding of the process that will be controlled.

- Review of the plant diagram or PFD (process flow diagram)
  - What is this unit producing?
  - What is the primary energy source (raw materials, etc.) for the plant?
  - Where do the raw materials for this energy come from?
  - Where does the plant output lead to?
- Analysis of the P&ID (piping and instrumentation diagram)
- Questions for the different system operators
  - What is it like working with this process?
  - Are there any unusual circumstances to consider?
  - Are there any special configurations or operating conditions to consider?
  - Where is the system generally operated?
  - Which process constraints are the most important and which ones are the trickiest?
- Questions for the project technicians
  - How large are the deviations between the raw material feeds and the product outputs?
  - How is the process affected over the course of a day? Over the course of a year (different seasons)?
  - What type of energy is used to power the plant? What is the source of this energy and how much does it cost?
  - Are there any existing problems with the plant or the existing closed loop control? If so, what are they?
  - What are the specifications for the assets (economic assets)?
  - What have you learned from working with the plant so far?
  - How should the plant perform after the MPC controller has been commissioned?
  - What are the expectations for the new closed loop control concept?
- Questions about the efficiency of the unit
  - Are there any environmental regulations or special taxes influencing how the process is performed?
  - Do any higher-level cost-optimized calculation methods already exist?
  - *How high is the throughput or the financial benefit of the unit?*
  - *How much money would the owner like to invest to increase efficiency?*
- What is the target increase in productivity in terms of the current controller?
- Do any mathematical models already exist for the unit (even stationary)?
- Review of system documentation and/or operating instructions
- Analysis of documented process data (trends)

22.9.2 Degree of freedom for MPC to modify the process

The following points can only be determined in close cooperation with the project technicians, equipment operators and planning personnel.

- What plant sections or parts of the process is the MPC able to influence?
- What plant sections or parts of the process which cannot be influenced should be taken into account by the MPC?
- The structure can be extended to include additional variables depending on the degree of freedom awarded to the processes being controlled (e.g. fictitious controlled variables, non-measurable controlled variables).

22.9.3 Specifying the MPC configuration

If there is already PID control of the system part, the next step is to decide which PID control loops to replace directly by the MPC and which control loops should be subordinate to the MPC. PID control loops are generally better at quickly correcting unknown faults or operating certain parts of the plant at faster sampling times (e.g. closed loop control of flow, level, temperature, pressure). For example, PID control loops can be removed from control loops with a very high response time or dead time (in connection with the MPC process sampling) or if parts of the MIMO (multiple-input multiple-output) system can be better decoupled by removing the primary controller and this decoupling is also necessary.
Definition of the manipulated variables, the controlled variables and the manipulated variables of the MPC (based on the entire plant or the objective of the entire plant).

Check the MPC configuration.
It is important that the manipulated variables have a direct, fast influence on their respective controlled variables (e.g. the right choice of the measurement location can reduce dead time).

22.9.4 Designing the test scenario for modeling

In order to calculate the most reliable, accurate process model possible, the following points should be considered.

- Several different test runs should be performed for each manipulated variable (test signal planning to determine which test signals are possible => responsible technicians or plant operators)
- The changes to the manipulated variables should be set in a way that prevents correlations between test runs.
- The changes to the manipulated variables should be as large as possible (as allowed by the respective process).
- Product specifications and process limits should be kept in consideration through all the entire testing process.
- The nominal values and the constraints must be present for the respective manipulated variables in the MPC.
- The test runs with the different variables and the various changes to the manipulated variable can be integrated into a test scenario. This test scenario should be discussed with the plant operators and technicians.
- Specify the sampling time for recording process data when testing the plant. (e.g. depending on the sampling times of the respective measurements).
- The data should be recorded unfiltered => any existing filters turned off (vector band, etc) in the trend block.
- Possible creation or estimation of stationary transfer function matrices for the manipulated variables / controlled variables and disturbance variables / controlled variables.

22.9.5 Determining the process data

The following points should be ensured before the actual test runs are carried out.

- Are all measurements, actuators working properly?
- Are all subordinate control loops in the desired mode?
- How is the plant currently being operated? Are the material flows in the desired range? Is the plant currently in normal operation or are different process reactions expected due to predictable events which could disrupt the test sequences? Are all units in the desired operating state?
- If possible a step response could be included for each manipulated variable. The data could then be compared with the stationary transfer function matrices, which would add to the knowledge of the process.

The tests should be carried out in a plant operating condition that is as close as possible to the nominal, optimal operating points of the process. For example, the optimum operating points can be determined by stationary operating point optimization in connection with estimation of the stationary transfer function matrices. The optimum operating points must then be discussed with the project technicians. It is important for the responsible plant operators to be informed in order to prevent anyone from intervening in the tests (don't forget about shift changes). Any disturbances which cannot be measured should be avoided as best as possible during the recording phase in the interest of subsequent model identification. Examples of disturbances which cannot be measured: Flow adjustment in recirculation ducts, changes in the product properties of the process outputs, changes in the controller settings of the subordinate controllers, various changes to the process operating points or different boundary conditions.

22.9.6 Analysis and model identification

The data obtained can be used for different identification methods. Based on knowledge of the process, purely physical limitations can be placed on the individual model dynamics in the respective order. On the other hand, the dead times of the models must be calculated separately. Identification of the process models should then take place based on estimated system dynamics (1 power buffer = > PT1) with the data sets corrected using the respectively estimated dead time. An initial comparison can then be made with the stationary models. Based on the recorded series of tests, the models can be directly compared with the measured controlled variables. The differences between positive and negative changes in manipulated variables should also be compared (e.g. to detect non-linearity). Additionally, the linear models can be adapted to the actual process through controlled variable transformation.

In the next step, you can link the MPC-controller to the process in passive mode (with all manipulated and disturbance variables provided). The controller supplies various metrics that can be used to analyze the model quality (see section 22.10.11). In addition, analyzing prediction errors between calculated and measured controlled variables makes it possible to identify stationary deviations. It is also possible to identify deviations in the model dynamics using correlation coefficients between the rates of the calculated and measured controlled variables.

22.9.7 Configuring the MPC

Initial configuration of the controller should also be performed when the controller is in passive mode. The MPC must be setup in such a way so that all constraints, system dynamics and economic aspects can be adhered to and/or satisfied.
It is important that the controller is designed or configured so that the amount of time for calculating the variables is always less than the selected process sampling time. The calculation time is increased mostly by the number of active constraints in conjunction with the control horizon and the number of manipulated variables.
Depending on the process to be controlled, the MPC has grades of freedom available. During configuration, the controlled variables should be assigned with the associated tuning factors depending on how important control performance is.
Configuring the manipulated variable dynamics by the "Y Move suppression". The dynamics of the individual manipulated variable changes should be set so that not only set value changes meet the requirements for the process dynamics but also any occurring disturbance inputs can be accordingly corrected with the appropriate dynamics (=> different tests in offline mode). The plant operators know about the required dynamics for most of the manipulated variables.
In addition, all MPC errors or warning messages should be discussed and accordingly added to the alarms in the APROL system. A "fail strategy" should also be established in the event that the MPC is shut off.

22.9.8 Initial startup

The most important factor after a new controller has been installed is the plant operator. If the MPC is not accepted, the likelihood increases that it will be taken offline. Therefore, a training course should be scheduled for plant operators which is clear and easy to understand. This training should point out the differences between the MPC and the previous solution.
A more detailed look into the MPC algorithm could alleviate any concerns those with authority to configure the MPC might have about making fine-tuning adjustments on their own.
After extensive testing in passive mode, the controller can then go online. The following points should be observed.

- The tuning factors for influencing the manipulated variable dynamics should be set conservatively. In addition, the change rate for the manipulated variables can be directly limited in the MPC using the constraints.
- Check to see whether the manipulated variables of the MPC match any operating point connections that might exist.

- Check to determine whether the controlled variables are also valid for the respective operating points.
- Some manipulated variables can be specified as stationary for the time being. During the commissioning phase, limitations can be loosened for the active manipulated variables and locked manipulated variables also released with increasing control quality.
- The standard deviations of the controlled variables should decrease. In addition, efforts should also be made to determine whether the variables are subject to high fluctuations (standard deviations).
- Check whether all controller errors are low or whether the set-points have been achieved.

22.9.9 Evidence of improved control quality

After successful commissioning, efforts should be made to ensure that the standard deviations of the control errors have been reduced and that the control performance of the controlled variables has been increased by commissioning of the MPC. The increase in control quality can also be represented in terms of something like weekly cost savings based on economic estimates (data from any existing stationary cost optimization).

22.10 Set-up and commissioning of the control module "ModelPredictiveController10_01"

This section explains the parameter setting options for the MPC.

22.10.1 Parameter settings for the process model / process sampling time/ prediction horizon

The process sampling time must be adapted to the process model. If the process models show a very large difference in their system time constants, it must be ensured that the sampled impulse or step responses fall within a maximum of 200 sample values in steady state vibration.
The sampling time can be selected using the plant characteristics in conjunction with the limits of the MPC function blocks (model horizon). Good controller performance is achieved if the minimum rise time min(Tf) for a process fulfills the following relationship during the sampling time Ts.

$$\frac{\min(Tf)}{Ts} \approx 4-10.$$

Equation 72, Process sampling time as a function of the system time constants

(See also Limitations of Dynamic Matrix Control, Morari). The MPC function blocks account for the step/impulse responses with a maximum model horizon of 400 sample values.
The prediction horizon is then adjusted to the number of sample values needed to reach stationary state. Be aware that the prediction horizon is limited to 200 sample values. For this reason, the stationary states (if stability => final value theorem) for all step responses of the system must be recorded before the end of the given sequence.

Standard transfer functions, with which many processes within the process technology industry can be described, are already prepared in the MPC module. As shown in (Figure 146, Parameter settings for MPC process models), the prepared transfer functions for describing the respective subprocess can be selected from a "drop down" menu. The parameter settings are then applied by directly specifying the transfer function parameters in the user interface.

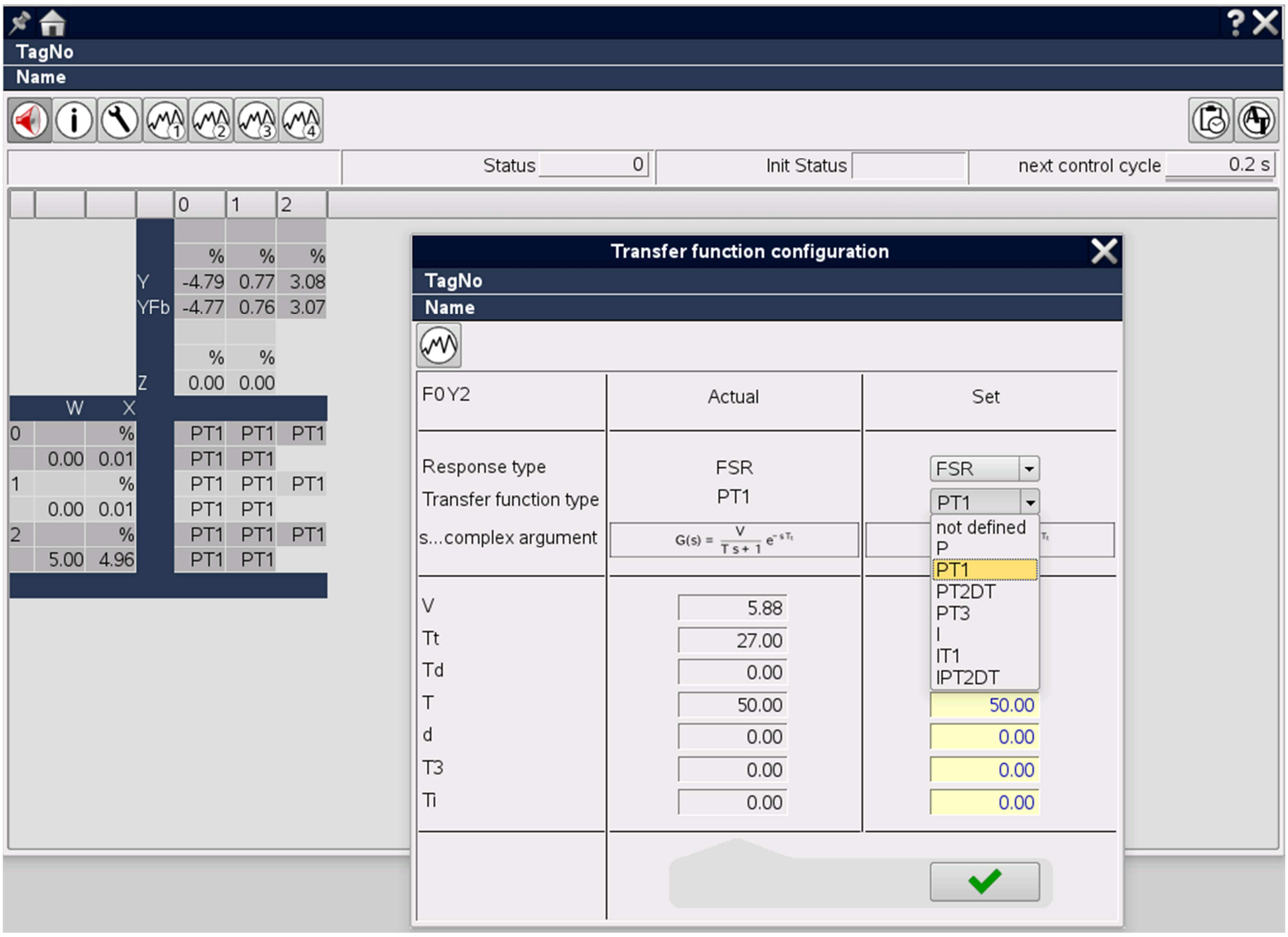


Figure 146, Parameter settings for MPC process models

In Figure 146, Parameter settings for MPC process models at the very top (under "Response type"), the type and process model can be saved in the MPC, and you can change between FSR ("finite step response") and FIR ("finite impulse response"). For systems with inherent regulation the "Response type" should always be set as "FSR". If the process to be controlled has a process part with integrating characteristics (without inherent regulation), the "Response type" must be set as "FIR" for all sub-transfer functions between the corresponding controlled variable and all manipulated inputs, as well as between the controlled variable and all disturbance variable inputs.

22.10.2 Checking model quality

A more accurate model of the process being controlled increases the success rate of the model-based controller. To check the model quality, the controlled variables are calculated in the controller based on the saved pulse- and/or step responses as well as the system history (manipulated and disturbance variables) and then compared with the measured data (see section 22.10.11). The prediction horizon determines the period of time which the system history should account for. As a result, the quality of the model-based calculated controlled variables depends on the model itself and on the defined prediction horizon. The model can be reviewed when the MPC controller is in passive mode. Different statistical parameters are calculated in the MPC (see section 22.10.11) which can be used to verify the process model.

22.10.3 Subordinate PID control loops

If PID control loops are intended as slave controllers for the MPC, attention must be paid to the fact that, depending on the dynamics of the lower-level control loop, this needs to be taken into account in the modeling of the MPC. If subordinate control loops do have to be taken into consideration, they must be optimally configured before the control dynamics are applied to the entire process model. It is very important for the subordinate controller to be configured as best as possible at this point because the dynamic part of the subordinate closed loop control accounted for in the MPC should not be re-configured again later. This is because model/system deviations would occur in this case.

22.10.4 The control horizon

The length of the control horizon can be determined by the system properties (e.g. taking into account the settling time for the slowest system behavior). The length of the prediction horizon "PrdHor" can be selected as the maximum possible control horizon "CtrlHor", however, the following points should be observed.

- The computational complexity increases non-linearly with the number of calculated open loop controlled variable changes (by at least O(["CtrlHor"]^2)). This can significantly increase the computing time.
- The robustness of the closed loop control, particularly in the case of high deviations between the model and the plant, is negatively affected by a high "CtrlHor" control horizon.
- The property comparable to the "i-component" in PID control must be estimated lower as the "CtrlHor" control horizon rises because the correction of the predication can take place over multiple open loop control variable changes.

The choice of control horizon is very highly dependent on the type or sampling of the model stored in the controller. For fast MPC configurations for processes without dead time, just a few sampling sequences are enough (Control horizon Y > 5). In other cases, control horizons such as Control horizon (Y) = Control horizon (X) will be reached. An initial check is needed whether the control horizon has been set too high in conjunction with the set process sampling time, if the optimization takes more time than the set process sampling time (see Figure 147, Optimization time for MPC exceeds process sampling time (Lunze, 2010) (King, 2011) (Maciejowski, 2002)). In this case, error 36506 will be given by the module status output.

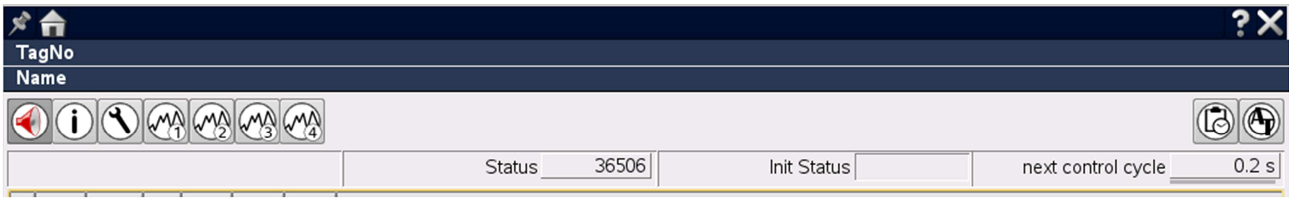


Figure 147, Optimization time for MPC exceeds process sampling time

The prediction horizon, control horizon and process sampling time can be set in the following graphical user interface.

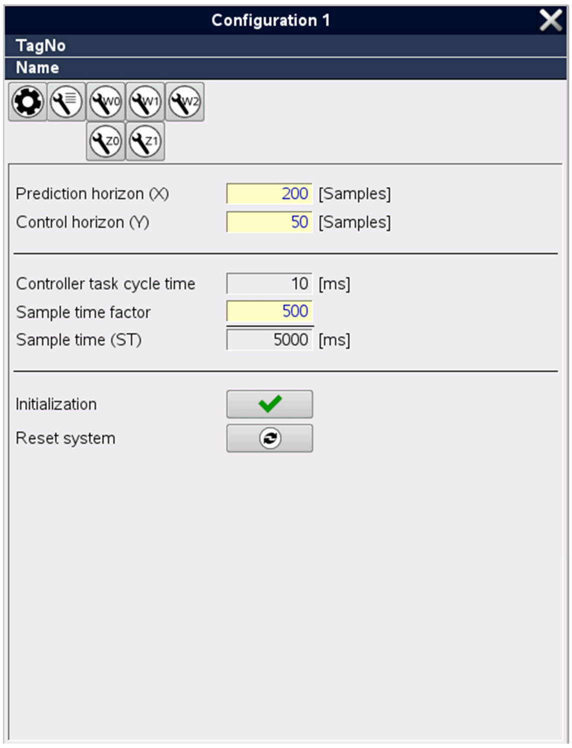


Figure 148, Setting the horizons and sampling time

22.10.5 Weighting of manipulated variables

All other settings, such as the controller dynamics, the limitations, the operating points, or the various operating settings can be set in the following user interface.

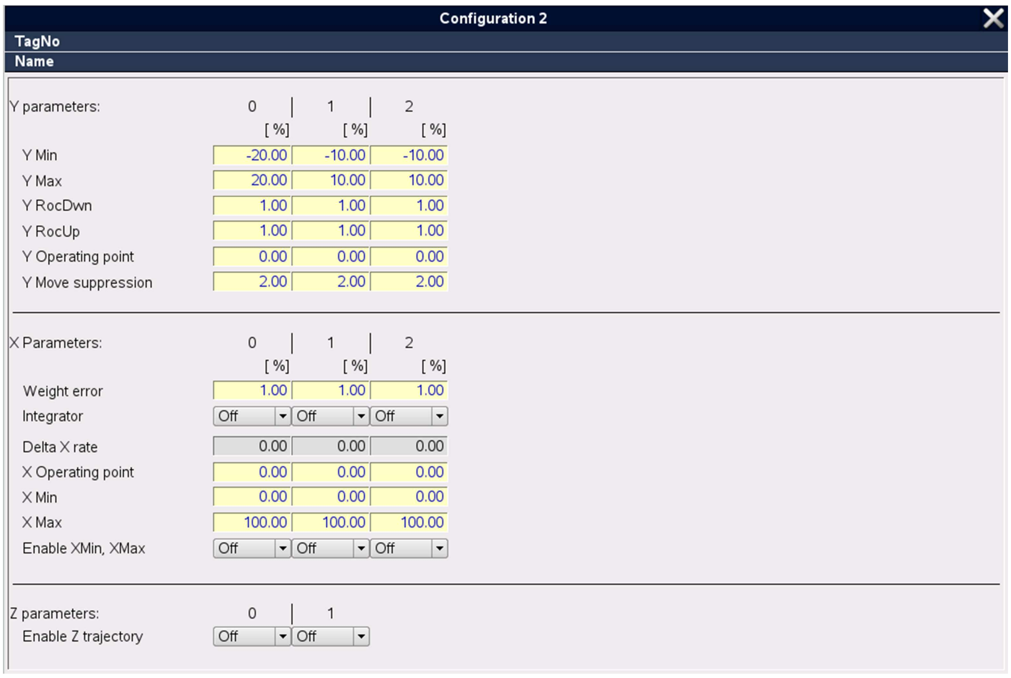


Figure 149, MPC parameter settings

The maximum possible deviations between the process model in the controller and the real model should be known (estimated) when setting up the controller on the real model. This allows the weighting factor for the manipulated variable, "OutMoveSuppression", which impacts the calculation quadratically, to be adapted to the individual problem. This ensures more robust control behavior under the highest manipulated variable dynamics in the "worst-case" scenario.
In principle, an increase in "OutMoveSuppression" will decrease the conditional number of the Hessian matrix of the process being optimized. Generally, reducing the conditional number will increase the robustness of the controller. The weighting of the controller error with "WeightError" is proportional to the dynamics of the controller.

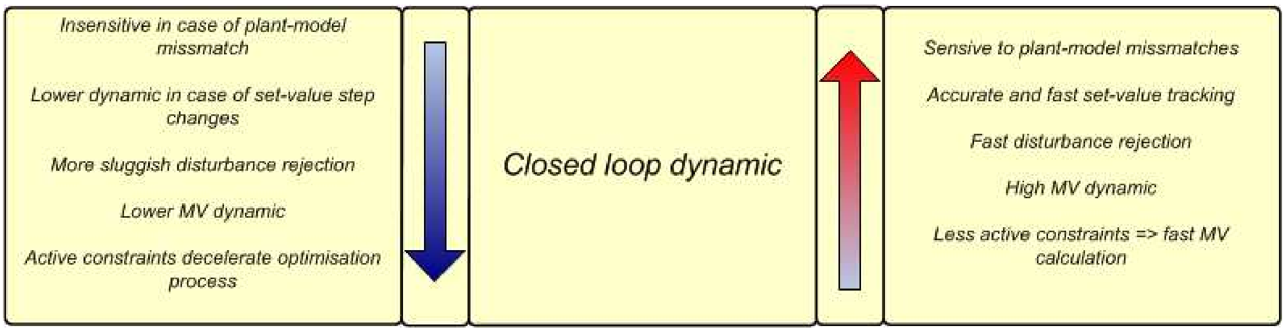


Figure 150, MPC parameter settings, dynamics of closed loop

22.10.6 Weighting of controlled variable errors

The weighting of the control error can be set separately depending on the controlled variables. This provides particular benefits for ("fully-coupled, under-actuated") MIMO systems, in which the relative meaning of the individual controlled variables can be specifically adjusted (e.g. temperature is more important than fill level). In order to achieve an appropriate weighting of the controlled variables, either all controlled variables must be normalized or prioritization will be given using the relative intervals between the weighting factors.
In principle, an increase in "WeightError" will increase the conditional number of the Hessian matrix of the process being optimized.

22.10.7 Limitations on manipulated variables

The controller directly accounts for limitations of the manipulated variables while being calculated. The constraints of the quadratic control loop quality factor can result in non-linear behavior of the controller. Caution should be used when setting the MPC because exaggerated configuration settings can dramatically reduce the robustness and is not always easy to identify (due to active constraints).

Important:

The calculation time of the optimization increases with the number of active constraints.

22.10.7.1 Limitation of manipulated variable change

The parameters "OutDeltaMin" and "OutDeltaMax" can be used to limit the minimum and the maximum value for the manipulated variable change depending on the scan cycle. The value can be assigned in accordance with Figure 149, MPC parameter settings.

22.10.7.2 Limitation of absolute manipulated variables

The parameters "OutMin" and "OutMax" are used to limit the minimum and maximum absolute value of the respective manipulated variables. This value is also assigned in accordance with Figure 149, MPC parameter settings.

22.10.7.3 "Soft constraints" to limit the controlled variables

In addition to the manipulated variable limits, restrictions on the controlled variables can also be introduced in the calculation of the open loop control variable changes by quadratic programming. If the constraints must be observed, there is a danger that the optimization problem cannot be solved. Introducing so-called soft constraints makes it possible to proceed without preventing the optimization problem from being solved. Depending on the controlled variable, the use of soft-constraints extends the optimization process variable only by one unknown (slip) variable ε with a corresponding tuning factor ρ

$$\min_{\Delta y,\varepsilon}\begin{bmatrix}\Delta y^T F \Delta y + f^T \Delta y \\ \rho\varepsilon^2 + 2\rho\varepsilon\end{bmatrix},$$

Equation73, optimization problem with soft constraints

and the following additional constraints

$$\begin{aligned}&\vec{x}_{\min}(k) - \vec{1}\varepsilon(j) \le \vec{x}(k+j) \le \vec{x}_{\max}(k) - \vec{1}\varepsilon(j)\\ &\varepsilon(j) \ge 0.\end{aligned}$$

Equation74, additional constraints for optimization problem with soft constraints

The tuning factor ρ is automatically set to equal the largest eigenvalue of the Hessian matrix F so that the conditional number is not negatively affected. The user cannot change the weighting of the soft constraints. Adjustment to the overall problem can be changed by relatively changing the constraints of the controlled variables in relation to the desired operating limit.

Important:

Normally, soft constraints are observed; however, as soon as an inconsistency were to occur in the algorithm, these soft constraints can be violated to the benefit of the robustness of the solution procedure.

22.10.8 Compensation for future faults

The MPC-controller offers the possibility of accounting directly for future faults that are determined for example during production planning. The compensation is carried out over the area of the prediction horizon. This compensation is particularly beneficial if the dead time of the manipulated variable transfer functions is greater than that of the disturbance variable transfer functions. In addition, the control horizon should not be set too small because the controller error caused by the compensation is compensated for by the manipulated variable changes over the control horizon.

22.10.9 Initializing the controller parameter settings

After all parameters have been set, the control module must be re-initialized. Initialization can be carried out at any point. If the initialization time is longer than the process sampling time, the manipulated variable output remains constant throughout the entire calculation time. When the initialization button is pressed, (see Figure 148, Setting the horizons and sampling time) the process can start.

The initialization will be indicated by the status information 36500 on the user interface (see Figure 151, initialization process MPC). In addition the user also has access to an information bar, which displays the progress of the initialization ("Init Status".

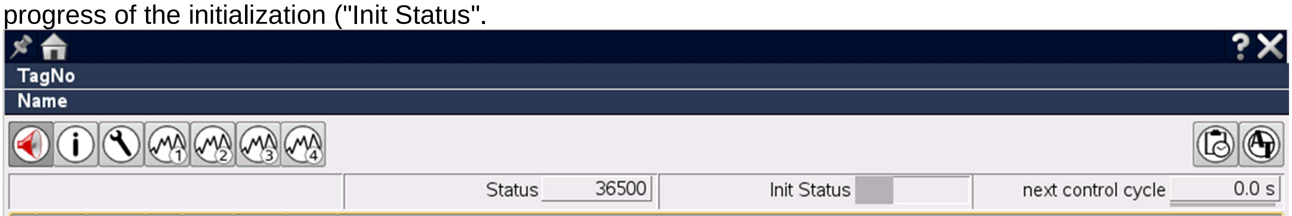


Figure 151, initialization process MPC

22.10.10 Multiple-input Multiple-output (MIMO) closed loop control

The scope of the MIMO process being controlled does not always have to be constant. This means that the number of input/output variables taken into account can be changed, for example by the subordinate controller or through specific adjustment. The MPC-controller must be able to correctly interpret these circumstances and to account for them accordingly when calculating the manipulated variable.

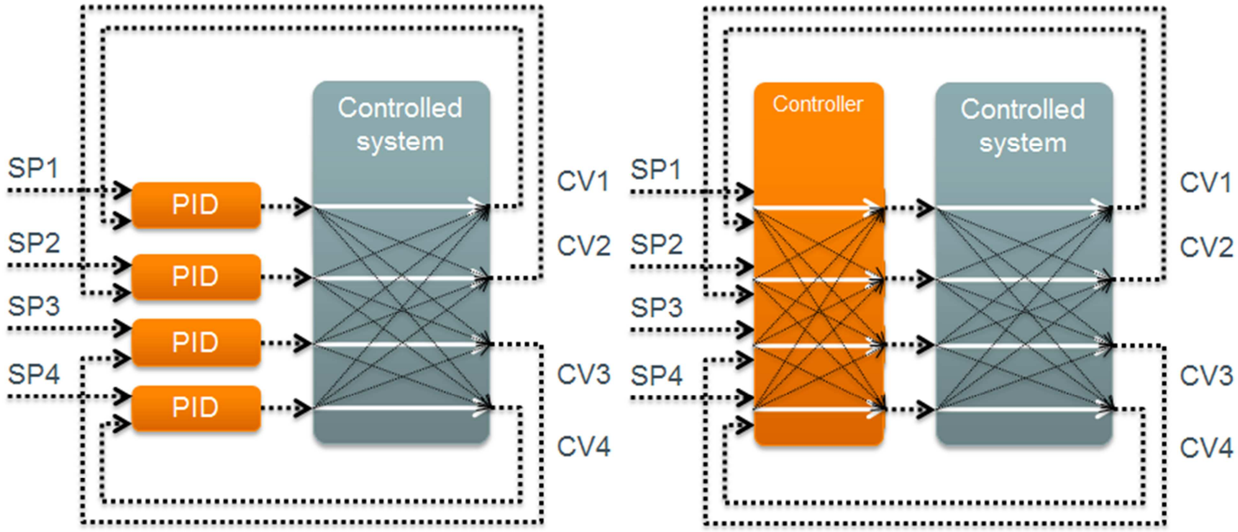


Figure 152, Comparison of single/multivariable variable control

22.10.10.1 Under-actuated systems

Under-actuated (over specified) systems can be distinguished by the fact that the process has more controlled variables than manipulated variables. The controller has more control targets than actuators. In these kinds of systems, a permanent control deviation cannot be avoided in stationary situations. Using "WeightError" to configure the controller makes it possible to prioritize the controlled variables and the levels of control quality for the respective controlled variable. For example, the controller can be configured so that a certain temperature is corrected better than a material concentration.

22.10.10.2 Determinate systems

Determinate multivariable systems have just as many controlled variables as manipulated variables. Systems of this type can be considered stationary without permanent control differences.

22.10.10.3 Over-actuated systems

In over-actuated systems, the number of manipulated variables is higher than the number of controlled variables. Systems of this type offer an extremely large number of potential combinations for manipulated variables in order to achieve the control targets. The surplus degrees of freedom can be reduced by the following measures. The locked degrees of freedom can can be used at the same time for cost-optimized settings of the manipulated variable.

- One additional controlled variable could be added for each control input. The relationship between the manipulated variable and controlled variable could for example be specified by a non step-capable proportional element (delayed by one scan cycle) with the gain factor 1. (Advantage: stationary exact solution. Disadvantage: considerable expansion of the MIMO system structure => not possible at low scan rates by increasing the computation time).
- Individual manipulated variables, which were able to be held stationary at a operating point anyway, were able to be removed from the MIMO structure. Changes to the respective manipulated variable could be recorded in the MPC via disturbance inputs. For example, taking

into account future disturbance variables by specifying the set value of externally controlled manipulated variables. (Advantage: Reduction of the MIMO system size. Disadvantage: Additional programming measures)

- Specific manipulated variables in a system that should be kept stationary at a specific operating point can be added to the MIMO system. The MPC could be configured so that an operating point has a permanent definition. The factor for reducing the manipulated variable dynamics (T move suppression) is increased significantly in the MPC-programming in contrast to all other manipulated variables. When changes occur to the controlled variable set values, the manipulated variables with extremely low manipulated variable dynamics will not be used (since the manipulated variables with higher dynamics have already corrected the control deviation).

The respective operating points could be calculated through stationary process optimization by specifying respective cost factors (LP-problem e.g. via Simplex algorithm).

22.10.11 Monitoring model quality

The MPC function block returns information via the manipulated variable inputs and disturbance variable inputs about the model quality through permanent calculation of the model-based controlled variable on the basis of the system history. The calculation is done with convolution of the individual inputs with the corresponding step or pulse responses followed by linear combination. The system history accounted for was linked directly to the prediction horizon. To analyze the model quality, the controlled variables calculated based on the model can be permanently compared with the measured controlled variables. The graphical comparison between the controlled variables calculated based on the model and those that were measured can be used be used to detect changes between the model and the plant. In addition, the following variables can be used to estimate the model quality.

Parameter	Data type	Description
ActDeltaPrd	LREAL[0...9]	Change rate of controlled variables calculated based on the model. The measure is based on the difference between the current and the past controlled variable calculated based on the model.
ActIDeltaPrd	LREAL[0...9]	Integral of the change rate of controlled variables calculated based on the model. The sum of the change rates can be used to validate the dynamic part of the model. With set-point changes, a higher correlation coefficient between the measured controlled variable and "ActIDeltaPrd" results in a good match between the model-based process dynamics and the actual process dynamics (see "CorrActMeasuredPrd").
PrdError	LREAL[0...9]	Prediction error (errors between measured controlled variables and those calculated based on the model).
CorrActMeasuredPrd	LREAL[0...9]	Correlation coefficients between ActIDeltaPrd and the measured controlled variables.

Table 16, Monitoring model quality

22.10.12 Control quality monitoring

The control quality of the model-based predictive closed loop control does depend greatly on the model quality, but the controller should be configured according to the environmental conditions. The following metrics can be used for general analysis of the control quality.

Parameter	Data type	Description
CtrlError	LREAL[0...9]	Control error as difference between measured controlled variables and set-points.
Slack	LREAL[0..9]	Slack variables that can occur when calculating soft controlled variable constraints. If soft constraints are used, the corresponding slack variables are exactly equal to 0 if the constraints of the controlled variables must be violated on the prediction horizon in order for a valid solution of the optimization criteria to be guaranteed. Slack variables can also

		be used to monitor the control quality. The more often the slack variables are unequal to 0, the more likely the controller will be confronted with situations that negatively influence the control quality or that drive the process outside the desired operating range. Slack variables that are unequal to 0 can be caused by unexpectedly strong disturbances, faulty controller configuration or starting and stopping procedures in the process among other things.

Table 17, Monitoring control performance

22.10.13 Summary

Setting the MPC can primarily be done using the horizons "prediction horizon", "control horizon", "sampling time (ST)" and the factors "weight error" and "Y move suppression".
When there is a large deviation between the model and the system, stable controller behavior may only be possible by setting high "Y move suppression" values. However, this will limit the dynamics of the closed loop control.

Notes

23 REVISION HISTORY

Manual version	Date	Change
V1.00	2014-03-19	Start of history

Perfection in Automation
www.br-automation.com