

PROCESS CONTROL

1. Objective of Process Control

Process control is used in order to maximize production while maintaining a desired level of product quality and safety and making the process more economical. Because these goals apply to a variety of industries, process control systems are used in facilities for the production of chemicals, pulp and paper, metals, food, and pharmaceuticals. While the methods of production vary from industry to industry, the principles of automatic control are generic in nature and can be universally applied, regardless of the size of the plant (1–3).

A block diagram of a process with a single manipulated variable and a single controlled variable is shown in Figure 1. This diagram includes feedforward, feedback, and supervisory control. The main aim of the feedback controller of this process is to keep the controlled variable Y , which is measured by some instrument as close as possible to the desired set point Y_{SP} . The controlled variable may be the quality of the final product, which could be pressure, temperature, liquid level, composition, or any other inventory, environmental, or quality variable. Set points are often determined by a supervisory control system using real-time numerical optimization techniques. Several different types of final control elements exist but the most common one is a control valve for controlling some flow of material. The disturbance variable D , also called the load variable, can cause the controlled variable to deviate from its set point, requiring control action in order to bring it back to its desired operating point. Both feedback and feedforward control can reduce the effects of disturbances, where each method has its own advantages and drawbacks that are discussed in more detail below. Disturbances can result from a variety of sources, including external environmental variables; they can occur randomly or have an underlying pattern to them, but in any case a disturbance variable cannot be influenced by the controller of the process. The error, or deviation, E , between the controlled variable Y and its set point Y_{SP} is the input to the feedback controller, which changes the manipulated variable U in order to decrease the error. In a typical process plant, there may be hundreds or even thousands of control loops such as the one shown in Figure 1.

1.1. Feedback Control. The purpose of feedback control is to keep the controlled variable close to its set point. This task can be achieved by computing the difference between the set point and the controlled variable and passing this as the input to the feedback controller. By its design, the feedback controller takes corrective action to reduce the deviation. Most often the manipulated variable moves in a direction opposite in sign to the error, which is called negative feedback. For example, if the temperature in a shower were to be controlled with the cold water flow, an increase in temperature above the set point gives a negative error between set point and controlled variable; hence, the manipulated variable should move in the opposite (positive) direction to compensate for the error, which in this case is an increase in cold water flow. Feedback controllers have user-specified parameters that can be adjusted to achieve desirable dynamic performance. The design of different types of feedback controllers is discussed in more detail in Sections 5 and 6.

1.2. Feedforward Control. A feedback controller can only take action after the controlled variable deviates from its desired set point and generates a non-zero error. However, the response to disturbances can be very sluggish if the process or measurement changes very slowly. In such a situation a feedforward controller can improve the performance. The feedforward controller predicts the effect that the disturbances will have on the controlled variable and takes control action that will counteract the influence of the disturbances. Since this control action is taken based upon model predictions, it can minimize the effect that the disturbances have on the controlled variable before any unwanted deviations occur. However, in order to make these predictions, the disturbances must be measurable and a model for the effect that the disturbances will have on the controlled variable is required. Because it is not possible to predict and measure every disturbance that affects a process, feedforward control is usually combined with feedback control. In such a configuration, the feedforward controller can counteract the effect of the measurable disturbances quickly, while the feedback controller will eliminate offset resulting from unmeasured load disturbances. Additional details on feedforward controllers can be found in Section 6.

1.3. Regulation and Set Point Changes. As mentioned earlier, the purpose of feedback control is to keep the controlled variable close to its set point. There are two reasons why the controlled variable may deviate from its set point: the set point is changed deliberately in order to achieve better performance or load disturbances drive the operation away from its desired set point. Controllers designed to reject load disturbances are called regulators while controllers designed to track set point changes are called servomechanisms. For most continuous processes set point changes occur infrequently, typically only if the supervisory controller computes a more favorable operating point. Due to this, regulators are the most common form of feedback controllers in continuous plants. In contrast, controllers for servo problems are common in batch plants, where frequent changes in the set points occur. While any controller should be designed such that it can be used for servo as well as for regulatory problems without causing the system to become unstable it may not exhibit good performance on both types of problems. Therefore, it is an important design task to determine which type of problem is dominant in a process and design the controller accordingly.

2. Control Hardware and Software

Process control as practiced in the process industries has undergone significant changes since it was first introduced in the 1940s. In the early 1960s, electrical analogue control hardware replaced much of the pneumatic analogue control hardware in many process industries. Certain control elements, ie, control valve actuators, have remained pneumatic even today. Electrical analogue controllers of the 1960s were single-loop controllers in which each input was first brought from the measurement point in the process to the control room where most of the controllers were located. The output from the controller was then sent from the control room to the final control element. The operator interface consisted of a control panel having a combination of display faceplates and

chart recorders for single-loop controllers and indicators. Control strategies primarily involved feedback control, usually with a proportional-integral (PI) controller.

During the late 1950s and early 1960s, a few firms introduced process control computers to perform direct digital control (DDC) and supervisory process control. In cases where the system made extensive use of DDC, the DDC loops often had close to 100% analogue control backup making the systems costly. Other early systems primarily used process control computers for supervisory process control. Regulatory control was provided by analogue controllers, which did not require backup, but the operator's attention was split between the control panel and the computer screens. The terminal displays provided the operator interface when supervisory control was being used, but the control panels were still located in the control room for the times when the analogue backup was necessary. Within this environment, some firms began to broaden the use of advanced control techniques such as feedforward control, multivariable decoupling control, and cascade control.

The functionalities of these early control systems were designed around the capabilities of the computers rather than the process characteristics. These limitations, coupled with inadequate operator training and an unfriendly user interface, led to designs that were hard to operate, maintain, and expand. In addition, many different systems had customized specifications, making them extremely expensive to assemble. Although valuable experience was gained in systems design and implementation, the lack of financial success hindered the infusion of digital system applications into the process industries until about 1970, when inexpensive microprocessors became commercially available.

2.1. Distributed Control Systems. Starting in the mid-1970s, control vendors introduced microprocessor-based distributed control systems (DCSs) and programmable logic controllers (PLCs). A digital distributed control system consists of many elements, as shown in Figure 2. Host computers perform computationally intensive tasks like optimization and advanced control strategies. Data highways, consisting of a digital transmission link, connect all other components in the system. Redundant data highways reduce the possible loss of data. Operator control stations provide video consoles for operator communication with the system, in order to supervise and control processes. Many control stations contain printers for alarm logging, report printing, or hard-copying of process graphics. Remote control units implement basic control functions like PID algorithms and sometimes provide data acquisition capability. Programmer consoles develop application programs for the distributed control system. Mass storage devices store the process data for control purposes as well as corporate decisions. Storage devices can be in the form of hard disks or databases. Communications and interactions between controllers, inputs, and outputs were realized by software, not by hardwiring. DCSs, therefore, revolutionized many aspects of process control, from the appearance of the control room to the widespread use of advanced control strategies.

Since the early 1980s, the capabilities of DCSs have improved dramatically (4–6). There has been a general increase in the use of digital communications technology within process control. Some advanced control strategies are implemented within the DCS. Most local control units perform their own analogue-

to-digital (A/D) and digital-to-analogue (D/A) conversion and can be located in equipment rooms closer to the process. Digital communications via a coaxial or fiber optic cable send information back to the control room, thus saving on wiring costs. With this trend toward increased use of digital communications technology, smart transmitters and smart actuators are also gaining in popularity. These devices, equipped with their own microprocessor, perform tasks such as autoranging, autocalibration, characterization, signal conditioning, and self-diagnosis at the device location. Thus tasks required of the local control unit or the data acquisition unit are reduced.

The features that have made DCSs popular are (1) reduction in wiring and installation costs through the use of data highways and remotely located local control units; (2) reduction in the space requirements for panels in the control room (3) improved operator interface with customized screens; (4) ease of expansion because of modularity of the DCSs; (5) increased flexibility in control configuration, allowing control strategies to be modified without any need for rewiring; and (6) improved reliability and redundancy.

Client–Server (Personal Computer) Configuration. Most distributed control systems today are open solutions based on component object models in order to interact with other programs. These technologies allow tying together the best applications of each kind from different vendors in order to optimize plant-wide control. This is in contrast to earlier control systems where a single vendor had to supply the whole plant automation system because no interaction with other vendor products was possible. The emergence of software technologies such as CORBA (Common Object Request Broker Architecture) and COM (Component Object Model) as well as the increasing use of the programming language Java and the internet/intranets has driven this development toward open solutions (or “plug and play”). Personal computers are replacing panel boards as operator stations, which makes it easier to exchange data from the control system with other applications running on personal computers or in the network.

2.2. Programmable Logic Controllers. Initially, programmable logic controllers were dedicated, stand-alone, microprocessor-based devices executing straightforward binary logic for sequencing and interlocks. These were originally intended for applications which, prior to that time, had been implemented with hardwired electromechanical and electrical relays, switches, pushbuttons, and timers. The PLCs significantly improved the ease with which modifications and changes could be implemented to such logic. Although many of the early applications were in the discrete manufacturing industries, the use of PLCs quickly spread to the process-related industries. The PLCs have become increasingly more powerful in terms of calculational capabilities (7,8), eg, PID algorithms; data highways to connect multiple PLCs; improved operator interfaces; and interfaces with personal computers and DCSs. Batch process control is dominated by logic-type controls, and PLCs are a preferred alternative to a DCS. Because of the availability of relatively smooth integrated interfaces between DCSs and PLCs, current practice is generally to use an integrated combination of a DCS and PLCs.

Up to several thousand discrete (binary) inputs and outputs can be accommodated with PLCs, which may also have several hundred analogue inputs and outputs for data logging and/or continuous PID control. All PLCs are designed to

handle Boolean (binary) logic operations efficiently. Because the logical functions are stored in main memory, one measure of a PLC's capability is its memory scan rate; typical values range from 10 to 50 ms. At the faster speeds, thousands of steps can be processed by a single unit. Most PLCs also handle sequential logic and are equipped with internal timing capability to delay an action by a prescribed amount of time, to execute an action for a prescribed time, etc. Newer PLC models often are networked to serve as one component of a DCS control system with operator I/O provided by a separate component in the network.

A distinction is made between configurable and programmable PLCs. The term configurable implies that logical operations (performed on inputs to yield a desired output) are located in the PLC memory, potentially in the form of ladder diagrams by selecting from a PLC menu or by direct interrogation of the PLC. Usually, the logical operations are put into PLC memory in the form of a higher level programming language. Most control engineers prefer the simplicity of configuring the PLC to the alternative of programming it. However, some batch applications, particularly those involving complex sequencing, are best handled by a programmable approach, perhaps through a higher level, computer control system.

2.3. Safety and Shutdown Systems. Plant safety is an important consideration in operating a plant. Furthermore, increasingly stringent government regulations require special attention to process safety during the design process. Process control plays an important role in the safety considerations of a plant. When automated procedures replace manual procedures for routine operations, the probability of human errors leading to hazardous situations is lowered. Additionally, the enhanced capability for presenting information to the process operators in a timely manner and in the most meaningful way increases the operator's awareness of the current plant condition. This reduces the time in which abnormal conditions can be recognized and minimizes the likelihood that the situation will progress to a hazardous state.

A protective system must be provided for processes where hazardous conditions can develop. One possible solution to this is to provide logic for the specific purpose of taking the process to a state where this condition cannot exist, called a safety interlock system. Because the process control system and the safety interlock system serve different purposes, they should be physically separated (9). The process control system will require more modifications due to changing process conditions than the safety interlock system and having a separate system reduces the risk of unintentionally changing the safety system as well.

Special high reliability systems have been developed for safety shutdowns, eg, triple modular redundant systems, which are designed to be fault-tolerant. This permits the system to have an internal failure and still perform its basic function. Basically a triple modular redundant system consists of three identical subsystems actively performing identical functions simultaneously. The results of the three subsystems are compared in a two-of-three voting network prior to sending the signals to the output devices. If any one of the subsystems experiences a failure, the overall system can still function properly as long as two of the subsystems are working. This setup allows the identification of components suspected of failure. To further increase reliability, multiple sensor and output

devices may be used. When multiple sensors are used, the system is often implemented along with a two-of-three voting network.

Alarms. The purpose of an alarm is to alert the process operator to a process condition that requires immediate attention (9). An alarm is activated whenever the abnormal condition is detected and the alert is issued. The alarm returns to normal when the abnormal condition no longer exists. Alarms can be defined on measured variables, calculated variables, and controller outputs. A variety of different classes of alarms exist. The most important ones are discussed in this section. A high alarm is generated when the value is greater than or equal to the value specified for the high-alarm limit. A low alarm is generated when the value is less than or equal to the value specified for the low alarm limit. A high deviation alarm is generated when the measured value is greater than or equal to the target plus the deviation alarm limit. A low deviation alarm is generated when the value is less than or equal to the target minus the deviation alarm limit. A trend alarm is generated when the rate of change of the variable is greater than or equal to the value specified for the trend alarm limit. One operational problem with alarms is that noise in the variable can cause multiple alarms whenever its value approaches a limit. This problem can be avoided by defining a deadband on the alarm. The high alarm is generated when the process variable is greater than or equal to the value specified for the high alarm limit. The high alarm return to normal is generated when the process variable is less than or equal to the high alarm limit less the deadband. Because the degree of noise varies from one input to the next, the deadband must be individually configured for each alarm.

2.4. Smart Transmitters, Valves, and Fieldbus. There is a clearly defined trend in process control technology toward increased use of digital technology. The trend, which started with digital controllers, has increasingly spread from that portion of the overall control system outward toward field elements such as smart transmitters and smart control valves. Digital communication occurs over a fieldbus, ie, a coaxial or fiber optic cable, to which intelligent devices are directly connected and transmitted to and from the control room or remote equipment rooms as a digital signal. The fieldbus approach reduces the need for twisted pairs and associated wiring (see Fig. 3).

Various field network protocols such as Foundation Fieldbus and Profibus provide the capability of transferring digital information and instructions among field devices, instruments, and control systems (10). The fieldbus software mediates the flow of information among the components. Multiple digital devices can be connected and communicate with each other via the digital communication line, which greatly reduces wiring cost for a typical plant. Manufacturers of instruments are focusing on interoperability among different fieldbus vendor products.

2.5. Process Control Software. The most widely adopted user-friendly approach is the fill-in-the-forms or table-driven process control languages (PCL). Popular PCLs include function block diagrams, ladder logic, and programmable logic (5). The core of these languages is a number of basic function blocks or software modules, such as analogue in, digital in, analogue out, digital out, PID, summer, splitter, etc. Using a module is analogous to calling a subroutine in conventional Fortran or C programs.

In general, each module contains one or more inputs and an output. The programming involves commuting outputs of blocks to inputs of other blocks via the graphical-user interface. Users are required to fill in templates to indicate the sources of input values, the destinations of output values, and the parameters for forms/tables prepared for the modules. The source and destination blanks may specify process I/O channels and tag names when appropriate. To connect modules, some systems require filling in the tag names of modules originating or receiving data. User specified fields include special functions, selectors (min or max), comparators (less than or equal to), and timers (activation delays). Most DCSs allow function blocks to be created.

2.6. Facility Control Hierarchy. Figure 4 shows the five levels in the manufacturing process where various optimization, control, monitoring, and data acquisition activities are employed (1). The relative position of each block in Figure 4 is intended to be conceptual, because there can be overlap in the functions carried out, and often several levels may utilize the same computing platform. The relative time scales where each level is active are also shown. Data from the plant (flows, temperatures, pressures, compositions, etc.) as well as so-called enterprise data, consisting of commercial and financial information, are used with the methodologies shown in order to make decisions in a timely fashion.

Advances in the capabilities of DCSs have made the incorporation of advanced controls within the DCS feasible. Modern process facilities are often designed with a relatively high degree of process integration in order to minimize the theoretical cost of producing the product. From an operation's standpoint, however, this integration gives rise to relatively complex interactions between the operating variables making it difficult to determine the plant adjustments to optimize the operation.

Each of the five conceptual control levels has its own requirements and needs in terms of hardware, software, techniques, and customization. Because information flows up in the hierarchy and control decisions flow down, effective control at a particular level occurs only if all the levels beneath the level of concern are working well. The highest level (planning and scheduling) sets production goals to meet supply and logistics constraints and addresses time-varying capacity and manpower utilization decisions. This is called enterprise resource planning (ERP) and the term supply chain in level 5 refers to the links in a web of relationships involving retailing (sales), distribution, transportation, and manufacturing. Planning and scheduling usually operate over relatively long time scales and tend to be decoupled from the rest of the activities in lower levels. In petroleum refining usually all of the refineries owned by an oil company are included in a comprehensive planning and scheduling model. This model can be optimized to obtain target levels and prices for interrefinery transfers, crude and product allocations to each refinery, production targets, inventory targets, optimal operating conditions, stream allocations, and blends for each refinery (1).

Generally, the various levels of control applications are aimed at one or more of the following objectives: (1) determining and maintaining the plant at a practical optimal operating point given the current conditions and economics; (2) maintaining safe operation for the protection of personnel and equipment; (3)

minimizing the need for operator attention and intervention; and (4) minimizing the number, extent, and propagation of upsets and disturbances.

At level 4, the plant-wide optimization level, the primary goal is to determine the optimal operating point of the plant's mass and energy balance and to adjust the relevant set points in an appropriate manner. There are many possible process interactions and combinations of constraints involved, thus these control applications help identify the optimal operating point. The appropriate optimization objective is dependent on the market situation, such that, at times, the optimization objective is to maximize production; at other times it is to minimize operating costs for a fixed production rate. The control applications at this level often utilize steady-state mathematical models for the process or portions thereof. These models must be tuned to a specific plant's operation and constraints in order to ensure that the key aspects of the overall plant operation are characterized. Execution frequency of the control applications at this level are therefore from hours to days, depending on the frequency of relevant variations. Standard mathematical techniques and heuristics based on experience are used to determine the operating point that best accomplishes the selected optimization objective. Once the optimum operating point has been determined, the relevant set points are passed down to the lower control levels. Because of the time constants and dynamics associated with the level 4 controls, set points are usually ramped incrementally to their new values in a manner such that the process is not disturbed and the proximity to constraints can be periodically checked before the next increment is made.

The plant optimization control level applications determine the values of key variables that optimize the overall plant material and energy balance. The control applications at the local optimization and supervisory control level, on the other hand, focus on subsystems within the overall plant. These subsystems usually consist of a single or, at most, a few highly interactive pieces of equipment. Most of the applications at this level are aimed at optimizing the process variables within an operating window defined by hard constraints, eg, equipment and material limits. Often the optimal operating point of the subsystems resides on one of the constraints of the operating window (see Fig. 5).

Hence, many of these control applications employ a constraint control strategy, ie, a strategy that pushes the subsystem against the closest active constraint. Typically the closest currently active constraint changes with time and situations, eg, between day and night, different weather conditions, different operating states of upstream equipment. The constraint control strategies continually make minor adjustments to keep the state of the system along the active constraint near its optimal point. Because such adjustments are made continually, these applications can generate significant benefits over the course of a year, although these benefits may appear minor when viewed at a single point in time. In addition to local optimization applications, the control level also includes multivariable, predictive, and model-based control strategies such as model predictive control and inferential control. The control applications at the local optimization and supervisory level typically provide set points for the controls at the advanced regulatory and basic regulatory control levels (Level 3a). The general objective of the advanced regulatory control level applications is to improve the performance of basic regulatory control level controllers (Level 3b).

The execution frequency of applications at this level is typically in the range of seconds to minutes and differ from the controls at the basic regulatory level in that the former controls are often multivariable and anticipatory in nature. The level of control closest to the process is the basic regulatory control level. Good performance of this level is crucial for the benefits of higher levels of control.

Level 2 (safety, environment, and equipment protection) includes activities such as alarm management and emergency shutdowns. While software implements the tasks shown, there is also a separate hardwired safety system for the plant. Level 1 (process measurement and actuation) provides data acquisition and on-line analysis and actuation functions, including some sensor validation. Ideally there is bidirectional communication between levels, with higher levels setting goals for lower levels and the lower levels communicating constraints and performance information to the higher levels. The time scale for decision making at the highest level (planning and scheduling) may be of the order of months, while at lower levels (eg, process control), decisions affecting the process can be made frequently, eg, in fractions of a second.

3. Instrumentation

3.1. Components of a Control Loop. Instrumentation, which provides the direct interface between the process and the control hierarchy, serves as the fundamental source of information about the process state and the ultimate means by which corrective actions are transmitted to the process. Figure 6 illustrates the hardware components of a typical modern digital control loop. The function of the process measurement device is to sense the value, or changes in value, of process variables. The choice of a specific device typically requires considerations of the specific application, economics, and reliability requirements. The actual sensing device may generate a physical movement, pressure signal, millivolt signal, etc. A transducer transforms the measurement signal from one physical or chemical quantity to another, eg, pressure to milliamps. The transduced signal is then transmitted to a control room through the transmission line. The transmitter is therefore a signal generator and a line driver. Often the transducer and the transmitter are contained in the same device. Most modern control equipment require a digital signal for displays and control algorithms, thus the analogue-to-digital converter (ADC) transforms the transmitter analog signal to a digital format. Because the ADCs may be relatively expensive if adequate digital resolution is required, the incoming digital signals are usually multiplexed. Prior to sending the desired control action, which is often in a digital format, to the final control element in the field, the desired control action is usually transformed by a D/A converter (DAC) to an analog signal for transmission. The DACs are relatively inexpensive and are not normally multiplexed. Widespread use of digital control technologies has made ADCs and DACs standard parts of the control system.

Once the desired control action has been transformed to an analog signal, it is transmitted to the final control element over the transmission lines. However, the final control element's actuator may require a different type of signal and

thus another transducer may be necessary. Many control valve actuators utilize a pressure signal so a current–pressure (I/P) transducer is used to provide a pressure signal to the actuator.

3.2. Process Measurements. The most commonly measured process variables are temperatures, flows, pressures, levels, and composition (see TEMPERATURE MEASUREMENT, FLOW MEASUREMENT; LIQUID-LEVEL MEASUREMENT; PRESSURE MEASUREMENT). When appropriate, other physical properties are also measured. The selection of the proper instrumentation for a particular application is dependent on factors such as the type and nature of the fluid or solid involved; relevant process conditions; rangeability, accuracy, and repeatability required; response time; installed cost; and maintainability and reliability. Various handbooks are available that can assist in selecting sensors for particular applications (9,11,12). Table 1 lists measurement options for temperature, flow, pressure, level and composition (1).

General Considerations for Measurements. In the selection of a measurement device, the required measurement range for the process variable must lie entirely within the instrument's range of performance. Accuracy, repeatability, or some other measure of performance are appropriate specifications depending on the application. Where closed-loop control is to be implemented, speed of response must be included as a specification. Data available from the manufacturers provide baseline conditions on reliability. Previous experience with the measurement device is very important. Materials of construction are selected so that the instrument must withstand the process conditions, such as operating temperatures, operating pressures, corrosion, and abrasion. For some applications, seals or purges may be necessary. For the first installation of a specific measurement device at a site, training of maintenance personnel and purchases of spare parts might be necessary. The potential for releasing process materials to the environment must be evaluated. Exposure to fugitive emissions for maintenance personnel is important when the process fluid is corrosive or toxic. If the measurement device is not inherently compatible with possible exposure to hazards, suitable enclosures must be purchased and included in the installation costs.

Instrument accuracy refers to the difference between the measured value and the true value of the measured variable. Since the true value is never known, accuracy usually refers to the difference between the measured value and a standard value of the measured variable. For process measurements, accuracy is usually expressed as a percentage of the span of the measured variable. Repeatability refers to the difference between measurements when the process conditions are the same. Repeatability is a very important factor for process control, because the main objective of regulatory control is to maintain uniform process operation. For process measurements it is important to distinguish between accuracy and repeatability. Some applications depend on the accuracy of the instrument, but other applications depend on repeatability. Excellent accuracy implies excellent repeatability; however, an instrument can have poor accuracy but excellent repeatability. In some applications this is acceptable.

Manufacturers of measurement devices always state the accuracy of the instrument. However, these statements specify reference conditions at which the measurement device will perform with the stated accuracy, with tempera-

ture and pressure most often appearing in the reference conditions. When the measurement device is applied at other conditions, the accuracy is affected. Manufacturers usually provide statements indicating how accuracy is affected when the conditions of use deviate from the reference conditions. Although appropriate calibration procedures can minimize some of these effects, rarely can they be totally eliminated. It is quite possible that a measurement device with a stated accuracy of 0.25% of span at reference conditions may only provide measured values with accuracies of 1% or less. Microprocessor-based measurement devices usually provide better accuracy than the traditional electronic measurement devices. In practice most attention is given to accuracy when the measured variable is the basis for billing a customer. Whenever a measurement device provides data for real-time optimization, accuracy is very important.

Temperature. Temperature sensor selection and installation should be based on the process-related requirements of a particular situation, ie, temperature level and range and process environment. For example, if the average temperature of a flowing fluid is to be measured, mounting the device nearly flush with the internal wall may cause the measured temperature to be affected by the wall temperature and the fluid boundary layer. Thermocouples are the most widely used means of measuring process temperatures and are based on the Seebeck effect. The emf developed by the hot junction is compared to the emf of a reference or cold junction, which is held at a constant reference temperature or has compensation circuitry. The difference between the hot junction and the reference junction temperature is thus determined. Depending on the temperature range and temperature level, various combinations of metals are used by the thermocouple, eg, Chromel/Alumel (Type K), iron/Constantan (Type J), and platinum–10% rhodium/platinum (Type S). Because thermocouple emfs are low level signals, it is important to prevent the contamination of the signal by stray currents or noise resulting from the proximity to other electrical devices and wiring.

Thermocouples are placed within protecting tubes, called thermowells, for protection against mechanical damage, vibration, corrosion, and stresses owing to flowing fluids. These thermowells impact the speed of response of the thermocouple by placing an additional lag in the control loop. Special thermowell designs do exist, which minimize this added lag. In hostile process environments where the reliability of a temperature measurement is a concern, multiple temperature sensors are sometimes used in conjunction with a majority voting system which can be implemented in software or hardware. Where exceptional accuracy and repeatability are required, resistance thermometry detectors (RTDs) are sometimes used although these are more expensive than thermocouples. The RTDs are based on the principle that as the temperature increases, the electrical resistance of conductors also increases. The RTDs can experience many of the same problems as thermocouples, so considerations such as thermowells and protection from electrical noise contamination are also appropriate in the case of RTDs.

Pyrometers are mostly used at very high temperatures ($>700^{\circ}\text{C}$) and estimate the temperature by measuring radiation emitted from the object whose temperature is to be determined. This can be done at several ranges of wavelengths, depending on the pyrometer type in order to achieve a high level of accu-

racy for the measurement. Pyrometers can measure higher temperatures than thermocouples or resistance thermometers.

Flow. The principal types of flow rate sensors are differential pressure, electromagnetic, vortex, turbine, and coriolis (13). Orifice plates and Venturi-type flow tubes are the most popular differential pressure flow rate sensors, where the pressure differential measured across the sensor is proportional to the square of the volumetric flow rate. Orifice plates are relatively inexpensive and are available in many materials to suit particular applications. This type of sensor is generally preferred for measuring gas and liquid flows. However, orifice plates have a relatively high unrecoverable pressure drop and limited range. Investment cost for a Venturi-type flow tube is generally higher than for an orifice plate for the same application but the accuracy is better. The higher unrecoverable pressure drop of the orifice plate sometimes dictates the use of a Venturi-type flow tube because of the overall cost. The proper installation of both orifice plates and Venturi-type flow tubes requires a length of straight pipe upstream and downstream of the sensor. The pressure taps and connections for the differential pressure transmitter should be located so as to prevent the accumulation of vapor when measuring a liquid and the accumulation of liquid when measuring a vapor.

Magnetic flow meters are sometimes utilized in corrosive liquid streams or slurries where a low unrecoverable pressure drop and high rangeability is required. The fluid is required to be electrically conductive. Magnetic flow meters, which use Faraday's law to measure the velocity of the electrically conductive liquid, are relatively expensive. Their use is therefore reserved for special situations where less expensive meters are not appropriate. Installation recommendations usually specify an upstream straight run of five pipe diameters, keeping the electrodes in continuous contact with the liquid.

Vortex meters have gained popularity where the unrecoverable pressure drop of an orifice meter is a concern. The vortex shedding meter is based on the principle that fluid flow about a bluff body causes vortices to be shed from alternating sides of the body at a frequency proportional to the fluid velocity. The vortex shedding meter can be designed to produce either a linear analog or a digital signal. The meter range, depending on process conditions, is relatively large (10:1). Because of the lack of moving parts and the absence of auxiliaries such as additional manifolds or valves, the reliability and safety of this meter are relatively good. The main restrictions on its use are (1) it should not be used for dirty or very viscous fluids; (2) the Reynolds number should be $>10,000$ but less than a process condition-dependent maximum set by cavitation, compressibility, and unrecoverable pressure drop; and (3) meters in pipes >20 cm (8 in.) in diameter tend to have limited applicability because of relatively high cost and limited resolution. Similar to other flow meters, the vortex shedding meter requires a fully developed flow profile, and therefore a run of straight pipe.

The trend in the chemical process industries is towards increased usage of mass flowmeters that are independent of changes in pressure, temperature, viscosity, and density (14). These devices include the thermal mass meter and the coriolis meter. Thermal mass meters are widely used in semiconductor manufacturing for control of low gas flow rates (called mass flow controllers, or MFCs).

The MFCs measure the heat loss from a heated element, which varies with flow rate, with an accuracy of $\pm 1\%$. Coriolis meters use a vibrating flow loop that undergoes a twisting action due to the coriolis effect. The amplitude of the deflection angle is converted to a voltage that is nearly proportional to the liquid mass flow rate, with an accuracy of $\pm 0.5\%$. Sufficient space must be available to accommodate the flow loop and pressure losses of 10 psi should be allowable. Capacitance probes measure the dielectric constant of the fluid and are useful for flow measurements of slurries and other two-phase flows.

Pressure. There are three distinct groups of pressure measurement devices. One is based upon the measurement of the height of a liquid column, another is based on the measurement of the distortion of an elastic pressure chamber and a third encompasses electrical sensing devices.

In liquid column pressure measuring devices the pressure is balanced against the pressure exerted by a column of a liquid with known density. The height of the liquid column directly correlates to the pressure to be measured. Most forms of liquid-column measurement devices are called manometers. Elastic element pressure measuring devices are those in which the measured pressure deforms an elastic material. The magnitude of the deformation is approximately equivalent to the applied pressure. Different types of elastic element measuring devices include Bourdon tubes, bellows, and diaphragms. Electrical sensing devices are based on the fact that when electrical conductors are stretched elastically, their length increases while the diameter decreases. Both of these dimensional changes result in an increase in electrical resistance of the conductor. Strain gauges and piezoelectric transducers are examples of electrical pressure sensing devices.

To avoid maintenance problems, the location of pressure measurement devices must be carefully considered to protect against vibration, freezing, corrosion, temperature, overpressure, etc. For example, in the case of a hard-to-handle fluid, an inert gas is sometimes used to isolate the sensing device from direct contact with the fluid. Optical fiber can be used for pressure measurement in high temperature environments (15).

Liquid Level. The location of a phase interface between two fluids is referred to as level measurement. Most often this is applied to liquid–gas interfaces, but interfaces between two liquids are not uncommon. Level measurement devices can be classified as either float-actuated devices, displacer devices, head devices, and fluid characteristics. The most widely used devices for measuring liquid levels involve detecting the buoyant force on an object or the pressure differential created by the height of liquid between two taps on the vessel. Consequently, care is required in locating the tap. Other less widely used techniques utilize concepts such as the attenuation of radiation, changes in electrical properties, and ultrasonic wave attenuation.

Chemical Composition. Onstream analyzers measure various physical and chemical properties as well as component compositions (16,17). Compared to most other instrumentation, analyzers are relatively expensive, more complex and sensitive, and require more regular maintenance by trained personnel. Therefore, the expense for onstream analyzers needs to be justified by the benefits generated through their use. Improvements in analyzer technology, digital control systems, and process control technology have led to increasing use of ana-

lyzers in closed-loop automatic control applications. The most common physical and chemical properties measured by analyzers include density, viscosity, vapor pressure, boiling point, flash point, cloud point, moisture, heating value, thermal conductivity, refractive index, and pH. Some of these analyzers are continuous while others are discrete.

In order to obtain quantitative composition measurements, specific instruments must be chosen depending on the nature of the species to be analyzed. Measuring a specific concentration requires a unique chemical or physical attribute. In infrared (IR) spectroscopy, the vibrational frequency of specific molecules like CO and CO₂ can be probed by absorbing electromagnetic radiation. Ultraviolet (UV) radiation analyzers operate similarly to IR analyzers in that the degree of absorption for specific compounds occurs at specific frequencies and can be measured. Magnetic resonance analysis (formerly called nuclear magnetic resonance) uses magnetic moments to discern molecular structure and concentrations.

Significant advances have occurred during the past decade to obtain lower cost measurements, in some cases miniaturizing the size of the measurement system in order to make on-line analysis feasible and reducing the time delays that often are present in analyzers. Recently, chemical sensors have been placed on microchips, even those requiring multiple physical, chemical, and biochemical steps (such as electrophoresis) in the analysis. This device has been called “lab-on-a-chip” (18). The measurements of chemical composition can be direct or indirect, the latter case referring to applications where some property of the process stream is measured (such as refractive index) and then related to composition of a particular component.

In gas chromatography (GC), usually the thermal conductivity is used to measure concentration. The GC can measure many components in a mixture at the same time, whereas most other analyzers can only detect one component, hence GC is very popular. A gas sample (or a vaporized liquid sample) is carried through the GC by an inert gas, and components of the sample are separated by a packed bed. Because each component has a different affinity for the column packing, it requires a distinct time to pass through the column, allowing individual concentrations to be measured. Typically all components can be analyzed in a five to ten minute time period (although miniaturized GCs are faster). The GC can measure concentrations ranging from parts per billion (ppb) to tens of percent, depending on the compound (17).

Mass spectroscopy (MS) determines the partial pressures of gases in a mixture by directing ionized gases into a detector under a vacuum (10^{-6} Torr), and the gas-phase composition is then monitored more or less continuously based on the molecular weight of the species (17). Sometimes GC is combined with MS in order to obtain a higher level of discrimination of the components present. Other on-line analyzers types include UV photometer, UV fluorescence, chemiluminescence, IR, and paramagnetic. Fiber-optic sensors are attractive options (although at a higher cost) for acquiring measurements in harsh environments such as high temperature or pressure. The transducing technique used by these sensors is optical and does not involve electrical signals, so they are immune to electromagnetic interference. Raman spectroscopy uses fiber-optics

and involves pulsed light scattering by molecules. It has a wide variety of applications in process control (15).

Key considerations in using an analyzer for closed-loop control are repeatable, reliable analyzer measurements and an appropriate analyzer system response time. In control applications at the supervisory control level and above, accuracy as well as repeatability is often required. On-line process analyzers may be placed into one of three categories. In the first, *in situ*, the analysis is continuous and the probe mounted directly in the process stream. In this category, the measurement can be treated similarly to other process measurements, as long as some additional care is taken owing to the reliability issues. In the second category, the analysis is continuous but the sample is not naturally in a form required by the analyzer. Thus, the sample must be conditioned. In the third category, the analyzer takes a period of time to analyze a discrete sample that must usually be conditioned. Also, a sample-and-hold circuitry is required to keep an output signal at its last value between analyses. In the latter two categories, the analysis introduces dead time into the control loop. Therefore, to achieve good closed-loop control in these instances, the additional dead time introduced by the measurement should be minimized and the control algorithm should contain some form of dead-time compensation.

Most analyzers require sample takeoff, sample transport, sample conditioning and preparation, analysis, and sample return or disposal. These subsystems need to be carefully designed to ensure that the analyzer meets its intended purpose and is reliable. The sample transport subsystem is required because analyzers are often placed in the protected, controlled environment of an analyzer shelter remote from the sample takeoff. The sample location should be selected such that the sample is representative, the complexity of the sample conditioning subsystem is minimized, and the equipment is accessible. It is therefore preferable that the sample is a single phase, relatively clean, and that the takeoff location does not add significant dead time to the control loop. Furthermore, the pressure at the sample takeoff should be such that the pressure differential between the sample capture and sample return point is adequate to drive the sample through the fast loop at a sufficient velocity to avoid the need for a sample pump. The purpose of the fast loop is to bring the sample to the vicinity of the analyzer takeoff at a high velocity.

Many factors need to be considered for the proper selection, design, and specification of an analyzer, its sample handling subsystems, and its intended use within a process control strategy and hierarchy, including the interface to the control system. As a rule of thumb, the measurement effective dead time plus lag time should be no greater than one-sixth of the time constant of the process and other elements in the control loop if possible. Consequently, most analyzers are utilized at the supervisory control level and above. Thus, for example, a process gas chromatograph system having a cycle time, including the sample handling subsystems, of five minutes is required for a supervisory or local optimization control loop where the effective time constant is thirty minutes. Therefore, only *in situ* analyzers should be considered at the regulatory control level. Furthermore, there is usually a direct relationship between reduced analyzer system cycle time and increased analyzer system cost.

3.3. Signal Transmission and Conditioning. A wide variety of physical and chemical phenomena are used to measure the many process variables required to characterize the state of a process. Because most processes are operated from a control room, these values must be available there. Hence, the measurements are usually transduced to an electronic form, most often 4–20 mA, and then transmitted to a remote terminal unit and then to the control room (see Fig. 3). Wherever transmission of these signals takes place in twisted pairs, it is especially important that proper care is taken so that these measurement signals are not corrupted owing to ground currents, interference from other electrical equipment and distribution, and other sources of noise. Most instrument and control system vendors publish manuals giving advice and instructions for installation and engineering practices for the proper grounding, shielding, cable selection, cable routing, and wiring for control systems. The importance of these considerations should not be underestimated (16).

3.4. Final Control Elements. Good control at any hierarchical level requires good performance by the final control elements in the next lower level. At the higher control levels, the final control element may be a control application at the next lower control level. However, the control command must ultimately affect the process through the final control elements at the regulatory control level, eg, control valves.

Control Valves. Material and energy flow rates are the most commonly selected manipulated variables for control schemes. Thus, good control valve performance is an essential ingredient for achieving good control performance. A control valve consists of two principal assemblies, a valve body and an actuator (Fig. 7). Good control valve performance requires a consideration of the process characteristics and requirements such as fluid characteristics, range, shut-off, and safety, as well as control requirements, eg, installed control valve characteristics and response time. The proper selection and sizing of control valves and actuators is an extensive topic in its own right (2,9,12,19). Many control valve vendors provide computer programs for the sizing and selection of valves.

The valve body, the portion that contains the process fluid, consists of the internal valve trim, packing, and bonnet. The internal trim determines the relationship between the flow area and the stem position, which is usually proportional to the air signal. This relationship, referred to as the inherent valve characteristic, is often classified as linear, equal percentage, or quick opening. Actual valve characteristics generally fall within these three classifications. For a globe-type valve, the internal trim consists of plug, seat ring, plug stem, plug guide, and in some instances a cage. In rotary style valves, such as a ball or butterfly type, the internal trim consists of a ball or vane, seal ring, rotary shaft, and bushings. Although the internal trim fixes the relationship between flow area and stem position, the relationship between the control air signal and the flow area may be modified by the use of different cams in the case of rotary style valves, or via software in the case of smart valves or control systems having software capabilities.

The actuator provides the force to move the stem or rotary shaft in response to changes in the controller output signal. Actuators must provide sufficient motive force to overcome the forces developed by the process fluid and the valve assembly; be responsive in quick and accurate positioning corresponding

to changes in the control signal; and be responsive in automatically positioning the valve in a safe position when a failure (eg, instrument air or electrical, a safety interlock, or a shutdown) has occurred. Most actuators are either of the spring–diaphragm, piston, or motor type.

The general approach to specifying a control valve involves selecting a valve body type, trim characteristics, size, and material based on the process fluid characteristics, desired installed valve characteristics, process conditions, and process requirements. The actuator is then specified based on the valve selected, process flow conditions, and required speed of response. Meeting the requirements of a safe fail-position involves considering both the valve and the actuator. Figure 7 is configured as an air-to-open or fail-closed (F/C) valve. An air-to-close (fail-open) requirement may be met by using a spring-diaphragm actuator below the diaphragm and providing the air supply above the diaphragm.

Various accessories can be supplied along with the control valves for special situations. Positioners ensure that the valve stem is accurately positioned following small or slowly changing control signals or where unbalanced valve forces exist. The valve positioner may be mounted on the side of the valve actuator, and can reduce the valve deadband from about $\pm 5\%$ to $\pm 0.5\%$, a significant enhancement. Details on mechanical valve positioners are given in Perry's Handbook (9). A digital valve positioner is useful in computer control because the normal sampling interval of one second is not fast enough for most flow control loops. When a valve positioner is incorporated into a control system, it effectively becomes a cascade control system. Boosters, which are actually pneumatic amplifiers, can increase the speed of response or provide adequate force in high pressure applications. Limit switches are sometimes included to provide remote verification that the valve stem has actually moved to a particular position.

In addition to control valves which regulate the flow of one stream, process facilities sometimes use modulating three-way control valves to adjust simultaneously the flows of two streams, either to divert one stream into two streams or to combine two streams into one. The stable operation of these three-way control valves requires that the flow tends to open the plugs. The most common uses of three-way control valves are in controlling the heat transferred by regulating the flow through and around a piece of heat-transfer equipment and in controlling the blending of two streams. The choice of using a diverting or mixing service in a heat-transfer application is determined by pressure and temperature consideration. The upstream valve location (diverting service) is preferred if there is no overriding consideration. If there is a change of phase involved in the heat-transfer equipment, then a diverting valve should be used. Three-way control valves should not be used in services if the temperature is high ($>260^\circ\text{C}$), if there is a high temperature differential ($>150^\circ\text{C}$), or if there is a high pressure or pressure differential. If any of these conditions exist, two single-seated two-way valves should be used to implement the bypass control strategy even though this is a more expensive initial investment.

Adjustable Speed Pumps. Instead of using fixed speed pumps and then throttling the process with a control valve, an adjustable speed pump can be used as the final control element. In these pumps speed can be varied by using variable-speed prime movers such as turbines or electric motors. Adjustable speed pumps offer energy savings as well as performance advantages over throttling

valves. One of these performance advantages is that, unlike a throttling valve, an adjustable pump does not contain dead time for small amplitude responses. Furthermore, nonlinearities associated with friction in the valve are not present in electronic adjustable speed pumps. However, adjustable speed pumps do not offer shutoff capability like control valves and extra flow check valves or automated on/off valves may be required.

Other Final Control Elements. Devices other than control valves and adjustable speed pumps are also used as final control elements. Dampers are used to control the flow of gases and vapors. Louvers are also used to control the flow of gases, eg, flow of air in air fin coolers. Feeders such as screw feeders, belt feeders, and gravimetric feeders are often used to control the flow of solids. Metering pumps and certain feeders combine the functions of the measurement and final control element in some control loops.

4. Process Dynamics and Mathematical Models

A thorough understanding of the time-dependent behavior of physical/chemical processes is required in order to instrument and control a modern plant. This in turn requires an appreciation of how mathematical tools can be employed in analysis and design of process control systems. Due to this, several underlying mathematical principles that are utilized in automatic control are presented in this section.

4.1. Physical Models versus Empirical Models. Most design procedures for process controllers require a model of some form. These models can be either primarily based on first principles, resulting in a physical model, or they can be derived from process data that in turn will lead to an empirical model. A physical model is most often based upon conservation laws of mass, energy, and momentum, typically resulting in a set of differential and algebraic equations that need to be solved simultaneously. Contrary to this, empirical models are derived from experimental or process data. In order to build such a model, the model structure has to be postulated and then the model parameters are fitted to the available data. For the development of both physical and empirical models, the most time-consuming step normally involves verification of their accuracy for predicting future plant behavior. A semi-empirical model combines aspects of both physical and empirical models, in that some physical parameters (a heat-transfer coefficient in an energy balance, for instance) may be estimated from available data using numerical optimization. In practice, most models are semi-empirical in some form or another, because physical models will contain process parameters that were fitted from data while the structure for empirical models is usually chosen such that it can represent the physical behavior of the process.

4.2. Simulation of Dynamic Models. Because process control normally is concerned with time-dependent process behavior, the model employed by a process control system needs to include a dynamic component. A variety of different model types exists consisting of sets of ordinary differential equations, algebraic equations, or even partial differential equations. For the purpose of this text only the most common form consisting of a set of nonlinear ordinary

differential equations will be considered:

$$\dot{X}(t) = f(X(t), u(t), d(t)) \quad (1)$$

Here X refers to the vector of state variables whose value will be calculated as a function of time from the differential equations, $u(t)$ is the manipulated variable, and $d(t)$ represents the disturbance. The controlled variable Y can be one of the state variables or it can be related to the state variables through some algebraic function. In most cases analytical solutions of such models cannot be determined, so that numerical techniques are required for their solution.

While there exist several approaches for numerically solving differential equations, all of them are based on discretizing the model equations in some form. One common approach is to approximate the differential equations with finite difference equations.

$$X(t + \Delta t) = X(t) + \Delta t \cdot f(X(t), u(t), d(t)) \quad (2)$$

Given the initial conditions for X , the time step Δt , and values for the inputs $u(t)$ and $d(t)$, this system of equations can be solved for $X(t + \Delta t)$. The simplest algorithm to perform this computation holds the time step constant during integration and is called the Euler integration method. However, in order to also achieve good accuracy, in addition to simplicity of the approach, very small time steps have to be chosen. Due to this, the Euler method is not necessarily the best approach for numerical integration of nonlinear differential equations and a number of more sophisticated approaches are available that allow much larger step sizes to be taken. In general these approaches use a varying step size for each integration step resulting in more efficient computations. One widely used approach is the fourth-order Runge Kutta method (20).

4.3. Laplace Transforms, Transfer Functions, and Block Diagrams.

When a process is represented by a set of nonlinear differential equations, it is a nontrivial task to analyze or to design a controller with specific properties. Making some simplifying assumptions is a requirement for controller design and implementation. Since a process is usually operated within a certain neighborhood of its normal operating point (steady state), the process model can be closely approximated by a linearized version of the model. This has the benefit that linear models permit the use of more convenient and compact methods for representing process dynamics, one of which is called Laplace transforms. The main advantage of Laplace transforms is that they convert differential equations into algebraic ones by a domain change resulting in a system which is easier to solve. The resulting Laplace-transformed system is a set of algebraic equations in the new variable s , called the Laplace variable. The Laplace transform is given by

$$F(s) = L[f(t)] = \int_0^{\infty} f(t)e^{-st} dt \quad (3)$$

where $F(s)$ is the symbol for the Laplace transform, $f(t)$ is some function of time, and L is the Laplace operator, defined by the integral. Tables of Laplace trans-

forms are well documented for common functions (1). With such transforms the concept of transfer functions and block diagrams can now be introduced.

Using Laplace transforms a linear differential equation with a single input u and single output y can be converted into a transfer function as follows:

$$Y(s) = G(s)U(s) \quad (4)$$

where

$U(s)$ is the transform of the input variable $u(t)$

$Y(s)$ is the transform of the output variable $y(t)$

$G(s)$ is the transfer function, obtained from transforming the differential equation

The transfer function $G(s)$ describes the dynamic characteristic of the process. For linear systems it is independent of the input variable and so it can readily be applied to any time-dependent input signal.

As an example, the first order differential equation

$$\tau \frac{dy(t)}{dt} + y(t) = Ku(t) \quad (5)$$

can be Laplace transformed to

$$Y(s) = \frac{K}{\tau s + 1} U(s) \quad (6)$$

Note that the parameters K and τ , known as the process gain and time constant, respectively, map into the transfer function as unspecified parameters. These parameters can then be estimated for controller design or simulation purposes. For more details on development of various transfer functions, see Seborg and co-workers (1).

When several pieces of equipment are connected, their dynamic behavior can be described by a transfer function for each process. It is possible to represent the overall process with a block diagram where each transfer function is placed in the corresponding block. Each block then describes how changes in the input variables of the block will affect the output variables of the block. Since it is possible to interconnect blocks, complex systems can be decomposed and represented in block diagram form. One example of such a block diagram is shown in Figure 1 where the figure contains the most important elements of a control system.

4.4. Fitting Dynamic Models to Experimental Data. When a model is expressed in transfer function notation, it may still include unspecified model parameters that appear in the mathematical expressions. One example of this is the gain K and the time constant τ in the above equations. Numerical values for these parameters have to be determined for controller design or for simulation purposes. Several different methods for the identification of model parameters in transfer functions are available. The most common approach is to

perform a step test on the process and collect the data along the trajectory until it reaches steady state. Figure 8 shows a typical step response for a linear process.

In order to identify the parameters, the form of the transfer function needs to be postulated and it needs to be determined which are the parameters of the transfer function that should be estimated. These parameters can then be calculated by a nonlinear least squares approach.

For illustration purposes, assume that y_i are the experimentally determined values and $y(t_i)$ refers to the calculated outputs of a system generated by the differential equation; then the objective function to be minimized is

$$\min_{K, \tau} \sum_{i=0}^n [y_i - y(t_i)]^2 \quad (7)$$

This approach will compute values for the parameters K and τ such that the predicted values of the outputs lie close to the experimentally determined ones.

One operational problem that step forcing causes is the fact that the process under study is moved away from its steady-state operating point, disturbing normal production in an operating plant. In order to address this, alternative methods like pulse testing and pseudo random binary signals (PRBS) have been developed. In pulse testing, a step is introduced for a certain period of time and then the system is forced back to its original steady state. While this method also excites the system, it does so only for a short period of time. The PRBS utilizes a series of input pulses of fixed height (either $+\Delta u$ or $-\Delta u$), but of random duration. The advantage PRBS has over conventional pulse testing is that the forcing can be concentrated on specific frequency ranges that are important for control system design, by changing the frequency of the input changes.

Generally, a trade-off exists between exciting the system in such a way that the most information can be gained and disturbing regular plant operations. The best compromise between these two objectives has to be determined on a case-by-case basis.

5. Feedback Control Systems.

Measurements of the controlled variable are available in many industrial control problems. Specifically, this is the case when temperatures, pressure, or flows are to be controlled. In these situations the controlled variable can be directly measured and the manipulated variable is adjusted via a final control element. As has been presented in previous sections, a feedback controller will take action when the controlled variable deviates from its set point, as detected by the non-zero value of the error signal. The focus of this section is on commonly used feedback controllers in commercial applications followed by a discussion of their impact on process performance.

5.1. On/Off Control. A variety of different feedback controllers exists. The simplest controller can only exhibit two settings and is called an on/off controller. One common application of this type of controller is for temperature con-

trol for home heating and air conditioning systems. The output of this controller is either at its maximum or its minimum value, depending on the sign of the error e :

$$u(t) = \begin{cases} u_{\max}, & e \geq 0 \\ u_{\min}, & e < 0 \end{cases} \quad (8)$$

While this type of controller is simple, it is seldom used in industrial plants. This is due to two reasons: it causes excessive wear in the control valve and it does not provide adequate disturbance rejection. These problems can be addressed with controllers that offer more flexibility than the simple on/off controller.

5.2. Proportional Control. Another type of simple controller is the proportional controller. These controllers offer more flexibility than the on–off controller because the manipulated variable is related not just to the sign of the error but also to its magnitude and is given by the following expression

$$u(t) = \bar{u} + K_c e(t) \quad (9)$$

where \bar{u} is the bias, K_c is the proportional constant, and $e(t)$ is the error that is being transmitted to the controller. The proportional constant serves as a tuning parameter for this controller. Choosing a larger value of K_c results in a more aggressive controller that in turn will usually lead to a faster response and less steady-state error (offset), as long as the closed-loop system remains stable. The effect that different values of the proportional constant have on the closed-loop behavior of a process are discussed in more detail in Section 5.6. However, no matter how large the tuning parameter is chosen, it is not possible to completely eliminate offset due to constant load disturbances with a proportional controller for most processes. If a performance requirement specifies that no offset can be present then a controller with integral action needs to be implemented (see Section 5.3).

It should be pointed out that the input–output behavior of an actual proportional controller has upper and lower bounds; the output will saturate when the control limits are reached. Standard limits on the controller output are 3–15 psig for pneumatic controllers, 4–20 mA for electric controllers, and 0–10 VDC for digital controllers.

5.3. Proportional plus Integral Control. Integrating action needs to be included in the control loop if an offset-free response in the presence of constant load disturbances or for set point changes is required. If the process does not exhibit integrating behavior itself then it is possible to implement a proportional plus-integral controller to achieve the desired performance. The PI controller is given by the following equation

$$u(t) = \bar{u} + K_c \left(e(t) + \frac{1}{\tau_I} \int_0^t e(t') dt' \right) \quad (10)$$

where the controller includes proportional as well as integrating action. This also results in a controller with two tuning parameters, which are the proportional

constant K_C and the integral time constant τ_I . While the effect of K_C is similar to what has been described in Section 5.2, the magnitude of τ_I determines how much integral action is added to the proportional controller.

As with any other method there are advantages as well as disadvantages associated with integral action in a controller. The integral action will eliminate offset for constant load disturbances but it can potentially lead to a phenomenon known as reset windup. This will be the case if a sustained error occurs resulting in a large integral term causing the controller output to saturate. Reset windup is a common phenomenon during start-up of batch processes, after large set point changes or resulting from large sustained load disturbances which are beyond the range of the manipulated variable. Due to this, many commercially available controllers provide a feature called anti-reset windup, where the integral mode is disabled when the controller is at its saturation limit (21). PI controllers make up the vast majority of controllers that are currently used in the chemical process industries.

5.4. Proportional plus Integral plus Derivative (PID) Control. One disadvantage of a PI controller is that the integral action might cause it to react more sluggishly than a proportional controller. If it is important to achieve a faster response that should be offset-free then this can be accomplished by including both derivative and integral action in the controller. In order to anticipate the future behavior of the error signal, a PID controller computes the rate of change of the error, thus the directional trend of the error signal influences the controller output. The manipulated variable is then computed by the following expression:

$$u(t) = \bar{u} + K_c \left(e(t) + \frac{1}{\tau_I} \int_0^t e(t') dt' + \tau_D \frac{de(t)}{dt} \right) \quad (11)$$

The PID controller of equation 11 contains three tuning parameters because the derivative mode adds a third adjustable parameter τ_D . Due to the addition of derivative action the process response time can be decreased. However, if the process measurement is noisy, the value of the derivative of the error may change rapidly and derivative action will amplify the noise. As a result of this, derivative action is seldom used in flow controllers because flow dynamics are relatively fast and flow measurements tend to be noisy. When the tuning parameter τ_D is chosen to be equal to zero then the PID controller will simplify to a PI controller that has been described in the previous subsection.

5.5. Digital PID. While many controllers have traditionally been analogue PI/PID controllers, the trend toward digital control systems has also had an influence on controller implementation. In many modern process plants, the analogue PI/PID controllers have been replaced by their digital counterparts. These often form a part of a digital control system. The discrete form of the PID controller equation is given by

$$u_k = \bar{u} + K_c \left(e_k + \frac{\Delta t}{\tau_I} \sum_{i=0}^k e_i + \tau_D \frac{e_k - e_{k-1}}{\Delta t} \right) \quad (12)$$

where Δt is the sampling period for the control calculations and k represents the current sampling time. If the process and the measurements permit to choose the sampling period Δt to be small then the behavior of the digital PID controller will approximate the one for an analog PID controller.

5.6. Open-Loop versus Closed-Loop Dynamics. Open-loop dynamics refers to the behavior of a process if no controller is acting on it. Similarly, if the controller is turned off by setting the proportional constant K_C to zero, the system will exhibit open-loop behavior and the system's dynamics are solely determined by the process. Therefore, it is not possible to reach a new set point for a plant in open-loop unless the input is changed manually. It is also not possible to reject disturbances when the process is operated without a controller. Some plants will also exhibit unstable behavior and cannot be operated in open-loop, either.

The purpose of using closed-loop control is to achieve a desired performance for the system. This can result in the system being stabilized, in a faster system response to set point changes, or in the ability to reject disturbances. The choice of the controller type as well as the values of the controller tuning parameters influence the closed-loop behavior. This is illustrated in the following example where the effect of the proportional constant on the closed-loop dynamics of the system is investigated.

Assume that a stable linear system is disturbed by a constant load change. A proportional-only controller in the form of equation 9 is then applied to the system. If the controller is turned off ($K_C = 0$) the system exhibits open-loop dynamics. This is shown in Figure 9, where the system will have steady-state offset resulting from this output disturbance. When the proportional constant is increased, the system responds faster and exhibits a smaller steady-state offset. Both of these are desirable effects, because the negative effects that disturbances have on the system need to be minimized. However, the controller gain cannot be increased indefinitely and result in a more desirable system response. This result can also be seen in Figure 9, where a further increase in the controller gain results in a trajectory that starts to oscillate around the desired operating point. There exists an ultimate controller gain K_U at which the oscillations will not die out and continue indefinitely. This behavior is undesirable and should be avoided. At even larger values for the proportional constant, the closed-loop system will be unstable.

For a controlled process one needs to find controller settings that result in a fast system response with little or no offset, while at the same time the system should be robust to changes in process characteristics. Finding the appropriate settings is called "tuning" the controller.

5.7. Controller Tuning and Stability. It can be concluded from the last example that finding optimum tuning parameters for a controller is an important task. Unsuitable parameters can result in not achieving the desired closed-loop performance (eg, slowly decaying oscillations or a slow acting process). It is also possible that a closed-loop process with a badly tuned controller might result in performance that is worse than for the open-loop case or that the process can even become unstable.

In order to address these issues, numerous approaches for controller tuning with varying degrees of complexity have been developed over the past 50 years

(1,3). One popular controller tuning method is based on process reaction curves, ie, the response of the open-loop system to a step input (eg, Figure 8). Key parameters such as dead time, time constant, and gain can be determined from the data resulting from a step response. These process parameters can then be used to calculate the controller parameters. One such method that has become increasingly popular is controller design based on internal model control (IMC) (21–23).

5.8. Mathematical Software for Process Control. A variety of different software packages is available that support the controller design, controller testing, and implementation process. Probably, the most widely used program for this purpose is MATLAB (The MathWorks, Inc.; www.mathworks.com), due to its flexibility. It allows to implement and test controllers by either solving differential equations, using Laplace transforms, or with block diagrams. MATLAB also provides a variety of routines that are commonly used for different controller design problems, eg, optimal control, nonlinear control, optimization, etc. One of the main advantages of MATLAB is that it is a programming language which provides control-related subroutines. This gives the process engineer flexibility with regard to the use of the software as well as how to extend or reuse already existing routines. It is also possible to exchange data with other software packages from within MATLAB.

6. Advanced Control Techniques

While the single-loop PI/PID feedback controller is satisfactory for many process applications, there are cases for which advanced control techniques can result in a significant improvement in closed-loop performance. These processes often exhibit one or more of the following phenomena:

- Slow dynamics.
- Time delays.
- Frequent disturbances.
- Multivariable interaction.

Since a large number of advanced control strategies are used in industry, the most important ones are briefly discussed below.

6.1. Feedforward Control. One of the disadvantages of conventional feedback control with large time lags or delays is that disturbances are not recognized until after the controlled variable deviates from its set point. However, if it is possible to measure the load disturbance directly then feedforward control can be applied in order to minimize the effect that this load disturbance will have on the controlled variable. The block diagram of the implementation of a feedforward controller is shown in Figure 10. Note that this is part of the block diagram presented earlier in Figure 1.

In addition to being able to measure the load disturbance, it is also required to determine a mathematical correlation for the effect that the load disturbance will have on the controlled variable in order to apply a feedforward controller.

The reason for this is that the feedforward controller will invert this model in order to cancel the effect that the disturbance will have. That way it is possible to reject a load disturbance before it has a visible impact on the controlled variable Y . The accuracy of the measurement of the disturbance variable as well as the quality of the model that describes the effect of D on Y are the main factors for the achievable performance of feedforward control.

A feedforward controller can be designed either based on the steady-state or dynamic behavior of the model. Assuming that perfect measurements and a perfect model were available, feedforward control could lead to perfect disturbance rejection, ie, the controlled variable would be kept precisely at its set point at all times. However, these are not realistic assumptions for an industrial setting and therefore feedforward control is usually combined with feedback control. This results in a controller implementation as was shown in Figure 1. For such a combination the controller tuning for the feedback loop can be performed independently of the feedforward controller, since the feedforward controller does not introduce instability in the closed-loop response. For more information on feedforward/feedback control applications, refer to Shinskey (24).

6.2. Cascade Control. Another possibility of controlling processes with multiple or slow-acting disturbances is to implement cascade control. The main idea behind cascade control is that more than just one controller is used to reject disturbances, instead a secondary controller is added to take action before the slow-acting disturbance will have an effect on the primary controlled variable. In order to achieve this the secondary controller also requires a secondary measurement point which needs to be located so that it recognizes the upset condition before the primary controlled variable is affected.

An example of such a control system is shown in Figure 11, where the reactor temperature is regulated by a cascade controller. The cascade control scheme consists of two controllers for this example: the master controller that regulates the temperature of the reactor content and a slave controller that controls the temperature of the jacket. This is in stark contrast to a single-loop feedback controller where only the reactor temperature is measured/controlled and therefore no control of the jacket temperature is possible. The main advantage of the cascade control strategy is that a second measured variable, in this case the jacket temperature, is located close to a potential disturbance and thus the controller will react faster to disturbances than a single-loop controller could.

Generally, a cascade control scheme consists of two nested loops: the inner loop that contains the secondary controller, and the outer loop that includes the primary controller. Additionally, the manipulated variable of the outer loop serves as the set point for the inner loop. The purpose of the inner loop is to quickly minimize the effect of the fast acting disturbance, whereas the primary controller eliminates offset and rejects slow disturbances. Because of this, the controller for the inner loop is usually chosen to be a proportional-only controller, since it responds quickly and there is no requirement for offset elimination in the inner loop. On the other hand, the primary controller needs to include integral action in order to eliminate offset.

Cascade control strategies are among the most popular process control strategies. Modern control systems have made their implementation and operation both easier from the standpoint of operations personnel, and cost effective as

they are implemented in software rather than hardwiring the connections. See Shinskey (24) and Seborg and co-workers (1) for a further discussion of this topic.

6.3. Selective and Override Control. Some processes have more controlled variables than manipulated variables. Such a situation does not allow an exact pairing of controlled and manipulated variables. A common solution is to use a device called a selector that chooses the appropriate process variable from among a number of valid measurements. The purpose of the selector is to improve control system performance as well as to protect equipment from unsafe operating conditions by choosing appropriate controlled variables for a specific process operating condition. Selectors can be based on either multiple measurement points, multiple final control elements, or multiple controllers.

One specific approach is auctioneering where the selector device chooses as its output signal the highest or lowest of two or more output signals. For example, a high selector can be used to determine the hot-spot temperature in a fixed bed chemical reactor: in a reactor in which exothermic catalytic reactions are occurring, the process can run away due to disturbances or changes in the reactor. Immediate action should be taken to prevent a dangerous rise in temperature. Multiple measurement points should be employed, because a hot-spot may potentially develop at one of several possible locations in the reactor. It is the purpose of the control scheme to monitor the highest temperature in the reactor and ensure that it does not go beyond a certain temperature limit. In order to address this, the selector will choose the highest temperature from the measurements and use it in a feedback loop to control the maximum reactor temperature (wherever it occurs). For this specific case, the output from the high selector is the input to the temperature controller. This approach minimizes the time required to identify a temperature excursion at some point in the bed.

A related approach is called an override, which uses high or low limits for process variables. One specific example of this is anti-reset windup in feedback controllers. Another example is a distillation column for which the heat input to the column reboiler has lower as well as upper limits. These limits are determined by physical constraints for operation of the column because the minimum level ensures that sufficient liquid remains on the tray, while the upper limit is determined by flooding.

Other types of selective systems employ multiple final control elements or multiple controllers. This is the case in applications where several manipulated variables are used to control a single process variable. Typical examples include the adjustment of both inflow and outflow from a chemical reactor in order to control reactor pressure or the use of both acid and base to control pH in waste-water treatment. For these application the selector has to choose which final control element should be used (25) in order to keep the controlled variable at its set point.

6.4. Adaptive Control and Autotuning. Operating conditions of a process can frequently change during plant operations. This does lead to the process behaving differently from the model that was used for controller design. Therefore, the controller does not have accurate knowledge of the process at the current operating point and might not be able to provide adequate disturbance rejection or set point tracking. One possibility to circumvent this is to use an

adaptive control system which will automatically adjust the controller parameters to compensate for changing process conditions (26).

Autotuning is a related method where the closed-loop system is periodically tested, and the test characteristics automatically determine new controller settings. An important procedure is to excite the system such that enough information becomes available to determine controller tuning parameters, while at the same time to not disturb regular plant operations.

A variety of adaptive controllers has been field-tested and commercialized. Many modern controllers have some type of autotuning feature. A common feature is that the process is placed in a controlled oscillation with very small amplitude, comparable to that of the noise level of the process. This is done via a relay-type step function with hysteresis. The autotuner is then used to identify the ultimate gain and period of the controlled cycle and automatically calculates K_C , τ_I , and τ_D using empirical tuning rules. It is also possible to implement gain scheduling with this controller, using several sets of PID controller parameters (1).

6.5. Fuzzy Logic Control. For many processes, it is very time consuming to determine accurate process models. However, at the same time it might be intuitive to get a rough estimate of how the manipulated variable should react to a process condition. For such a case, fuzzy logic controllers can offer an advantage over conventional PID controllers. The reason for this is that fuzzy controllers do not require an exact mathematical description of a process. Instead they do classify the controller inputs and output as belonging to one of several groups (ie, low, normal, and high). Fuzzy rules are then used to compute the output category from the given inputs. These rules either have to be provided by the control engineer or they have to be identified from plant operations by autotuning.

It is also possible to combine fuzzy logic controllers with neural networks in order to form neuro-fuzzy controllers (27). This type of controller can offer significant advantages over conventional PID when applied to nonlinear systems whose characteristics change over time.

6.6. Statistical Process Control. Statistical process control (SPC), also called statistical quality control (SQC), has found widespread application in recent years due to the growing focus on increased productivity. Another reason for its increasing use is that feedback control cannot be applied to many processes due to a lack of on-line measurements as is the case in many microelectronics fabrication processes. However, it is important to know if these processes are operating satisfactorily. While SPC is unable to take corrective action while the process is moving away from the desired target, it can serve as an indicator that product quality might not be satisfactory and that corrective action should be taken for future plant operations.

For a process that is operating satisfactorily the variation of product quality will fall within acceptable bounds. These bounds usually correspond to the minimum and maximum values of a specified product property. Normal operating data can be used to compute the mean and the standard deviation σ of a given process variable from a series of n observations y_1, y_2, \dots, y_n as follows:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad (13)$$

$$\sigma = \left[\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2 \right]^{1/2} \quad (14)$$

The standard deviation is a measure for how the values of y spread around their mean. A large value of σ indicates that wide variations in y occur. Assuming the process variable follows a normal probability distribution, then 99.7% of all observations should lie within an upper limit of $y + 3\sigma$ and a lower limit of $y - 3\sigma$. These limits can be used to determine the quality of the control. If all data from a process lie within the $\pm 3\sigma$ limits, then it can be concluded that nothing unusual has happened during the recorded time period, the process environment is relatively unchanged, and the product quality lies within specifications. On the other hand, if repeated violations of the limits occur, then the conclusion can be drawn that the process is out of control and that the process environment has changed. Once this has been determined, the process operator can take action in order to adjust operating conditions to counteract undesired changes that have occurred in the process conditions.

There are several specific rules that determine if a process is out of control. Some of the more widely used ones are the Western-Electric rules that state that a process is out of control if process data include:

- One measurement outside the $\pm 3\sigma$ control limit.
- Any seven consecutive measurements lying on the same side of the mean.
- A decreasing or increasing trend for any seven consecutive measurements.
- Any nonrandom pattern in the process measurements.

These rules can be applied to data in a control chart, such as is shown in Figure 12, where pressure measurements are plotted over a time horizon. It is then possible to read off from the control chart if the process is out of control or if it is operating within normal parameters.

Since a process that is out of control can have important economic consequences, such as product waste and customer dissatisfaction, it is important to keep track of the state of the process. Statistical process control provides a convenient method to continuously monitor process performance and product quality. However, it differs from automatic process control (such as feedback control) in that it serves as an indicator that the process is not operating within normal parameters but does not provide a controller setting that will bring the process back to its desired operating point. More details on SPC can be found in References (1), (3), and (28).

6.7. Multivariable Control. Most industrial chemical engineering processes contain several manipulated as well as controlled variables. These processes are called multivariable control systems. It is possible to analyze the

interactions among the control loops with techniques like the relative gain array (29). If it turns out that there are only small interactions between the loops then it is possible to pair the inputs and outputs in a favorable way and use single-loop controllers that can be tuned independently from one another. However, if strong interactions exist, the controllers might need to be detuned in order to reduce oscillations. An alternative approach is to utilize multivariable control techniques like model predictive control, which is discussed below. These techniques will take the interactions among the control loops into account and can result in excellent performance even for loops that have a strong effect on one another.

6.8. Model Predictive Control. Model predictive control (MPC) is a model-based control technique that has been widely used in the process industries. It is the most popular technique for handling multivariable control problems with multiple inputs and multiple outputs (MIMO) and can also accommodate inequality constraints on the inputs or outputs such as upper and lower limits. All of these problems are addressed by MPC by solving an optimization problem and therefore no complicated override control strategy is required (1). One formulation of an objective function to be minimized for a SISO case is shown in the following:

$$J = \sum_{k=0}^{N-1} (y_{sp} - y_k)^T Q (y_{sp} - y_k) + \sum_{k=0}^{M-1} \Delta u_k^T R \Delta u_k \quad (15)$$

where y_k refers to the output at time k , y_{sp} is the set point for this output, $\Delta u_k = u_k - u_{k-1}$ is the change of the input between time steps $k-1$ and k , and Q and R are weighting matrices. This objective function will penalize deviations of the controlled variable from the desired set point over the prediction horizon of length N as well as changes in the manipulated variable over the control horizon of length M . The optimization determines a future input trajectory that minimizes the objective function subject to constraints on the manipulated as well as the controlled variables. The predictions y_k are made based upon a model of the plant. This plant model will be updated using deviations between the predicted output and the real plant output from past data. An illustration of the trajectories for the manipulated (discrete trajectory) as well as the controlled variables (continuous trajectory) are shown in Figure 13.

A variety of different types of models can be used for the prediction. Choosing an appropriate model type is dependent on the application to be controlled. The model can be based upon first principles or it can be an empirical model. Also, the supplied model can be either linear or nonlinear, as long as the model predictive control software supports this type of model. Most industrial applications of MPC have relied on linear empirical models, since they can more easily be identified and solved and approximate most processes fairly well.

Usually, MPC implementations change set points in order to move the plant to a desired steady state; the actual control changes are implemented by PID controllers in response to the set points. MPC is one of the most popular advanced control strategies and has a substantial impact on oil refineries and petrochemical plants, with a number of commercial software packages in use.

Refer to Reference 30 for further details on model predictive control.

6.9. Real-Time Optimization. Operating objectives for process facilities are set by economics, product orders, availability of raw materials and utilities, etc. At different points in time it may be advantageous or necessary to operate a process in different ways to meet a particular operating objective. A process plant, however, is a dynamic, integrated environment where external and internal conditions can cause the optimal operating point for each operating objective to vary from time to time. These operating points can be computed by real-time process optimization (RTO), where the optimization can be performed on several levels, ranging from optimization within model predictive controllers, to supervisory controllers that determine the targets for optimum production of the plant, to optimization of production cycles. The plant-wide problems that can be solved by optimization techniques on a daily or hourly basis can be large containing thousands or even tens of thousands of variables.

The foundation of RTO is made up of the economic and process models of a plant. The processes are often described by steady-state models because set point selection is traditionally considered part of steady-state process optimization and thus depends on steady-state material and energy balances. Therefore, traditional control theory is not normally employed in RTO. Instead, set point trajectories are re-evaluated periodically to incorporate changes in process conditions.

Often the optimal operating points at the local optimization level lie along a constraint. For such a case a RTO strategy is a straightforward way to achieve and maintain operation close to the optimal operating point. RTO monitors the proximity to relevant constraints, identifies which are active, and incrementally adjusts the set points to control and maintain operation close to a constrained condition.

Many different optimization techniques are available nowadays for RTO, including linear programming (LP), nonlinear programming (NLP), and mixed integer nonlinear programming (MINLP); a variety of different solution algorithms for each type of problem exists. See Reference 31 for more details on optimization.

7. Batch and Sequence Control

In batch processes the product is made in discrete batches by sequentially performing a number of processing steps in a defined order on the raw materials and intermediate products. For example, fixed amounts of reactants may be charged to a vessel, mixed and heated to a reaction temperature, reacted for a fixed period of time, drained from the vessel, separated, dried, and packaged. Large production runs are achieved by repeating the process. The term recipe has a range of definitions in batch processing, but in general a recipe is a procedure with the set of data, operations, and control steps to manufacture a particular grade of product. A formula is the list of recipe parameters, which includes the raw materials, processing parameters, and product outputs. A recipe procedure has operations for both normal and abnormal conditions. Each operation contains resource requests for certain ingredients (and their amounts). The operations in the recipe can adjust set-points and turn equipment on and off. The complete production run for a specific recipe is called a campaign (multiple batches). A production

run consists of a specified number of batches using the same raw materials and making the same product to satisfy customer demand. The accumulated batches are called a lot (32,33).

In multigrade batch processing, the instructions remain the same from batch to batch, but the formula can be changed to yield modest variations in the product. In emulsion polymerization, different grades of polymers are manufactured by changing the formula. In flexible batch processing, both the formula (recipe parameters) and the processing instructions can change from batch to batch. The recipe for each product must specify both the raw materials required and how conditions within the reactor are to be sequenced in order to make the desired product.

Many batch plants, especially those used to manufacture pharmaceuticals, are certified by the International Standards Organization (ISO). The ISO 9000 standard (and related standards 9001–9004) states that every manufactured product should have an established, documented procedure, and the manufacturer should be able to document that the procedure was followed. Companies must pass periodic audits to maintain ISO 9000 status. Both ISO 9000 and the United States Food and Drug Administration (FDA) require that only a certified recipe can be used. Thus, if the operation of a batch becomes “abnormal”, performing any unusual corrective action to bring it back within the normal limits is not an option. Also, if a slight change in the recipe apparently produces superior batches, the improvement cannot be implemented unless the entire recipe is recertified. The FDA typically requires product and raw material tracking, so that product abnormalities can be traced back to their sources.

7.1. Batch Process Control Hierarchy. Functional control activities for batch process control are presented in Figure 14 and are summarized below in four categories: batch sequencing and logic control, control during the batch, run-to-run control, and batch production management (1,34).

In batch sequencing and logic control, sequencing of control steps that follow the recipe involves, for example: mixing of ingredients, heating, waiting for a reaction to complete, cooling, or discharging the resulting product. Transfer of materials to and from batch tanks or reactors includes metering of materials as they are charged (as specified by each recipe), as well as transfer of materials at the completion of the process operation. In addition to discrete logic for the control steps, logic is needed for safety interlocks to protect personnel, equipment, and the environment from unsafe conditions. Process interlocks ensure that process operations can only occur in the correct time sequence.

Feedback control of flow rate, temperature, pressure, composition, and level, including advanced control strategies, falls under control during the batch, which is also called “within-the-batch” control (34). In sophisticated applications, this requires defining an operating trajectory for the batch (ie, temperature or flow rate as a function of time). In simpler cases, it involves tracking of set points of the controlled variables, which includes ramping the controlled variables up and down and/or holding them for a prescribed period of time. Detection of when the batch operations should be terminated (end point) may be performed by inferential measurements of product quality, if direct measurement is not feasible.

Run-to-run control (also called batch-to-batch) is a supervisory function based on off-line product quality measurements at the end of a run. Operating conditions and profiles for the batch are adjusted between runs to improve the product quality using tools such as optimization. Batch production management entails advising the plant operator of process status and how to interact with the recipes and the sequential, regulatory, and discrete controls. Complete information (recipes) is maintained for manufacturing each product grade, including the names and amounts of ingredients, process variable set points, ramp rates, processing times, and sampling procedures. Other database information includes batches produced on a shift, daily, or weekly basis, as well as material and energy balances. Scheduling of process units is based on availability of raw materials and equipment and customer demand.

The ability to handle recipe changes after a recipe has started is a challenging aspect of batch control systems. Often it is desirable to change the grade of the batch to meet product demand, or to change the resources used by the batch after the batch has started. Because every batch of product is not always good, there needs to be special-purpose control recipes to fix, rework, blend, or dispose of bad batches, if that is allowable. It is important to be able to respond to unusual situations by creating special-purpose recipes and still meet the demand. This procedure is referred to as reactive scheduling. When ample storage capacity is available, the normal practice has been to build up large inventories of raw materials and ignore the inventory carrying cost. However, improved scheduling can be employed to minimize inventory costs, which implies that supply chain management techniques may be necessary to implement the schedule (35).

7.2. Sequential Function Charts. Compared to a continuous process, batch process control requires a greater percentage of discrete logic and sequential control than regulatory control loops. Batch control applications must control the timing and sequencing of the process steps based on discrete input and outputs as well as analog outputs (36). The complexity of the interactive logic within and between the various control levels, the required interactions with operators, and the need for ongoing application modification and maintenance are reasons why organization, functional design, and clear documentation are so important to the successful use of batch control applications. In order to describe what must be done, structural models are commonly used to represent the required batch processing actions, the batch equipment, and the combination of components (see Fig. 15). Various formats have been proposed for describing the batch control applications, eg, how the batch processing steps are carried out with the batch equipment and instrumentation, interfaces between the various levels of control, interfaces between the batch control and the operator actions and responses, and interactions and coordination with the safety interlocks. The formats proposed include flow charts, state charts, decision tables, structured pseudocode, state transition diagrams, petri nets, and sequential function charts.

A sequential function chart (SFC) describes graphically the sequential behavior of a control program. More sophisticated than the information flow chart, it is derived from two earlier approaches, Petri Nets and Grafcet (37). The SFCs consist of steps that alternate between action blocks and transitions. Each step corresponds to a state of the batch process. A transition is associated with a condition that when true activates the next step and deactivates the pre-

vious step. Steps are linked with action blocks that perform a specific control action. SFC and Grafcet are standard languages established by the International Electrotechnical Commission (IEC) and are supported by an association of vendors and users called PLCopen; see www.plcopen.org. Figure 16 gives a simple illustration of the SFC notation. The steps are denoted as rectangles (the initial step corresponds to the double rectangle) and the transition symbol is a small horizontal bar on the line linking control steps. A double bar is used for branching, and it can precede a transition when two or more paths can be followed. Likewise a double bar indicates where two or more parallel paths join together into a single path. In Figure 16, both steps 5 and 6 must be completed before moving to step 7. The active steps are shown with a black dot in the box.

Since 1989, the Instrument, Systems, and Automation Society (ISA) has been working on developing a standard, referred to as ISA-S88, for standardizing batch control concepts, structures, terminology, data structures, and batch control configuration and programming languages (38,39).

8. Economic Aspects: Process Instrumentation and Control Systems

Investment for instrumentation and control systems and their installation can make up a significant percentage of the total installed cost for a grassroots continuous process facility. Instrumentation and control systems also represent a substantial percentage of the overall facility maintenance costs. Investment costs may be placed in one of two categories, ie, nondiscretionary and discretionary. For example, a control application that requires certain instrumentation and control system features, and that increases the average production rate, normally has little economic value in a market-limited environment, but may have significant value within a production-limited environment. A list of the benefits of improved control follows:

- Raw material costs and utilization.
 - Yield improvement.
 - Reduced losses.
 - Improved product and raw material recovery.
 - Ability to operate closer to constraints and specifications consistently and safely.
 - Reduced quality giveaway.
- Operating and inventory cost.
 - Reduction in utility requirements.
 - Decreased need for conservative operation.
 - Reduced need for building inventory and associated provisions.
 - Consistent operation closer to optimum operation.
 - Reduced reprocessing of off-spec material.
- Other
 - Improvement in service factor.
 - Reduction of off-grade product when shifting operations and control strategies for different situations such as grade changes and feedstock

changes.

Improved data for operations and control performance engineering analysis.

Generally, the installed cost of instrumentation and control systems is typically in the range of 3–10% of the total installed cost of the facility. The engineering and detailed design costs are typically ~35% of the purchased equipment cost of the process-related equipment or ~13% of the total installed cost. Of these engineering and detailed design costs, those associated with instrumentation and control systems are typically ~11% in a full engineering and procurement project for a grassroots facility. It is often possible, once a rough design has been developed, to work with a number of vendors to obtain preliminary costing for their systems. Instrumentation costs can vary significantly based on materials of construction, service conditions, etc.

BIBLIOGRAPHY

“Instrumentation” in *ECT* 1st ed., Vol. 7, pp. 908–926, by J. Procopi, Minneapolis-Honeywell Regulator Co.; in *ECT* 2nd ed., Vol. 11, pp. 739–774, by N. A. Fiorino, Honeywell, Inc.; “Instrumentation and Control” in *ECT* 3rd ed., Vol. 13, pp. 485–512, by T. J. Williams, Purdue University; “Process Control” in *ECT* 4th ed., Vol. 20, pp. 124–174, by John Paul San Giovanni, Jockey Hollow Technologies; “Process Control” in *ECT* (online), posting date: December 4, 2000, by John Paul San Giovanni, Jockey Hollow Technologies.

CITED PUBLICATIONS

1. D. E. Seborg, T. F. Edgar, and D. A. Mellichamp, *Process Dynamics and Control*, 2nd ed., John Wiley & Sons Inc., New York, 2003.
2. C. A. Smith and A. B. Corripio, *Principles and Practice of Automatic Process Control*, 2nd ed., John Wiley & Sons Inc., New York, 1999.
3. B. A. Ogunnaike and W. H. Ray, *Process Dynamics, Modeling, and Control*, Oxford, New York, 1994.
4. C. D. Johnson, *Process Control Instrumentation Technology*, 6th ed., Prentice-Hall, Upper Saddle River, N.J., 2000.
5. B. G. Liptak, *Instrument Engineers Handbook*, Process Software and Digital Networks, 3rd ed., CRC Press, New York, 2002.
6. S. M. Herb, *Understanding Distributed Processor Systems for Control*, ISA, Research Triangle Park, North Carolina, 1999.
7. J. W. Webb and R. A. Reis, *Programmable Logic Controllers*, 5th ed., Prentice-Hall, Upper Saddle River, N.J., 2002.
8. T. A. Hughes, *Programmable Controllers*, ISA, Research Triangle Park, N.C., 1997.
9. T. F. Edgar, C. L. Smith, F. G. Shinskey, G. W. Gassman, P. J. Schafbuch, T. J. McAvoy, and D. E. Seborg, “Process Control”, Section 8 in *Perry's Chemical Engineering Handbook*, 6th ed., McGraw-Hill, New York, 1997.
10. J. Berge, *Fieldbuses for Process Control Engineering, Operation and Maintenance*, ISA, Research Triangle Park, North Carolina, 2002.

11. R. Connell, *Process Instrumentation Applications Manual*, McGraw-Hill, New York, 1996.
12. B. G. Liptak, ed., *Instrument Engineers Handbook*, Vol. 1 (Process Measurement) Chilton Books, Philadelphia, Pa., 2000.
13. D. W. Spitzer, *Flow Measurement*, 2nd ed., ISA, Research Triangle Park, North Carolina, 2001.
14. J. W. Dolenc, *Chem Engr. Prog.*, **92**, 22(1996).
15. J. Dakin and B. Culshaw, eds., *Optical Fiber Sensors: Applications, Analysis, and Future Trends*, Vol. IV, Artech House, Norwood, Mass., 1997.
16. S. Soloman, *Sensors Handbook*, McGraw-Hill, New York, 1999.
17. G. D. Nichols, *On-Line Process Analyzers*, John Wiley & Sons Inc., New York, 1988.
18. F. McLennan and B. Kowalski, eds., *Process Analytical Chemistry*, Blackie Academic and Professional, London, U.K., 1995.
19. G. Borden, ed., *Control Valves*, ISA, Research Triangle Park, N.C., 1998.
20. A. Constantinides and N. Mostoufi, *Numerical Methods for Chemical Engineers with MATLAB Applications*, Prentice-Hall, Englewood Cliffs, N.J., 1999.
21. B. W. Bequette, *Process Control: Modeling, Design and Simulation*, Prentice-Hall, Upper Saddle River, N.J., 2003.
22. M. Morari and E. Zafiriou, *Robust Process Control*, Prentice-Hall, Englewood Cliffs, N.J., 1989.
23. I. L. Chien and P. S. Fruehauf, *Chem. Engr. Prog.* **86**(10), 33 (1990).
24. F. G. Shinskey, *Process Control Systems*, 4th ed., McGraw-Hill, New York, 1996.
25. T. E. Marlin, *Process Control*, McGraw-Hill, New York, 2000.
26. K. J. Åström and B. Wittenmark, *Adaptive Control*, Addison-Wesley, Reading, Mass., 1988.
27. J.-S. R. Jang, C.-T. Sun, and E. Mizutani, *Neuro-Fuzzy and Soft Computing*, Prentice-Hall, Englewood Cliffs, N.J., 1997.
28. D. C. Montgomery, *Introduction to Statistical Quality Control*, 4th ed., John Wiley & Sons Inc., New York, 2001.
29. T. J. McAvoy, *Interaction Analysis*, ISA, Research Triangle Park, North Carolina, 1983.
30. E. F. Camacho and C. Bordons, *Model Predictive Control*, Springer-Verlag, New York, 1999.
31. T. F. Edgar, D. M. Himmelblau, and L. Lasdon, *Optimization of Chemical Processes*, McGraw-Hill, New York, 2001.
32. T. G. Fisher, *Batch Control Systems: Design, Application, and Implementation*, ISA, Research Triangle Park, North Carolina, 1990.
33. H. P. Rosenof and A. Ghosh, *Batch Process Automation: Theory and Practice*, Van Nostrand Reinhold, New York, 1987.
34. D. Bonvin, *J. Process Contr.* **8**, 355(1998).
35. J. F. Pekny and G. V. Reklaitis, Towards the Convergence of Theory and Practice: A Technology Guide for Scheduling/Planning Methodology, in Foundations of Computer-Aided Process Operations (FOCAPO) Conference Proceedings, *AIChE Symp. Ser.* **94**, 91,(1998).
36. K. T. Erickson and J. L. Hedrick, *Plantwide Process Control*, John Wiley & Sons Inc., New York, 1999.
37. R. David, *IEEE Trans. Control Syst. Technology*, **3**, 253 (1995).

38. *Batch Control I: Models and Terminology*, ANSI/ISA-88.01, ISA, Research Triangle Park, North Carolina, 1995.
39. *Batch Control 2: Data Structures and Guidelines for Languages*, ANSI/ISA-88.00.02, ISA, Research Triangle Park, North Carolina, 2001.

JUERGEN HAHN
Texas A&M University

THOMAS F. EDGAR
The University of Texas at Austin

Table 1. On-Line Measurement Options for Process Control

Temperature	Flow	Pressure	Level	Composition
thermocouple	orifice	liquid column	float-activated	gas-liquid chromatography (GLC)
resistance	venturi	diaphragm	head devices	mass spectrometry (MS)
thermometer	turbine	strain gauges	electrical (conductivity)	magnetic resonance analysis (MRA)
detector (RTD)	vortex-shedding	piezoelectric transducers	radiation	infrared (IR)
pyrometer	magnetic		ultrasonic	raman
laser	thermal mass			ultraviolet (UV)
	coriolis			thermal conductivity
	capacitance			refractive index (RI)
	probes			capacitance probe
				electrophoresis
				electrochemical
				paramagnetic
				chemi/
				bioluminescence

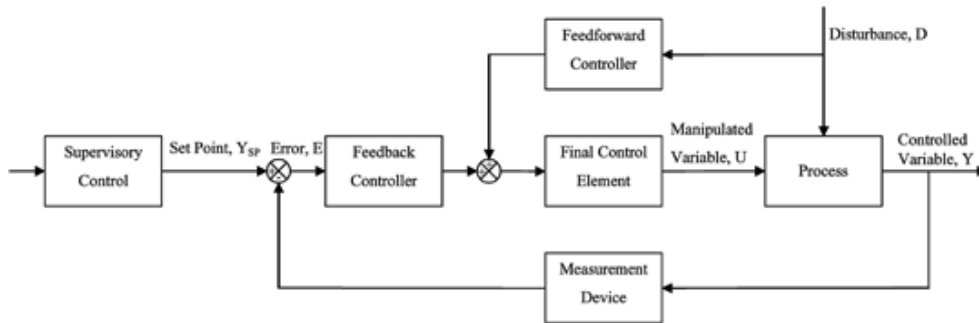


Fig. 1. Block diagram of a process.

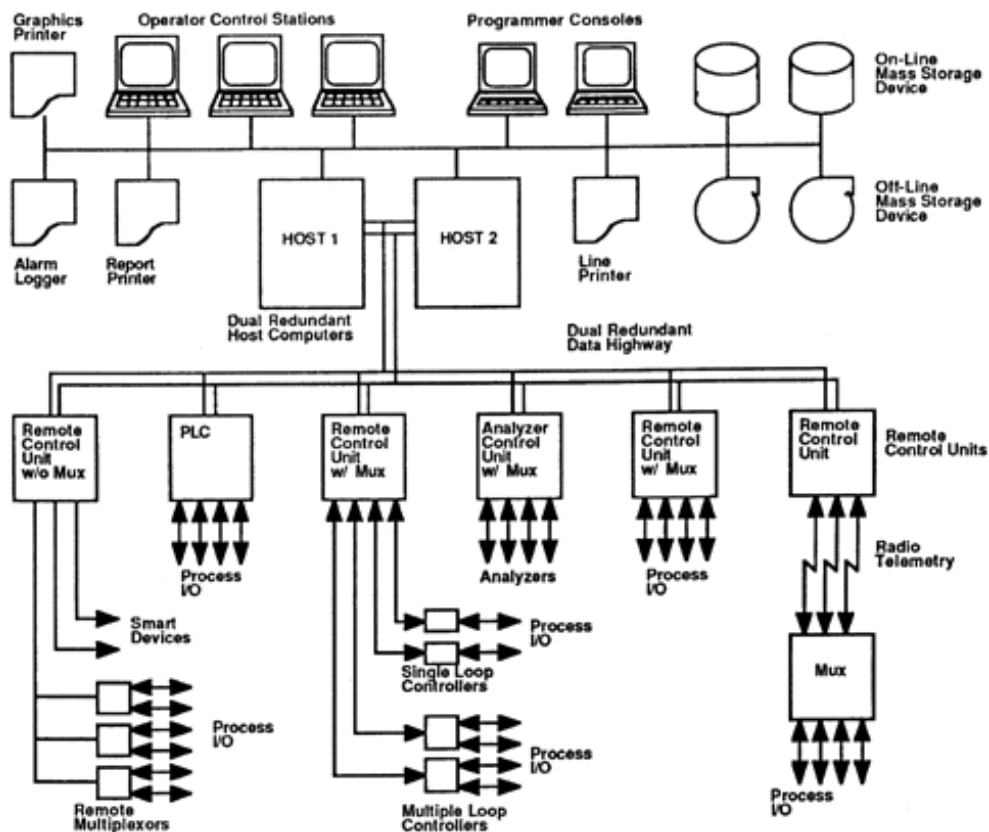


Fig. 2. Distributed control system structure.

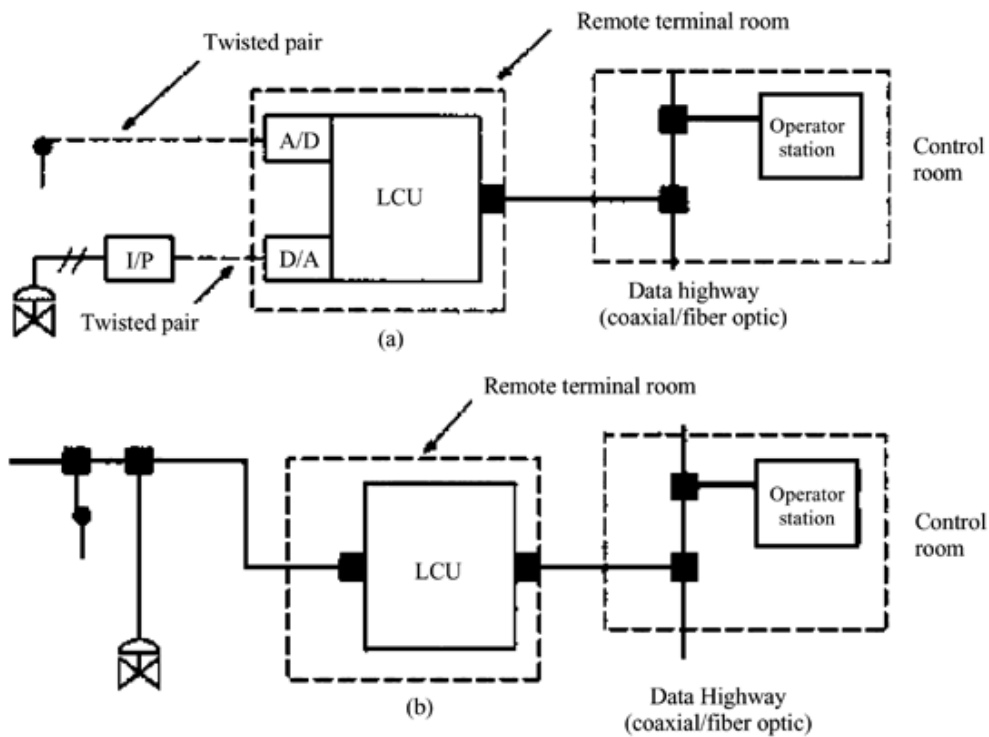


Fig. 3. Impact of field bus on wiring costs, where A/D and D/A correspond to analogue-to-digital and digital-to-analogue converts, respectively; LCU=local control unit; and I/P= current-to-pressure transducer. **(a)** DCS with remote terminal rooms; and **(b)** DCS with remote terminal rooms and fieldbus.

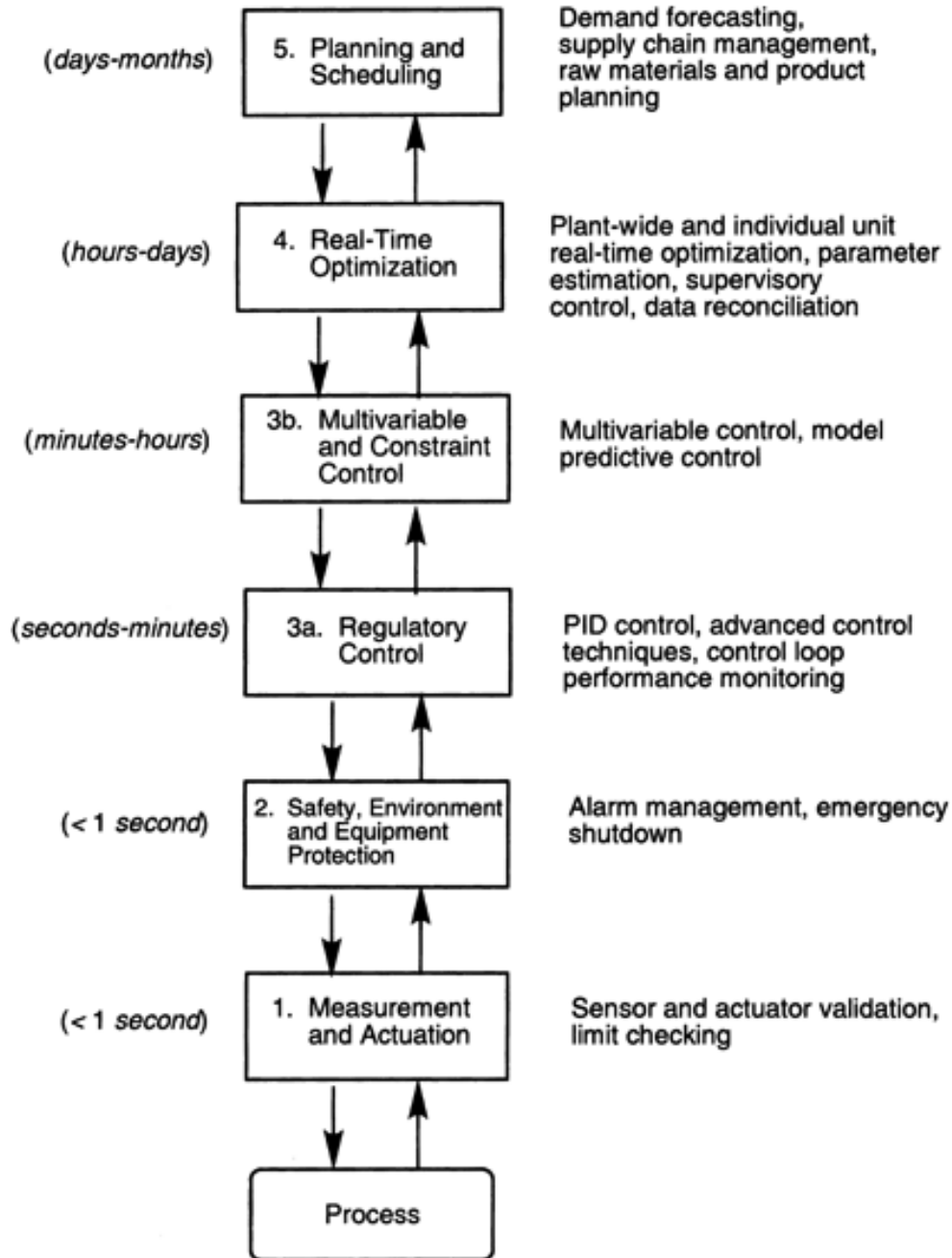


Fig. 4. The five levels of process control and optimization in manufacturing. Time scales are shown for each level (1).

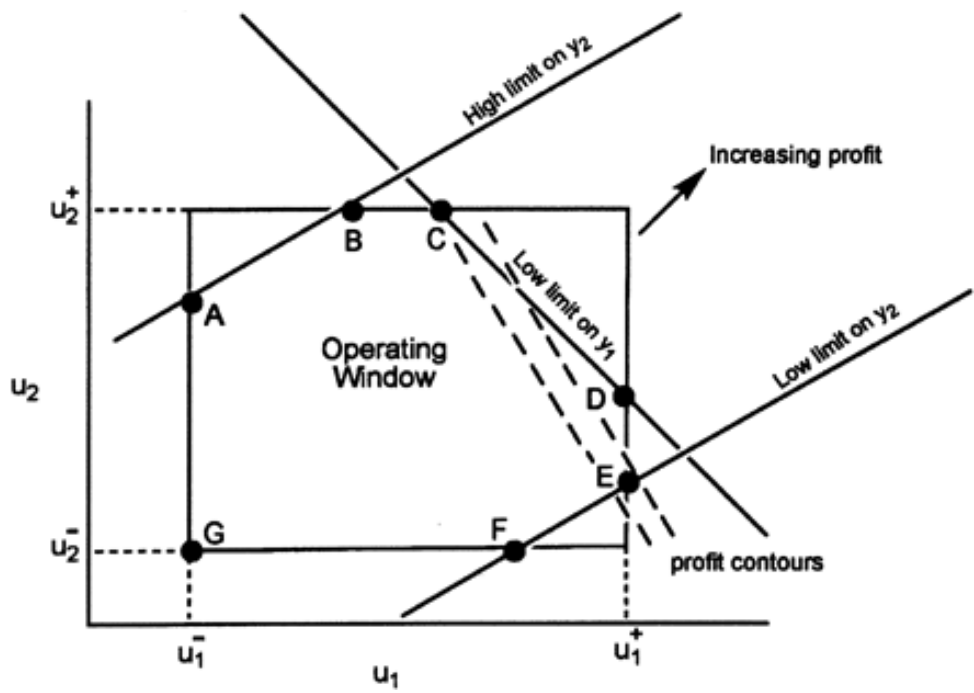


Fig. 5. Operating window for two control variables (u_1 and u_2). The dashed lines are objective function contours, increasing from left to right. The maximum profit occurs where the profit line intersects the constraints at vertex D (1).

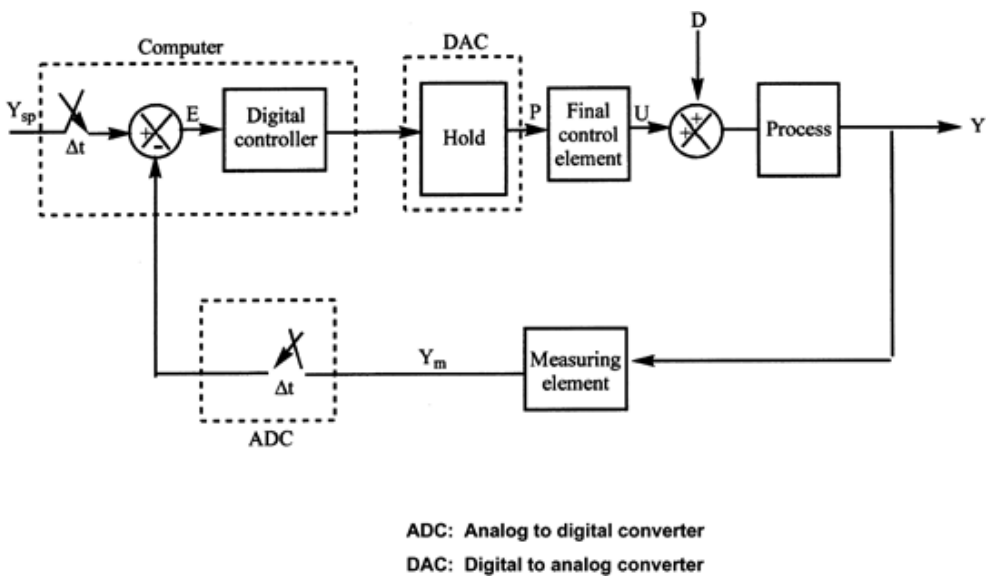


Fig. 6. Components of a computer control loop.

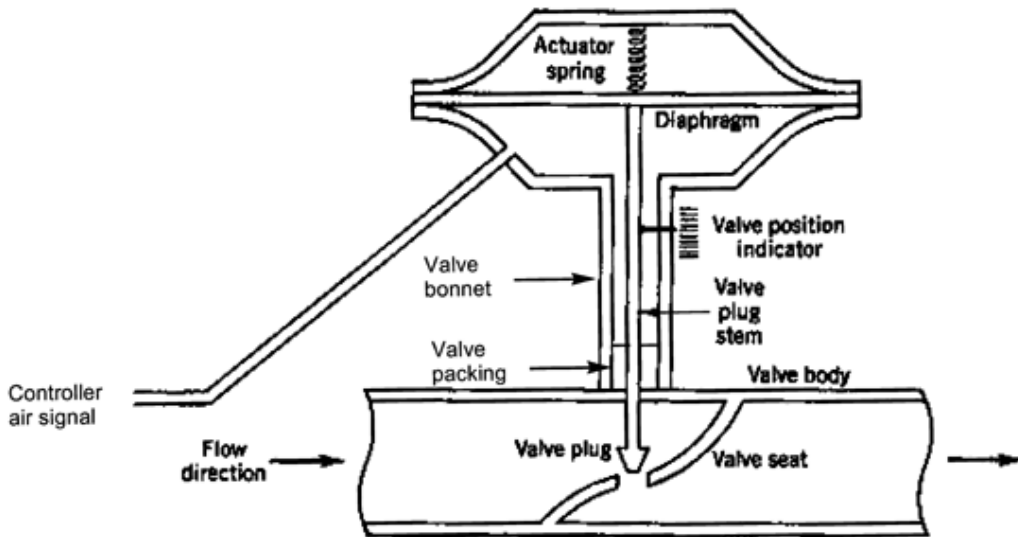


Fig. 7. Control valve and actuator.

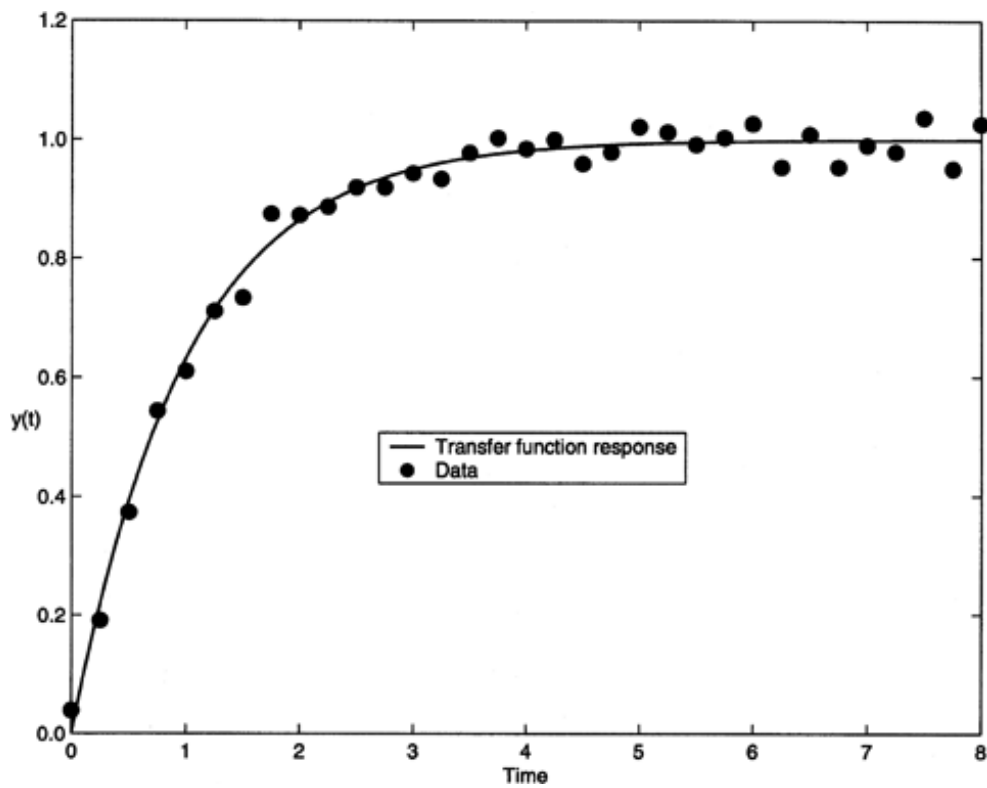


Fig. 8. Step response of a first order model.

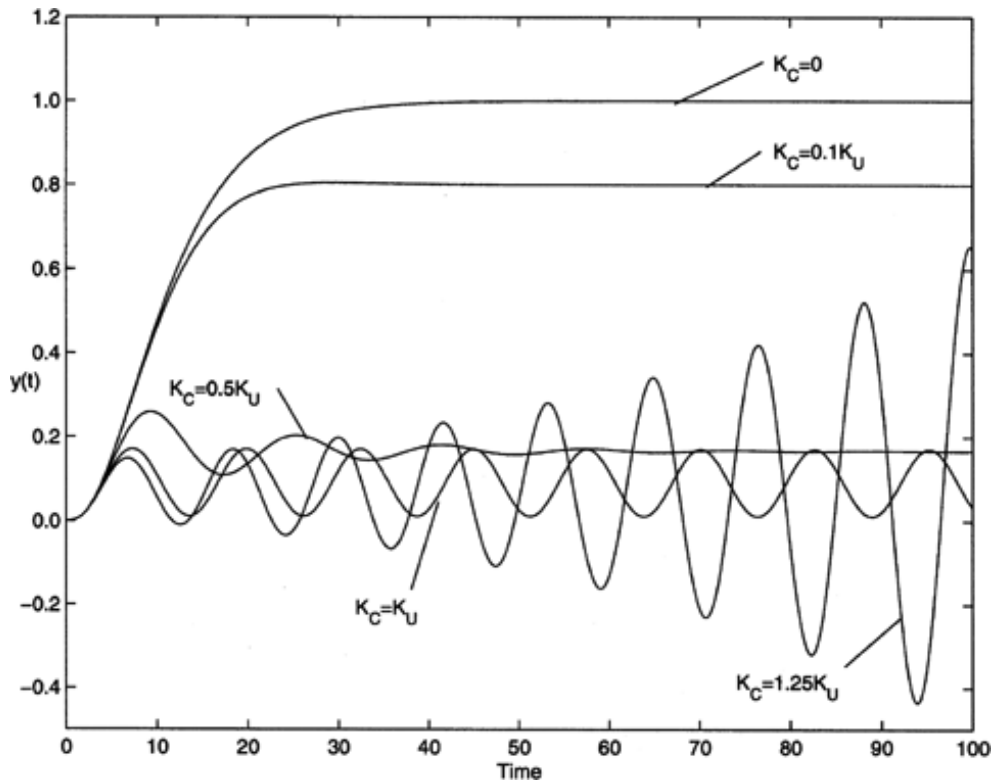


Fig. 9. Dynamic response of open and closed-loop systems.

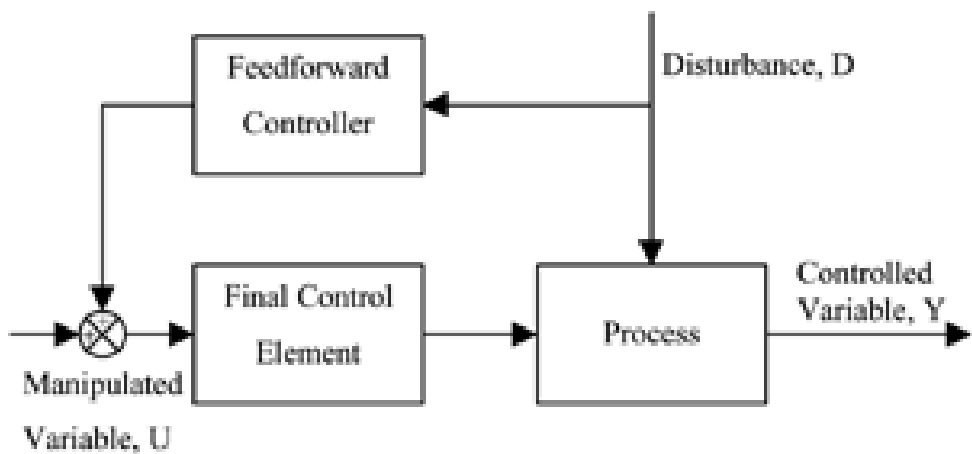


Fig. 10. Implementation of a feedforward controller.

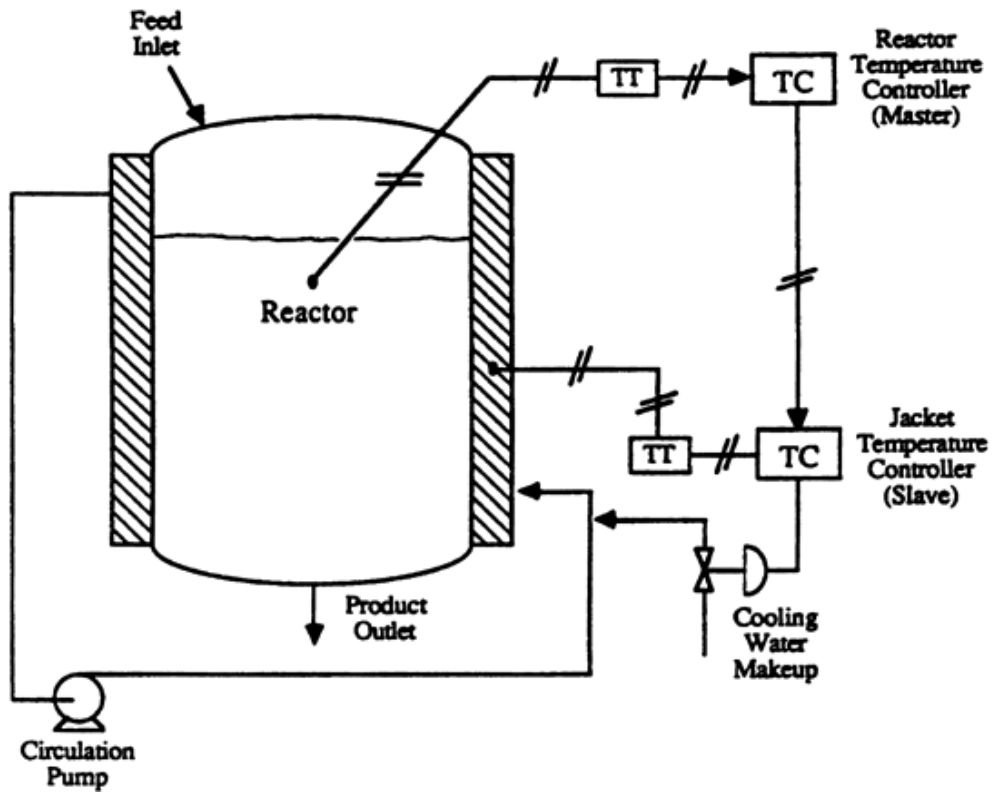


Fig. 11. Cascade control of an exothermic chemical reactor.

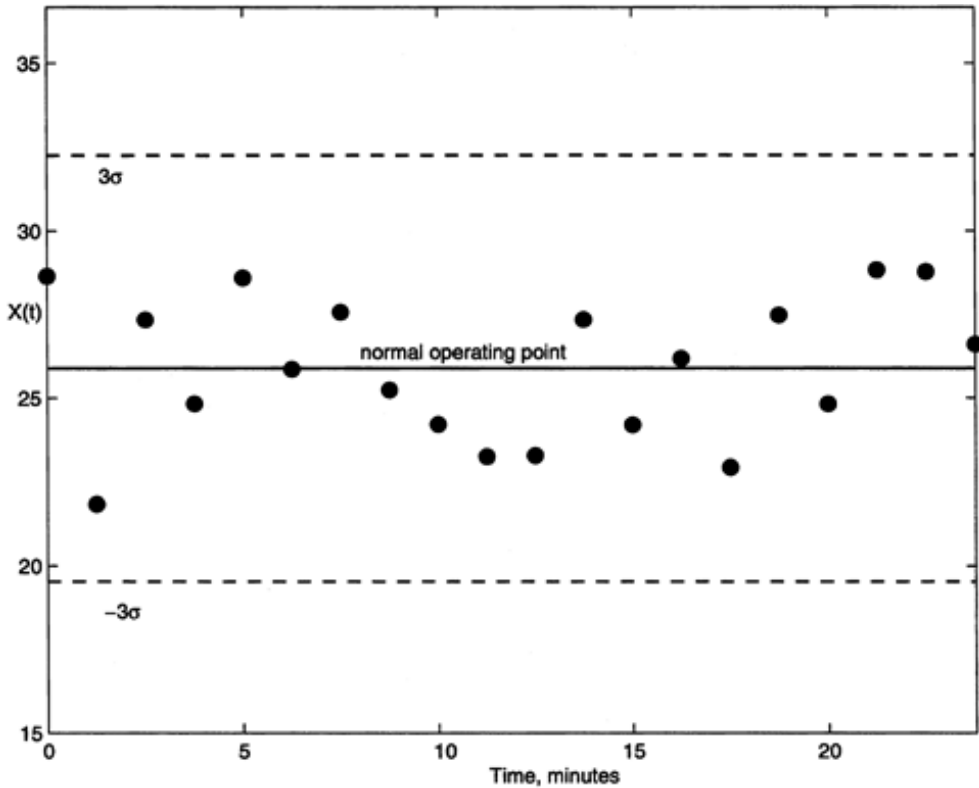


Fig. 12. Control chart for SPC. Dashed lines indicate the $\pm 3 \sigma$ levels for the process variables.

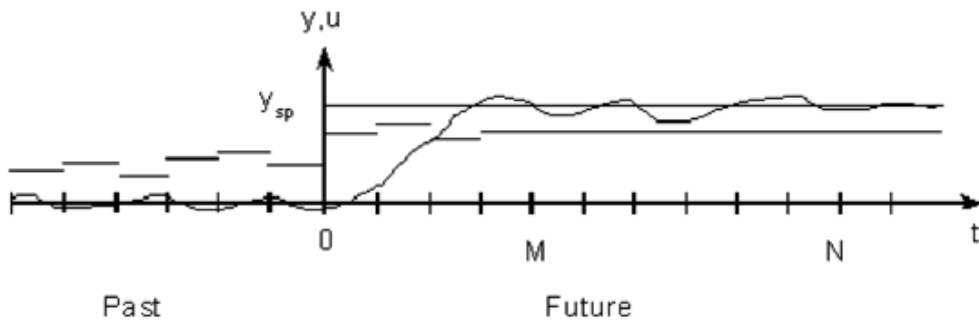


Fig. 13. Past inputs and trajectories are used for forecasting future control action.

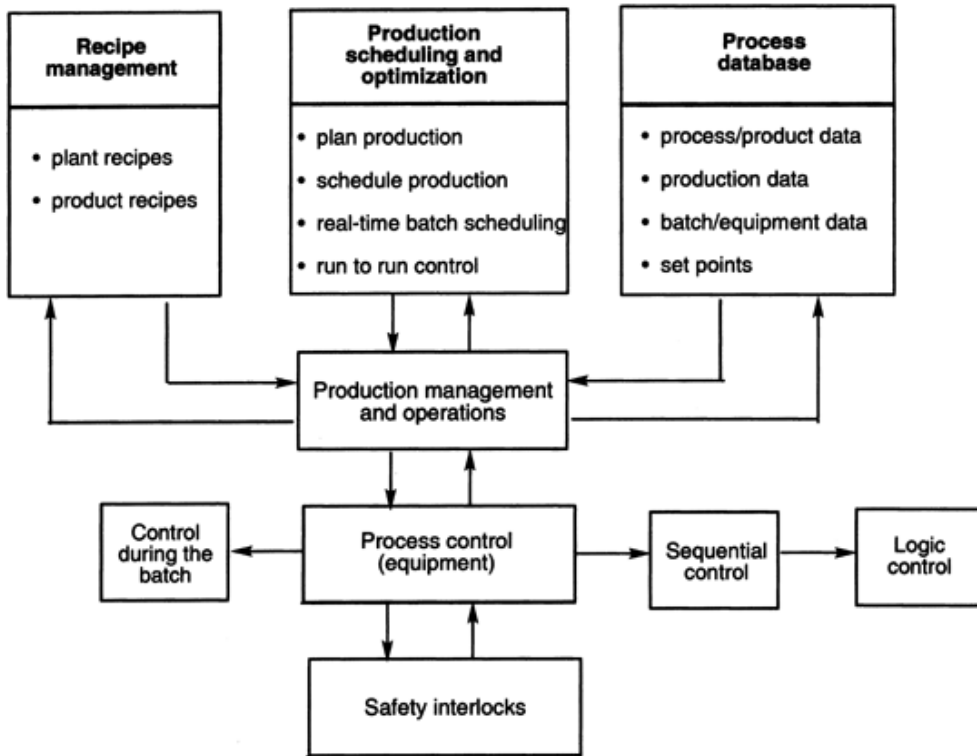


Fig. 14. Batch control system information and control system.

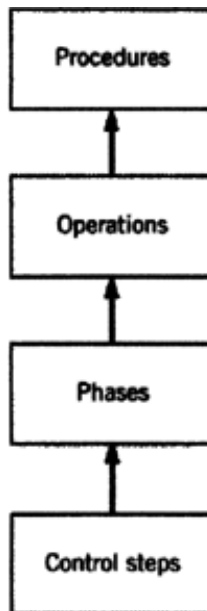


Fig. 15. Batch structural models and definitions.

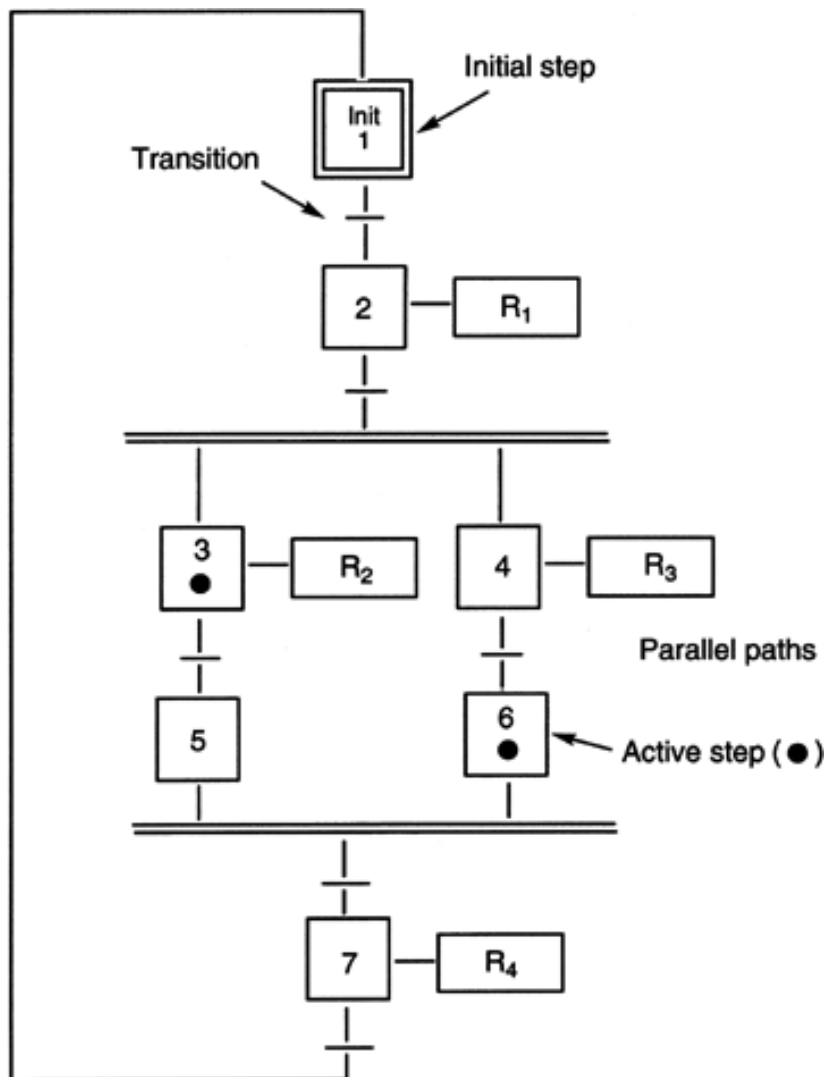


Fig. 16. A generic sequential function chart (active steps are 3 and 6).