Control Engineering

A guide for beginners

Manfred Schleicher
Frank Blasinger

Control Engineering

A guide for beginners

Manfred Schleicher
Frank Blasinger

Preface

This work is intended to be of practical assistance in control engineering technology. It will help you to select and set up a suitable controller for various applications. It describes the different types of controller and the options for setting them up. The explanations and definitions are provided without using advanced mathematics, and are mainly applied to temperature-control loops.

In this new and revised edition, Chapters 3 and 5 have been extensively updated.

We wish to thank our colleagues for their valuable support in writing this book.

Fulda, January 2003

Manfred Schleicher Frank Blasinger

JUMO GmbH & Co. KG, Fulda, Germany

3rd Edition

Part number: 00323761
Book number: FAS 525
Printing date: 02.04
ISBN: 3-935742-01-0

1 Basic concepts

1.1 Introduction

Automatic control is becoming more and more important in this age of automation. In manufacturing processes it ensures that certain parameters, such as temperature, pressure, speed or voltage, take up specific constant values recognized as the optimum, or are maintained in a particular relationship to other variables. In other words, the duty of control engineering is to bring these parameters to certain pre-defined values (setpoints), and to maintain them constant against all disturbing influences. However, this apparently simple duty involves a large number of problems which are not obvious at first glance.

Modern control engineering has links with almost every technical area. Its spectrum of application ranges from electrical engineering, through drives, mechanical engineering, right up to manufacturing processes. Any attempt to explain control engineering by referring to specialized rules for each area would mean that the control engineer has to have a thorough knowledge of each special field in which he has to provide control. This is simply not possible with the current state of technology.

However, it is obvious that there are certain common concepts behind these specialized tasks. It soon becomes clear, for example, that there are similar features in controlling a drive and in pressure and temperature control: these features can be described by using a standard procedure. The fundamental laws of control engineering apply to all control circuits, irrespective of the different forms of equipment and instruments involved.

A practical engineer, trying to gain a better understanding of control engineering, may consult various books on the subject. These books usually suggest that a more detailed knowledge of control engineering is not possible, without extensive mathematical knowledge. **This impression is completely wrong**. It is found again and again that, provided sufficient effort is made in presentation, a clear understanding can be achieved, even in the case of relationships which appear to demand an extensive mathematical knowledge.

The real requirement in solving control tasks is not a knowledge of many formulae or mathematical methods, but a **clear grasp of the effective relationships** in the control circuit.

1.2 Concepts and designations

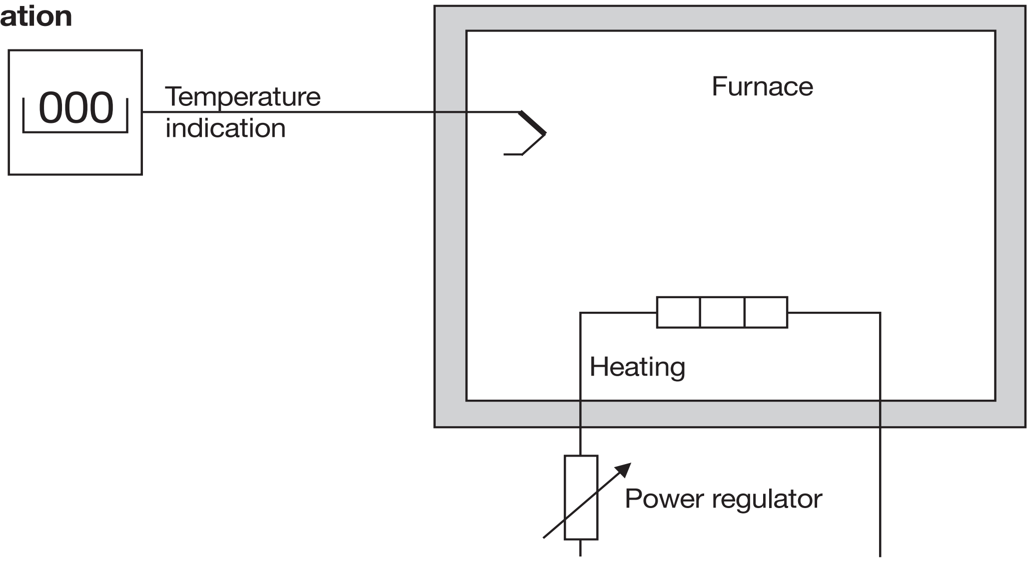
Today, thanks to increasing standardization, we have definite concepts and designations for use in control engineering. German designations are laid down in the well-known DIN Standard 19 226 (Control Engineering, Definitions and Terms). These concepts are now widely accepted in Germany. International harmonization of the designations then led to DIN Standard 19 221 (Symbols in control engineering), which permits the use of most of the designations laid down in the previous standard. This book keeps mainly to the definitions and concepts given in DIN 19 226.

1.3 Operation and control

In many processes, a physical variable such as temperature, pressure or voltage has to take up a specified value, and maintain it as accurately as possible. A simple example is a furnace whose temperature has to be maintained constant. If the energy supply, e.g. electrical power, can be varied, it is possible to use this facility to obtain different furnace temperatures (Fig. 1). Assuming that external conditions do not change, there will be a definite temperature corresponding to each value of the energy supply. Specific furnace temperatures can be obtained by suitable regulation of the electrical supply.

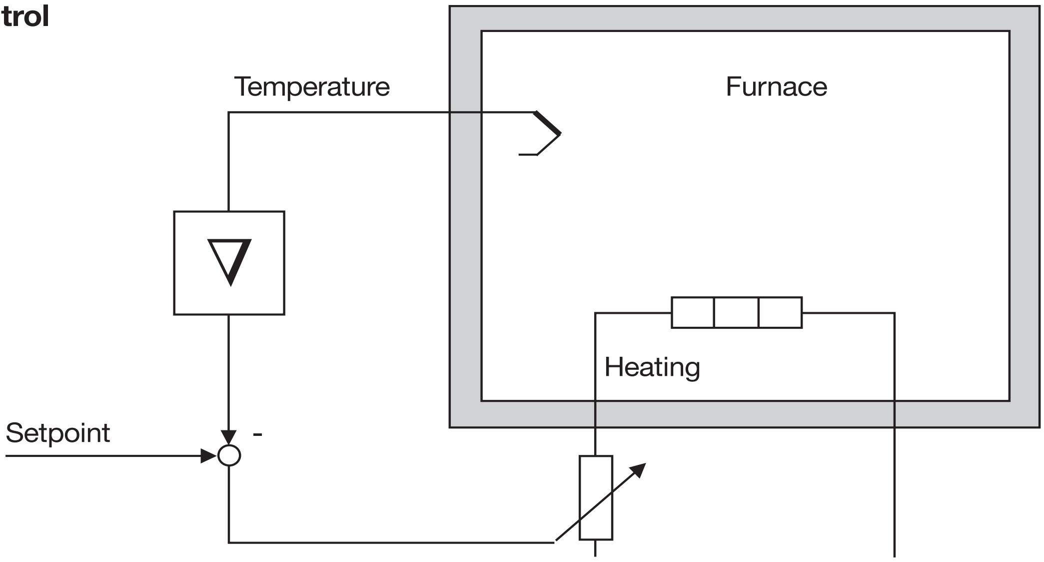
However, if the external conditions were to change, the temperature will differ from the anticipated value. There are many different kinds of such disturbances or changes, which may be introduced into the process at different points. They can be due to variations in external temperature or in the

Operation



Operation (as defined in DIN 19 226) is the procedure in a system whereby one or more variables, as input variables, influence an output variable, according to rules inherent in the system.
The distinguishing feature of operation is the open action sequence, through the individual transfer elements or the operating chain.

Control



Control (as defined in DIN 19 226) is the procedure whereby a variable (the process variable) is continuously monitored and compared with another variable (the setpoint). As a result of this comparison, it is then influenced to make it equal to the setpoint.
The resulting action sequence takes place in a closed circuit, the control loop.

Fig. 1: Operation and control

heating current, or caused by the furnace door opening. This type of temperature control takes no account of the actual furnace temperature, and a deviation from the required value may not be noticed by the operator.

Some form of control is necessary if the furnace temperature has to maintain its value in spite of changes in external conditions, or non-constant disturbances which cannot be predicted. In its simplest form the control may just be a thermometer which measures and indicates the actual furnace temperature. The operator can now read the furnace temperature, and make appropriate adjustments to the energy supply, in the event of a temperature deviation (Fig. 1).
The energy supply is now no longer pre-determined, but is linked to the furnace temperature. This measure has converted furnace operation into furnace control, with the operator acting as the controller.

Control involves a comparison of the actual value with the desired value or setpoint. Any deviation from the setpoint leads to a change to the energy supply. The energy input is no longer fixed, as is the case with simple operation, but depends on the actual process value attained. We refer to this as a closed control loop (Fig. 2)

If the connection to the temperature probe is broken, the control loop is open-circuited. Because there is no feedback of the process value, an open control loop can only be used for operation.

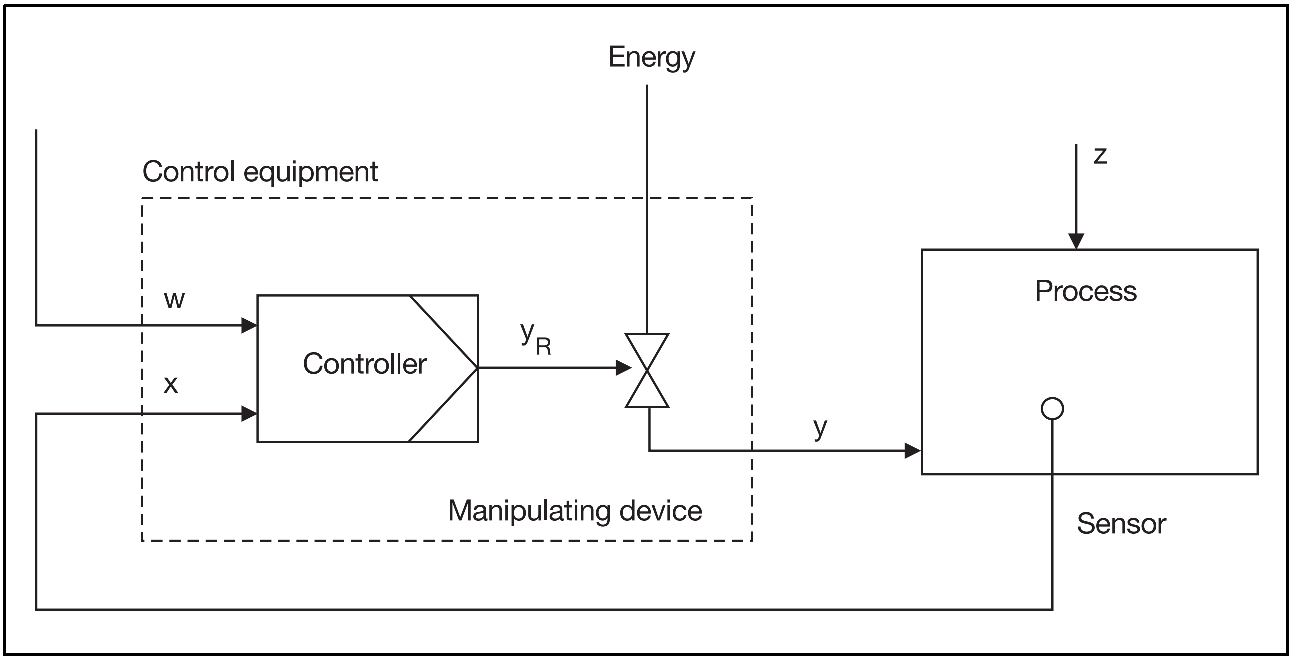


Fig. 2: The closed control loop

The control loop has the following control parameters (the abbreviations conform to DIN 19 226):

Process variable (process value, PV) x: the process value is the control loop variable which is measured for the purpose of control and which is fed into the controller. The aim is that it should always be made equal to the desired value through the action of the control (example: actual furnace temperature).

Desired value (setpoint, SP) w: the predetermined value at which the process variable has to be maintained through the action of the control (example: desired furnace temperature). It is a parameter which is not influenced by the control action, and is provided from outside the control loop.

Control difference (deviation) e: difference between desired value and process variable $e = w - x$ (example: difference between required and actual furnace temperature).

Disturbance z: an effect whose variation exerts an unfavorable influence on the process value (influence on the controlled variable through external effects).

Controller output Y_R: it represents the input variable of the manipulating device (the manipulator or actuator).

Manipulating variable y: a variable through which the process value can be influenced in the required way (e.g. heating power of the furnace). It forms the output of the control system and, at the same time, the input of the process.

Manipulation range Y_h: the range within which the manipulating variable can be adjusted.

Control loop: connection of the output of the process to the input of the controller, and of the controller output to the process input, thus forming a closed loop.
It consists of controller, manipulator and process.

The physical units involved can differ widely:
process value, setpoint, disturbance and deviation usually have the same physical units such as °C, bar, volts, r.p.m., depth in metres etc. The manipulating variable may be proportional to a heating current in amps or gas flow in m^3/min, or is often a pressure expressed in bar. The manipulation range depends on the maximum and minimum values of the manipulating variable and is therefore expressed in the same units.

1.4 The control action

The basic task of the controller is to measure and prepare the process value PV, and compare it with the setpoint SP; as a result it produces the appropriate manipulating variable MV. The controller has to perform this action in a way which compensates for the dynamic characteristics of the controlled process. This means that the process value PV should reach the setpoint SP as rapidly as possible, and then fluctuate as little as possible about it.

The action of the controller on the control loop is characterized by the following parameters:

- the overshoot: X_o,
- the approach time: T_a, the time taken for the process value PV to reach the new setpoint SP for the first time,
- the stabilization time: T_s,
- and also agreed tolerance limits ± Δx (see Fig. 3)

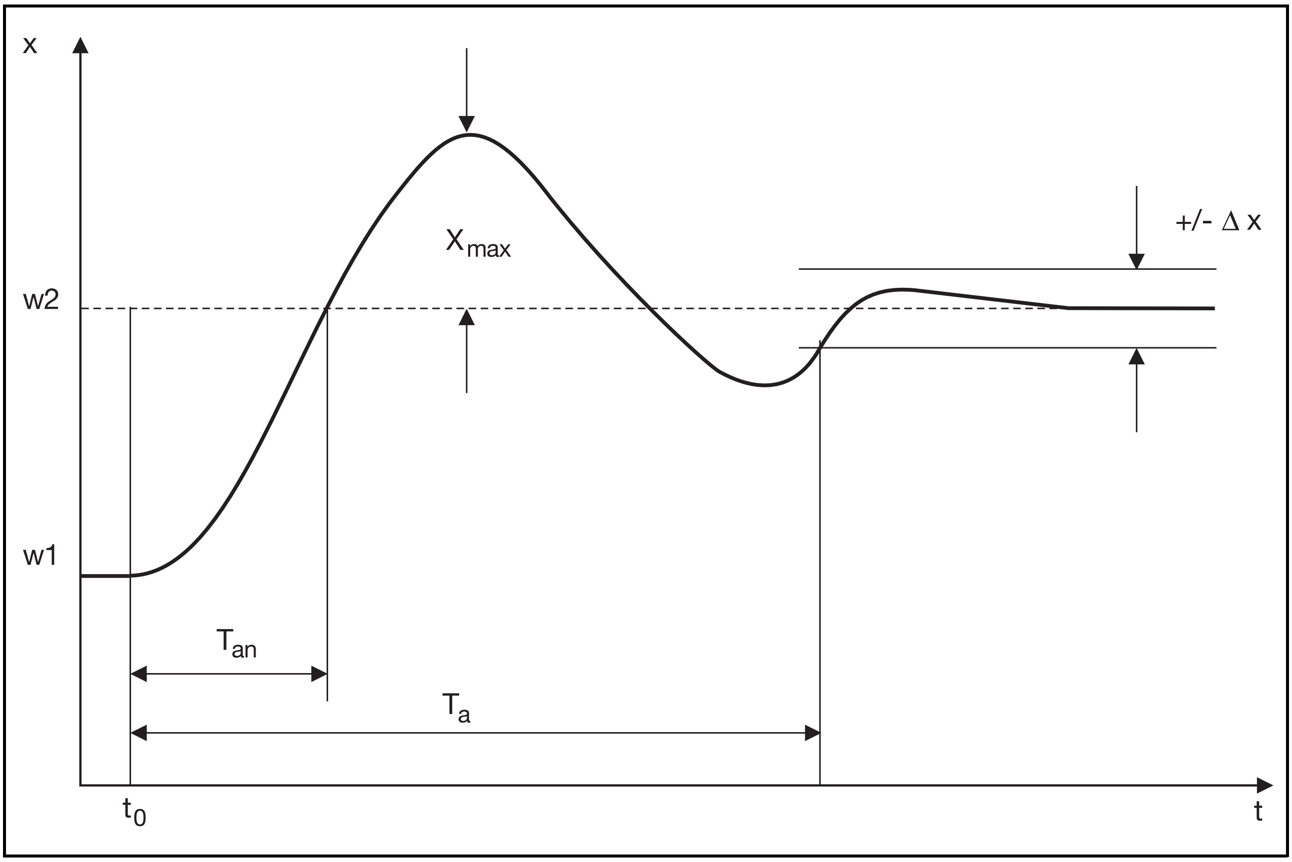


Fig. 3: Criteria for control action

The controller is said to have “stabilized” when the process is operating with a constant manipulating variable MV, and the process value PV is moving within the agreed tolerance band ± Δx.

In the ideal case the overshoot is zero. In most cases this cannot be combined with a short stabilization time. In certain processes, e.g. speed controls, rapid stabilization is important, and a slight overshoot beyond the setpoint can be tolerated. Other processes, such as plastics production machinery, are sensitive to a temperature overshoot, since this can quite easily damage the tool or the product.

1.5 Construction of controllers

The choice of a suitable controller depends essentially on its application. This concerns both its mechanical features and its electrical characteristics. There is a wide range of different designs and arrangements, so only a few will be discussed here. The discussion is limited to electronic controllers, and excludes mechanical and pneumatic control systems. The user, who is faced with choosing a controller for his particular application, will be shown initially which types are available. The listing is not intended to be comprehensive.

Mechanical variations:

- **Compact controllers** (process controllers) contain all the necessary components (e.g. display, keypad, input for setpoint etc.) and are mounted in a case which includes a power supply. The housing usually has one of the standard case sizes, 48mm x 48mm, 48mm x 96mm, 96mm x 96mm or 72mm x 144mm.
- **Surface-mounting controllers** are usually installed inside control cabinets and mounted on a DIN-rail or the like. Indicating devices such as process value display or relay status LEDs are not usually provided, as the operator does not normally have access to these controllers.
- **Rack-mounting controllers** are intended for use in 19-inch racks. They are only fitted with a front panel and do not have a complete housing.
- **Card-mounted controllers** consist of a microprocessor with suitable peripherals, and are used in various housing formats. They are frequently found in large-scale installations in conjunction with central process control systems and PLCs. These controllers again have no operating or indicating devices, since they receive their process data via an interface from the central control room through software programs.

Functional distinctions

The terms that are used here are covered and explained in more detail in later chapters (see Fig. 4).

- **Continuous controllers**
  (usually referred to as proportional or analog controllers)
  Controllers which receive a continuous (analog) input signal, and produce a controller output signal that is also continuous (analog). The manipulating signal can take on any value within the manipulation range. They usually produce output signals in the range 0 — 20mA, 4 — 20mA or 0 — 10V. They are used to control valve drives or thyristor units.

- **Discontinuous controllers**

2-state controllers (single-setpoint controllers) with one switching output are controllers that produce a discontinuous output for a continuous input signal. They can only switch the manipulating variable on and off, and are used, for instance, in temperature-control systems, where it is only necessary to switch the heating or cooling on or off.

3-state controllers (double-setpoint controllers) have two switching control outputs. They are similar to 2-state controllers but have two outputs for manipulating variables. These controllers are used for applications such as heating/cooling, humidifying/dehumidifying etc.

- **Quasi-continuous controllers**

Quasi-continuous controllers with one switching output are controllers that achieve a quasi-continuous action. The average value of the controller output over a defined time interval shows approximately the same time-dependent variation as a continuous controller. Applications are, for instance, temperature control (heating or cooling), where improved control-loop performance is required. In practice, quasi-continuous controllers with one switching output are also described as 2-state controllers.

Quasi-continuous controllers with two switching outputs can steer a process in opposing directions (for example, heating/cooling or humidifying/dehumidifying). These controllers also achieve a quasi-continuous action, by pulsing the switched outputs. In practice, all controllers that use two outputs to steer a process in opposing directions are referred to as 3-state controllers. Here the outputs need not necessarily be switched, but can be continuous.

- **Modulating controllers**

Modulating controllers have two switching outputs and are specially designed for motorized actuators which are used, for instance, to drive a valve to the open and closed positions.

- **Actuating controllers**

Actuating controllers are also used for motorized actuators and again have two switching outputs. They differ from modulating controllers by requiring feedback of the actuator position (stroke retransmission).

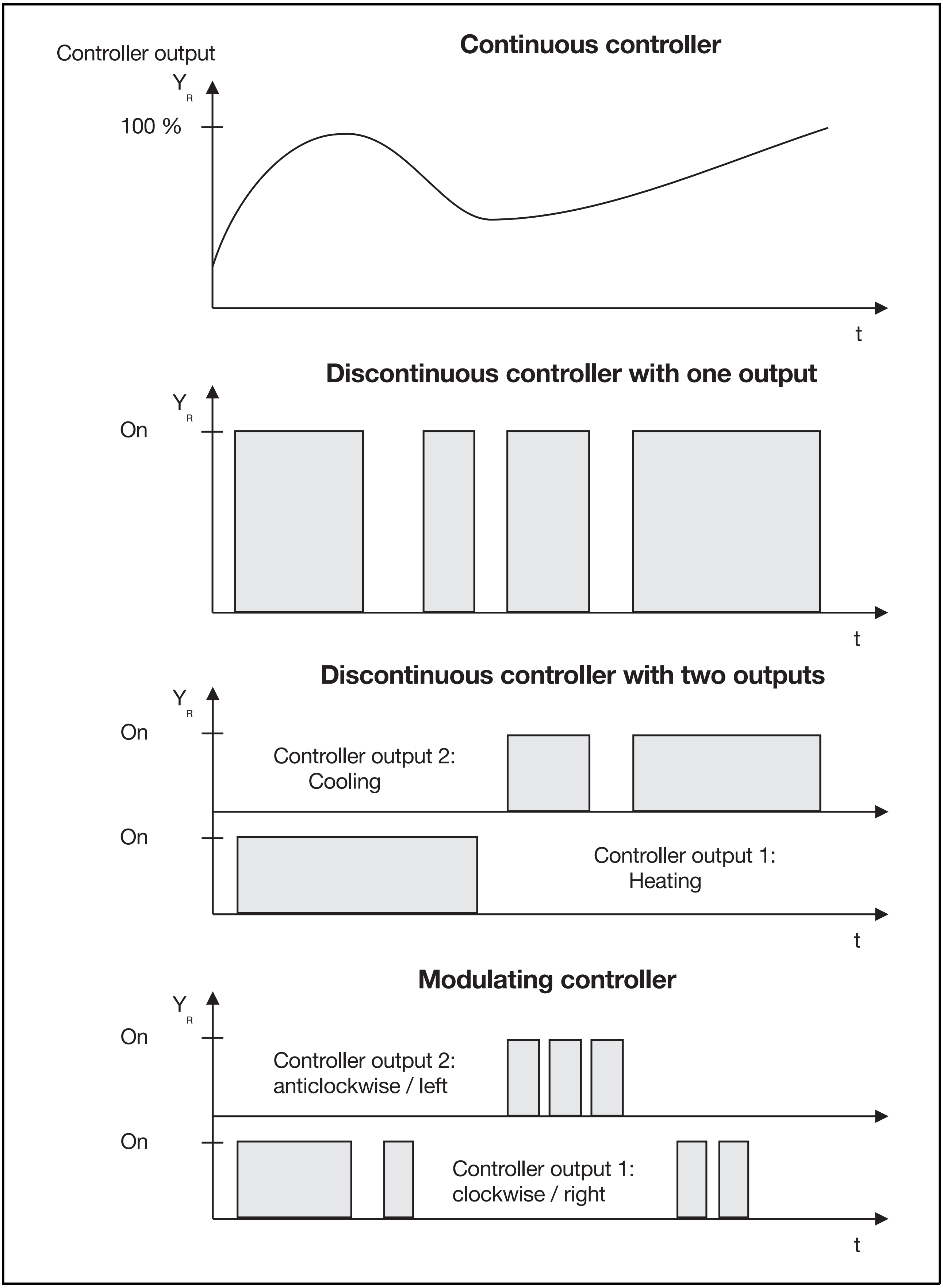


Fig. 4: Difference in controller functions

All these types of controller (apart from the discontinuous controller) can be implemented with different forms of dynamic response. This is often referred to as the “controller structure”. The terms used are P, PI, PD or PID controllers (see Fig. 5).

Different setpoint arrangements
The setpoint can be set manually on the controller by means of a potentiometer, or by using keys to input digital values. The setpoint is indicated in either analog form (pointer of a setpoint knob), or digitally as a numerical value.

Another possibility is the use of an external setpoint. The setpoint is then fed in as an electrical signal (e.g. 0 — 20 mA) from some external device. As well as these analog signals, it is also possible to use digital signals for setting the setpoint. The signals are fed into the controller through a digital interface and can be derived from another digital instrument, or from a computer linked to the controller. If this external setpoint operates according to a fixed time sequence (program), this is also referred to as program or sequence control.

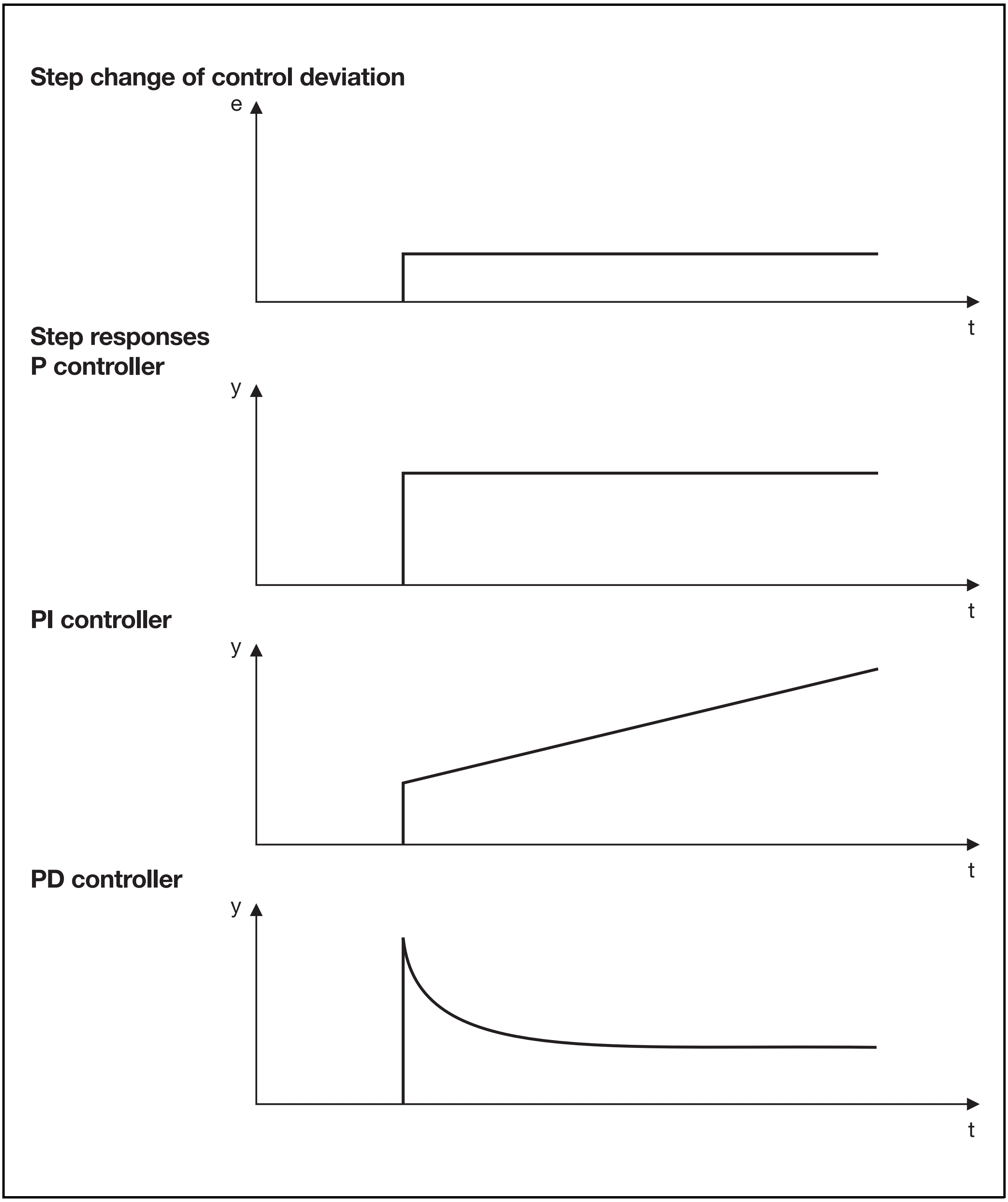


Fig. 5: Typical step responses

Evaluation of the process variable

The process variable must be available as an electrical signal. Its form depends on the sensor used and on the processing of this signal. One possibility is to connect the transducer signal (sensor, probe) directly to the controller input. The controller must then be capable of processing this signal; in many temperature probes the output signal is not a linear function of the temperature, and the controller must have a suitable linearization facility.

The other possibility is the use of a transmitter.
The transmitter converts the sensor signal into a standard signal (0 — 20mA, 0 — 10V) and usually also linearizes the signal. In this case the controller need only have an input for standard signals.

The process value is normally displayed on the controller. This can be in the form of a digital display (numerical indication), which has the advantage of being readable from a longer distance. The advantage of the analog display (pointer movement) is that trends such as rising or falling of the process variable are clearly visible, as well as the position within the control range.

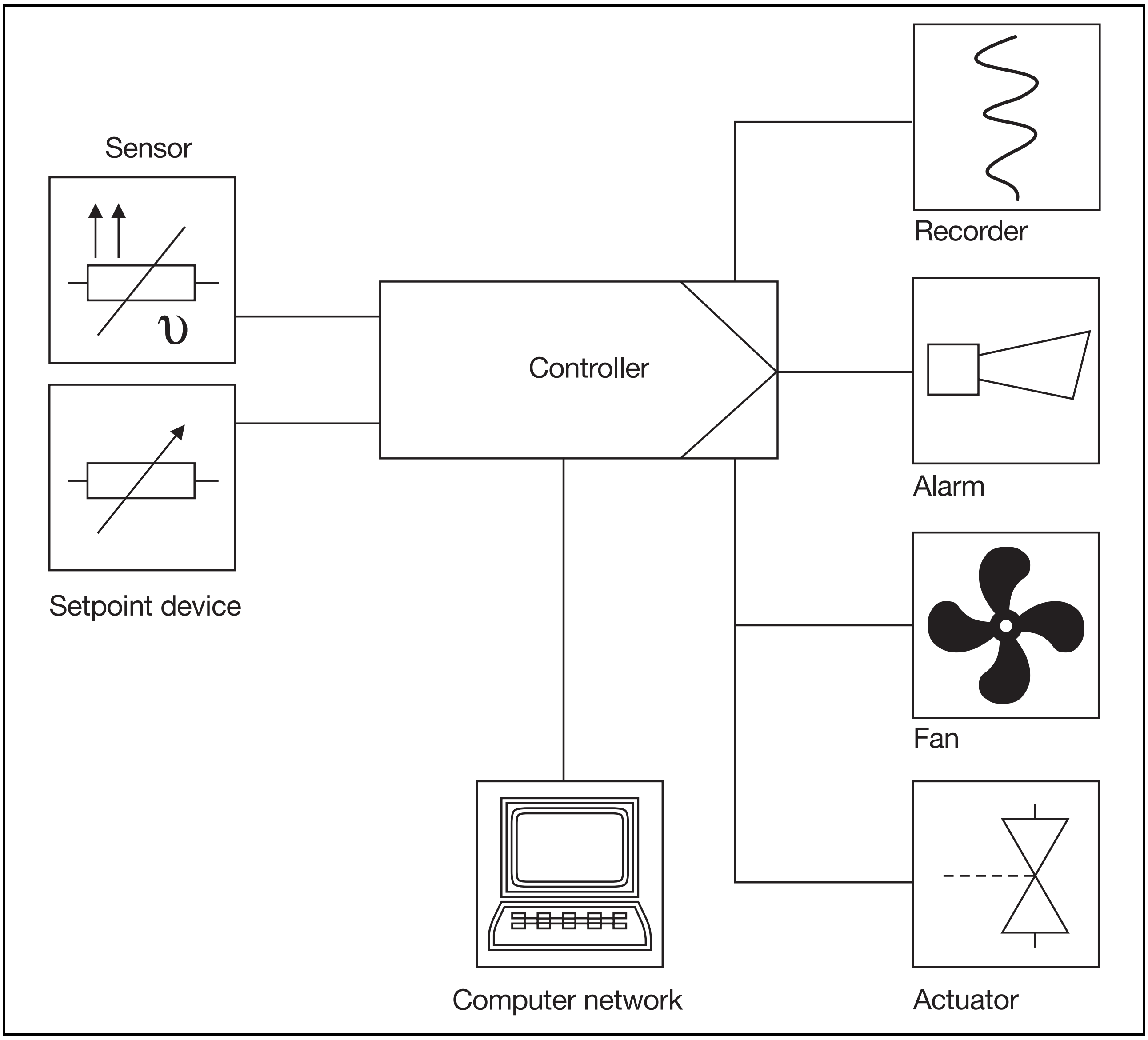


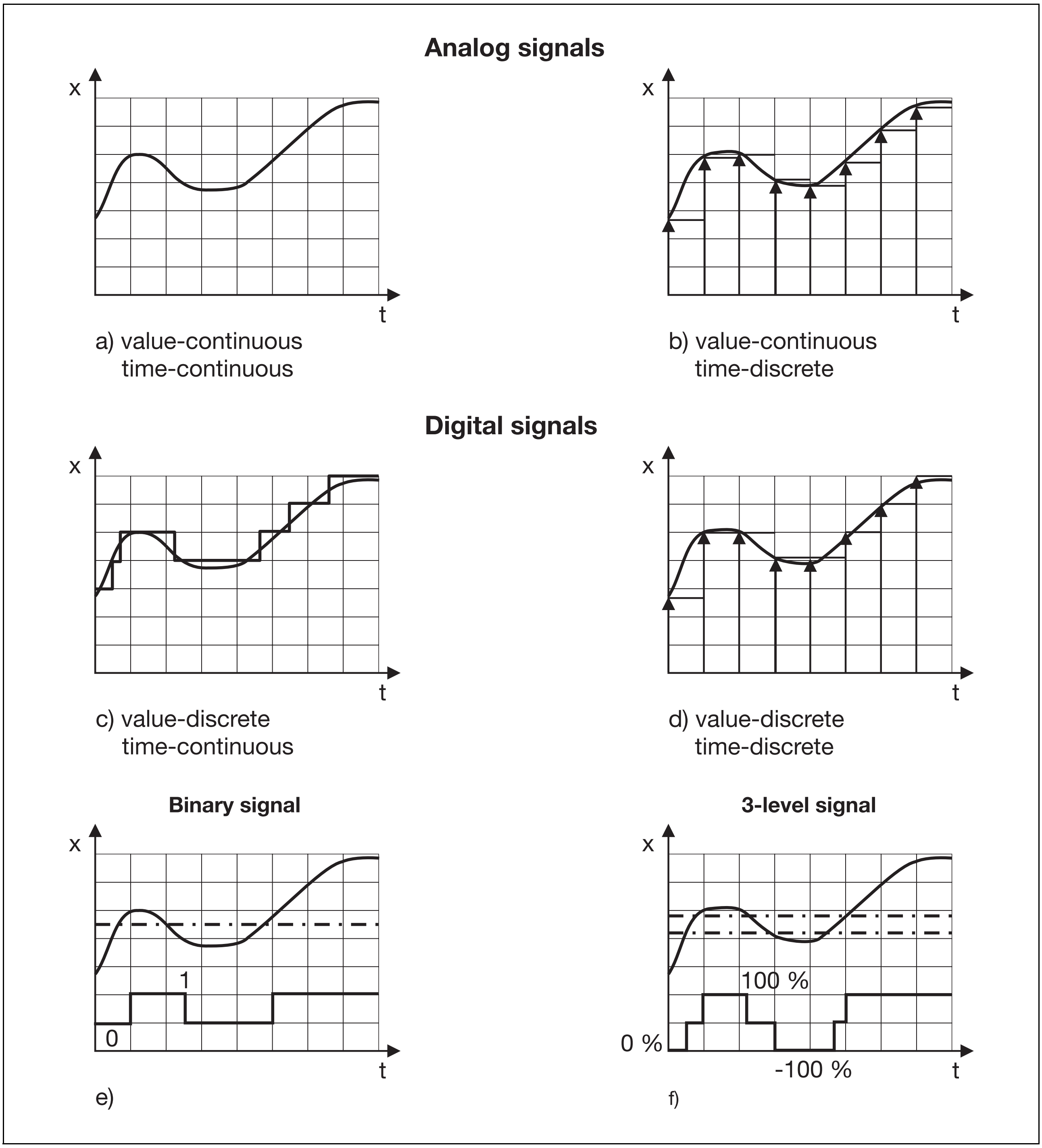
Fig. 6: Example for external connections to a controller

In many cases the process value requires further processing, e.g. for a recorder or for remote indication. Most controllers provide a process value output where the process variable is given out as a standard signal.
In order to signal movements of the process variable above or below certain values, the controllers are provided with so-called limit comparators (limit value or alarm contacts), which provide a signal if the process value infringes set limits. This signal can then be used to trigger alarms or similar equipment.

1.6 Analog and digital controllers

1.6.1 Signal types

Technical systems can be classified according to the type of signals at their inputs and outputs. The signals differ in their technical nature. In control systems we often find temperature, pressure, current or voltage as signal carriers which, at the same time, determine the units of measurement. The signals can be divided into different types, depending on their range of values and variation with time.


Analog signals

a) value-continuous
time-continuous

b) value-continuous
time-discrete

Digital signals

c) value-discrete
time-continuous

d) value-discrete
time-discrete

Binary signal

e)

3-level signal

f)

Fig. 7: Various signal forms

Analog signals

Analog signals have the greatest number of possible signal levels. The measuring device converts the process variable PV, for example a temperature, into a signal corresponding to this temperature. Each temperature value corresponds to a value of the electrical signal. If the temperature now varies continuously, the signal will also vary continuously. We call this a value-continuous signal.

The essential element in defining analog signals is that such signals pass continuously through a full range of values.

The time course is also continuous; at every instant the signal value corresponds to the temperature at this instant. It is therefore also a time-continuous signal (see Fig. 7a). In an application where the measuring device operates through a channel selection switch in which the contact arm is rotating continuously, the measured signal is only sampled at certain discrete times. The signal is then no longer time-continuous, but time-discrete (see Fig. 7b). On the other hand, the measurement remains value-continuous, since the measured signal is fully reproduced at each sampling instant.

Digital signals

Digital signals belong to the group of discrete signals. Here the individual signal levels are represented by numerals (digitally). This means that discrete signals can only take up a limited number of values. The variation of such discrete signals with time always appears as a series of steps.

A simple example of a system with discrete signals is the control system of a passenger lift or elevator, which can only take up discrete values for the height. This type of signal appears in control systems using computers, or digital controllers. The important feature here is that the analog signals can only be converted into digital signals by discretization of the signal level. There are no longer any intermediate values. However, assuming that the conversion takes place at an effectively unlimited speed, it is still possible to have a time-continuous signal (see Fig. 7c). In practice, the technical methods available limit the conversion to a time-discrete form. In other words, the analog/digital converter, used in digital control, only carries out the conversion process at discrete time intervals (sampling time). From the analog signal we obtain a result which is both value-discrete and time-discrete (see Fig. 7d).

It is quite evident that conversion of analog to digital signals in this way leads to a loss of information about the measured signal.

Binary signals

In their simplest form the signals can only have two states, and are therefore called binary signals. The control engineer is already familiar with this type of signal. The two states are normally described as “0” and “1”. Every switch used to turn a voltage on and off produces a binary signal as its output variable. Binary signals are also referred to as logic values and are assigned the values “true” and “false”. Virtually all digital circuits in electrical engineering work with this type of logic signals. Microprocessors and computers are built up from such elements, which only recognize these two signal states (see Fig. 7e).

3-state signals

Signals with the next higher information content after binary signals are 3-state signals (sometimes called tri-state signals). They are often used in connection with motors. Essentially, a motor can have three operating states. The motor can be stationary, or it can rotate clockwise or anticlockwise. Corresponding elements with a 3-state action are frequently found in control engineering, and are of great interest. Each of the three signal levels can have any desired value; in certain cases each signal level can be a positive signal, or the magnitude of the positive and negative signals can be different (see Fig. 7f).

1.6.2 Fundamental differences

A controller produces a relationship between the process variable PV and the setpoint SP, and derives from it the manipulating variable MV. There are a number of ways to carry out this task: mechanical, pneumatic, electrical, mathematical. The mechanical controller, for example, alters a signal through a lever system, the electronic controller through operational amplifiers. With the introduction of more powerful and low-cost microprocessors, another type of electrical controller has cornered the market in recent years, the microprocessor controller (digital controller). The measurement signal is no longer processed in an operational amplifier, but is now calculated using a microprocessor. The different structures found in these digital controllers can be described directly in mathematical terms.

The term “digital” means that the input variable, the process value, must initially be digitized, i.e. converted into a numerical value, as described in Chapter 1.6.1, before the signal can be processed by the microprocessor. The calculated output signal (the manipulating variable) then has to be converted back to an analog signal, by a digital to analog converter, to control the process, or alternatively, fed directly to a digital actuator. There is very little functional difference between digital and analog controllers, so this is not covered in-depth in the context of this book.
Use of a digital display is, in itself, not an adequate criterion for calling an instrument a digital controller. There are instruments which work on analog principles, but which have a digital display. They do not have an internal microprocessor to calculate the signals, and are therefore still referred to as analog controllers.

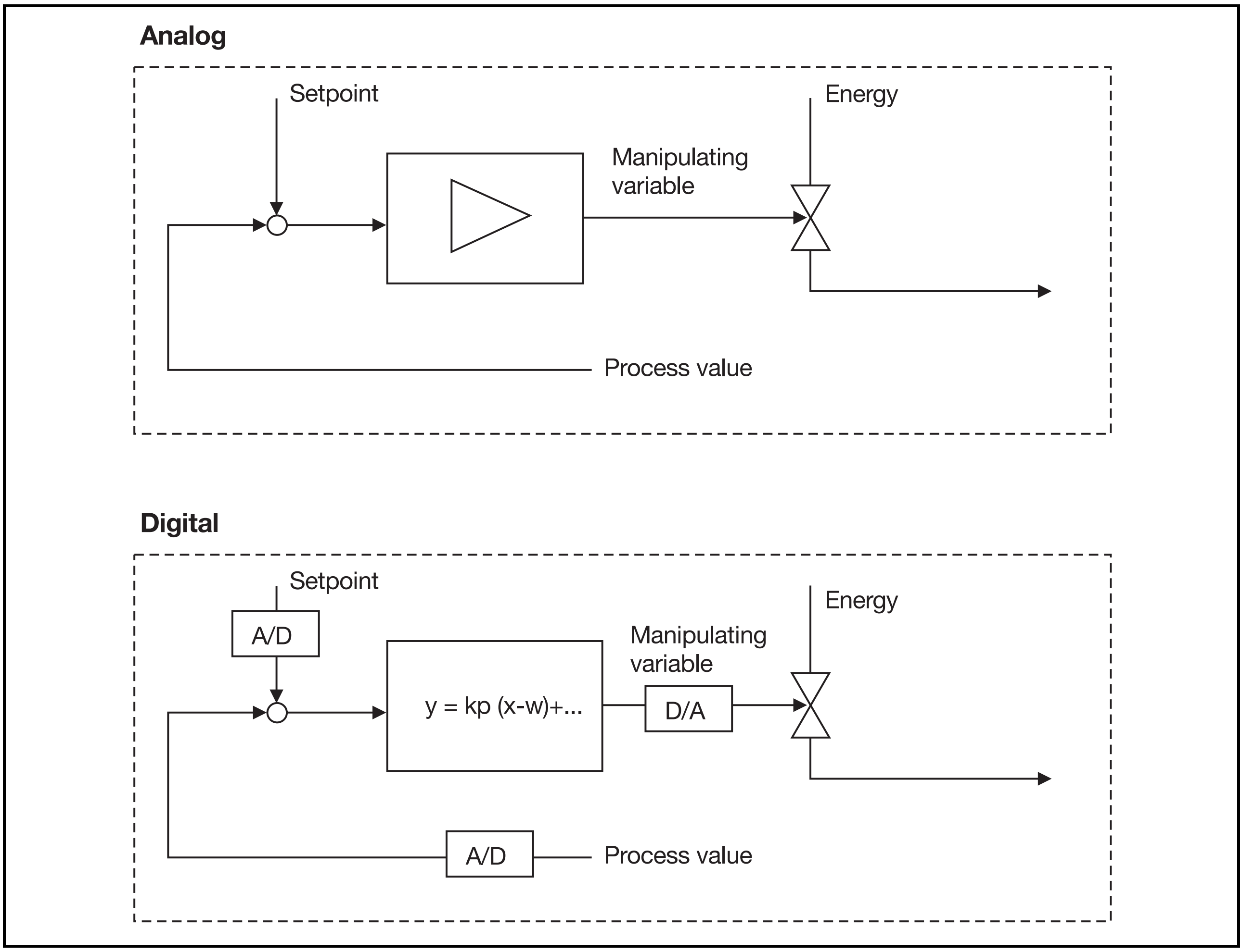
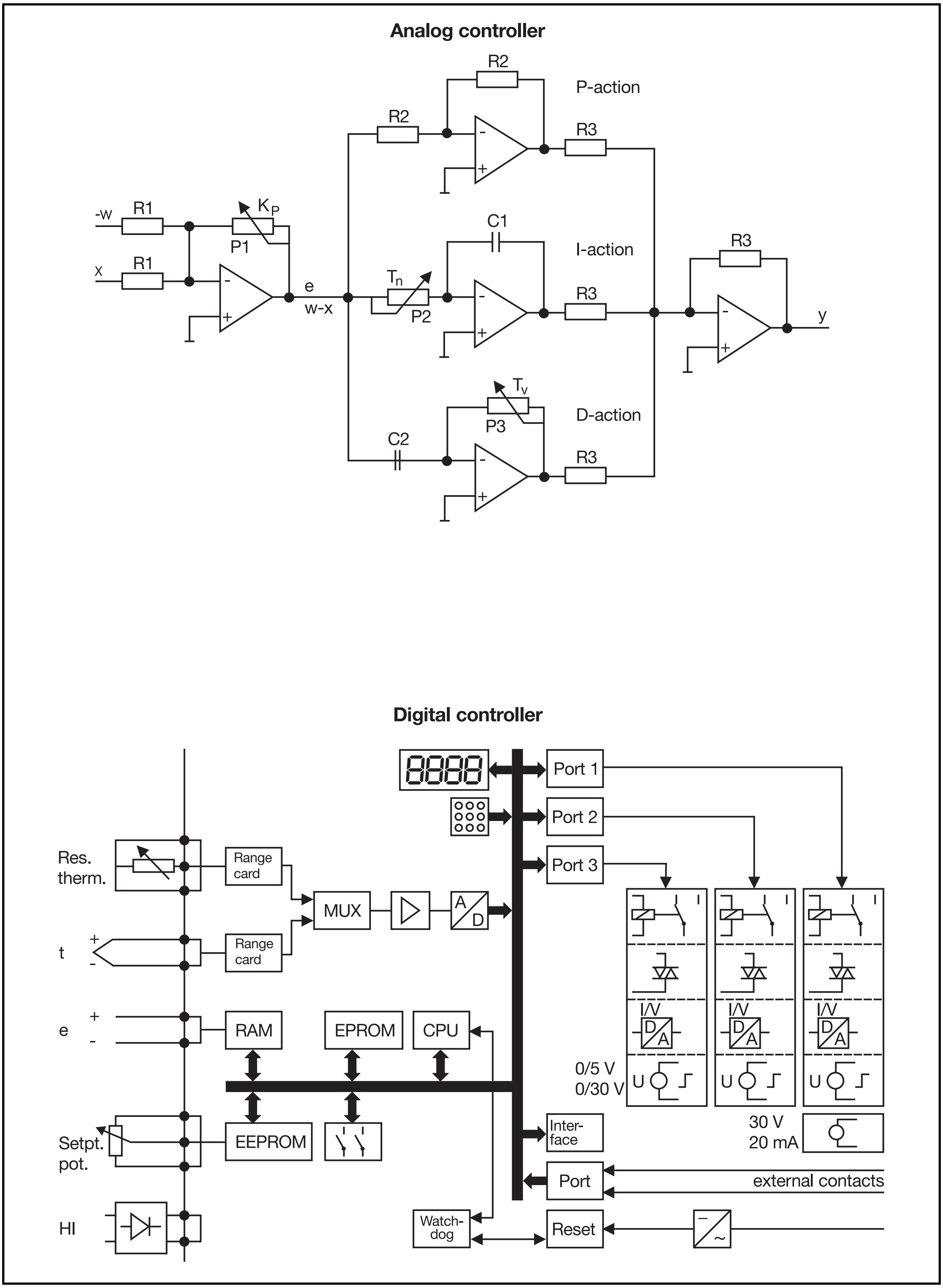


Fig. 8: Principle of analog and digital controllers



Analog controller

Digital controller

Fig. 9: Arrangement of analog and digital controllers

Advantages and disadvantages of digital controllers
Analog controllers are built up from operational amplifiers. The control parameters are set by means of potentiometers, trimmers or solder links. Controller structure and characteristics are largely predetermined by the design and construction. They are used where there is no requirement for very high accuracy, and where the required features of the controller, such as its dynamic action, are already known at the planning stage. Because of its speed of reaction, the analog controller has clear advantages in extremely fast control loops.

In digital controllers a microprocessor converts all analog inputs into numerical values, and uses them to calculate the manipulating variable. This has certain advantages compared with analog processing:

- increased accuracy of control, depending on the measurement signal and the technology used (e.g. A/D converter). Unlike components which are affected by tolerances and drift, the mathematical relationships used have a constant accuracy and are unaffected by ageing, variations in components and temperature effects.
- high flexibility in the structure and characteristics of the controller. Instead of having to adjust parameters or unsolder components, as in analog controllers, a digital controller can be modified by simply programming a new linearization, controller structure etc. by inputting numerical values
- facility for data transfer. There is often a need to modify or store information about process status variables, or pass it on for different uses, and this is very simple to achieve using digital technology. Remote setting of parameters through data systems, such as process management systems via a digital interface, is also quite simple.
- control parameters can be optimized automatically, under certain conditions.

Digital controllers also have disadvantages compared with controllers operating on analog principles. The digital display, normally standard with digital controllers, makes it more difficult to identify trends in process values. Digital instruments are more sensitive to electromagnetic interference. The processor needs a certain time to calculate parameters and to carry out other tasks, so that process values can only be read in at certain time intervals. The time interval between two successive readings of the process variable is referred to as the sampling time, and the term “sampling controller” is often used. Typical values of the sampling time in compact controllers are in the range 50 — 500msec. There are no technical reasons why controllers with sampling times less than 1 msec could not be built. If the process is relatively slow compared with the sampling time, the behavior of a digital controller is similar to that of an analog controller, since the sampling action is no longer noticeable.

1.7 Manipulating devices

The purpose of the manipulating device is to influence the process variable. Its main task is to regulate a mass or energy flow. Mass flows may have either gaseous or liquid state, e.g. natural gas, steam, fuel oil etc.

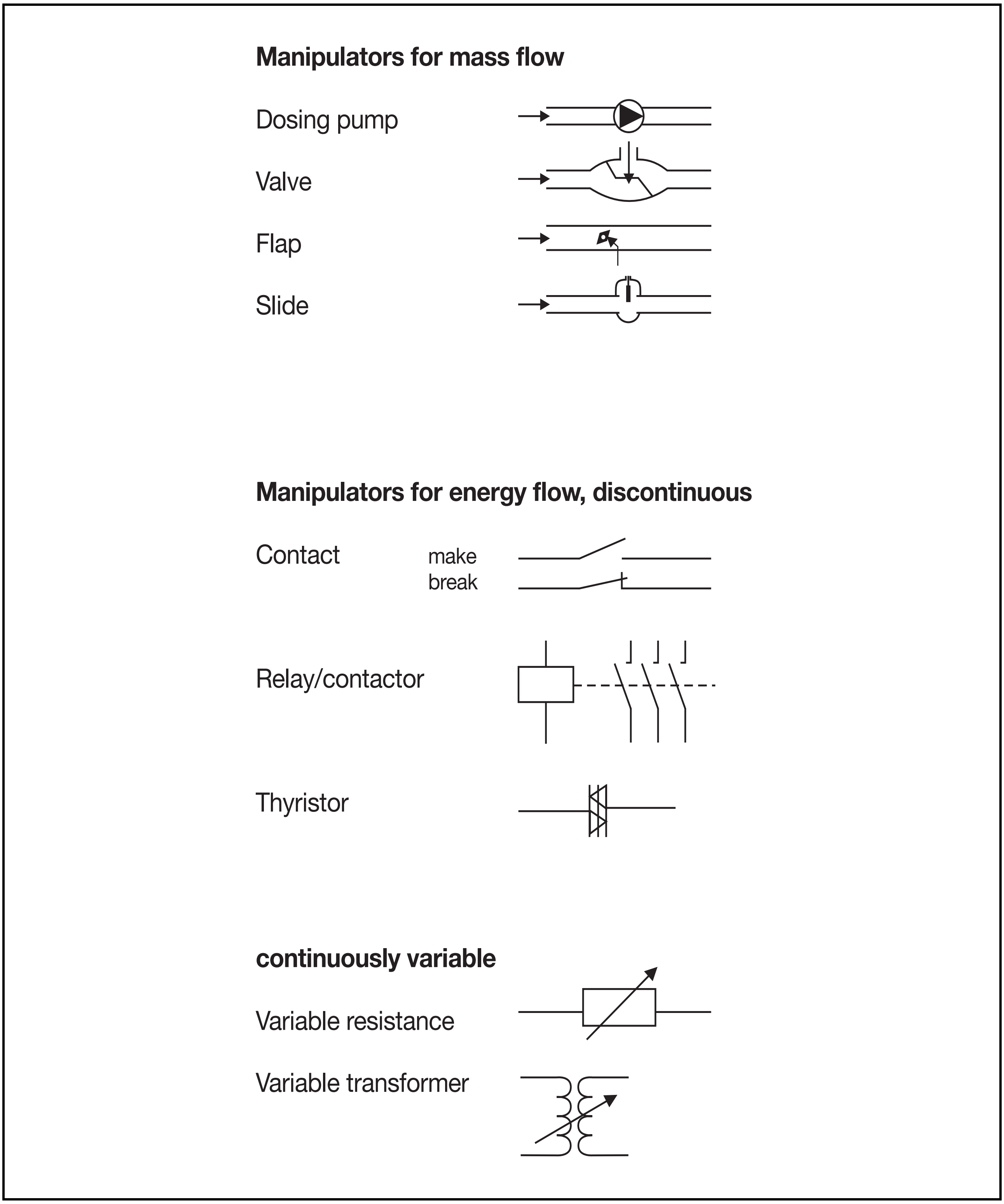


Fig. 10: Overview of different manipulators

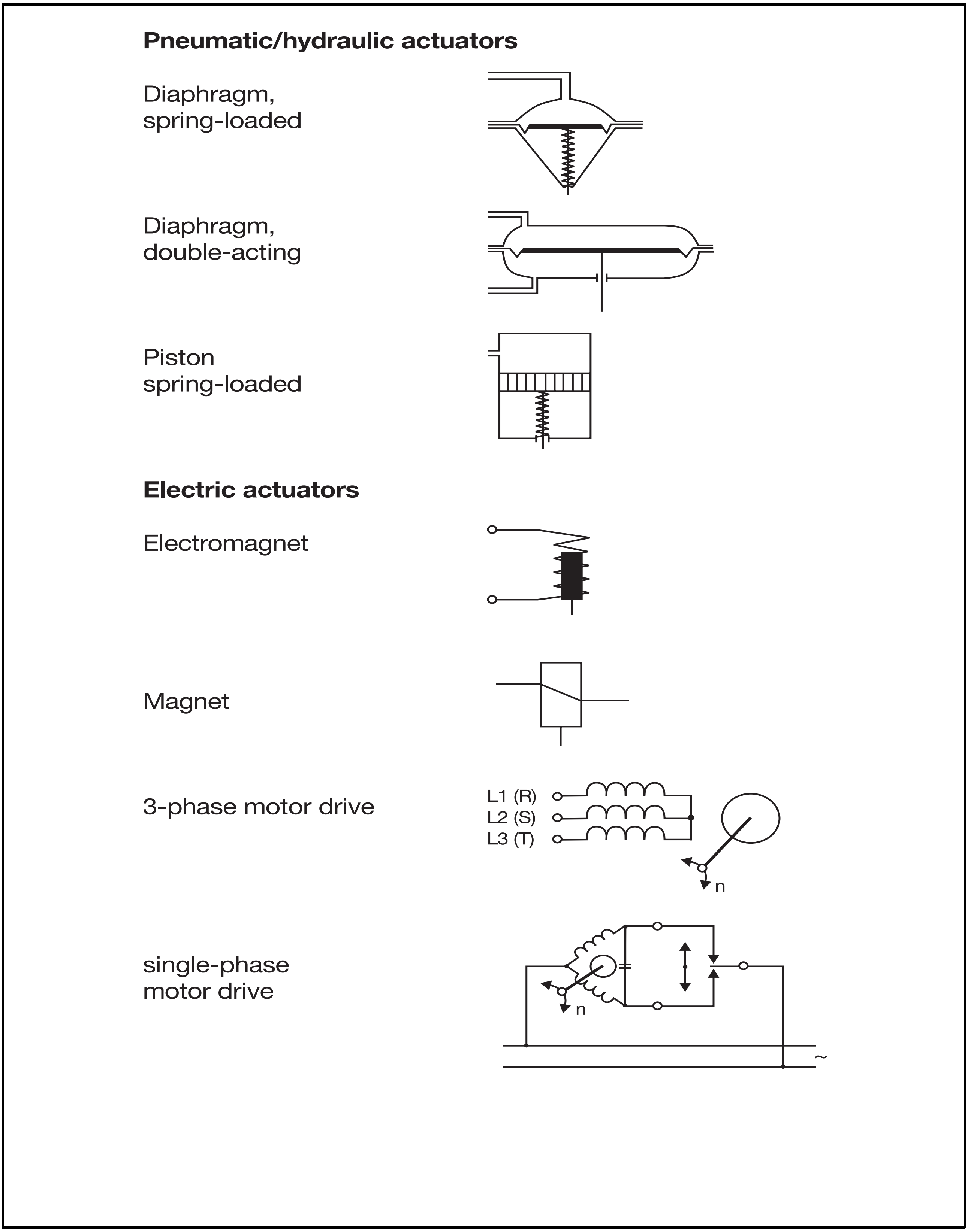


Fig. 11: Overview of different actuators

Energy flows often take the form of electrical energy. The energy supply can be varied discontinuously through contacts, relays or contactors, or continuously by means of variable transformers, variable resistors or thyristor units.

The manipulating device is frequently operated by an actuator where the controller cannot operate it directly, for instance, if it cannot provide sufficient power, or where the output of the controller is in the wrong energy form for driving the manipulator. The controller then operates either a mechanical-pneumatic or electrically powered driver. For example, the relays built into switching controllers can normally only handle currents up to 5A; external contactors or solid-state relays are then used to control the higher power required by the process.

Table 1 gives a brief overview of the various manipulators/drivers and their operation from suitable controllers.

Controller type	Operated manipulators/drivers
Continuous controllers	Adjustable resistor Thyristor unit Valves, flaps, slides Speed-controlled motors
2-state controllers	Contact Relay, contactor, solenoid valve Solid-state relay for heating, cooling etc.
3-state controllers (switching)	Heating, cooling, relays etc.
Modulating controllers	Actuating motors (AC, DC, 3-phase etc.)

Table 1: Controller types and manipulators/drivers

1.8 Other methods of achieving constant values

Automatic control, i.e. measurement of the process variable PV, comparison with the setpoint SP, and production of the manipulating variable MV, is not the only possible way of ensuring that a parameter is kept constant. There are several other methods of achieving this, which often offer a more cost-effective solution, as an alternative to automatic control.

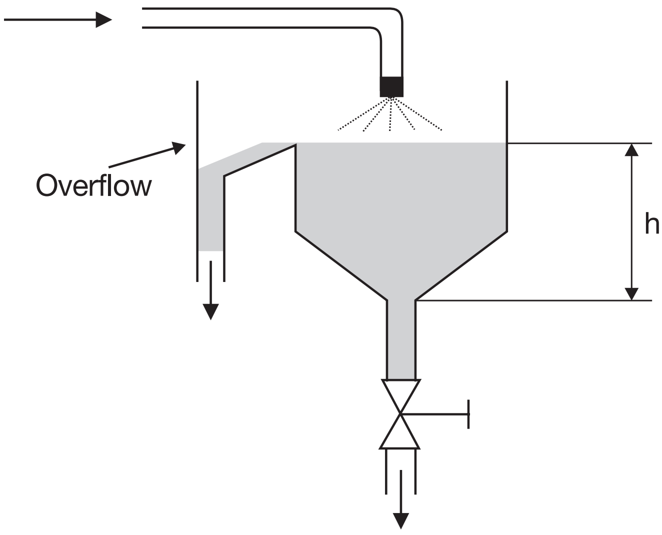
1.8.1 Utilizing physical effects

There are a number of physical values which remain constant over a wide range even when subjected to varying external influences. They include, for example, the melting point of a substance. While ice is melting, the temperature remains constant at 0°C. Physical effects like this are successfully used in many measurements, particularly in the laboratory. In this way, a temperature can be maintained constant to a high degree of accuracy, without the expense of sophisticated control equipment.

1.8.2 Constructional measures

To some extent, parameters can be held constant through suitable constructional features. For example, a constant liquid level can be maintained in a container or tank, in spite of variations in the inflow rate, just by providing an overflow (see Fig.12a). Another example is a swimming pool, where the water level can be maintained constant by providing an overflow all round the pool.

a) Constant value by constructional measures



b) Constant value by operation

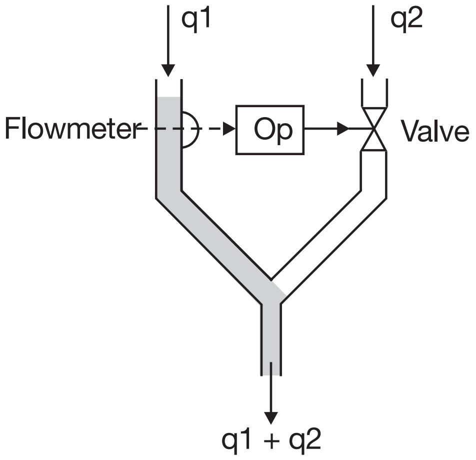


Fig. 12: Methods of achieving constant values

1.8.3 Maintaining constant values by operation

As already discussed in Chapter 1.3, “Operation and control”, a parameter can be kept constant by suitable operation. An example could be to maintain a constant furnace temperature. Assuming a constant voltage, i.e. a steady power supply to an electrically heated furnace, the setting of an energy regulator can be varied to provide different furnace temperatures. By noting these temperatures, i.e. by producing a temperature scale and attaching this to the energy regulator, we can then set any desired furnace temperature. As the adjustment is made by hand, we refer to this as manual operation. The input parameter in this form of temperature control is the setting of the energy regulator, the output variable is the furnace temperature, which can be displayed on a suitable indicating instrument (see Fig. 1).

Adjustment of the input parameters need not take place manually, but can be automated: this is then called automatic operation. As an example, take the control of a mixing process. The task consists of producing a flow Q_2 which is proportional to an externally determined flow Q_1 in order to achieve a particular mixture ratio (see Fig. 12b). Here the flow Q_1 is determined as the input variable, and is applied to the operating equipment. The output of the operating equipment operates a manipulator which changes the flow Q_2.

From this it is clear that a process variable can also be kept constant by simple operation. However, it should be borne in mind that operation has considerable disadvantages compared with automatic control. If the process is subjected to a disturbance, or there is a change in the transfer characteristic of the manipulating device, there can be undesirable changes in output, even with a fixed transfer action between input and output variables.

1.9 Main areas of control engineering

Today, control engineering has applications in almost every area of technology. In Chapter 1.1 we have already seen that these different applications have certain common features, which can be described through a standard procedure. A number of main application groups have evolved as a result of differing process variables, stabilization rates, types of machinery and equipment, and certain special features of the application field.

- industrial process control
- drive control (speed control)
- control of electrical variables
- positional control
- course control

Fig. 13: Main areas of control engineering

Industrial process control
This heading covers the control of temperature, pressure, flow, level etc. in many different industrial applications. If we look at the criterion “stabilization time”, this can have an order of magnitude ranging from milliseconds, e.g. in pressure control, up to several hours in the case of temperature control of larger installations (industrial furnaces).

Drive control (speed control)
This group includes speed control of motors on different machines and installations, such as in plastics manufacture, paper production or textile machinery. Specially designed controllers are normally used for these applications, since they have to remain stable during fast disturbances in the range of tenths of seconds.

Control of electrical variables
This refers to stabilization of electrical parameters, e.g. voltage, current, power or even frequency. This type of equipment is used in power generation or to stabilize characteristic values in supply networks. Here again there are very fast disturbances, in the range of tenths of seconds or even shorter.

Position control
This involves the positioning of tools, workpieces or complete assemblies, either in two or three dimensions. Examples include a milling machine and the positioning of guns on ships and tanks. Once again, stabilization at the setpoint must be very rapid and very accurate.

Course control
The control of the course of ships or planes. Here the controller has to satisfy special demands, such as high processing speed and operational safety, combined with low weight.

1.10 Tasks of the control engineer

So far we have discussed various concepts and designations, the differences between operation and control and the various forms of controllers and manipulators. We can now summarize the tasks a control engineer has to face in practice.

The most important tasks for a control engineer are as follows:

- Determining the process variable
- Checking whether automatic control offers significant advantages
- Determining the measurement site
- Assessing the disturbances
- Selecting the manipulator
- Selecting a suitable controller
- Installation of the controller in accordance with applicable regulations
- Starting up, adjusting parameters, optimizing

Fig. 14: Tasks of a control engineer

By control engineer, we don’t mean specialist engineers and technicians from universities or research departments, who work in the laboratory developing controllers, control algorithms or special control circuits. Specialists such as these require a much more extensive knowledge. Instead we are addressing people working on site who may have to optimize an unsatisfactory control loop or convert from manual operation to automatic control, or those involved in the design of a control loop for a new installation. In most cases these operations can be tackled without using advanced mathematics. All that is really needed is a basic understanding, pragmatic rules and knowledge gained from past experience.

As a general principle for planning a control system, it should be borne in mind that when high-performance demands are placed on a controller, the costs will increase considerably.

2.1 Dynamic action of technical systems

The process is the element of a system which has to be controlled in accordance with the application duty. In practice, the process represents either an installation or a manufacturing process which requires controlling. Normally, the process covers a number of elements within a system. The input is the manipulating variable y received from the control device. The output is represented by the process variable x. As well as these two variables there are the disturbances z which affect the process to some extent, through external influences or process-dependent variations.

An example of a process is a gas-fired furnace (see Fig. 15). At the start of the process is the valve, which has as its input the manipulating variable of the controller. The valve controls the gas flow to the burner. The burner produces heat energy by burning the gas, which brings the charge up to a higher temperature. If the temperature in the charge is measured (process value), this also forms part of the process. The final component of the process here is the sensor, which has the job of converting the temperature into an electrical signal. Disturbances here are all the variables in the process which, when they change, result in a different temperature for the same valve setting.

Example: If the manipulating variable is just large enough to give the required temperature in the charge, and a disturbance occurs due to a fall in outside ambient temperature, then, if the manipulating variable is not changed, the temperature in the charge will also be lower.

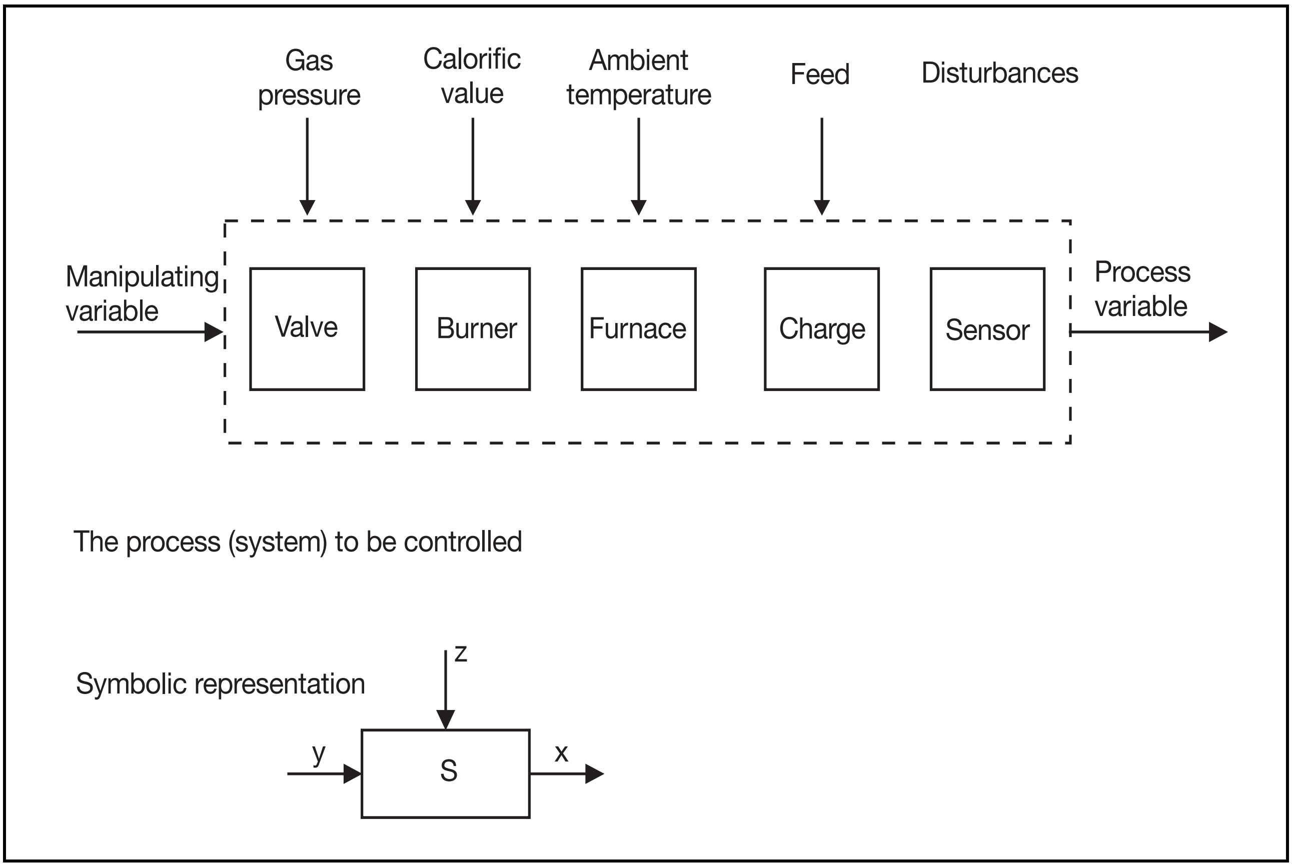


Fig. 15: Input and output variables of a process

When designing a control loop, it is important to know how the process responds when there is a change in one of the influencing variables mentioned above. On the one hand, it is of interest to know the new process value reached when stable conditions have been attained, following such changes. On the other hand, it is also important to find out how the process value varied with time during the transition to the new steady-state value. A knowledge of the characteristics determined by the process is essential and can help to avoid difficulties later on, when designing the process.

2 The process

Although processes have many different technical arrangements, they can be broadly categorized by the following features:

- with and without self-limitation,

- with and without dead time or timing elements,

- linear or non-linear.

In most cases, however, a combination of individual characteristics will be present.

An accurate characterization and detailed knowledge of the process is a prerequisite for the design of controls and for the optimum solution of a control task. It is not possible to select suitable controllers and adjust their parameters, without knowing exactly how the process behaves. The description of the dynamic action is important to achieve the objective of control engineering, i.e. to control the dynamic behavior of technical dynamic systems and to impose a specific transient response on the technical system.

Static characteristic
The static behavior of a technical system can be described by considering the output signal in relation to the input signal. In other words, by determining the value of the output signal for different input signals. With an electrical or electronic system, for instance, a voltage from a voltage source can be applied to the input, and the corresponding output voltage determined. When considering the static behavior of control loop elements, it is of no importance how a particular control element reaches its final state. The only comparison made is limited to the values of the input and output signals at the end of the stabilization or settling time.
When measuring static characteristics, it is interesting to know, amongst other things, whether the particular control loop element exhibits a linear behavior, i.e. whether the output variable of the control element follows the input proportionally. If this is not the case, an attempt is made to determine the exact functional relationship. Many control loop elements used in practice exhibit a linear behavior over a limited range. With special regard to the process, this means that when the manipulating variable MV is doubled, the process value PV also doubles; PV increases and decreases equally with MV.
An example of a transfer element with a linear characteristic is an RC network. The output voltage U2 follows the applied voltage U1 with a certain dynamic action, but the individual final values are proportional to the applied voltages (see Fig. 16). This can be expressed by stating that the process gain of a linear process is constant, as a change in the input value always results in the same change in the output value.

However, if we now look at an electrically heated furnace, we find that this is in fact a non-linear process. From Fig. 16 it is clear that a change in heater power from 500 to 1000W produces a larger temperature increase than a change in power from 2000 to 2500W. Unlike the behavior of an RC network, the furnace temperature does not increase to the same extent as the power supplied, as the heat losses due to radiation become more pronounced at higher temperatures. The power must therefore be increased to compensate for the energy losses. The transfer coefficient or process gain of this type of system is not constant, but decreases with increasing process values. This is covered in more detail in Chapter 2.8.

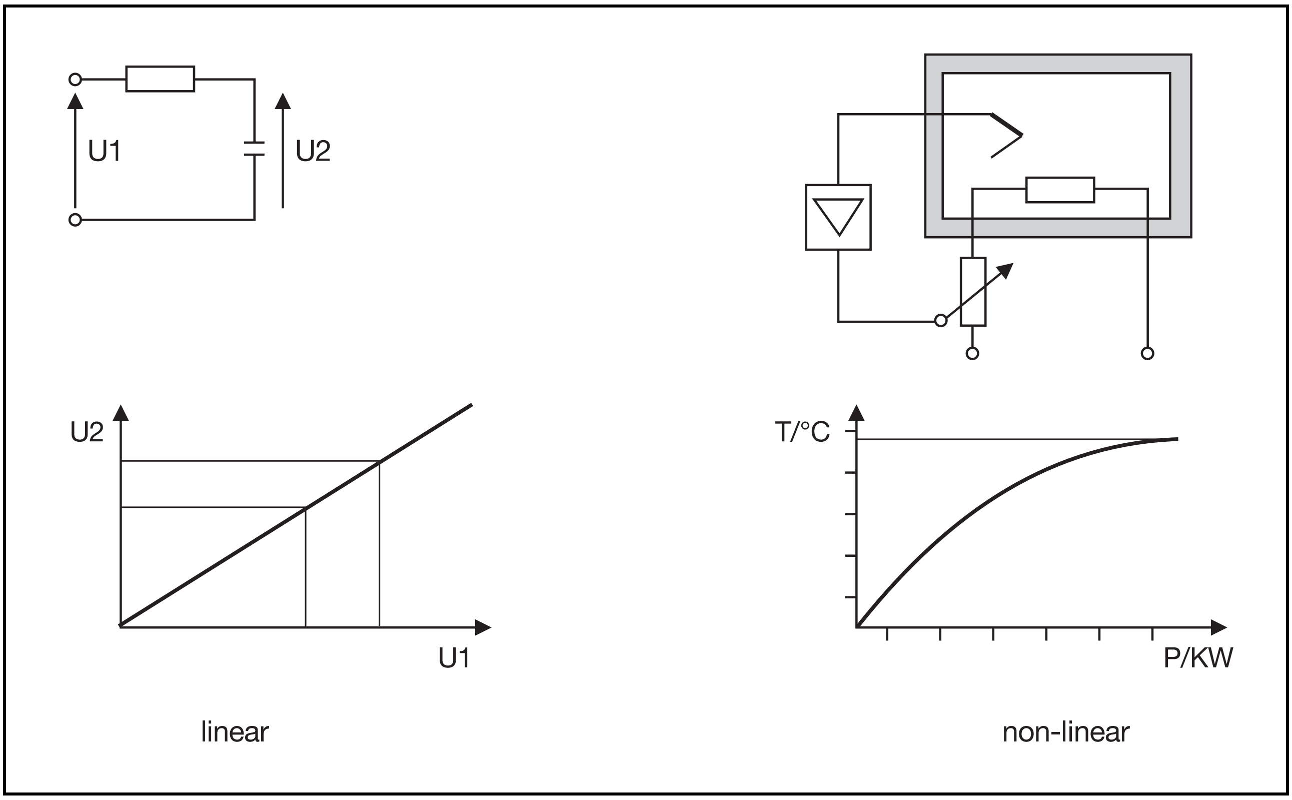


Fig. 16: Linear and non-linear characteristics

Dynamic characteristic

The dynamic response of the process is decisive for characterizing the control loop. The dynamic characteristic describes the variation in the output signal of the transfer element (the process) when the input signal varies with time. In theory, it is possible for the output variable to change immediately and to the same extent as the input variable changes. However, in many cases, the system responds with a certain delay.

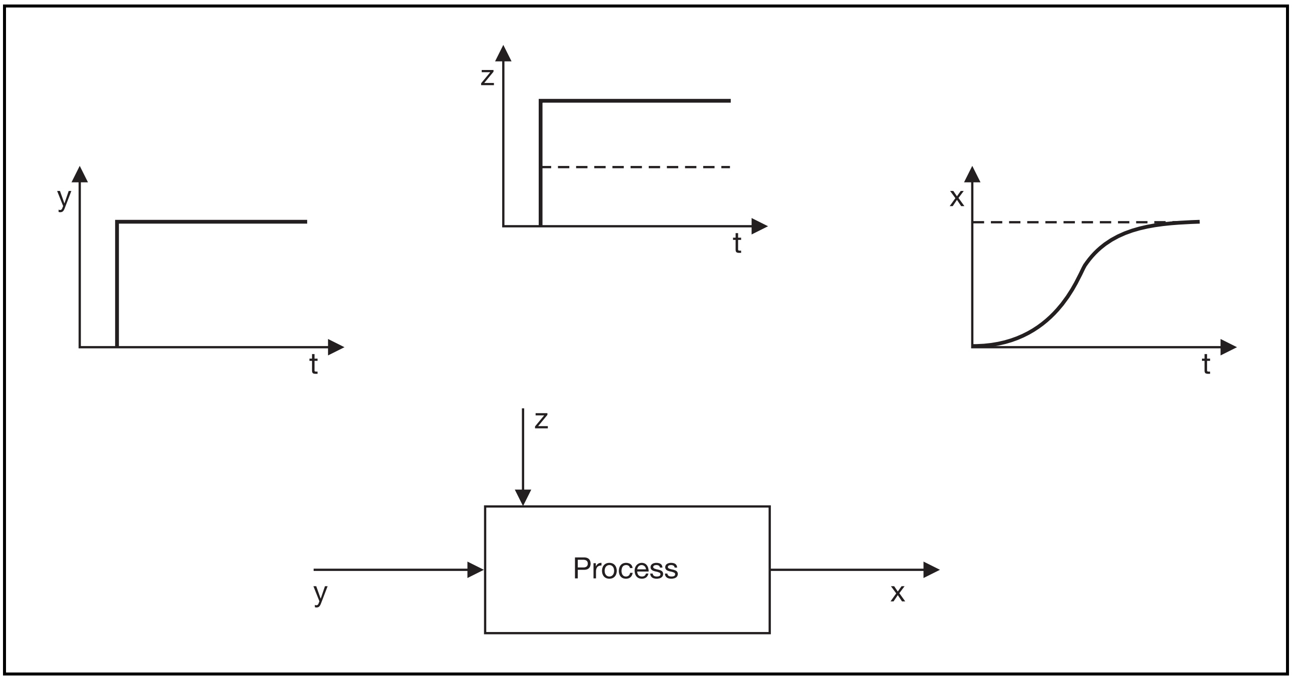


Fig. 17: Step response of a process with self-limitation

The simplest way of establishing the behavior of the output signal is to record the variation of the process value PV with time, after a step change in the manipulating variable MV. This “step response” is determined by applying a step change to the input of the process, and recording the variation of PV with time. The step change need not necessarily be from 0 to 100 %; step changes over smaller ranges can be applied, e.g. from 30 to 50 %. The dynamic behavior of processes can be clearly predicted from this type of step response, which will be discussed in more detail in Chapter 2.6.

2.2 Processes with self-limitation

Processes with self-limitation respond to a change in the manipulating variable or to a disturbance by moving to a new stable process value. This type of process can dissipate the energy supplied and achieve a fresh equilibrium.

A classic example is a furnace where, as the heating power is increased, the temperature rises until a new equilibrium temperature is reached, at which the heat lost is equal to the heat supplied. However, in a furnace, it takes some time to achieve the new equilibrium following a step change in the manipulating variable. In processes without delays, the process value immediately follows the manipulating variable. The step response of such a process then has the form shown in Fig. 18.

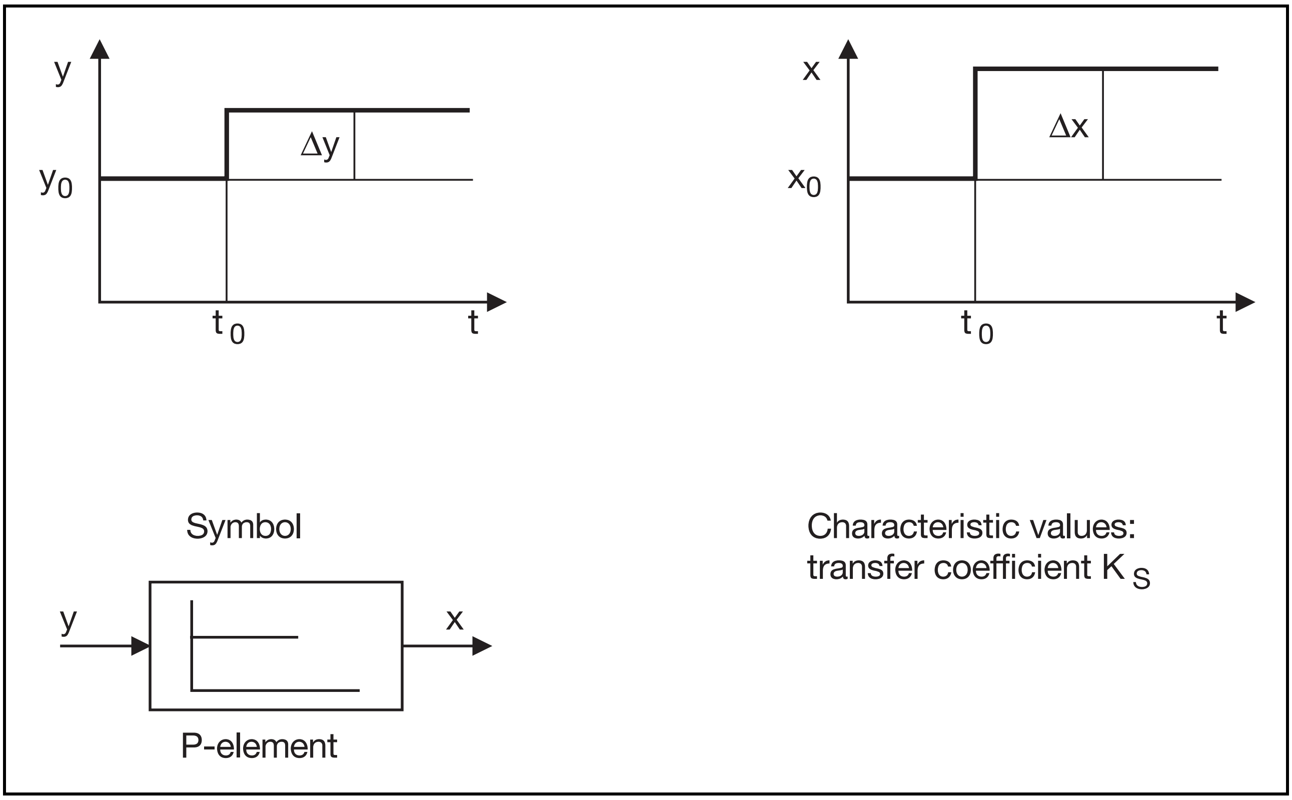


Fig. 18: Process without delay; P process

In this type of process with self-limitation, the process value PV is proportional to the manipulating variable MV, i.e. PV increases to the same extent as MV. Such processes are often called proportional processes or P processes. The relationship between process value x and manipulating variable y is given by:

$$\Delta x = K_S \bullet \Delta y$$

The factor K_S is known as the process gain (transfer coefficient). The relationship will be discussed in more detail in Chapter 2.8.

Examples of proportional processes are:

- mechanical gearing without slip
- mechanical transmission by lever
- transistor (collector current I_c follows the base current I_B with virtually no delay)

2.3 Processes without self-limitation

A process without self-limitation responds to a change in the manipulating variable or to a disturbance by a permanent constant change in the process value. This type of process is found in the course control of an aircraft, where a change in the manipulating variable (rudder deviation) produces an increase in the process value deviation (course deviation) which is proportional to time. In other words, the course deviation continually increases with time (see Fig. 19).

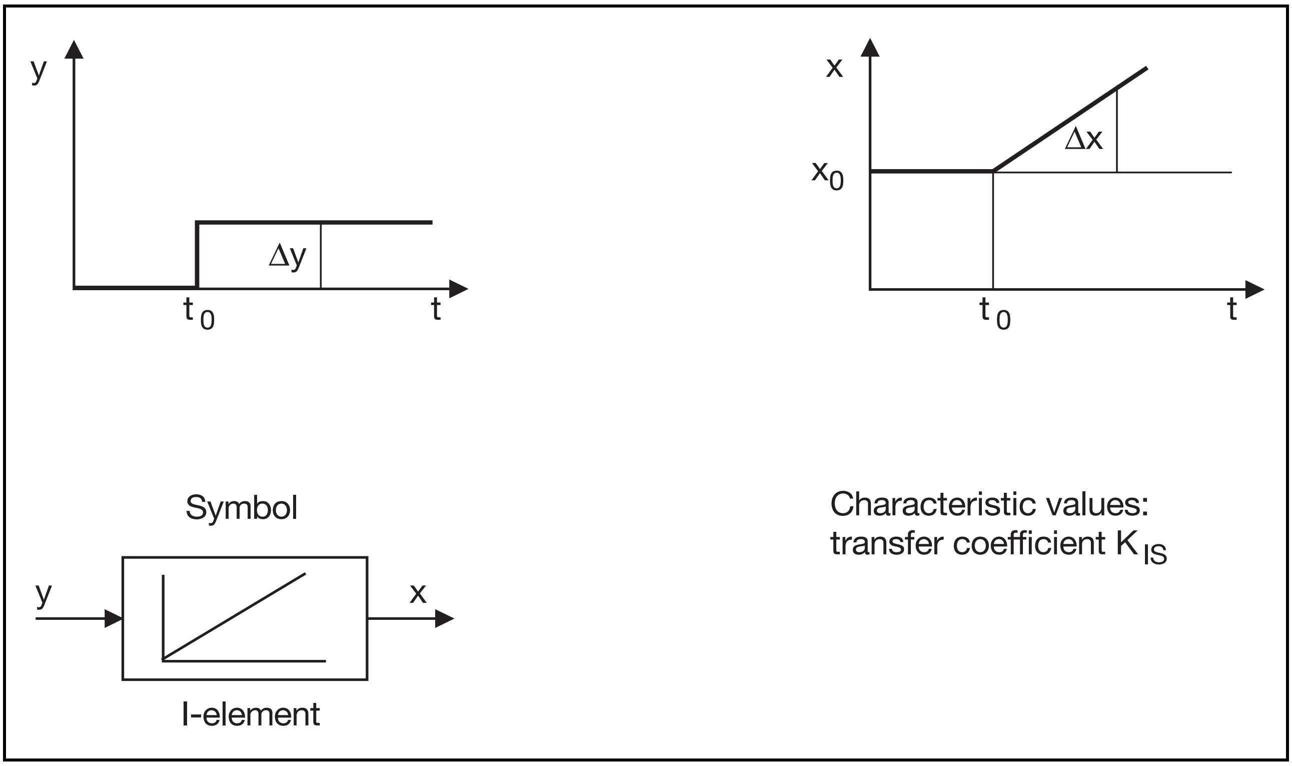


Fig. 19: Process without self-limitation; I process

Because of this integrating effect, such processes are also called integral processes or I processes. In this type of process, the process value increases proportionally with time as a result of a step change Δy in the manipulating variable. If the change in MV is doubled, the process value will also double after a certain time.

If Δy is constant, the following relationship applies:

$$\Delta x = K_{IS} \bullet \Delta y \bullet t$$

K_{IS} is called the transfer coefficient of the process without self-limitation. The process value now increases proportionally with both the manipulating variable change Δy, as in a process with self-limitation, and also with time t.

Additional examples of processes without self-limitation are:

- an electric motor driving a threaded spindle
- the liquid level in a tank (see Fig. 20)

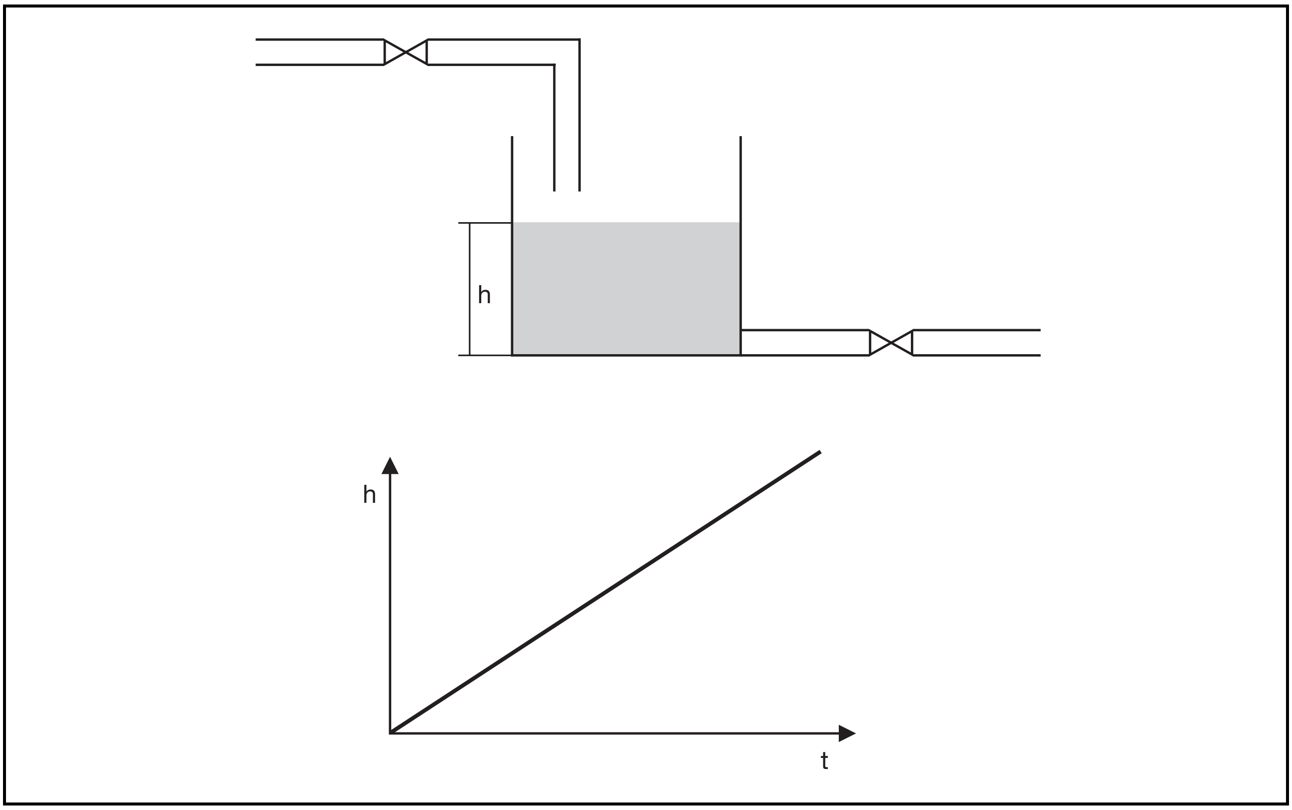


Fig. 20: Liquid level in a tank; I process

Probably the best known example of a process without self-limitation is a liquid container with an inflow and an outflow. The outlet valve, which here represents the disturbance, is assumed to be closed initially. If the inlet valve is now opened to a fixed position, the liquid level (h) in the container will rise steadily at a uniform rate with time.

The level in the container rises faster as the inflow rate increases. The water level will continue to rise until the container overflows. In this case, the process does not self-stabilize. Taking the effect of outflow into consideration, no new equilibrium is reached after a disturbance (except when inflow = outflow), unlike the case of a process with self-limitation.

In general, processes without self-limitation are more difficult to control than those with self-limitation, as they do not stabilize. The reason is, that following an overshoot due to an excessive change in MV by the controller, the excessive PV cannot be reduced by process self-limitation. Take a case where the rudder is moved too far when making a course adjustment, this can only be corrected by applying an opposing MV. An excessive change in MV could cause the process value to swing back below the desired setpoint, which is why control of such a process is more difficult.

2.4 Processes with dead time

In processes with a pure dead time the process only responds after a certain time has elapsed, the dead time T_t. Similarly, the response of the process value is delayed when the manipulating variable changes back (see Fig. 21).

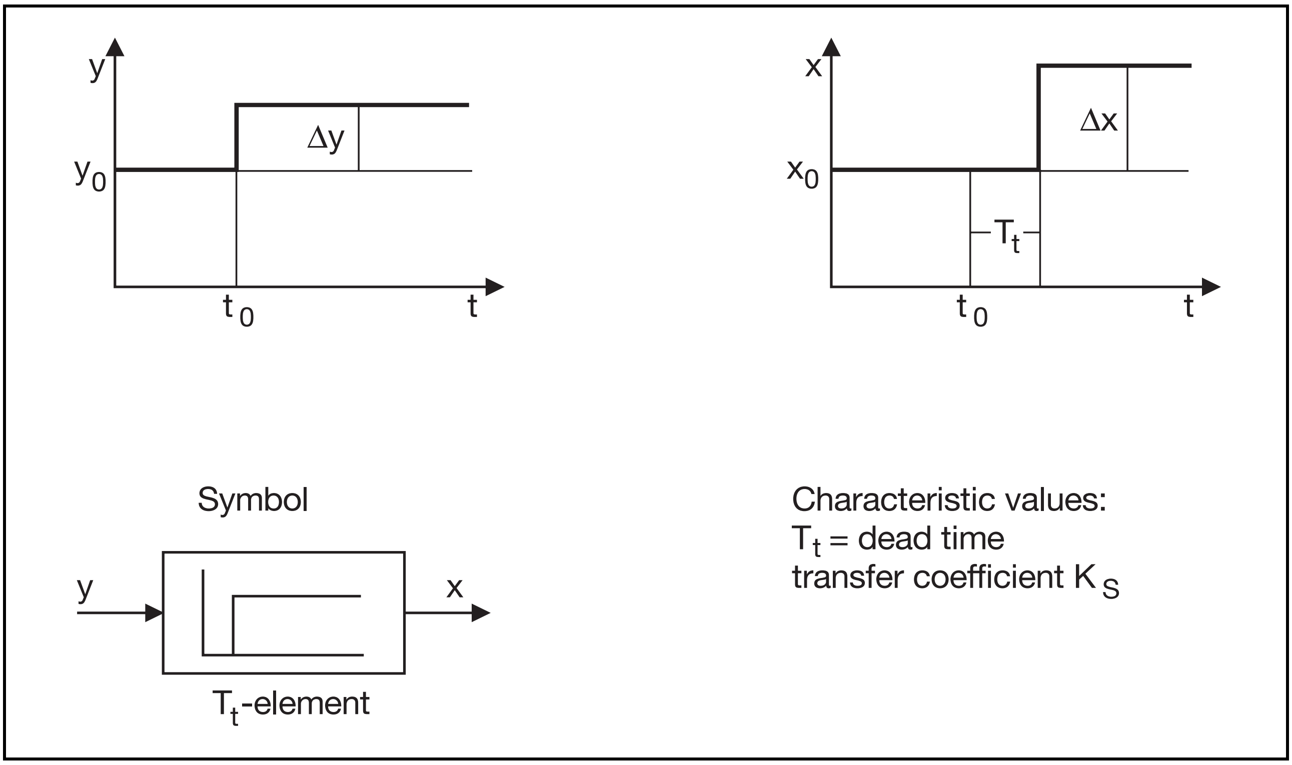


Fig. 21: Process with dead time; T_t process

A typical example here is a belt conveyor, where there is a certain time delay before a change in the chute feed rate is recorded at the measurement location (see Fig. 22).

Systems like this, which are affected by a dead time, are called T_t processes. The relationship between process value x and manipulating variable y is as follows:

$$\Delta x = K_S \bullet \Delta y$$

but delayed by the dead time T_t.

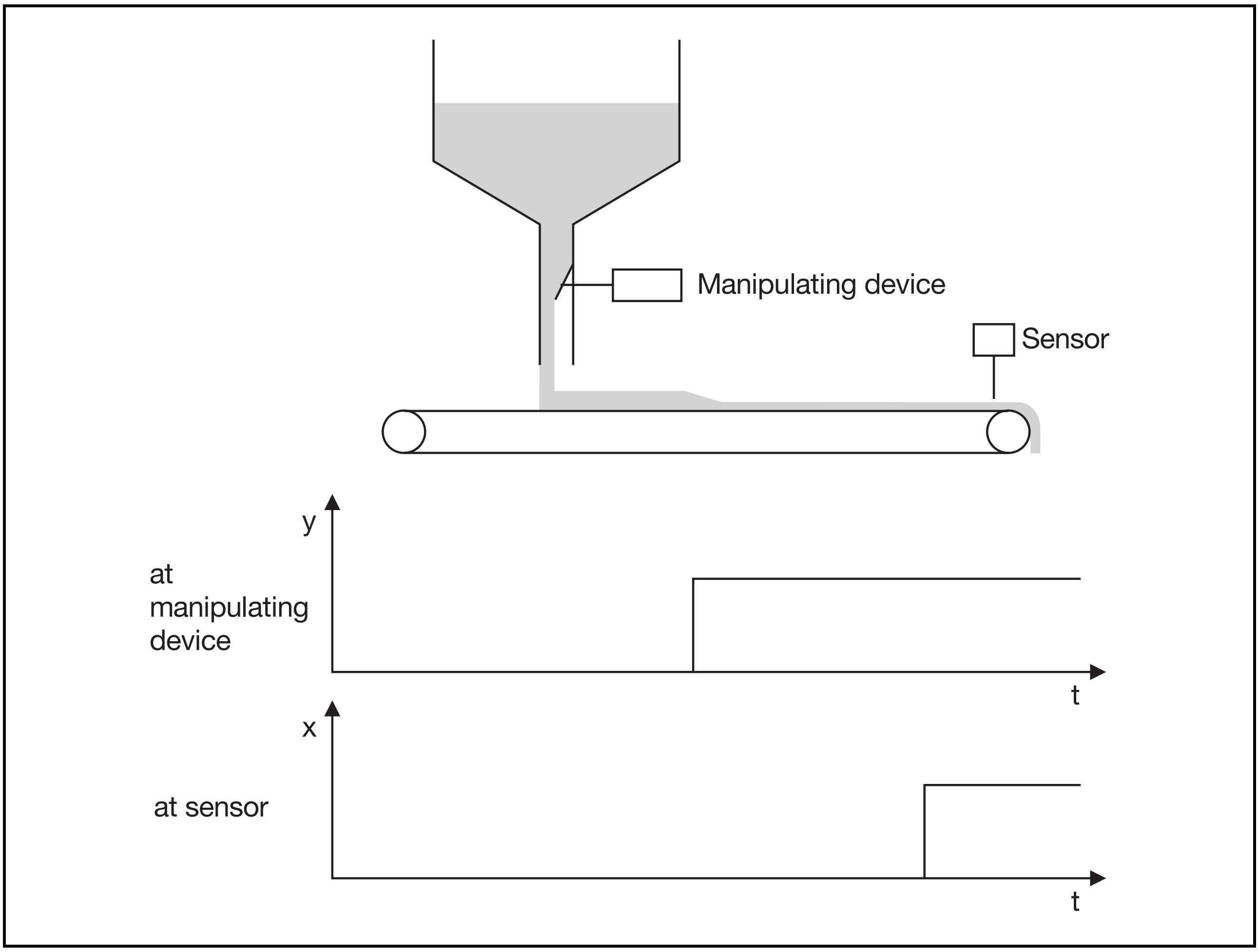


Fig. 22: Example of a process with dead time; belt conveyor

Another example is a pressure control system with long gas lines. Because the gas is compressible, it takes a certain time for a pressure change to propagate. By contrast, liquid-filled pipelines have virtually no dead time, since any pressure change is propagated at the speed of sound. Relay switching times and actuator stroke times also introduce delays, so that such elements in the control loop frequently give rise to dead times in the process.

Dead times pose a serious problem in control engineering, since the effect of a change in manipulating variable is only reproduced in the process variable after the dead time has elapsed. If the change in manipulating variable was too large, there is a time interval before this is noticed and acted on by reducing the manipulating variable. However, if this process input is then too small, it has to be increased once more, again after the dead time has elapsed, and so the sequence continues. Systems affected by dead time always have a tendency to oscillate. In addition, dead times can only really be compensated for by the use of very complex controller designs. When designing and constructing a process, it is very important that dead times are avoided wherever possible. In many cases this can be achieved by a suitable arrangement of the sensor and the application point of the manipulating variable. Thermal and flow resistances should be avoided or kept to a minimum. Always try to mount the sensor at a suitable location in the process where it will read the average value of the process conditions, avoiding dead spaces, thermal resistances, friction etc.

Dead times can occur in processes with and without self-limitation.

2.5 Processes with delay

In many processes there is a delay in propagation of a disturbance, even when no dead time is present. Unlike the case explained above, the change does not appear to its full extent after the dead time has elapsed, but varies continuously, even following a step change in the disturbing influence.

Continuing with the example of a furnace, and looking closely at the internal temperature propagation:

If there is a sudden change in heating power, the energy must first of all heat up the heating element, the furnace material and other parts of the furnace until a probe inside the furnace can register the change in temperature. The temperature therefore rises slowly at first until the temperature disturbance has propagated and there is a constant flow of energy. The temperature then continues to rise. Over a period of time the temperature of the heating element and the probe come closer and closer together; the temperature increases at a lower rate and approaches a final value (see Fig. 23).

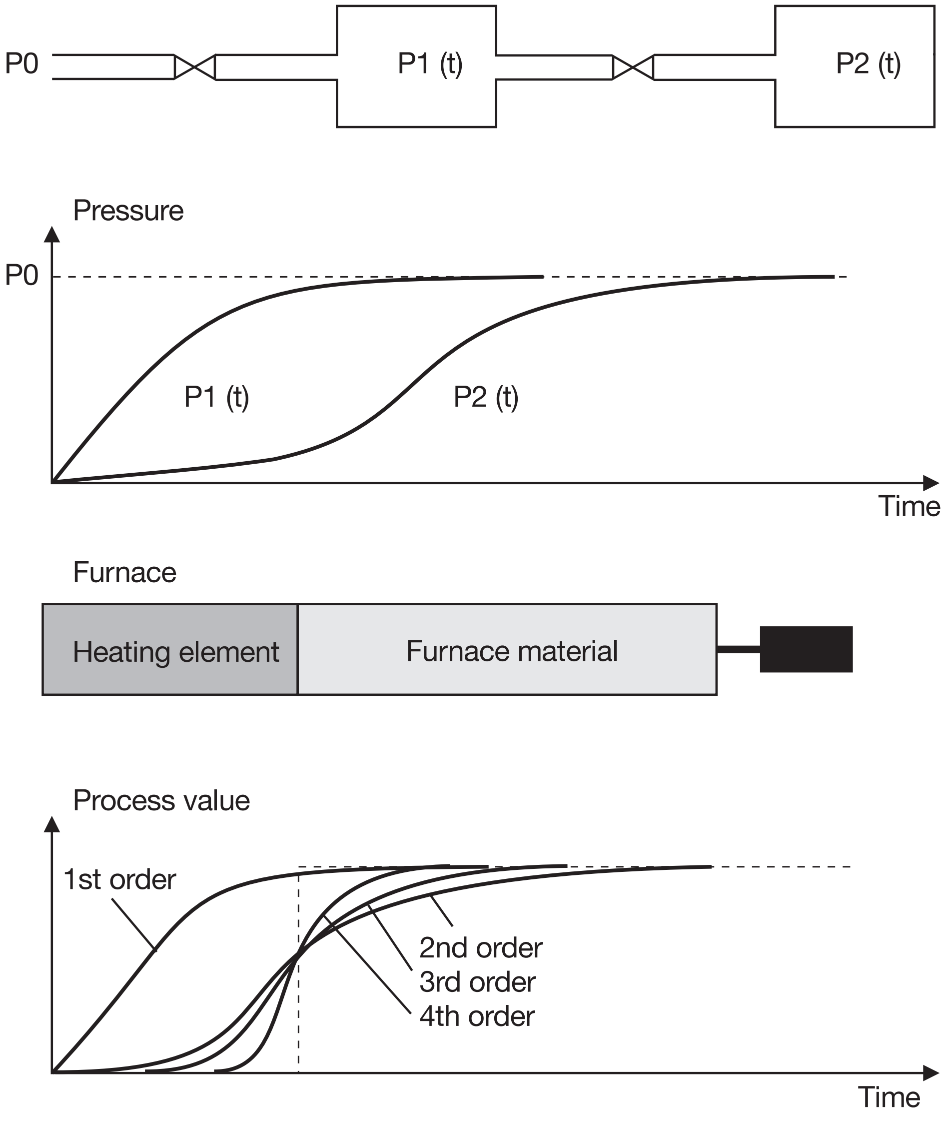


Fig. 23: Processes with delay

As an analogy, consider two pressure vessels which are connected by a throttle valve. In this case, the air must flow into the first vessel initially, and build up a pressure there, before it can flow into the second vessel. Eventually, the pressure in the first vessel reaches the supply pressure, and no more air can flow into it. As the pressures in the two vessels slowly come into line with each other, the pressure equalization rate between the two vessels becomes slower and slower, i.e. the pressure in the second vessel rises more and more slowly. Following a step change in the manipulating variable (in this case the supply line pressure) the process value (here the pressure in the second vessel) will take the following course: a very slow rise to begin with until a certain pressure has built up in the first vessel, followed by a steady rise and then finally an asymptotic or gradual approach to the final value.

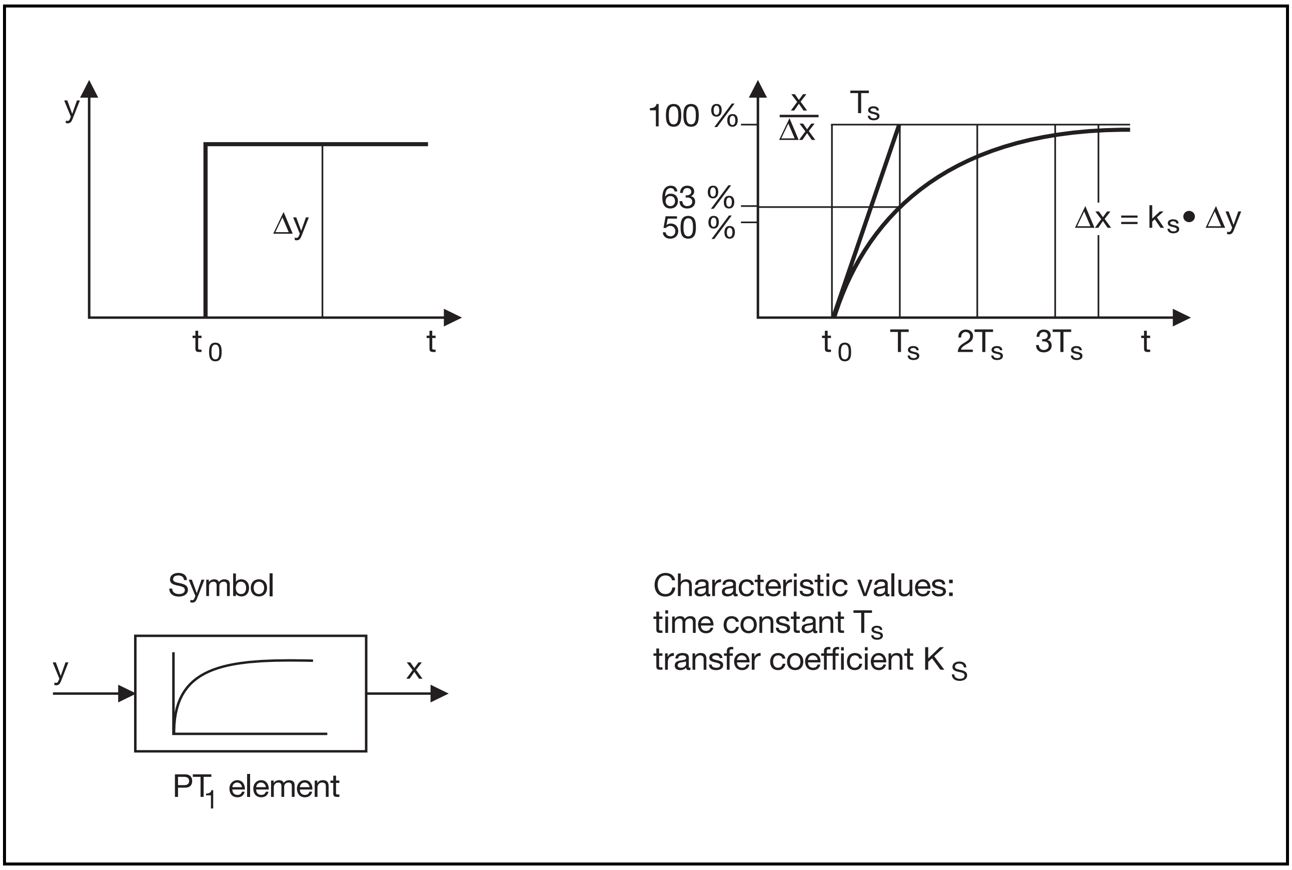
The transfer function of this type of system is determined by the number of energy stores available which are separated from each other by resistances. This concept can also be used when referring to the number of delays or time elements present in a process.

Such processes can be represented mathematically by an equation (exponential function) which has an exponential term for each energy store. Because of this relationship, these processes are designated as first-order, second-order, third-order processes, and so on.

The systems may be processes with or without self-limitation, which can also be affected by dead time.

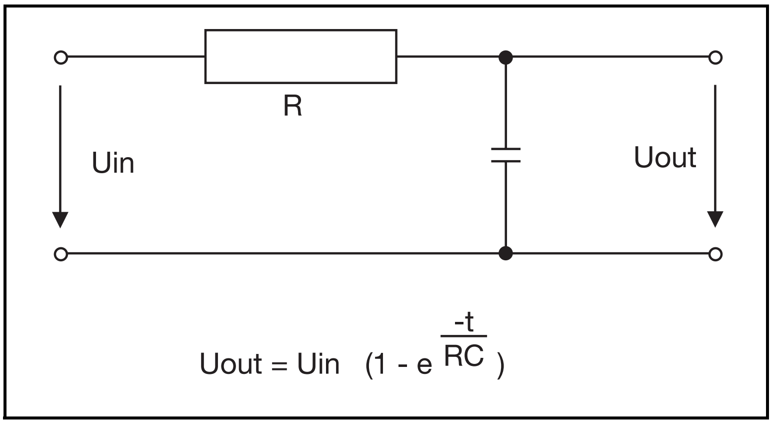
2.5.1 Processes with one delay (first-order processes)

In a process with one delay, i.e. with one available energy store, a step change in MV causes the PV to change immediately without delay and at a certain initial rate of change: PV then approaches the final value more and more slowly (see Fig. 24).



Characteristic values:
time constant T_s
transfer coefficient K_S

Fig. 24: First-order process; PT_1 process



$$\mathrm{Uout} = \mathrm{Uin}\ \left(1 - e^{\frac{-t}{RC}}\right)$$

Example:
A typical example of a first-order process is the charge or discharge of a capacitor through a resistor. The plot of the process variable (capacitor voltage) follows a typical exponential function.

For a step change Δy the relationship is as follows:

$$\Delta x = K_S \bullet \Delta y \bullet \left(1 - e^{\frac{-t}{T}}\right)$$

The term in brackets shows that a step change in MV does not produce a corresponding immediate change in PV. Instead PV slowly approaches the final value in a characteristic manner. As the time t increases (large value of t/T), the value of the expression in the brackets tends towards 1, so that for the final value, $\Delta x = K_S \bullet \Delta y$.

As shown in Fig. 24, after a time t = T (time constant), the PV has reached 63 % of the final value. After a time t = 5 T, the PV has almost reached 100 % of the final value.

Such processes are also referred to as T_1 processes. If it is a process with self-limitation, it is referred to as a PT_1 process; a process without self-limitation is an IT_1 process. Processes with one delay (first-order) occur very frequently. Examples are:

- heating and cooling of a hot water tank
- filling a container with air or gas via a throttle valve or a small bore pipe

2.5.2 Processes with two delays (second-order processes)

In a process with two delays there must be two storage elements, connected together by a resistance. Such processes can be characterized by specifying the transfer coefficient K_S and the time constants T_1 and T_2. Here, in contrast to a first order process, the step response of the process value starts with a horizontal phase and also has a point of inflection (see Fig. 25).

The course of the step response cannot be drawn by simply combining T_1 and T_2. For a step change Δy and for T_1 not equal to T_2 , the relationship is as follows:

$$\Delta x = K_S \bullet \Delta y \bullet \left(1 - \frac{T_1}{T_1 - T_2} e^{\frac{-t}{T_1}} + \frac{T_2}{T_1 - T_2} e^{\frac{-t}{T_2}}\right)$$

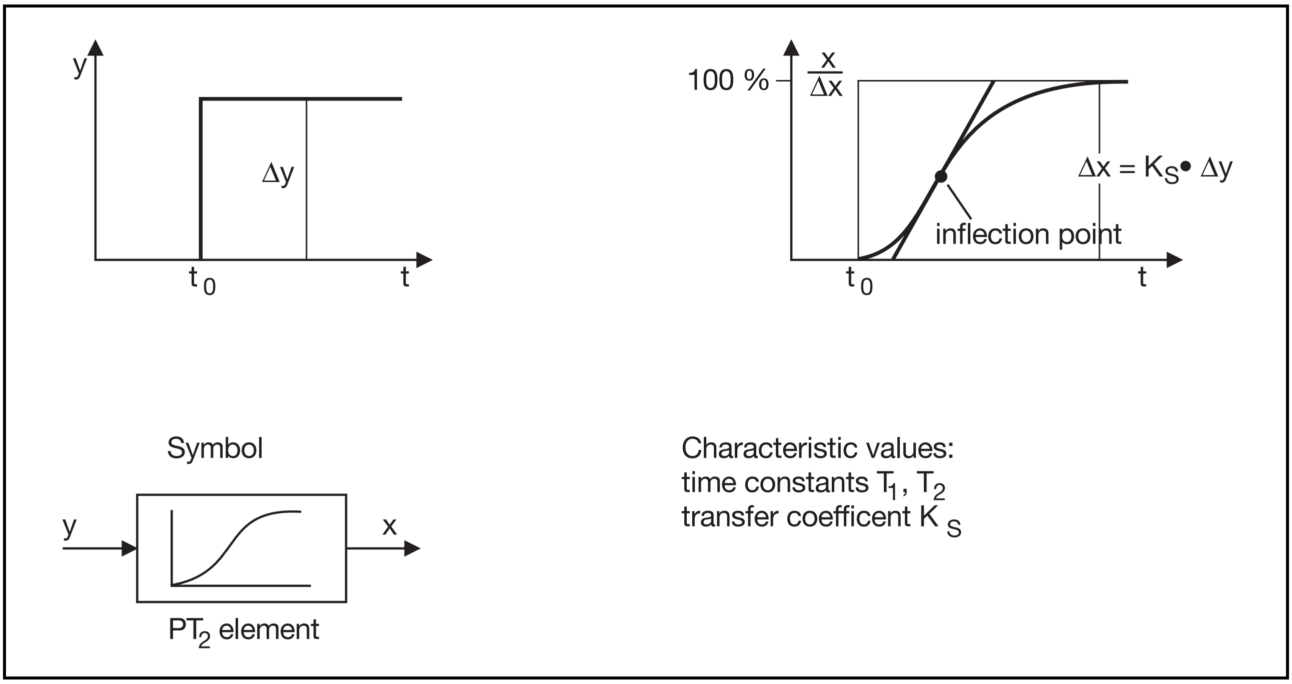


Fig. 25: Second-order process; PT$_2$ process

Such a process is normally called a PT$_2$ process. As already discussed, second-order processes always have a point of inflection, where the radius of curvature changes from a left-hand to a right-hand curve. First-order processes do not have this point of inflection.

Typical examples of this type of action are:

- filling two containers in series with air or gas through restrictors (see Fig. 23)
- charging up two RC networks in series
- temperature rise in a heated hot-water tank, where the thermometer is mounted in a pocket.

2.5.3 Processes with several delays (higher-order processes)

If there are more than two storage elements, the process has a correspondingly higher order. Interestingly enough, the transfer function characteristic of a higher-order process shows very little change from that of a second-order process. The rise of the curve does, however, become increasingly steeper and more delayed, until, with an infinite number of time delay elements, it approaches a pure dead time (see Fig. 26).

The order of a process is an important characteristic, particularly when describing it mathematically. In practice, almost every process is made up of a large number of widely differing energy storage elements, such as protective fittings, filling materials for temperature probes, dead spaces in manometers, etc. As a result, it is quite impossible to give an accurate mathematical description of an actual process.

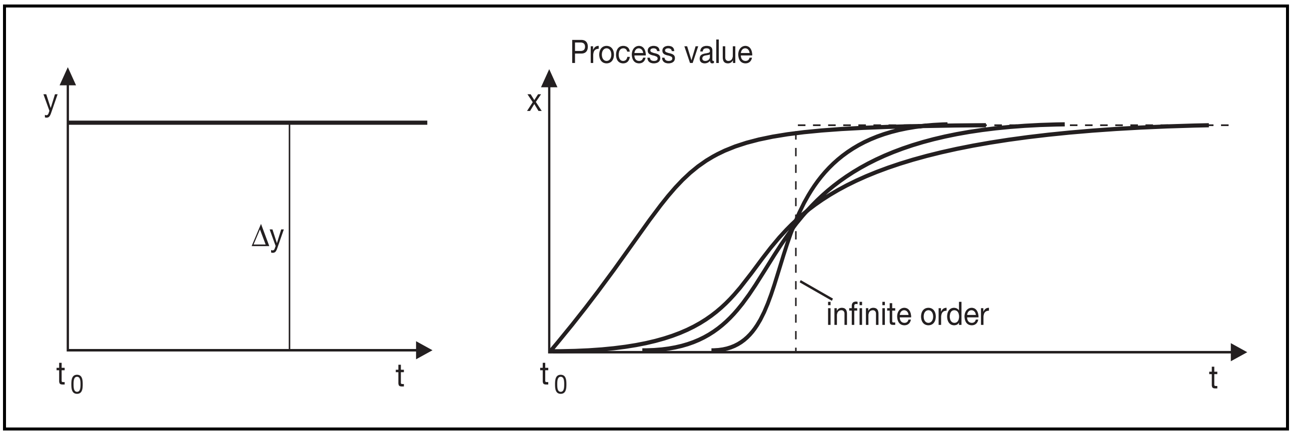


Fig. 26: Processes with several delays

In practice, the exact order of the process is not as important as might appear at first glance. Of much greater significance are the longest delay times, which determine the nature of the process.

As the order of the process increases, it becomes more and more difficult to control, since it approximates more and more to a system with dead time. A combination with a pure dead time is also possible, when the controllability deteriorates even more. Controllability is improved when there are significant differences between the time constants of the individual process elements. The worst case occurs when the time constants have the same value.

2.6 Recording the step response

The step response of a process, i.e. the course of the process value PV following a step change in manipulating variable MV can be characterized by two time values:

- the delay time $T_{u,}$ and the
- response time T_g

If these times are known, a quick estimate of the controllability of a process can be made, and the control parameters determined in a simple way, as explained later. The order of the process is ignored when using this approach, where it is assumed that any process is made up of a dead time T_u and a first-order process with a time constant T_g.

To determine such a transfer function and the resulting delay and response times, a recorder is connected to the transducer (sensor) and the manipulating variable (e.g. heating current) changed suddenly. Obviously, the change in MV should be limited to a value such that the new setpoint can be reached without damaging the system. The course of the process value is recorded, a tangent is drawn to the curve at the point of inflection, and T_u and T_g are determined as shown in Fig. 27.

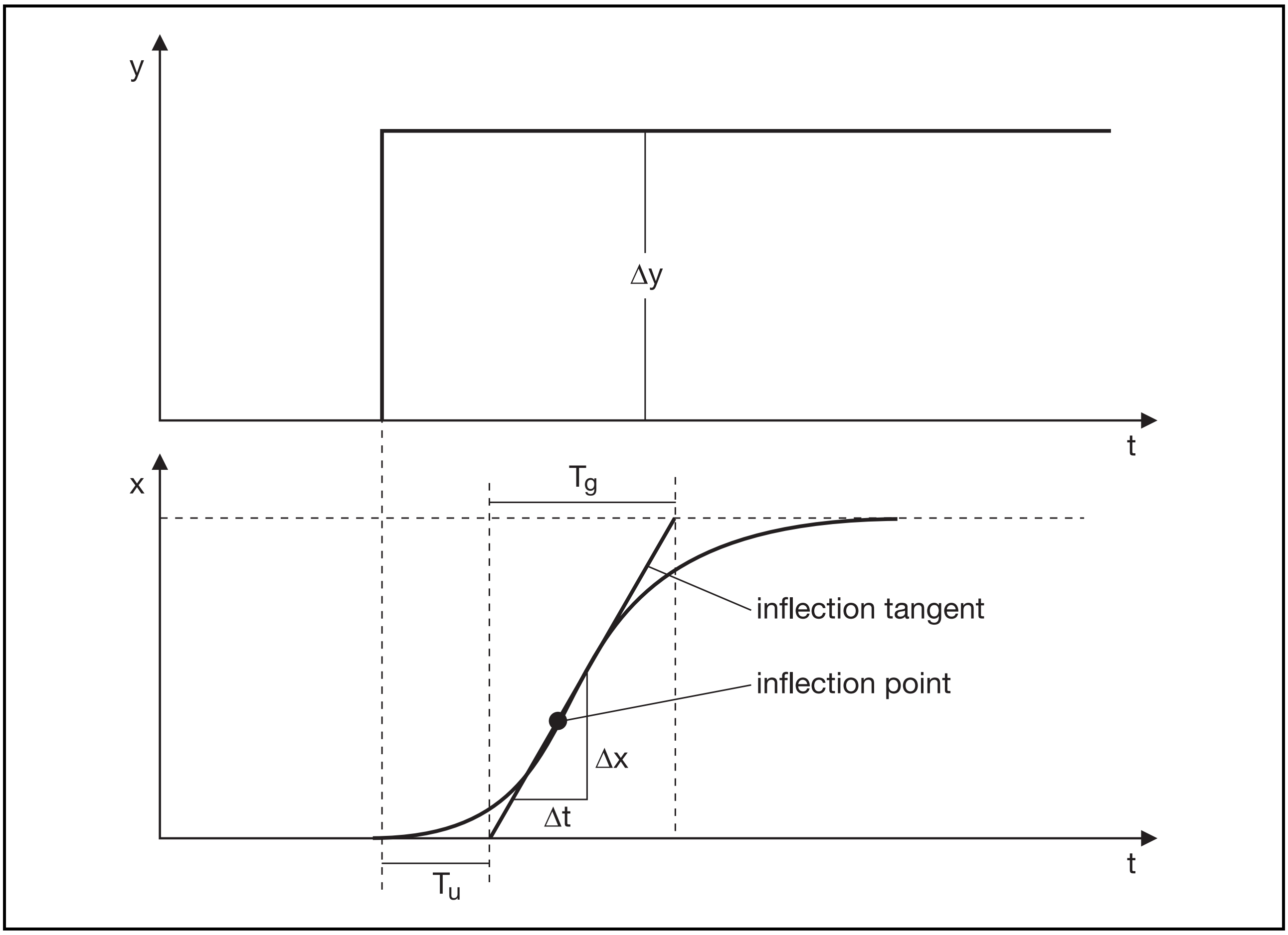


Fig. 27: Determining the delay time and response time

The ratio of delay time to response time gives information about the character of the process and its controllability:

T_g/T_u more than 10: process easy to control

T_g/T_u between 10 and 3: process can be controlled

T_g/T_u less than 3: process difficult to control

As the ratio of response time T_g to delay time T_u reduces, there is an increasing delay before the change in manipulating variable is communicated to the controller, and the controllability is progressively reduced. As explained in Chapter 2.5.3, a low T_g/T_u ratio corresponds to a steep gradient on the graph, representing a higher-order process which is difficult to control because of its tendency to overshoot.

Fast processes with T_g/T_u less than 3 are comparatively rare in furnaces, for example, since the temperature disturbance propagates relatively slowly and continuously through the furnace material. One exception is the type of furnace where the heating acts directly on the charge. The situation is quite different with pressure control: opening an air-lock can lead to a sudden drop in pressure, to which the controller must respond with an equally fast increase in supply pressure. Pressure equalization in the system takes place just as quickly, so that the entire control process is completed within a short space of time. In these processes the dead times are long in relation to the delay times. Certain chemical processes (reactions, neutralization) can sometimes proceed very quickly.

As well as the delay and response times, another important characteristic of the process can be determined, the maximum rate of rise V_{max}. It is obtained from the slope of the tangent at the point of inflection (see Fig. 27):

$$V_{max} = \frac{\text{change in process variable}}{\text{unit time}} = \frac{\Delta x}{\Delta t}$$

As shown later in the section on optimization, the maximum rate of rise is also referred to when setting control parameters.

DIN Specification 19 226 refers to the start-up value A instead of the rate of rise. This start-up value is the reciprocal of the maximum rate of rise of the process value PV for a sudden change in the manipulating variable from 0 — 100%:

$$A = \frac{1}{V_{max}}$$

2.7 Characteristic values of processes

The delay times and response times (standard values) of some typical processes are shown in Table 2 below.

Process variable	**Type of process**	**Delay time T_u**	**Response time T_g**
Temperature	small electrically-heated oven large electrically-heated furnace large gas-fired reheating furnace autoclave high-pressure autoclave injection molding machine extruder packaging machine	0.5 — 1 min 1 — 5 min 0.2 — 5 min 0.5 — 0.7 min 12 —15 min 0.5 — 3 min 1 — 6 min 0.5 — 4 min	5 — 15min 10 — 20min 3 — 60min 10 — 20min 200 — 300min 3 — 300min 5 — 60min 3 — 40min
Pressure	drum boiler, gas or oil-fired drum boiler, solid fuel-fired	0sec 0 — 2min	150sec 2.5 — 5min
Flow	gas pipelines liquid pipelines	0 — 5sec 0sec	0.2 — 10sec 0sec

Table 2: Delay times and response times (standard values) for some processes

The values given in the table should be taken as average values and serve only as a rough guide. For practical applications, the values of delay time and response time should be determined by carrying out a step response test.

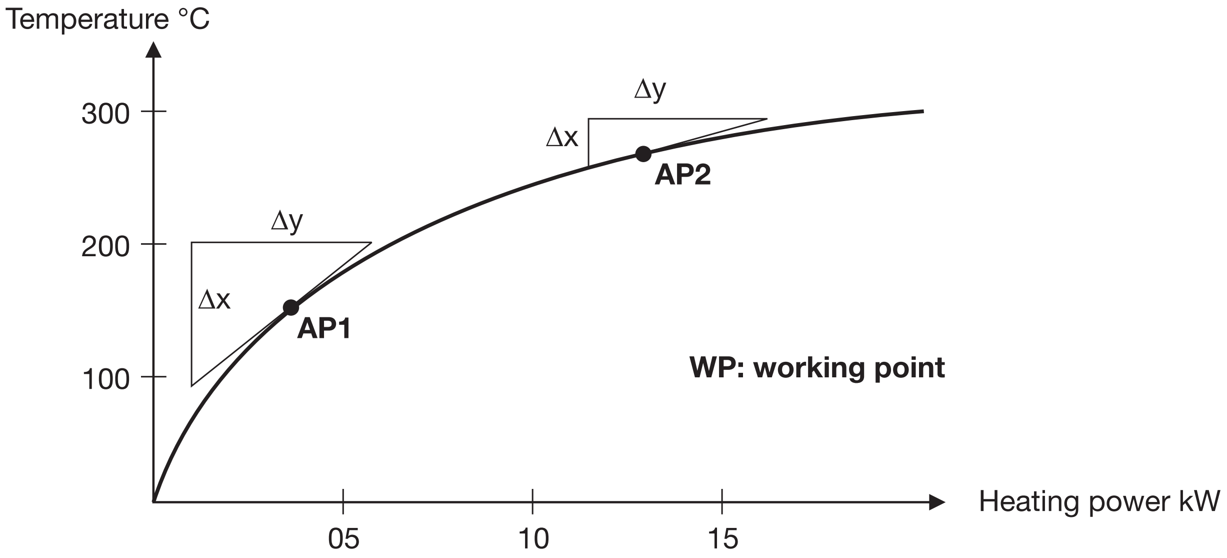
2.8 Transfer coefficient and working point

Previous sections have dealt mainly with the dynamic characteristic of the process (course of the step response), i.e. its behavior with respect to time. Chapter 2.1 has already mentioned the static characteristic, and described the final values for various manipulating variables. No account is taken of changes in the process value with respect to time.

The transfer coefficient is given by the ratio of output to input value, in this case the ratio of the change in process variable to the change in manipulating variable.

$$K_S = \frac{\text{change in process variable}}{\text{change in manipulating variable}} = \frac{\Delta x}{\Delta y}$$

The transfer coefficient K_S is also called the process gain.
In many cases the process gain K_S is not constant over the entire range of the process variable, as the following case will explain: in a furnace, a small increase in heating power is sufficient to produce a large increase in temperature at the lower end of the temperature range; at the upper end of the temperature range, however, a much larger change in energy flow is required to achieve the same effect (see Fig. 28).



⇨ K_S depends on the working point

$$\text{WP1: } K_S = \frac{\Delta x}{\Delta y} = \frac{100\ °\text{C}}{5\ \text{kW}} = 20\ \frac{°\text{C}}{\text{kW}}$$

$$\text{WP2: } K_S = \frac{\Delta x}{\Delta y} = \frac{30\ °\text{C}}{5\ \text{kW}} = 6\ \frac{°\text{C}}{\text{kW}}$$

Fig. 28: Process gain and working point

The reason for this is that the furnace used in the example represents a non-linear process. In addition to temperature processes, such processes also include processes where the friction is proportional to speed, relationships between motor power and speed, etc.

The delay time and response times in non-linear processes also depend on the working point. Process gain K_S, delay time, response time and other such values must be referred to a working point, i.e. to a pair of values of MV and PV, at which they have been evaluated. In general, they are not valid for other working points; in non-linear processes, a new set of values must be determined for each different working point, since the controller setting depends on the process parameters. In such processes the action of the controller is only optimized at the working point of the process for which the values were evaluated. If this is changed, for example, if a different process temperature is required, the controller has to be re-tuned to achieve optimum control.

Generally, the working point should lie in the range from middle to upper third of the transfer function at full power. A working point in the lower third is less satisfactory, because of the large excess power. Although the desired value (setpoint) is reached more rapidly in this case, the controllability is made worse. A working point in the upper part of the characteristic is also unsatisfactory, due to the lack of reserve power and resultant slow stabilization, and is also unsatisfactory from the point of view of disturbances.

3.1 Introduction

After discussing processes in Chapter 2, we now turn to the second important element of the control loop, the controller. The controller has already been described as the element which makes the comparison between process variable PV and setpoint SP, and which, depending on the control deviation, produces the manipulating variable MV. The output of a continuous controller carries a continuous or analog signal, either a voltage or a current, which can take up all intermediate values between a start value and an end value.

The other form of controller is the discontinuous or quasi-continuous controller in which the manipulating variable can only be switched on or off.

Continuous controllers offer advantages for certain control systems since their action on the process can be continuously modified to meet demands imposed by process events. Common industry standard output signals for continuous controllers are: 0 — 10V, 0 — 20mA, 4 — 20mA. On a continuous controller with a 0 — 20 mA output, 10% manipulating variable corresponds to an output of 2mA, 80% corresponds to 16mA, and 100% equals 20mA.

As discussed in Chapter 1, continuous controllers are used to operate actuators, such as thyristor units, regulating valves etc. which need a continuous signal.

3.2 P controller

In a P controller the control deviation is produced by forming the difference between the process variable PV and the selected setpoint SP; this is then amplified to give the manipulating variable MV, which operates a suitable actuator (see Fig. 29).

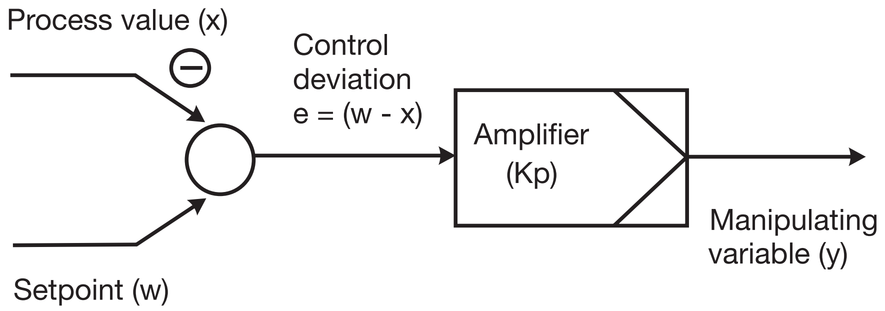


Fig. 29: Operating principle of a P controller

The control deviation signal has to be amplified, since it is too small and cannot be used directly as the manipulating variable. The gain (Kp) of a P controller must be adjustable, so that the controller can be matched to the process.

The continuous output signal is directly proportional to the control deviation, and follows the same course; it is merely amplified by a certain factor. A step change in the deviation e, caused for example by a sudden change in setpoint, results in a step change in manipulating variable (see Fig. 30).

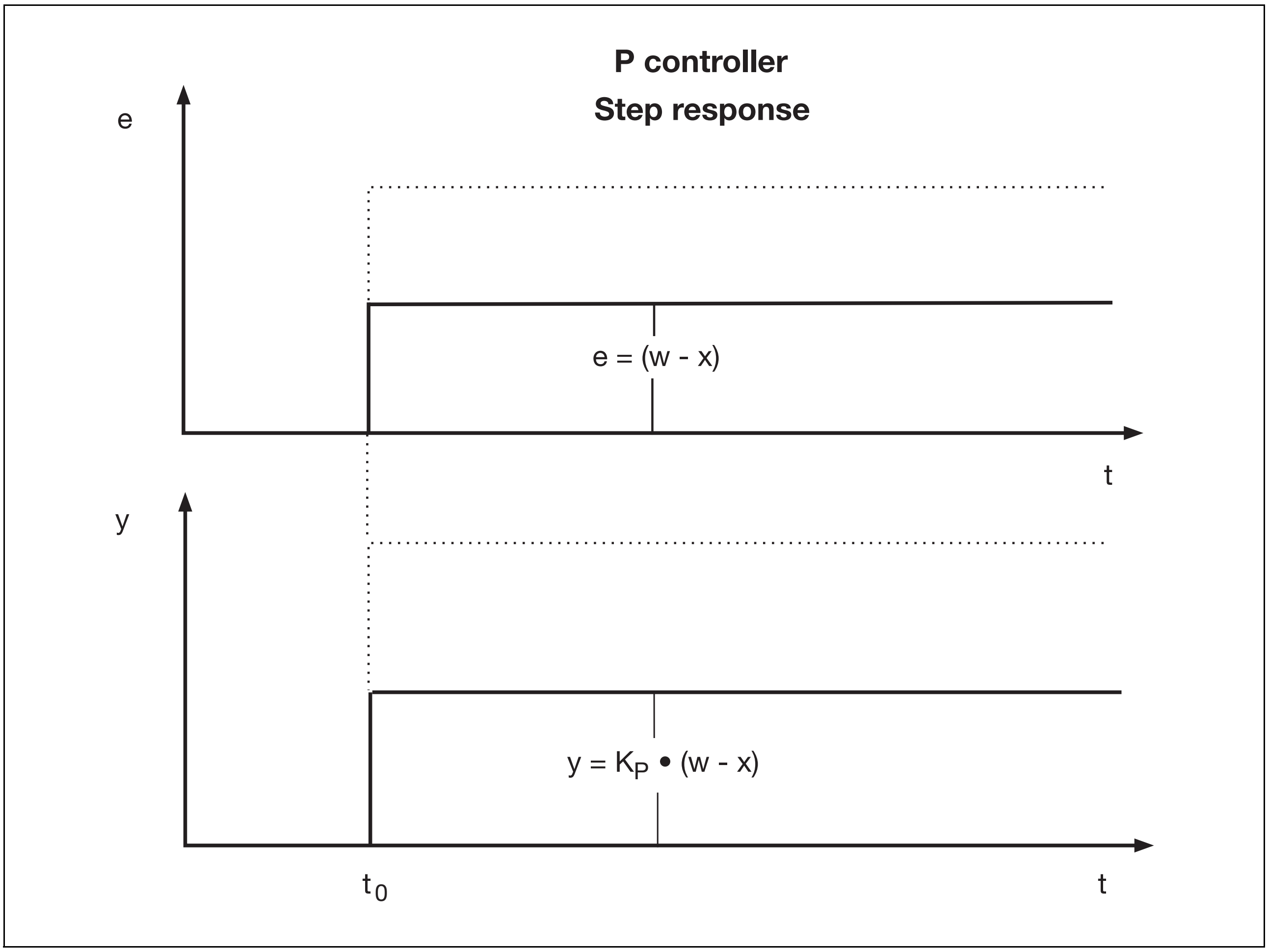


Fig. 30: Step response of a P controller

The step response of a P controller is shown in Fig. 30.

In other words, in a P controller the manipulating variable changes to the same extent as the deviation, though amplified by a factor. A P controller can be represented mathematically by the following controller equation:

$$y = K_P \bullet (w - x)$$

The factor K_P is called the proportionality factor or transfer coefficient of the P controller and corresponds to the control amplification or gain. It should not be confused with the process gain K_S of the process.

So, in an application where the user has set a K_P of 10 %/°C, a P controller will produce a manipulating variable of 50 % in response to a control difference of 5 °C.

Another example would be a P controller for the regulation of a pressure, with a K_P set to 4 %/bar. In this case, a control difference of 20 bar will produce a manipulating variable of 80 %.

3.2.1 The proportional band

Looking at the controller equation, it follows that, in a P controller, any value of deviation would produce a corresponding value of manipulating variable. However, this is not possible in practice, as the manipulating variable is limited for technical reasons, so that the proportional relationship between manipulating variable and control deviation only exists over a certain range of values.

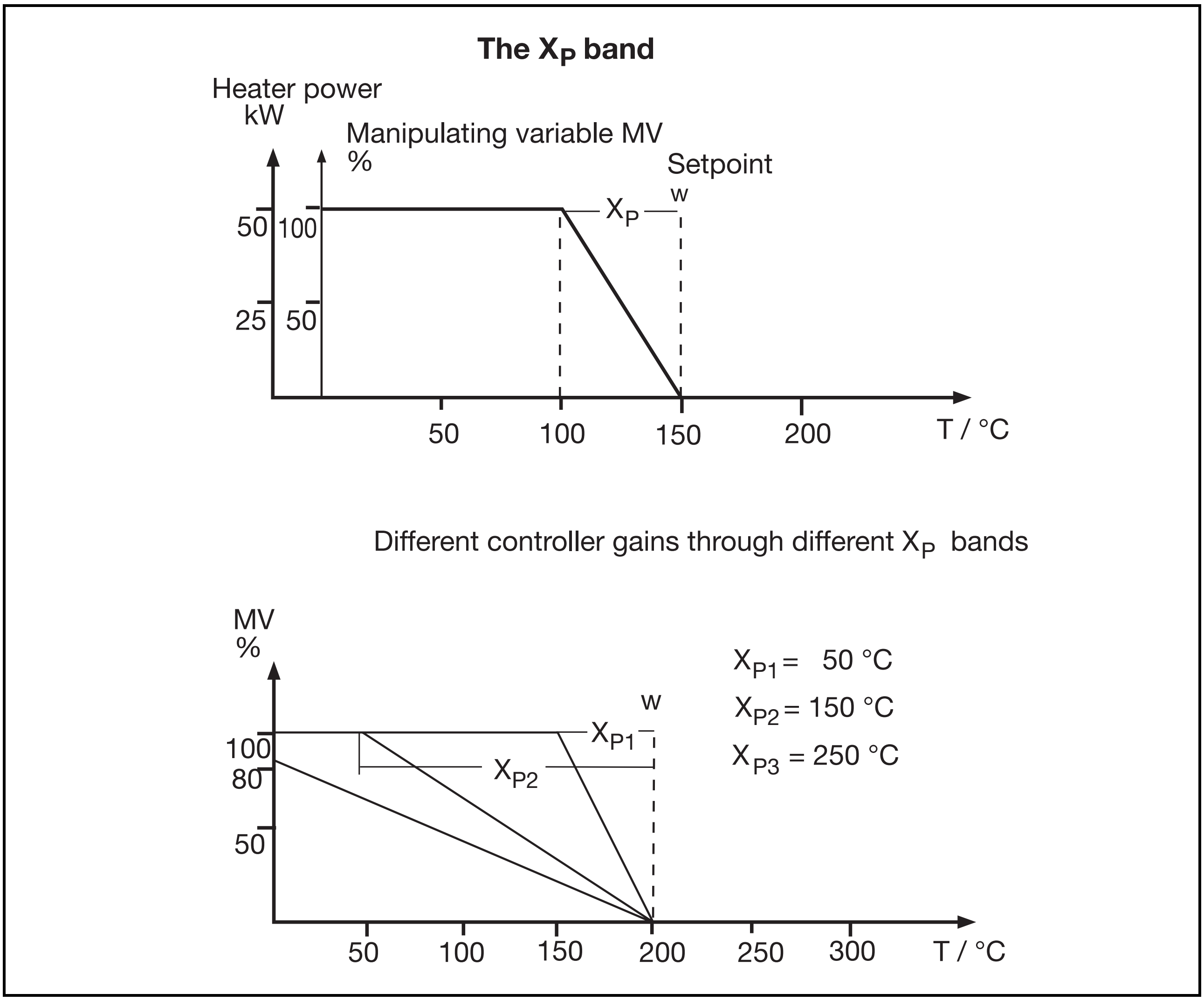


Fig. 31: The position of the proportional band

The top half of Fig. 31 shows the characteristic of a P controller, which is controlling an electrically heated furnace, with a selected setpoint w = 150°C.

The following relationship could conceivably apply to this furnace
The manipulating variable is only proportional to the deviation over the range from 100 to 150°C, i.e. for a deviation of 50°C from the intended setpoint of 150°C. Accordingly, the manipulating variable reaches its maximum and minimum values at these values of deviation, and the highest and lowest heater power is applied respectively. No further changes are possible, even if the deviation increases.

This range is called the proportional band X_P. Only within this band is the manipulating variable proportional to the deviation. The gain of the controller can be matched to the process by altering the X_P band. If a narrower X_P band is chosen, a small deviation is sufficient to travel through the full manipulating range, i.e. the gain increases as X_P is reduced.

The relationship between the proportional band and the gain or proportionality factor of the controller is given by the following formula:

$$X_P = \frac{1}{K_P} \bullet 100\%$$

Within the proportional band X_P, the controller travels through the full manipulating range y_H, so that K_P can be determined as follows:

$$K_P = \frac{y_H}{X_P} = \frac{\text{max. manipulating range}}{\text{proportional band}}$$

The unit of the proportionality factor K_P is the unit of the manipulating variable divided by the unit of the process variable. In practice, the proportional band X_P is often more useful than the proportionality factor K_P and it is X_P rather than K_P that is most often set on the controller. It is specified in the same unit as the process variable (°C, V, bar etc.). In the above example of furnace control, the X_P band has a value 50 °C. The advantage of using X_P is that the value of deviation, which produces 100 % manipulating variable, is immediately evident. In temperature controllers, it is of particular interest to know the operating temperature corresponding to 100 % manipulating variable. Fig. 31 shows an example of different X_P bands.

An example
An electric furnace is to be controlled by a digital controller. The manipulating variable is to be 100 % for a deviation of 10 °C. A proportional band $X_P = 10$ is therefore set on the controller.

Until now, for reasons of clarity, we have only considered the falling characteristic (inverse operating sense), in other words, as the process variable increases, the manipulating variable decreases, until the setpoint is reached. In addition, the position of the X_P band has been shown to one side of and below the setpoint.

However, the X_P band may be symmetrical about the setpoint or above it (see Fig. 32). In addition, controllers with a rising characteristic (direct operating sense) are used for certain processes. For instance, the manipulating variable in a cooling process must decrease as the process value increases.

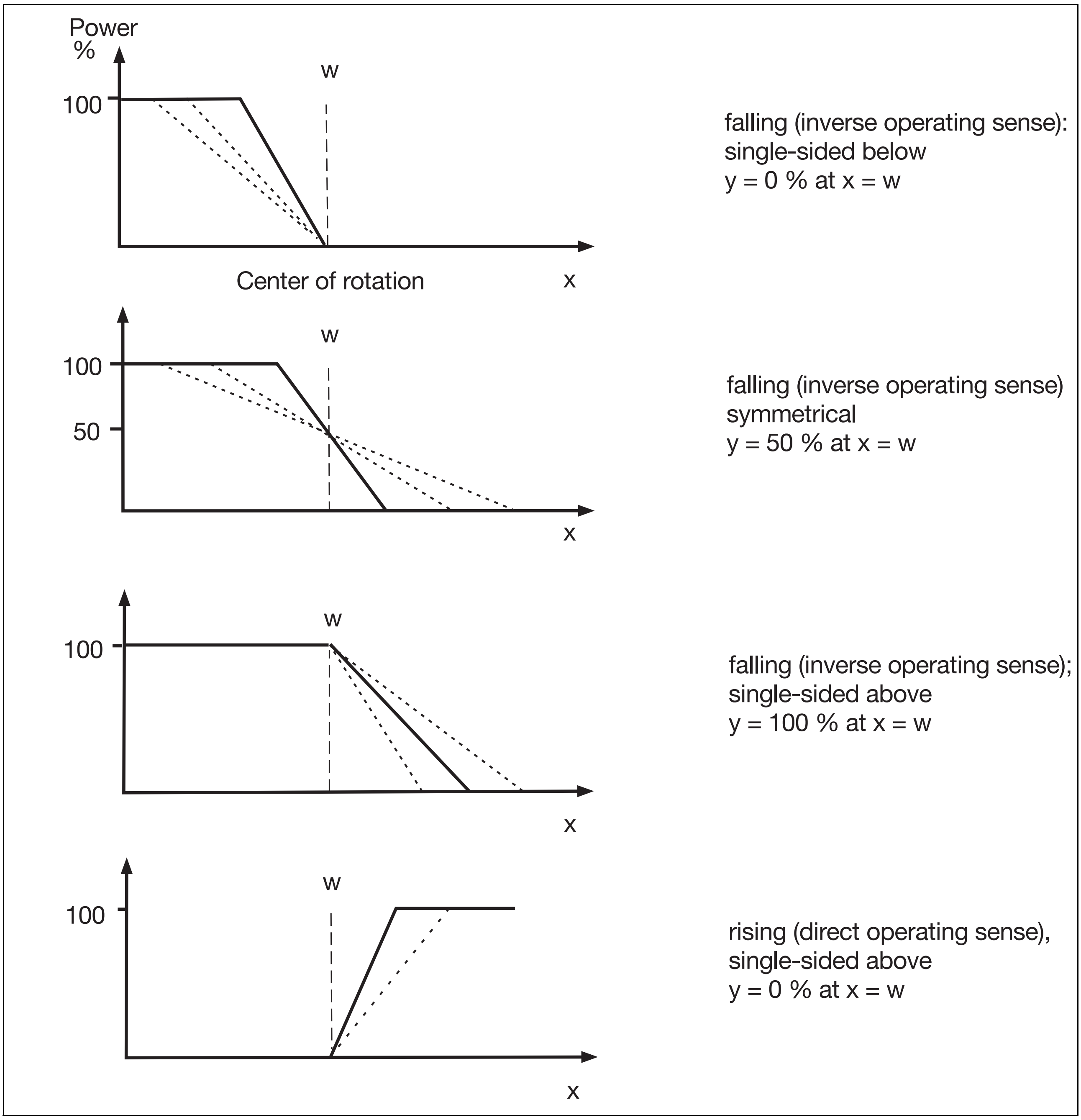


Fig. 32: Position of the proportional band about the setpoint

The advantages of X_P bands which are symmetrical or asymmetrical about the setpoint will be discussed in more detail under 3.2.2.

3.2.2 Permanent deviation and working point

A P controller only produces a manipulating variable when there is a control deviation, as we already know from the controller equation. This means that the manipulating variable becomes zero when the process variable reaches the setpoint. This can be very useful in certain processes, such as level control. However, in our example of the furnace, it means that heating power is no longer applied when the control deviation is zero. As a consequence, the temperature in the furnace falls. Now there is a deviation, which the controller then amplifies to produce the manipulating variable; the larger the deviation, the larger the manipulating variable of the controller. The deviation now takes up a value such that the resulting manipulating variable is just sufficient to maintain the process variable at a constant value.

A P controller always has a permanent control deviation or offset

This permanent deviation can be made smaller by reducing the proportional band X_P. At first glance, this might seem to be the optimal solution. However, in practice, all control loops become unstable if the value of X_P falls below a critical value - the process variable starts to oscillate.

If the static characteristic of the process is known, the resulting control deviation can be found directly. Fig. 33 shows the characteristic of a P controller with an X_P band of 100°C. A setpoint of 200°C is to be held by the controller. The process characteristic of the furnace shows that a manipulating variable of 50% is required for a setpoint of 200°C. However, the controller produces zero manipulating variable at 200°C. The temperature will fall, and, as the deviation increases, the controller will deliver a higher manipulating variable, corresponding to the X_P band. A temperature will be reached here, at which the controller produces the exact value of manipulating variable required to maintain that temperature. The temperature reached, and the corresponding manipulating variable, can be read off from the point of intersection of the controller characteristic and the static process characteristic: in this case, a temperature of 150°C with a manipulating variable of 40%.

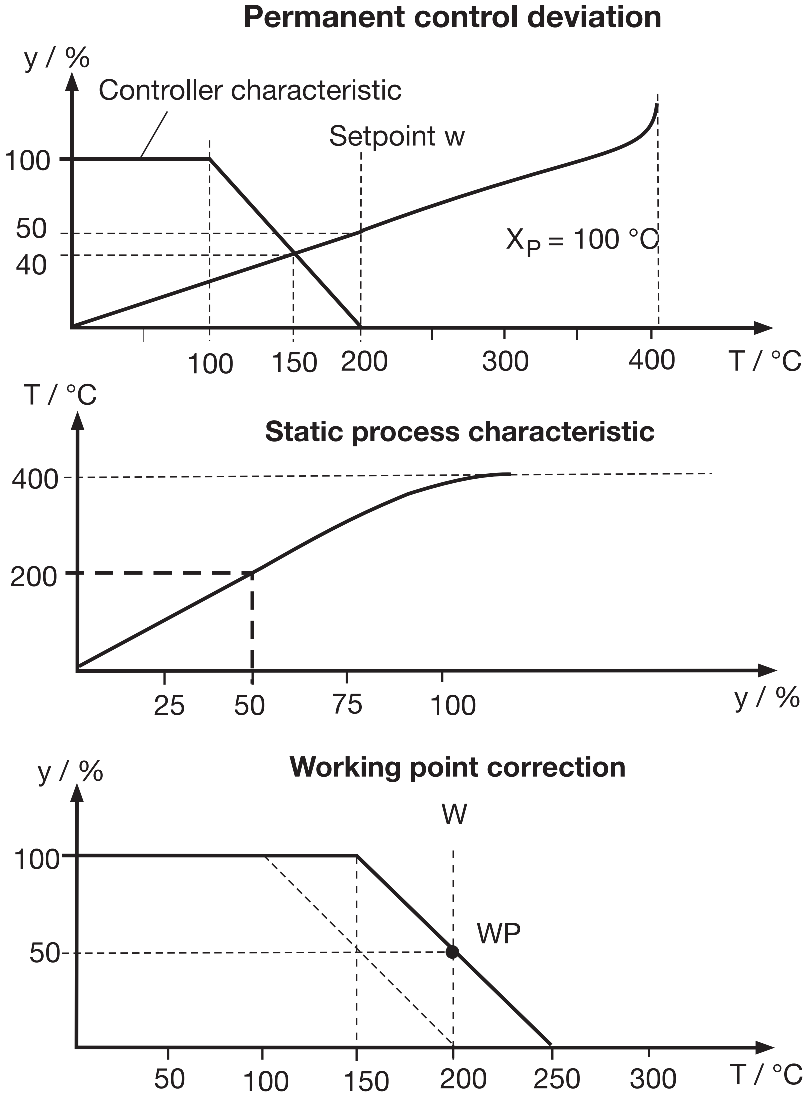


Fig. 33: Permanent deviation and working point correction

It is clear that in a furnace, for instance, a certain level of power must be supplied in order to reach and maintain a particular setpoint. So it makes no sense to set the manipulating variable to zero when there is no control deviation. The manipulating variable is usually set to a specific percentage value for a control difference of 0. This is called working point correction, and can be adjusted on the controller, normally over the range of 0 — 100 %. This means that with a correction of 50 %, the controller would produce a manipulating variable of 50 % for zero control deviation. In the example given, see Fig. 33, this would lead to the setpoint w = 200 °C being reached and held. We can see that the proportional band exhibits a falling characteristic that is symmetrical about the setpoint. If the process actually requires the manipulating variable set at the working point, as in our example, the control operates without deviation.

Setting the working point in practice

In practice, the process characteristic of a process is not usually known. However, the working point correction can be determined by manually controlling the process variable at the setpoint value that the controller is to hold later. The manipulating variable required for this is also the value for the working point correction.

Example

In a furnace where a setpoint of 200 °C is to be tracked, the controller would be set to manual mode and the manipulating variable slowly increased by hand, allowing adequate time after the change for the end temperature to be reached. A certain value of manipulating variable will be determined, for example 50 %, which is sufficient for a process variable of 200 °C. This manipulating variable is then fed in as the value for the working point correction.

After feeding in the value for the working point correction, the controller will only operate without control difference at the particular setpoint for which the working point correction was made. Furthermore, the external conditions must not change. If other disturbances did affect the process, (for example, a fall in the temperature outside a furnace), a control difference would be set once again, although this time the value would be smaller.

We can summarize the main points about the control deviation of a P controller as follows (controller with falling characteristic, process with self-limitation):

Without working point WP

- The process variable remains in a steady state below the setpoint.

With working point WP (see Fig. 33)

- below the working point (in this case 0 — 50 % manipulating variable)
  process variable is above the setpoint
- at the working point (in this case 50 % manipulating variable)
  process variable = setpoint
- above the working point (in this case 50 — 100 % manipulating variable)
  process variable is below the setpoint

In a P controller, the output signal has the same time course as the control deviation, and because of this it responds to disturbances very rapidly. It is not suitable for processes with a pure dead time, as these start to oscillate due to the P controller. On processes with self-limitation, it is not possible to control exactly at the setpoint; a permanent deviation is always present, which can be significantly reduced by introducing a working point correction.

3.2.3 Controllers with dynamic action

As we saw in the previous chapter, the P controller simply responds to the magnitude of the deviation and amplifies it. As far as the controller is concerned, it is unimportant whether the deviation occurs very quickly or is present over a long period.

When a large disturbance occurs, the initial response of a machine operator is to increase the manipulating variable, and then keep on changing it until the process variable reaches the setpoint. He would consider not only the magnitude of the deviation, but also its behavior with time (dynamic action).

Of course, there are control components that behave in the same way as the machine operator mentioned above:

- The D component responds to changes in the process variable. For example, if there is 20 % reduction in the supply voltage of an electric furnace, the furnace temperature will fall. This D component responds to the fall in temperature by producing a manipulating variable. In this case, the manipulating variable is proportional to the rate of change of furnace temperature, and helps to control the process variable at the setpoint.
- The I component responds to the duration of the deviation. It summates the deviation applied to its input over a period of time. If this controller is used on a furnace, for example, it will slowly increase the heating power until the furnace temperature reaches the required setpoint.

In the past, dynamic action was achieved in analog controllers by feeding part of the manipulating variable back to the controller input, via timing circuits. The feedback changes the input signal (the real control deviation) so that the controller receives a simulated deviation signal that is modified by a time-dependent factor. In this way, using a D component, a sudden change in process variable, for example, can be made to have exactly the same initial effect as a much larger control deviation. In this connection, because of this reverse coupling, we often talk about feedback. In modern microprocessor controllers, the manipulating variable is not produced via feedback, but derived mathematically direct from the setpoint and process variable.

We will avoid using the term feedback in this book, as far as possible.

The components described above are often combined with a P component to give PI, PD or PID controllers.

3.3 I controller

An I controller (integral controller) integrates the deviation signal applied to its input over a period of time. The longer there is a deviation on the controller, the larger the manipulating variable of the I controller becomes. How quickly the controller builds up its manipulating variable depends firstly on the setting of the I component, and secondly on the magnitude of the deviation.

The manipulating variable changes as long as there is a deviation. Thus, over a period of time, even small deviations can change the manipulating variable to such an extent that the process variable corresponds to the required setpoint.

In principle, an I controller can fully stabilize after a sufficiently long period of time, i.e. setpoint = process variable. The deviation is then zero and there is no further increase in manipulating variable.

Unlike the P controller, the I controller does not have a permanent control deviation

The step response of the I controller shows the course of the manipulating variable over time, following a step change in the control difference (see Fig. 34).

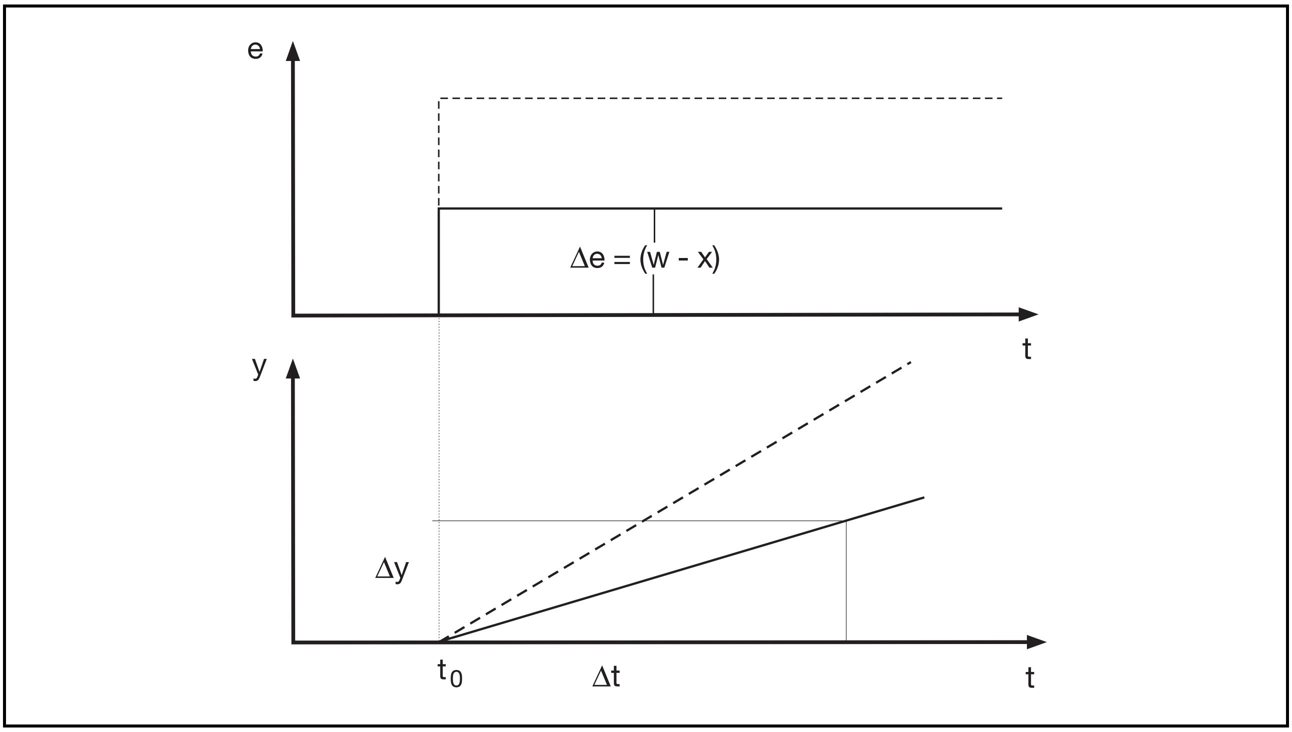


Fig. 34: Step response of an I controller

For a constant control deviation Δe, the equation of the I controller is as follows:

$$\Delta y = \frac{1}{T_I} \bullet \Delta e \bullet t$$

Here T_I is the integral time of the I controller and t the duration of the deviation. It is clear that the change in manipulating variable y is proportional not only to the change in process variable, but also to the time t.

If the control deviation is varying, then:

$$y = \frac{1}{T_I}\int e \bullet dt \bullet \frac{s}{K}$$

The integral time of the I controller can also be evaluated from the step response (see Fig. 34):

$$T_I = \frac{\Delta e \bullet \Delta t}{\Delta y}$$

If the process variable is below the setpoint on an I controller with a negative operating sense, as used, for example, in heating applications, the I controller continually builds up its manipulating variable. When the process variable reaches the setpoint, we now have the possibility that the manipulating variable is too large, because of delays in the process. The process variable will again increase slightly; however, the manipulating variable is now reduced, because of the sign reversal of the process variable (now above the setpoint).

It is precisely this relationship that leads to a certain disadvantage of the I controller
If the manipulating variable builds up too quickly, the control signal which arises is too large, and too high a process variable is reached. Now the process variable is above the setpoint and the sign of the deviation is reversed, i.e. the control signal decreases again. If the decrease is too sudden, a lower process value is arrived at, and so on. In other words, with an I controller, oscillations about the setpoint can occur quite frequently. This is especially the case if the I component is too strong, i.e. when the selected integral time T_I is too short. The exception to this is the zero-order process where, because there are no energy storage possibilities, the process variable follows the manipulating variable immediately, without any delay; the control loop forms a system which is not capable of oscillation.

To develop a feel for the effect of the integral time T_I, it can be defined as follows: The integral time T_I is the time that the integral controller needs to produce its constant control difference at its output (without considering sign). Imagine a P controller for a furnace, where the response time T_I is set at 60sec and the control difference is constant at 2°C. The controller requires a time T_I = 60sec for a 2% increase in manipulating variable, if the control difference remains unchanged at 2°C.

Summarizing the main points, the I controller removes the control deviation completely, in contrast to the P controller.

An I controller is not stable when operating on a process without self-limitation, and is therefore unsuitable for control of liquid levels, for example. On processes with long time constants, the I component must be set very low, so that the process variable does not tend to oscillate. With this small I component, the I controller works much too slowly. For this reason, it is not particularly suitable for processes with long time constants (e.g. temperature control systems). The I type of controller is frequently used for pressure regulation, and in such a case T_n is set to a very low value.

3.4 PI controller

As we have found in the I controller, it takes a relatively long time (depending on T_i) before the controller has built up its manipulating variable. Conversely, the P controller responds immediately to control differences by immediately changing its manipulating variable, but is unable to completely remove the control difference. This would seem to suggest combining a P controller with an I controller. The result is a PI controller. Such a combination can combine the advantage of the P controller, the rapid response to a control deviation, with the advantage of the I controller, the exact control at the setpoint.

We can obtain the step response of a PI controller simply by superimposing the step responses of a P and an I controller, as shown in Fig. 35.

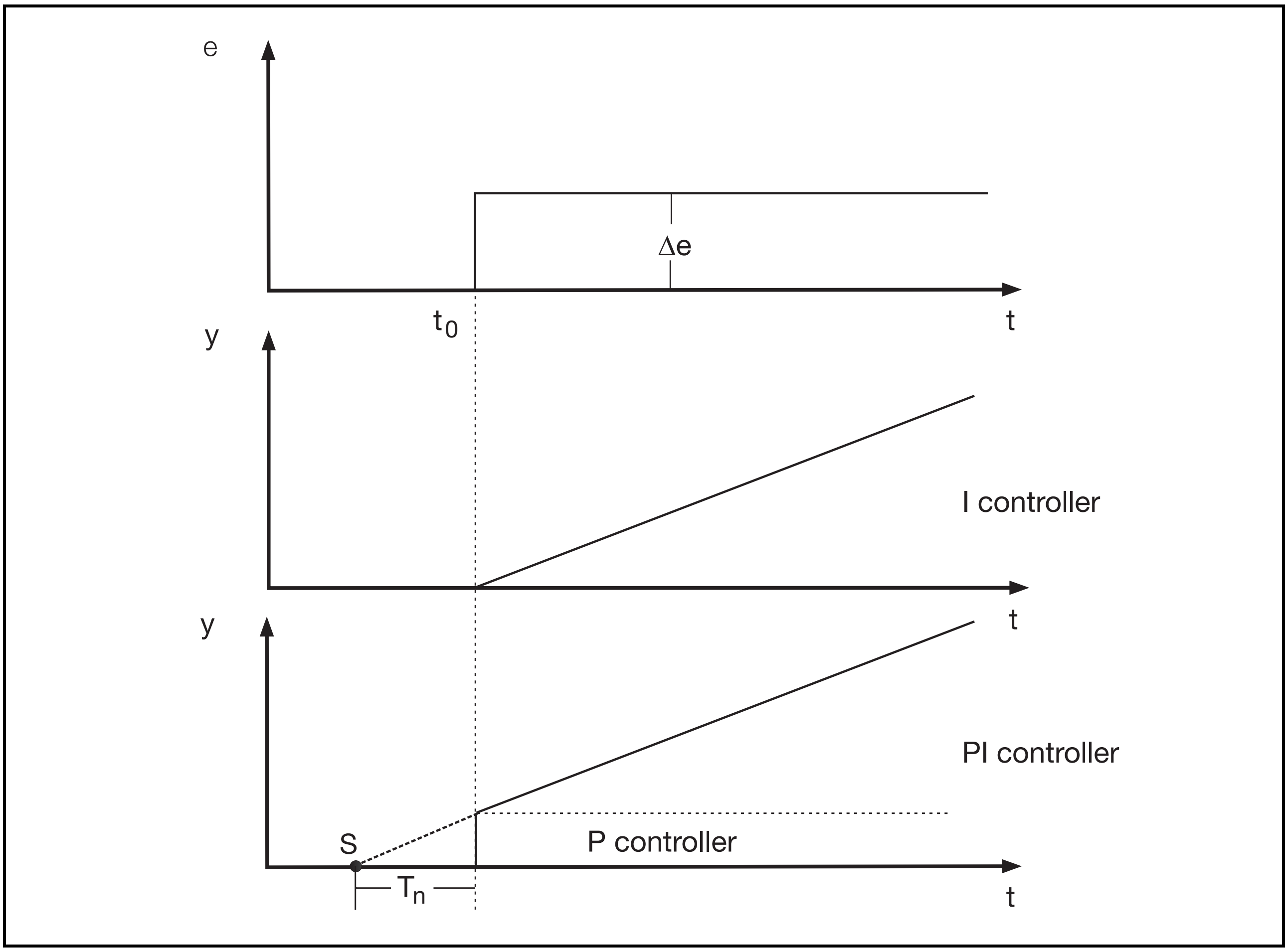


Fig. 35: Step response of a PI controller

If the diagonally rising straight line of the PI manipulating variable is projected back to its point of intersection S with the time axis, it intercepts a length of time there. With a PI controller, this corresponds to the reset time T_n .

For a control deviation e = Δe = constant, we obtain the following equation for the PI controller:

$$\Delta y = \frac{1}{X_P} \bullet 100\% \bullet \left(\Delta e + \frac{1}{T_n} \bullet \Delta e \bullet t\right) = \frac{1}{X_P} \bullet 100\% \bullet \Delta e \bullet \left(1 + \frac{1}{T_n} \bullet t\right)$$

The reset time is a measure of the extent to which the duration of the control deviation affects the control function. A long reset time means that the I component has little influence, and vice versa. From the equation above, it is evident that the real amplification of the I component is the factor

$$\frac{1}{X_P} \bullet 100\% \bullet \frac{1}{T_n}$$

With a PI controller, therefore, a change in proportional band X_P also causes a change in the integral action. If the proportional gain of a PI controller is increased by reducing X_P, the integral action will also be increased, so the controller will make a faster integration of the control difference.

It is also possible to interpret T_n as the time interval required for the I component to produce the same manipulating variable y (for a given deviation), as that already produced by the P component (following a step change). The formula given above is only valid when the deviation remains constant during the time interval t. If this is not the case, the relationship is as follows:

$$\Delta y = \frac{1}{X_P} \bullet 100\% \bullet e + \frac{1}{X_P} \bullet 100\% \bullet \frac{1}{T_n} \bullet \int e \bullet dt$$

As mentioned earlier, a PI controller can, in principle, be built up by combining a P controller and an I controller. With a sudden deviation, the manipulating variable is initially formed by the P component (see Fig. 36). Because of the changed manipulating variable, the process variable moves towards the setpoint, i.e. the deviation is reduced, and with it the manipulating variable produced by the P controller. Now the manipulating variable produced by the I component ensures exact control. Whereas the P component of the manipulating variable steadily decreases as the setpoint is approached, the I component continues to build up. Here, however, the increase is also smaller, because of the reducing deviation, until finally, when the setpoint is reached, nothing more is added to the current manipulating variable. When the system has stabilized, the manipulating variable of the PI controller is produced solely by the I component.

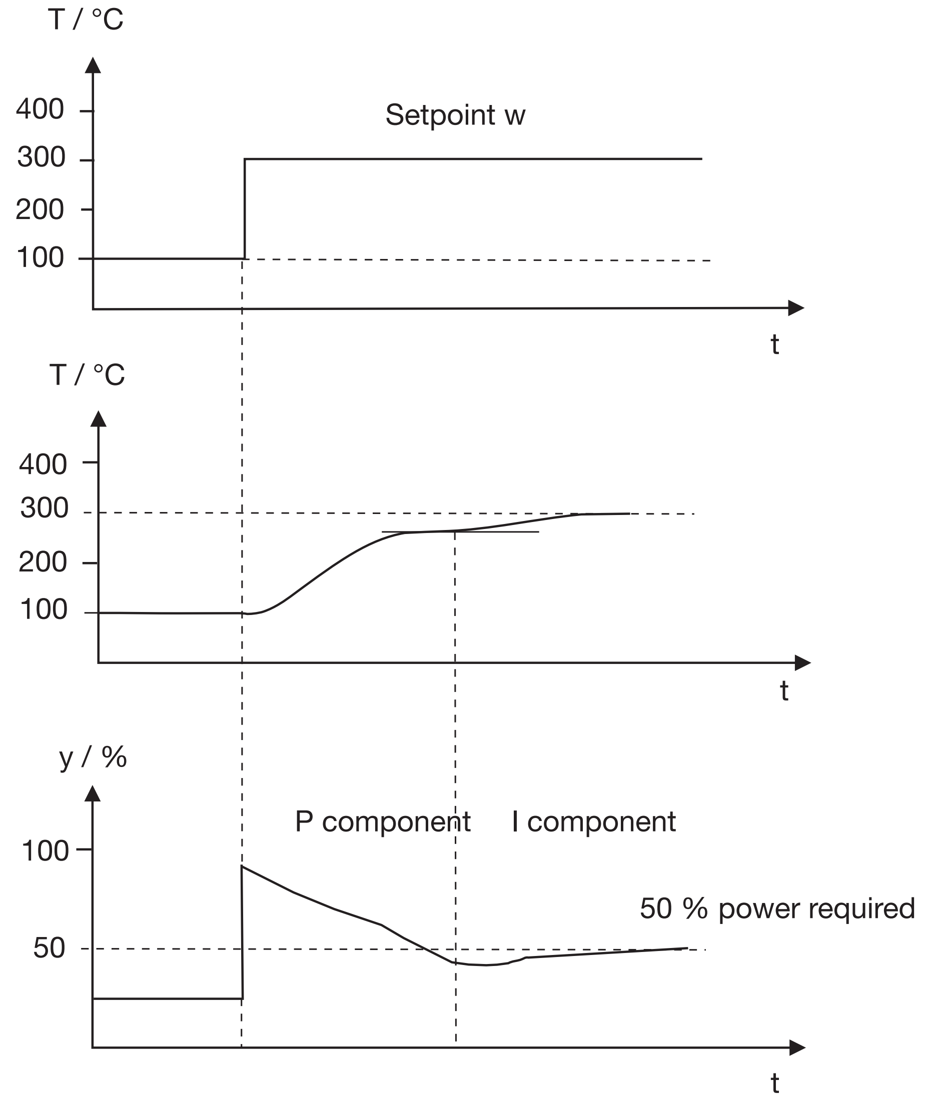


Fig. 36: Formation of the manipulating variable in a PI controller

Summarizing the main points:
In a PI controller, the P component causes the manipulating variable to respond immediately to the control deviation. The PI controller is therefore much faster than an I controller. The I component integrates the control deviation at the output of the controller, so that the closed control loop acts to reduce the remaining deviation.

3.5 PD controller

If a large disturbance occurs in a control loop which is being controlled manually, bringing with it a change in the manipulating variable, the operator will try to cushion the effect of the disturbance by making a large initial adjustment of the actuator. He then quickly reduces the adjustment, so that the new equilibrium of the control loop can be approached gradually. A controller which responds in a similar way to the above operator is the PD controller: it consists of a P component with a known proportional action, and a D component with a derivative action. This D component responds not to the magnitude or duration of the control deviation, but to the rate of change of the process variable. Fig. 37 shows how such a PD controller builds up its manipulating variable.

Fig. 37 explains how the PD controller works. If a new setpoint is applied, the manipulating variable is increased by the P component; this component of the manipulating variable is always proportional to the deviation. The process variable responds to the increased manipulating variable, for example, a furnace temperature rises. As soon as the process variable changes, the D component starts to take effect: while the process variable increases, the D component forms a negative manipulating variable, which is subtracted from the manipulating variable of the P component, finally producing the manipulating variable at the controller output. When the process variable is tracking the setpoint, the D component “brakes”, thus preventing the manipulating variable overshooting above the setpoint.

If the process variable has reached its maximum value after an overshoot above the setpoint, and is now reducing, the D component gives out a positive manipulating variable. In this case, the D component counteracts the change in process variable.

The D component only intervenes in the process when there is a change in process variable. The size of the manipulating variable of the D component depends on the rate of change of the process variable, that is on the magnitude of $\Delta x/\Delta t$ (see the gradient triangle in Fig. 37). In addition, the effect can be changed at the controller via the time T_d (derivative time), which we will get to know in this chapter. A pure D controller is not suitable for control, as it does not intervene in the process when there is a constant deviation, or when the process variable remains constant.

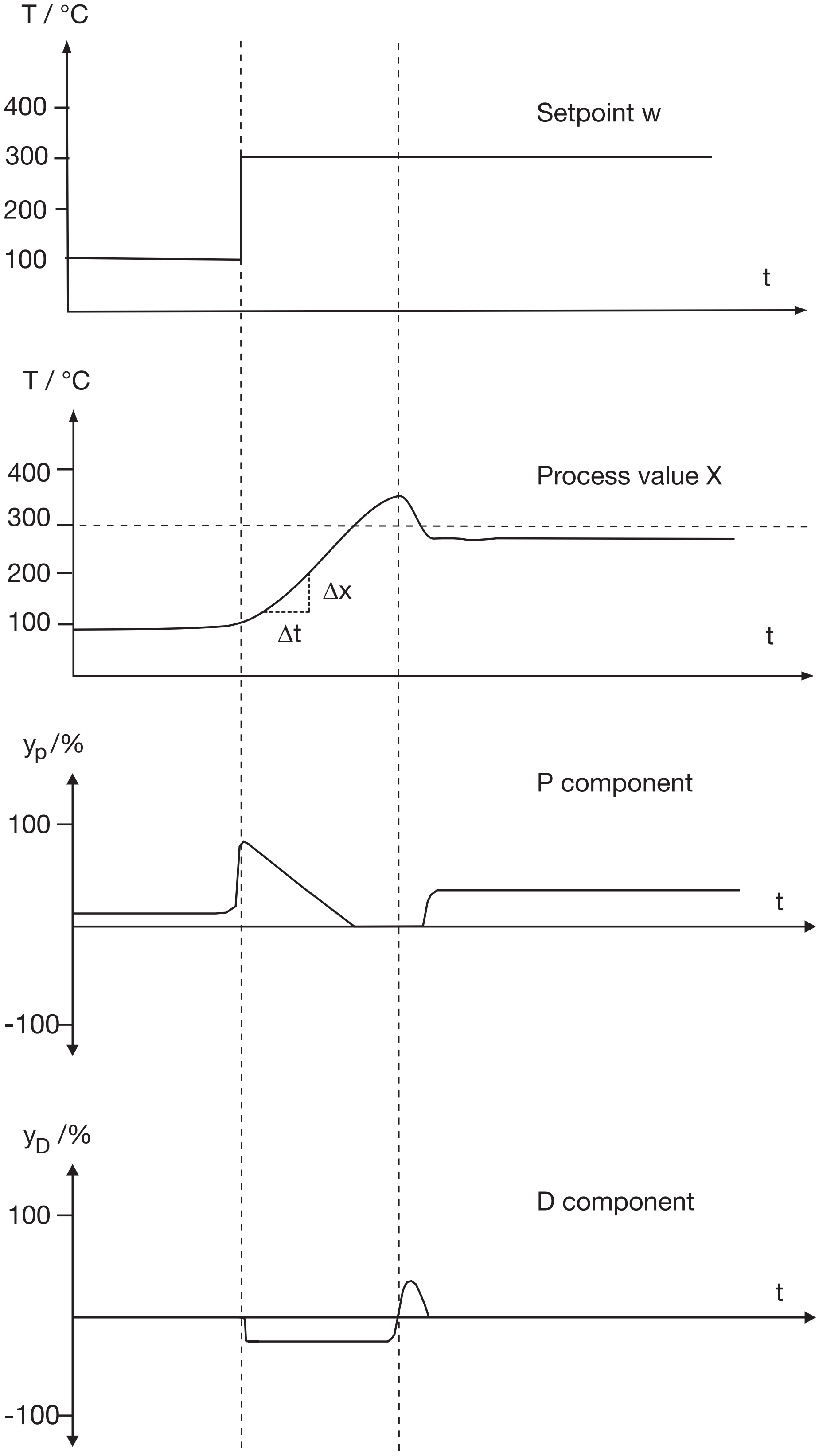


Fig. 37: Formation of the manipulating variable in a PD controller

Fig. 38 shows the ramp function response for a PD controller, where we can imagine the increasing control deviation resulting from a falling process variable.

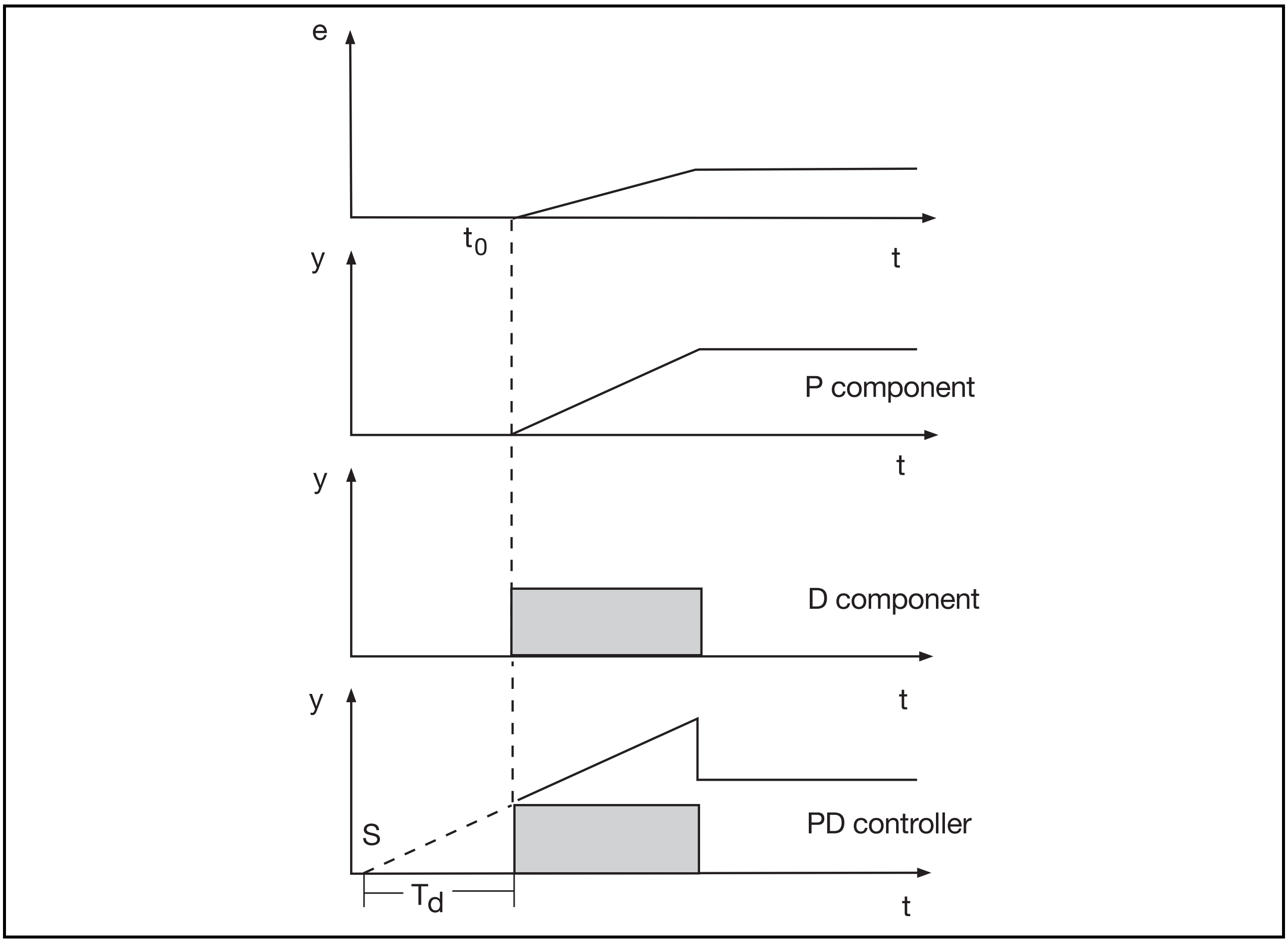


Fig. 38: Response of a PD controller to a ramp function

From Fig. 38 we can see that there is a noticeable manipulating variable from the D component at the start of the ramp function, since this manipulating variable is proportional to the rate of change of the process value. The P component needs a certain time, namely the derivative time T_d , to reach the same value manipulating variable as the D component has built up. The derivative time is obtained by projecting the diagonally rising line back to its point of intersection S with the time axis.

Mathematically, the rate of change v is obtained from the change in control deviation “de” per unit time “dt”:

$$v = \frac{de}{dt}$$

For the PD controller, this leads to the following control equation:

$$y = \frac{1}{X_P} \bullet 100\% \bullet \left(e + T_v \bullet \frac{de}{dt}\right)$$

In principle, the D component has the following effects:
As soon as the process variable changes, the D component counteracts this change.
For a controller with an inverse operating sense (i.e. for heating), this means for example:

- If the process variable reduces as a result of a disturbance in the process, the D component forms a positive manipulating variable, which counteracts the reduction in the process variable.
- If the process variable increases as a result of a disturbance in the process, the D component forms a negative manipulating variable, which counteracts the increase in the process variable.

3.5.1 The practical D component - the DT_1 element

In principle, we could also consider the step response of a PD controller in the same way as previously for P and PI controllers. Now, however, the rate of change at a step is infinitely large. In theory, the D signal derived from a step would therefore be an infinitely high and infinitely narrow spike (see Fig. 39). Theoretically, this means that the manipulating variable has to take up an infinitely high value for an infinitely short time, and then return immediately to the value produced by the P component. This is simply not possible, for both electrical and mechanical reasons. Furthermore, such a short pulse would scarcely influence the process. In practice, the immediate decay is prevented by forming the D component through a DT_1 element. This element consists of a D component, which we have already met in this chapter, in series with a T_1 element. The T_1 element behaves like a first-order process with a transfer coefficient of 1.

Fig. 39 shows the step response of the “practical” D component. T_1 is the time constant of the T_1 element. In practice, this time constant is set at $T_d/4$, and when T_d is changed, the time constant is changed by the same ratio. The derivative time T_d can be determined from the step response of the “practical” D component, on the basis of the ratio $T_1 = T_d/4$.

T_1 is specified by the manufacturer, and cannot be altered by the user.

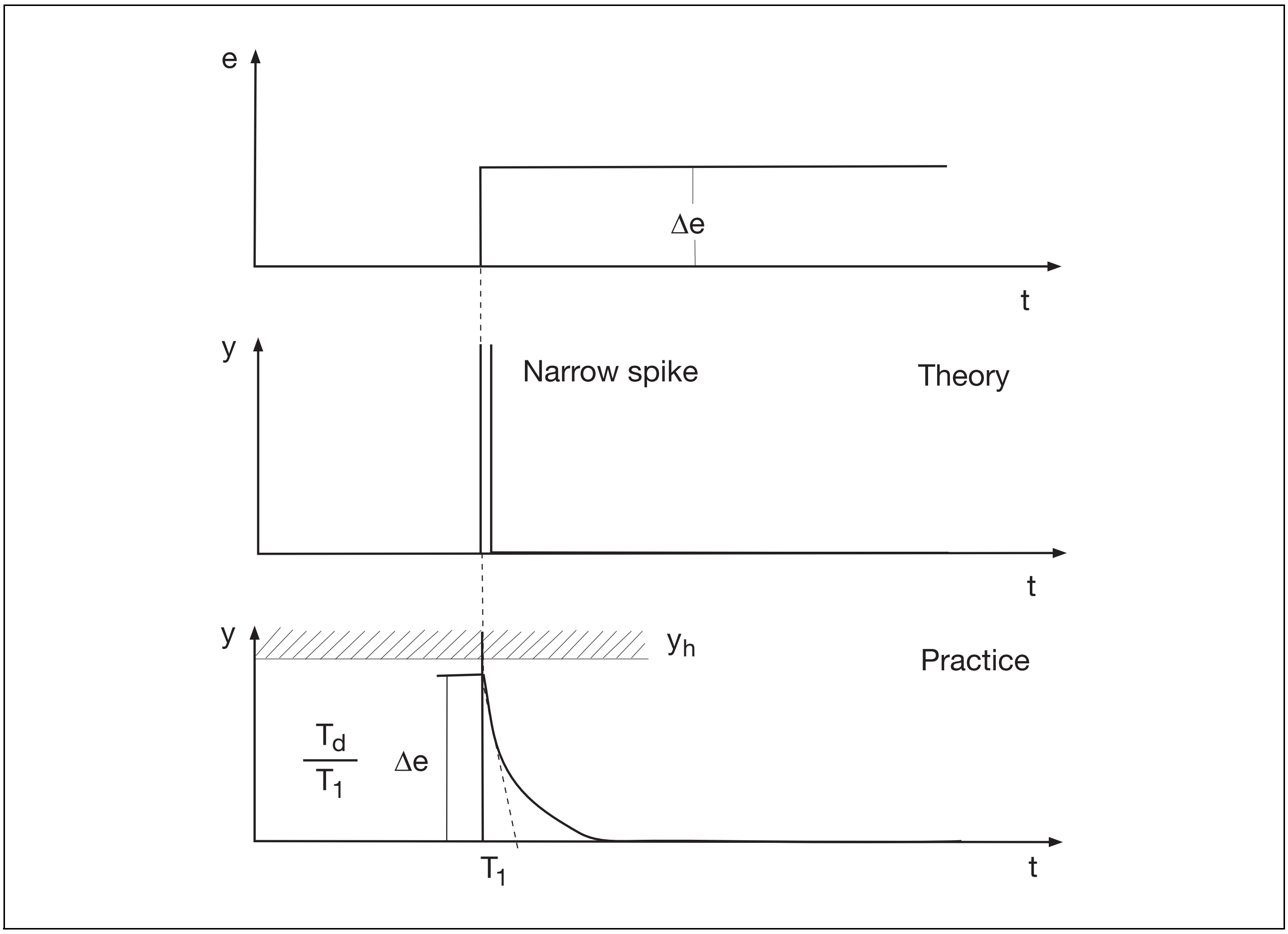


Fig. 39: Step response of a DT_1 element

Summarizing the main points:
A pure D controller has no practical importance since it takes no account of a permanent deviation, and simply responds to the rate of change of the process variable. By comparison, the PD controller is very widely used. The D component ensures a fast response to disturbances, whereas its "braking behavior" also stabilizes the control loop. The D component is not suitable for processes with pulsating variables, such as pressure and flow control.
The main application for the PD controller is where tools or products are prone to damage if the setpoint is exceeded. This applies particularly to plastics processing machines. However, PD controllers, like the P controller, always have a permanent deviation, when controlling processes with self-limitation.

3.6 PID controller

We have seen earlier that the combination of a D component or an I component with a P controller offered certain advantages in each case. Now it seems logical to combine all three structures, resulting in the PID controller.

With this controller, the X_P, T_n, T_d parameters are adjusted for the P, I and D action. These three components can be seen in the step response of a PID controller (see Fig. 40).

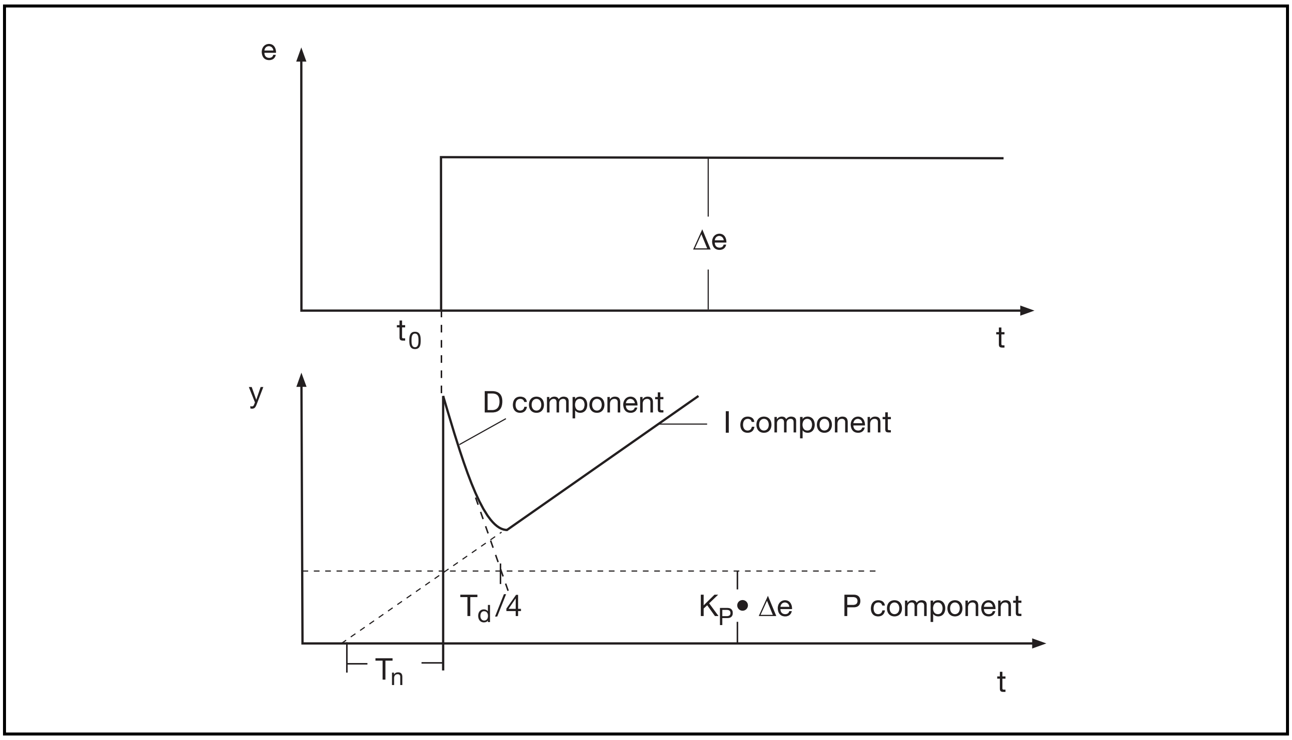


Fig. 40: Step response of a PID controller

According to DIN 19 225, such a controller obeys the following controller equation:

$$\Delta y = K_P \bullet \left(e + \frac{1}{T_n} \int e \bullet dt + T_d \bullet \frac{de}{dt} \right)$$

(ideal PID controller)

As already discussed in the previous section, the individual parameters (K_P, T_d, T_n) have different effects on the individual components.

larger X_P (corresponds to smaller K_P): corresponds to smaller P component

larger T_n : corresponds to reduced I component

larger T_d : corresponds to increased D component

On some controllers with PID action, T_d and T_n cannot be adjusted separately. Practical experience has shown that optimum performance is obtained with a ratio $T_d = T_n / 4$ to 5. This ratio is frequently a fixed setting on the controller, and only one parameter can be varied (usually T_n).

We can summarize by noting that the PID controller brings together the best characteristics of the P, I and D controllers. The P component responds with a suitable manipulating variable when a deviation occurs. The D component counteracts changes in the process variable, and increases the stability of the control loop. The permanent deviation is removed by the I component. The PID type of controller is used for most applications.

3.6.1 Block diagram of the PID controller

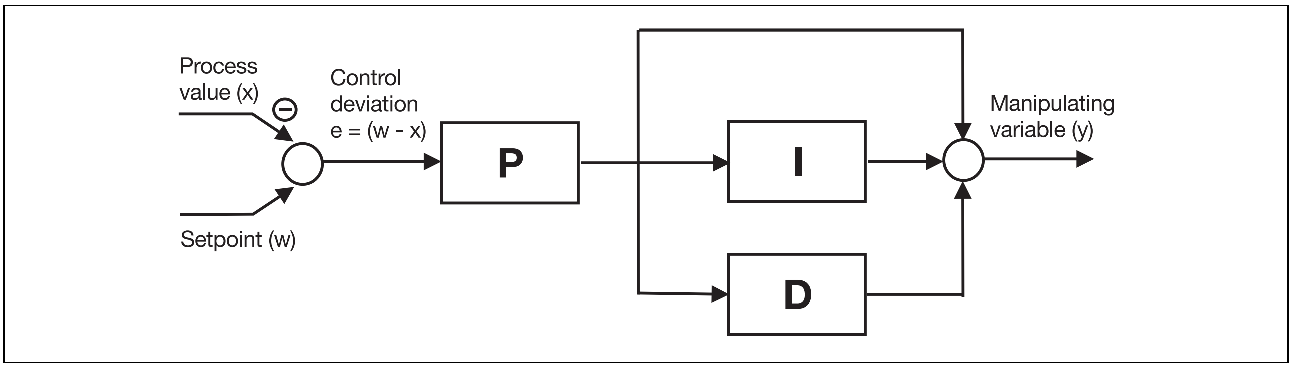


Fig. 41: Block diagram of the PID controller

As we have already seen in this chapter, from the controller equations for the PI, PD and PID controllers, the I and D actions of a PID controller are influenced not only by the adjustment of the T_n and T_d parameters, but also by the proportional gain with X_P. If the proportional gain of a PID controller is doubled (by halving X_P), the controller not only has double the proportional action, but the I and D components are also increased to double the value.

An example
The PID controller shown in Fig. 41 has settings T_n = 10sec and X_P = 100 (the D component should be disregarded in this example). The control deviation is 2.

When K_P and X_P are given as percentage values, the P component has a gain of:

$$1\left(K_P = \frac{1}{X_P} \bullet 100\ \%\right)$$

The control deviation is thus offered directly to the I component. We already know from Chapter 3.3 “I controller”, that an I controller requires a time equal to T_n to fully reproduce the input signal at its output (percentage values). The I component would thus require 10sec before it has increased its manipulating variable by 2%. X_P is now set to 50, so that the gain of the P component is 2.

Now the control difference is first amplified by a factor of 2, before it is offered to the I component. The I component now increases its manipulating variable by 4% every 10 seconds. The effect of the I component was also amplified by a factor of 2.

Changing the proportional gain in a PID controller changes the I and D action to the same extent

4.1 Operating methods for control loops with continuous controllers

The previous chapters dealt with the individual elements of a control loop, the process and the controller. Now we consider the interaction between these two elements in the closed control loop.

Amongst other things, the stable and unstable behavior of a control loop should be examined, together with its response to setpoint changes and disturbances. In the section on “Optimization”, we will come across the various criteria for adjusting the controller to the process.

We also often refer to the static and dynamic behavior of the control loop. The static behavior of a control loop characterizes its steady state on completion of all dynamic transient effects, i.e. its state long after any earlier disturbance or setpoint change. The dynamic behavior, on the other hand, shows the behavior of the control loop during changes, i.e. the transition from one state of rest to another. We have already discussed this kind of dynamic behavior in Chapter 2 “The process”.

When a controller is connected to a process, we expect the process variable to follow a course like that shown in Fig. 42.

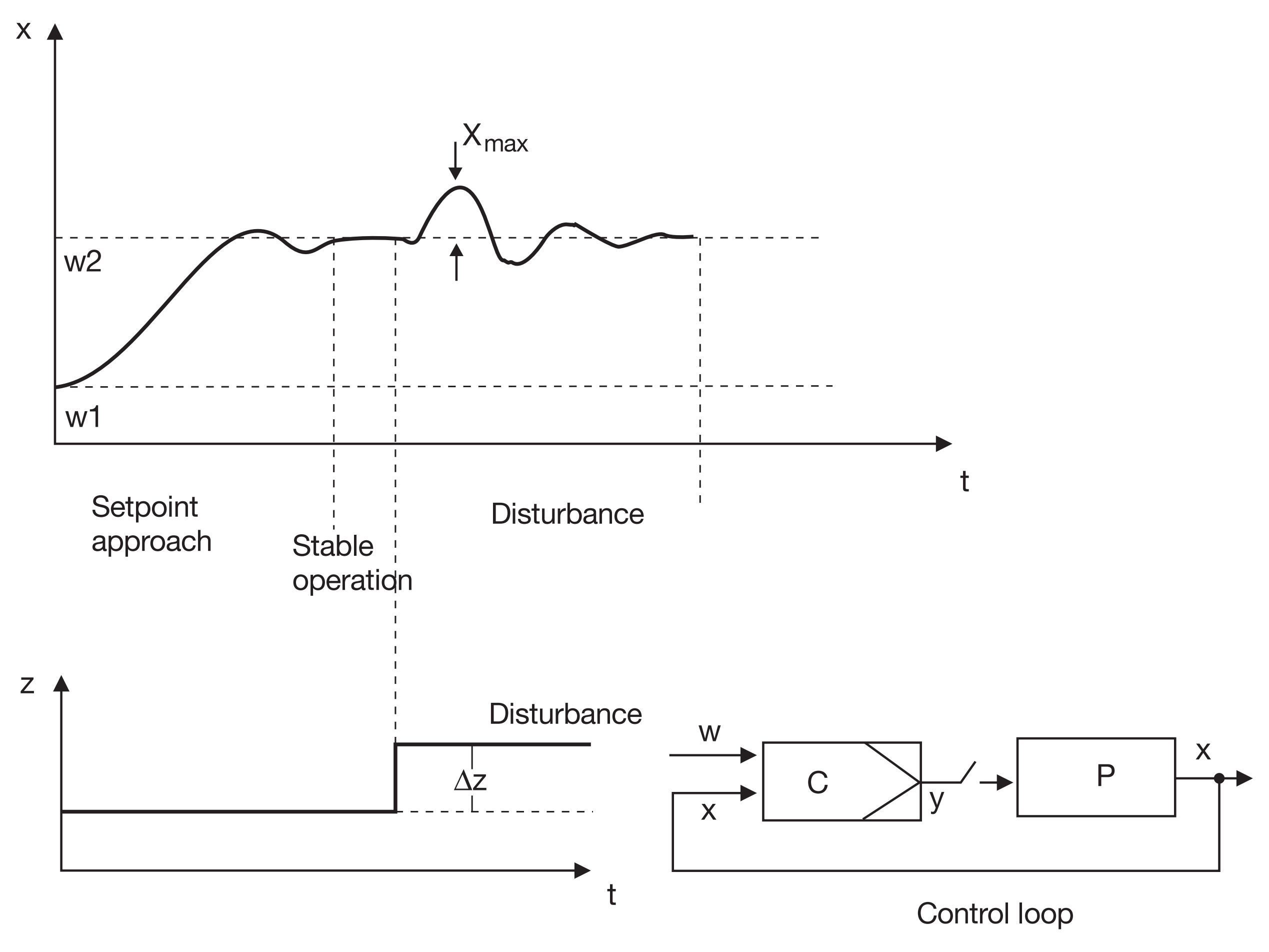


Fig. 42: Transition of the process variable in the closed control loop

- After the control loop is closed, the process variable (x) should reach and hold the predetermined setpoint (w) as quickly as possible, without appreciable overshoot. In this context, the run-up to a new setpoint value is also called the setpoint response.
- After the start-up phase, the process variable should maintain a steady value without any appreciable fluctuations, i.e. the controller should have a stable effect on the process.
- If a disturbance occurs in the process, the controller should again be able to control it with the minimum possible overshoot, and in a relatively short response time. This means that the controller should also exhibit a good disturbance response.

4.2 Stable and unstable behavior of the control loop

After the end of the start-up phase, the process variable should take up the steady value, predetermined by the setpoint, and enter stable operation. However, it could happen that the control loop becomes unstable, and that the manipulating variable and process variable perform periodic oscillations. Under certain circumstances, this could result in the amplitude of these oscillations not remaining constant, but instead increasing steadily, until it fluctuates periodically between upper and lower limit values. Fig. 43 shows the two cases of an unstable control loop.

Here, we often talk about the self-oscillation of a control loop. Such unstable behavior is mostly caused by low noise levels present in the control loop, which introduce a certain restlessness into the loop. Self-oscillation is largely independent of the construction of the control loop, whether it be mechanical, hydraulic or electrical, and only occurs when the returning oscillations have a larger amplitude than those sent out, and are in phase with them.

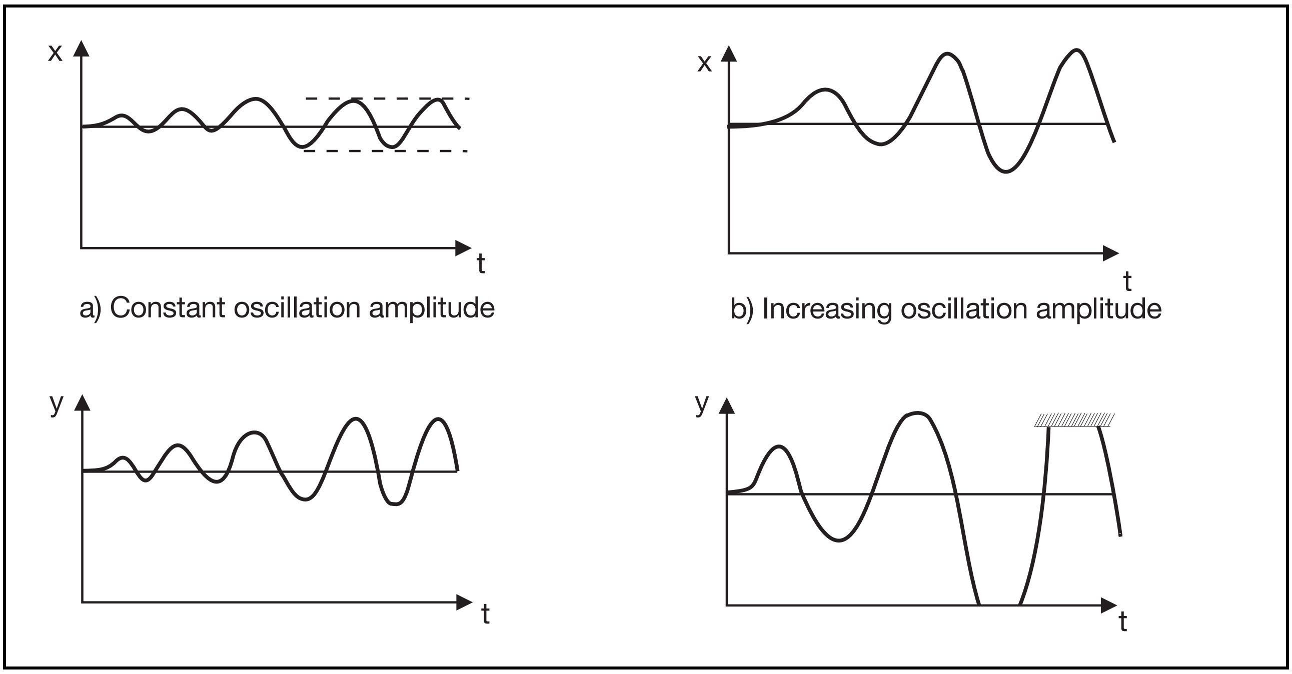


Fig. 43: The unstable control loop

If certain operating conditions, (e.g. new controller settings), are changed in a control loop that is in stable operation, there is always a possibility of the control loop becoming unstable. However, in practical control engineering, the stability of the control loop is an obvious requirement. We can generalize by stating that stable operation can be achieved in practice by choosing a sufficiently low gain in the control loop and a sufficiently long controller time constant.

4.3 Setpoint and disturbance response of the control loop

As already mentioned, there are basically two cases which result in a change in the process variable. When describing the behavior of a process in the control loop, we use the terms setpoint response or disturbance response, depending on the cause of the change:

Setpoint response
The setpoint has been adjusted and the process has reached a new equilibrium.

Disturbance response
An external disturbance affects the process and alters the previous equilibrium, until a stable process value has developed once again.

The setpoint response thus corresponds to the behavior of the control loop, following a change in setpoint. The disturbance response determines the response to external changes, such as the introduction of a cold charge into a furnace. In a control loop, the setpoint and disturbance responses are usually not identical. One of the reasons for this is that they act on different timing elements or at various intervention points in the control loop.

In many cases, only one of the two types of process response is important.

When a motor subjected to continuously variable shaft loading still has to maintain a constant speed, it is clearly only the disturbance response which is of importance. Conversely, in the case of a furnace, where the charge has to be brought to different temperatures over a period of time, in accordance with a specific setpoint profile, the setpoint response is of more interest.

The purpose of control is to influence the process in the desired manner, i.e. to change the setpoint or disturbance response. It is impossible to satisfactorily correct both forms of response in the same way. A decision must therefore be made whether to optimize the control for disturbance response or setpoint response. More about this in the section on “Optimization”.

4.3.1 Setpoint response of the control loop

As already explained, the main objective in a control loop with a good setpoint response is that, when the setpoint is changed, the process variable should reach the new setpoint value as quickly as possible and with minimal overshoot. Overshoot can be prevented by a different controller setting, but only at the expense of the stabilization time (see Chapter 4.1, Fig. 42). After closing the control loop, it takes a certain time for the process variable to reach the setpoint value predetermined at the controller. This approach to the setpoint can be made either gradually (creep) or in an oscillatory manner (see Fig. 44).

Which particular control loop response is considered most important varies from one case to another, and depends on the process to be controlled.

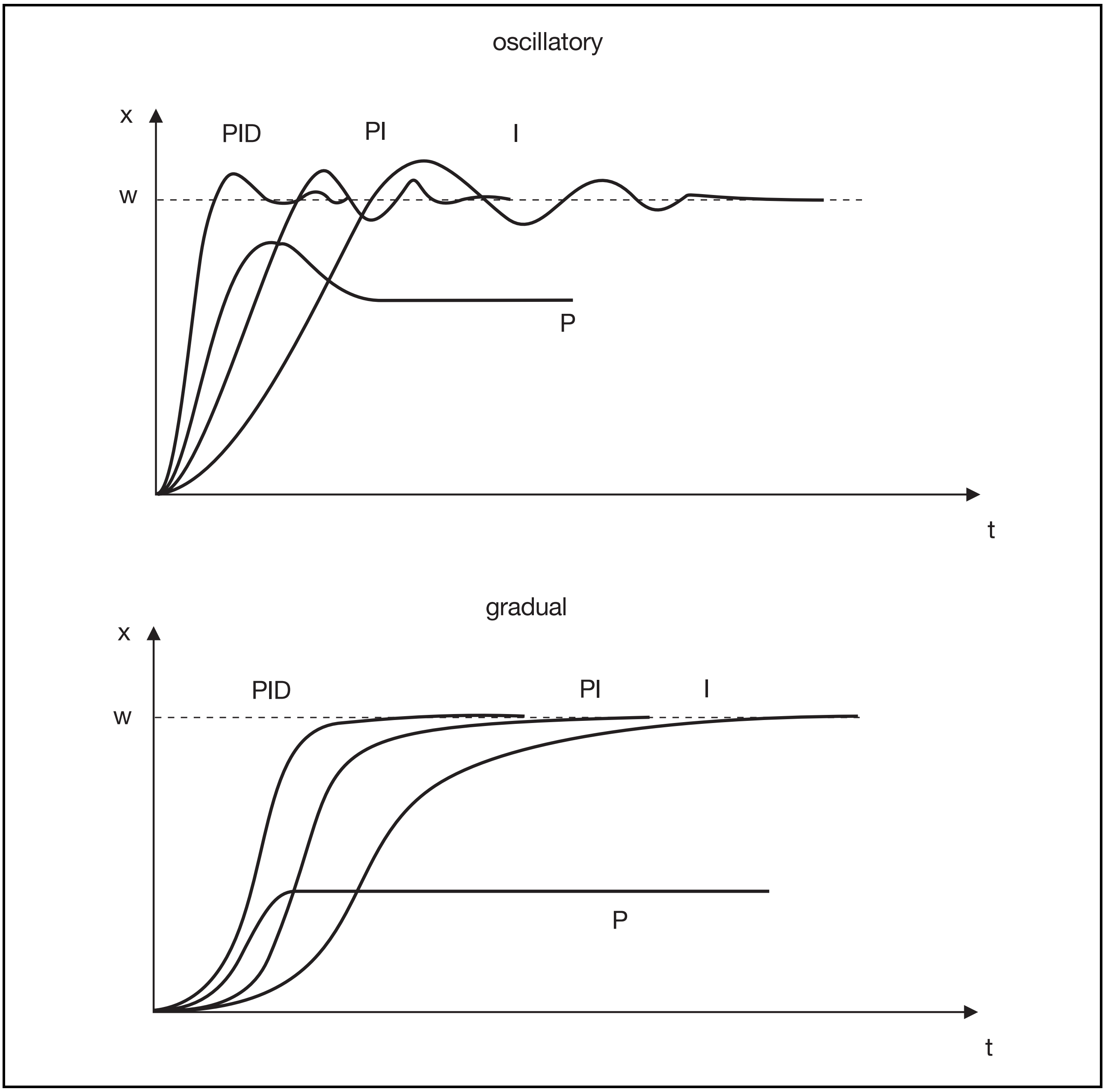


Fig. 44: Approach to the setpoint

4.3.2 Disturbance response

When the start-up phase is complete and the control loop is stable, the controller now has the task of suppressing the influence of disturbances, as far as possible. When a disturbance does occur, it always results in a temporary control deviation, which is only corrected after a certain time. To achieve good control quality, the maximum overshoot, the permanent control deviation and the stabilization time should be as small as possible (see Chapter 1.4, Fig. 3). As the size of disturbances of the characteristics in a control loop normally has to be accepted as given, good control quality can only be achieved by a suitable choice of controller type and an appropriate optimization.

The disturbances can act at different points in the process. Depending on the point of application of the disturbance, its effect on the dynamic transition of the process variable will differ. Fig. 45 shows the course of a disturbance step response of the process, when a disturbance acts at the beginning, in the middle and at the end of the process.

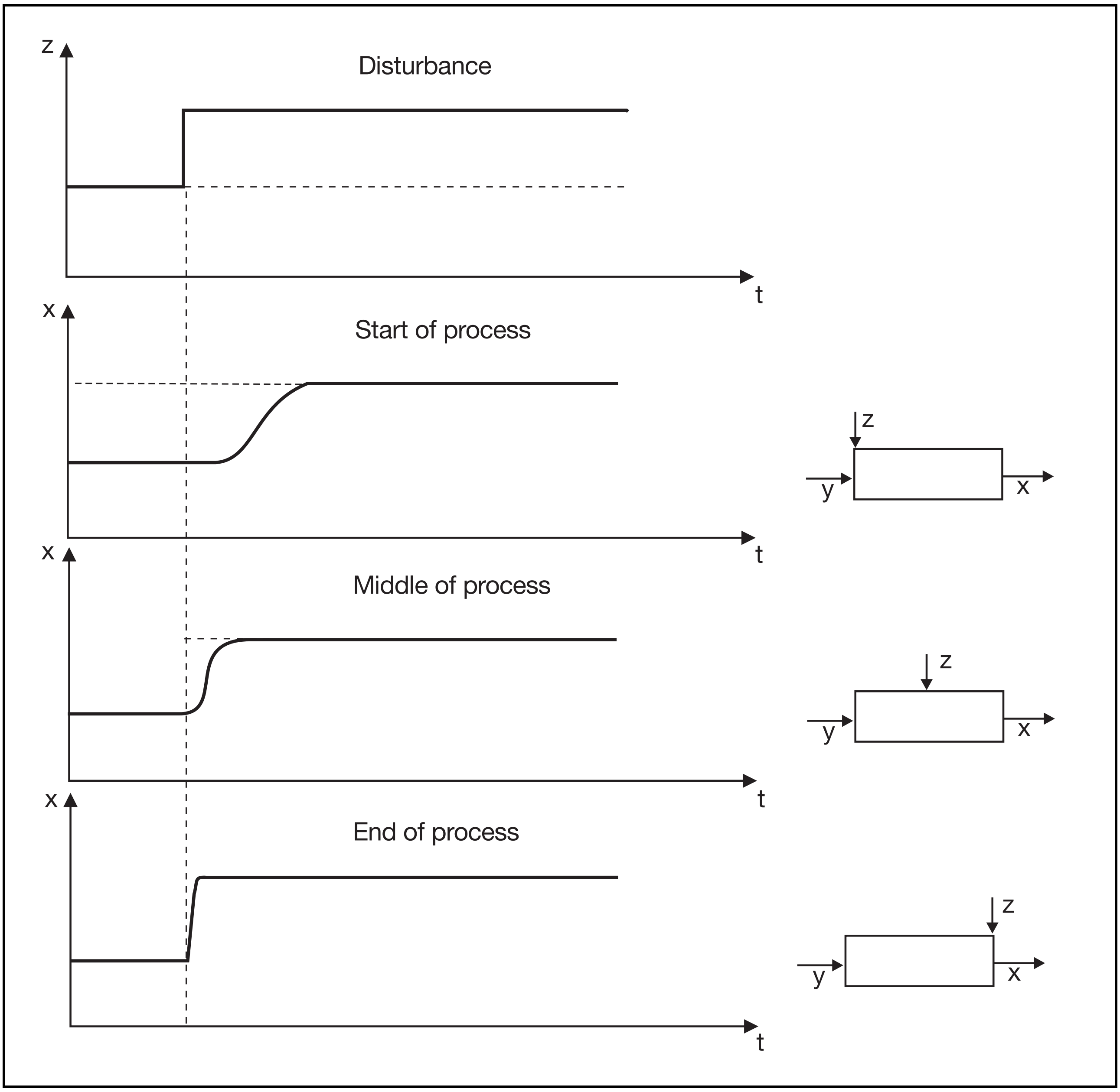


Fig. 45: Disturbance step response of a process

4.4 Which controller is best suited for which process?

After selecting a suitable controller according to type, dimensions etc. (see Chapter 1.5), the problem now arises of deciding which dynamic response should be employed to control a particular process. With modern microprocessor controllers, the price differentials between P, PI and PID controllers have been eroded. Hence it is no longer crucial nowadays, whether a control task can still be solved with just a P controller.

Regarding dynamic action, the following general points can be made:
P controllers have a permanent deviation, which can be removed by the introduction of an I component. However, there is an increased tendency to overshoot, because of this I component, and the control becomes a little more sluggish. Accurate stable control of processes affected by delays can be achieved by a P controller, but only in conjunction with an I component. With a dead time, an I component is always required, since a P controller, used by itself, leads to oscillations. An I controller is not suitable for processes without self-limitation.

The D component enables the controller to respond more quickly. However, with strongly pulsating process variables, such as pressure control etc., this leads to instabilities. Controllers with a D component are very suitable for slow processes, such as those found in temperature control. Where a permanent deviation is unacceptable, the PI or PID controller is used.

The relationship between process order and controller structure is as follows:
For processes without self-limitation or dead time (zero-order), a P controller is adequate. However, even in apparently delay-free processes, the gain of a P controller cannot be increased indefinitely, as the control loop would otherwise become unstable, because of the small dead times that are always present. Thus, an I component is always required for accurate control at the setpoint.

For first-order processes with small dead times, a PI controller is very suitable.

Second-order and higher-order processes (with delays and dead times) require a PID controller. When very high standards are demanded, cascade control should be used, which will be discussed in more detail in Chapter 6. Third-order and fourth-order processes can sometimes be controlled satisfactorily with PID controllers, but in most cases this can only be achieved with cascade control.

On processes without self-limitation, the manipulating variable must be reduced to zero after the setpoint has been reached. Thus, they cannot be controlled by an I controller, since the manipulating variable is only reduced by an overshoot of the process variable. For higher-order processes without self-limitation, a PI or PID controller is suitable.

Summarizing the selection criteria results in the following tables:

	Permanent deviation		**No permanent deviation**	
	P	**PD**	**PI**	**PID**
Temperature	simple process for low demands	simple process for low demands	suitable	highly suitable
Pressure	mostly unsuitable	mostly unsuitable	highly suitable; for processes with long delay time I controller as well	suitable, if process value pulses not too much
Flow	unsuitable	unsuitable	suitable, but I controller frequently better	suitable
Level	with short dead time suitable	suitable	suitable	highly suitable
Conveyor	unsuitable because of dead time	unsuitable	suitable, but I controller mostly best	nearly no advantages compared with PI

Table 3: Selection of the controller type for controlling the most important process variables

Process	**Controller structure**			
	P	**PD**	**PI**	**PID**
pure dead time	unsuitable	unsuitable	very suitable, or pure I controller	
first-order with short dead time	suitable if deviation is acceptable	suitable if deviation is acceptable	highly suitable	highly suitable
second-order with short dead time	deviation mostly too high for necessary X_P	deviation mostly too high for necessary X_P	not as good as PID	highly suitable
higher-order	unsuitable	unsuitable	not as good as PID	highly suitable
without self-limitation with delay	suitable	suitable	suitable	suitable

Table 4: Suitable controller types for the widest range of processes

4.5 Optimization

Controller optimization (or “tuning”) means the adjustment of the controller to a given process. The control parameters (X_P, T_n , T_d) have to be selected such that the most favorable control action of the control loop is achieved, under the given operating conditions. However, this optimum action can be defined in different ways, e.g. as a rapid attainment of the setpoint with a small overshoot, or a somewhat longer stabilization time with no overshoot.

Of course, as well as very vague phrases like “stabilization without oscillation as far as possible”, control engineering has more precise descriptions, such as examining the area enclosed by the oscillations and other criteria. However, these adjustment criteria are more suitable for comparing individual controllers and settings under special conditions (laboratory conditions). For the practical engineer working on the installation, the amount of time taken up and the practicability on site are of greater significance.

The formulae and control settings given in this chapter are empirical values from very different sources. They refer to certain idealized processes and may not always apply to a specific case. However, anyone with a knowledge of the various adjustment parameters, on a PID controller, for example, should be able to adjust the control action to satisfy the relevant demands.

Apart from the mathematical derivation of the process parameters and the controller data derived from them, there are various empirical methods. One method consists of periodically changing the manipulating variable and investigating how the process variable follows these changes. If this test is carried out for a range of oscillation frequencies of the setpoint, the amplitude and phase shift of the resulting process variable fluctuations can be used to determine the frequency response curve of the process. From this it is possible to derive the control parameters. Such test methods are very expensive, involve increased mathematical treatment, and are not suitable for practical use.

Other controller settings are based on empirical values, obtained in part from lengthy investigations. Such methods of selecting controller settings (especially the Ziegler and Nichols method and that of Chien, Hrones and Reswick) will be discussed in more detail later.

4.5.1 The measure of control quality

Standard text book instructions for controller optimization are usually based on step changes in, for example, a disturbance or the setpoint. Disturbances are usually assumed to act at the start of the process.

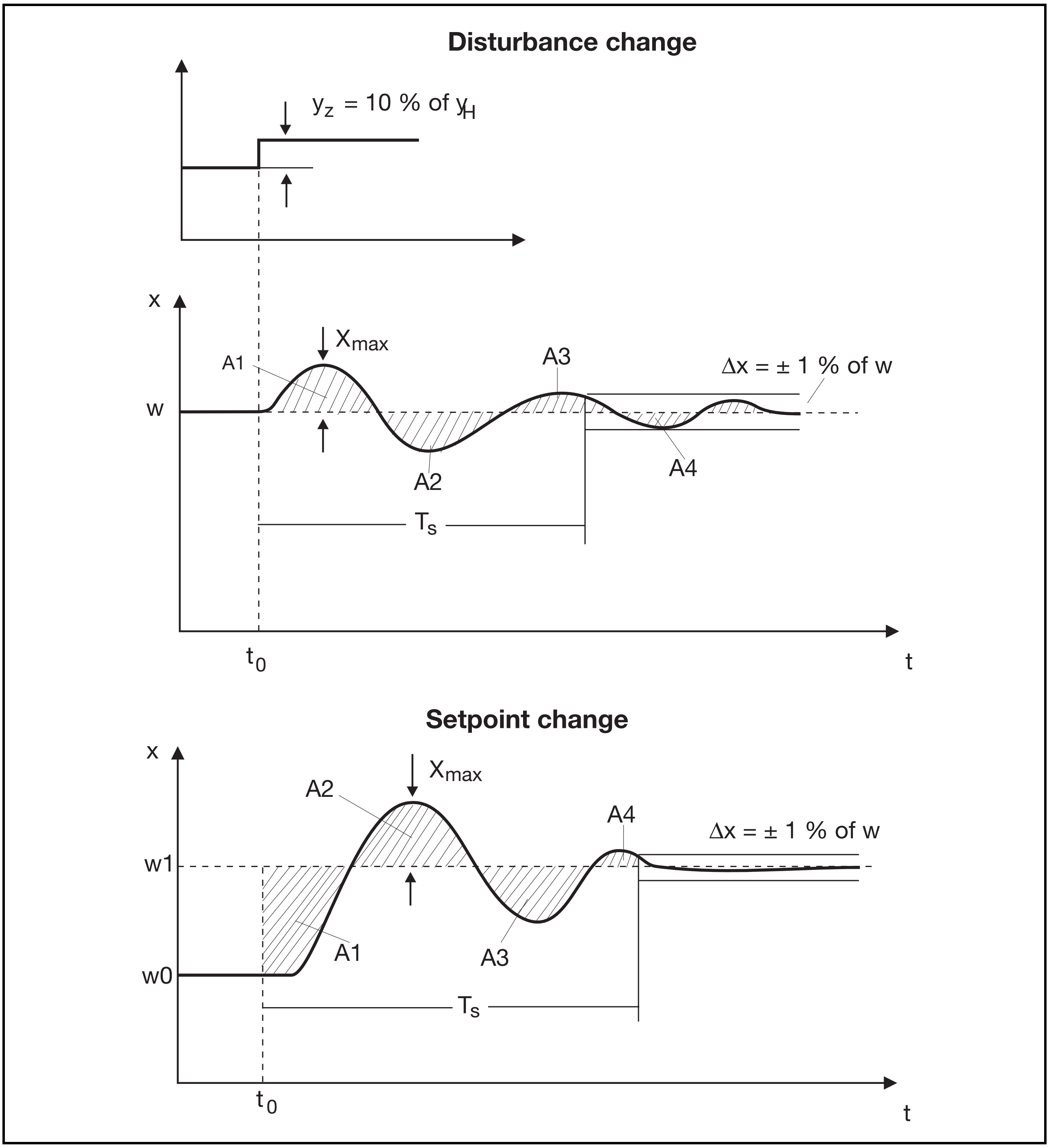


Fig. 46: The measure of control quality

This type of disturbance is also the most important one, as it frequently occurs in normal operation, testing is very feasible and because of its clear mathematical analysis. Fig. 46 shows that for a step change disturbance, the overshoot amplitude X_o and the stabilization time T_s offer a measure of quality. For a more exact definition of the stabilization time, we have to establish when the control action is regarded as complete. It is convenient to regard stabilization after a disturbance as being complete, when the control difference remains within ±1 % of the setpoint w. For expediency, the size of the disturbance is taken as 10 % of y_H .

In addition to the overshoot amplitude and the stabilization time, for mathematical analysis, the area of the control error is also used as a measure of control quality (see Fig. 46).

Linear control area (linear optimum): $[A]_{min} = A_1 - A_2 + A_3 \ldots$

Magnitude control area (magnitude optimum): $[A]_{min} = | A_1 | + | A_2 | + | A_3 | + \ldots$

Squared control area (squared optimum): $[A]_{min} = A1^2 + A2^2 + A3^2 + \ldots$

Without doubt, quite apart from any other considerations, one controller setting can be said to exhibit better control quality than another, if the resulting overshoot amplitudes are smaller and the stabilization time is shorter. Some tests indicate, however, that within certain limits it is possible to have a small overshoot at the expense of a longer stabilization time, and vice versa. For the given control error area, there is a definite controller setting at which the areas are at a minimum.

As mentioned several times previously, differing levels of importance are attached to the various measures of control quality, depending on the type of process variable and the purpose of the installation (see also Chapter 4.3 “Setpoint and disturbance response of the control loop”).

4.5.2 Adjustment by the oscillation method

In the oscillation (or limit cycle) method, devised by Ziegler and Nichols, the control parameters are adjusted until the stability limit is reached, and the control loop formed by the controller and the process starts to oscillate, i.e. the process variable performs periodic oscillations about the setpoint. The controller setting values can be determined from the parameters found from this test. The procedure can only be used in processes that can actually be made unstable and where an overshoot does not cause danger. The process variable is made to oscillate by initially reducing the controller gain to its minimum value, i.e. by setting the proportional band to its maximum value. The controller must be operating as a pure P controller; for this reason, the I component (T_n) and the D component (T_d) are switched off. Then the proportional band X_P is reduced until the process variable performs undamped oscillations of constant amplitude.

This test produces:

- the critical proportional band X_{Pc} , and
- the oscillation time T_c of the process variable (see Fig. 47)

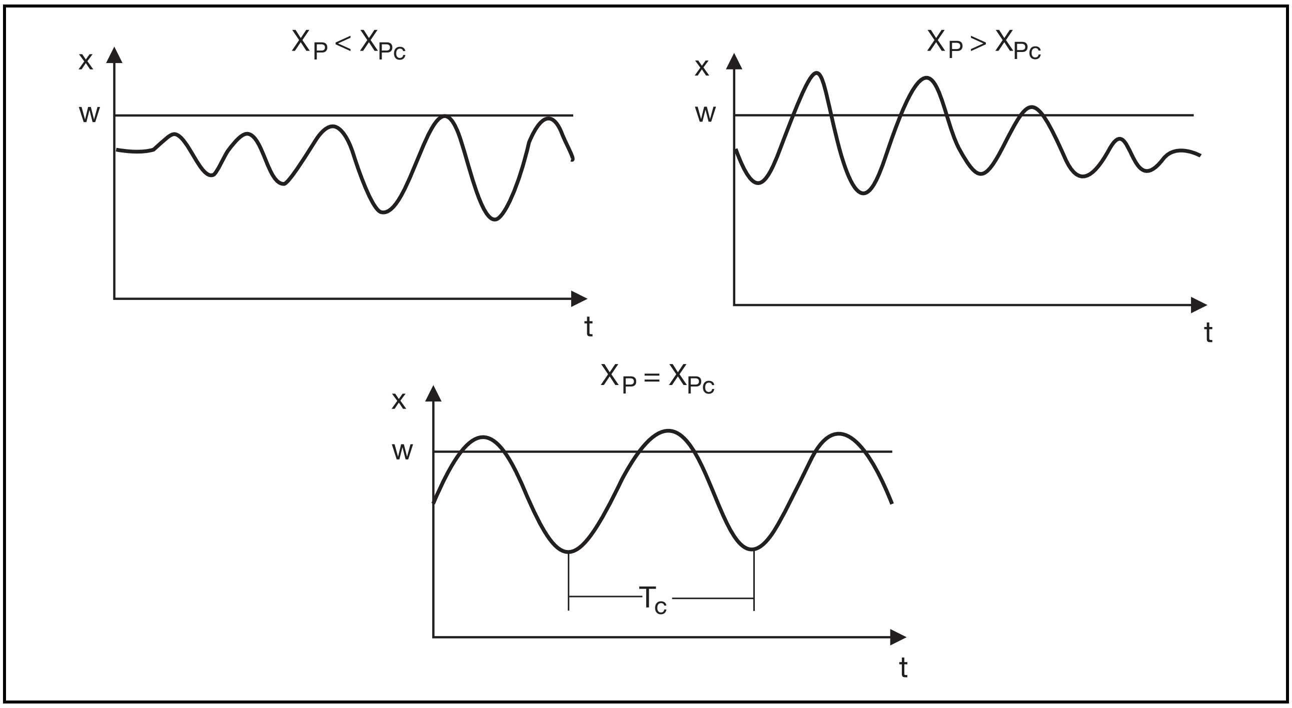


Fig. 47: Oscillation method after Ziegler and Nichols

The controller can then be set to the following values:

Controller structure	
P	$X_P = X_{Pc} / 0.5$
PI	$X_P = X_{Pc} / 0.45$ $T_n = 0.85 \cdot T_c$
PID	$X_P = X_{Pc} / 0.6$ $T_n = 0.5 \cdot T_c$ $T_d = 0.12 \cdot T_c$

Table 5: Adjustment formulae based on the oscillation method

Without doubt, the advantage of this process is that the control parameters can be studied under operational conditions, as long as the adjustments described succeed in achieving oscillations about the setpoint. There is no need to open the control loop. Recorder data is easily evaluated; with slow processes, the values can even be determined by observing the process variable and using a stopwatch. The disadvantage of this method is that it can only be used on processes which can be made unstable, as mentioned above.

The Ziegler and Nichols adjustment rules apply mainly to processes with short dead times and with a ratio T_g / T_u greater than 3.

4.5.3 Adjustment according to the transfer function or process step response

Another method of determining the parameters involves measuring process-related parameters by recording the step response, as already described in Chapter 2.6. It is also suitable for processes which cannot be made to oscillate. However, it does require opening the control loop, for instance, by switching the controller over to manual mode in order to exert a direct influence on the manipulating variable. If possible, the step change in manipulating variable should be made when the process variable is close to the setpoint.

A method that can be used to determine the control parameters when the process parameters are known has been developed by Chien, Hrones, and Reswick (CHR). This approximation method yields good control parameters, not only for disturbances, but also for setpoint changes, and is suitable for processes with PTn structure (with n equal to 2 or greater).

The step response is used to determine the delay time T_u, the response time T_g and the process transfer coefficient K_S (see Fig. 46).

$$K_S = \frac{\text{change in process variable}}{\text{change in manipulating variable}} = \frac{\Delta x}{\Delta y}$$

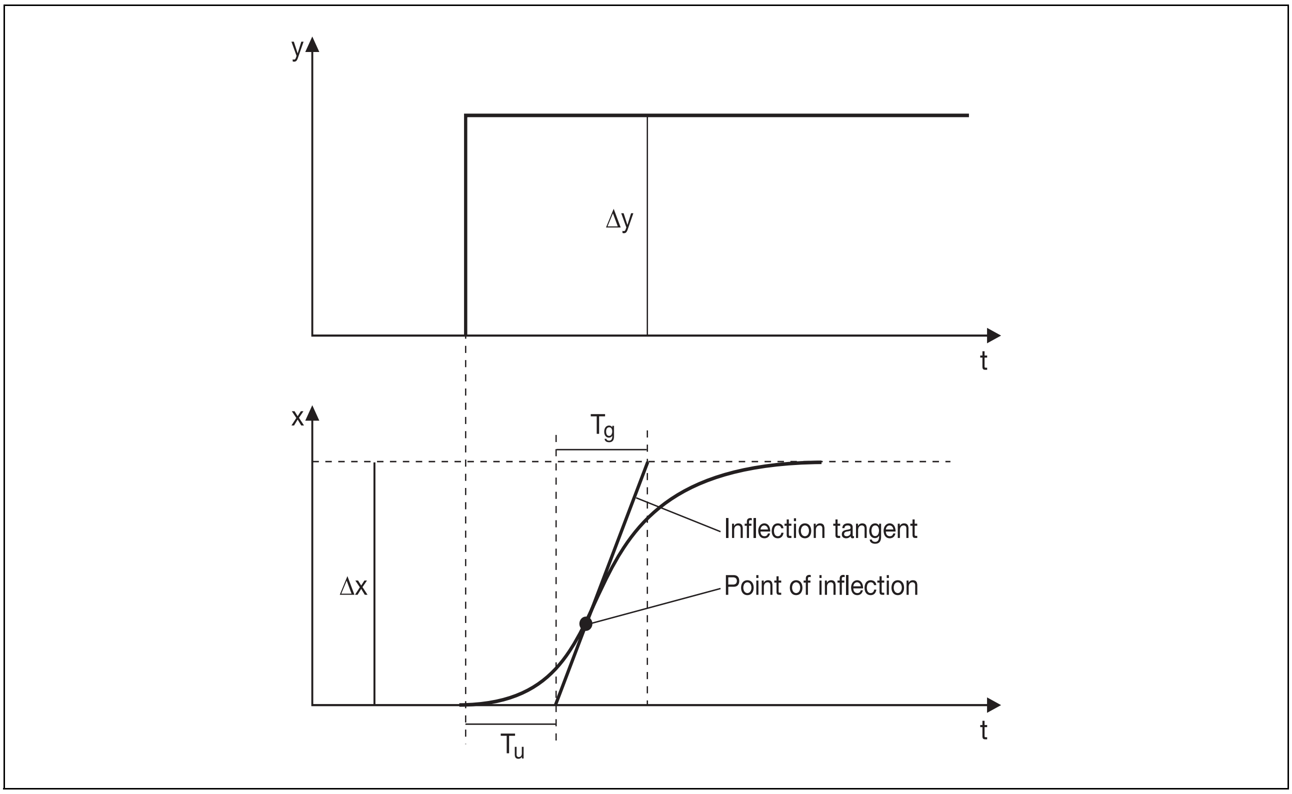


Fig. 48: Adjustment according to the step response

The values found are applied using the following setting rules:

Controller structure	Setpoint	Disturbance
P	$X_P = 3.3 \cdot K_S \cdot (T_u/T_g) \cdot 100\%$	$X_P = 3.3 \cdot K_S \cdot (T_u/T_g) \cdot 100\%$
PI	$X_P = 2.86 \cdot K_S \cdot (T_u/T_g) \cdot 100\%$ $T_n = 1.2 \cdot T_g$	$X_P = 1.66 \cdot K_S \cdot (T_u/T_g) \cdot 100\%$ $T_n = 4 \cdot T_u$
PID	$X_P = 1.66 \cdot K_S \cdot (T_u/T_g) \cdot 100\%$ $T_n = 1 \cdot T_g$ $T_d = 0.5 \cdot T_u$	$X_P = 1.05 \cdot K_S \cdot (T_u/T_g) \cdot 100\%$ $T_n = 2.4 \cdot T_u$ $T_d = 0.42 \cdot T_u$

Table 6: Formulae for adjustment according to the step response

Example:
T_n, T_d and X_P have to be determined for a temperature control process. The future operating range is at 200°C. The heater power can be continuously adjusted using a variable transformer, and the maximum output is 4kW. The disturbance response parameters for a PID structure have to be evaluated.

First the heater power is set to give a temperature close to the future working point, for example 180°C at 60% heater power. Now the heater power is suddenly increased to 80% and the variation in temperature recorded. The inflection tangent is then drawn, giving T_u as 1 min and T_g as 10 min. If it is difficult to determine the point of inflection, the change in manipulating variable must be increased, i.e. by starting the test at a lower heater power and ending at a higher heater power. The final temperature in the case illustrated here is 210°C.

This gives the following values:

$$K_S = \frac{\Delta x}{\Delta y} = \frac{210\ °C - 180\ °C}{80\ \% - 60\ \%} = \frac{30\ °C}{20\ \%} = 1.5\ °C/\%$$

Using the values obtained for T_u and T_g , the parameters are calculated as follows:

$$T_n \approx 2.4 \bullet T_u \approx 2.4 \bullet 1 \approx 2.4\ \text{min} \approx 144\ \text{sec}$$

$$T_d \approx 0.42 \bullet T_u \approx 0.42 \bullet 1 \approx 0.42\ \text{min} \approx 25\ \text{sec}$$

$$X_P \approx 1.05 \bullet K_S \bullet \frac{T_u}{T_g} \approx 1.05 \bullet 1.5\frac{°C}{\%} \bullet \frac{1\ \text{min}}{10\ \text{min}} \bullet 100\ \% \approx 15{,}75\ °C$$

We should not overlook a certain disadvantage of this process. In practice, the graph very rarely shows a very clear point of inflection. Hence, drawing the tangent at the point of inflection can lead to errors in determining the values of T_u and T_g , which may or may not be significant. The method illustrated is still very useful for forming a first impression of the controller settings. Other criteria can then be used to tune the settings.

4.5.4 Adjustment according to the rate of rise

In some cases, there can be difficulties in determining the response time T_g when using the methods described above. Very often, the manipulating variable can only be set to either 0 or 100%. Operating the process continuously at 100% manipulating variable can be highly destructive.

A more constructive alternative is to avoid determining T_g, and instead to evaluate the maximum rate of rise V_{max}. To do this, the manipulating variable is suddenly set to 100% and the output of the process observed (see Fig. 49). The process variable will only start to change after a certain time, following the change in manipulating variable. The rate of change will increase continuously until the point of inflection is reached. At the point of inflection, the process variable approaches its final value more and more slowly. Using this method, it is necessary to wait until the point of inflection is reached, and then set the manipulating variable back to 0% again. It is important to remember that, especially in processes with long delays, such as furnaces, the process variable can continue to increase considerably, even after the heating has been switched off.

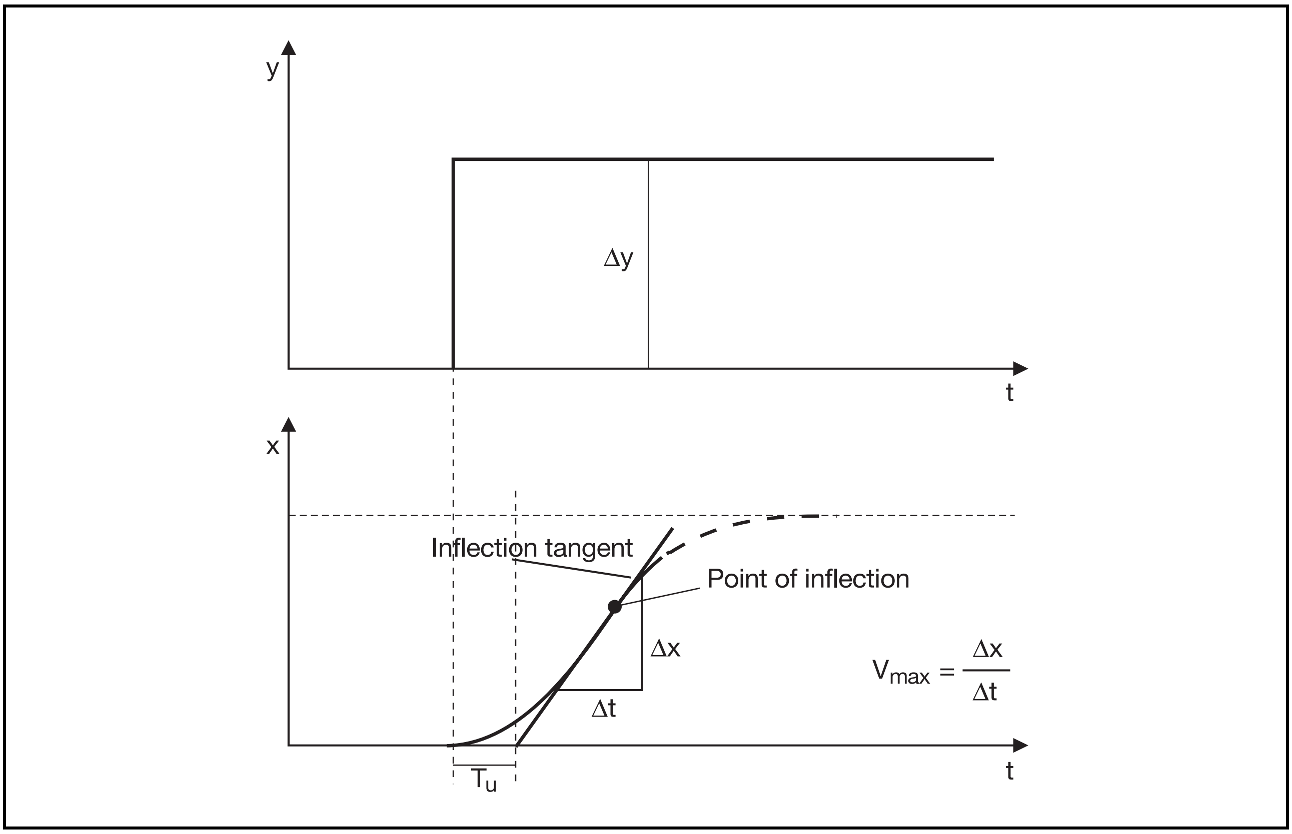


Fig. 49: Adjustment according to the rate of rise

The tangent at the point of inflection is now drawn, and V_{max} determined from the gradient triangle. Using the delay time determined in a similar manner from the step response, the controller settings can be implemented in accordance with the table which follows later.

The method described yields even better values if the controller and any manipulating device that might be present allow the manipulating variable to be set to any value. In this case, the step change in manipulating variable should be made close to the setpoint required later:

Example:
The future setpoint value of a furnace is 300°C. The existing controller is set to manual mode and the manipulating variable manually increased until the furnace temperature reaches 280°C; this temperature is reached at, say, 60% manipulating variable. Now the manipulating variable is suddenly set to 100%, and the point of inflection awaited. To apply the adjustment according to the

rate of rise, for this example, we also need the height of the step change in manipulating variable Δy (40%) and the manipulation range y_H (in this case 100%, as the controller can be set to 100% manipulating variable) for use in the formulae which follow later.

Controller structure	100% step change in MV	Any changes in MV
P	$X_P = V_{max} \cdot T_u$	$X_P = V_{max} \cdot T_u \cdot y_H/\Delta y$
PI	$X_P = 1.2 \cdot V_{max} \cdot T_u$ $T_n = 3.3 \cdot T_u$	$X_P = 1.2 \cdot V_{max} \cdot T_u \cdot y_H/\Delta y$ $T_n = 3.3 \cdot T_u$
PID	$X_P = 0.83 \cdot V_{max} \cdot T_u$ $T_n = 2 \cdot T_u$ $T_d = 0.5 \cdot T_u$	$X_P = 0.83 \cdot V_{max} \cdot T_u \cdot y_H/\Delta y$ $T_n = 2 \cdot T_u$ $T_d = 0.5 \cdot T_u$
PD	$X_P = 0.83 \cdot V_{max} \cdot T_u$ $T_d = 0.25 \cdot T_u$	$X_P = 0.83 \cdot V_{max} \cdot T_u \cdot y_H/\Delta y$ $T_d = 0.25 \cdot T_u$

Table 7: Formulae for adjustment according to the slew-rate response, for processes with self-limitation

4.5.5 Adjustment without knowledge of the process

Occasionally, a controller has to be adjusted to a process where it is simply not possible to record a transfer function or to open the control loop. If the process is not overly slow, the controller is initially set to a pure P structure with the largest proportional band possible, so that pure P action is achieved.

The setpoint is set close to the future operating point, and the process value indication on the controller is observed. After some time, the process value will stabilize at a value quite some way from the setpoint. This is because of the low gain through the large proportional band setting. X_P is now reduced, as a result of which the deviation from the setpoint becomes smaller and smaller. As X_P is further reduced, a point is eventually reached at which the process value starts to oscillate periodically. There is no point in reducing X_P any further, as it would only increase the amplitude of these oscillations. These oscillations are not usually symmetrical about the setpoint; their mean value is either above or below the setpoint. The reason for this, as we have already established, is the continued presence of the permanent deviation that a P controller produces.

The proportional band is now increased once again, until the process value becomes stable.

Next, the I component is added (PI structure), and the reset time T_n is reduced step by step. The process variable slowly approaches the setpoint, as a result of the I component. Reducing T_n still further accelerates the approach, but also leads to oscillations. We now apply a disturbance to the process, either by changing the setpoint or an external disturbance. The approach to the new setpoint is monitored. If the process value overshoots, we have to increase T_n . If the approach is only very slow, the reset time setting can be reduced still further.

The D component can be activated next, if required, (PID structure), by setting T_d to a value of approximately T_n /4.5.

The procedure described above is a widely used practical method, suitable for simple processes.

4.5.6 Checking the controller settings

We cannot expect the control loop to achieve optimum performance with the initial parameter settings. Some readjustment will usually be required, particularly on processes that are difficult to control, with a T_g / T_u ratio less than 3. The step response of the process variable to a setpoint change clearly shows any mismatch of the control parameters. The resulting transient response can be used to draw conclusions about any necessary corrections. Alternatively, an external disturbance can be applied to the process, for example, by opening a furnace door, and then analyzing the effects of the disturbance. A recorder is used to monitor the process variable, and the controller setting adjusted if necessary (see Fig. 50).

Increasing the proportional band X_P - corresponding to a reduction in controller gain - leads to a more stable transient response. Without an I component, a permanent deviation can be detected. Reducing X_P reduces the deviation, but a further reduction in proportional band eventually leads to undamped oscillations. Setting the controller setting just below self-oscillation, by setting a small X_P, leads to a small deviation, but this is not the optimum setting, as in this case the control loop is only very lightly damped. As a consequence, even small disturbances cause the process variable to oscillate.

The I component reduces the permanent deviation in accordance with the reset time T_n . If the I component is too low (T_n too large), the visible effect is that the process variable only creeps gradually towards the setpoint. A larger I component (T_n small) acts like an excessive control gain, and makes the control loop unstable, resulting in oscillations.

A large derivative time T_d has an initial stabilizing effect, but, with a pulsating process variable, it can also make the control loop unstable.

Fig. 50 indicates possible incorrect settings. It uses as an example the setpoint response of a third-order process with a PID controller.

When optimizing a controller, only one parameter should be adjusted at a time, then the effect of this change awaited before changing further parameters. Furthermore, we have to consider whether the controller should be optimized for disturbance response or setpoint response.

It is found, for example, that a “tight” controller setting with a high controller gain may indeed give a fast approach to the setpoint, but the control loop is poorly damped because of the high gain. This could mean that a short duration disturbance produces oscillation. In other words, a lower controller gain slows down the approach to the setpoint somewhat, but makes the entire control loop more stable.

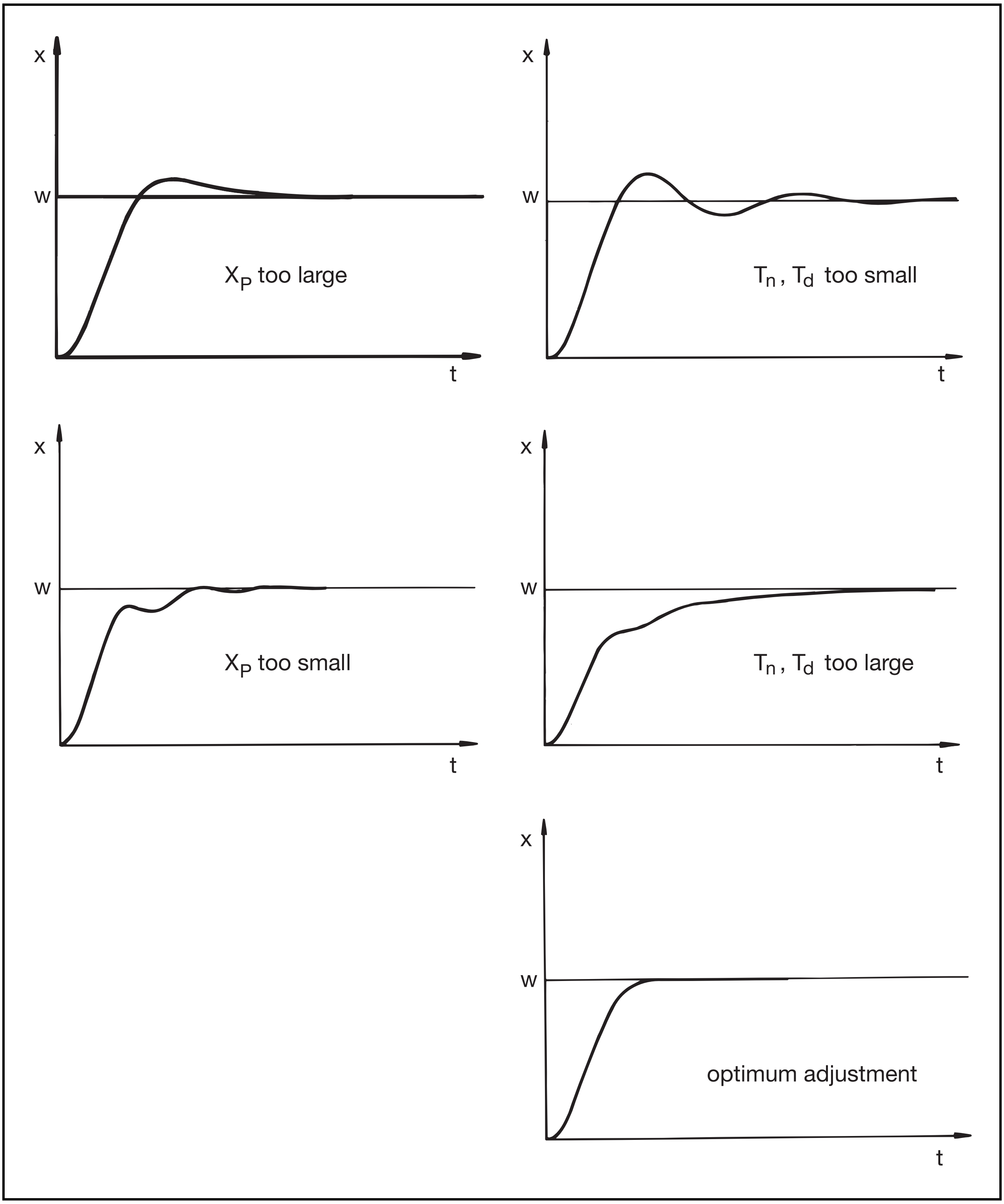


Fig. 50: Indications of possible incorrect adjustments

5.1 Discontinuous and quasi-continuous controllers

With the continuous controllers described previously, with P, PD, I, PI and PID actions, the manipulating variable y can take on any value between the limits y = 0 and $y = y_H$. In this way, the controller is always able to keep the process variable equal to the setpoint w.

In contrast to continuous controllers, discontinuous and quasi-continuous controllers do not have a continuous output signal, but one that can only have the state ON or OFF. The outputs from such controllers are frequently implemented as relays, but voltage and current outputs are also common. However, unlike the continuous controller, these are binary signals that can only have a value of 0 or the maximum value. These signals can be used to control devices such as solid-state relays.

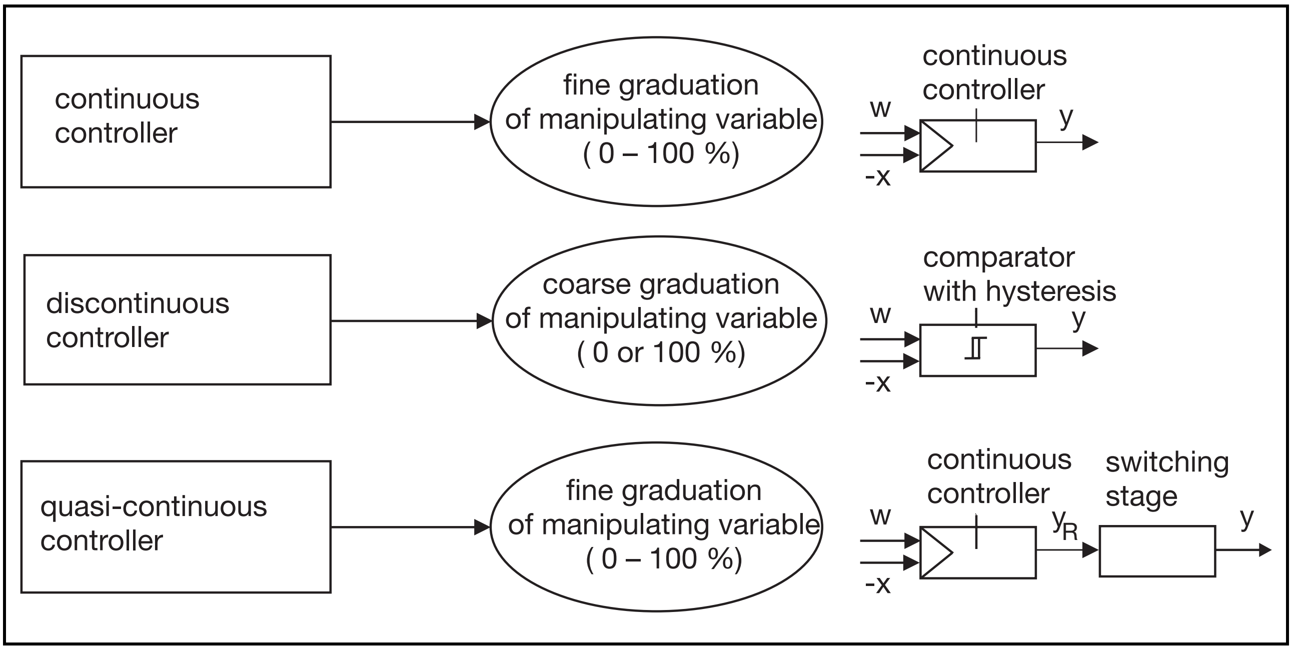


Fig. 51: Continuous, discontinuous and quasi-continuous controllers

In addition to these controller types with binary outputs, there are also 3-state and multi-state controllers, where the manipulating variable output can have 3 or more levels. A tri-state controller would, for instance, be used for heating and cooling tasks, or humidification and dehumidification.

It might be assumed that controllers with outputs which can only be in the ON or OFF state would only produce an unsatisfactory control action. But surprisingly enough, satisfactory results for the intended purposes can be achieved in many control processes, particularly with quasi-continuous controllers. Discontinuous and quasi-continuous controllers are very widely used, because of the simple construction of the output stage and the actuators that are required, resulting in lower costs. They are found universally in those areas of process control where the processes are relatively slow and can be readily controlled with switching actuators.

The simplest controller with a binary output is the discontinuous controller, which is effectively a limit switch that simply switches the manipulating variable on or off, depending on whether the process variable goes below or above a predetermined setpoint. A simple example of such a controller is the two-state bimetallic temperature controller in an electric iron, or a refrigerator thermostat.

Quasi-continuous controllers can be put together, for example, by adding a switching stage to the output of a continuous controller (see Fig. 51), thus converting the continuous output signal into a switching sequence. P, PD, I, PI and PID actions can also be implemented for these controllers (Fig. 51) and the foregoing remarks about continuous controllers are also applicable.

5.2 The discontinuous controller

The discontinuous controller has only 2 switching states, i.e. the output signal is switched on and off, depending on whether the process variable goes below or above a predetermined limit or setpoint. These devices are also often used as limit monitors, which initiate an alarm message when a setpoint is exceeded.

A simple example of a mechanical discontinuous controller is, as previously explained, the bimetallic switch of an electric iron, which switches the heating element off when the set temperature is reached and switches it on again when the temperature falls by a fixed switching differential (hysteresis). There are other examples in the field of electronic controllers. For example, a resistance thermometer (Pt 100), whose electronic circuitry switches heating on if the temperature falls below a certain value, say 5°C, to provide frost protection for an installation. In this case, the resistance thermometer together with the necessary electronic circuitry takes the place of the bimetallic switch.

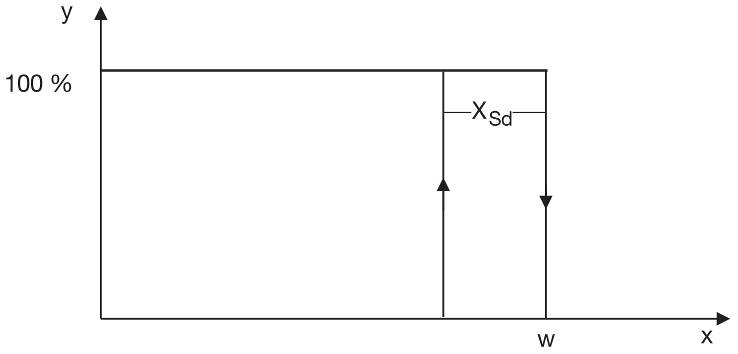


Fig. 52: Characteristic of a discontinuous controller

The discontinuous controller shown here supplies 100% power to the process until the setpoint is reached. If the process variable rises above the setpoint, the power is taken back to 0%. Apart from the hysteresis, we see that the discontinuous controller corresponds to a continuous controller with no proportional band ($X_P = 0$) and therefore “infinite” gain.

5.2.1 The process variable in first-order processes

If we connect a discontinuous controller, such as a rod thermostat, to a first-order process (e.g. a thermostatic bath with water circulation, warmed by an immersion heater), we find that the course of the process variable and manipulating variable is as shown in Fig. 53. In theory, the controller should switch off the energy when the setpoint is reached, the process variable would fall immediately and once again go below the setpoint. The controller would immediately switch on again, and so on. Because an idealized first-order process has no delay time, the relay would switch on and off continuously, and would be destroyed in a very short time.

For this reason, a discontinuous controller usually incorporates a switching differential X_{Sd} (also known as hysteresis) about the setpoint, within which the switch status does not change. In practice, the switching differential is often to one side of the setpoint, either below (for example with heating) or above (for example with cooling). Fig. 53 shows a case where the switching differential lies below the setpoint. The switch-off point of the controller is the setpoint w. In practice, as the process is not ideal (it has some delay time), the higher and lower values of the process variable do not coincide exactly with the switching edges of the differential (X_{Sd}).

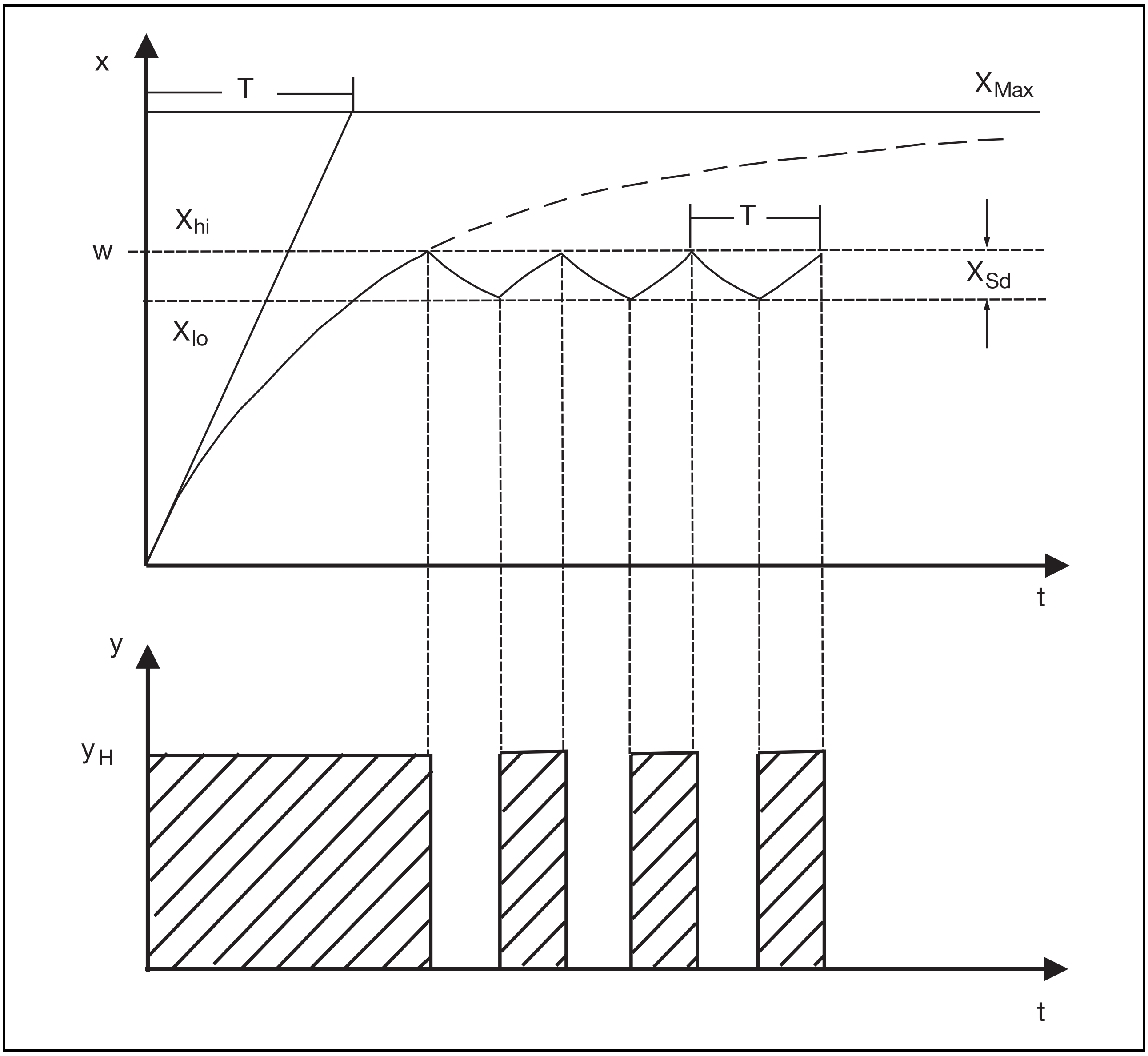


Fig. 53: Discontinuous controller in a first-order process

What matters however, is that the controller only switches when the process variable has moved outside the differential band that has been set. The process variable continually fluctuates, at least between the values X_{hi} and X_{lo}. The fluctuation band of the process variable is therefore influenced by the switching differential.

In a process with delay, the discontinuous controller can only maintain the process variable constant between the values X_{hi} and X_{lo}. The on-off switching is due to the manipulating variable being too large to maintain the process variable constant when it is switched on, and too small when it is switched off. In a large number of control tasks, where the process variable only needs to be maintained approximately constant, these fluctuations are not a problem. An example of this is a domestic electric oven, where it does not matter if the actual temperature fluctuates between 196°C and 204°C for a baking temperature of 200 °C.

If these continuous fluctuations of the process variable do cause problems, they can be minimized to a limited extent by selecting a smaller switching differential Xsd. This automatically leads to more switching operations per unit time, i.e. the switching frequency increases. This is not always desirable, as it affects the life of the controller relay.

It can be shown (mathematical details are not entered into here) that the following relationship exists between the switching frequency (f_{sw}) and the parameters T, X_{max} and X_{Sd} :

$$f_{sw} = \frac{1}{T_{osc}} = \frac{1}{4} \bullet \frac{X_{max}}{X_{Sd}} \bullet \frac{1}{T} \qquad \text{valid for} \qquad x \approx \frac{X_{max}}{2}$$

f_{sw} : switching frequency
T_{osc} : period of oscillation
X_{max} : max. process variable reached with the controller output permanently switched on
X_{Sd} : switching differential
T : time constant of the first-order process

We can see from this relationship that the shorter the time constant (T), the higher the switching frequency. A control process with short time constants will therefore produce a high switching frequency, which would contribute to rapid wear of the switching stage of the controller. For this reason, a discontinuous controller is unsuitable for this type of process.

5.2.2 The process variable in higher-order processes

In a process with delay, we have seen that under ideal conditions the fluctuation band is determined only by the switching differential X_{Sd} of the controller. The process itself has no effect here. In a process with several delays, which can be described as delay time, response time and transfer coefficient, this is no longer the case. As soon as there are any delays the process variable will continue to rise or fall after switch-off and will only return after reaching a maximum. Fig. 54 shows how the process variable overshoots the response threshold of the relay when the manipulating variable is switched on and off.

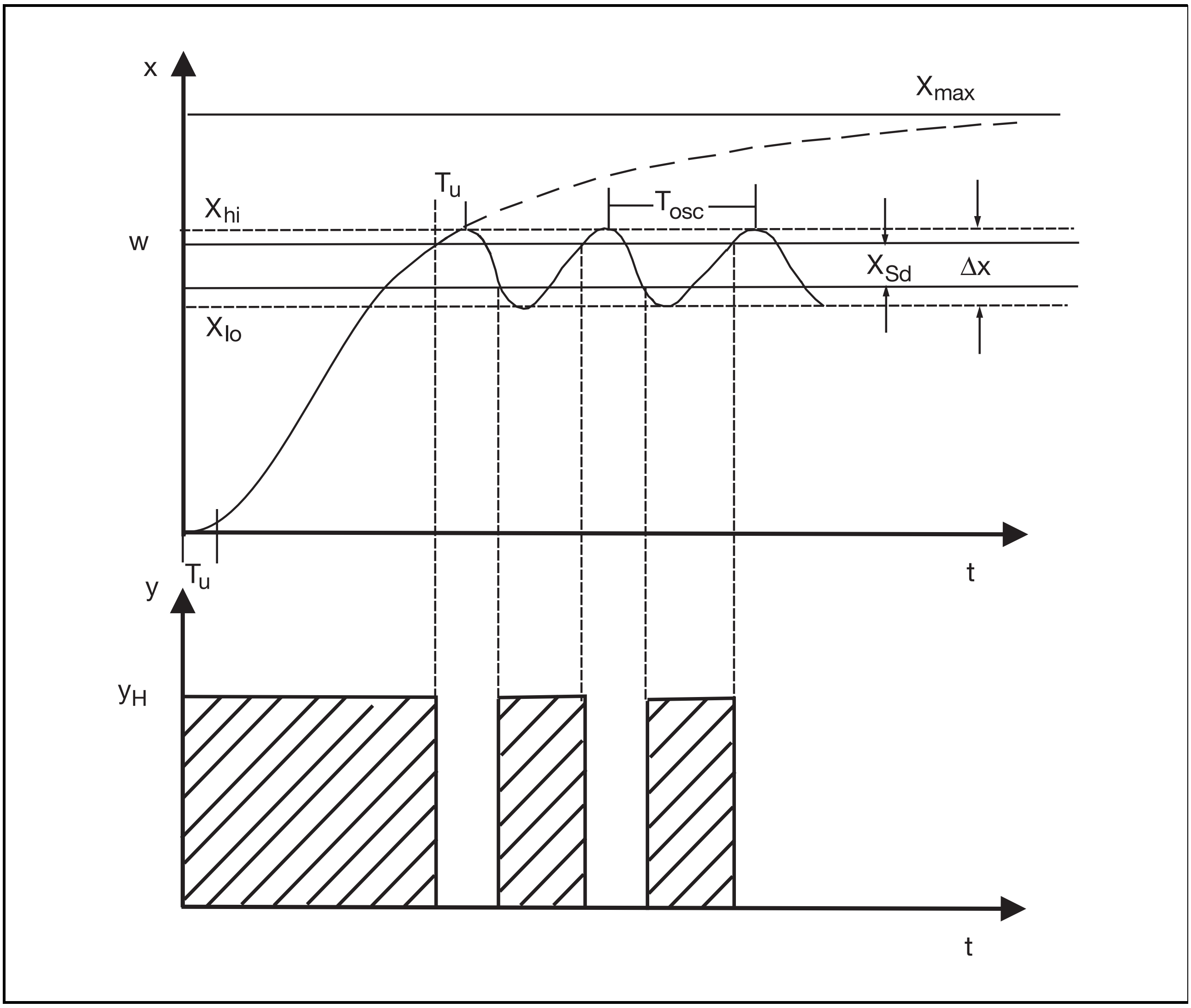


Fig. 54: Discontinuous controller in a higher-order process

This produces an overshoot of the process variable, with limits given by the values X_{hi} and X_{lo}. This means that the process variable fluctuates even when the controller has zero switching differential, as the process only reacts to the change in manipulating variable after the end of the delay time.

Once again, take the electrically heated furnace as an example. If the energy supply is switched off when the setpoint is reached, the temperature still continues to rise. The reason is that the temperature in the furnace only permeates slowly, and when the setpoint is reached, the heater rod is already at a higher temperature than that reported by the sensor. The rod and furnace material continue to supply additional heat. Similarly, when the heating is switched on again, heating-up is rather sluggish and initially the temperature continues to fall a little further after switch-on.

The more powerful the heater, the greater is the temperature difference between the heater rod and the sensor during heating-up, because of the process delay, and the process variable will overshoot the setpoint even more during heating-up. We use the term excess power in this connection, meaning the percentage by which the maximum power of a furnace is greater than the power required to approach a setpoint.

Example: A furnace which requires a manipulating variable of 2kW on average to stabilize at a setpoint of 200°C, but has a 4 kW continuous output rating, has an excess power of 100% at the working point of 200°C.

This means that the higher the excess power, the wider is the fluctuation band Δx of the process variable about the setpoint.

Now the present (but unwanted) fluctuation band of the process variable can be estimated for the case where 100% excess power is available:

It is assumed that the switching differential $X_{Sd} = 0$

$$\Delta x = X_{max} \bullet \frac{T_u}{T_g} \qquad \text{valid for} \qquad x \approx \frac{X_{max}}{2}$$

As we can see, the fluctuation band is dependent not only on X_{max} (with a linear process this is proportional to the excess power) but also on the ratio T_u/T_g, whose reciprocals we are already familiar with from Chapter 2, and which give a measure of how good the controllability of a process is. The shorter the delay time in comparison with the response time, the narrower is the fluctuation band. The formula given for the fluctuation band Δx is valid for $X_{Sd} = 0$. If there is a switching differential, this is also added to the fluctuation band.

This gives us the formula:

$$\Delta x = X_{max} \bullet \frac{T_u}{T_g} + X_{Sd}$$

The formula for the period of oscillation is: $T_{osc} = 4T_u$ (valid for $X_{Sd} = 0$)

If a switching differential X_{Sd} has been set, then the period of oscillation is slightly longer. From this we can derive the maximum switching frequency, which can be used to predict the expected contact life:

$$f_{osc} = \frac{1}{4T_u}$$

5.2.3 The process variable in processes without self-limitation

Because the step responses of an integrating process are linear, the behavior of a discontinuous controller is easy to describe and calculate. Here again the process value fluctuates between the given limits X_{hi} and X_{lo} (Fig. 55). In an ideal process without delay time T_u, the limit values are equal to the switching differential X_{Sd}.

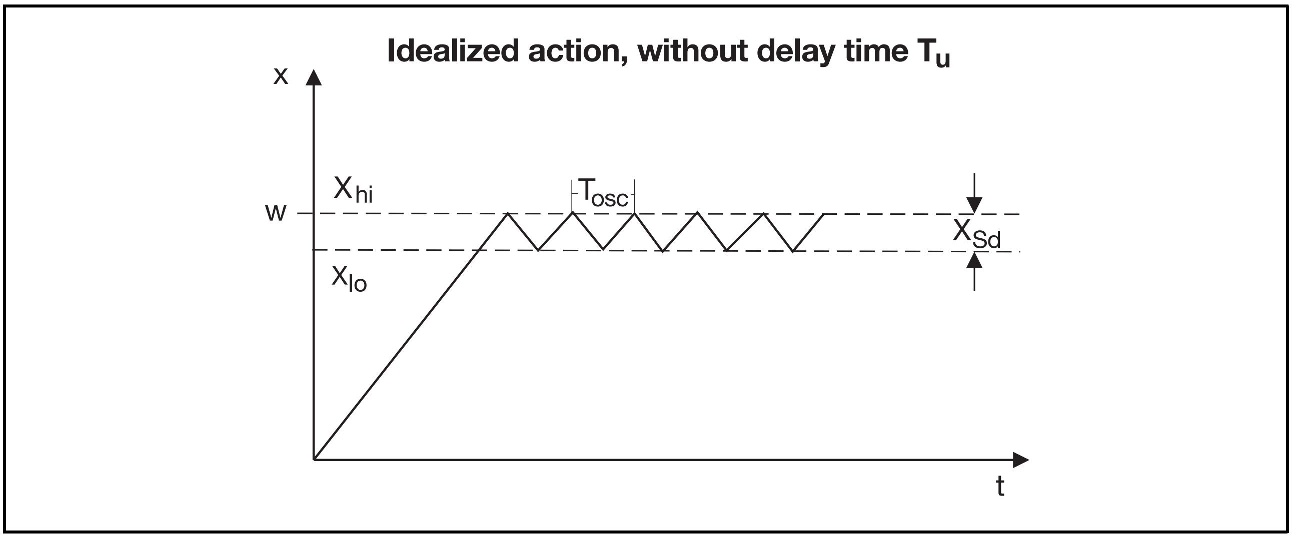


Fig. 55: Discontinuous controller in a process without self-limitation

The switching frequency f_{sw} is given by:

$$f_{sw} = \frac{1}{T_{osc}} = \frac{K_p \bullet y_H}{2 \bullet X_{Sd}}$$

K_p : proportionality factor of the process
y_H : maximum value of the manipulating variable

An example of such an application is a discontinuous controller used as a limit switch for level control of a water tank. The tank is used as a storage reservoir, from which water is drawn to meet demand or into which a constant amount flows.

Summarizing, we can say that the discontinuous controller offers the advantage of simple construction and few parameters which have to be set. The disadvantage is the fluctuation of the process variable about the setpoint. In non-linear processes these fluctuations can be wider in the lower operating range than in the upper, because the process has excess power here. Approaching the setpoint in the lower operating range will often result in wider fluctuations than in the upper operating range. The area of application for such discontinuous controllers is limited to applications where precise control is not required. In practice, these controllers are implemented through mechanical thermostats, level switches etc. If an electronic controller with a sensor is used, the controller is almost always provided with a dynamic action.

5.3 Quasi-continuous controllers: the proportional controller

As we have already seen, a quasi-continuous controller consists of a continuous controller and a switching stage. If this controller is operated purely as a proportional controller, then the characteristics which we have already met in Chapter 3.2.1 “The proportional band” apply equally here.

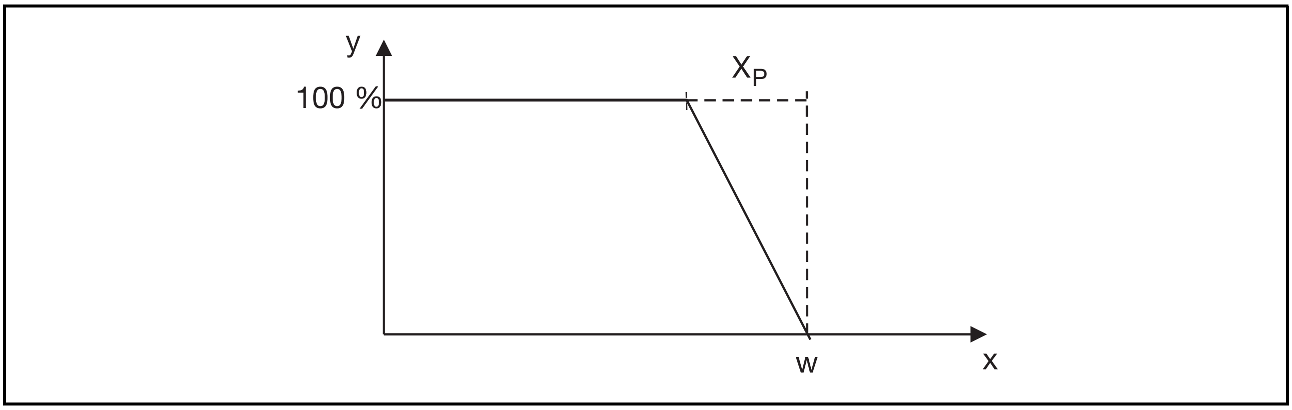


Fig. 56 : Proportional band of a quasi-continuous proportional controller

The quasi-continuous controller whose characteristic is shown in Fig. 56 always gives out a 100% manipulating variable, as long as the process value lies below the proportional band. As the process value enters the proportional band and approaches the setpoint, so the manipulating variable becomes progressively lower.

How can a controller with a switched output provide a virtually constant energy supply i.e. stepless dosage?

In the end it is immaterial whether a furnace is operated at 50% heating power all the time or at 100% heating power for only half the time. The quasi-continuous controller changes the switch-on ratio or ON-time ratio (also known as duty-cycle) of the output signal instead of changing the size of the output signal. An ON-time ratio of 1 corresponds to 100% of the manipulating variable, 0.25 corresponds to 25% of the manipulating variable, and so on.

The ON-time ratio, or duty-cycle R is defined as follows:

$$R = \frac{T_{on}}{T_{on} + T_{off}}$$

T_{on} = ON time
T_{off} = OFF time

Multiplying the ratio R by 100 gives the relative ON-time in % of R, which corresponds to the manipulating variable in %.

$$R(\%) = y = R \bullet 100\%$$

With a quasi-continuous controller the characteristic of the process (especially the time constants) exerts a strong influence on the course of the process variable. In a process where a disturbance is transmitted relatively slowly (a process with long time constants) and where energy can be stored, there is a smoothing effect on any pulses. With a suitable switching frequency, the use of a quasi-continuous controller with these processes achieves a similar result to that achieved using a continuous controller.

The situation is different with a very fast process, where there is hardly any smoothing of the constantly changing flow of energy, and the process variable fluctuates accordingly. Hence quasi-continuous controllers are preferably used where the process is comparatively slow, and are especially popular in temperature control systems.

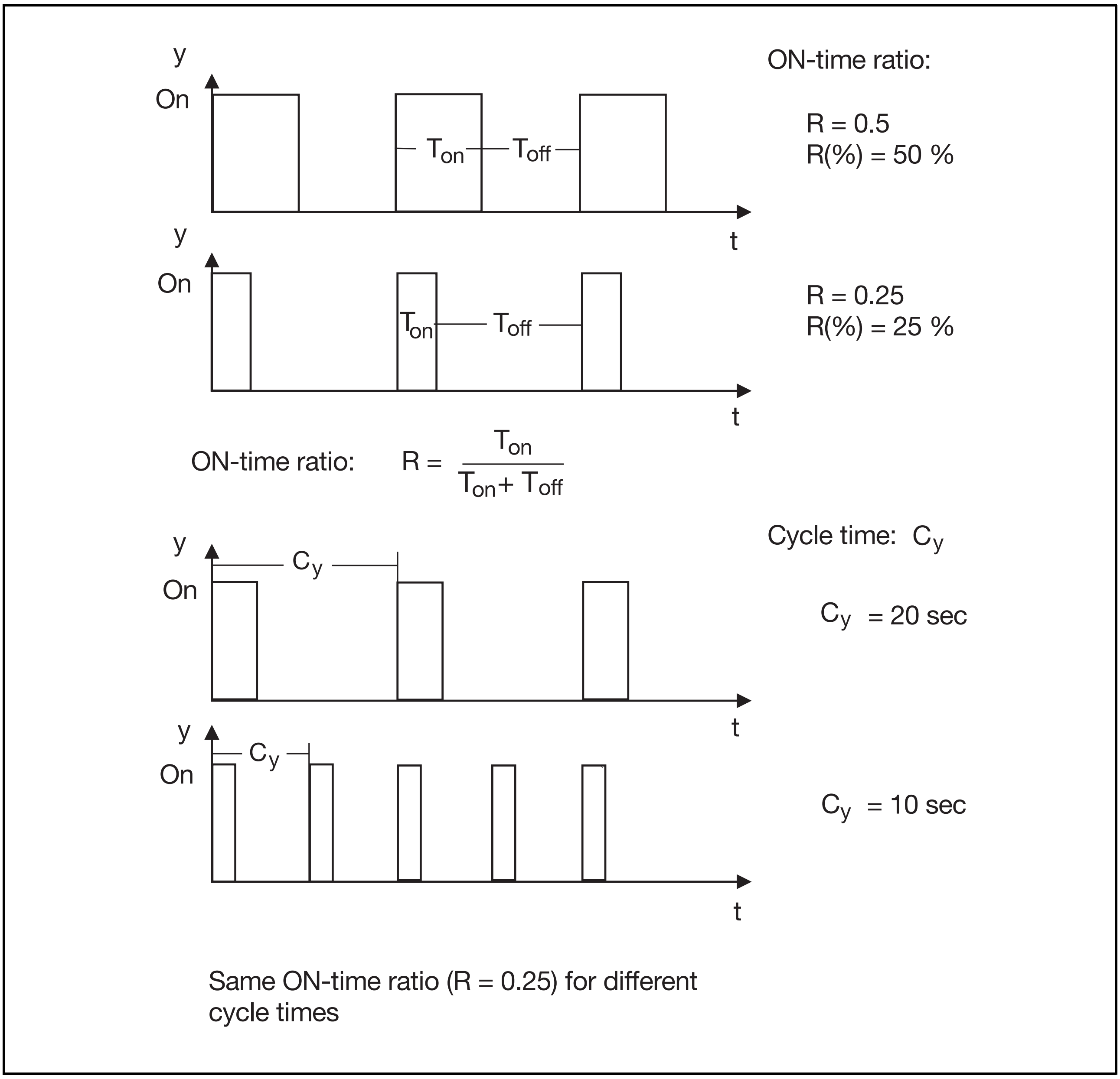


Fig. 57: Power control

The definition of ON-time ratio (or duty-cycle) means the ratio of the switch-on time of a controller output to the sum of the switch-on and switch-off times, e.g. an ON-time ratio of 0.25 means that the power supply is switched on for 25% of the total time. It gives no information on the actual duration of the periods during which the switching cycles take place.

For this reason, the so-called cycle time (C_y) is defined, which fixes this time period. It represents the period during which switching on and off takes place once, i.e. it is equal to the sum of the switch-on and switch-off times (Fig. 57). The switching frequency is the reciprocal of the cycle time.

Fig. 57 shows the same ON-time ratio (R = 0.25) for different cycle times.

For a given ON-time ratio of 0.25 and a cycle time of C_y = 20 sec, this means that the energy supply is switched on for 5 seconds and switched off for 15 seconds. If the cycle time is 10 sec, the energy supply is switched on for 2.5 seconds and switched off for 7.5 seconds. In both cases, the power supplied is 25 %, but with a finer dosage with C_y = 10 sec. The fluctuations of the process variable are smaller in the second case.

Theoretically, the ON-time of the controller is given by the following relationship:

$$T_{on} = \frac{y \bullet C_y}{100\ \%}$$

T_{on} = ON time
y = manipulating variable in %
C_y = cycle time

This means that a shorter cycle time results in a finer dosage of the energy supply. On the other hand, there is increased switching of the actuating device (relay or contactor). The switching frequency can easily be determined from the cycle time.

Example:
The cycle time of a controller used for temperature control is C_y = 20 seconds. The relay used has a contact life of 1 million switching operations. The value given for C_y results in 3 switching operations per minute, i.e. 180 per hour. For 1 million operations, this gives a life of 5555 hours = 231 days. Based on an operating time of 8 hours per day, this represents approx. 690 days. Assuming around 230 working days per year we arrive at an operating life of approx. 3 years.

Generally, the cycle time is selected so that the control process is able to smooth out the energy bursts supplied, to eliminate periodic fluctuations of the process variable as far as possible. At the same time, the number of switching operations must always be taken into account. With a microprocessor controller however, the value set for the cycle time C_y is not held constant over the whole of its working range. A detailed discussion of this point is rather complicated and would be too advanced at this stage. If it is possible to operate a switching P controller in manual mode, the influence on C_y can be observed by direct input of a manipulating variable.

When C_y is matched to the dynamic action of the process, the behavior of a quasi-continuous controller (as a proportional controller with dynamic action) can definitely be comparable with that of a continuous controller, which also explains its name. With quasi-continuous controllers the different manipulating variables are the result of a variation of the ON-time ratio, but there is no discernible difference in the course of the process variable when compared to that of a continuous controller.

5.4 Quasi-continuous controllers: the controller with dynamic action

A quasi-continuous controller, operated as a pure proportional controller and with C_y suitably matched, shows almost the same behavior in a process as does a continuous controller with P action. Although it reacts very quickly to changes in the control deviation, it cannot reduce the control deviation to zero, which is also the case with a proportional controller. A quasi-continuous controller can also be configured as a PID controller, which means that it slows down as the setpoint is approached and stabilizes accurately at the setpoint.

A quasi-continuous controller (and also a P controller) can be pictured as a combination of a continuous controller and a switching stage connected to the output. The continuous controller calculates its manipulating varaible from the course of the process variable deviation and controls the switching stage accordingly. The switching stage calculates the relative ON-time of the switching stage output. The output of the switching stage is pulsed in accordance with the ON-time ratio and the value set for C_y.

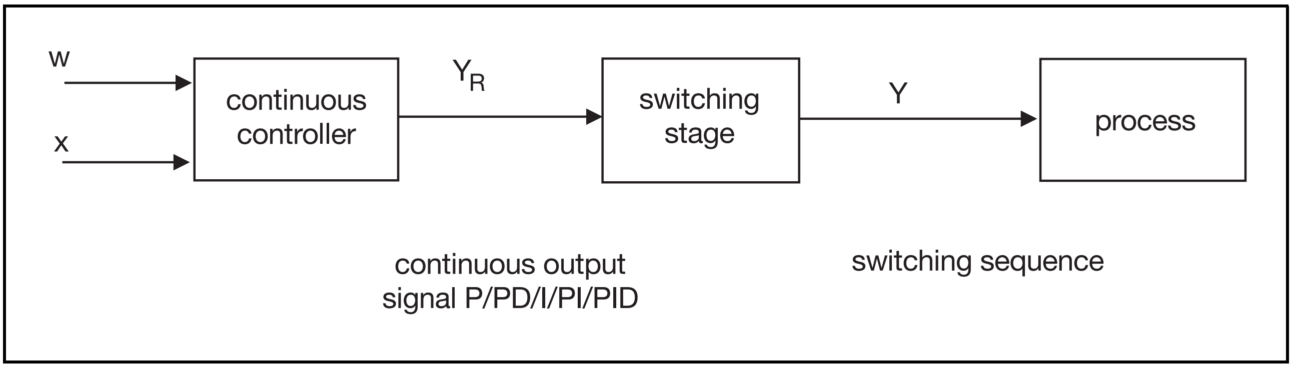


Fig. 58: Quasi-continuous controller with dynamic action, as a continuous controller with a switching stage

Example: The continuous controller produces a manipulating variable of 50%. Likewise, for the switching stage, 50% manipulating variable means an ON-time of 50%. Let us assume that the value set for C_y is 10 seconds, then the switching stage will turn the input on and off every 5 seconds.

5.4.1 Special features of the switching stages

As already described, the switching stage calculates the relative ON-time of the switching stage output on the basis of the manipulating variable of the continuous controller. The output of the switching stage is pulsed in accordance with the relative ON-time and the value set for C_y.

Using analog technology, an additional D action stems from the technical implementation of the switching stage. If the continuous controller preceding the switching stage is operated as a proportional controller, then combining it with the switching stage results in a PD action. This means that if the analog controller referred to has a pure P action (i.e. only a proportional band X_P can be set), then the quasi-continuous controller, obtained by connecting a switching stage to its output, has an additional D action which is not adjustable. As this characteristic proved to be very useful, it has also been retained in the microprocessor controllers. The value of the derivative component is normally set by the manufacturer and therefore cannot be adjusted.

The following table shows the control configurations of a quasi-continuous controller and the actual control actions:

Set control action	Actual control action
P	PD
PD	PDD
I	PI
PI	PID
PID	PD/PID

Table 8: Control configurations of a quasi-continuous controller, and actual control actions

Example: A quasi-continuous controller configured as a proportional controller actually has PD action, because of the switching stage connected to its output.
If the continuous controller has a PID structure, the existing switching controller has a P, an I and two D components. PD/PID controllers of this type are widely used with switched temperature controllers, as they offer the best start-up action for this application.

Depending on the controller structure, various setting parameters are obtained for the quasi-continuous controller:

PD	PDD	PI	PID	PD/PID
X_P	X_P	-	X_P	X_P
C_y	C_y	C_y	C_y	C_y
-	-	T_n	T_n	T_n
-	T_d	-	-	T_d

Table 9: Setting parameters for differing dynamic actions

5.4.2 Comments on discontinuous and quasi-continuous controllers with one output

In practice, discontinuous and quasi-continuous controllers are often brought together under the concept of the 2-state controller, on the basis that the output of the controller can only assume two conditions, either on or off. These controllers are configured at the instrument by defining the control type, namely 2-state controller. If the proportional band X_P is now set to 0, we have a discontinuous controller. If a proportional band X_P greater than 0 is selected, a quasi-continuous controller is obtained, on which the appropriate control parameters (T_n, T_d and C_y) can be set.

5.5 Controller with two outputs: the 3-state controller

5.5.1 Discontinuous controller with two outputs

A discontinuous controller with two outputs can be thought of in simple terms as a combination of two discontinuous controllers, each with one output, but linked together, one below the other. They can be used for heating, for example, when below the setpoint and for cooling when above the setpoint. Other applications are, for instance, humidification and dehumidification in a climatic cabinet, etc. Each output of the controller is assigned a manipulating variable. E.g. the first controller output would often be used for heating and the second output for cooling. All parameters associated with the “heating controller” are identified by an Index_1 and those associated with the “cooling controller” by Index_2.

If the process variable varies within a fixed interval about the setpoint – the contact spacing X_{Sh} – then neither output is active (Fig. 59). This contact spacing is necessary to prevent continual switching between the two manipulating variables e.g. heating and cooling registers, when the process variable is unsteady.

As well as the contact spacing, discontinuous controllers with two outputs also have a hysteresis for each of the heating and cooling contacts, which are normally indicated by the switching differentials X_{Sd1} and X_{Sd2}. These two parameters eliminate any contact “chatter”, when the process variable moves from the heating zone or cooling zone into the contact spacing.

With regard to the switching differentials X_{Sd1} and X_{Sd2}, and the related switching frequency and control quality in connection with the process characteristics, the same considerations apply as for a discontinuous controller with only one output.

With this controller, the control accuracy which can be achieved is limited by the switching hysteresis values and the contact spacing (Fig. 59).

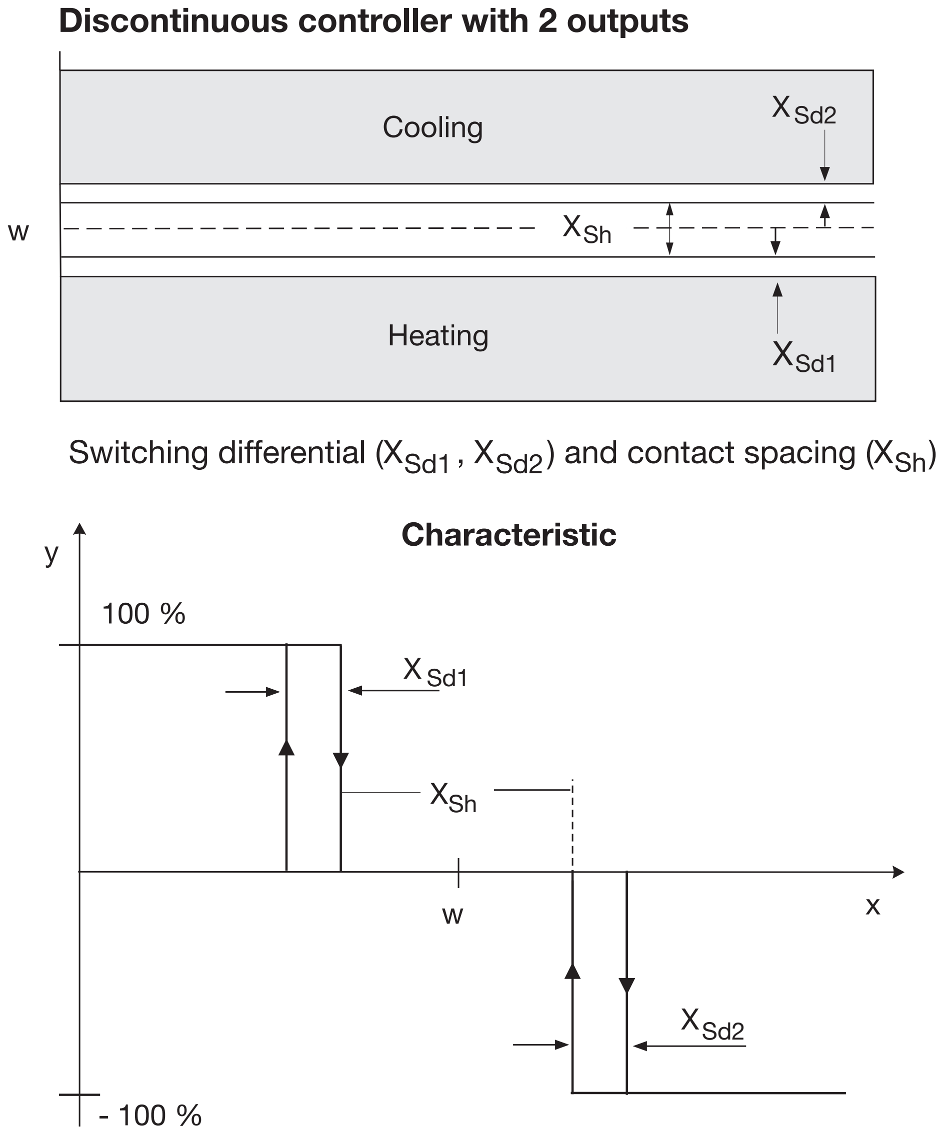


Fig. 59: Characteristic of a discontinuous controller with two outputs

5.5.2 Quasi-continuous controller with two outputs, as a proportional controller

Even a quasi-continuous 3-state controller, where each of the two outputs is operated by a proportional controller, can be thought of in simple terms as a combination of two quasi-continuous controllers linked together, one below the other. The switching differential does not apply here, but a contact spacing can be set.

Fig. 60 shows the characteristic of a quasi-continuous 3-state controller used to control an climatic cabinet.

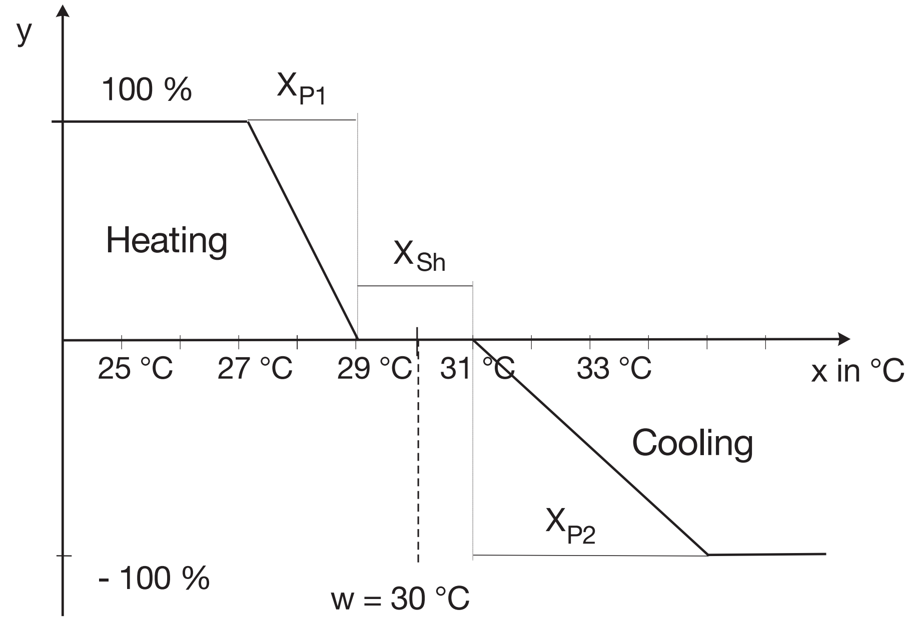


Fig. 60: Characteristic of a quasi-continuous controller with two outputs, as a proportional controller

As shown in Fig. 60, the two proportional bands X_{P1} and X_{P2} can be adjusted independently for a quasi-continuous controller with two outputs. This is necessary, because in general the process gain is different for the two manipulating variables, as the heating register influences the process differently from the cooling (through a fan, for example).

The way in which this controller works is described below: The process value in the cabinet is 26°C. Now the control is switched on:

Heating: The heating relay is energized and the heating heats up with 100% manipulating variable, whereupon the process value increases. The heating manipulating variable continually reduces from a process value above 27°C (on reaching the proportional band), the heating relay starts to pulse and the switch-on times become progressively shorter. The control deviation and hence the manipulating variable become smaller, until a manipulating variable is obtained which is just sufficient to maintain the process value. Now the process value is below the contact spacing, and we obtain a positive manipulating variable (for example 28.5°C and 25% manipulating variable).

Cooling: Now the ambient temperature increases (disturbance), whereupon the inner chamber of the cabinet is heated. The process value increases – on entering the contact spacing (29°C) the manipulating variable is 0%, and there is neither heating nor cooling. The cooling relay only starts to pulse above a temperature of 31°C (the manipulating variable becomes negative). Likewise, the control deviation reaches a value such that the manipulating variable produced is just sufficient to maintain the resulting process value.

5.5.3 Quasi-continuous controller with two outputs and dynamic action

With a quasi-continuous controller with two outputs and dynamic action, it is additionally possible to set a T_n and a T_d for each of the two manipulating variables. With this controller, the three control components (P, I and D) are only all effective together outside the contact spacing. If a process enters the contact spacing, the P component is not effective. In the contact spacing, only the I and D components are effective, and in principle the process variable can therefore be stabilized exactly.

Because there is no P component in the contact spacing, a large contact spacing has an adverse effect on the dynamic response, as the controller is slow in the contact spacing. The size of the contact spacing chosen should be no larger than necessary, but no unwanted changeover between the manipulating variables should occur, either when approaching the setpoint (possible overswing of the process variable above the setpoint) or when stabilizing at the setpoint (fluctuations of the process variable about the setpoint).

Too small a contact spacing can lead to a pointless waste of energy in an installation

Table 10 shows the setting parameters of a quasi-continuous controller with two outputs and dynamic action:

	set	effective						
Discontinuous			$X_{P1} = 0$ $X_{P2} = 0$	–	–	–	X_{Sh}	X_{d1}; X_{d2}
Quasi-continuous	P	PD	X_{P1}; X_{P2}	–	–	C_{y1}; C_{y2}	X_{Sh}	–
	PI	PID	X_{P1}; X_{P2}	T_{n1}; T_{n2}	–	C_{y1}; C_{y2}	X_{Sh}	–
	PID	PD/PID	X_{P1}; X_{P2}	T_{n1}; T_{n2}	T_{d1}; T_{d2}	C_{y1}; C_{y2}	X_{Sh}	–
	PD	PDD	X_{P1}; X_{P2}	–	T_{d1}; T_{d2}	C_{y1}; C_{y2}	X_{Sh}	–
	I	PI	–	T_{n1}; T_{n2}	–	C_{y1}; C_{y2}	X_{Sh}	–

Table 10: Setting parameters for a quasi-continuous controller, with two outputs and dynamic action

5.5.4 Comments on controllers with two outputs

In Chapter 5.5 we met controllers where the two outputs could influence a process variable in two directions. The controllers described always had two outputs of the same type (discontinuous or quasi-continuous). In practice, it may turn out that the outputs are different. It may well be, for example, that a controller has to provide a discontinuous output for cooling and a quasi-continuous output for heating. Such a controller cannot be classified under any of Chapters 5.5.1, 5.5.2 or 5.5.3.

To limit the number of names, in practice all controllers with two outputs which can influence a process variable in two directions are referred to as...

. . . 3-state controllers

irrespective of whether the outputs are discontinuous, quasi-continuous or continuous. As an example, mention should be made of a controller which operates a thyristor-controlled power unit and a refrigerator unit and thus maintains constant temperature in a climatic cabinet. The two plants require two controller outputs – but the controller must provide a continuous output for heating and a switched output for cooling.

5.6 The modulating controller

Actuators (actuator drives) have three operating conditions: opening, holding, closing.

Electrical drives in particular are widely used for this application, where a motor controlled for clockwise or anticlockwise rotation drives a worm gear to operate a valve, throttle, variable transformer or similar device. Both DC and 3-phase motors are used, with single phase motors used for smaller drives, switched by contactors or relays (see Chapter 1.7).

These drives stand out from the controller applications already discussed in one particular way. When the heating is switched on in a furnace, it operates immediately at full power, and when it is switched off, the supply of power stops immediately. In contrast, actuators require a certain time to reach the maximum manipulating variable (valve opening etc.). In addition, an electrical actuator holds the position it has reached, even when there is no signal from the controller. The actuator can, for example, stay at 60 % open, even though it is not operated by the controller at this time. The controller must take these properties into account.

Modulating controllers are used for this type of actuator drive.

The modulating controller consists of a continuous controller (P or PD) and a switching element. If we regard the valve position as the manipulating variable, the combination of modulating controller and regulating valve exhibit PI or PID action.

To understand the operation of the modulating controller take a look at Fig. 61:

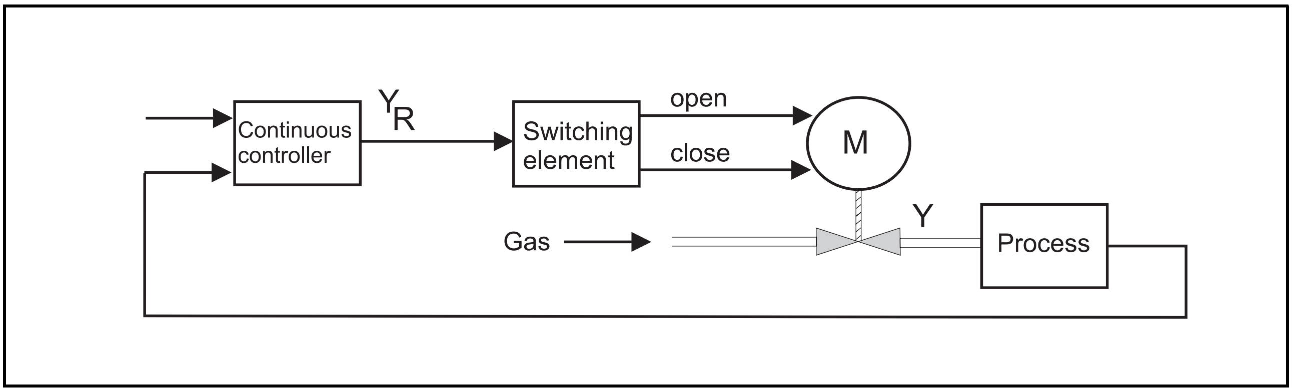


Fig. 61: The modulating controller, with a regulating valve in the control loop

The modulating controller shown controls the temperature in a furnace via a regulating valve in the gas flow. The switching stage provides two relay outputs which drive the valve open and closed over the range 0 to 100%.

The controller, the switching element and the regulating valve must now be thought of as a single unit. The modulating controller (meaning here the continuous controller and the switching element) can be configured for PI or PID action. If a control deviation occurs, the valve will exhibit the corresponding PI or PID action. If then, for example, PI action is set on the controller, then the combined modulating controller and regulating valve (with the valve opening as manipulating variable) will have PI action.

Example: The modulating controller of Fig. 61 was configured as a PI controller. The proportional band X_P was set to 25°C, and the reset time T_n to 120 seconds. The associated regulating valve has an actuator stroke time T_y (the time required by the actuator to travel from 0 to 100% or from 100 to 0% manipulating variable) of 60 seconds. Fig. 63 shows the step response of the system.

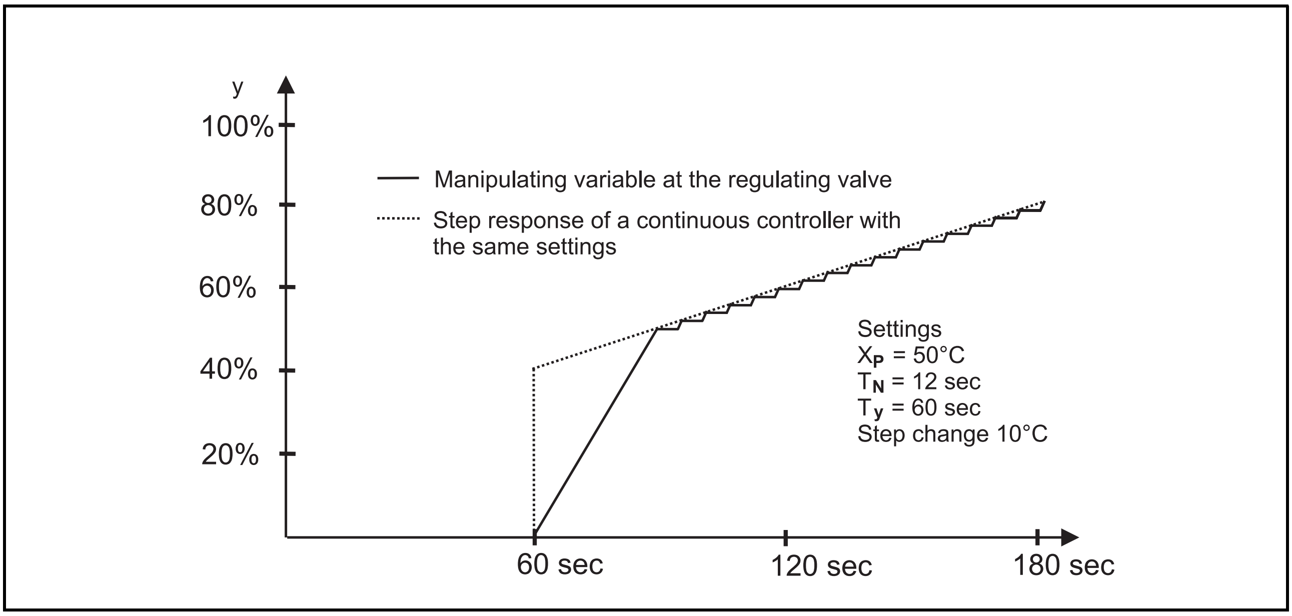


Fig. 62: Step response of the modulating controller and regulating valve system

Fig. 62 shows the step response of a PI controller with the parameters X_P = 25°C and T_n = 120sec. The step change in control deviation occurs at t = 60 sec and amounts to 10°C. The modulating controller has the following settings: X_P = 25°C, T_n = 120sec, and the actuator stroke time T_y = 60seconds. By operating via the “Open” output of the modulating controller, the PI action of the combined modulating controller and regulating valve will be implemented. The opening of the valve will, of course, lag behind the manipulating variable of a PI controller, as it has a stroke time of 60 seconds.

The modulating controller receives no indication of the exact position of the valve. It assumes that the valve opens and closes at exactly the same speed. The modulating controller calculates the time for which the “Open” contact must be closed, until, in theory, the valve position corresponds to the manipulating variable of the corresponding PI controller. For this to work, the modulating controller must have knowledge of the actuator stroke time.

The modulating controller also has its contact spacing set so as to lie symmetrically about the setpoint. Within the contact spacing, no control operation occurs on the actuator, which means that if the process variable enters the contact spacing, the valve will remain in its old position.

With modulating controllers, a minimum pulse duration T_{Mmin} can be taken into account. This may be necessary because of minimum switch-on times of the actuator drive (e.g. play in the gears). With a microprocessor controller, however, it is at least the sampling time or cycle time of the controller. The minimum pulse duration T_{Mmin} can be set directly on many controllers.

The minimum pulse duration has a direct influence on the positioning accuracy of the actuator, and consequently on the expected control accuracy.

The following relationship generally applies for linear processes:

$$\Delta x = X_{Max} \bullet \frac{T_{Mmin}}{T_y}$$

Δx : control accuracy
X_{Max} : maximum process value
T_{Mmin} : minimum pulse length
T_y : actuator stroke time

It is important that the contact spacing X_{Sh} of a modulating controller is not set smaller than the control accuracy Δx, calculated from the minimum pulse duration. Choosing a contact spacing smaller than this value will result in permanent fluctuations of the process variable as the actuator continually changes over from clockwise to anticlockwise rotation, making excessive demands on the actuator.

If the correct contact spacing is chosen, the true control deviation will be smaller than the set contact spacing, because the final pulse runs the actuator into the contact spacing and thereby reduces the control deviation.

As the actuator drive has the same characteristic for clockwise and anticlockwise rotation, there is only one setting each for X_P, T_n and T_d. The setting parameters are then as follows:

Dynamic action	**PI**	**PID**
Setting parameters	X_P	X_P
	T_n	T_n
	-	T_d
	T_y	T_y
	X_{Sh}	X_{Sh}

Table 11: Setting parameters with the modulating controller

5.7 Continuous controller with integral motor actuator driver

A “continuous controller with integral motor actuator driver” or, for short, an actuating controller, is much more suitable for operating a motorized actuator than is a modulating controller. It forms a type of cascade structure (see Chapter 6.6), and consists of a continuous controller and a subordinate actuating controller (Fig. 63).

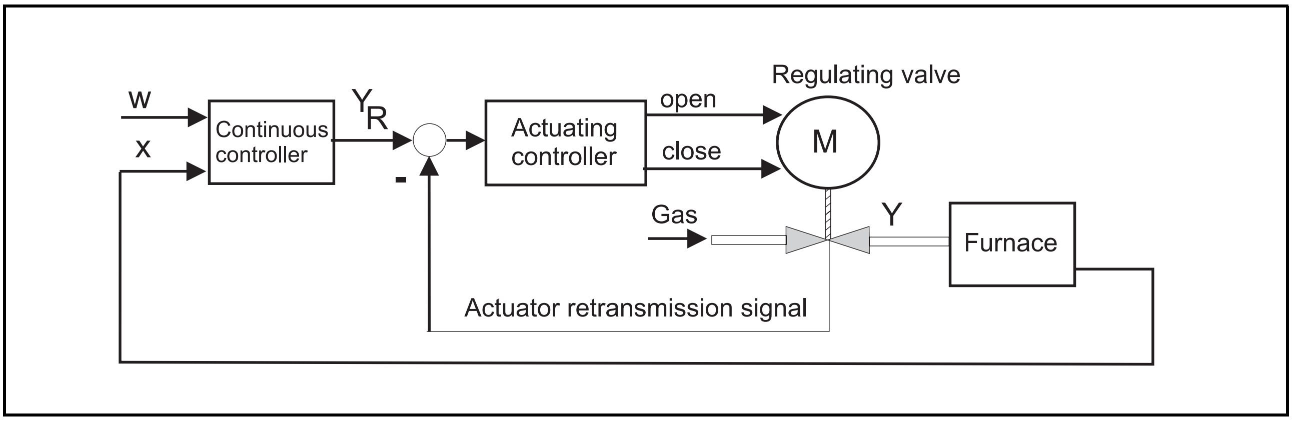


Fig. 63: The actuating controller with a regulating valve in the control loop

The continuous controller outputs the manipulating variable, based on the course of the control deviation and the parameters set on the controller (Fig. 63). The usual control structures, i.e. P, PI, I, PD, PID can be set for the continuous controller. The duty of the actuating controller is now to regulate this manipulating variable on the regulating valve. The actuating controller operates the actuator via two switching outputs, and receives an actuator retransmission signal (usually a standard signal 0/4 — 20 mA, 0/2 — 10V etc.), which feeds the actuator position back to the controller.

Example: The continuous controller determines a manipulating variable of 20% from the course of the control deviation. The actuating controller now controls the valve at 20% opening. The valve provides a 0 — 10V actuator retransmission signal that corresponds to 0 — 100% opening of the valve. If the actuating controller has controlled the valve to 20% opening, the actuator retransmission signal would be 2V.

An actuator stroke time must also be fed into the actuating controller, which the controller then uses to optimize its control parameters.

Where a motor is being operated which has an appreciable overrun (poor braking action), juddering of the actuator motor can be avoided by increasing the contact spacing (X_{Sh}).

Advantages of the actuating controller in comparison with the modulating controller:
Unlike the modulating controller, the actuating controller offers the advantage of a subordinate controller structure. If a control deviation occurs, the actuating controller ensures that the motor is driven directly to a new position. This is achieved by comparing the actuator position with the manipulating variable (y_R) of the continuous controller.

A modulating controller does not receive an actuator retransmission signal and must always assume a linear actuator action. If the actuator has a non-linearity, or play is present in the actuator mechanism, this assumption will only be an approximation.

An actuating controller offers the following setting parameters for the corresponding control action:

Dynamic action	PD	PDD	PI	PID	PD/PID
Setting parameters	X_P	X_P	-	X_P	X_P
	-	-	T_n	T_n	T_n
	-	T_d	-	-	T_d
	T_y	T_y	T_y	T_y	T_y
	X_{Sh}	X_{Sh}	X_{Sh}	X_{Sh}	X_{Sh}

Table 12: Setting parameters with the actuating controller

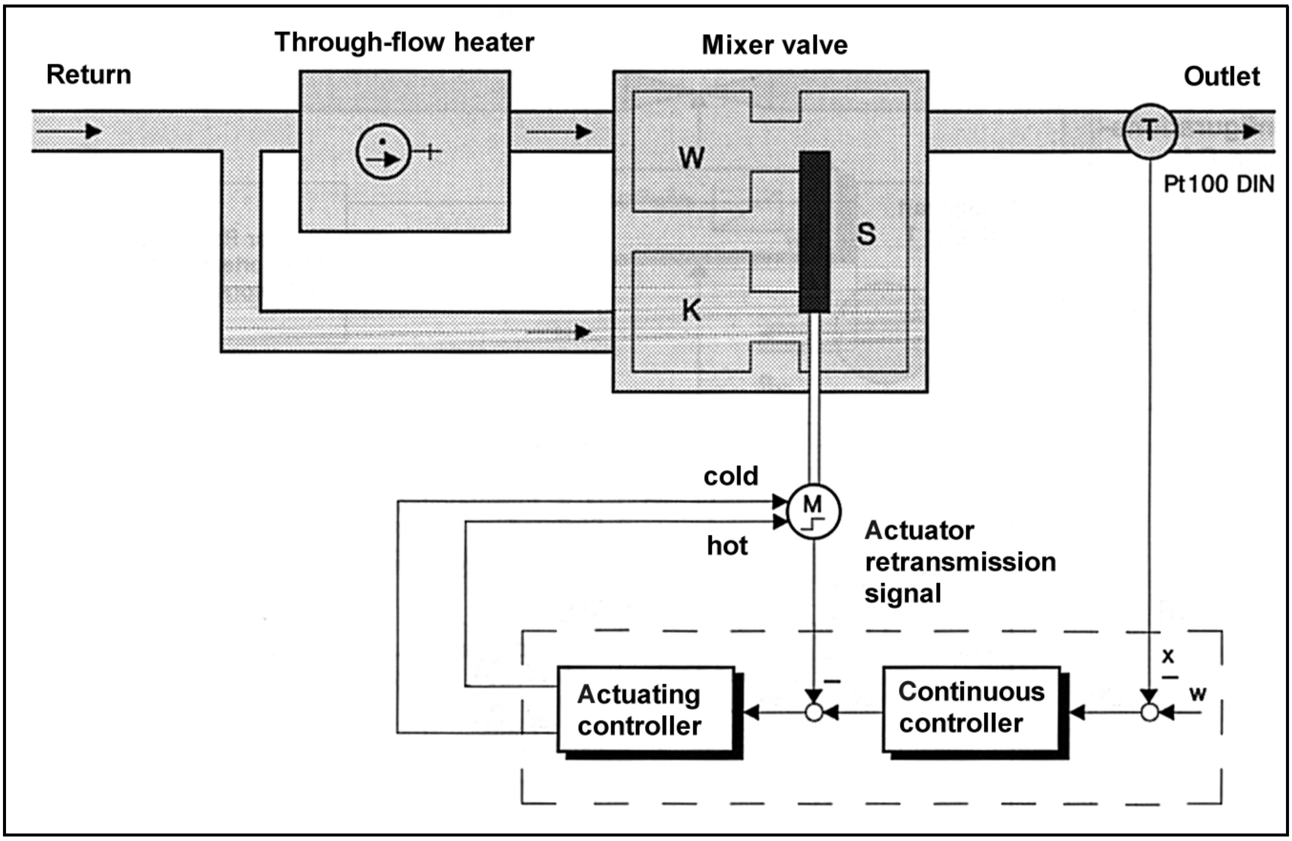


Fig. 64: Example of an application for an actuating controller

Example:
The actuating controller described is used to control the outflow temperature of a heating system. The main element is a mixer valve whose chambers “C” and “W” for cold and warm water are linked through piping to the water return pipe. The mix temperature is measured by a Pt100 and can be varied by adjusting the position of the slider “S”, which is operated by an actuator motor. The input variable of the actuator motor is in the form of switching pulses for opening and closing the outflow opening.

6 Improved control quality through special controls

So far, we have only considered single-loop control circuits, where controller and process form a closed signal loop. However, when using such single-loop control circuits, there are limits to the control quality which can be achieved in certain control processes. It is possible to go beyond the control quality limits imposed by the single-loop control circuit by using multi-loop control circuits, or by switching auxiliary variables on and off. To some extent, relatively simple solutions can lead to considerable improvements in control quality.

6.1 Base load

With a base load setting, the controller only influences part of the total manipulating variable, and a fixed proportion is continuously supplied to the process (combination of control and operation). It could then be the case that, for example in an electrically heated furnace, one section of the heating elements is controlled by the controller, whereas another section is supplied at full supply voltage (see Fig. 65).

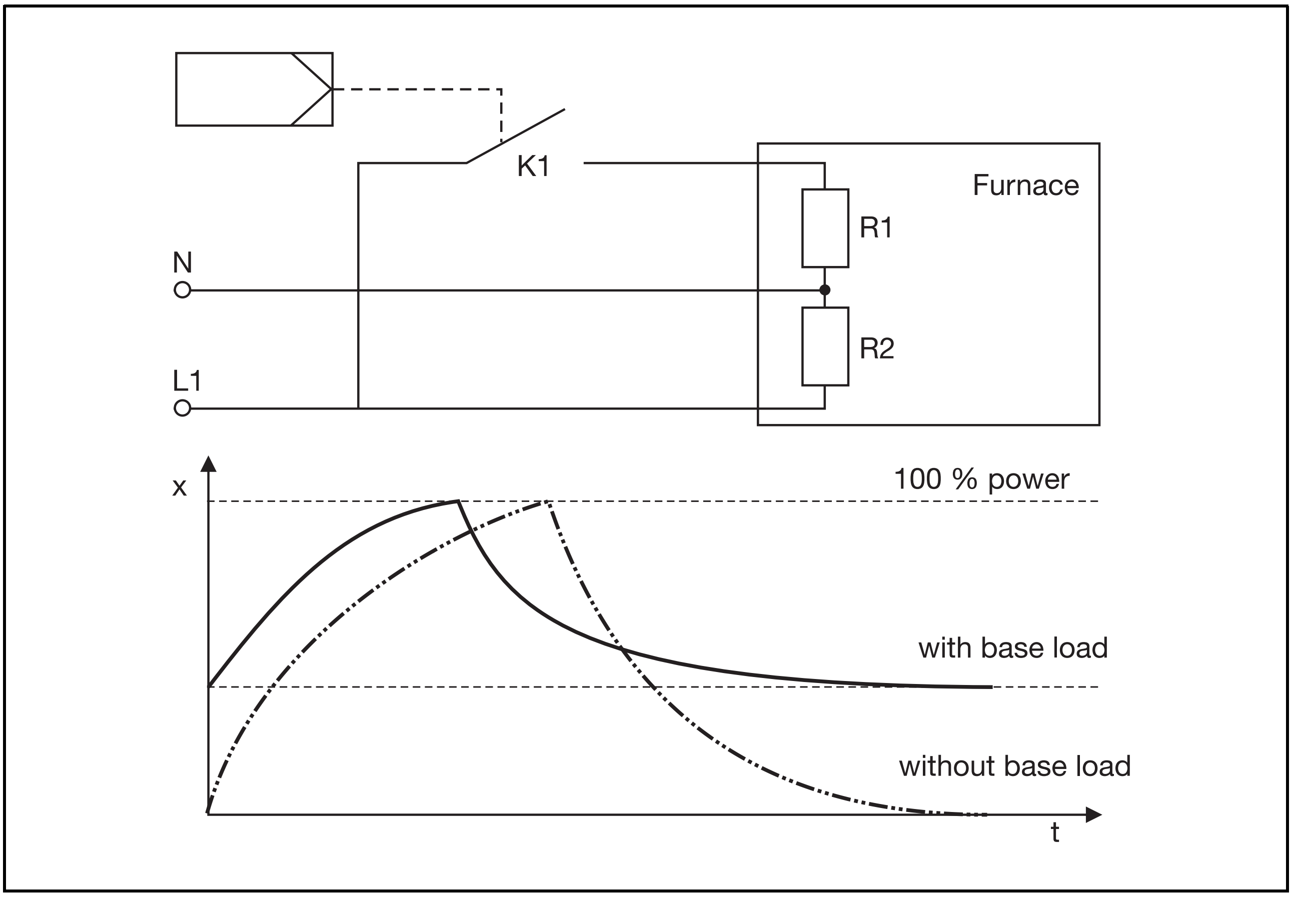


Fig. 65: Base load setting

Essentially, base load setting offers the following advantages:

- The actuator, e.g. a thyristor controller, can be more compact and less expensive, as it only needs to control low power.
- Load fluctuations on the supply network, as a result of power consumption in bursts caused by the switching controller, are similarly reduced.
- If the controller fails, the process can still be operated with the base load component of the total power.

Against the advantages shown, there are also a number of disadvantages:

- The dynamic control action is impaired, especially with regard to disturbances. As the controller now no longer provides the full manipulating variable, the cooling curve, for instance, is not only shifted by the amount of the base load setting, but is also clearly flatter (see Fig. 65). If, for any reason, the power requirement suddenly becomes less than the base load setting, the controller is helpless in this situation, as it cannot reduce the manipulating variable below the value of the base load.
- In addition, the base load setting must also be matched to the setpoint. If the setpoint is changed downwards, for instance, the excess power could suddenly be too large; with an upward change, the excess power could be too small. In such cases, the base load should be changed at the same time as the setpoint.

6.2 Power switching

If a process is being operated with different setpoints or working points, it is better to switch the applied power rather than use a base load. In an electrically heated furnace, this can be achieved by switching part of the heating elements through a limit value switch (limit comparator as an advance contact), in order to facilitate furnace operation at full power in the upper temperature range (see Fig. 66).

This gives the following advantages:

- At any one time, the process can be operated in the upper third of the characteristic valid at that time (see Fig. 66). In this way, the excess power at small values of the process variable can be minimized.
- The dynamic control action is rather better here when compared with the base load method, as in this case the control power can be reduced to zero (after falling below the changeover point).

There can be a disadvantage with this circuit if it operates with a setpoint close to the changeover point, as the process has two different values of process gain here.

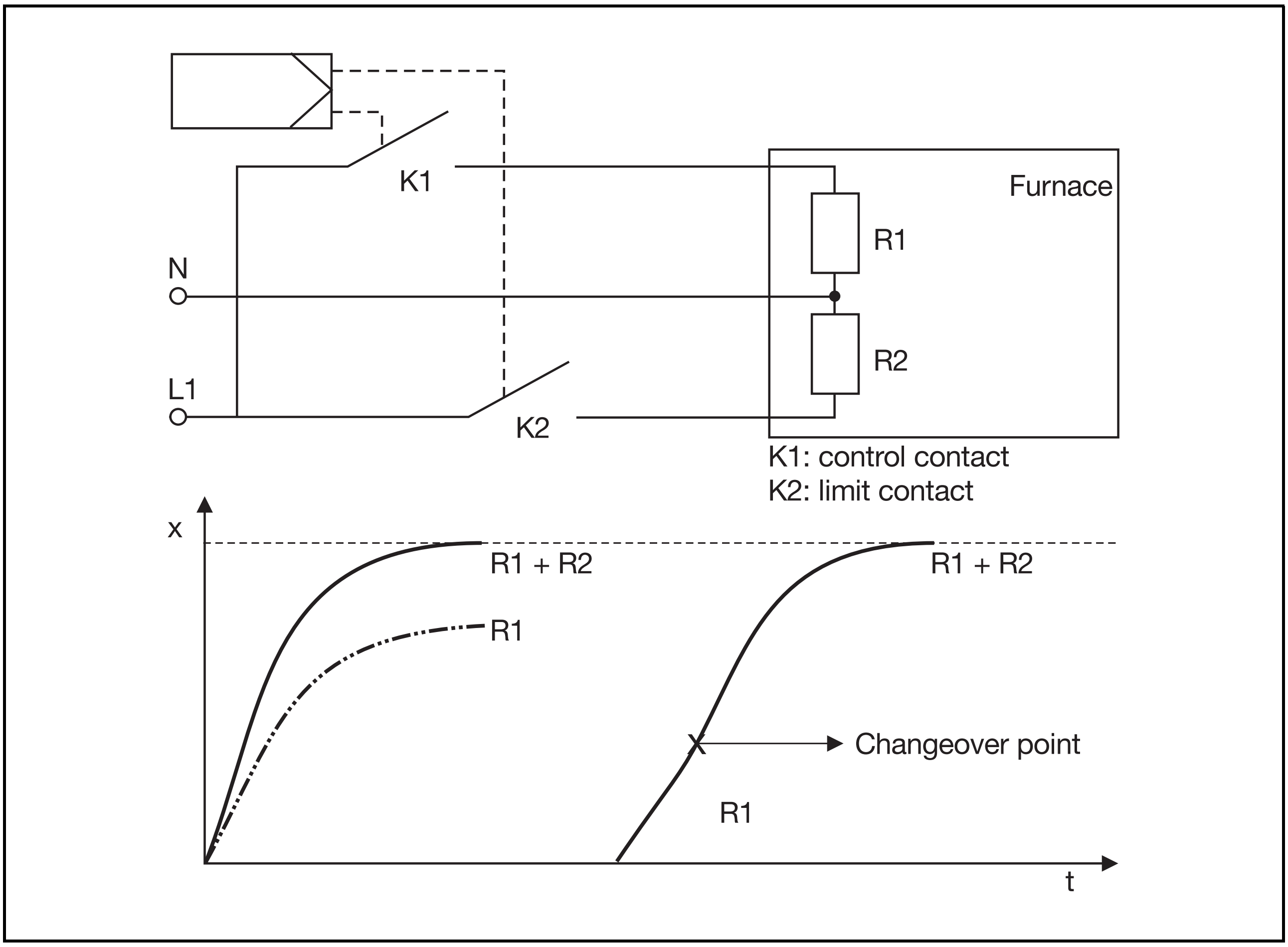


Fig. 66: Power switching

6.3 Switched disturbance correction

The effect of a disturbance can often be predicted within certain limits. For example, opening a furnace door leads to a fall in temperature of 30°C. Instead of first waiting for the process to respond to this disturbance and then for the controller to take corrective action, the disturbance can be responded to directly. To do this, the furnace door is fitted with a position switch that increases the manipulating variable (e.g. heating power) by several percent when the furnace door opens.

This principle is known as switched disturbance correction. It is useful when the cause and effect of a disturbance are known, and where the disturbance occurs frequently and reproducibly. The disturbance is quickly compensated by the rapid response made possible without time delays caused by the controller and process.

We will now look at three different possibilities of switched disturbance correction:

Maintaining the disturbance constant
The effect of the disturbance on the process variable is eliminated by maintaining the disturbance constant by means of an auxiliary control loop (see Fig 67 a). Maintaining the disturbance constant should only be used when suitable technology is available to measure disturbances and maintain them constant.

An example of this is the temperature control of a gas-fired annealing furnace. Here, the main disturbance, gas pressure, can be maintained constant by an in-line pressure controller, which at the same time can also reduce the higher supply pressure to the lower burner pressure. The block diagram of this method can be applied to our own example:
The controller has the job of bringing the process variable x of the process (the temperature of the annealing furnace) to the setpoint w, by giving out the manipulating variable y. If the disturbance z (the gas pressure) is not maintained constant, then, when the gas pressure fluctuates, the controller has to change its output repeatedly, if it is to hold the same setpoint. The auxiliary controller (the pressure controller) now maintains a constant gas pressure, so that this disturbance no longer influences the annealing furnace.

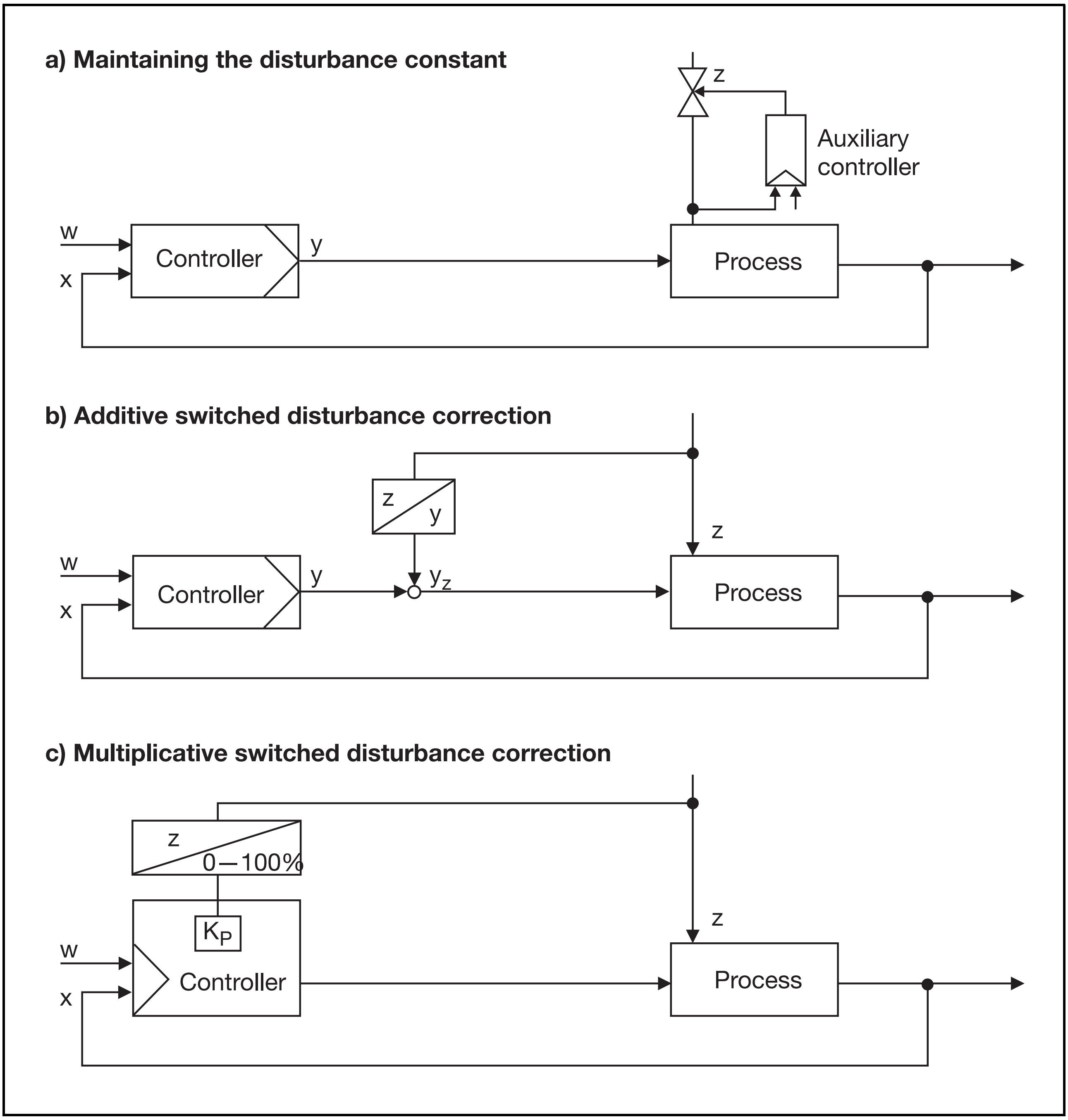


Fig. 67: Forms of switched disturbance correction

Additive/multiplicative switched disturbance correction
With both these methods, when the disturbance changes, the manipulating variable y of the controller is manipulated to counteract the effect of the disturbance (see Fig. 67 b, c).

With **additive switched disturbance correction** (Fig. 67 b) the manipulating variable (y) is increased by an amount proportional to the disturbance. In other words, this type of switched disturbance correction takes into account any offset shifts in the process. Controllers that allow such a switched disturbance correction to be implemented (compact controllers), normally provide an input for the switching signal. A signal proportional to the disturbance is applied to the controller input, which influences the manipulating variable in accordance with the setting. To illustrate this, we can take the example above where the furnace door is opened. When the door is opened, the manipulating variable is increased by a fixed amount.

A multiplicative switched disturbance correction exerts an influence on the controller gain K_P. As the measured disturbance changes its value, so the value of K_P set at the controller is changed in the same ratio, in the range 0 — 100 % (see Fig. 67 c). This method is suitable for use in processes where the manipulating signal (controller output) must be changed to the same extent as any disturbance which may occur.

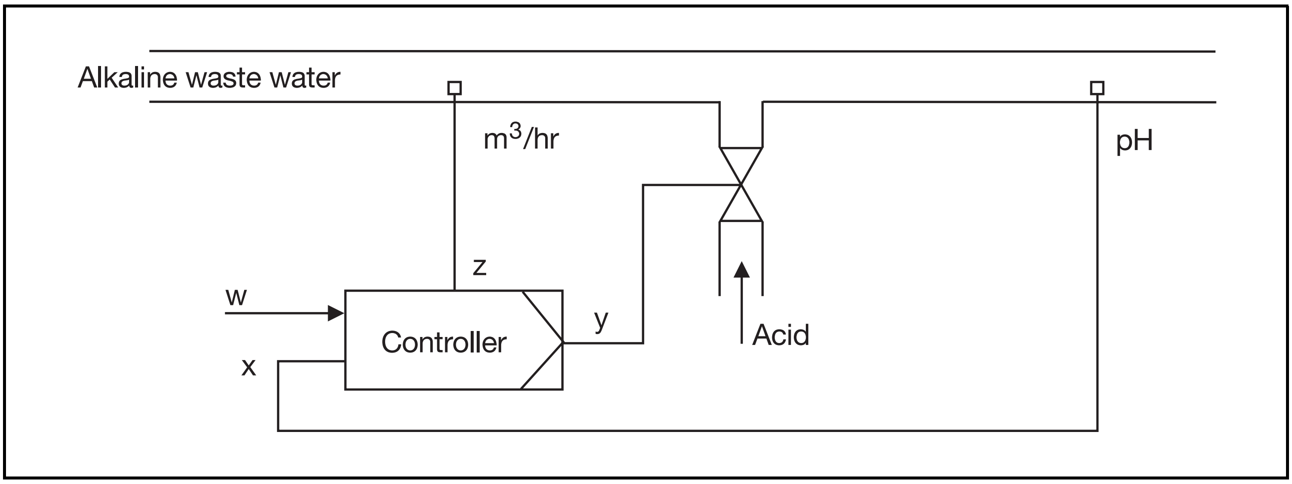


Fig. 68: Neutralization plant

As an example, a neutralization plant could be quoted, in which alkaline waste water is neutralized with acid (see Fig. 68). The process variable is the pH value, which should be in the neutral range. The controller exerts an influence on the pH value by changing the inflow of acid (y). First of all, let us consider how the plant operates without multiplicative switched disturbance correction. Assume that the controller has stabilized at a defined flow rate with, say, 30 % manipulating variable. Now, the disturbance (flow) changes, and the quantity of waste water per unit time is now twice as large. The pH value will now increase, and the controller will increase its manipulating variable until the process variable reaches the setpoint again. This will be the case with 60 % manipulating variable (double the quantity of acid). We can see that the manipulating variable must be kept proportional to the disturbance to maintain the same setpoint, other conditions remaining unchanged. This can be achieved by measuring the disturbance (flow) and applying multiplicative switching. The disturbance is scaled at the controller over the range from zero to the maximum disturbance value which could occur; the controller now changes its proportional action to the same extent, over the range 0 — 100 %.

If we now look at our example again:
Assume here that the controller has stabilized again with, say, 30 % manipulating variable. Now the disturbance (flow) changes to twice the value. Likewise, through the multiplicative switched disturbance correction, the proportional gain (that corresponds to the overall gain, see also Fig. 41) is set to double its value. The manipulating variable of the controller immediately increases to 60 % and there are no larger control deviations.

The examples of switched disturbance correction shown here apply to discontinuous controllers with 2-state action and continuous controllers. The relationships for 3-state, modulating and actuating controllers are more complex, and will not be discussed here.

6.4 Switched auxiliary process variable correction

Where a disturbance cannot be measured or localized, it is possible to derive an auxiliary process variable X_{aux} from the process, where X_{aux} has a shorter time delay than the main process variable x, and apply it to the controller input, after suitable conversion (see Fig. 69). In this way, the disturbances at the process input (e.g. supply disturbances) are quickly reported to the controller.

However, X_{aux} is normally applied through an adaptive timing element, so that the process variable is not distorted under stabilized conditions. With this arrangement, two control loops, each with its own complete signal path, are coupled together. It should be noted that the control loop can possibly become unstable as a result of overly strong switching of the auxiliary process variable and an unsuitable controller setting.

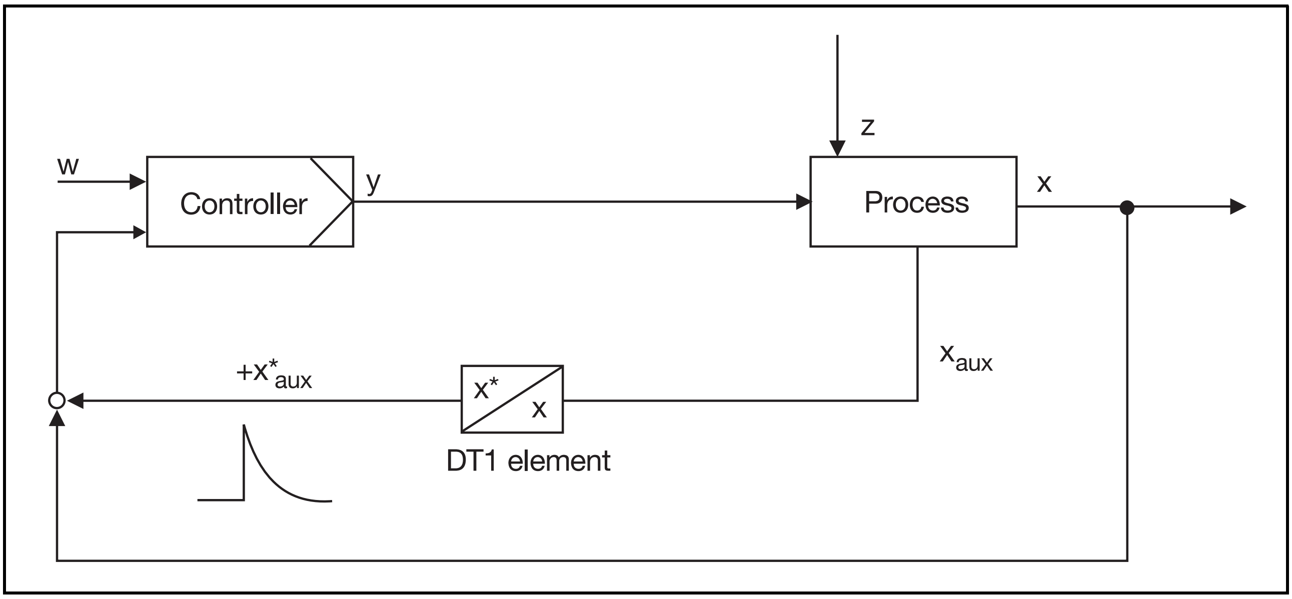


Fig. 69: Switched auxiliary process variable correction

6.5 Coarse/fine control

Two control loops in series are used to maintain some parameter of a mass flow or energy flow constant. The residual deviation from the first controller, the coarse controller (C1), is corrected by the second, fine controller (C2) – see Fig. 70.

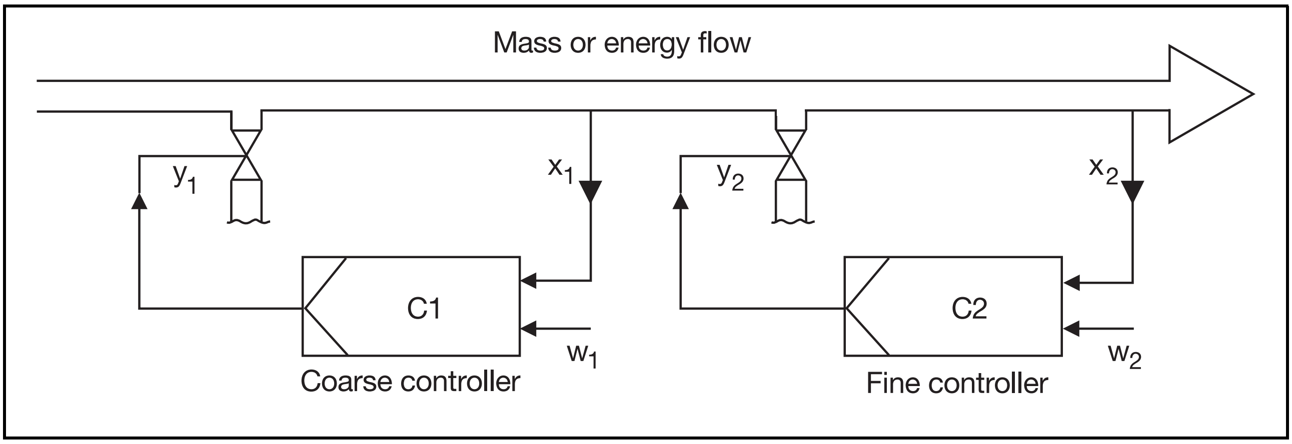


Fig. 70: Coarse/fine control

Here again we can use as an example a pH control system for neutralizing industrial waste water. Because of the large variations in inflow normally present, and the changing composition, it is often appropriate to connect two control loops in series, so that the variations in pH value are maintained within the permissible tolerances.

6.6 Cascade control

Cascade control can significantly improve the control quality. This applies in particular to the dynamic action of the control loop, in other words, the transition of the process variable following setpoint changes or disturbances. Processes with a T_g / T_u ratio less than 2 or 3 can be difficult to control with a simple control system; because of the relatively long delay time, the controller does not become aware of how it should respond until a very late stage. We therefore try to split the control loop into several partial loops (usually two), which are controlled separately. Control of these partial loops is much easier, as each has only a fraction of the overall delay time. This arrangement is also known as multi-loop or networked control.

Fig. 71 shows the block diagram for cascade control.

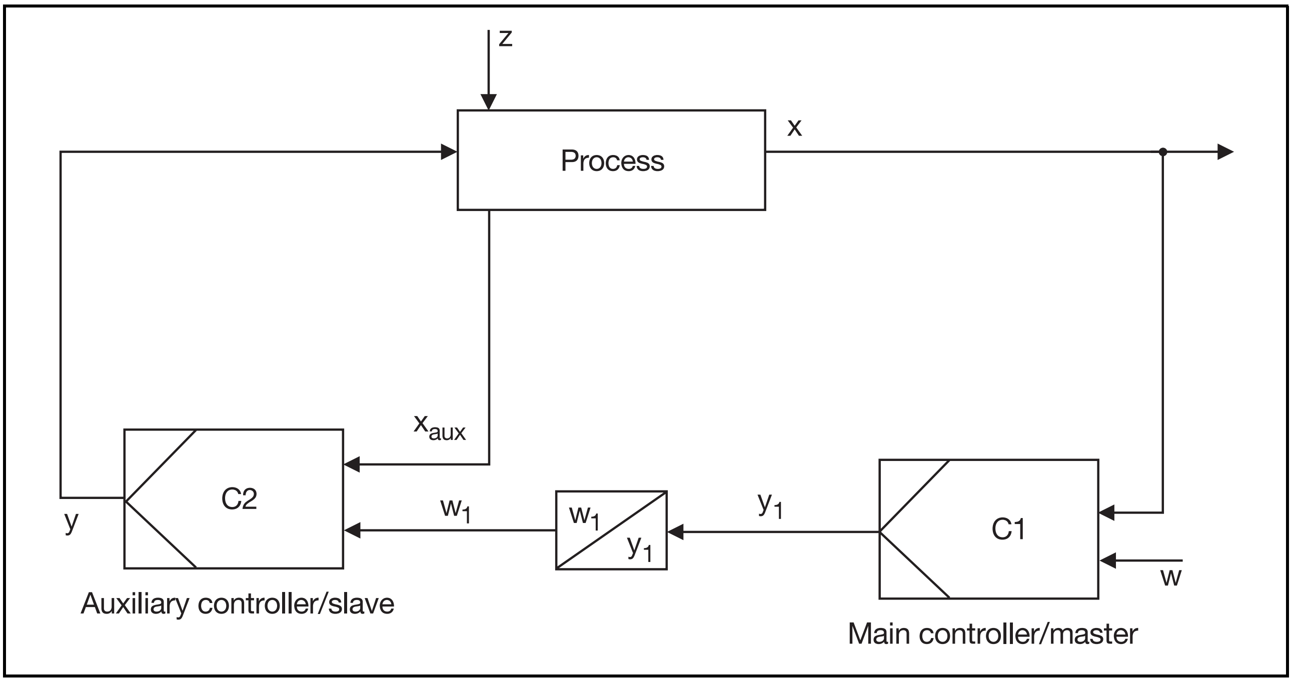


Fig. 71: Cascade control

An auxiliary process variable x_{aux} is derived from the process and applied to the input of an auxiliary controller, the output of which controls the manipulating variable y. The setpoint w_1 of the auxiliary controller is determined by the manipulating variable of the main controller, such that the process variable reaches the set value. The auxiliary control loop can be set to respond more rapidly, and quickly eliminates all disturbances at the input to the process.

The subordinate auxiliary controller is constructed in the same way as an ordinary controller. However, it must have an input for an electrical setpoint signal, as its setpoint is set by the supervisory controller. In other respects, it must be matched to the demands of its duty, with regard to input, output etc. The auxiliary controller has the job of changing the auxiliary process variable very quickly, in proportion to the manipulating variable of the main controller; hence P or PD controllers are normally used for this application, or also, less frequently, a PI controller. The master controller, set for setpoint response, is usually a PI or PID controller.

For cascade control, it is important that the subordinate loop is at least 2 — 3 times faster than the outer loop, as otherwise the overall control loop will tend to oscillate. One advantage of cascade control is that the dynamic response of the control loop is much improved. Another advantage is that the controllers are much easier to adjust. The master controller is switched to manual mode, and the slave controller is optimized. Then the master controller is optimized, with the slave controller kept in automatic mode.

An example of cascade control is the temperature control of a furnace heated by a gas burner (see Fig. 72).

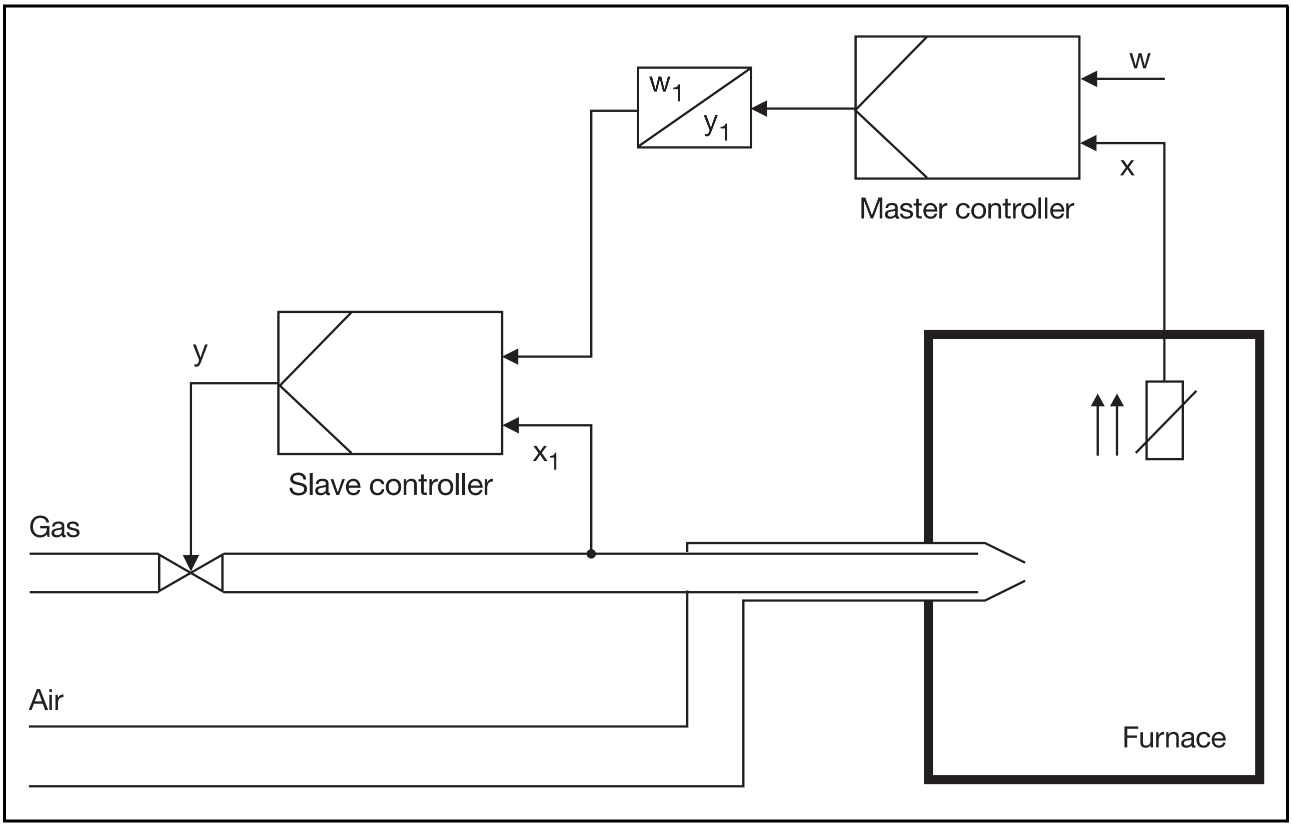


Fig. 72: Cascade temperature control for a burner

The master controller outputs a manipulating variable y_1 in the range 0 — 100 %, on the basis of the control difference applied to it. The slave controller now receives this manipulating variable as its setpoint, but only after the signal is normalized: on the basis of the normalization, the setpoint of the slave controller (w_1) amounts to 0 — “maximum gas flow”, corresponding to 0 — 100 % manipulating variable of the master controller. With its manipulating variable, the master controller practically presets the desired gas quantity per unit time. The slave controller has the job of controlling the gas flow accurately. The slave controller now takes over part of the timing elements and corrects disturbances at the input to the process, for example, fluctuations in the gas pressure. The control action is improved on this basis, and, in certain cases, processes can only be controlled by introducing cascade control.

6.7 Ratio control

Ratio controllers are used in burner controls (control of the gas/air mixture ratio), analytical techniques (mixing of reagents) and in process engineering (preparation of mixtures). These controllers have two process value inputs. The ratio of the two input variables is the real process variable. The value required for this ratio is set as the setpoint, directly at the controller.

A ratio controller is frequently used as a slave controller. Here, the controller has the task of controlling the quantities of two substances in such a way that the mixing ratio stays constant when differing total quantities of the mixture are required. With this kind of slave control, there are two setpoints: the mixing ratio and the total quantity. Accordingly, two controllers are used, one of which controls the total quantity of the mixture per unit time, whilst the other influences the mixing ratio, by adjusting the dosage of the separate components. As the total quantity per unit time is the ultimately decisive setpoint, this controller is designated as master controller, whilst the subordinate controller controls the substance mixing ratio to meet the requirements of the master controller.

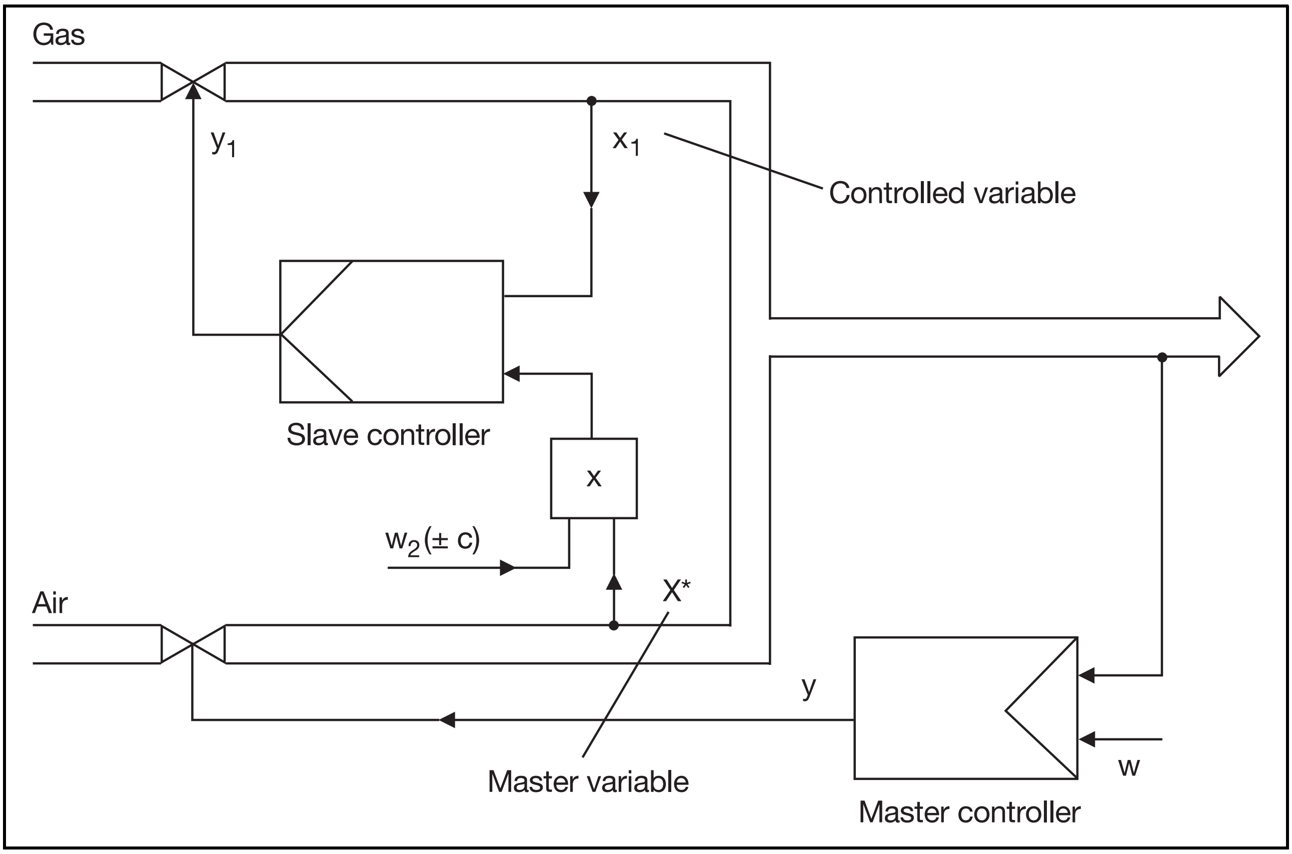


Fig. 73: Ratio control

An example of this is the mixture control shown in Fig. 73: two substances have to be mixed in a fixed ratio to each other, whilst the demand for the quantity of the mixture fluctuates according to production requirements. Two control circuits are required for this, one to control the total quantity of both substances after mixing, the other to control the mixing ratio. In controlling the total quantity, it is sufficient to influence only one component, since the other is made to follow according to the set ratio. However, the mixing ratio is controlled independently of the master controller, so that the master controller and its associated valve have been fitted purely to control the air flow and hence the total quantity. Without the master controller, only the mixing ratio remains constant, whereas the total quantity of the two substances is disregarded.

A ratio controller is a standard controller whose input stage has two inputs to suit this modified specification. With regard to the dynamic action, all the variations of the standard controller could conceivably be used. Because of the nature of the process, the controllers are usually continuous, modulating or actuating controllers, with PI or PID action. With microprocessor controllers, functions such as ratio control can usually be configured directly.

6.8 Multi-component control

In a multi-component control system, various process-dependent variables produce the process value for the controller and determine the control deviation, as in a steam/feed water control, for instance (see Fig. 74).

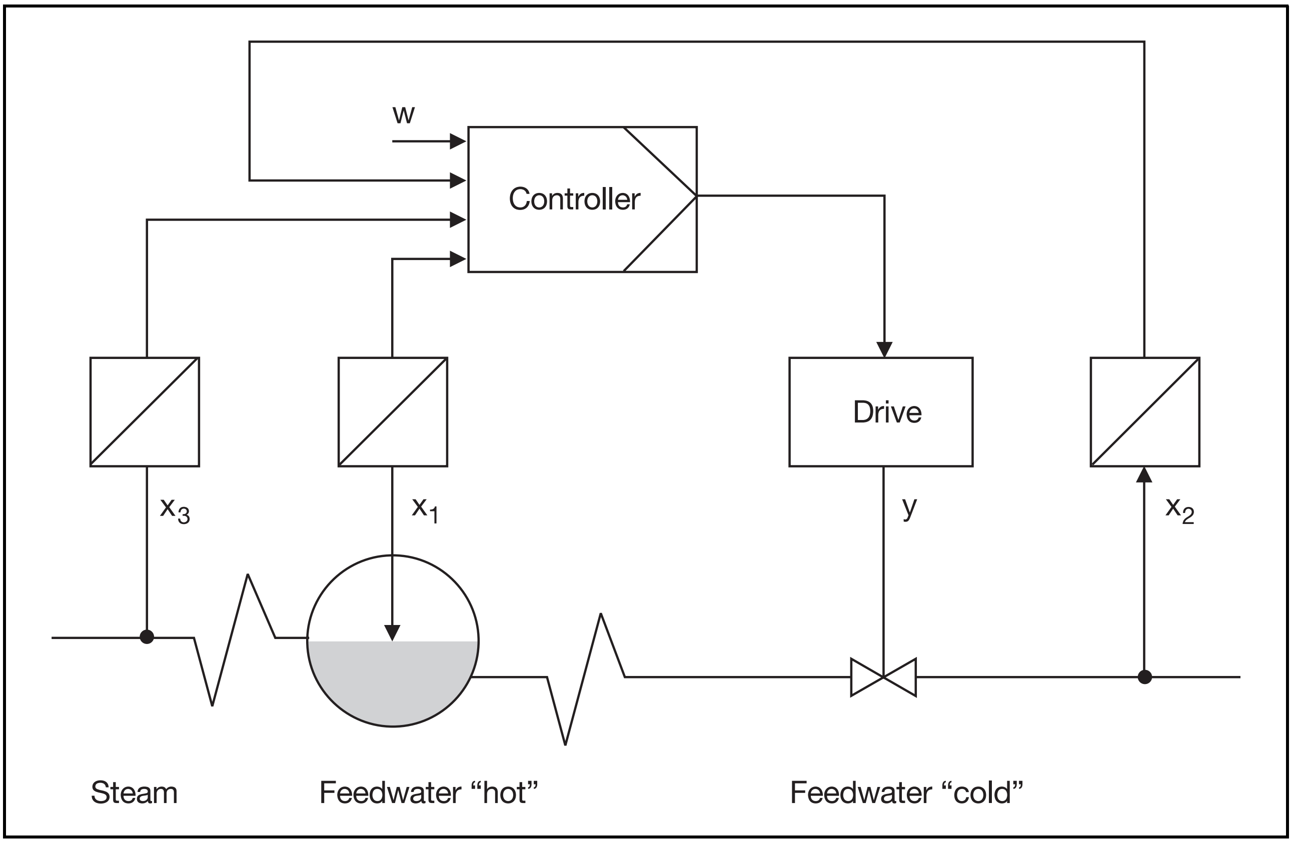


Fig. 74: Multi-component control

In this case, the individual process values can each be allocated a different weighting factor, so that they affect the control deviation to different extents; the main process variable is normally allocated the highest weighting factor.

In the example given, the following relationship might apply:

$$x = x_1 + 2b(x_2 - (2ax_3 + c))$$

x_3 = steam flow
x_2 = feedwater flow
x_1 = water level

If we now consider the relationship in the above formula without the constants, we have:

$$x = x_1 + (x_2 - x_3)$$

We can see that when the steam draw-off x_3 is equal to the cold water supplied x_2, the expression in brackets becomes zero, and the process value x now depends only on the water level x_1. Such a multi-component controller must have suitable number of inputs, and, if necessary, a facility for combining the individual signals via certain computations. In the example shown, there are three components, hence this circuit has been given the name “3-component control”.

This chapter introduces a number of additional functions, i.e properties which are specific to particular controllers and are required for certain control tasks. With modern microprocessor-based instruments, such options can be implemented using suitable software.

7.1 Control station / manual mode

With a control or auto-manual station, the manipulating variable can be altered directly. The control loop is interrupted so that the process variable has no significance. The controller now acts purely as an operating device (see Fig. 75).

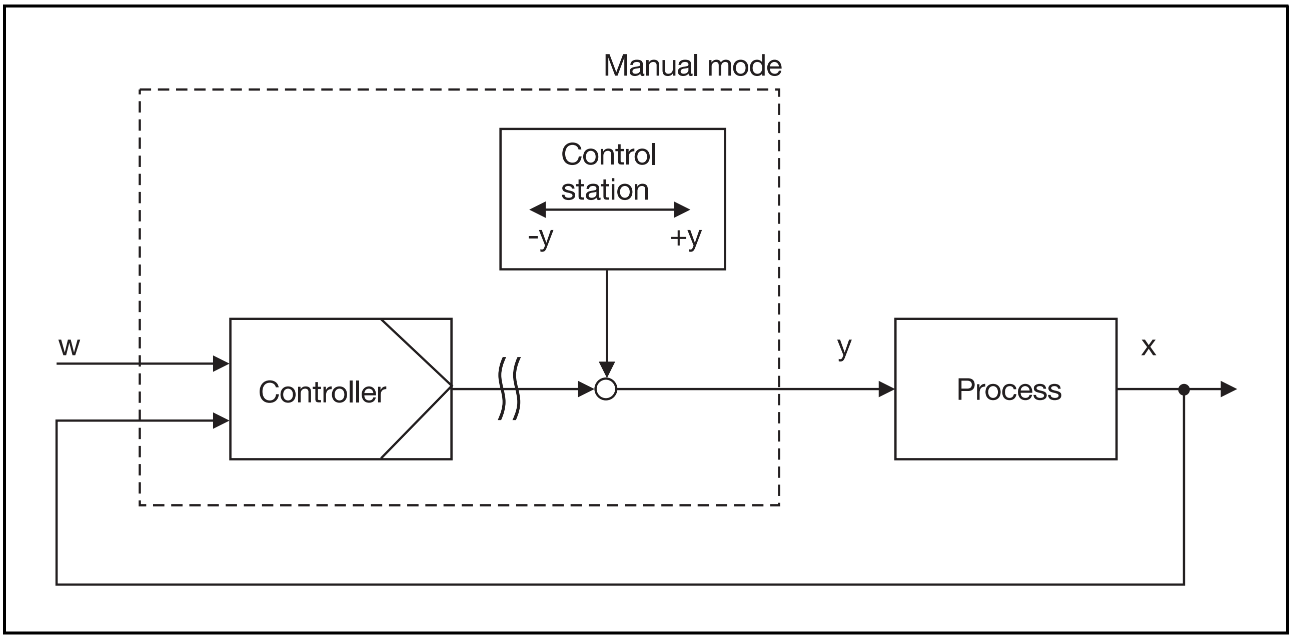


Fig. 75: Use of a control station

This feature is useful where valves or actuators have to be fully opened or closed for cleaning, as this would not otherwise be possible. Control stations are also invaluable for trials and test runs, when the manipulating variable has to be operated in manual mode, i.e. without automatic control. They can be integrated in the controller or arranged as a separate instrument.

Nowadays, with many microprocessor controllers, the function of a control station is provided by the manual mode setting. If a controller is switched from automatic mode (where the controller is trying to control the process value at the setpoint) to manual mode, the control function is disabled. Now the actuator, the thyristor power regulator, the cooling etc., can be controlled by setting the manipulating variable manually. Manual operation is possible with all types of controller.

In manual mode, the operator could set a value for the actuator very different to the value of the current manipulating variable set by the controller. However, such an extreme change could have a destructive effect on the actuator. This problem is overcome by the provision of bumpless transfer from automatic to manual mode, where the manipulating variable remains at its current value, and can then be changed manually.

In case of a broken probe, caused, for example, by a cable fault or mechanical damage to the sensor, automatic control is no longer possible, and the controller switches off the energy supply for safety reasons. In certain processes this could destroy the batch or cause a lengthy loss of production, owing to the long warm-up times or similar such conditions. Here, controllers offering manual operation have the advantage that, in such cases, the process can still be operated manually, albeit with reduced accuracy, and brought to completion.

7.2 Ramp function

Ramp functions are often used in processes with delicate products or to protect heating elements, as they ensure that the process variable approaches the setpoint slowly. In processes with large excess power, it also makes sense to limit the rate of rise of the process variable. With this function, the user sets a maximum rate of change (e.g. in °C/min, °C/hr etc.) at which a ramp setpoint travels towards the main setpoint. The process variable continuously tracks this changing value (see Fig. 76).

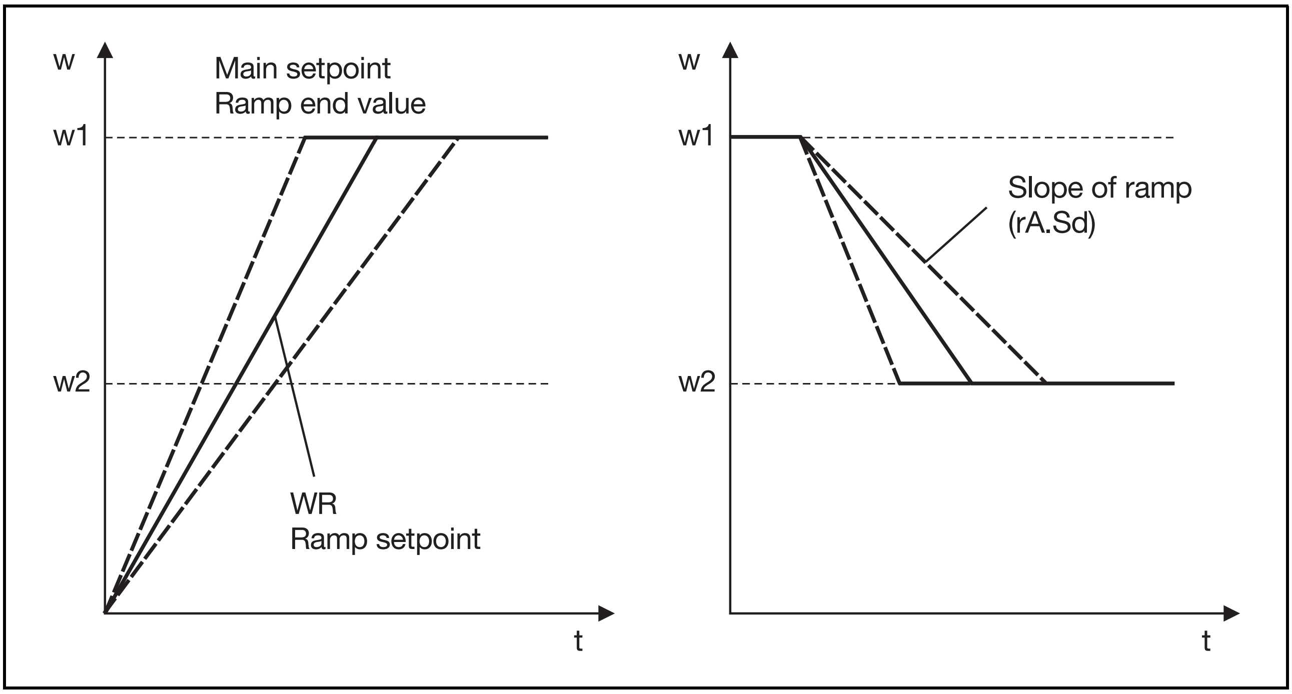


Fig. 76: Diagram of the ramp function

When the ramp function is activated, the ramp setpoint is normally made equal to the current process value, and then altered towards the main setpoint at the set gradient. Once the setpoint is reached, the ramp function is terminated, and the instrument controls at the set value until, for instance, the main setpoint is changed. If it does change, the newly activated value will once again be approached by a ramp. In this way, both rising and falling ramps can be implemented.

7.3 Limiting the manipulating variable

A manipulating variable limit can be used to limit the controller output signal at either a maximum or a minimum value. One application is where the actuating device fitted (e.g. pump, electric heating etc.) is over-sized; it avoids excess power and its associated problems, such as the process variable overshooting the setpoint. Further, a minimum manipulating variable limit can be a wise precaution in the control of gas burners, for instance. Setting a minimum manipulating variable (e.g. 3%) avoids the gas supply being interrupted and thus the burner going out. The controller then only gives out a manipulating variable within the range of the set minimum or maximum values.

In a continuous controller with a 0 — 10V output signal and a 5 — 95% range limitation, the output signal is restricted to the range 0.5 — 9.5V.

7.4 Program controller

So far, we have always assumed that the process variable has to be maintained constant at a fixed setpoint value. Such controls are also called fixed setpoint controls. In comparison, there are also a number of manufacturing processes where the setpoint does not represent a fixed value for the control system. Instead it represents a parameter which varies with time, i.e. a specific profile for the variation of the setpoint with time is required. Controllers used for this application are called program or profile controllers. They are often encountered as temperature controls, e.g. in annealing furnaces for metallic materials, in ceramic kilns, in chemical engineering for the batch manufacture of products, in climatic chambers etc.

The selectable program time is normally of the order of minutes, hours or even days. Fig. 77 shows such a profile for the course of the setpoint. Conditioned by the process, the course of the process variable does not normally show the same sharp transitions as the setpoint profile.

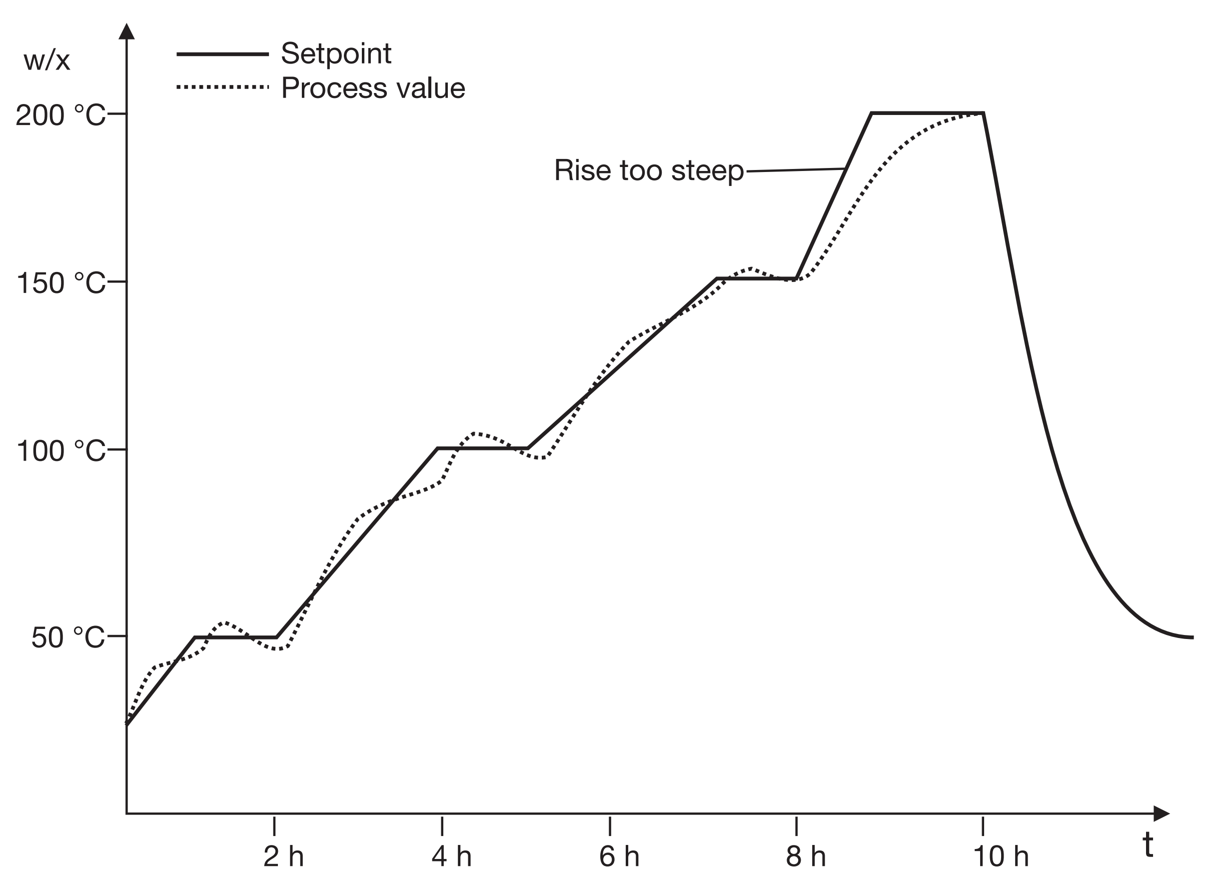


Fig. 77: Diagram of a setpoint profile

The setpoint profiles are usually predetermined by an external programming device or a computer; they are fed into the controller via the interface or a separate setpoint input. As the process variable has to follow the continuously varying setpoint, the arrangement is also called follower control.

The rise of the setpoint profile must not exceed the rise of the process characteristic, otherwise the process variable no longer follows the program profile, but rises at the maximum rate set by the process characteristic (see Fig. 77). This is particularly important in non-linear processes such as temperature controls, where the process gain decreases with higher process values. This can sometimes be avoided by programming the instrument so that the program run is terminated as soon as a certain deviation occurs, i.e. the process variable can no longer follow the program.

7.5 Self-optimization

Optimum adjustment of a controller to the process can be a time-consuming affair, particularly if the process concerned is rather slow. Furthermore, as we saw in Chapter 2.8, the optimal values for X_P, T_d and T_n are dependent on the working point. So it is quite likely that several sets of control parameters will have to be found for one process.

It seems an obvious step to let the controller itself make automatic adjustments to the control parameters. With microprocessor controllers, the facility for automatic controller setting (self-optimization) is available.

Basically, the choice lies between controllers which determine the control parameters simply on the basis of user requirements, and adaptive controllers, where the settings are continually checked and changed.

The methods used for self-optimization are usually built into the controller as software functional blocks, and identify the process on the basis of the response of the process variable to step changes in manipulating variable, and from this determine the best control parameters. The values determined for X_P, T_d, T_n etc. can be refined by the user (see Fig. 78).

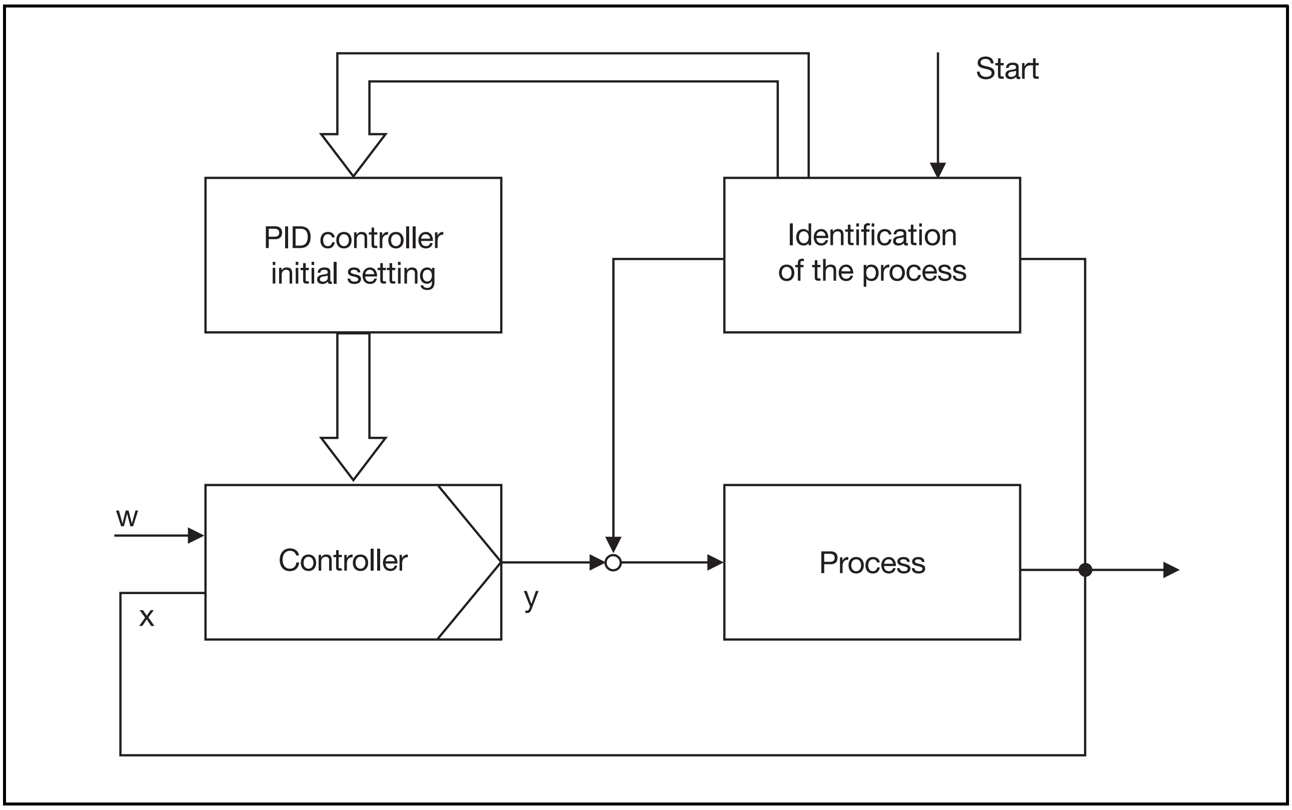


Fig. 78: Operating principle of self-optimization

In principle, a self-optimization can be arranged, for instance, to take place about the setpoint: if the controller has stabilized the process variable at the setpoint, the self-optimization can then be started, and the controller outputs 100% and 0% manipulating variable alternately. The controller determines the best parameters by examining the oscillations of the process variable about the setpoint, and then automatically accepts the parameters. With this method, it should be ensured that no damage can be caused when the process variable exceeds the setpoint.

Fig. 79 shows another type of self-optimization: in this case, the self-optimization process is started when the process value is below the setpoint. The controller determines the delay time from the initial response of the process variable. The controller calculates a switching level based jointly on the delay time and the gradient of the response of the process variable, and if this level is exceeded it sets the manipulating variable to 0%. With a linear process, interposing the switching level prevents the process variable from overshooting the setpoint. With non-linear processes, the overshoot is not completely prevented, but is at least reduced. In all, the controller outputs 100% manipulating variable twice, interrupted by a 0% manipulating variable output. Afterwards, the controller accepts the optimized parameters and controls accurately at the setpoint.

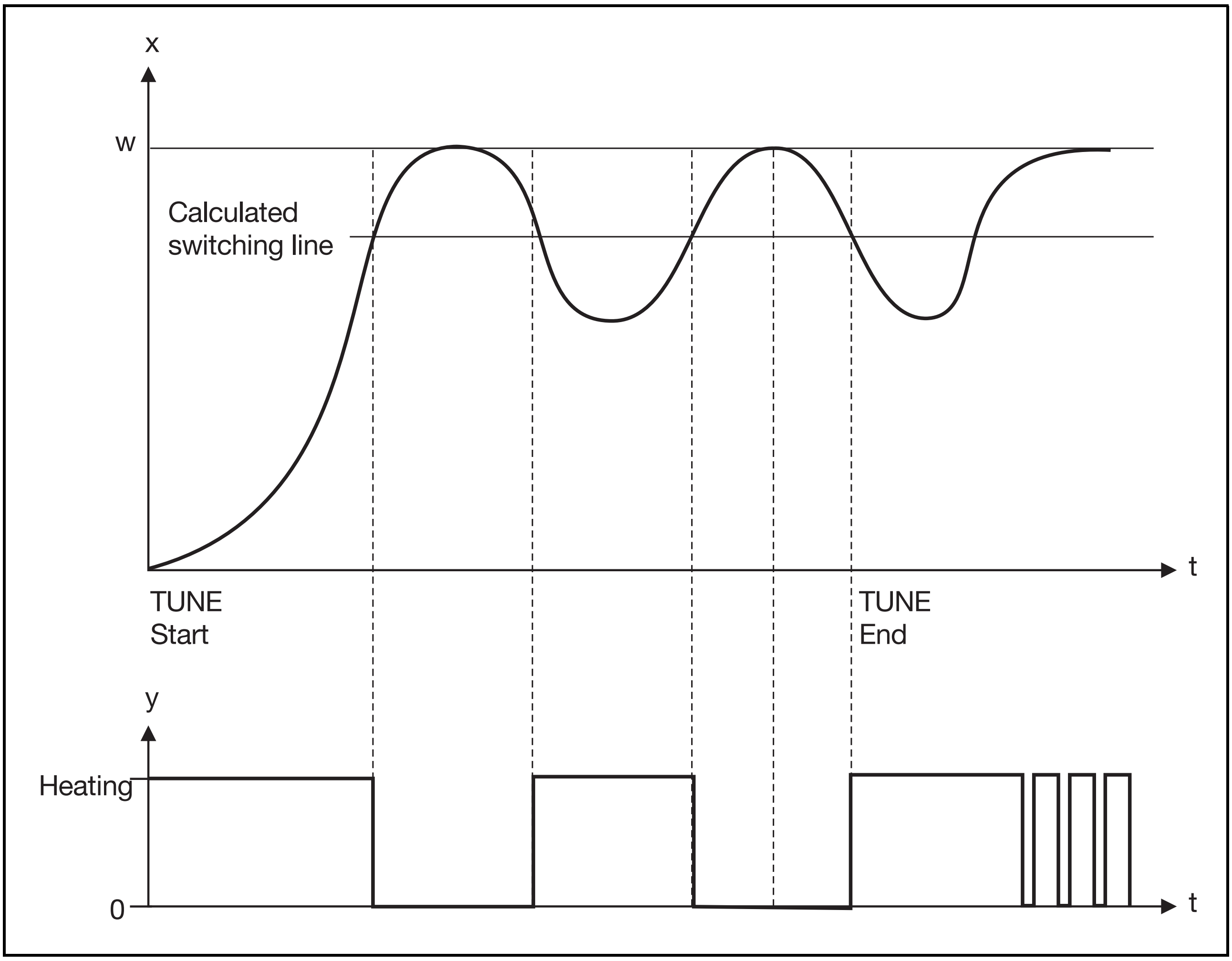


Fig. 79: Fluctuation of the process variable about the switching line

Manufacturers normally assume a process with self-limitation and without dead time elements as the basis for determining the control parameters. The closer the actual process corresponds to this model, the more effective is the self-optimization.

7.6 Parameter/structure switching

By switching between parameter sets, it is possible to operate the controller in a control process which requires different settings as the conditions vary. A set may comprise various parameters, such as dynamic response, X_P bands, cycle time C_y etc. Some controllers even allow switching between complete controller structures, such as switching from PD to PID structure.

Switching between sets is initiated either via the controller's logic inputs or via programmed values which depend on the control deviation or various setpoint ranges (see Fig. 80).

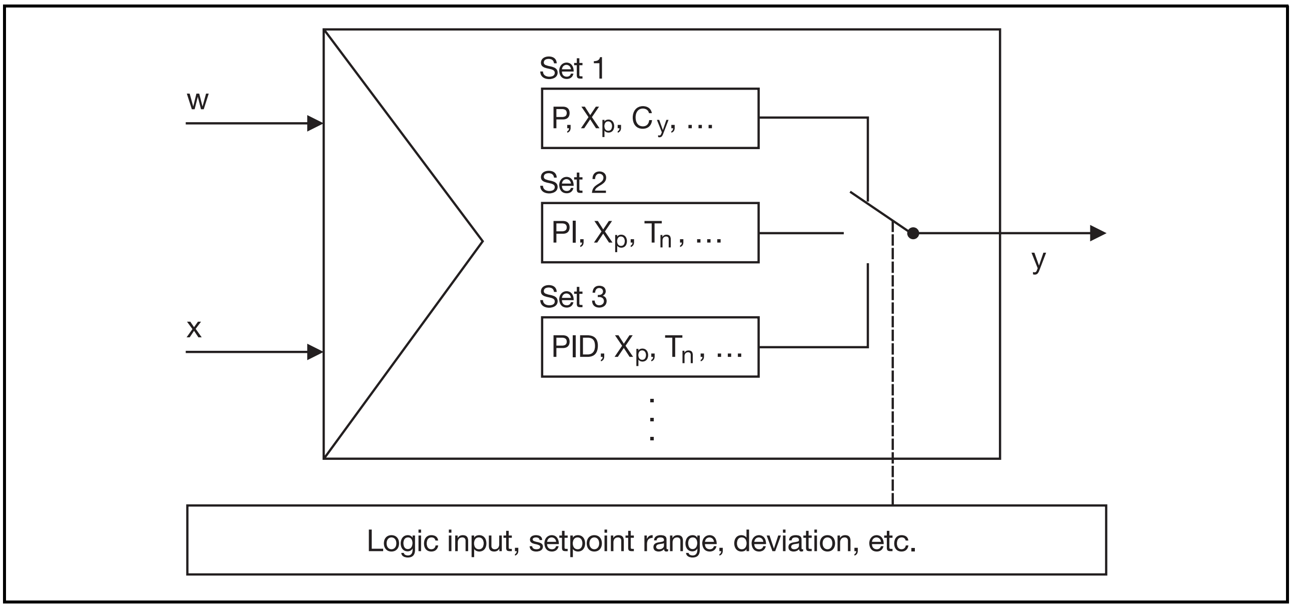


Fig. 80: Schematic diagram of parameter set switching

Parameter set switching of this type finds application, for instance, when there are repeated start-ups during automatic operation, where the setpoint must be reached in minimum time, and where the process variable x must not overshoot the setpoint. This is not normally possible, however, with the parameters used for stable operation. As an example, depending on the control deviation, the controller switches from P action with a somewhat smaller X_P band to PI action before the setpoint is reached. Another application where switching can help is when running different charges in an industrial furnace (annealing furnace). The control action of the process changes according to the loading (half charge, full charge), and the controller must adapt to each individual case. Various parameter sets can be allocated, based on certain preliminary tests. Moreover, this special function can be used to advantage where different working points are operated over the full range of the characteristic, and where a non-linear process is concerned, in which variable margins of excess power are to be expected.

7.7 Fuzzy logic

Fuzzy logic was developed in the USA by Lotfi A. Zadeh in 1965. This mathematical method is based on the fuzzy set theory, which represents an extension of Boolean algebra. This technology has undergone continuous development, and its advantages were quickly recognized in Japan, where they began to apply the technology to a multitude of control tasks (fuzzy control). This application of fuzzy logic to automation and control engineering represents a logical extension of traditional control technology.

A detailed coverage of this technology is outside the scope of this book; only its application in combination with a PID controller will be illustrated. It should be mentioned briefly that fuzzy logic deals with the subjective uncertainty of expressions, such as “temperature too high”, “pressure too low” or “humidity too low”, as well as decision-making based on such linguistic variables. In this way, it reconstructs the imprecise concepts of human thought.

The following illustrates how fuzzy logic and a PID control algorithm work together to improve the disturbance and setpoint responses of a control loop. Fig. 81 shows the simplified block diagram of a PID controller combined with a fuzzy module. The input variables of the fuzzy controller are the control deviation and the time derivative of the process variable, as well as information on whether the controller should operate for setpoint or disturbance response. The output manipulating variable of the fuzzy controller is weighted by a parameter Fc1 and added to the manipulating variable of the PID controller. In this way, the manipulating variable acting on the process is made up of the manipulating variable of the PID controller and that of the fuzzy controller. A second output from the fuzzy controller is a control output which influences the PID control parameters according to parameter Fc2. Thus, with just two parameters, Fc1 and Fc2, the fuzzy controller can be matched to the process.

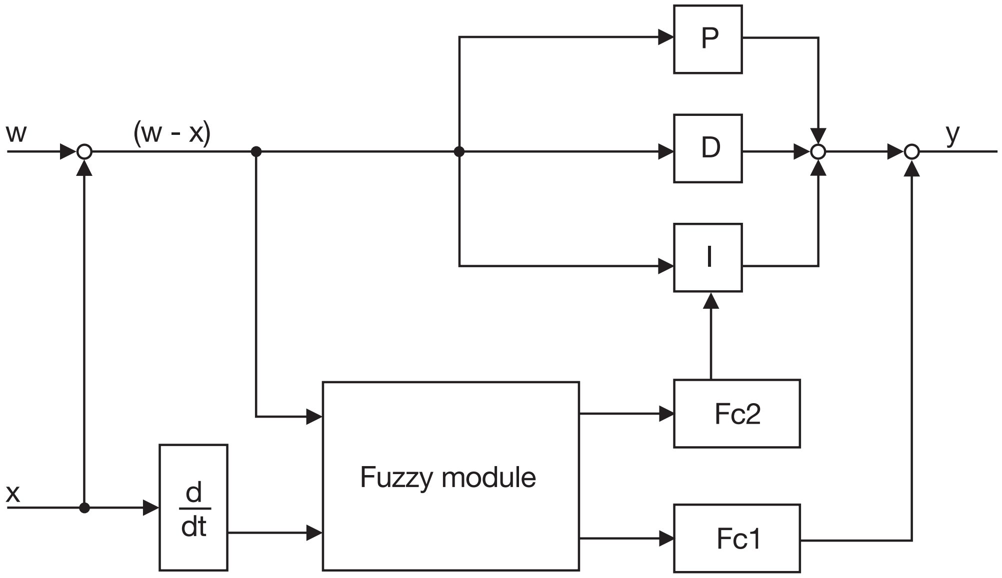


Fig. 81: Simplified block diagram of a PID controller with fuzzy module

The fuzzy controller incorporates linguistic control blocks based on “IF-THEN” rules. These linguistic rules determine the transient response of the fuzzy controller to setpoint changes and disturbances.

The combination of a fuzzy controller in parallel with a conventional controller offers several advantages:

- With a non-linear process behavior or on higher-order processes, the fuzzy controller can compensate for specific inaccuracies of the PID controller by supportive intervention under critical operating conditions.

This is exemplified by systems whose behavior changes over the operating time. The fuzzy controller is less sensitive to process parameter changes than the controller with its fixed parameter settings. This means that variable process characteristics are dealt with more effectively by the combination of fuzzy logic and the PID controller.

- For setpoint response, the integral fuzzy controllers increase the damping factor of the control loop in certain conditions. This means that any tendencies to oscillate during the run-up to the setpoint are reduced and overshoots are damped considerably.
- The problem of optimizing the controller specifically for setpoint or disturbance response is lessened by the fuzzy components, as fuzzy control combines the control theory of both optimization criteria.
- When disturbances occur, the fuzzy controller responds more dynamically than the PID controller and hence stabilizes the disturbance more rapidly. The criterion of control effectiveness is improved, i.e. the resulting control error area is smaller than with stabilization by the PID controller alone (see Fig. 82).

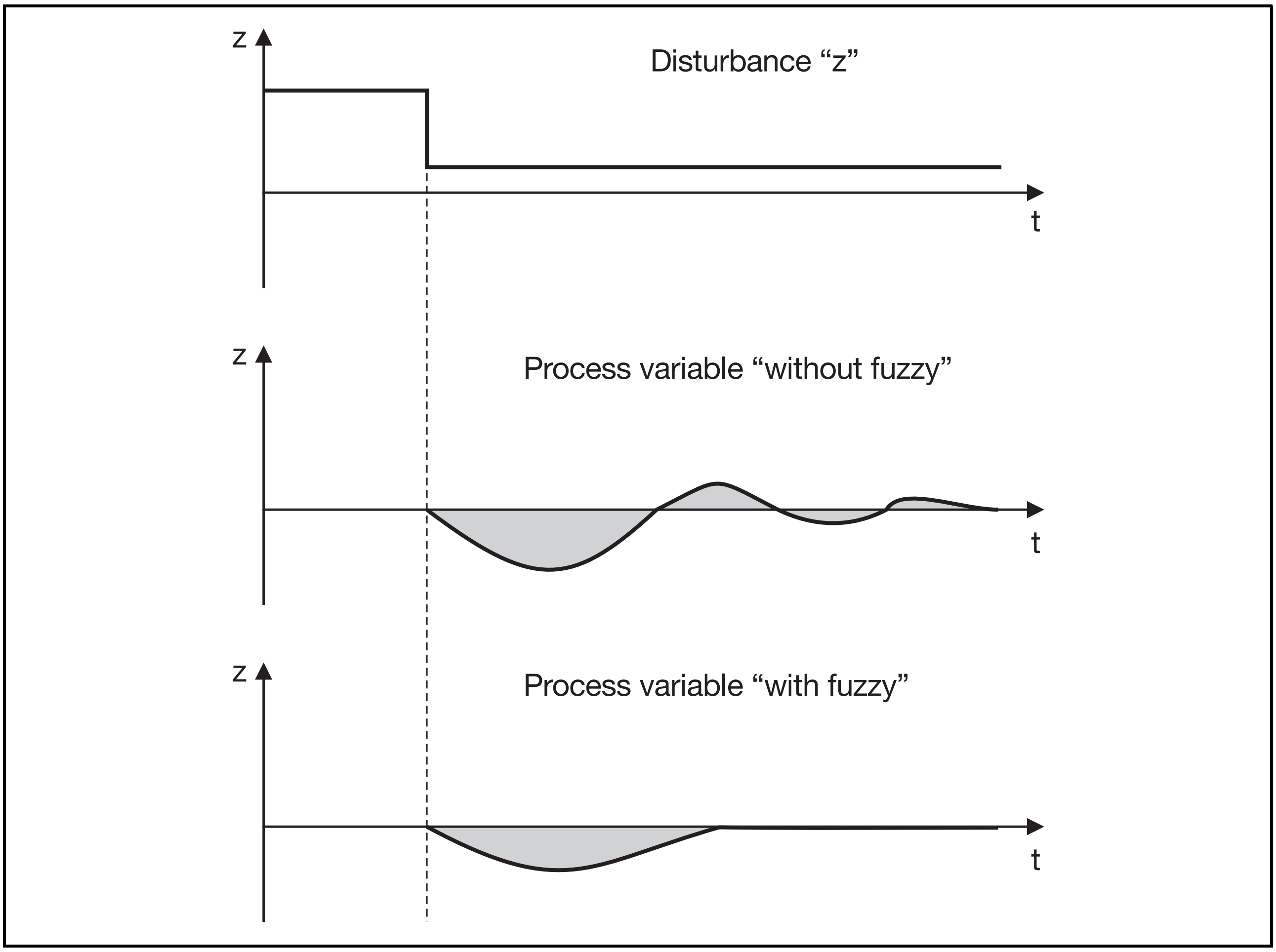


Fig. 82: Disturbance response of a third-order process, using a controller with and without fuzzy module

We can summarize by saying that fuzzy logic, in its application as fuzzy control in automation and control engineering, is highly regarded today, despite initial problems in accepting this technology. The fuzzy concept has proved to be a powerful tool, particularly in complex control tasks. Traditional control technology will certainly not be replaced by fuzzy logic, but technological processes will emerge that can be controlled more economically and more safely using fuzzy logic. The combination of the fuzzy controller with the PID controller is a sensible example of this.

From this we can record that the fuzzy controller in general, as well as the combination of a traditional conventional controller and fuzzy module, bring about increased flexibility, so that the controller can be better matched to different processes and control objectives.

8 Standards, symbols, literature references

Standards

Standards and guidelines are very important in engineering, including control engineering. They lay down concepts and designations, component values, dimensions, as well as important numerical values and numerical ranges.

From the large number of published standards and guidelines, the following list below has been limited to those which are useful for the basic principles of control engineering and have been used in the preparation of this book.

DIN 19 221	1993-05	Messen, Steuern, Regeln; Formelzeichen der Regelungs- und Steuerungstechnik (Measurement, operation, control; symbols in control engineering)
DIN 19 225	1981-12	Messen, Steuern, Regeln; Bennenung und Einteilung von Reglern (Measurement, operation, control; control engineering concepts)
DIN 19 226	(1-5)	Regelungs- und Steuerungstechnik; Begriffe und allgemeine Grundlagen (Control Engineering; Definitions and Terms)
DIN 19 236	1977-01	Messen, Steuern, Regeln; Optimierung, Begriffe (Measurement, operation, control; optimization, definitions)
Guideline VDI/VDE 2189 Sheet 1		Beschreibung und Untersuchung von Zwei- und Mehrpunktreglern ohne Rückführung (Description and analysis of two-state and multi-state controllers without feedback)
Guideline VDI/VDE 2189 Sheet 2		Beschreibung und Untersuchung von Zwei- und Dreipunktreglern mit Rückführung (Description and analysis of two-state and three-state controllers with feedback)
Guideline VDI/VDE 2189 Sheet 3		Beschreibung und Untersuchung von Dreipunkt-Schrittreglern (Description and analysis of modulating controllers)
Guideline VDI/VDE 2189 Sheet 4		Beschreibung und Untersuchung digital arbeitender Kompaktregler (Description and analysis of digital compact controllers)
Guideline VDI/VDE 2190 Sheet 1		Beschreibung und Untersuchung stetiger Regelgeräte Grundlagen (Description and analysis of continuous control instrument fundamentals)

Symbols

C_y	cycle time, quasi-continuous controller	T_{on}	switch-on time, quasi-continuous controller
e	deviation	T_{osc}	oscillation time, discontinuous controller
f_{sw}	switching frequency, discontinuous controller	T_s	stabilization time, quasi-continuous controller
K_I	integral coefficient of the controller	T_t	dead time

8 Standards, symbols, literature references

K_{IS}	transfer coefficient of process without self-limitation
K_S	transfer coefficient (process gain) of the process
K_P	proportionality factor of the controller
R	ON-time ratio, quasi-continuous controller
$R(\%)$	ON-time ratio, quasi-continuous controller in%
t	time (continuous)
T	time constant of first-order process
T_a	approach time
T_C	oscillation time at X_{Pc}
T_d	derivative time of the controller
T_g	response time
T_{Mmin}	minimum pulse duration, modulating and actuating controllers
T_n	reset time of the controller
T_{off}	switch-off time, quasi-continuous controller
T_u	delay time
T_I	integral time
T_y	stroke time
V_{max}	maximum rate of rise
w	setpoint (SP)
x	process variable, process value (PV)
X_{Max}	maximum process value (process variable)
X_o	overshoot
X_P	proportional band of the controller
X_{Pc}	critical X_P, at which the control loop with P controller oscillates uniformly
X_{Sd}	switching differential, discontinuous controller
X_{Sh}	contact spacing
y	manipulating variable (MV)
y_H	manipulation range
z	disturbance

References

- E. SAMAL / W. BECKER ‘Grundriss der Praktischen Regelungstechnik 18’
  Basics of practical control engineering 18 (Oldenburg)
- LENZ /OBERST / KOEGST ‘Grundlagen der Steuerungs- und Regelungstechnik’
  Fundamentals of control engineering (Hüthig Verlag)
- P. BUSCH ‘Elementare Regelungstechnik’
  Elementary control engineering (Vogel Verlag)
- BÖTTLE / BOY / GROTHUSMANN ‘Elektrische Mess- und Regeltechnik’
  Electrical measurement and control engineering (Vogel Verlag)
- D. WEBER ‘Regelungstechnik- Wirkunsweise und Einsatz elektronischer Regler’
  Control engineering - operation and application of electronic controllers (Expert Verlag)
- D. WEBER ‘Elektronische Regler- Grundlagen, Bauformen und Einstellkriterien’
  Electronic controllers - fundamentals, types and adjustment criteria (JUMO)
- JUMO GmbH & Co. KG ‘Betriebsanleitung- Universeller Programmregler
  JUMO DICON 401/501’
  Operating Instructions - universal profile controller JUMO DICON 401/501

Index

measurement • control • recording

JUMO GmbH & Co. KG · 36039 Fulda, Germany · Phone +49 661 6003-396
Fax +49 661 6003-500 · e-mail: manfred.schleicher@jumo.net · Internet: www.jumo.net

02.04/00323761

ISBN 3-935742-01-0