SECTION 2
CONTROL SYSTEM FUNDAMENTALS*

E. H. Bristol
Research Department, The Foxboro Company (a Siebe Company), Foxboro, Massachusetts. (Basic Control Algorithms)

G. A. Hall, Jr.
Westinghouse Electric Corporation, Pittsburgh, Pennsylvania. (Control Principles—prior edition)

Peter D. Hansen
Systems Development and Engineering, The Foxboro Company (a Siebe Company), Foxboro, Massachusetts. (Techniques for Process Control)

Stephen P. Higgins, Jr.
Retired Senior Principal Software Engineer, Honeywell Inc., Phoenix, Arizona. (Techniques for Process Control—prior edition)

Richard H. Kennedy
The Foxboro Company (a Siebe Company), Foxboro, Massachusetts. (Techniques for Process Control—prior edition)

E. C. Magison
Ambler, Pennsylvania. (Safety in Instrumentation and Control Systems)

Joe M. Nelson
Applications Systems, Honeywell Inc., Billerica, Massachusetts. (Techniques for Process Control—prior edition)

Robert L. Osborne
Manager, Diagnostics and Control, Westinghouse Electric Corporation, Orlando, Florida. (Techniques for Process Control—prior edition)

John Stevenson
West Instruments, East Greenwich, Rhode Island. (Control Principles)

* *Persons who authored complete articles or subsections of articles, or who otherwise cooperated in an outstanding manner in furnishing information and helpful counsel to the editorial staff.*

CONTROL PRINCIPLES

by John Stevenson*

In this article the commonly measured process variable temperature is used for the basis of discussion. The principles being discussed apply also to other process variables, although some may require additional sophisticated attention, as discussed in the subsequent article, which deals with techniques for process control.

In contrast with manual control, where an operator may periodically read the process temperature and adjust the heating or cooling input up or down in such a direction as to drive the temperature to its desired value, in automatic control, measurement and adjustment are made automatically on a continuous basis. Manual control may be used in noncritical applications, where major process upsets are unlikely to occur, where any process conditions occur slowly and in small increments, and where a minimum of operator attention is required. However, with the availability of reliable low-cost controllers, most users opt for the automatic mode. A manual control system is shown in Fig. 1.

In the more typical situation, changes may be too rapid for operator reaction, making automatic control mandatory (Fig. 2). The controlled variable (temperature) is measured by a suitable sensor, such as a thermocouple, a resistance temperature detector (RTD), a thermistor, or an infrared pyrometer. The measurement signal is converted to a signal that is compatible with the controller. The controller compares the temperature signal with the desired temperature (set point) and actuates the final control device. The latter alters the quantity of heat added to or removed from the process. Final control devices, or elements, may take the form of contactors, blowers, electric-motor or pneumatically operated valves, motor-operated variacs, time-proportioning or phase-fired silicon-controlled rectifiers (SCRs), or saturable core reactors. In the case of automatic temperature controllers, several types can be used for a given process. Achieving satisfactory temperature control, however, depends on (1) the process characteristics, (2) how much temperature variation from the set point is acceptable and under what conditions (such as start-up, running, idling), and (3) selecting the optimum controller type and tuning it properly.

PROCESS (LOAD) CHARACTERISTICS

In matching a controller with a process, the engineer will be concerned with process reaction curves and the process transfer function.

Process Reaction Curve

An indication of the ease with which a process may be controlled can be obtained by plotting the process reaction curve. This curve is constructed after having first stabilized the process temperature under manual control and then making a nominal change in heat input to the process, such as 10%. A temperature recorder then can be used to plot the temperature versus time curve of this change. A curve similar to one of those shown in Fig. 3 will result.

Two characteristics of these curves affect the process controllability, (a) the time interval before the temperature reaches the maximum rate of change, A, and (2) the slope of the maximum rate of change of the temperature after the change in heat input has occurred, B. The process controllability

* Engineering Manager, West Instruments, East Greenwich, Rhode Island.

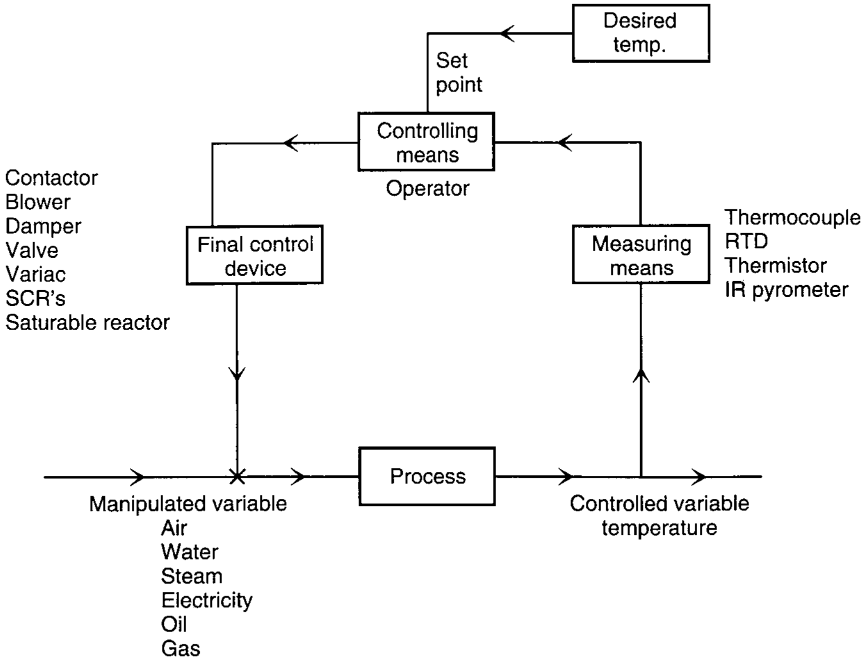


FIGURE 1 Manual temperature control of a process. (*West Instruments.*)

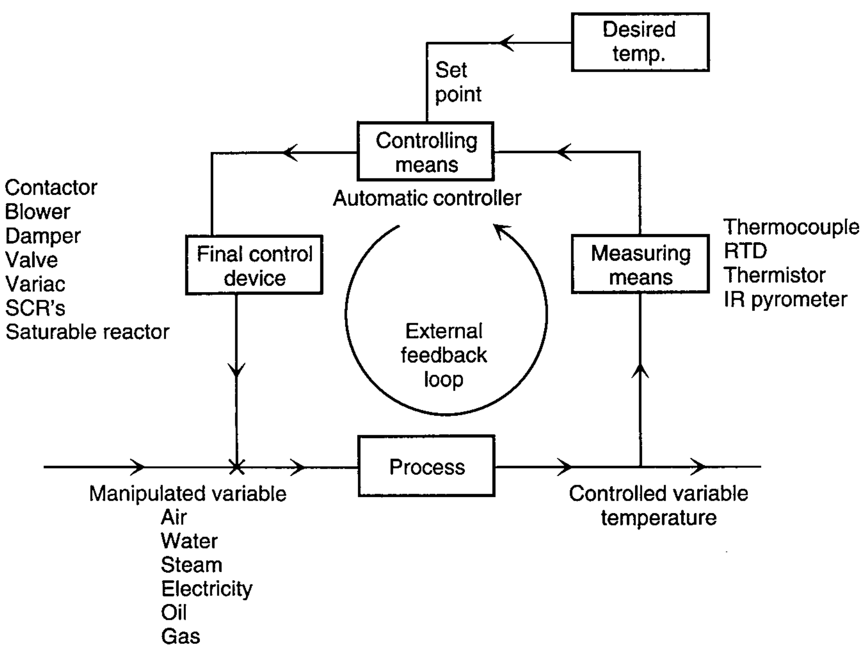


FIGURE 2 Automatic temperature control of a process. (*West Instruments.*)

decreases as the product of *A* and *B* increases. Such increases in the product *AB* appear as an increasingly pronounced S-shaped curve on the graph. Four representative curves are shown in Fig. 3.

The time interval *A* is caused by dead time, which is defined as the time between changes in heat input and the measurement of a perceptible temperature increase. The dead time includes two components, (1) propagation delay (material flow velocity delay) and (2) exponential lag (process thermal time constants). The curves of Fig. 3 can be related to various process time constants. A single time-constant process is referred to as a first-order lag condition, as illustrated in Fig. 4.

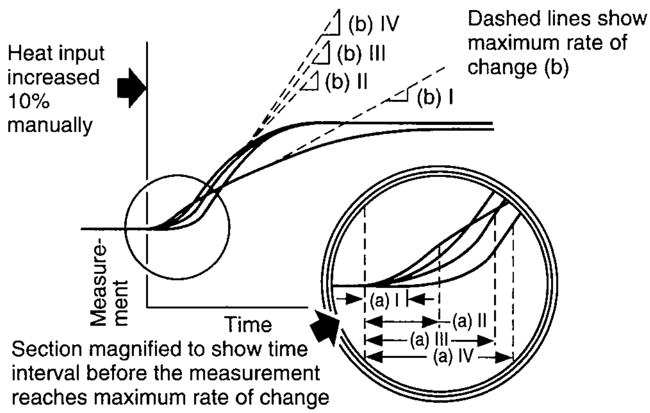


FIGURE 3 Process reaction curves. The maximum rate of temperature rise is shown by the dashed lines which are tangent to the curves. The tangents become progressively steeper from I to IV. The time interval before the temperature reaches the maximum rate of rise also becomes progressively greater from I to IV. As the S curve becomes steeper, the controllability of the process becomes increasingly more difficult. As the product of the two values of time interval *A* and maximum rate *B* increases, the process controllability goes from easy (I) to very difficult (IV). Response curve IV, the most difficult process to control, has the most pronounced S shape. Similar curves with decreasing temperature may be generated by decreasing the heat input by a nominal amount. This may result in different *A* and *B* values. (*West Instruments*.)

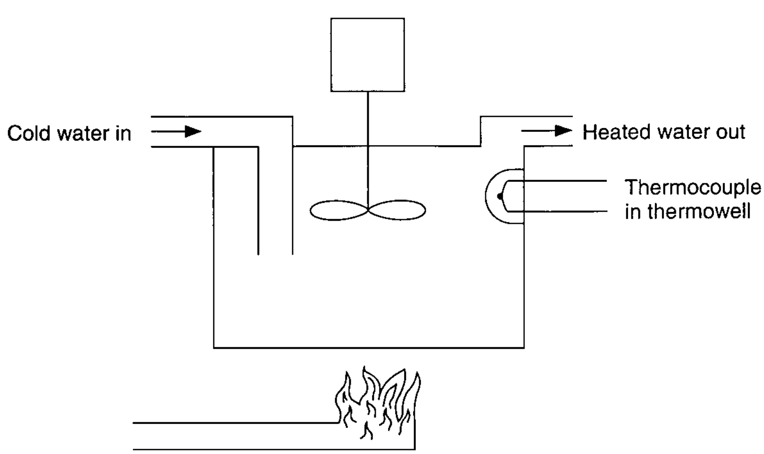


FIGURE 4 Single-capacity process. (*West Instruments*.)

This application depicts a water heater with constant flow, whereby the incoming water is at a constant temperature. A motor-driven stirrer circulates the water within the tank in order to maintain a uniform temperature throughout the tank. When the heat input is increased, the temperature within the entire tank starts to increase immediately. With this technique there is no perceptible dead time because the water is being well mixed. Ideally, the temperature should increase until the heat input just balances the heat taken out by the flowing water. The process reaction curve for this system is shown by Fig. 5.

The system is referred to as a single-capacity system. In effect, there is one quantity of thermal resistance R_1 from the heater to the water and one quantity of thermal capacity C_1, which is the quantity of water in the tank. This process can be represented by an electrical analog with two resistors and one capacitor, as shown in Fig. 6. R_{LOSS} represents the thermal loss by the flowing water plus other conduction, convection, and radiation losses.

It should be noted that since the dead time is zero, the product of dead time and maximum rate of rise is also zero, which indicates that the application would be an easy process to control. The same process would be somewhat more difficult to control if some dead time were introduced by placing the temperature sensor (thermocouple) some distance from the exit pipe, as illustrated in Fig. 7. This propagation time delay introduced into the system would be equal to the distance from the outlet of the tank to the thermocouple divided by the velocity of the exiting water. In this case the reaction curve would be as shown in Fig. 8. The product *AB* no longer is zero. Hence the process becomes increasingly more difficult to control since the thermocouple no longer is located in the tank.

A slightly different set of circumstances would exist if the water heater were modified by the addition of a large, thick metal plate or firebrick on the underside of the tank, between the heater and the tank bottom, but in contact with the bottom. This condition would introduce a second-order lag, which then represents a two-capacity system. The first time constant is generated by the thermal resistance from the heater to the plate and the plate heat capacity. The second time constant comes from the thermal resistance of the plate to the water and the heat capacity of the water. The system is shown in Fig. 9. The reaction curve for the system is given in Fig. 10. There is now a measurable

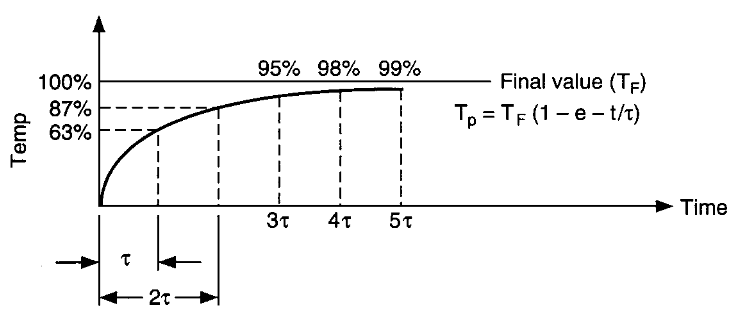


FIGURE 5 Reaction curve for single-capacity process. (*West Instruments.*)

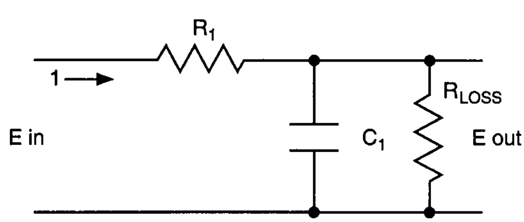


FIGURE 6 Electrical analog for single-capacity process. (*West Instruments.*)

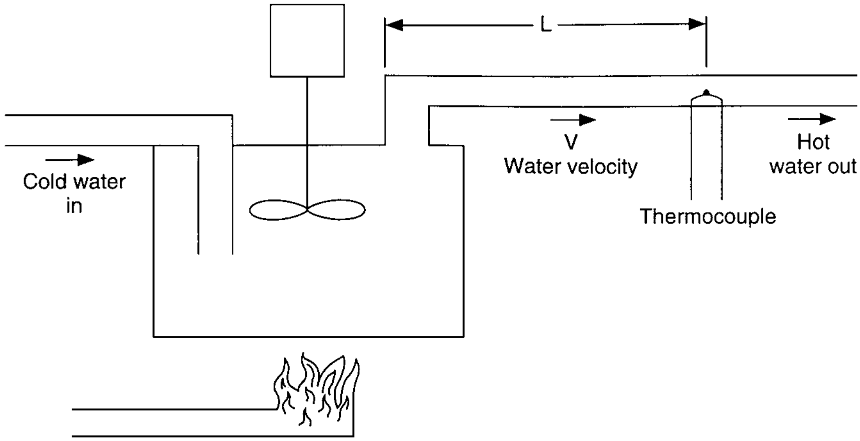


FIGURE 7 Single-capacity process with dead time. (*West Instruments*.)

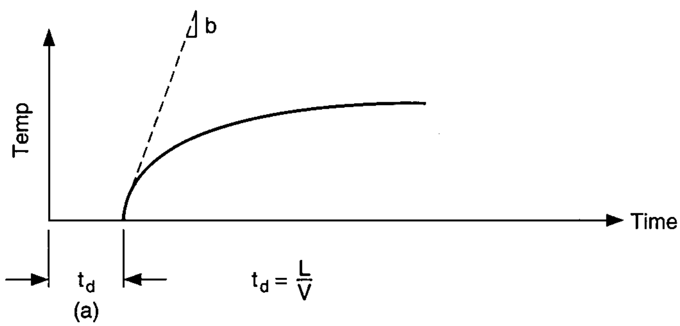


FIGURE 8 Reaction curve for single-capacity process with dead time. (*West Instruments*.)

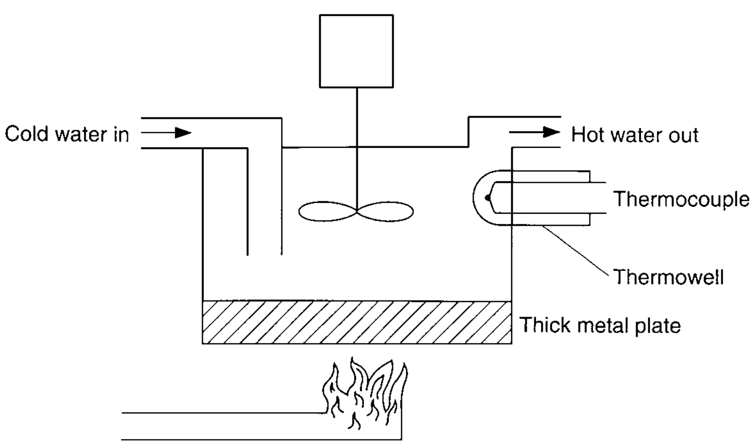


FIGURE 9 Two-capacity process. (*West Instruments*.)

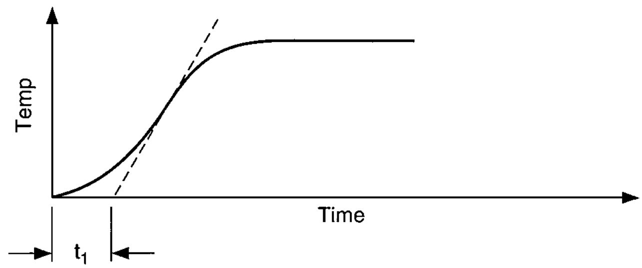


FIGURE 10 Reaction curve for two-capacity process. (*West Instruments.*)

time interval before the maximum rate of temperature rise, as shown in Fig. 10 by the intersection of the dashed vertical tangent line with the time axis. The electrical analog equivalent of this system is shown in Fig. 11. In the diagram the resistors and capacitors represent the appropriate thermal resistances and capacities of the two time constants. This system is more difficult to control than the single-capacity system since the product of time interval and maximum rate is greater.

The system shown in Fig. 9 could easily become a third-order lag or three-capacity system if there were an appreciable thermal resistance between the thermocouple and the thermowell. This could occur if the thermocouple were not properly seated against the inside tip of the well. Heat transfer from the thermowell to the thermocouple would, in this case, be through air, which is a relatively poor conductor. The temperature reaction curve for such a system is given in Fig. 12, and the electric analog for the system is shown in Fig. 13. This necessitates the addition of the R_3, C_3 time constant network.

Process Transfer Function

Another phenomenon associated with a process or system is identified as the steady-state transfer-function characteristic. Since many processes are nonlinear, equal increments of heat input do not

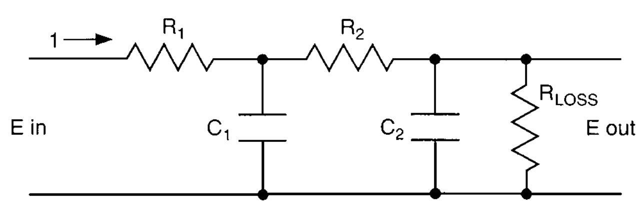


FIGURE 11 Electrical analog for two-capacity process. (*West Instruments.*)

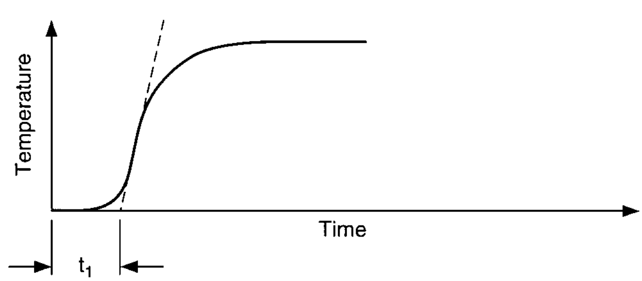


FIGURE 12 Reaction curve for three-capacity process. (*West Instruments.*)

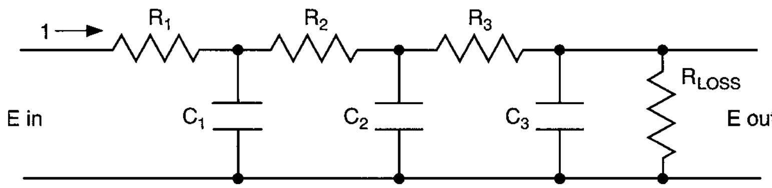


FIGURE 13 Electrical analog for three-capacity process. (*West Instruments*.)

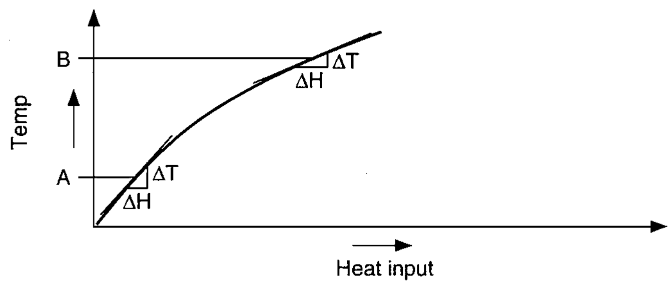


FIGURE 14 Transfer curve for endothermic process. As the temperature increases, the slope of the tangent line to the curve has a tendency to decrease. This usually occurs because of increased losses through convection and radiation as the temperature increases. This process *gain* at any temperature is the slope of the transfer function at that temperature. A steep slope (high $\Delta T/\Delta H$) is a high gain; a low slope (low $\Delta T/\Delta H$) is a low gain. (*West Instruments*.)

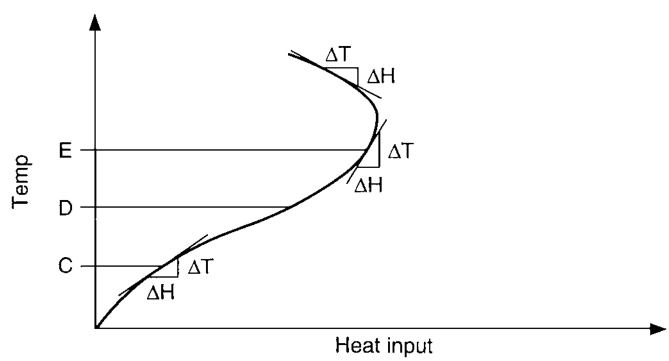


FIGURE 15 Transfer curve for exothermic process. This curve follows the endothermic curve up to the temperature level D. At this point the process has the ability to begin generating some heat of its own. The slope of the curve from this point on increases rapidly and may even reverse if the process has the ability to generate more heat than it loses. This is a negative gain since the slope $\Delta T/\Delta H$ is negative. This situation would actually require a negative heat input, or cooling action. This type of application is typical in a catalytic reaction process. If enough cooling is not supplied, the process could run away and result in an explosion. Production of plastics from the monomer is an example. Another application of this type is in plastics extrusion, where heat is required to melt the plastic material, after which the frictional forces of the screw action may provide more than enough process heat. Cooling is actually required to avoid overheating and destruction of the melt material. (*West Instruments*.)

necessarily produce equal increments in temperature rise. The characteristic transfer-function curve for a process is generated by plotting temperature against heat input under constant heat input conditions. Each point on the curve represents the temperature under stabilized conditions, as opposed to the reaction curve, which represents the temperature under dynamic conditions. For most processes this will not be a straight-line, or linear, function. The transfer-function curve for a typical endothermic process is shown in Fig. 14, that for an exothermic process in Fig. 15.

CONTROL MODES

Modern industrial controllers are usually made to produce one, or a combination of, control actions (modes of control). These include (1) on-off or two-position control, (2) proportional control, (3) proportional plus integral control, (4) proportional plus derivative (rate action) control, and (5) proportional plus integral plus derivative (PID) control.

On-Off Control Action

An on-off controller operates on the manipulated variable only when the temperature crosses the set point. The output has only two states, usually fully on and fully off. One state is used when the temperature is anywhere above the desired value (set point), and the other state is used when the temperature is anywhere below the set point.

Since the temperature must cross the set point to change the output state, the process temperature will be continually cycling. The peak-to-peak variation and the period of the cycling are mainly dependent on the process response and characteristics. The time-temperature response of an on-off controller in a heating application is shown in Fig. 16, the ideal transfer-function curve for an on-off controller in Fig. 17.

The ideal on-off controller is not practical because it is subject to process disturbances and electrical interference, which could cause the output to cycle rapidly as the temperature crosses the set point. This condition would be detrimental to most final control devices, such as contactors and valves. To

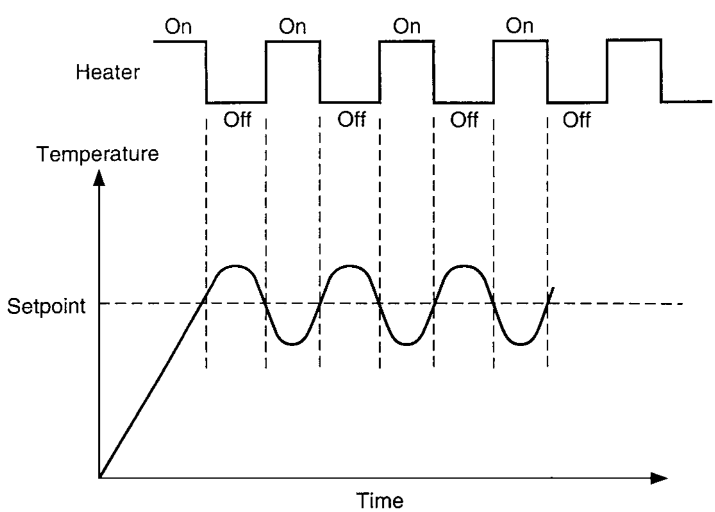


FIGURE 16 On-off temperature control action. (*West Instruments*.)

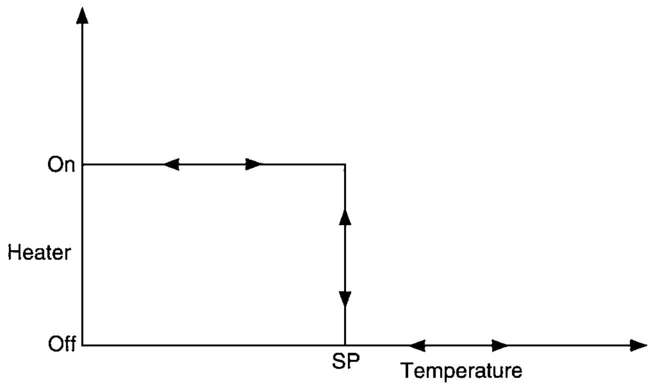


FIGURE 17 Ideal transfer curve for on-off control. (*West Instruments*.)

prevent this, an on-off differential or "hysteresis" is added to the controller function. This function requires that the temperature exceed the set point by a certain amount (half the differential) before the output will turn off again. Hysteresis will prevent the output from chattering if the peak-to-peak noise is less than the hysteresis. The amount of hysteresis determines the minimum temperature variation possible. However, process characteristics will usually add to the differential. The time-temperature diagram for an on-off controller with hysteresis is shown in Fig. 18. A different representation of the hysteresis curve is given in the transfer function of Fig. 19.

Proportional Control

A proportional controller continuously adjusts the manipulated variable so that the heat input to the process is approximately in balance with the process heat demand. In a process using electric heaters, the proportional controller adjusts the heater power to be approximately equal to the process

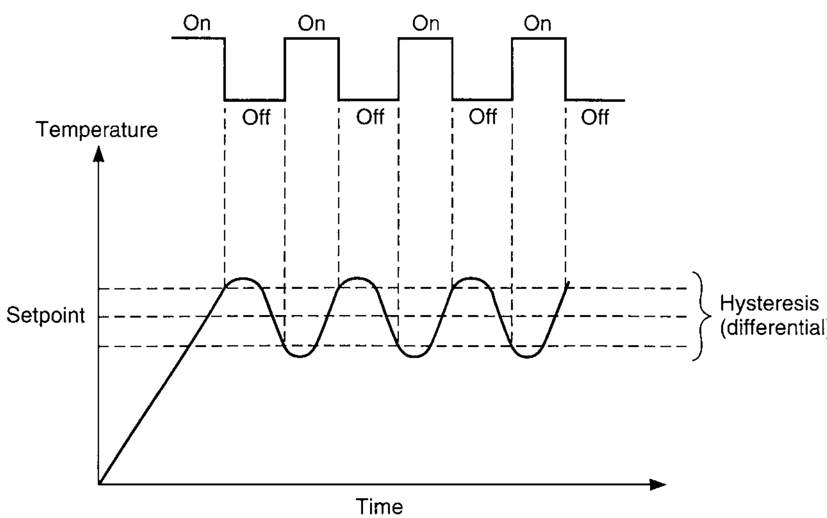


FIGURE 18 Time-temperature diagram for on-off controller with hysteresis. Note how the output changes state as the temperature crosses the hysteresis limits. The magnitude, period, and shape of the temperature curve are largely process-dependent. (*West Instruments*.)

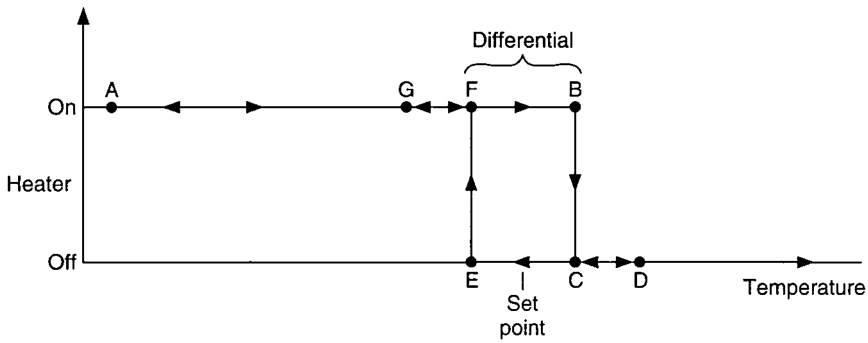


FIGURE 19 Another representation of the hysteresis curve—transfer function of on-off controller with hysteresis. Assuming that the process temperature is well below the set point at start-up, the system will be at *A*, the heat will be on. The heat will remain on as the temperature goes from *A* through *F* to *B*, the output turns off, dropping to point *C*. The temperature may continue to rise slightly to point *D* before decreasing to point *E*. At *E* the output once again turns on. The temperature may continue to drop slightly to point *G* before rising to *B* and repeating the cycle. (*West Instruments*.)

heat requirements to maintain a stable temperature. The range of temperature over which power is adjusted from 0 to 100% is called the proportional band. This band is usually expressed as a percentage of the instrument span and is centered about the set point. Thus in a controller with a 1000°C span, a 5% proportional band would be 50°C wide and extend 25°C below the set point to 25°C above the set point. A graphic illustration of the transfer function for a reverse-acting controller is given in Fig. 20.

The proportional band in general-purpose controllers is usually adjustable to obtain stable control under differing process conditions. The transfer curve of a wide-band proportional controller is shown in Fig. 21. Under these conditions a large change in temperature is required to produce a small change in output. The transfer curve of a narrow-band proportional controller is shown in Fig. 22. Here a

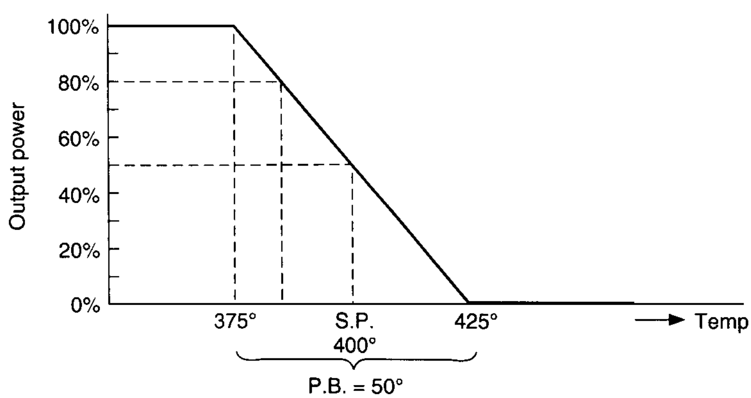


FIGURE 20 Transfer curve of reverse-acting controller. The unit is termed reverse-acting because the output decreases with increasing temperature. In this example, below 375°C, the lower edge of the proportional band, the output power is on 100%. Above 425°C the output power is off. Between these band edges the output power for any process temperature can be found by drawing a line vertically from the temperature axis until it intersects the transfer curve, then horizontally to the power axis. Note that 50% power occurs when the temperature is at the set point. The width of the proportional band changes the relationship between temperature deviation from set point and power output. (*West Instruments.*)

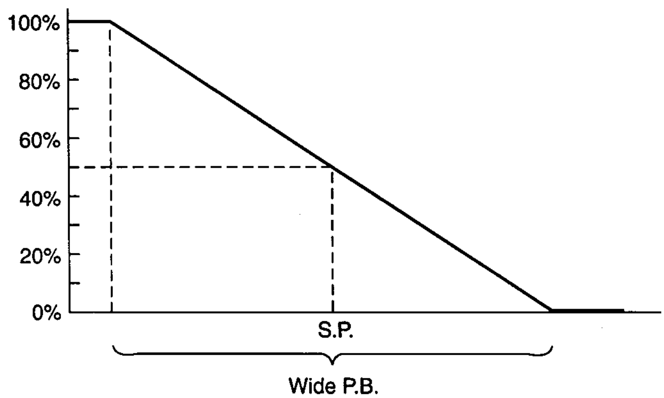


FIGURE 21 Transfer function for wide-band proportional controller. (*West Instruments.*)

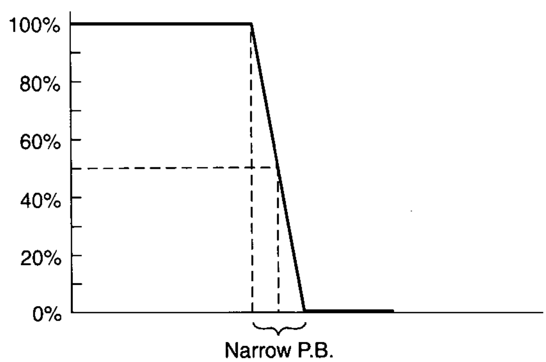


FIGURE 22 Transfer function for narrow-band proportional controller. (*West Instruments.*)

small change in temperature produces a large change in output. If the proportional band were reduced to zero, the result would be an on-off controller.

In industrial applications the proportional band is expressed as a percent of span, but it may also be expressed as controller gain in others. Proportional band and controller gain are related inversely by the equation

$$\text{Gain} = \frac{100\%}{\text{proportional band (\%)}}$$

Thus narrowing the proportional band increases the gain. For example, for a gain of 20 the proportional band is 5%. The block diagram of a proportional controller is given in Fig. 23. The temperature signal from the sensor is amplified and may be used to drive a full-scale indicator, either an analog meter or a digital display. If the sensor is a thermocouple, cold junction compensation circuitry is incorporated in the amplifier. The difference between the process measurement signal and the set point is taken in a summing circuit to produce the error or deviation signal. This signal is positive when the process is below the set point, zero when the process is at the set point, and negative when the process is above the set point. The error signal is applied to the proportioning circuit through a potentiometer gain control.

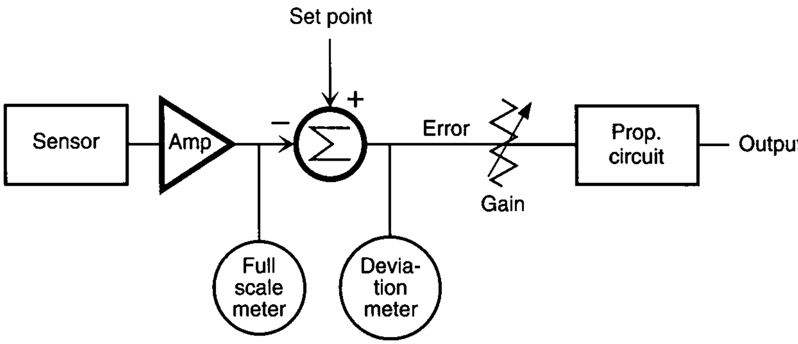


FIGURE 23 Block diagram of proportional controller. (*West Instruments.*)

The proportional output is 50% when the error signal is zero, that is, the process is at the set point.

Offset

It is rare in any process that the heat input to maintain the set-point temperature will be 50% of the maximum available. Therefore the temperature will increase or decrease from the set point, varying the output power until an equilibrium condition exists. The temperature difference between the stabilized temperature and the set point is called offset. Since the stabilized temperature must always be within the proportional band if the process is under control, the amount of offset can be reduced by narrowing the proportional band. However, the proportional band can be narrowed only so far before instability occurs. An illustration of a process coming up to temperature with an offset is shown in Fig. 24. The mechanism by which offset occurs with a proportional controller can be illustrated by superimposing the temperature controller transfer curve on the process transfer curve, as shown in Fig. 25.

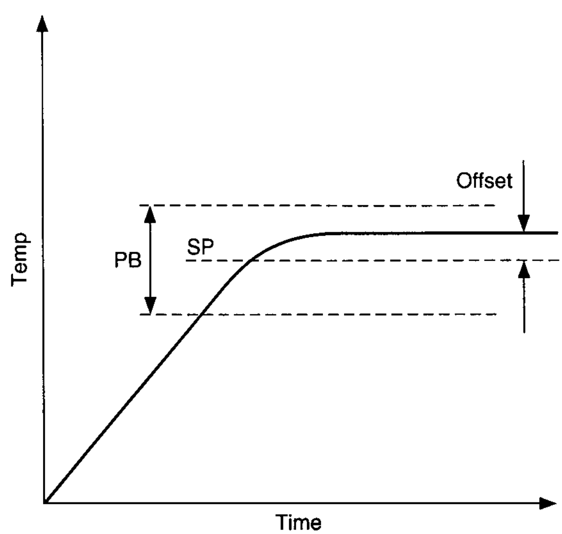


FIGURE 24 Process of coming up to temperature with an off-set (*West Instruments.*)

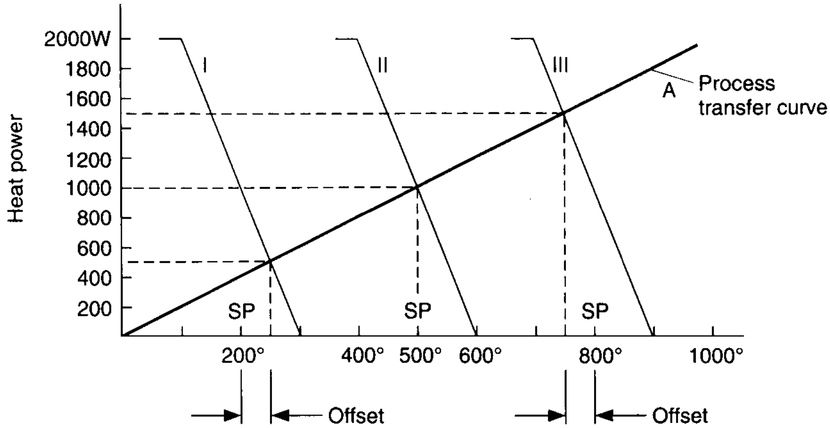


FIGURE 25 Mechanism by which offset occurs with a proportional controller. Assume that a process is heated with a 2000-watt heater. The relationship between heat input and process temperature, shown by curve *A*, is assumed to be linear for illustrative purposes. The transfer function for a controller with a 200°C proportional band is shown for three different set points in curves I, II, and III. Curve I with a set point of 200°C intersects the process curve at a power level of 500 watts, which corresponds to a process temperature of 250°C. The offset under these conditions is 250 to 200°C, or 50°C high. Curve II with a set point of 500°C intersects the process curve at 1000 watts, which corresponds to a process temperature of 500°C. There is no offset case since the temperature corresponds to the 50% power point. Curve III with a set point of 800°C intersects the process curve at 1500 watts, which corresponds to a temperature of 750°C. The off-set under these conditions is 750 to 800°C, or 50°C low. These examples show that the offset is dependent on the process transfer function, the proportional band (gain), and the set point. (*West Instruments.*)

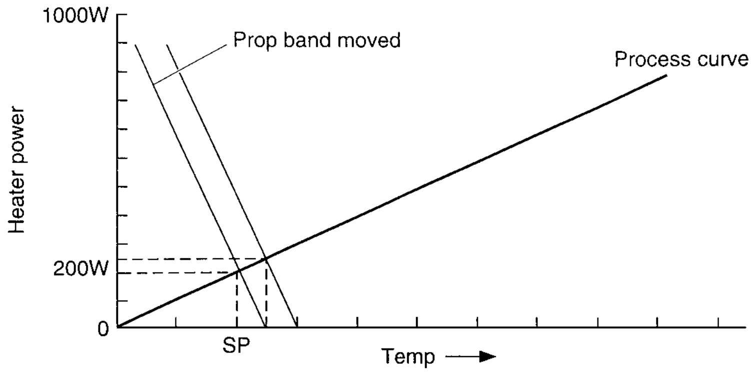
Manual and Automatic Reset

Offset can be removed either manually or automatically. In analog instrumentation, manual reset uses a potentiometer to offset the proportional band electrically. The amount of proportional band shifting must be done by the operator in small increments over a period of time until the controller power output just matches the process heat demand at the set-point temperature (Fig. 26). A controller with manual reset is shown in the block diagram of Fig. 27.

Automatic reset uses an electronic integrator to perform the reset function. The deviation (error) signal is integrated with respect to time and the integral is summed with the deviation signal to move the proportional band. The output power is thus automatically increased or decreased to bring the process temperature back to the set point. The integrator keeps changing the output power, and thus the process temperature, until the deviation is zero. When the deviation is zero, the input to the integrator is zero and its output stops changing. The integrator has now stored the proper value of reset to hold the process at the set point. Once this condition is achieved, the correct amount of reset value is held by the integrator. Should process heat requirements change, there would once again be a deviation, which the integrator would integrate and apply corrective action to the output. The integral term of the controller acts continuously in an attempt to make the deviation zero. This corrective action has to be applied rather slowly, more slowly than the speed of response of the load. Otherwise oscillations will occur.

Automatic Reset—Proportional plus Integral Controllers

Automatic reset action is expressed as the integral time constant. Precisely defined, the reset time constant is the time interval in which the part of the output signal due to the integral action increases



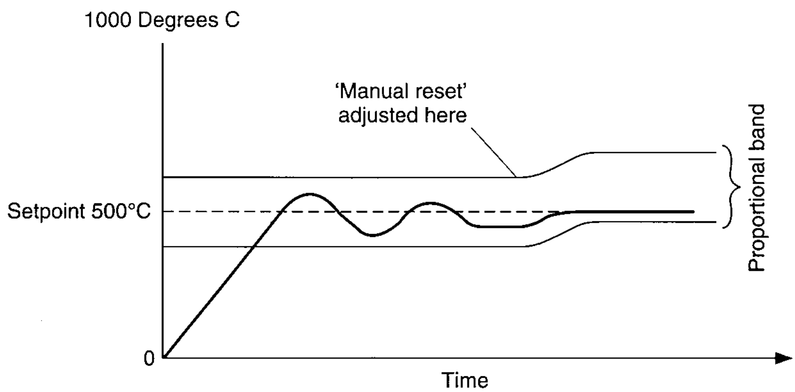


FIGURE 26 Manual reset of proportional controller. (*West Instruments.*)

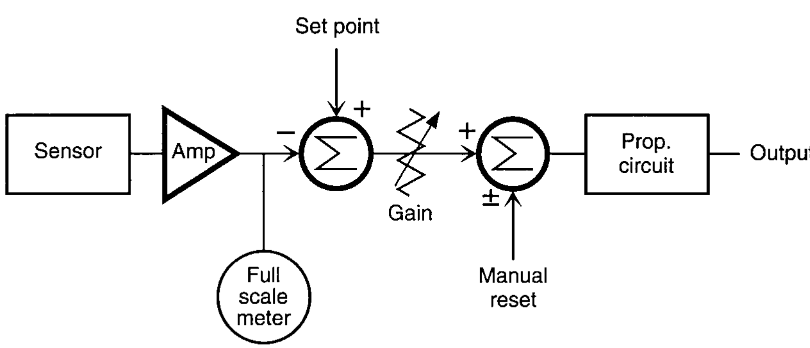


FIGURE 27 Block diagram of proportional controller with manual reset. (*West Instruments.*)

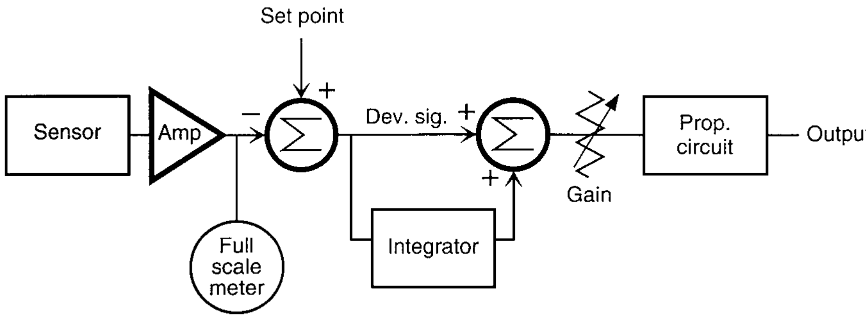


FIGURE 28 Block diagram of proportional plus integral controller. (*West Instruments.*)

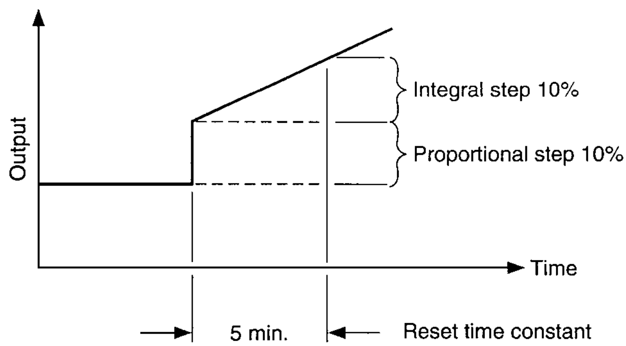


FIGURE 29 Reset time definition. (*West Instruments.*)

by an amount equal to the part of the output signal due to the proportional action, when the deviation is unchanging. A controller with automatic reset is shown in the block diagram of Fig. 28.

If a step change is made in the set point, the output will immediately increase, as shown in Fig. 29. This causes a deviation error, which is integrated and thus produces an increasing change in controller output. The time required for the output to increase by another 10% is the reset time—5 minutes in the example of Fig. 29.

Automatic reset action also may be expressed in repeats per minute and is related to the time constant by the inverse relationship

$$\text{Repeats per minute} = \frac{1}{\text{integral time constant (minutes)}}$$

Integral Saturation

A phenomenon called integral saturation is associated with automatic reset. Integral saturation refers to the case where the integrator has acted on the error signal when the temperature is outside the proportional band. The resulting large output of the integrator causes the proportional band to move so far that the set point is outside the band. The temperature must pass the set point before the controller output will change. As the temperature crosses the set point, the deviation signal polarity changes and the integrator output starts to decrease or desaturate. The result is a large temperature overshoot. This can be prevented by stopping the integrator from acting if the temperature is outside the proportional band. This function is called integral lockout or integral desaturation.

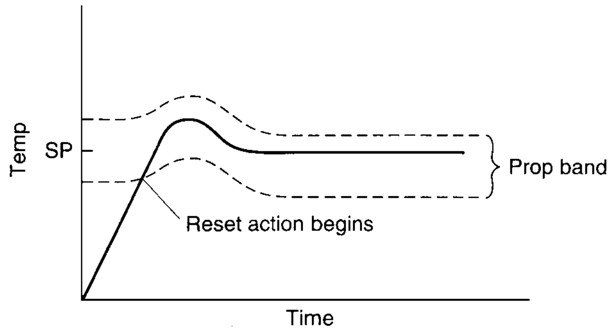


FIGURE 30 Proportional plus integral action. (*West Instruments.*)

One characteristic of all proportional plus integral controllers is that the temperature often overshoots the set point on start-up. This occurs because the integrator begins acting when the temperature reaches the lower edge of the proportional band. As the temperature approaches the set point, the reset action already has moved the proportional band higher, causing excess heat output. As the temperature exceeds the set point, the sign of the deviation signal reverses and the integrator brings the proportional band back to the position required to eliminate the offset (Fig. 30).

Derivative Action (Rate Action)

The derivative function in a proportional plus derivative controller provides the controller with the ability to shift the proportional band either up or down to compensate for rapidly changing temperature. The amount of shift is proportional to the rate of temperature change. In modern instruments this is accomplished electronically by taking the derivative of the temperature signal and summing it with the deviation signal (Fig. 31(*a*)). [Some controllers take the derivative of the deviation signal, which has the side effect of producing upsets whenever the set point is changed (Fig. 31(*b*)).]

The amount of shift is also proportional to the derivative time constant. The derivative time constant may be defined as the time interval in which the part of the output signal due to proportional action increases by an amount equal to that part of the output signal due to derivative action when the deviation is changing at a constant rate (Fig. 32).

Derivative action functions to increase controller gain during temperature changes. This compensates for some of the lag in a process and allows the use of a narrower proportional band with its lesser offset. The derivative action can occur at any temperature, even outside the proportional band, and is not limited as is the integral action. Derivative action also can help to reduce overshoot on start-up.

Proportional plus Integral plus Derivative Controllers

A three-mode controller combines the proportional, integral, and derivative actions and is usually required to control difficult processes. The block diagram of a three-mode controller is given in Fig. 33. This system has a major advantage. In a properly tuned controller, the temperature will approach the set point smoothly without overshoot because the derivative plus deviation signal in the integrator input will be just sufficient for the integrator to store the required integral value by the time the temperature reaches the set point.

Time- and Current-Proportioning Controllers

In these controllers the controller proportional output may take one of several forms. The more common forms are time-proportioning and current-proportioning. In a time-proportioning output,

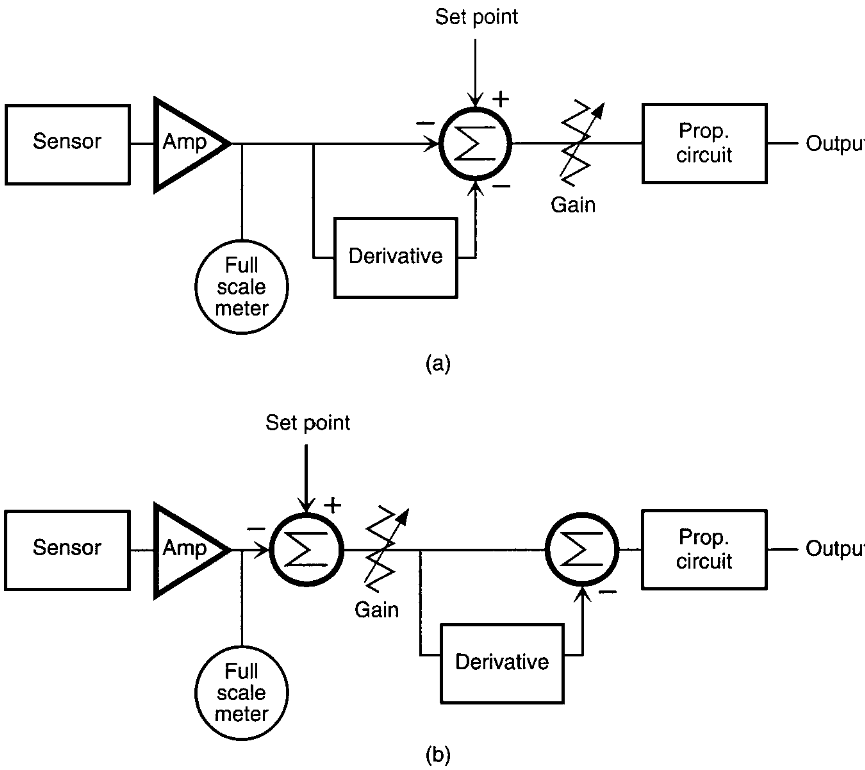


FIGURE 31 Block diagram or proportional plus rate controller. (*a*) The derivative of the sensor (temperature) signal is taken and summed with the deviation signal. (*b*) The derivative of the deviation signal is taken. (*West Instruments.*)

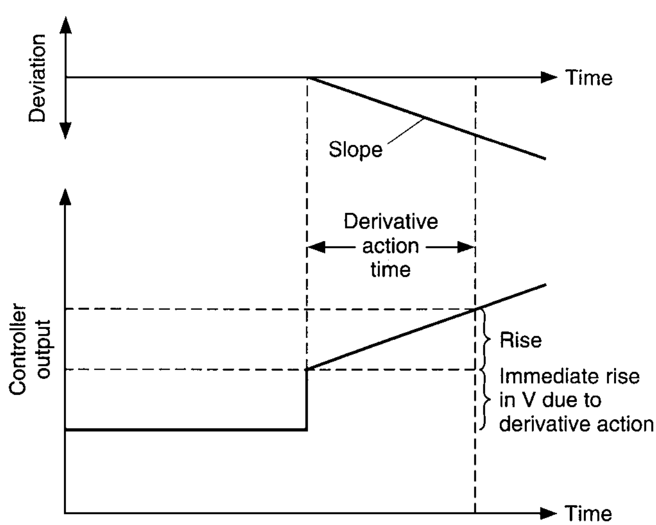


FIGURE 32 Derivative time definition. (*West Instruments.*)

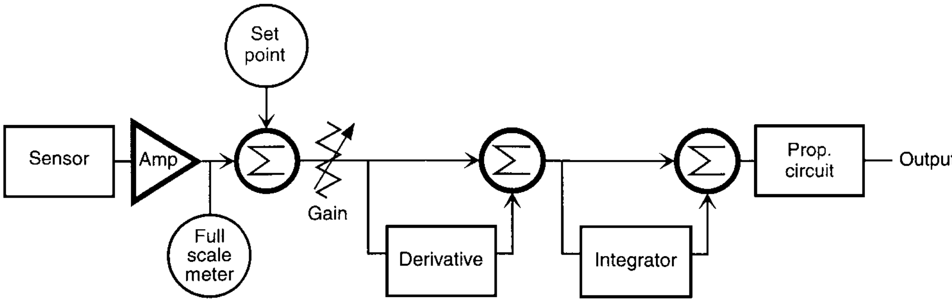


FIGURE 33 Proportional plus integral plus derivative controller. (*West Instruments.*)

power is applied to the load for a percentage of a fixed cycle time. Figure 34 shows the controller output at a 75% output level for a cycle time of 12 seconds.

This type of output is common with contractors and solid-state devices. An advantage of solid-state devices is that the cycle time may be reduced to 1 second or less. If the cycle time is reduced to one-half the line period (10 ms for 50 Hz), then the proportioning action is sometimes referred to as a stepless control, or phase-angle control. A phase-angle-fired output is shown in Fig. 35.

The current output, commonly 4 to 20 mA, is used to control a solid-state power device, a motor-operated valve positioner, a motor-operated damper, or a saturable core reactor. The relationship between controller current output and heat output is shown in Fig. 36.

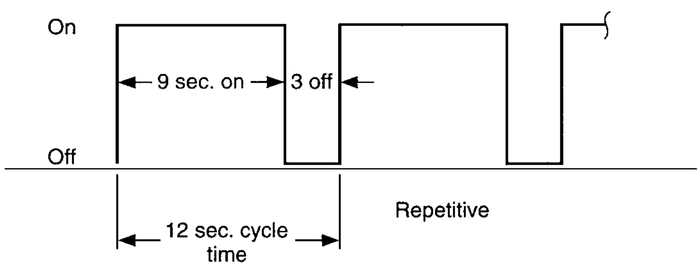


FIGURE 34 Time-proportioning controller at 75% level. (*West Instruments.*)

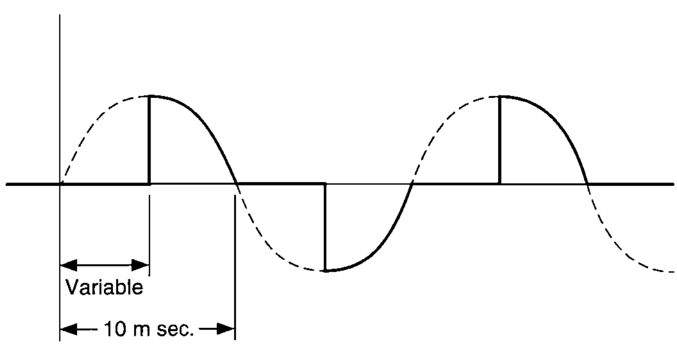


FIGURE 35 Phase-angle-fired stepless control ouput. (*West Instruments.*)

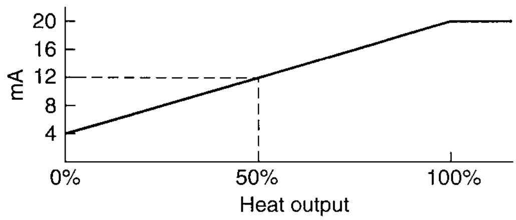


FIGURE 36 Current-proportioning controller. (*West Instruments.*)

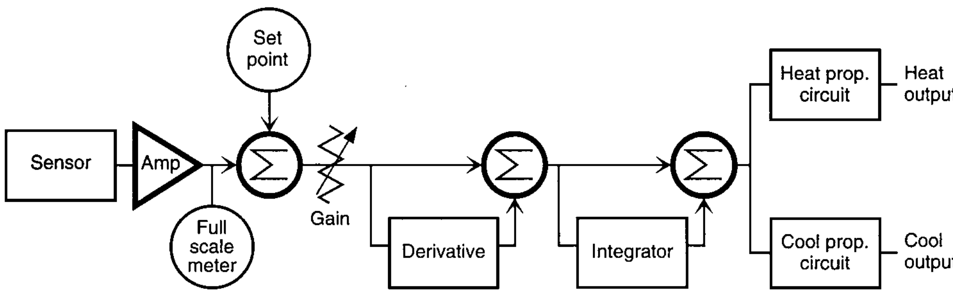


FIGURE 37 Heat-cool PID controller. (*West Instruments.*)

Heat-Cool PID Control

Certain applications that are partially exothermic demand the application of cooling as well as heating. To achieve this, the controller output is organized as shown in Fig. 37. The controller has two proportional outputs, one for heating and one for cooling.

The transfer function for this type of controller is shown in Fig. 38. Below the proportional band, full heating is applied; above the proportional band, full cooling is applied. Within the proportional band (X_{p1}) there is a linear reduction of heating to zero, followed by a linear increase in cooling with increasing temperature. Heating and cooling can be overlapped (X_{sh}) to ensure a smooth transition between heating and cooling. In addition, to optimize the gain between heating and cooling action, the cooling gain is made variable (X_{p2}).

PROCESS CONTROL CHARACTERISTICS AND CONTROLLER SELECTION

The selection of the most appropriate controller for a given application depends on several factors, as described in the introduction to this article. The process control characteristics are very important criteria and are given further attention here. Experience shows that for easier controller tuning and lowest initial cost, the simplest controller that will meet requirements is usually the best choice. In selecting a controller, the user should consider priorities. In some cases precise adherence to the control point is paramount. In other cases maintaining the temperature within a comparatively wide range is adequate.

In some difficult cases the required response cannot be obtained even with a sophisticated controller. This type of situation indicates that there is an inherent process thermal design problem. Thermal design should be analyzed and corrected before proceeding with controller selection. A good thermal

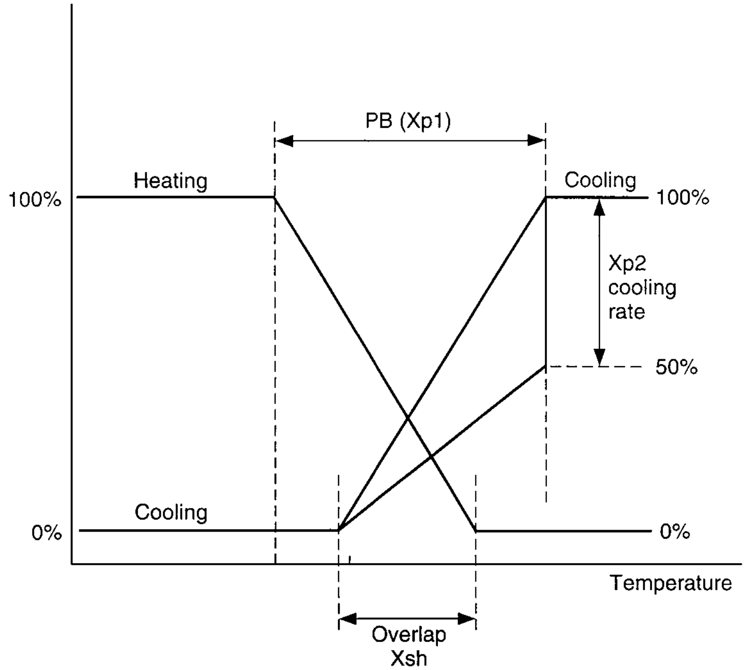


FIGURE 38 Transfer function for heat-cool PID controller. The controller has two proportional outputs, one for heating and one for cooling. (*West Instruments.*)

design will provide more stable control and allow the use of a less complicated and usually less expensive controller.

Controller Selection

Selection of the controller type may be approached from several directions:

1. Process reaction curve
2. Physical thermal system analysis
3. Previous experience
4. Experimental testing

The process reaction curve may be generated and observed to classify the process as easy or difficult to control, single capacity, or multicapacity. This knowledge should be compared with the process temperature stability requirements to indicate which type of controller to use.

The process controllability may be estimated by observing and analyzing the process thermal system. What is the relative heater power to load heat requirements? Are the heaters oversized or undersized? Oversized heaters lead to control stability problems. Undersized heaters produce slow response. Is the thermal mass large or small? What are the distance and the thermal resistance from the heaters to the sensor? Large distances and resistances cause lag and a less stable system. Comparing the controllability with the process temperature stability requirements will indicate which type of controller to use. This same system of analysis can be applied to process variables other than temperature.

Prior experience often is an important guideline, but because of process design changes, a new situation may require tighter or less stringent control. A method often used is to try a simple controller, such as proportional plus manual reset, and to note the results compared with the desired system response. This will suggest additional features or features that may be deleted.

Single-Capacity Processes

If the process reaction curve or system examination reveals that the process can be classified as single-capacity, it may be controlled by an on-off controller. However, two conditions must be met: (1) a cyclical peak-to-peak temperature variation equal to the controller hysteresis is acceptable, and (2) the process heating and cooling rates are long enough to prevent too rapid cycling of the final control devices. Controller hysteresis also has an effect on the period of temperature cycling. Wider hysteresis causes a longer period and greater temperature variation. A narrow hysteresis may be used with final control devices, such as solid-state relays, triacs, and SCRs, which can cycle rapidly without shortening their life. Typical system responses for oversized and undersized heater capacity are shown in Figs. 39 and 40, respectively.

If the previously mentioned two conditions are not acceptable, then the use of a proportional controller is indicated. A proportional controller would eliminate the temperature cycling. In a controller with adjustable proportional band, the band usually may be adjusted quite narrow and still maintain stability so that offset will not be a problem. If the controller has a fixed proportional band, at a value much larger than optimum, the resulting offset may be undesirable. Manual reset may be added to reduce the offset. A narrow proportional band will make the offset variations minimal with changes in process heat requirements so that automatic reset usually will not be required.

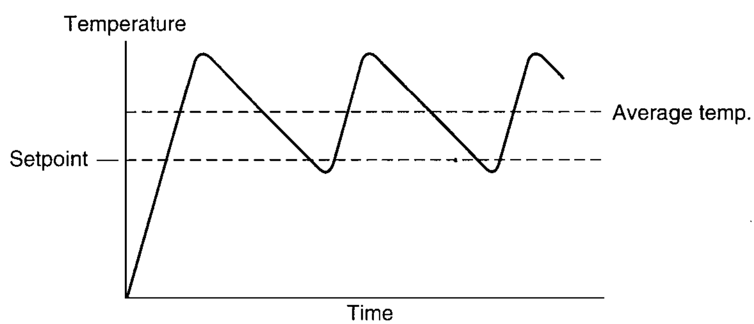


FIGURE 39 Typical system response for oversized heater condition. (*West Instruments.*)

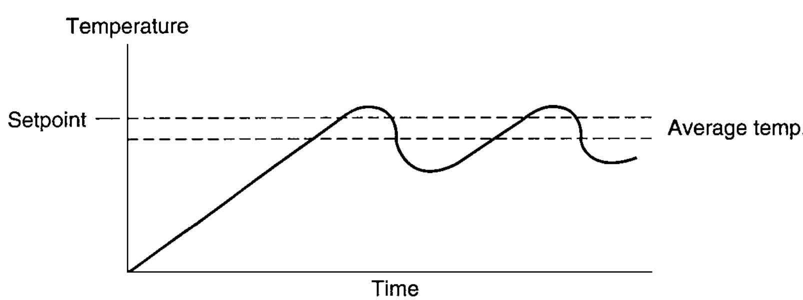


FIGURE 40 Typical system response for undersized heater condition. (*West Instruments.*)

A single-capacity process usually will not require derivative action. However, control action during process upsets may be improved by the addition of some derivative action. Adding too much derivative action (too long a derivative time constant) can cause instability with some controllers.

Multicapacity Processes

A multicapacity process or a single-capacity process with transport delay is generally not suited to on-off control because of the wide temperature cycling. These processes require proportional control. Depending on the process difficulty, as evidenced by the process reaction curve and the control precision requirements, a proportional controller or one with the addition of derivative and integral action will be required.

Proportional controllers must be "tuned" to the process for good temperature response. The question is—what is good response? Three possible temperature responses under cold start-up conditions are shown in Fig. 41.

Control systems usually are tuned under operating conditions rather than for start-up conditions. Tuning a controller requires first that the process temperature be stable near the operating point with the system in operation. Then a known process disturbance is caused and the resulting temperature response observed. The response is best observed on a recorder. The proper disturbance for tuning is one which is likely to occur during actual operation, such as product flow change or speed change. However, this may be impractical and thus a small set-point change usually is used as the disturbance. The optimum tuning for set-point changes may not produce optimum response for various process disturbances.

Process characteristics for a proportional-only controller are given in Fig. 42. The curves show the resulting temperature change after decreased process heat demand. Similar curves would result if the set point were decreased several degrees.

Processes with long time lags and large maximum rates of rise, such as a heat exchanger, require a wide proportional band to eliminate oscillation. A wide band means that large offsets can occur with changes in load. These offsets can be eliminated by the automatic reset function in a proportional plus integral controller. The system response curve will be similar to those shown in Fig. 43 for a decrease in heat demand.

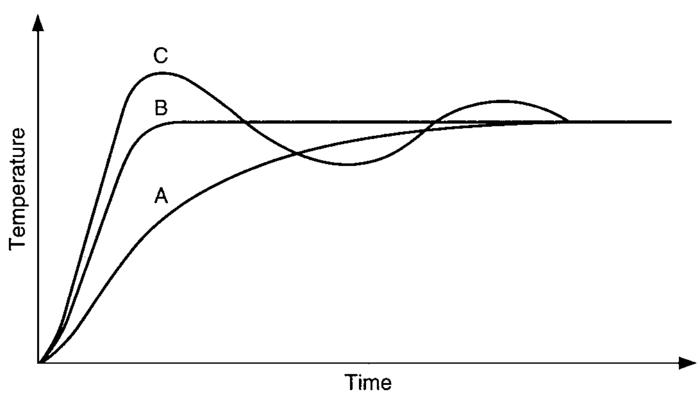


FIGURE 41 Three possible temperature responses of proportional controllers under cold start-up conditions. Curve *A* could be considered good response if a slow controlled heat-up is required. Curve *B* would be considered good response if the fastest heat-up without overshoot is required. Curve *C* could be good response if the fastest heat-up is required. The definition of "good response" varies with the process and operational requirements. (*West Instruments.*)

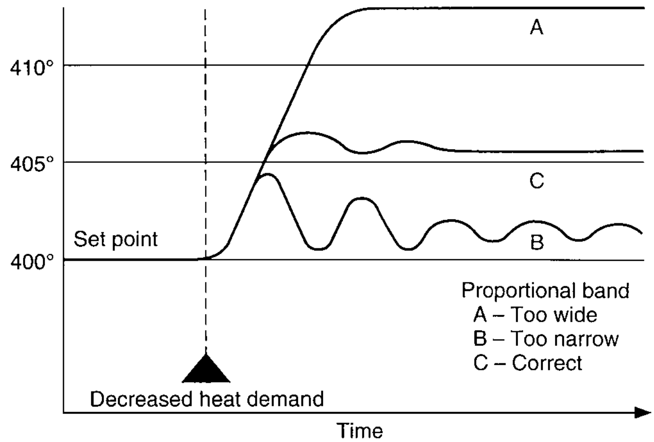


FIGURE 42 Process characteristics for a proportional-only controller. Curve *A* results when the proportional band is too wide. Note the large offset. The offset can be reduced by narrowing the proportional band. Instability results if the proportional band is too narrow, as shown by curve *B*. Optimum control, as shown by curve *C*, is achieved at a proportional band setting slightly wider than that which causes oscillation. If process parameters change with time or if operating conditions change, it will be necessary to retune the controller or avoid this by using a proportional band wider than optimum to prevent future instability. (*West Instruments.*)

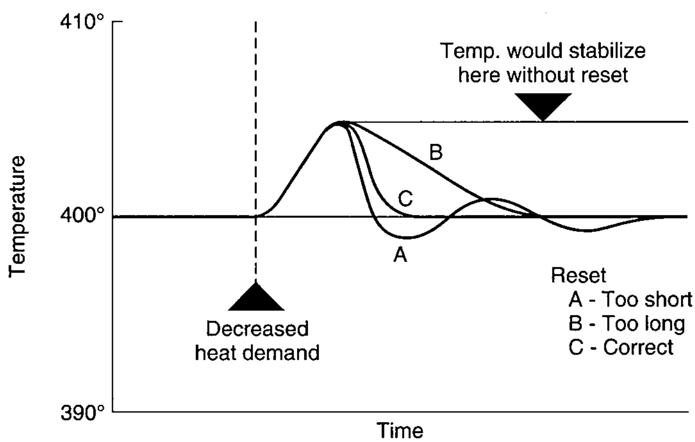


FIGURE 43 System response curves for proportional plus integral controller for an application such as a heat exchanger when there is a decrease in heat demand. An integral time constant which is too long for the process will take a long time to return the temperature to set point, as shown in curve *B*. An integral time constant that is too short will allow integral to outrun the process, causing the temperature to cross the set point with damped oscillation, as shown by curve *A*. If the integral time is much too long, continuous oscillation results. The integral time constant usually considered optimum is that which returns the temperature to set point as rapidly as possible without overshooting it, as shown by curve *C*. However, a damped oscillation (curve *A*) may be more desirable if the temperature must return to set point faster and some overshoot can be allowed. (*West Instruments.*)

Derivative (rate) action may be used to advantage on processes with long time delays, speeding recovery after a process disturbance. The derivative provides a phase lead function, which cancels some of the process lag and allows the use of a narrower proportional band without creating instability. A narrower proportional band results in less offset. The response of a proportional plus derivative controller in a system is dependent not only on the proportional band and the derivative time constant, but also on the method used to obtain the derivative signal. The curves of Fig. 44 show some typical results of the proportional plus derivative control algorithm as used in the West MC30 and previous controllers as well as most other brands of proportional plus derivative controllers. Figure 45 shows the response to a decrease in heat demand for the controllers described previously (Fig. 44).

As noted from the aforementioned illustrations, the problem of superimposed damping or continuous oscillation remains. This condition may be corrected by either decreasing the derivative time constant or widening the proportional band. The oscillations result from a loop gain that is too great at the frequency of oscillation. The total loop gain is the process gain times the proportional gain times the derivative gain. Decreasing any one of these gains will decrease the total loop gain and return stability.

The proportional plus derivative controller may be used to advantage on discontinuous processes such as batching operations involving periodic shutdown, emptying, and refiling. Here the proportional plus integral controller would not perform well because of the long time lags and intermittent operation. Derivative action also reduces the amount of overshoot on start-up of a batch operation.

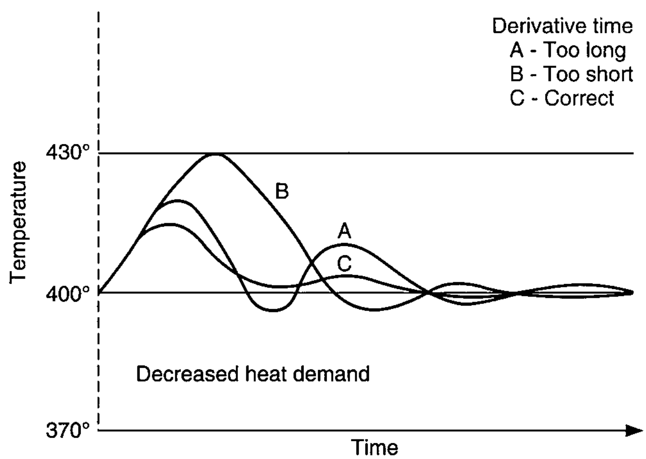


FIGURE 44 Response curves for controller using proportional plus derivative algorithm. A derivative time constant that is too long causes the temperature to change too rapidly and overshoot the set point with damped oscillation (curve *A*). A derivative time constant that is too short allows the temperature to remain away from the set point too long (curve *B*). The optimum derivative time returns the temperature to set point with a minimum of ringing (curve *C*). The damped oscillation about the final value in curve *A* can be due to excessive derivative gain at frequencies above the useful control range. Some controllers have an active derivative circuit which decreases the gain above the useful frequency range of the system and provides full phase lead in the useful range. The results of too short a derivative time constant remain the same as shown, not enough compensation for process lags. However, this method improves response at the optimun derivative time constant. It also produces two more possible responses if the time constant is longer then optimum. (*West Instruments.*)

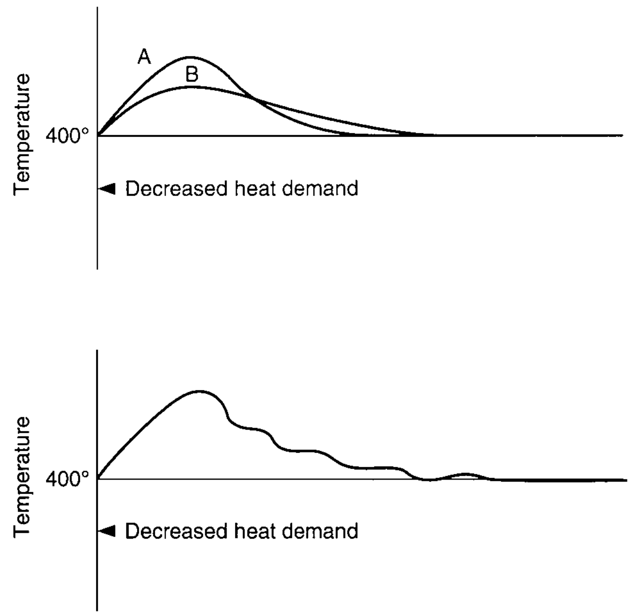


FIGURE 45 Response to a decrease in heat demand for the controller of Fig. 44. Curve *A* (top diagram) is for the optimum derivative time constant. The temperature returns smoothly to set point. Curve *B* shows one possibility for a derivative time constant which is too long. The temperature deviation is less, but the temperature returns to set point on the derivative time constant curve. The curve in the bottom diagram shows another possibility for a derivative time constant which is too long. The temperature returns to set point on the derivative time constant curve, but has either damped or continuous oscillation superimposed. (*West Instruments.*)

The most difficult processes to control, those with long time lags and large maximum rates of rise, require three-mode or proportional plus integral plus derivative (PID) controllers. The fully adjustable PID controller can be adjusted to produce a wide variety of system temperature responses from very underdamped through critically damped to very overdamped.

The tuning of a PID control system will depend on the response required and also on the process disturbance to which it applies. Set-point changes will produce a different response from process disturbances. The type of process disturbance will vary the type of response. For example, a product flow rate change may produce an underdamped response while a change in power line voltage may produce an overdamped response.

Some of the response criteria include rise time, time to first peak, percent overshoot, settling time, decay ratio, damping factor, integral of square error (ISE), integral of absolute error (IAE), and integral of time and absolute error (ITAE). Figure 46 illustrates some of these criteria with the response to an increase in set point.

In many process applications a controller tuning that produces a decay ratio of ¼ is considered good control. However, this tuning is not robust enough. Also the tuning parameters that produce a decay ratio of ¼ are not unique and neither are the responses, as illustrated in Fig. 47. In some applications the deviation from set point and the time away from set point are very important. This leads to imposing one of the integral criteria, as shown in Fig. 48.

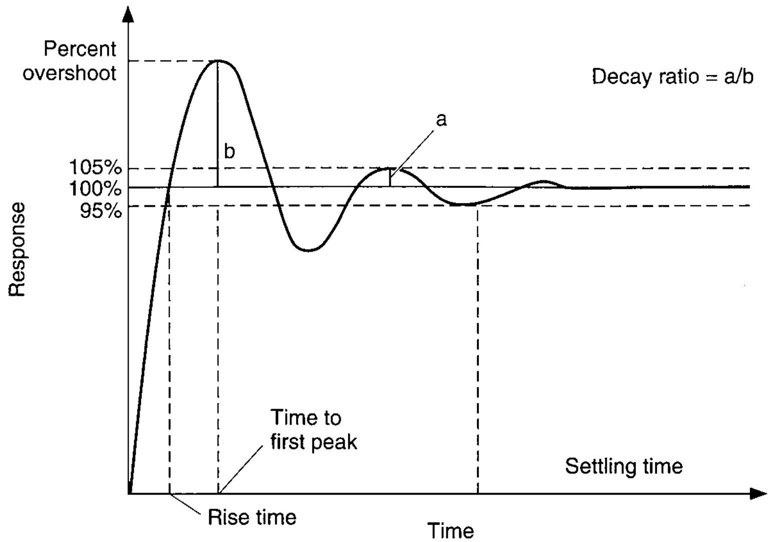


FIGURE 46 Tuning of PID control system depends on response required as well as on process disturbance to which it applies. (*West Instruments.*)

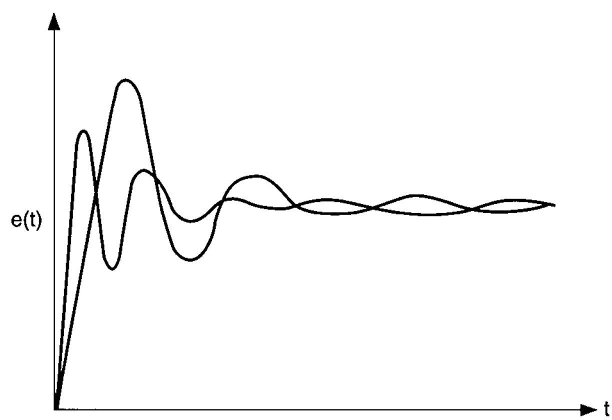


FIGURE 47 Nonunique nature of quarter decay ratio. (*West Instruments.*)

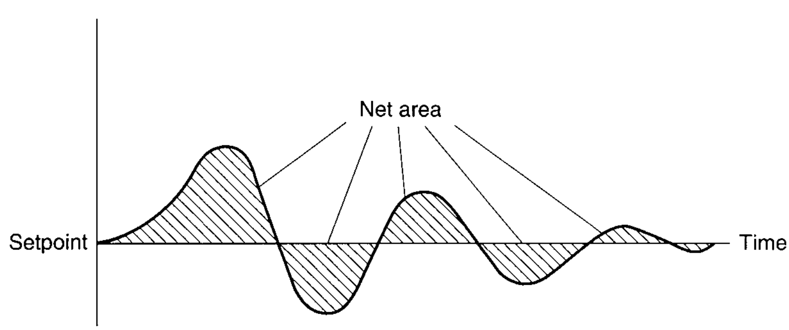


FIGURE 48 Applications of integral criteria. (*West Instruments.*)

REFERENCE

1. Astrom, K., and T. Hagglund, *PID Controllers: Theory, Design, and Tuning*, 2nd ed., Instrument Society of America, Research Triangle Park, North Carolina, 1995.

TECHNIQUES FOR PROCESS CONTROL

by Peter D. Hansen*

This article presents a number of techniques useful in the analysis and design of modern process control systems. Particular emphasis is given to transfer-function and adaptive methods, which lead to designs that cope with process delay (dead time), loop interaction, nonlinearity, and unmeasured disturbances.

The mathematical approaches described here can be used (1) by manufacturers of industrial controllers to achieve an improved and more versatile design, (2) by control engineers seeking a solution to difficult control applications, and (3) by students and researchers to achieve a more thorough understanding of control system dynamics.

An effort has been made to use consistent notation throughout this article. Uppercase letters represent transfer operators, functions of the differential operator s or the backward shift operator z^{-1}. Matrices are in boldface type. Scalar or vector variables and parameters are represented by lowercase letters. Key equations referenced throughout are the process equation [Eq. (5)], the target performance equation [Eq. (8)], the design equation [Eq. (10)], the open-loop controller equation [Eq. (11)], and the feedback controller equation [Eq. (13)].

DIGITAL CONTROL

A digital controller is generally considered to be superior to an analog controller. However, if it is used to emulate an analog controller, the digital device may be less effective because of phase (or delay) and resolution errors introduced by sampling and converting. The digital controller's advantage is its algorithmic flexibility and precision with respect to both calculations and logic, thereby facilitating on-line restructuring and parameter adaptation.

A digital control algorithm utilizes samples of its input signals which are discrete in both magnitude and time. Usually, continuous signals are sampled at a constant rate. Sampling the controlled variable introduces phase lag (effective delay) into the feedback loop because of

1. Low-pass filtering
2. Computation and transmission
3. Output holding between updates

* Bristol Fellow, Systems Development and Engineering, The Foxboro Company (a Siebe Company), Foxboro, Massachusetts.

Effective delay or parasitic lag, whether in the digital or the analog portion of a feedback loop, has an adverse effect on performance.

Loop delay, measurement noise, and output saturation determine the performance achievable with feedback control. Minimum integrated absolute error in response to an unmeasured load increases in proportion to the delay time for a dominant-delay process and in proportion to the square of delay for a dominant-lag process. Consequently the sampling-related delays should be made a small fraction of the total loop delay by using a small sampling interval.

State-Space Representation

Linear dynamic systems can be represented in terms of a state-variable vector x as a set of simultaneous first-order difference equations,

$$\begin{aligned} x\{t+h\} &= \mathbf{M}x\{t\} + \mathbf{N}u\{t\} \\ y\{t\} &= \mathbf{C}x\{t\} \end{aligned} \tag{1}$$

where h is the computing interval, or as differential equations,

$$\begin{aligned} \frac{dx}{dt} &= \mathbf{A}x + \mathbf{B}u \\ y &= \mathbf{C}x \end{aligned} \tag{2}$$

In these equations u is a vector of inputs, and y is a vector of measured variables. Matrices, but not vectors, are capitalized and in boldface type. **M** and **A** are square matrices and **N**, **C**, and **B** are rectangular (noninvertible). When a zero-order hold drives the continuous process, the representations are related at sampling instants by

$$\begin{aligned} \mathbf{M} &= e^{\mathbf{A}h} = \sum_{n=0}^{\infty} \frac{(\mathbf{A}h)^n}{n!} \\ \mathbf{N} &= (\mathbf{M} - \mathbf{I})\,\mathbf{A}^{-1}\mathbf{B} = \left[\sum_{n=0}^{\infty} \frac{(\mathbf{A}h)^n}{(n+1)!} \right] \mathbf{B}h \end{aligned} \tag{3}$$

The inversion of **A** can be avoided by replacing **M** with its Taylor series, which converges (possibly slowly) for all $\mathbf{A}h$.

The state-space approach may be used to model multivariable systems whose characteristics are time-varying and whose controlled variables are not measured directly. However, the representation may be inefficient because the matrices are often sparse. The approach can be generalized to characterize nonlinear systems by considering the right-hand sides of Eqs. (1) or (2) to be vector functions of the state variables and inputs. However, a process with a time delay cannot be represented directly with a finite differential-equation form. The difference-equation form introduces an extra state variable for each time step of delay.

Methods of analyzing observability, controllability, and stability of state-space representations are discussed in many control texts [1]–[3], as are design methods for predictors and controllers. The state-space feedback-controller design procedures lead to inflexible global control structures, which are usually linear. All manipulated variables are used to control each controllable state variable, and all measured variables are used to calculate each observable state variable. Consequently an on-line redesign (adaptive) capability may be needed to retune for process nonlinearity and to restructure following either an override condition or a loss of a measurement or manipulator.

Transfer-Operator Representation

The state-space equations can be expressed in transfer-function form, using algebraic operators to represent forward shift z and differentiation s,

$$\begin{aligned} y &= \mathbf{C}(z\mathbf{I} - \mathbf{M})^{-1}\,\mathbf{N}u \\ y &= \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\,\mathbf{B}u \end{aligned} \tag{4}$$

For a single-input, single-output time-invariant system, these equations can be expressed as

$$Ay = BDu + Ce \tag{5}$$

where y is the controlled or measured variable, u is the manipulated variable, and e is a load (or disturbance) variable. A, B, C, and D are polynomial functions of s or the backward shift $z^{-1} = e^{-hs}$. B contains stable zeros and is therefore cancelable. D may be noncancelable and has unity steady-state gain. A zero of a polynomial is a root, a value of its argument (s or z^{-1}) that causes the polynomial to be zero. Unstable (noncancelable, nonminimum phase) zeros are in the right half of the complex s plane or inside the unit circle in the complex z^{-1} plane. The delay operator, whose zeros are at the origin of the z^{-1} plane, is noncancelable and nonminimum phase. Its inverse, a time advance, is physically unrealizable.

For sinusoidal signals the differentiation operator becomes $s = j\omega$, and the backward shift becomes $z^{-1} = e^{-j\omega h}$. In steady state the radian frequency ω is zero, allowing s to be replaced with 0 and z^{-1} with 1 in the polynomial operators. The role of the C polynomial is played by an "observer" in state-space design. When e is dominated by measurement noise, e appears unfiltered at y; hence C is (almost) equal to A. When e is a load upset, e appears at y filtered by the process dynamics $1/A$; hence C is (nearly) 1. When e is considered an impulse, a process that would have a nonzero steady-state response to a steady-state e input has an additional zero at $s = 0$ or $z^{-1} = 1$ in its A and B polynomials.

An example of the conversion from the s domain to the z^{-1} domain is shown in [1]. A sampled lag $\{\tau_L\}$–delay $\{\tau_D\}$ process with gain $\{k\}$, whose input u is constant between sampling instants (because of a zero-order hold), is represented as

$$(1 - z^{-1}e^{-b})y = kz^{-n}[1 - e^{-a} + z^{-1}(e^{-a} - e^{-b})]u \tag{6}$$

where the delay is between n and $n+1$ sampling intervals, $nh < \tau_D < (n+1)h$, and

$$a = \frac{(n+1)h - \tau_D}{\tau_L}$$

$$b = \frac{h}{\tau_L}$$

When $e^{-a} - e^{-b} < 1 - e^{-a}$, the first-order factor in parentheses on the right of Eq. (6) is cancelable and can be part of B. Otherwise it must be part of D.

When b is very small, because the sampling interval is very small, Eq. (6) becomes

$$[1 - z^{-1}(1 - b)]y = \left(\frac{k}{\tau_L}\right) z^{-n}[h(n+1) - \tau_D + z^{-1}(\tau_D - nh)]u \tag{7}$$

Except for the b term on the left, this is indistinguishable from an integral $\{\tau_L/k\}$–delay$\{\tau_D\}$ process, signaling the likelihood of numerical difficulty in applications such as parameter (k and τ_L) identification.

Presuming that the desired behavior is a function of the measured variable y, the target closed-loop performance can be expressed as

$$Hy = Dr + Fe \tag{8}$$

where H is a (minimum-phase) polynomial with unity steady-state gain and stable zeros. H and F may be totally or partially specified polynomials. If e may have an arbitrary value in steady state, the steady-state value of F must be zero for y to converge to the set point (or reference input) r. Eliminating y from Eqs. (5) and (8) results in

$$ADr + AFe = HBDu + HCe \tag{9}$$

Because D is not cancelable, this equation cannot be solved directly for u. However, the product HC may be separated into two parts, the term on the left-hand side AF and a remainder expressed as the product DG, so that D becomes a common factor of Eq. (9):

$$HC = AF + DG \tag{10}$$

This equation is key to the controller design: selecting some and solving for other coefficients of H, F, and G when those of A, B, C, and D are known or estimated.

OPEN-LOOP CONTROL

The open-loop controller design results when Eq. (10) is used to eliminate AFe from Eq. (9), since D is not zero:

$$u = \frac{Ar - Ge}{BH} \tag{11}$$

Of course, if e is unmeasured, open-loop control will not reduce e's effect on y. This causes F to be an infinite-degree polynomial HC/A and G to be zero. If e is measured, G is a feedforward operator. Substituting u from Eq. (11) back into the process equation, Eq. (5), and not canceling terms common to both numerator and denominator, results in the open-loop performance equation

$$y = \frac{BA(Dr + Fe)}{HBA} \tag{12}$$

To avoid (imperfect) canceling of unstable roots, A as well as H and B must contain only stable zeros.

High-performance ($H \approx 1$) open-loop control applies the inverse of the process characteristic A/B to a set-point change. Because a dominant-lag process has low gain at high frequencies, its controller has high gain there. A rapid set-point change is likely to saturate the manipulated variable, but otherwise leaves its trajectory unchanged. The early return of this variable from its limit causes slower than optimal controlled variable response. This can be avoided by using nonlinear optimization (such as quadratic programming suggested in [4]) to compute the optimal controller-output trajectory, taking into account output limits, load level, and other process equality and inequality constraints.

The performance of an open-loop controller may be degraded by an unmeasured load or by mismatch between the process and the inverse controller at low frequencies. Mismatch at high frequency will not cause significant difficulty, however.

FEEDBACK CONTROL

Combining Eqs. (10) and (11) with the target equation, Eq. (8), to eliminate e results in the closed-loop control law

$$u = \frac{Cr - Gy}{BF} \tag{13}$$

This equation also results when e is eliminated from the process and target equations, Eqs. (5) and (8), and D is made a common factor with the design equation, Eq. (10). From Eq. (13) it is clear that the disturbance polynomial C and its Eq. (10) decomposition terms F and G play a key role in the feedback-controller design. Various methods for determining these polynomials will be discussed. Except in the special case where $C = G$, the control output u does not depend exclusively on the control error $r - y$. However, the controller will provide integral action, eliminating steady-state error, if the steady-state values of G and C are equal and not zero, and either those of $e\{t\}$, A, and B are zero or that of F is zero.

Substituting u from Eq. (13) back into the process equation and not canceling terms common to both numerator and denominator results in the closed-loop performance equation

$$y = BC\frac{Dr + Fe}{HBC} \tag{14}$$

To avoid (imperfect) canceling of unstable roots, C as well as H and B must contain only stable zeros. However, it is not necessary that all of the zeros of A be stable when the control loop is closed. Zeros of C not common to A correspond to unobservable modes of the disturbance variable e. Zeros of B or D not common to A correspond to uncontrollable modes of y.

When the manipulated variable saturates, it is necessary to stop (or modify) the controller integral action to prevent "windup." If the integration were allowed to continue, the prelimited controller output would continue to rise (wind up) above the limit value, requiring a comparable period after the control error reverses sign before the manipulated variable could recover from saturation. This would cause a significant (avoidable) controlled-variable overshoot of the set point. A controller of a dominant-lag process, designed for good unmeasured-load rejection, employs significant proportional (and derivative) feedback G. When responding to a large set-point change, this feedback keeps the output saturated and the controlled variable rate limited longer than would a linear open-loop controller, resulting in a faster response [5]. Halting the integral action, while the manipulated variable remains limited, prevents appreciable overshoot.

The performance of the feedback loop is most sensitive to the process behavior in the frequency range where the absolute loop gain is near 1. Performance at significantly lower frequencies is often quite insensitive to the process characteristics, load, and controller tuning.

Robustness

The ability of a feedback loop to maintain stability, when the process parameters differ from their nominal values, is indicated with robustness measures. Denoting the locus of combined shifts of the process gain by the factor b and the process delay by the factor d, which make the loop marginally stable, is a useful indicator of robustness, providing, in a more physical form, the information contained in gain and phase margins. The use of the two parameters b and d is based on the idea that the process behavior, in the frequency range critical for stability, can be approximated with an n-integral-delay two-parameter model. The integrals, whose number n may range from 0 to 1 plus the number of measurement derivatives used in the controller, contribute to the gain shift b. The phase, in excess of the fixed contribution of the integrals, and the shift in phase $d\omega$ can be considered to be contributed by "effective" delay.

At marginal stability the return difference (1 plus the open-loop gain) is zero:

$$1 + \frac{G}{FB}\frac{bBD^d}{A} = 1 + \frac{bGD^d}{AF} = 0 \tag{15}$$

Here it is assumed that D is effective delay, which may include small lags not included in A and a $(1 - \tau s)/(1 + \tau s)$ factor for each nonminimum-phase zero.

Figure 1 is a plot of d versus b, using logarithmic scales, for PID control of pure-delay, integral-delay, and double-integral-delay processes. The proportional band (PB) and the integral time (IT) were determined for minimum overshoot using the algebraic PID design method described in a later section. In [6] a single number, characterizing robustness, is derived from a robustness plot. This robustness index is -1 plus the antilog of the length of the half-diagonal of the diamond-shaped box centered at the nominal design point ($d = b = 1$) that touches the d versus b curve at its closest point. A value of 1 indicates that the product or ratio of d and b can be as large as 2 or as small as 0.5 without instability.

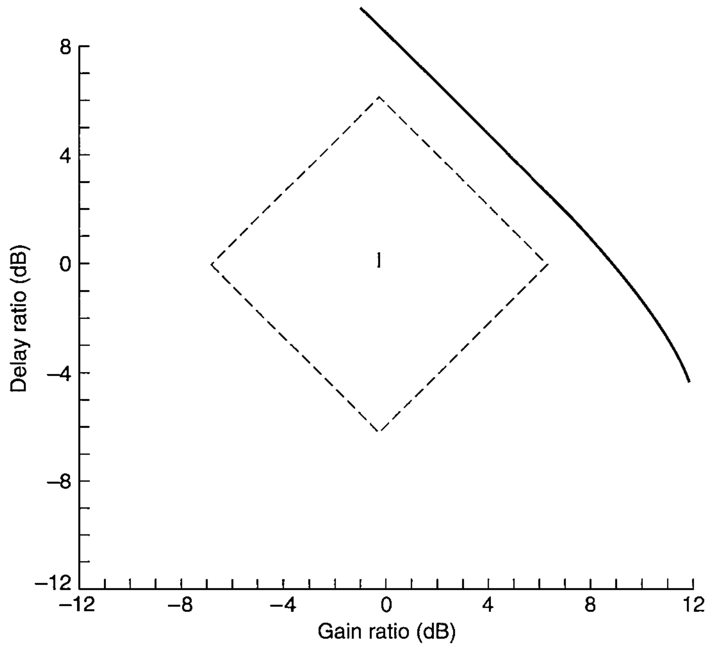


FIGURE 1(*a*) Robustness plots. The diamond corresponds to a factor of 2 changes in the product or quotient of delay and gain from their nominal values. PI control of pure delay process. (*Foxboro.*)

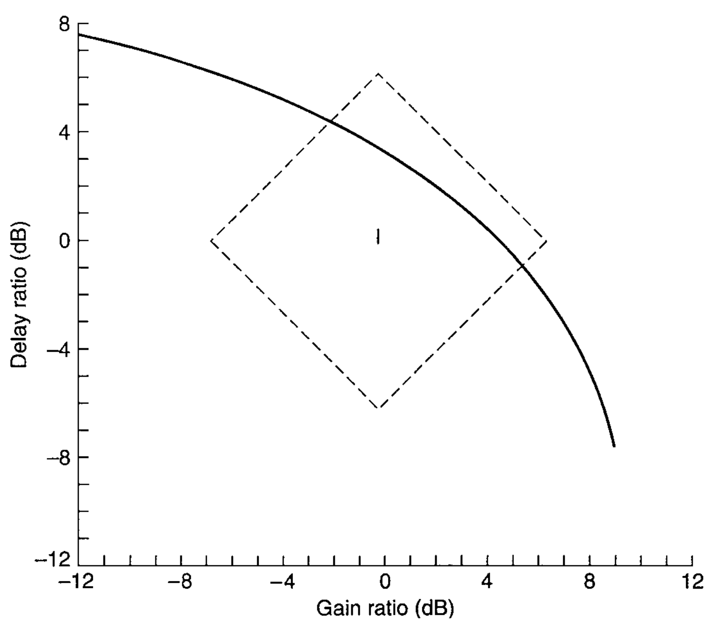


FIGURE 1(*b*) PID control of integral-delay process. (*Foxboro.*)

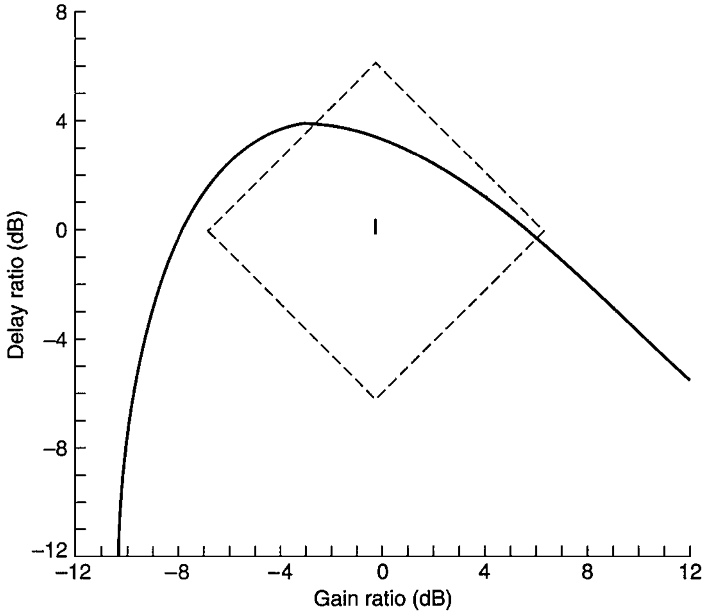


FIGURE 1(*c*) PID control of double-integrator-delay process. (*Foxboro.*)

For the three cases of Fig. 1 the robustness index is 1.67, 0.47, and 0.30, each determined by sensitivity to delay shift. The diamond-shaped box in the figure would correspond to a robustness index of 1.

Most control schemes capable of providing high performance provide poor robustness (a robustness index near 0). Adaptive tuning may be required to keep a controller that is capable of high performance current with changing process conditions.

Digital simulation provides a useful means for exploring robustness experimentally. Nonlinearities can be included naturally in the time domain. It may not be necessary to use an exotic integration algorithm if the process can be modeled in real-factored form. The factored form can be much less sensitive to roundoff error than unfactored polynomial and state-space forms.

The simulation equations should be solved in causal sequence. Each equation's dependent variable should be updated based on the most current computed value of its independent variables (as is done in the Gauss-Seidel iterative solution of algebraic equations). A useful such model for a first-order factor, $y/x = 1/(1 + \tau s)$, is

$$y\{t\} = y\{t-h\} + \frac{h}{(\tau + h)}(x\{t\} - y\{t-h\}) \tag{16}$$

and for a damped second-order factor, $y/x = 1/(1 + Ts + T\tau s^2)$, is

$$\begin{aligned} v\{t\} &= v\{t-h\} + \frac{h}{\tau + h}(x\{t\} - v\{t-h\} - y\{t-h\}) \\ y\{t\} &= y\{t-h\} + \frac{h}{T+h}v\{t\} \end{aligned} \tag{17}$$

The internal variable v is a measure of the derivative of the output y,

$$\frac{dy}{dt} \approx \frac{v\{t\}}{T+h} \tag{18}$$

These models both give the correct result when $T = \tau = 0$: $y\{t\} = x\{t\}$. When the sampling interval h is very small compared with T and τ, it may be necessary to compute with double precision to avoid truncation, because the second term on the right of Eqs. (17) and (18) may become much smaller than the first before the true steady state is reached.

A fixed time delay may be modeled as an integer number of computing intervals, typically 20 to 40. At each time step an old data value is discarded and a new value added to a storage array. Incremented pointers can be used to keep track of the position of the delay input and output, as in a ring structure. This avoids shifting all the stored data each time step, as in a line structure.

FEEDFORWARD CONTROL

Feedforward control, to counteract the anticipated effect of a measured load e_M, combined with feedback control to mitigate the effect of an unmeasured load e_U, makes use of two design equations like Eq. (10), one for each load type,

$$\begin{aligned} HC_U &= AF_U + DG_U \\ HC_M &= AF_M + DG_M \end{aligned} \tag{19}$$

C_M need not be cancelable and may include a delay factor. Combining the process equation, like Eq. (5), with the target equation, like Eq. (8), with the design equations (19), like Eq. (10), to eliminate e_U and D results in the combined feedback and feedforward control law, like Eqs. (11) and (13):

$$u = \frac{C_U r - G_U y}{BF_U} - \left(G_M - \frac{F_M G_U}{F_U} \right) \frac{e_M}{BH} \tag{20}$$

The second (e_M) term is an additive feedforward correction. If $F_M G_U / F_U = G_M$, feedforward control is not capable of improving upon feedback performance. The $F_M G_U / F_U$ term represents the reduction, from the open-loop feedforward correction G_M, needed to prevent redundant (overcorrecting) contributions. F_M can be made (nearly) zero, at least at low frequencies, when there is no more effective delay in the manipulated-variable path D to the controlled variable y than in the measured disturbance path C_M. Then from Eqs. (19), $G_M = HC_M / D$ because D is a factor of C_M. The measured disturbance e_M is (almost) perfectly rejected with the feedforward correction $u_{FF} = -(C_M / BD) e_M$, provided the controller output does not limit.

Feedforward provides a means for this single-output transfer function approach to be applied to a process with interacting loops. Unlike the state-space approach, it is necessary to associate each controlled variable with a particular manipulated variable. Then the effect of other manipulated variables on that controlled variable can be removed or reduced with feedforward corrections. This approach has the advantage that the appropriate compensation can be applied to active loops even when other loops are saturated or under manual control.

Furthermore, feedforward compensation may be structured to multiply the feedback correction. Multiplicative compensation is particularly effective for a temperature or composition loop manipulated with a flow. This configuration is intended to make the process appear more linear, as seen from the feedback controller. Thus the feedforward can be considered to provide gain scheduling for the feedback controller. Alternatively, the feedback controller can be viewed as adaptively tuning the gain of the feedforward compensator.

Other nonlinearities, even though they may involve the feedback variable, may be removable with additive or multiplicative feedforwardlike corrections [7]. For example, consider a nonlinear dominantly second-order process, such as a robot arm with negligible actuator delay and linkage flexibility. The process

$$g\{y\} \frac{d^2 y}{dt^2} + f\left\{ y, \frac{dy}{dt} \right\} = u \tag{21}$$

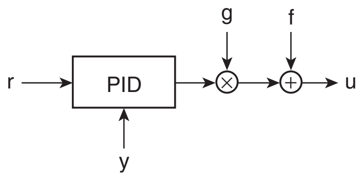


FIGURE 2 Feedback controller with multiplicative and additive compensation. (*Foxboro.*)

is controllable with u when the functions f and g are known,

$$u = f\left\{y, \frac{dy}{dt}\right\} + g\{y\}\left[K(r - y) - D_M \frac{dy}{dt}\right] \quad (22)$$

This control, shown in Fig. 2, achieves linear closed-loop response to set point r,

$$y = \frac{Kr}{K + D_M s + s^2} \quad (23)$$

Here the proportional (K) and derivative (D_M) feedback terms should be chosen to achieve desired closed-loop performance, taking the neglected effective delay and high-frequency resonances into account, perhaps adaptively. If the K and D_M terms can be made large compared with f/g, the closed-loop performance may be quite insensitive to imperfect compensation for f. An integral term added to the proportional plus derivative controller would help to adapt the effective gain of the multiplicative g term. Dominantly first- or zero-order processes can be linearized similarly.

MULTIPLE-LOOP CONTROL

A cascade of control loops, where the output of a primary (outer-loop) controller is the set point of the secondary (inner-loop) controller, may improve performance of the outer loop, particularly when the primary measurement responds relatively slowly. Nonlinearity, such as results from a sticking valve, and disturbances within the fast inner loop can usually be made to have little effect on the slow outer loop. Limits on the primary output constrain the set point of the secondary loop. Typical secondary controlled variables are valve position and flow. Jacket temperature may be the secondary variable for a batch reactor. The design of the primary controller should provide means of preventing integrator windup when the secondary controller limits or is in manual [8].

Controllers also may be structured in parallel to provide a safety override of a normal control function. For example, the normal controlled variable may be a composition indicative of product quality. In an emergency, a pressure controller may take over its manipulated variable. This may be done by selecting the controller with the smaller (or larger) output or error to drive the manipulated variable. Means for preventing integrator windup of the unselected controller should be provided.

When there are multiple interacting controlled and manipulated variables, every controlled variable should be paired with a controller output in order to make each control loop as insensitive as possible to the status of the others. Bristol's relative gain array (RGA) [9] can help in the evaluation of potential pairs. A good pair choice may result by considering some controller outputs to be the sum (or ratio) of measurable variables. Control is then implemented with a cascade structure. One of the summed (or ratioed) variables, acting as a feedforward, subtracts from (or multiplies) the primary output to get the secondary set point for the other variable. Again, means to prevent integrator windup, when the secondary saturates, should be provided.

For example, in distillation column control (Fig. 3), the distillate (DF) and reflux (LF) flows may be manipulated to control the distillate (impurity) composition and condenser level. Normally LF is much larger than DF. If LF and DF were the controller outputs, an RGA would show that the larger flow LF could be paired with level and DF with composition [8]. However, if the composition controller adjusts [DF/(LF + DF)] and the level controller adjusts (LF + DF), an RGA would indicate minimal interaction since the ratio has no effect on level when the sum is constant. In this case the distillate set point is found by multiplying the composition controller output by the measured LF + DF, and the reflux set point is found by subtracting the measured DF from the level controller output. Dynamic compensation of the feedforward terms will not improve performance, since LF affects the composition with no more effective delay than DF.

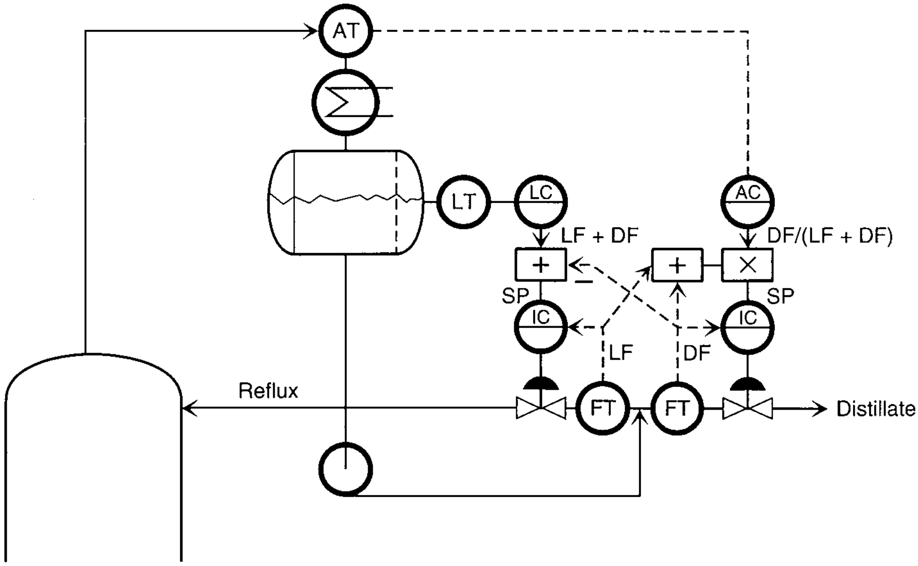


FIGURE 3 Level controller LC manipulates total flow and composition controller AC manipulates reflux ratio. (*Foxboro.*)

The RGA is an array where each element Γ_{ij} is a ratio of the sensitivities of a measurement y_i to an output u_j, the numerator having all other outputs fixed and the denominator having all other measurements fixed. For the process

$$\begin{aligned} A_1 y_1 &= b_{11} D_1 u_1 + b_{12} D_2 u_2 \\ A_2 y_2 &= b_{21} D_1 u_1 + b_{22} D_2 u_2 \end{aligned} \tag{24}$$

where A_i and D_j are dynamic operators and b_{ij} are constants, the RGA elements are also constants,

$$\text{RGA} = \begin{bmatrix} \Gamma & 1-\Gamma \\ 1-\Gamma & \Gamma \end{bmatrix} \tag{25}$$

with only one interaction parameter Γ,

$$\Gamma = \frac{b_{11} b_{22}}{b_{11} b_{22} - b_{12} b_{21}} \tag{26}$$

[The number of interaction parameters is $(n-1)^2$, where n is the number of interacting loops, because each RGA row and column sums to 1.]

When Γ is between 0.5 and 2, (u_1, y_1) and (u_2, y_2) could be pairs. When Γ is between -1 and 0.5, the opposite pairs could be used. Least interaction occurs when Γ is 1 or 0, which happens when one of the b_{ij} terms is zero. Values of Γ smaller than -1 or larger than 2 indicate that neither set of pairs should be used because the interaction is too severe. Saturating one of the loops, or placing it in manual, would change the gain in the other by more than a factor of 2.

When there is effective delay associated with each of the b_{ij} terms (here b_{ij} is not entirely cancelable), it is useful to compare the sum of the b_{11} and b_{22} delays with the sum of the b_{12} and b_{21} delays. If the interaction is significant (Γ not within 0.2 of either 0 or 1) and the combination with the smaller delay sum does not confirm the pairing based on Γ, a different choice of controller output variables (using a decoupling feedforward compensation) may be indicated.

DISTURBANCE REPRESENTATION

If the disturbance is a gaussian random variable, e is assumed to be a zero-mean white gaussian noise source. Colored noise is assumed to result from stably filtering the white noise e. The filter moving-average (numerator) characteristic is included in the C polynomial, and its autoregressive (denominator) characteristic is included in the A and B (or D) polynomials. The cross-correlation function of a linear filter's output with its input is equal to its impulse response convolved with the autocorrelation function of its input. When the input is white noise, its autocorrelation is an impulse function (the derivative of a step function). The cross correlation is then equal to the filter's impulse response [1]–[3].

Similarly, a deterministic disturbance may be considered to be the step or impulse response of a stable linear filter. A more complicated disturbance may be represented as the filter response to a sequence of steps (or integrated impulses). Any delay associated with an unmeasured disturbance is considered to determine the timing of the step in order that C_U be cancelable (stable zeros).

Pole-Placement Design

Pole-placement design [1] requires that A, B, C, and D be known and that a suitable H be selected. There may be difficulty in separating the BD product. As an expedient, B may be considered a constant and D allowed to include both stable and unstable zeros. If the degree of all of the polynomials is no greater than n, Eq. (10) provides $2n + 1$ equations, one for each power of s or z^{-1}, to solve for the coefficients of F and G. If the unmeasured disturbance e is considered an impulse, the degree of the G polynomial should be less than that of A, otherwise the degrees may be equal.

A design resulting in a B or F polynomial with a nearly unstable zero, other than one contributing integral action, should probably be rejected on the grounds that its robustness is likely to be poor. It may be necessary to augment H with additional stable factors, particularly if the degree of F or G is limited. The set point can be prefiltered to shift any undesired poles in the set-point function D/H to a higher frequency. However, the resulting uncompensatable unmeasured-load rejection function F/H may be far from optimal.

Linear-Quadratic Design

Linear-quadratic (LQ) design [1] provides a basis for calculating the H and F polynomials, but is otherwise like the pole-placement approach. In this case D contains both the stable and the unstable zeros and B is a constant. C/A is an impulse response function since e is specified to be an impulse. The H polynomial, which contains only stable zeros, is found from a spectral factorization of the steady-state Riccati equation,

$$\sigma H\{z\}H\{z^{-1}\} = \mu A\{z\}A\{z^{-1}\} + D\{z\}D\{z^{-1}\}$$

or

$$\sigma H\{s\}H\{-s\} = \mu A\{s\}A\{-s\} + D\{s\}D\{-s\} \tag{27}$$

The parameter σ is chosen to make the steady-state value of H unity and μ is an arbitrary parameter in the criterion function J,

$$J = \boldsymbol{E}\{(r - y)^2 + \mu u^2\} \tag{28}$$

$\boldsymbol{E}$ is the expectation operator. The u term imposes a soft constraint on the manipulated variable with a penalty factor μ. A polynomial X, satisfying

$$\begin{aligned} X\{z\}A\{z^{-1}\} + \sigma H\{z\}G\{z^{-1}\} &= D\{z\}C\{z^{-1}\} \\ X\{z\}H\{z^{-1}\} + \mu A\{z\}G\{z^{-1}\} &= D\{z\}F\{z^{-1}\} \end{aligned} \tag{29}$$

and Eq. (27) also satisfies Eq. (10) and minimizes Eq. (28) for an impulse disturbance e. The equations in s are similar. When the degree of A is n, the first equation of Eqs. (29) provides $2n$ equations, one for each power of z (or s). These can be solved for the $2n$ unknown coefficients of X and G, after H and σ are found from Eq. (27). G has no nth-degree coefficient and $X\{z\}$ and $D\{z\}$ have no zero-degree coefficient. [$X\{-s\}$ and $D\{-s\}$ have no nth-degree coefficient.] None of the polynomials is more than nth degree. F can then be found by polynomial division from Eq. (10) or the second of Eqs. (29).

For the optimization to be valid, the penalty factor μ must be large enough to prevent the manipulated variable u from exceeding its limits in responding to any input. However, if μ were chosen to be too large, the closed-loop response could be as sluggish as the (stabilized) open-loop response. The LQ problem may be solved leaving μ as a tuning parameter. Either experiments or simulations could be used to evaluate its effect on performance and robustness.

Despite the LQ controller being optimal with respect to J for a disturbance input, a switching (nonlinear) controller that takes into account the actual manipulated variable limits and the load can respond to a step change in set point r in less time and with less integrated absolute (or squared) error.

Minimum-Time Switching Control

The objective for the switching controller is to drive the controlled variable y of a dominant-lag process from an arbitrary initial value so that it settles at a distant target value in the shortest possible time. The optimal strategy is to maintain the manipulated variable u at its appropriate limit until y nears the target value r. If the process has a secondary lag, driving u to its opposite limit for a short time will optimally slow the approach of y to r, where it will settle after u is stepped to the intermediate value q needed to balance the load. Until the last output step, switching control is the same as "bang-bang" control. Determination of the output switching times is sufficient for open-loop control. The switching criteria must be related to y and its derivatives (or the state variables) in a feedback controller. Either requires solving a two-point boundary-value problem.

As an example, consider a linear integral $\{T\}$–lag$\{\tau_L\}$–delay$\{\tau_D\}$ process with constant manipulated variable (controller output) u. The general time-domain solution has the form

$$x\{t\} = at + b\exp\left\{-\frac{t}{\tau_L}\right\} + c$$
$$y\{t\} = x\{t - \tau_D\} \qquad (30)$$

where x is an unmeasured internal variable and a, b, and c are constants that may have different values in each of the regimes. At time zero the controlled variable y is assumed to be approaching the target value r from below at maximum rate,

$$\frac{dy\{0\}}{dt} = \frac{dx\{0\}}{dt} = \frac{u_M - q}{T} = a - \frac{b}{\tau_L}$$
$$x\{0\} = b + c \qquad (31)$$
$$y\{0\} = x\{0\} - \frac{(u_M - q)\tau_D}{T}$$

where u_M is the maximum output limit and q is the load. At that instant the output is switched to the minimum limit, assumed to be zero. If the next switching were suppressed, x and y would eventually achieve their negative rate limit,

$$\frac{dy\{\infty\}}{dt} = \frac{dx\{\infty\}}{dt} = -\frac{q}{T} = a$$

Combining with Eqs. (31) to eliminate a,

$$b = -\frac{u_M \tau_L}{T} \qquad c = x\{0\} - b \tag{32}$$

At time t_1, x reaches the target value r with zero derivative,

$$\frac{dx\{t_1\}}{dt} = 0 = -\frac{q}{T} + \frac{u_M}{T}\exp\left\{-\frac{t_1}{\tau_L}\right\} \qquad x\{t_1\} = r = x\{0\} - \frac{q(t_1 + \tau_L)}{T} + \frac{u_M \tau_L}{T} \tag{33}$$

Consequently,

$$t_1 = \tau_L \ln\left\{\frac{u_M}{q}\right\} \tag{34}$$

and switching from maximum to zero output occurred when

$$r - y\{0\} - \left(\tau_D + \tau_L - t_1 \frac{q}{u_M - q}\right)\frac{dy\{0\}}{dt} = 0 \tag{35}$$

If q is zero, as it may be when charging a batch reactor, the qt_1 product is zero, even though t_1 is infinite.

Optimal switching, from zero output to that required to match the load q, occurs at time t_1. If t_1 is less than the delay time τ_D,

$$r - y\{t_1\} - \left(\tau_D + \tau_L - t_1 \frac{u_M}{u_M - q}\right)\frac{dy\{t_1\}}{dt} = 0 \tag{36}$$

Otherwise

$$r - y\{t_1\} - \left(\tau_L - \frac{\tau_D}{\exp\left\{\frac{\tau_D}{\tau_L}\right\} - 1}\right) - \frac{dy\{t_1\}}{dt} = 0 \tag{37}$$

The controlled variable y will settle at the target value r at time $t_1 + \tau_D$. At this time the switching controller could be replaced by a linear feedback controller designed to reject unmeasured load disturbances and correct for modeling error. This combination of a switching controller with a linear controller, called dual mode, may be used to optimally start up and regulate a batch process [9].

Minimum-Variance Design

The minimum-variance design is a special case of the LQ approach where μ is zero. This may result in excessive controller output action with marginal improvement in performance and poor robustness, particularly when a small sampling interval h is used. Performance is assumed limited only by the nonminimum-phase (unstable) zeros and delay included in D. The H polynomial contains the stable zeros and the reflected (stable) versions of unstable zeros of D.

When D is a k time-step delay and e is considered an impulse, the minimum-variance solution for $F\{z^{-1}\}$ and $G\{z^{-1}\}$ can be found from Eq. (10) by polynomial division of HC by A. $F\{z^{-1}\}$ consists of the first $k-1$ quotient terms. The remainder $(HC - AF)$ is $DG\{z^{-1}\}$. From Eq. (27) H is unity. However, if H is arbitrarily assigned, this design becomes pole placement.

As a minimum-variance example, consider D to be a delay. Again, H is unity. The disturbance e is assumed to be a step applied downstream of the delay. C is unity. B is a gain b. The A polynomial represents a kind of lag,

$$A = 1 - aD \tag{38}$$

In one delay time the controlled variable can be returned to the set point r, hence $F = 1 - D$. Equation (10) becomes

$$1 = (1 - aD)(1 - D) + DG \tag{39}$$

Solving for G,

$$G = 1 + a - aD \tag{40}$$

and the controller from Eq. (13) becomes

$$u = \frac{(r - y)/(1 - D) - ay}{b} \tag{41}$$

When the delay is one computing step h, the process can be considered a sampled first-order lag with time constant $h/\ln\{1/a\}$ and a zero-order hold. Equation (41) has the form of a digital proportional plus integral controller with proportional band PB $= b/a$ and integral time IT $= ah$. When $a = 1$, the process can be considered a sampled integral with time constant h/b and a zero-order hold. When a is zero, the process is a pure delay and the controller is "floating" (pure integral, PB IT $= bh$). The controlled variable y has dead-beat response in one time step for either a set point or a load step,

$$y = Dr + (1 - D)e \tag{42}$$

The response is also optimum with respect to the minimum largest absolute error and the minimum integrated absolute error (IAE) criteria.

When the delay has k time steps, the $F = 1 - D$ factor in the controller equation has k roots equally spaced on the unit circle. One provides infinite gain at zero frequency (integral action). The others make the controller gain infinite at frequencies that are integer multiples of $1/kh$, not exceeding the Nyquist frequency $1/2h$. These regions of high gain cause the loop stability to be very sensitive to mismatch between the actual process delay and kh used in the controller. As k approaches infinity, the robustness index approaches zero, indicating no tolerance of delay mismatch. A low-pass filter that improves the robustness by attenuating in these regions of high gain also degrades the nominal performance.

The control loop may be structured as two loops, the outer-loop integral controller providing the set point r_I to the inner-loop proportional controller,

$$r_I = \frac{r - y}{1 - D} = r - y + Dr_I \tag{43}$$

and

$$u = \frac{r_I - ay}{b} \tag{44}$$

The effect of closing only the inner loop is to create a delay process as seen by the outer-loop controller,

$$y = Dr_I + e \tag{45}$$

The outer-loop controller can be considered model feedback in relation to the closed inner loop, since the difference between the controlled variable's measured (y) and predicted (Dr_I) values is fed back as a correction to an open-loop (unity-gain) controller.

Model-Feedback Control

Model-feedback control, whose variations include, among others, Smith predictor, Dahlin, dynamic matrix, and model predictive (unless a special unmeasured disturbance model is used [10]), consists of an open-loop controller with a feedback correction equal to the model prediction error. If model-feedback control is applied without first closing a proportional inner loop, there results

$$u = \frac{A}{BH}\left[r - \left(y - \left(\frac{BD}{A}\right)u\right)\right]$$
$$y = \frac{D}{H}r + (1 - D)\left(\frac{C}{AH}\right)e \tag{46}$$

assuming no process-model mismatch and no unstable zeros of A. H is a filter chosen to improve robustness. For the above process, assuming H to be unity, the response to a step disturbance can be easily calculated by polynomial division,

$$F = (1 - D)\frac{C}{A} = \frac{1 - D}{1 - aD}$$
$$= 1 - (1 - a)D[1 + (aD) + (aD)^2 + (aD)^3 + \cdots] \tag{47}$$

The maximum error occurs during the first delay interval when $F\{0+\} = 1$. After n delay steps $F\{n+\}$ is reduced to a^n. Even though this result is optimum with respect to the minimum largest absolute error criterion, the recovery from a load upset can be very slow when a is close to 1 (or divergent, when a is greater than 1). The ratio of the IAE to the optimum is $1/(1-a)$. Consequently model-feedback control may not adequately reject an unmeasured load disturbance, when the process has a dominant lag (a near 1), unless well-chosen inner-loop feedback is applied or the model deliberately mismatches the process, as recommended in [6]. However, design of the inner-loop controller or the model mismatch, for near optimal load rejection, may require more detailed high-frequency knowledge (for example, a spectral factorization of process polynomials) than is necessary for selecting an output trajectory (open-loop controller) to achieve good set-point tracking.

The stability of a tightly tuned matched-model feedback loop is very sensitive to mismatch between the model and process delays. To achieve adequate robustness, it may be necessary to detune the controller with H, further sacrificing unmeasured-load rejection capability.

Without the inner loop or deliberate model mismatch, early return of the output from saturation may cause excessively slow controlled-variable response of a dominant-lag process to a large set-point step. As with open-loop control, an on-line nonlinear optimizer may be used to avoid this suboptimal behavior.

Algebraic Proportional plus Integral plus Derivative Design

In this section a two-phase method for applying Eq. (10) to the design of analog (or fast-sampling digital) proportional plus integral plus derivative (PID) controllers is described. Unlike Bode and root-locus design methods, this method allows all of the controller parameters as well as the closed-loop performance parameters (time scale and load sensitivity) to be found directly, without trial and error.

The process is represented with an A polynomial in s. B, D, and C are assumed 1. The inverse of a delay or small numerator zero factor (if representable as a convergent Taylor series up to the frequency range critical for stability) is included in the A polynomial,

$$(a_0 + a_1 s + a_2 s^2 + a_3 s^3 + \cdots)y = u + e \tag{48}$$

A large stable zero is unusual and requires special consideration. It should be approximately canceled by a controller or process pole. Such a pole is not available from a PID controller. However, an effective process cancellation, without a factorization, may result by disregarding the zero-order terms of both the process numerator and the denominator before determining A by polynomial division. When the process zero and pole are sufficiently dominant, mismatch in zero-order terms, which affects the very low-frequency behavior, will be corrected by high controller gain in that frequency range.

This two-phase design process implicitly imposes the performance limitation that would normally be imposed by including delay and nonminimum-phase zeros in D. The first design phase prevents inner-loop feedback when the open-loop process already approximates a pure delay. This is done by selecting the inner-loop gain and derivative terms,

$$G_I = K_M + D_M s \tag{49}$$

to make the closed inner loop H_I^{-1} approximate a delay at low and moderate frequencies. As many low-order terms of Eq. (10) are matched as are needed to determine the unknown controller and performance parameters.

The H_I polynomial is chosen to be the Taylor-series expansion of an inverse delay whose time τ_I and gain h_0 is to be determined,

$$H_I = h_0\left[1 + \tau_I s + \frac{(\tau_I s)^2}{2} + \frac{(\tau_I s)^3}{6} + \cdots\right] \tag{50}$$

When F_I is chosen as 1 (instead of choosing h_0 as 1), Eq. (10) gives

$$H_I = A + G_I \tag{51}$$

The limited-complexity inner-loop proportional plus derivative controller can significantly influence only the low-order closed-loop terms and hence can shape only the low-frequency closed-loop behavior. Only the most dominant two poles (lowest in frequency) of the open-loop process may be unstable, since only they can be stabilized with proportional and derivative feedback. As a result the limiting closed inner-loop performance measures, the values of τ_I and h_0, are determined by a_2 and a_3, provided the latter have the same sign. Equating term by term and rearranging gives

$$\begin{aligned}
\tau_I &= \frac{3a_3}{a_2} \\
h_0 &= \frac{2a_2}{\tau_I^2} \\
D_M &= h_0\tau_I - a_1 \\
K_M &= h_0 - a_0
\end{aligned} \tag{52}$$

When the sign of D_M is different from that of K_M, or derivative action is not desired, the parameters should be calculated with

$$\begin{aligned}
\tau_I &= \frac{2a_2}{a_1} \\
h_0 &= \frac{a_1}{\tau_I} \\
D_M &= 0 \\
K_M &= h_0 - a_0
\end{aligned} \tag{53}$$

For a pure delay process both K_M and D_M are zero. The closed inner loop becomes

$$y = \frac{r_I + e}{H_I} \tag{54}$$

The outer loop uses gain and integral terms applied to the error,

$$r_I = \left(\frac{1}{I_E s} + K_E\right)(r - y) \tag{55}$$

Using this equation to eliminate r_I from the previous gives

$$[1 + I_E s(K_E + H_I)]y = (1 + K_E I_E s)r + I_E s e \tag{56}$$

The target closed-loop set-point behavior is chosen to approximate a nonovershooting delaylike model, n equal lags. The shape parameter n and the time constant τ_0 are to be determined, as are the controller parameters K_E and I_E,

$$\left[\left(1 + \frac{\tau_0 s}{n}\right)^n y\right] = r + \frac{I_E s e}{1 + K_E I_E s} \tag{57}$$

Equating term by term and solving the four simultaneous equations gives

$$\begin{aligned} n &= 10.4 \\ \tau_0 &= 1.54\tau_I \\ K_E &= 0.198 h_0 \\ I_E &= \tau_0 / h_0 \end{aligned} \tag{58}$$

A small value of I_E is desirable because its product with the output change is equal to the integrated error response to a load change. Since the controller is designed to achieve very small overshoot, the product is also nearly equal to the IAE for a step-load change. If the response shape parameter n were made infinite (corresponding to a pure delay target) instead of matching the fourth-degree terms, a faster but more oscillatory and less robust response would result (with near minimum IAE for a load step). Then the closed-loop time constant τ_0 would become $1.27\tau_I$ and K_E would equal $0.289h_0$. Equations (58) still would apply for I_E.

The resulting controller is a four-term noninteracting (sum of terms) type. The following equations may be used to convert these tuning values to those for a conventional three-term PID interacting (product of factors) type, adjusted for good load rejection,

$$K = K_E + K_M \tag{59}$$

If K^2 is greater than four times the ratio of D_M to I_E,

$$\begin{aligned} \frac{1}{\text{PB}} &= 0.5\left[K + \left(K^2 - \frac{4D_M}{I_E}\right)^{0.5}\right] \\ \text{IT} &= I_E/\text{PB} \\ \text{DT} &= D_M \text{PB} \end{aligned} \tag{60}$$

Otherwise,

$$\begin{aligned} \text{PB} &= \frac{2}{K} \\ \text{IT} = \text{DT} &= D_M\,\text{PB} \end{aligned} \tag{61}$$

To achieve the designed set-point response also, the set-point input should be applied to the controller through a lead-lag filter. The lag time is matched to the integral time IT. The ratio of lead-to-lag time α is made equal to K_E PB. The resulting controller equation is

$$u = \frac{(1 + \alpha \mathrm{IT}s)r - (1 + \mathrm{IT}s)(1 + \mathrm{DT}s)y}{\mathrm{PB\ IT}s} \tag{62}$$

Derivative action is applied only to the controlled measurement, not to the set point. To prevent excessive valve activity at frequencies beyond the closed-loop bandwidth, it is customary to condition the controlled measurement y with a low-pass filter whose time constant is a small fraction (≈ 0.1) of DT. In order that the sampling process of a digital controller not further diminish the effectiveness of the derivative term, the sampling interval should be less than the effective filter time.

For example, consider a thermal process with a 300-second lag and a 10-second effective delay. The algebraic design gives an integral time IT = 24 seconds, derivative time DT = 3.6 seconds, and the closed-loop time constant $\tau_0 = 23.1$ seconds, all more sensitive to the delay than the lag. The sampling interval should not exceed 0.4 second in order not to compromise closed-loop performance. This is a surprisingly small interval, considering the relatively slow open-loop response.

Antialias Filtering

When an analog signal contains a component with frequency higher than the Nyquist frequency (half the sampling frequency f_S), the sampled signal component appears to have a frequency less than the Nyquist frequency, as shown in Fig. 4. If the analog signal component frequency f lies between odd integers $(2n - 1)$ and $(2n + 1)$ times the Nyquist frequency,

$$(n - 0.5)f_S \le f < (n + 0.5)f_S \tag{63}$$

the sampled signal component has the same amplitude but appears to be shifted to the frequency $|f - nf_S|$. This frequency shifting is called aliasing [1].

To achieve good performance with digital control, it is necessary to sample the controlled variable at a rate faster than twice the highest frequency significant for control and to attenuate, before sampling, components with higher than Nyquist frequency. This should be done with minimum attenuation or phase shifting of the lower-frequency components that are important for control.

It is particularly important to remove, before sampling, a signal component that is an integer multiple of the sampling frequency, because this component would shift to zero frequency, causing a steady-state offset error. An analog filter that averages between sampling instants removes such components completely. This filter can be realized with an analog-to-frequency converter and a

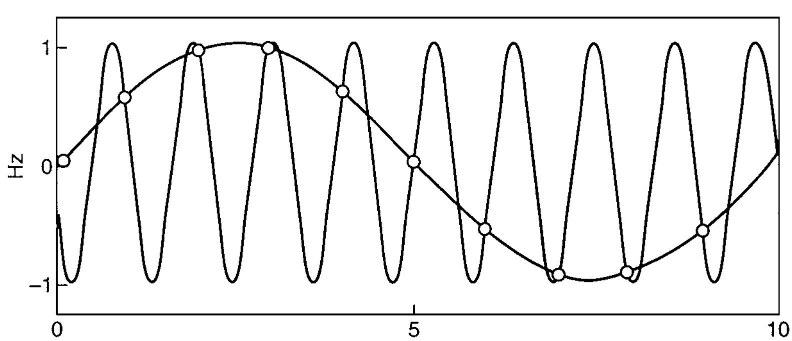


FIGURE 4 Two signals with different frequencies (0.1 and 0.9 Hz) have same values at 1-Hz sampling instants. The 0.9-Hz signal is aliased as 0.1 Hz after sampling. (*Foxboro.*)

sampled digital counter. The longer the sampling interval, the higher is the count and the greater is the digital signal resolution.

However, the analog averaging filter does not attenuate sufficiently near the Nyquist frequency. A stage of digital filtering, a two-sample average, completely removes a Nyquist-frequency component in the sampled signal. Its passband is flatter and its cutoff is sharper than those of a digital Butterworth (autoregressive) filter having the same low-frequency phase (effective delay).

This measurement antialias filter together with the manipulated variable zero-order hold adds an effective delay of $1.5h$ to the analog process. Calculation delay, which may be as large as the sampling interval h, is additional. Consequently the sampling interval should be small compared with the effective delay of the process, in order that feedback loop performance (unmeasured-load rejection) not be compromised. On the other hand, resolution is improved by averaging over a longer sampling interval.

An antialiasing filter removes or diminishes the effects of higher than Nyquist frequency process signals on the digital measurements. However, digital outputs apply a sequence of steps to the process that may excite a higher than Nyquist-frequency open-loop resonance (such as water hammer). This effect will be accentuated in high-performance loops with large output steps and in loops where the resonance is synchronized with a multiple (harmonic) of the output update frequency.

ADAPTIVE CONTROL

Time-varying or nonlinear process dynamics, variable operating conditions, slow response to upsets, dominant unmeasured disturbances, performance degradation resulting from deliberate upsets, and lack of tuning expertise are all reasons to consider adaptive (self-tuning) control. Economic incentives may result from improved control of product quality and yield, a higher production rate, less energy usage, improved plant safety, and less pollution.

Adaptive control schemes may be classified according to several design alternatives. Associated with each of these are concerns that should be addressed in order to assure robust adaptation.

First, the adaptor may be open- or closed-loop. An open-loop adaptor programs the controller tuning based on a model of the process. The model may be fixed and nonlinear (such as a neural net) or time-varying and linear, with parameters updated to reconcile measured process inputs and outputs. Mismatch between the model and the process may be caused by misidentification as a result of large nonstationary unmeasured disturbances, insufficiently rich inputs, or by model structural deficiencies. Mismatch may lead to inappropriate controller tuning that will not be corrected until (or unless) the model is revised. For example, even though the effective process delay limits how tightly the controller can be tuned, the model delay may be assigned arbitrarily (it is not easily identified) or the process delay masked by a larger sampling interval. On the other hand, a closed-loop adaptor monitors one or more performance measures for the closed control loop and adjusts controller parameters to drive these measures to target values. Desired performance is assured when the feedback adaptor converges. Its issues include the rate of convergence and the robustness of the adaptive loop.

Second, the adaptation may be based on observations of responses to deliberate or natural disturbances. For example, a performance-optimizing adaptor, using a hill-climbing strategy, requires the comparison of process performance measures for responses to identical, hence deliberate, disturbances. However, a deliberate disturbance degrades a well-tuned loop's performance. Therefore deliberate upsets should be applied infrequently and only when there is high likelihood that the controller is mistuned and is not likely to recover soon. A deliberate disturbance has the advantage that it is known and can be made large enough to dominate unmeasured disturbances and rich enough to excite important process modes. On the other hand, natural disturbances may be unmeasured and nonstationary, often consisting of significant isolated filtered steplike events and low-level noise. However, a response to a natural disturbance may contain incomplete information to observe all process modes and make an unambiguous adaptation.

Third, the target control-loop time scale may be fixed or optimal. One may choose to drive the closed-loop performance to that of a fixed model, as is done with pole-placement, fixed model-delay minimum-variance, or model-reference adaptors. A fixed time scale for all operating conditions must be user-selected to be as large as the largest minimum, and hence suboptimal for other operating conditions. Extensive knowledge of the process dynamics helps in choosing the time scale. Alternatively the control-loop time scale may be adapted in an open-loop minimum-variance adaptor by updating the effective delay time, or in a performance-feedback adaptor by choosing response shape parameters that are sensitive to the relative positions of the three (or more) most dominant closed-loop poles. On the other hand, a tightly tuned controller may not have an adequate stability margin, particularly if the (nonlinear) process may experience sudden large unmeasured load changes.

Fourth, the adaptations may be performed continuously, with updates each sampling interval, or aperiodically, following the response to each significant disturbance. When the controller tuning is updated each sampling instant, the update must recursively take into account a portion of past history. Over this time interval both the process parameters and the statistical measures of unmeasured disturbances are assumed to be stationary. However, real processes may be subjected to minimal load changes for extended intervals and to large unmeasured nonstationary load changes at other times. This makes choosing an adequate time interval difficult, particularly when effective adaptation requires that a time-varying linear model track a nonlinear process. On the other hand, an adaptive scheme designed to update controller parameters following significant isolated disturbance responses must also cope with cyclical and overlapping responses. It should distinguish among a loop instability, a stable limit cycle, and a cyclical disturbance.

Cycling can also be a problem for an adaptor based on model identification, because the signals may provide incomplete information for identifying more than two process parameters. An incorrect identification may lead to worse controller tuning. Furthermore, if the cycle amplitude is large, it may be impractical to make the identification unique by superimposing sufficiently significant deliberate disturbances.

An event-triggered adaptor provides an additional opportunity: the state and measured inputs, existing at the moment a new disturbance is sensed, can be used to select among several sets of stored tunings to determine the set to be used during, and updated following, the response interval. This capability (such as gain scheduling and multiplicative feedforward compensation) enables the controller to anticipate and compensate for the effect of process nonlinearity.

Three types of adaptive controllers will be discussed. First is the performance-feedback type, using expert system rules to cope with incomplete information, and using nonlinear dead-beat adaptation when information is complete. The second is an open-loop type that uses a recursive parameter identifier to update the parameters of a difference equation model. The controller is tuned as if the model were the process, thus invoking the "certainty equivalence" principle. Third is another open-loop type. This one identifies the low-order parameters of a differential equation model in order to update the coefficients in a feedforward compensator. It uses the moment-projection method on each isolated response.

Other types of adaptors, including the model reference type, may be more suitable for set-point tracking than for unmeasured load rejection. For example, the extended-horizon type uses a nonlinear optimizer to determine an open-loop controller's constrained output trajectory. The controlled-variable trajectory must be optimized over a prediction interval that exceeds that of the output by the process delay time. An extended-horizon optimizer can be used in conjunction with predictive-model feedback, useful (as explained earlier) for load rejection with a dominant-delay process and for model mismatch correction with set-point tracking. The nonlinear optimization calculations may require a large time step that compromises control loop performance with additional effective delay. Also, a linear moving-average process model may have so many degrees of freedom that on-line identification is impractical, both because the computational load is excessive and because a reasonable interval of natural-signal history does not persistently excite all of the model and process modes. A fixed nonlinear neural-net model may be more practical for a time-invariant process, even though its programming requires training with extensive signal records spanning the range of expected operating conditions.

PATTERN RECOGNITION AND EXPERT SYSTEMS, PERFORMANCE-FEEDBACK ADAPTOR

A performance-feedback adaptor monitors a single-loop variable control error [11]. Since the set point and process measurement are known, the controller output does not contain independent information, because it is determined by the controller equation. Pattern features are measured of the error response to a significant (typically unmeasured) disturbance. When enough of the response has been observed, the controller parameters are adjusted in order to make the feature values of the next response approach target values.

A significant error event is detected when the absolute control error exceeds a set threshold. The threshold value is chosen large enough so that an error event will not be triggered by low-level process or measurement noise. Since error peaks are the most prominent features of an oscillatory response, peak amplitudes and times are sought. Expert system (heuristic) rules may be used to distinguish response peaks from noise peaks.

Zero-to-peak $(-E_2/E_1)$ and peak-to-peak $[(E_3 - E_2)/(E_1 - E_2)]$ error amplitude ratios may be chosen as shape features. These are independent of the response amplitude and time scales and are called overshoot and decay, respectively. Target values for these ratios may be chosen to minimize a criterion function, such as minimum IAE, for a specific process and disturbance shape. The ratio of times between peaks is not recommended as a controlled-shape feature, because it is relatively sensitive to nonlinearity and noise and insensitive to the relative location of closed-loop poles.

The second and third peaks, E_2 and E_3, do not exist for an overdamped response. In this case a "knee" (or quasi-peak, defined by "expert" criteria) is sought to determine a response time scale used to terminate the peak search and to estimate the (negative) effective overshoot. Decay is zero in this case.

If the error response is a damped quadratic function, the first three error peaks E_i are related by

$$E_1 E_3 = E_2^2 \qquad \text{decay} = \text{overshoot} \tag{64}$$

This response, like the overdamped response, does not contain complete information for controller tuning [12]. A response containing both lag and underdamped quadratic terms provides complete information for tuning a PID controller. For the response

$$E\{t\} = \alpha e^{-at} + \beta e^{-bt} \cos\{\omega t\} \tag{65}$$

three shape ratios β/α, b/a, and b/ω provide sufficient information to update three controller parameters PB, IT, and DT. However, if either α or β were 0, values of a, or b and ω, would be unmeasurable and the information incomplete. It is desirable that the features used for adaptation reflect the relative pole positions indicated by the last two ratios, because the dominant error poles are usually closed-loop poles. Furthermore, the features should be insensitive to the first ratio, β/α, because this ratio is sensitive to the unmeasured disturbance shape and point of application, as indicated by the relative location of the error signal zeros.

When information is complete from the error response to a load step applied upstream of a dominant lag or delay,

$$E_1 E_3 > (E_2)^2 \qquad \text{decay} > \text{overshoot} \tag{66}$$

This corresponds to the lag (α) and quadratic (β) terms making contributions of the same sign to the first peak.

These inequalities can be reversed if the disturbance has a different shape or is applied at a different location. Reversal would result if the disturbance were a narrow pulse or if it were a step applied downstream of a dominant lag. Reversal also would result if the disturbance were statically compensated by a feedforward controller and could result if the disturbance were to affect interacting loops. The final part of this type of error response has a shape similar to the "usual" response so that, when the first peak, or peaks, is discarded and the remaining peaks are renumbered, decay becomes greater than overshoot.

Peak shifting desensitizes the pattern features to open-loop and disturbance-signal zeros, while maintaining sensitivity to the relative positions of the three most dominant error-signal poles. These poles usually are the closed-loop poles, but may include the poles of the disturbance signal if the disturbance is not applied suddenly.

The changes in the controller-parameter vector P are computed from the deviation of the measured-feature vector F from its target-value vector F_t, according to the adaptor's nonlinear gain function matrix **G**,

$$P\{i+1\} = P\{i\} + \mathbf{G}(F_t - F\{i\}) \tag{67}$$

The response feature vector $F\{i\}$, measured after the ith response, is a nonlinear function of the controller parameter vector $P\{i\}$ (existing during the ith response), the process type, and the disturbance shape. For a given process and disturbance shape, simulation can be used to map feature deviations as a function of control parameter deviations $\delta\mathbf{F}/\delta\mathbf{P}$, allowing feature deviations to be predicted with

$$F\{i+1\} = F\{i\} + \frac{\delta\mathbf{F}}{\delta\mathbf{P}}(P\{i+1\} - P\{i\}) \tag{68}$$

If the unique inverse of the function matrix $\delta\mathbf{F}/\delta\mathbf{P}$ exists, the latest response contains complete information. Then a dead-beat adaptation ($F\{i+1\} = F_t$) is possible with

$$\mathbf{G} = \left(\frac{\delta\mathbf{F}}{\delta\mathbf{P}}\right)^{-1} \tag{69}$$

This multivariable adaptive loop is quite robust, being particularly tolerant of smaller than optimum **G**. For example, if the optimum **G** is multiplied by a factor ranging between 0 and 2, the adaptive loop will remain stable. The eigenvalues of the $I - (\delta\mathbf{F}/\delta\mathbf{P})\mathbf{G}$ matrix, nominally located at the origin of the complex z plane, must stay within the unit circle.

If the process were a lag delay, a process-type variable, sensitive to the ratio of the lag-to-delay times, could be used to interpolate between the two extremes. The ratio of the controller IT to the response half-period is an indicator of the process type when decay and overshoot are fixed. This property may be used to identify the process type. The optimal IT-to-half-period ratio IT/T is smaller for the pure delay than for the integral delay.

When only two features, such as overshoot and decay, reliably characterize a response shape, only two controller parameters (PB and IT of a PID controller) are determined through performance feedback. However, because the optimal derivative-to-integral ratio is also a function of the process type, DT can be calculated after IT and the process type have been determined.

The process type, that is, the proportional-band ratio PB/PB_t and the integral-time ratio IT/IT_t, is determined by interpolating stored data from performance maps for the process-type extremes, given overshoot, decay, IT/T, and DT/IT. The half-period T and the controller parameters PB, IT, and DT are values for the latest response. PB_t and IT_t are the newly interpolated values of the controller parameters predicted to produce the target features on the next response.

When the error-shape information is incomplete, as when the response is overdamped, quadratic, nonisolated, or nonlinear (because the measurement or controller output has exceeded its range), expert system rules are used to improve the controller tuning. These rules, of the if-then-else type, invoke a special strategy for each of these contingencies. Several retunings may be needed before a response shape contains sufficient information to achieve the desired performance on the next response. Even when the information is incomplete, robust tuning rules are possible, provided derivative action is not essential for stabilizing the control loop.

A nonisolated response is recognized if its start is detected while waiting for the last response to settle. A nonisolated response may be caused by the failure of the preceding response to damp quickly enough. If the decay of a nonisolated response is sufficiently small, even though it may be bigger than the target, the existing tuning is retained. Typically a continuing oscillation will be dominated by a quadratic factor, giving rise to incomplete information for retuning.

A nonisolated response may also be caused by a rapid sequence of load changes, the next occurring before the response to the last has settled. Peak shifting tends to desensitize the adaptor to a strange sequence of peaks, allowing detuning only when a conservative measure of decay is excessive.

A marginally stable loop is distinguished from a limit cycle or response to a cyclical load by observing the improvement in decay caused by an adaptive retuning. If retuning fails to reduce the decay measure, the last successful tunings may be restored and adaptation suspended until settling or an operator intervention occurs.

DISCRETE-MODEL IDENTIFICATION, OPEN-LOOP ADAPTATION

Adaptation of a feedback controller based on an identification of an input-output process model is most effective when the important process inputs are measured and (at least partially) uncorrelated with one another. An unmeasured disturbance is assumed to come from a stationary filtered white gaussian-noise source uncorrelated with the measured inputs. When a large unmeasured disturbance violates this assumption, the identified model may be a poor match for the process and poor controller tuning may result. Process-model mismatch may also result when the disturbance fails to independently excite process or model modes, a condition called nonpersistent excitation. A poor model structure, such as one having an incorrect unadapted delay or insufficient model degrees of freedom, may also cause mismatch that leads to poor control. Coefficients for a model having both linear and nonlinear terms may not be uniquely identifiable if the process input and output changes are small.

Two types of models may be identified, called explicit and implicit. An explicit model relates the process inputs and output with parameters natural to the process, such as Eq. (5). The explicit model is most useful for the design of an open-loop or model-feedback controller. A complicated design process involving Eq. (10) would be needed to compute the feedback controller parameters of Eq. (13). An implicit model combines the target equation, Eq. (8), and the feedback control equation, Eq. (13), so that the parameters needed for control are identified directly. In either case the identification model may be put in the prediction form

$$\Omega\{t+k\} = \Phi\{t\}^T \Theta + \varepsilon\{t+k\} \tag{70}$$

which predicts the value of Ω, k time steps ahead, given present and past values of the process inputs and outputs concatenated in the vector Φ. Θ is a corresponding vector of parameters determined by the identifier; ε is the identification error.

For the explicit model,

$$Ay = DBu + Ce + e_0 \tag{71}$$

Here e_0 is the steady-state offset and

$$\begin{aligned} A &= 1 + a_1 z^{-1} + \cdots \\ C &= 1 + c_1 z^{-1} + \cdots \\ B &= b_0 + b_1 z^{-1} + \cdots \end{aligned} \tag{72}$$

The time step h is assumed to be one time unit. The time step should be chosen small enough that the antialias filter, the digital computation, and the output hold do not dominate the effective delay, but large enough that roundoff or data storage do not cause difficulty. Here D is assumed to be a known and fixed k-time-step delay. The value of k must be large enough that B is stably cancelable. Of course, other choices for D are possible. The prediction model variables become

$$\begin{aligned} \Omega\{t\} &= y\{t\} \\ \Phi\{t-1\}^T &= [u\{t-k\}, \ldots, -y\{t-1\}, \ldots, \varepsilon\{t-1\}, \ldots, 1] \\ \Theta^T &= [b_0, \ldots, a_1, \ldots, c_1, \ldots, e_0] \end{aligned} \tag{73}$$

If the model matched the process exactly, the prediction error $\varepsilon\{t\}$, which is uncorrelated with the variables in $\Phi\{t-1\}$, would equal the white-noise disturbance $e\{t\}$. In order to identify C, past values of the prediction error are needed, but these cannot be found until Θ is identified. This difficulty can be overcome by solving for Θ recursively. When Θ is updated each time step, $\varepsilon\{t\}$ can be calculated using the most recent Θ. C must be constrained to have stable zeros. An algorithm that identifies C is said to be "extended." When Φ and Ω are prefiltered by C^{-1}, the algorithm is "maximum likelihood." If the identifier inputs are prefiltered by E^{-1}, $(E/C) - 0.5$ must be positive real in order to ensure that Θ can converge to its true value [13]. Convergence also requires no structural mismatch.

The model form for implicit identification uses the target equation, Eq. (8), to eliminate $r\{t-k\}$ from the controller equation, Eq. (13),

$$Hy\{t\} = BFu\{t-k\} + Gy\{t-k\} - C'r\{t-k-1\} + e_0 + Fe\{t\} \tag{74}$$

where $C' = C - 1$ and D is a k-step delay ($h = 1$). The prediction model variables become

$$\Omega\{t\} = Hy\{t\} = h_0 y\{t\} + h_1 y\{t-1\} + \cdots \tag{75}$$

where H is specified.

$$\begin{aligned} BF &= \beta_0 + \beta_1 z^{-1} + \cdots \\ G &= \alpha_0 + \alpha_1 z^{-1} + \cdots \\ \Phi\{t-k\}^T &= [u\{t-k\}, \ldots, y\{t-k\}, \ldots, -r\{t-k-1\}, \ldots, 1] \\ \Theta^T &= [\beta_0, \ldots, \alpha_0, \ldots, c_1, \ldots, e_0] \end{aligned} \tag{76}$$

To identify the C polynomial, the set point r must be active, the updated control law implemented, and a recursive algorithm used. C must be constrained to have stable zeros. Also, k must be large enough that the zeros of BF are stable. If the model matched the controlled process exactly, the modeling error $\varepsilon\{t\}$, which is uncorrelated with any of the variables in $\Phi\{t-k\}$, would equal the closed-loop noise response $Fe\{t\}$. If $H = 1$, this implies a minimum-variance design, otherwise pole placement.

The same positive real and structural consistency requirements apply for convergence of an implicit parameter set as apply for the explicit set. The positive real condition ensures that the component of control error Fe, in phase with the model error ε, is positive and at least half as big.

Control based on either model will not have integral action when the identifier is turned off. Integral action depends on updating the offset e_0. The effective integral time constant depends on the quantity of past history taken into account in calculating Θ. Consequently it is likely to be significantly larger than optimal. On the other hand, integral action can be implemented explicitly in an outer-loop controller, such as

$$\delta r_I = \frac{r - y}{2k - 1} \tag{77}$$

without identifying e_0, if an incremental identification model is used to design the inner-loop controller,

$$\delta u = \frac{(C\delta r_I - G\delta y)}{BF} \tag{78}$$

For an incremental model, the values of variables in Ω and Φ are the changes δu and δy from one time step to the next. If C were 1 in such a model, the autocorrelation function of the unmeasured disturbance noise would be a step instead of an impulse. The inner-loop set point δr_I will be active, providing excitation to a mode (allowing identification of β_0 as well as α_0) that would not be excited in a single-loop structure when the set point r is fixed.

A restricted complexity model has fewer modes than the process. Consequently its modeling error will have components resulting from structural mismatch as well as unmeasured disturbances. It may have substantially fewer parameters than an "exact" model. For example, the C and β polynomials may be restricted to one term ($C = 1, BF = \beta_0$) to be certain that they have no unstable zeros. Less past history is needed to reliably identify a small number of parameters since fewer equations are needed to solve for fewer unknowns. Therefore a restricted complexity model can be updated more quickly, allowing it to better track a changing or nonlinear process. The identifier inputs Φ and Ω should be filtered in order to make the process-model match best in the critical frequency range, where the open-loop absolute gain is near 1 for feedback control or near steady state for open-loop, or feedforward, control.

The implicit model form can be used to identify the delay time. The same Φ vector can be used for a number of predictor models, each predicting Ω a different number of time steps k into the future. If d is the largest possible value of k, the identifier equations can be time-shifted to yield a common Φ [14]. At time step t,

$$\Omega_k\{t-d+k\} = \Phi\{t-d\}^T \Theta_k \tag{79}$$

The prediction model with the largest β_0 coefficient can be chosen for the controller design, since this model indicates the greatest sensitivity of the controlled (predicted) variable to the present manipulated variable. Hence it will result in the smallest controller gain. Furthermore, if more than one β coefficient is identified, the model with the largest β_0 is most likely to have stable zeros. The model with the smallest prediction error is most likely the one with the smallest k because of the autoregressive α terms, but this model would not necessarily be best for control. The identification filter time constants can be made proportional to the identified k, since k determines the closed-loop performance and the critical frequency range.

CONTINUOUS-MODEL IDENTIFICATION, OPEN-LOOP ADAPTATION

A continuous (differential equation) model, in contrast to a difference equation model, is insensitive to the computing interval h, provided h is small. A restricted complexity identifier, for a process that includes delay, can be based on the method of moments [15].

The Laplace transform $X\{s\}$ of each signal's derivative $x\{t\}$ can be expanded into an infinite series of moments,

$$X\{s\} = \int_0^\infty e^{-st} x\{t\}\,dt = M_0\{x\} - sM_1\{x\} + \cdots \tag{80}$$

Signal derivatives are used so that the moment integrals, for an isolated response, converge to near final values in the finite time τ from the disturbance start,

$$M_n\{x\} \approx \int_0^\tau t^n x\{t\}\,dt \approx \sum_{k=1}^{\tau/h} (kh)^n \times \{k\}h \tag{81}$$

The signal transforms are related to the model polynomials on a term-by-term basis. Choosing the B and D polynomials to be 1 and

$$\begin{aligned} A\{s\} &= a_0 - sa_1 + \cdots \\ C\{s\} &= c_0 - sc_1 + \cdots \end{aligned} \tag{82}$$

in the process equation

$$u\{s\} = A\{s\}y\{s\} - C\{s\}e\{s\} \tag{83}$$

gives

$$\begin{aligned} M_0\{u\} &= a_0 M_0\{y\} - c_0 M_0\{e\} \\ M_1\{u\} &= a_1 M_0\{y\} + a_0 M_1\{y\} - c_1 M_0\{e\} - c_0 M_1\{e\} \\ &\vdots \end{aligned} \tag{84}$$

For each additional equation there are one plus the number of measured disturbances e of additional unknown parameters. The projection algorithm can be used to find the smallest sum of weighted squared parameter changes that will satisfy the equations. Using projection, only those parameters weighting signals that are significantly active are updated. Equations (84) expressed in vector and matrix form, for use in the projection algorithm, are

$$
\begin{aligned}
\Omega &= \Phi^T \Theta \\
\Omega^T &= [M_0\{u\}, M_1\{u\}, \ldots] \\
\Theta^T &= [a_0, a_1, \ldots, c_0, c_1, \ldots] \\
\Phi^T &= \begin{vmatrix} M_0\{y\}, & 0, \ldots, -m_0\{e\}, & 0, \ldots \\ M_1\{y\}, & M_0\{y\}, \ldots, -M_1\{e\}, & -M_0\{e\}, \ldots \end{vmatrix}
\end{aligned}
\tag{85}
$$

The moment-projection approach is particularly suited for adapting feedforward gain and delay compensators, because the inputs need not be persistently excited. Only two moments need be computed for each signal and two moment equations solved by projection. However, when signals are cycling or responses overlap, the moment integrals do not converge and the adaptation must be frozen. Since an adaptive feedback controller should be capable of stabilizing a stabilizable unstable loop, the moment-projection method is not suited for adaptation of a feedback controller.

LEAST-SQUARES METHOD, BATCH PARAMETER IDENTIFICATION

A batch identifier calculates the model parameters that best fit a block of measured data. The start may be triggered when a significant disturbance is sensed and the end may follow the settling of an isolated response or the detection of a preset number of peaks for a cycling response. A least-squares identifier finds the parameter vector Θ that minimizes the sum of squared prediction errors,

$$\varepsilon\{t\} = \Omega\{t\} - \Phi\{t-k\}^T \Theta \tag{86}$$

When the inverse of the matrix $\mathbf{P}$ exists,

$$\mathbf{P}^{-1} = \sum_i \Phi\{i-k\}\Phi\{i-k\}^T \tag{87}$$

the result is given by

$$\Theta = \mathbf{P} \sum_i \Phi\{i-k\}\Omega\{i\} \tag{88}$$

In order that $\mathbf{P}^{-1}$ not be dominated by steady-state components of Φ, it is customary to choose Φ and Ω to have nearly zero mean. If the means were completely removed, $\mathbf{P}^{-1}$ would be the covariance of Φ and $\mathbf{P}$ would be the covariance of Θ. $\mathbf{P}$ also appears in the recursive algorithm of the next section.

For Θ to be calculable, $\mathbf{P}^{-1}$ must not be singular. Nonsingularity is difficult to guarantee. If any of the process inputs were quiescent over the identification period, $\mathbf{P}^{-1}$ would be singular. This could happen if the controller output were limited or if the controller were in the manual mode. When $\mathbf{P}$ exists, the process is said to be persistently excited. It may be necessary to add otherwise undesirable probing signals to the normal process inputs to achieve persistent excitation.

KALMAN FILTER, RECURSIVE PARAMETER IDENTIFICATION

The Kalman filter [16] provides a one-step-ahead prediction of the parameter vector Θ (treated as a state variable) in a model equation,

$$\Theta\{t\} = \Theta\{t-h\} + v\{t\} \tag{89}$$

modified by observations of a related measured variable scalar (or vector) Ω,

$$\Omega\{t\} = \Phi\{t-k\}^T \Theta\{t\} + \varepsilon\{t\} \tag{90}$$

The model equation, Eq. (89), has a zero-mean white gaussian noise source vector υ with covariance matrix $\mathbf{Q}$, which causes Θ to change randomly. The scalar (or vector) observation equation has a zero-mean white gaussian noise source scalar (or vector) ε, with covariance value (or matrix) R, in this case uncorrelated with υ. Depending on which noise source dominates, the Kalman filter weights the other equation more heavily,

$$\Theta\{t\} = \Theta\{t-1\} + K\{t\}(\Omega\{t\} - \Phi\{t-k\}^T \Theta\{t-1\}) \tag{91}$$

Θ is the predicted parameter vector and K is the time-varying Kalman gain vector (or matrix) which can be precalculated using

$$K\{t\} = \mathbf{P}\{t-1\}\Phi\{t-k\}(R + \Phi\{t-k\}^T \mathbf{P}\{t-1\}\Phi\{t-k\})^{-1} \tag{92}$$

$$\mathbf{P}\{t\} = (\mathbf{I} - K\{t\}\Phi\{t-k\}^T)\mathbf{P}\{t-1\} + \mathbf{Q} \tag{93}$$

When $\mathbf{Q}$ is a null matrix, this algorithm is the recursive least-squares algorithm, finding the parameter set Θ that minimizes the sum of squared model errors, equally weighted over all past observations. As time progresses the $\mathbf{P}$ matrix and gain K approach zero, so that eventually each new observation has almost no effect on the identified Θ. Therefore an unmodified recursive least-squares solver does not allow a model to adapt to a time-varying or nonlinear process.

Whereas R tends to reduce K and $\mathbf{P}$ by a factor each iteration, $\mathbf{Q}$ increases $\mathbf{P}$ by a fixed increment. Therefore $\mathbf{Q}$ has a relatively greater influence when $\mathbf{P}$ is small and vice versa when $\mathbf{P}$ is large. Thus when neither R nor $\mathbf{Q}$ is zero, $\mathbf{P}$ tends toward a midrange value. As a result, the Kalman gain K remains finite, so that the most recent observations affect the identified Θ, allowing the model to adapt to a time-varying or nonlinear process. In effect this method, in contrast to the variable forgetting factor approach, weights a different quantity of past history for each variable Φ_i, depending on its activity and the ratio of $\mathbf{Q}_{ii}$ to R.

When both R and $\mathbf{Q}$ are zero, the predictor becomes an orthogonal projection algorithm. Θ converges in a determinate number of iterations to a fixed vector, provided each of these observations contains some independent information. The number of iterations is the number of parameters in Θ divided by the number of components in Ω.

PROJECTION

If $\mathbf{P}$ is not updated and R is zero, the Kalman filter algorithm performs a projection each iteration, finding the smallest set of weighted squared parameter (state-variable) changes that satisfy the model equations exactly. The weighting matrix $\mathbf{P}$ is the a priori covariance of Θ.

REFERENCES

1. Astrom, K. J., and B. Wittenmark, *Computer Controlled Systems*, Prentice-Hall, Englewood Cliffs, New Jersey, 1984.

2. Takahshi, Y., M. J. Rabins, and D. M. Auslander, *Control Systems*, Addison-Wesley, Reading, Massachusetts, 1970.
3. Franklin, G. F., and J. D. Powell, *Digital Control*, Addison-Wesley, Reading, Massachusetts, 1980.
4. *Chemical Process Control Series*, American Institute of Chemical Engineers, New York, 1991.
5. Hansen, P. D., "Recovery from Controller Bounding," presented at the ASME Winter Annual Meeting, 1987.
6. Shinskey, F. G., "Putting Controllers to the Test," *Chemical Engineering*, pp. 96–106, Dec. 1990.
7. Slotine, J. J. E., and W. Li, *Applied Nonlinear Control*, Prentice-Hall, Englewood Cliffs, New Jersey, 1991.
8. Shinskey, F. G., *Process Control Systems*, 3rd ed., McGraw-Hill, New York, 1988.
9. Bristol, E. H., "On a New Measure of Interaction for Multivariable Process Control," *IEEE Trans. Automatic Control*, Jan. 1966.
10. Ricker, N. L., "Model Predictive Control: State of the Art," *Chemical Process Control Series*, American Institute of Chemical Engineers, New York, 1991.
11. Kraus, T. W., and T. J. Myron, "Self-Tuning PID Controller Uses Pattern Recognition Approach," *Control Engineering*, June 1984.
12. Hansen, P. D., "Recent Advances in Adaptive Control," presented at the 40th Chemical Engineering Conf., Halifax, Nova Scotia, Canada, July 1990.
13. Ljung, L., and T. Soderstrom, *Theory and Practice of Recursive Identification*, M.I.T. Press, Cambridge, Massachusetts, 1982.
14. Hansen, P. D., and T. W. Kraus, "Expert System and Model-Based Self-Tuning Controllers," in *Standard Handbook of Industrial Automation*, D. M. Considine, Ed., Chapman and Hall, New York, 1986, p. 216.
15. Bristol, E. H., and P. D. Hansen, "Moment Projection Feedforward Control Adaptation," in *Proc. 1987 American Control Conf.*, p. 1755.
16. Goodwin, G. C., and K. S. Sin, *Adaptive Filtering, Prediction and Control*, Prentice-Hall, Englewood Cliffs, New Jersey, 1983.
17. Ray, W. H., *Advanced Process Control*, New York, McGraw-Hill, 1981.
18. Hang, C. C. et al., *Adaptive Control*, Instrument Society of America, Research Triangle Park, North Carolina.
19. Wang, H. et al., *Advanced Adaptive Control*, Tarrytown, New York, Elsevier Science, 1995.
20. Luyben, W. L., and Luyben, M. L., *Essentials of Process Control*, McGraw-Hill, New York, 1997.

BASIC CONTROL ALGORITHMS

by E. H. Bristol*

INTRODUCTION

Continuous process control traditionally used analog control devices naturally related to the process. The modern era has largely replaced these with microprocessors and digital computations to reduce cost, but also for greater flexibility, supporting more diverse algorithmic forms. Initially digital control was implemented by the conversion of each analog control function into some (perhaps generalized) digital equivalent [1–4].

* The Foxboro Company, Dept. 0346, Bldg. C41–1H, Foxboro Massachusetts, 02035.

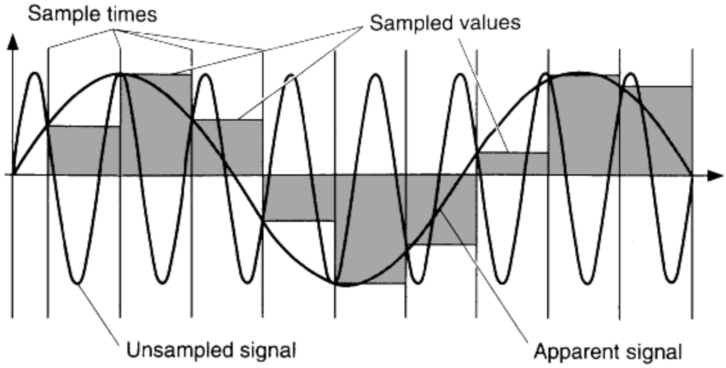


FIGURE 1 Aliasing of sampled to apparent signal.

This conversion requires an appropriate software architecture (more about that below) and the matching of the digital algorithms to analog behavior. Whereas the analog controller continuously sensed the process state and manipulated the actuators, the digital controller must repeatedly sample that state, convert it to a quantized number, use that number to compute control actions, and output those actions. Each of these steps involves its own problems and errors.

Standard digital control texts address the sampling problem thoroughly [5], but treat the broader control design in terms of neutral, computed parameters not clearly related to the process gains and time constants. Process control depends on standard control algorithms whose parameters may be set or tuned in terms of known process properties. The approach defined herein emphasizes control with traditional parameters, naturally related to the process and practice. This includes compensating sampled algorithms and their tuning properties for the sampling time, or even for irregular sampling.[1]

One aspect of sampling should be understood: aliasing. This is a bizarre effect in which sampled higher-frequency data appear to be at a much lower, more significant frequency (Fig. 1). The same effect limits the minimum apparent width in time of any disturbance pulse to one sampling period. A control loop attempting to control such a false disturbance causes a real one instead.

Aliasing is not normally a problem because most commercial controllers sample frequently and filter out the still higher frequencies. But users should be aware of aliasing if they construct home-brewed control algorithms of low sampling frequency.[2]

Except for aliasing, the choice of sampling time is not an issue today. Increasing the sample times faster than the dominant closed-loop time constants of the economically important process variables gives rapidly diminishing performance returns.[3]

For example, with end-product quality the sole criterion of a process that takes an hour to respond, even fast flow loops can be sampled once every 5 min. Of course, many "housekeeping functions," like flow control, may have constraining side effects whose violation would involve real costs if this logic were actually implemented. Sampling times faster than 1 s, well filtered, are rarely needed in continuous fluid process control, even when the local process dynamics are faster (as with flow loops).

The inexpensive microprocessor has caused a conservative design trade-off to favor faster sampling times, eliminating the issues for which the cost is so small. Nevertheless, a 10-to-1 reduction in

---

[1] The resulting parameter forms may look unnecessarily approximate. But recent standards efforts [6] have argued for even simpler control parameters. In either case it will always be possible to replace the proposed parameters by computations that support the more exact or simpler form. The chosen forms are based on the understanding that any use of digital control with analog process is inherently approximate and that intuitive tuning of parameters is the most important design consideration of these approximations. Apparently formal calculations of the parameters will always be misleading.

[2] Model-based control techniques, like internal model control and dynamic matrix control, are often constrained to operate at low sample times for reasons of modeling sensitivity.

[3] However, slow sampling frequency also causes tuning sensitivity, even for frequencies fast enough to give good control. Further doubling or tripling the frequency beyond this point should eliminate even this problem.

sampling time does correspond to a 10-to-1 reduction in computing resources if one has the tools and imagination to use the capability. Faster sampling times also require that internal dynamic calculations be carried out to a greater precision.

Process control algorithm design also differs from academic treatments in respect to accuracy considerations. Practice designs are for control and human operation, not for simulation or computation. When accuracy is important, a common thread through the discussion is the effect of differences among large numbers in exaggerating error and the role of multiplication in creating these large numbers.

Rarely is the parameter or performance precision (say, to better than 10%) important to control, even under adaptation. While there are sensitive high-performance, control situations, a 2:1 error in tuning is often unimportant. In contrast, there are situations within an algorithm in which a small error (say, a quantization error of 1 part in 100,000) will cause the control to fail. Computational perfection is unimportant; control efficacy is crucial.

A casual experimenter or designer of a one-use algorithm should be able to tailor a simple controller in FORTRAN floating point without any special considerations. The normal debugging and tuning should discover any serious deficiencies. A commercial design, applied in many applications, calls for a deeper understanding of algorithmic and fixed- versus floating-point issues and trade-offs. Fixed-point and machine language programming may become a lost art, even though their difficulties are overstated. The advantages, in speed and exactitude, are worth consideration, particularly on small control computing platforms, for which they may be essential. The problems are simply those of understanding and of effective specification and testing of the algorithm.

The discussion thus addresses the refined design of process-control-oriented continuous control algorithms and their software support. It emphases high-quality, linear, dynamic algorithms. This still includes representative examples of discrete computation's effects on modeled continuous control activities:

- imprecision, as it limits control performance,
- quantization, as it artificially disturbs the process,
- sampling, as it affects time continuous computations such as dead time.

In this environment, nonlinearity serves to compensate for a basically linear structure. One special concern, addressed below, is the backcalculation required for some aspects of these compensations.

The user interested in the more general computation of nonlinear functions should consult the general computing literature [7, 8]. The discussion briefly touches issues relevant to advanced forms like adaptive control, but their details are outside the intended scope.

STRUCTURE OF TRADITIONAL PROCESS CONTROL

The traditional process control structure is based on the combination of simple (PI/PID) controllers in simple single or cascaded loops, augmented with various nonlinear, feedforward, and dynamic compensators and constraint overrides.

Figure 2(*a*) illustrates the basic cascaded structure. In this structure, a primary controller, controlling a corresponding process variable, manipulates the set point to a secondary controller that acts to stabilize and control the corresponding secondary process variable.

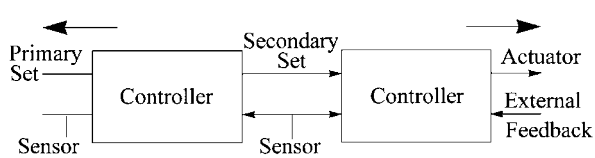


FIGURE 2(*a*) Basic cascaded structure.

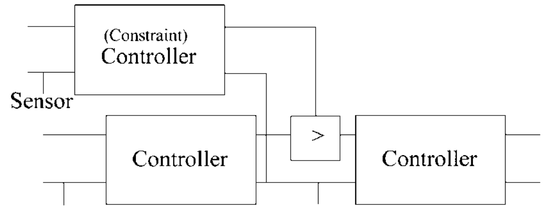


FIGURE 2(*b*) Constraint structure.

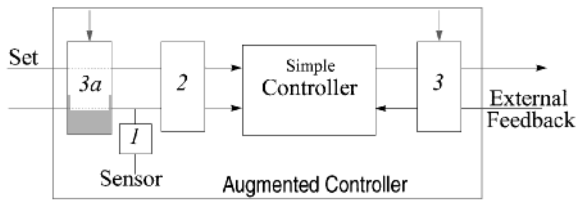


FIGURE 2(*c*) Augmented controller.

It will be noted that each controller has two connections on both the input side and the output side. On the input side these serve as the standard set point and measurement connection. The output side includes the normal output connection as well as the external and status feedback connections, described below, whose purpose is to adapt the control to loss of output control.

The structure can be continued with any number of controllers, defining a degree-of-freedom path of controlled variables. The control of the primary variable is then supported by successively more secondary variables until the ultimate control falls on the actuator. The single degree of freedom represented by that valve is then passed back along that path until it is the primary variable that is effectively free for manipulation.

The degree-of-freedom path can be overridden by any number of constraint controllers or limiters. Figure 2(*b*) illustrates the constraint structure. The selector symbol (>) represents a computation that outputs the greater of two inputs, thus imposing a low limit on the original degree-of-freedom path.

The constraint controller acts in feedback to set that limit value that will ensure that its measured value stays within the constraint limit defined by its set point. When the constraint controller takes over to enforce the constraint, it takes over the degree-of-freedom path, so that the path now runs from the constrained process variable through the secondary variable to the actuator. Any number of low- or high-limit or constraint override structures could similarly be included within the cascaded structure.

The traditional structure allows each controller in a cascade/constraint structure to be augmented in a variety of ways that enhance the control function without changing the underlying intent and structure. The figure shows the three distinct structural roles of this kind of compensation, with several variants:

1. Computation of the desired measurement form from the sensor signal.
2. Linearization compensation of the controller in terms of the controller input (applying the same computation to setpoint and measurement without changing their operationally displayed form).
3. Linearization compensation of the controller at its output, which requires the application of the compensation to the output, the inversion of that same compensation (as described below) into the external feedback, and the appropriate backpassage of any status information. A modified form of this structure is needed to support feedforward control, in which the disturbance signal (with any dynamic compensation) is brought into the compensating calculation as a kind of computing parameter.[4]

[4] A later distinction will be made about the bumpless application of tuning parameter changes. The feedforward signal will usually be intended to incorporate both any appropriate parameter change and the compensating disturbance bump!

3a. Partly for historical reasons, ratioing control is usually included as an operationally accessible setpoint parameter multiplying a feedforward measurement to generate the true controller reference input.

A discussion below will show a branching out of the degree-of-freedom paths, matching the constraint branching together. The collective structures allow the implementation of all normal single-loop control. And each of the structures has multivariable counterparts.

NUMBER SYSTEMS AND BASIC ARITHMETIC

Operation on analog data involves not simply converting to decimal numbers, but also converting to the format that those numbers take in the digital computer. There are two such formats: fixed point for representing integers and floating point for representing real numbers. Even though analog data are usually conceived of in terms of real numbers, the most effective processing of continuous control data, particularly for small microprocessors, is in fixed-point arithmetic.

In this case, the data must be (painfully) scaled in the same way that analog control systems and simulations were scaled. Scaling and multiple-precision computations and conversions are the explicit responsibility of the control algorithm programmer. Apart from computational necessity, scaling is an art that is inherently significant to proper control design, since the process is itself fixed point. Meaningful control actions can be guaranteed only if the user is aware that the scale of the calculations is in fact appropriate to the process.

Floating-point data are automatically scaled, in that they have a fractional part f or mantissa, corresponding to the integer part of fixed-point data, and an exponent part e, which defines an automatically adjusted scaling factor equal to a base value b (usually 2) raised to that exponent. Although the fractional part is usually viewed as a fraction, the discussion will be less confusing if e is chosen so that f is an appropriate integer; any such value is then expressible entirely in terms of integers, as $f \times b^e$.

In the past, floating-point formats were not standardized [9]. Moreover, the floating-point format inherently involves arbitrary truncations and uncertain relationships between single- and double-precision computation. For this reason, fixed-point algorithms still permit the most precise designs.

Fixed-Point Format

A fixed-point format represents an integer as a binary number stored in a word containing a fixed number (n, below) of bits, usually in what is called two's complement format. This format represents positive numbers in a range of 0 to $2^n - 1$ and signed numbers in a range of -2^{n-1} to $2^{n-1} - 1$. In two's complement arithmetic, negative numbers are represented in binary notation as if they were larger positive integers. As a result, when the numbers are added, the natural additive truncation achieves the effect of signed addition.

Thus, in Fig. 3, (with $n = 3$) negative 2 is represented by a binary number corresponding to a positive 6. And 6 added to 2 becomes an 8 that, with truncation of the carry, corresponds to 0. Thus the 6 is a perfectly good negative 2.

Two's complement arithmetic is also related to modular, or around-the-clock, arithmetic. Because of this behavior, results too large to fit into a word must be taken care of in one of three ways:

- Data values can be distributed in a sufficiently large set of data words to represent them. The arithmetic hardware or software then utilizes carry bits to pass the data between words when a carry is required.
- Results too large to fit a word may be saturated.[5]

[5] That is, a value too large to fit the word is replaced with the largest number of the right sign that will fit.

Sign and Value	Two's Complement	
Decimal	Binary	Positive Decimal Equivalent
0	000	0
1	001	1
2	010	2
3	011	3
–4	100	4
–3	101	5
–2	110	6
–1	111	7
(2)+(–2)	Carry bit → 1000	(3)+(5)=8

FIGURE 3 Two's complement arithmetic.

- One may ignore the problem in parts of the calculation for which it is clear that results will never overflow.[6]

The remaining discussion will be framed in decimal arithmetic, rather than in the unfamiliar binary arithmetic, for clarity's sake. Sufficient to say, actual designs are implemented by good fixed-point hardware including the basic binary operations and carry/overload bits to support effective single- and multiple-precision arithmetic. For this reason, multiplication and divisions of powers of 2 will be particularly efficient, to be preferred when there is a choice.

Fixed-Point Scaling

In the discussion of multiple precision and scaling, it is convenient to distinguish different tag ends to the variable name to represent different aspects of the variables under consideration. Thus if *V* is the name of a variable

- *V.S* will become the scaled representation of the variable, taken as a whole.
- *V*.0, *V*.1, *V*.2, . . . , *V.m* (or *V.S*.0, *V.S*.1, *V.S*.2, . . . , *V.S.m*) will be different storage words making up the multiple-precision representation of *V* (or *V.S*), with *V*.0 being the right-most word and *V*.1 being the next right-most word, etc., as shown in Fig. 4.
- When *V* is scaled by a fixed-point fraction, the numerator will be *V.N* and the denominator will be *V.D* [either of which may be itself in multiple-precision form (with *V.N*.0, *V.N*.1, . . . , or *V.D*.0, *V.D*.1, . . .]:

$$V.S = \frac{V \times V.N}{V.D}$$

V	V.0	V.1	V.2
123456	12	34	56

FIGURE 4 Decimal multiple precision.

Normally the actual scaling computations will take place only when the data is processed for input/output (I/O) or display; control calculations will generally take place with respect to *V.S.* For this reason, when no confusion arises, the discussion will refer to *V*, substituting for *V.S.* Scaling conversions, like binary and decimal arithmetic and conversions, are straightforward though tedious; they need to be addressed further.

The special power of this notation is that it allows all scaling and multiple-precision computations to be expressed entirely in conventional algebra. In particular, for analytical purposes, a

[6] For which they will never become too large or too small to fit the word. This is a dangerous approach.

multiple-precision value can be expressed and operated on simply as a sum of the normal values $V.0$, $V.1$, $V.2$, ..., $V.m$, each with its own scaling factor:

$$V = V.0 \times B^m + V.1 \times B^{m-1} + V.2 \times B^{m-2} + \cdots + V.m$$

The value B is one greater than the largest positive number represented in a single-precision word ($B = 2^n$ for unsigned binary data; $B = 2^{n-1}$ for signed data).

Control algorithm parameter scaling is chosen according to need, computational convenience, and best use of the available data range, working with single-precision representation as much as possible. For a controller measurement or value on a 16-bit word machine, 1/2% precision is normally minimally adequate. For positive data, a byte represents a data range of 0–255 ($2^8 - 1$), which is better than 1/2% precision.

On the other hand, with a 16-bit word, there are 8 more bits in the word that could be used to give a smoother valve action and simplify calculations. When only positive values are worked with, 16 bits correspond to the normally unnecessary precision of 1 part in 65,535; with signed data, 1 part in 32,767; or, with a 1-bit safety margin, 1 part in 16,384. This scaling is well above the minimum process data precision, while still not forcing multiple-precision arithmetic in most cases. The corresponding 14-bit analog-to-digital (A/D) and digital-to-analog (D/A) converters for the process data are still reasonable.

Control parameters may have different natural data ranges or scalings. A controller gain might be scaled so that the minimum gain is 1/256 and the maximum gain is 256. In this way the range of values equally spans high-gain and low-gain processes about a nominal unity gain. Time constants may require a certain resolution. When a minimum reset time of 1 s is adequate (1 h corresponds to 3600 s), 2^{16} corresponds to more than 18 h.

On the other hand, using the full range of data storage for the control parameters may require arithmetic routines that mix signed and unsigned arguments. At this level, there is a trade-off between the increased efficiency and the added complexity. Certainly it is more convenient to program uniformly in signed fixed-point (or even floating-point) arithmetic. But the costs of this convenience are also significant.

Range and Error in Fixed-Point Arithmetic

Good design for quantization and multiple-precision avoids poor control arising from inaccuracies not inherent in the real process. Normal control algorithms involve the standard combinations of additions, subtractions, multiplications, and divisions; rational functions of their data. Often the basic calculation allows many ways to order these calculations that are theoretically identical in their result as long as precision is indefinite.

Practical fixed-point programming requires a more careful understanding of the basic operations and the effect of ordering on the calculation. First, one should examine how each operation affects the worst-case range[7] and the error accumulation of the result. As will be seen, range effects are typically more important.

Error can be considered in terms of absolute error, i.e., the actual worst-case error, or relative error, i.e., the worst-case error as a percentage of the nominal value or scale. Usually the relative error is important in final results and products (or quotients), whereas the absolute error is important in determining how the error accumulates in a sum.

When two numbers are added or subtracted, their worst-case range doubles, either requiring a carry (into a multiple-precision result) or requiring a sign bit[8] [see Fig. 5(*a*)]. Additions and subtractions also cause the absolute error to increase. Addition of numbers of the same sign can never increase the relative error over the worst error of the added numbers. But their subtraction increases it, as does addition of numbers whose sign is uncertain. In Fig. 5(*b*),[9] two 10% relative error numbers, when

[7] The range of result for all possible combinations of the input data.

[8] When positive numbers are subtracted.

[9] In Fig. 5(*b*) the underlined numbers represent an error term being added or subtracted from the "ideal" value of a computational input or output.

$$99 + 99 = 198$$
$$99 + (-99) = 0$$
$$99 - 0 = 99$$
$$0 - 99 = -99$$

(*a*) Range effects

$$(20 \pm \underline{2}) + (10 \pm \underline{1}) = (30 \pm \underline{3})$$
$$\frac{(20 \pm \underline{2}) + (10 \pm \underline{1})}{20 + 10} = \left(\frac{30}{30} \pm \frac{\underline{3}}{30}\right)$$
$$\frac{(20 \pm \underline{2}) - (10 \pm \underline{1})}{20 - 10} = \left(\frac{10}{10} \pm \frac{\underline{3}}{10}\right)$$

(*b*) Error effects

FIGURE 5 Addition and subtraction.

$$\overline{99} \times \overline{99} = \overline{98}\ \overline{01}$$
$$\overline{98}\ \overline{01} / \overline{99} = \overline{99}$$
$$\overline{98}\ \overline{00} / \overline{99} = \overline{98} \text{ Rem.: } \overline{98}$$
$$\text{but: } \overline{98}\ \overline{00} / \overline{1} = \overline{98}\ \overline{00}$$

(*a*) Range effects

$$(10 \pm 1) \times (10 \pm 1) = (100 \pm 21)$$
$$\frac{(100 \pm 8)}{(10 \pm 1)} = 10 \pm 2$$

(*b*) Error effects

FIGURE 6 Multiplication and division.

added, result in a 10% relative error. But when subtracted, the relative error can be arbitrarily large (30%) in this case. In this way, differences between large numbers explain most computing accuracy problems.

Fixed-point multiplication does more than just increase the range of the result; it doubles the required storage of the result [see Fig. 6(*a*)]. For this reason, most hardware implements single-precision multipliers to return a double-word result. By analogy, most hardware division divides a double-precision dividend by a single-precision divisor to obtain a single-precision quotient with a single-precision remainder. But, as Fig. 6(*a*) shows, such a division can still give rise to a double-precision quotient. The hardware expresses this as an overload (similar to division by zero).

Fixed-point hardware is designed to permit the effective programming of multiple-precision arithmetic. One of the strengths of fixed-point arithmetic, with its remainders, overflow and carry bits, and multiple-precision results is that no information is lost except by choice. The precision, at any point in the calculation, is entirely under the control of the programmer. This is not true of floating-point arithmetic. As shown in Fig. 6(*b*), multiplication and division always increase (e.g., double) relative error.

Fixed-Point Multiplication and Division

Multiplication's range explosion is its most problematic aspect. Among other consequences, it generates large numbers whose differences may cause large errors. It makes smaller numbers still smaller compared with errors caused by other large numbers. The important issue is the ratio of the large numbers (which make the errors) to the small numbers (which end up as the result). Multiplication's range expansion forces a choice: precision can be preserved or intermediate values can be rescaled and truncated back to their original data size.

Thus it is usually desirable to avoid repeated multiplications. Often multiplications and divisions occur together in a manner in which they can be alternated, the multiplication generating a double-precision value and the division returning that value to a single-precision quotient. The basic proportional controller calculation is a good example:

$$\text{Output} = \frac{100 \times \text{error}}{\text{Proportional Band}} + \text{bias}$$

Common combined operations such as this may usefully call for specially designed routines. In general, one should try to maintain a constant data range and size throughout the computation. Often a product is naturally intended to return a value in the same range as a process variable (output). There are two natural cases: either a value is multiplied by a gain or by a proper fraction (Fig. 7). In the controller calculation the 100/Proportional Band gain might be limited to the range of 1/256–256. If the gain is expressed as a single-precision integer (with a decimal point effectively in the middle of its digits) rather than as a fraction, one can still generate an appropriate single-precision result by taking the middle single-precision set of digits out of the double-precision scaled result [Fig. 7(*a*)], saturating and truncating the extraneous data.

(*a*): $\overbrace{\text{XX.XX}}^{\text{Gain}} \times \overbrace{\text{XXXX}}^{\text{Natural Range}} = \underline{\text{XX}}\overbrace{\text{XXXX}}^{\text{Natural Range}}.\underline{\text{XX}}$

(*b*): $\overbrace{\text{.XXXX}}^{\text{Fraction}} \times \overbrace{\text{XXXX}}^{\text{Natural Range}} = \overbrace{\text{XXXX}}^{\text{Natural Range}}.\underline{\text{XXXX}}$

FIGURE 7 Scaled multiplication of (*a*) gain, (*b*) fraction.

As a common case of proper fractional multiplication, a lag calculation (time constant T) can calculate its output X as a weighted average of the past output Y' and the current input X:

$$Y = \frac{1}{T+1}X + \frac{T}{T+1}Y' = Y' + \frac{1}{T+1}(X - Y')$$

In this case, when the proper fractions are multiplied (scaled as integers), the proper result can nevertheless be returned as the left-most part of the double-precision result [Fig. 7(*b*)]. It is often appropriate to combine constants or parameters together into a common effective parameter that is the above kind of gain or proper fraction. This ensures that the linear operation with the process data is truncated by only the one final multiplication.

The combined parameters can be made to act consistently on the data even if in error as calculated; the errors can be reinterpreted as (presumably insignificant) errors in the original parameters. Note that the right-most expression within the above equation involves a feedback between Y and the difference between X and Y'. Often such a feedback can be included in the calculation to improve an otherwise error-prone algorithm, making it self-corrective, like any other feedback system.

There is a parallel between the above range discussion and traditional dimensional analysis [10]: sums must be of data elements with identical units and similar ranges. Multiplications change the units and also change the range. The ultimate purpose of any of our calculations is to convert data of limited range from the process to data with limited range to drive the process. Thus, whatever the internal gyrations, the process constrains the results to be reasonable. The challenge is to carry out the calculations so that the inherently limited practical range prevails throughout the calculation.

Digital Integration for Control

Figure 8 shows a different kind of division problem: process control integration. The object is to integrate the process error in a control (reset) calculation. This can be done by computation of a shared fractional multiplier $\Delta t/T$, which has been scaled to give a good range [which might also include the proportional band action *PB*, as in $100 \times \Delta t/(PB \times T)$]. This result would be multiplied by the error and added in double precision to the previous integrated (i.e., summed) value to get the current control value.

Double precision is essential here, since a large value of T corresponds to a small $\Delta t/T$ and a truncation loss of significant process errors. For example, suppose that the error and the sum are both scaled as signed single-precision integers, ranging from $-10{,}000$ to $+10{,}000$. With $\Delta t = 1$ and the minimum $T = 1$ (corresponding to 1 s), the corresponding maximum $\Delta t/T$, which equals 1, must be scaled to a value of 10,000. The natural scaling of the controller output can be achieved by dividing by 10,000 (see later scaling discussion below). That is, the scaled equations should be

$$\text{Sum}.S = \text{oldSum}.S + \left[\frac{(\Delta t/T).S \times \text{Error}.S}{10{,}000}\right]$$

$$\sum_{i=0}^{n} \frac{Error\ (i)\ \Delta t}{T} =$$

$$\frac{\sum_{i=0}^{n} Error\ (i)\ \Delta t}{T} = \frac{\sum_{i=0}^{n-1} Error\ (i)\ \Delta t}{T} + Error(n)\ \frac{\Delta t}{T}$$

$$= \sum_{i=0}^{n-1} \frac{Error\ (i)\ \Delta t}{T} + Quotient\left(\frac{Error(n)\ \Delta t + Remainder}{T}\right)$$

Note:

$$Error\ (n)\ \Delta t = T \cdot Quotient\left(\frac{Error(n)\ \Delta t + Remainder}{T}\right)$$

FIGURE 8 An integration trick.

In normal single-precision division, only the integer quotient is considered. Thus a product $[(\Delta t/T).S \times \text{Error}.S]$ less than 10,000 is truncated as zero: The error is ignored and the integration stops, causing a permanent offset. For T large, equal to 10,000 (which corresponds to $\sim$3 h) and $(\Delta t/T).S$ small (equal to 1.0 in this case), any error less than 100% (scaled to 10,000) will be lost. Under control the result is a 100% offset. Double precision alleviates the offset but still loses some information to truncation.[10]

The better method, shown in Fig. 8, preserves the division and achieves an exact result with the same storage as double precision. In this case, the product of the error and the sample time (the sample time may equal 1) is formed as a double-precision value, then added to the remainder from the previous sample calculation's division. This net value is divided by T to get back a single-precision quotient to be added to the old sum, and a remainder to be saved for the next sample time. Data results truncated from any quotient are never lost, but preserved and accumulated in the remainder, to show up in some later quotient.[11]

Floating-Point Format

Modern microprocessors are often supported with high-speed floating-point processors. Without these, floating-point calculations run an order of magnitude slower than fixed-point calculations. With the floating-point format, all remainders, carries, and multiple-precision results of single-precision computations disappear. Instead, the user chooses a fixed level of precision whose truncations show up in guard bits [9].

Floating-point errors introduced in small differences of large numbers become more severe, and at the same time less obvious because their processing is automated and invisible. For example, when a large number L (e.g., 10,000, stored to three-digit accuracy) is added to a small number S (e.g., 1), the smaller number totally disappears. The sum $L - L + S$ should equal S. If the calculation is carried out in the natural ordering, $(L - L) + S$, then S will result. But the nominally equivalent $(L + S) - L$ will generate zero.

[10] One alternative to multiple-precision integration uses random numbers (or dither, in mechanical engineering terms). The computation is carried out in normal double precision. However the higher-precision (lower-order) data word is ignored and not saved; it is replaced with a random number in the next sample time's calculation when it is needed again. This method works well practically and theoretically. A similar method was incorporated into the Digital Equipment Corp. PDP-1 floating-point package. But who would have the courage to use it on a real process?

[11] As pointed out in the next section, floating point lacks the support for such refined tricks. It is particularly subject to integration problems because a large sum may be big enough to prevent the addition of a small term, even though that term is rescaled to avoid its being truncated to zero.

There are a number of intricate ways of avoiding this problem:

- Convert all floating-point numbers to ratios of integers and operate in the fixed-point format. This is one way to study the properties of the algorithm.
- Use adequate precision. Practically this is often unpredictable, and theoretically it is impossible because repeating decimals require infinite data. This suggests also deferring division to the last operation.
- Reorder the calculation for the most favorable computation, either statically based on algorithm properties or by using an arithmetic package that continually reorders the operands. It is useful, in the following discussion, to consider every list of terms to be added (subtraction being taken as addition of a negated number) to be ordered by magnitude. The individual operations are carried out according to these rules.

Under addition or subtraction

- combine small numbers together before large numbers (to let them accumulate before being lost),
- take differences (subtractions or additions of numbers of opposite sign) before sums (to achieve all possible subtractive cancellation between larger numbers before these can lose the small numbers),
- from a list of numbers to be multiplied and divided, cancel or divide out most nearly equal numbers first (to minimize floating-point overload[12]),
- when all denominator terms are gone, multiply largest with smallest numbers (to minimize floating-point overflow).

One other advantage of following such a set of ordering rules is that it will give identical results to identical data even when they originally occurred in a different programmed order.

Generalized Multiple-Precision Floating-Point

Normally, the multiple-precision floating-point format is the same as the single precision with larger fraction and exponent data fields. The author has experimented with a more open-ended multiple-precision floating-point, illustrated in Fig. 9. The multiple-precision floating-point number is represented by a summed, ordered set of signed single-precision floating-point numbers, each mantissa having M digits.[13]

The numbers in the set are chosen so that the set can be truncated at any point to give the best possible truncated approximation. In this case, the value is considered to be made up of the truncated value and a signed (±) remainder that expresses the part cut off in truncation.[14] The generalized floating point would be supported by the following operations:

$$12340.36789 \Rightarrow 1234 \times 10^{1} + 3679 \times 10^{-4} - 1000 \times 10^{-8}$$

Addition:

$$A = 3057 \times 10^{6} \qquad B = 4263 \times 10^{4}$$

$$A + B = 309963 \times 10^{4} = 3100 \times 10^{6} - 3700 \times 10^{2}$$

FIGURE 9 Generalized multiple-precision floating point ($M = 4$).

[12] The process of the floating-point exponent's getting too large for the provided data space.

[13] Analogous to multiple-precision fixed-point data represented as a summed scaled set of single-precision numbers. Each element is normalized (shifted), to use the full range of M digits. Recall that the general representation of a single-precision floating-point number is $f \times b^{e}$, where f, e, and b are integers. As above, the format is described in terms of a general b and illustrated with $b = 10$. In practice b will equal 2. Note that the different members in the summed set may have different signs! However, the sign of the total (set) value is still the sign of its initial and largest element.

[14] The remainder magnitude is always less than half the unit value of the next larger element in the set, since the best approximation requires that the remainder round to zero.

- Text conversion, converting to or from text general precision floating point to the internal format consisting of the summed set of single-precision values.
- (Set) Normalization, reprocessing the set of single-precision values so that they are ordered in magnitude, with largest first,[15] so that the magnitude of each mantissa is between b^{M-1} and b^M and so that consecutive values have exponents whose difference is greater than M (or, if the difference equals M,[16] then the magnitude of the second mantissa is less than or equal to $b^M/2$) [11].
- Addition and subtraction, merging entries into a final value set followed by a normalization of that set.
- Multiplication, multiplying of every pair of elements, one from each of the multiplicand and multiplier, to generate a double-precision result, followed by the merging and the normalization of the accumulated set of results.
- Division, dividing the largest elements in the dividend by the divisor, subtracting the divisor multiplied by the quotient from the dividend to obtain the remainder. This division is designed to select that quotient that returns the remainder with the smallest magnitude (of whatever sign). The remainder can be redivided by the divisor to compute any desired level of multiprecision quotients, with the final resulting quotient and remainder being normalized. The remainder so developed can be used in the earlier integration procedure.

Such a generalized floating point can be used to develop calculations whose precision expands indefinitely as needed. Such a system could give absolutely error-free results, without any special care.

SPECIFICATION OF FIXED-POINT ALGORITHMS

Clear fixed-point specification includes the proper statement of computation order and of scaling of intermediate and final results. This can be superposed on a conventional algebraic notation. For example, in the following control computation, the parentheses define any required ordering:

$$\text{Output}_5 = \left(\left(\frac{100 \times \text{Error}_1}{\text{Proportional Band}_2}\right)_4 + \text{Bias}_3\right)_5$$

Unparenthesized addition and subtraction are assumed to be from left to right and multiplication and division are assumed to alternate, as described above. Any constants that can be combined would certainly be combined in a working system. This should include any scaling constants. The subscripts refer to scaling specifications in the table of Fig. 10 below.

Such a table would completely specify all scalings, saturations, and conversions appropriate to the values within the calculation. Remembering that the internal scaling is reflected in the relation

$$V.S = \frac{V.N \times V}{V.D}$$

Subscript	I/O	Saturation Hi	Saturation Lo	*V.N*	*V.D*	Sign	Precision
1	IN	100	-100	16384	100	±	1
2	IN	16384	0	2	1	+	1
3	IN	100	-100	16384	100	±	1
4	–	YES		16384	200	±	1
5	OUT	100	0	16384	100	+	1

FIGURE 10 Table of variable scaling-related properties.

[15] In the form $f_o \times b^{e_o}, f_1 \times b^{e_1}, f_2 \times b^{e_2}, \ldots$.
[16] Not possible after normalization if M equals 2.

we find that the conversion from functional algebraic equations to internal computational form is then carried out by the computation

$$V = \frac{V.D \times V.S}{V.N}$$

When all of these scalings are incorporated back into the original proportional controller calculations, its internal scaled form becomes

$$\text{Output}.S \times \frac{100}{16{,}384} = \frac{100 \times \dfrac{100}{16{,}384}\text{Error}.S}{\dfrac{1}{2} \times \text{PropBand}.S} + \frac{100}{16{,}384} \times \text{Bias}.S$$

or

$$\text{Output}.S = \left(\left(\frac{200 \times \text{Error}.S}{\text{PropBand}.S}\right) + \text{Bias}.S\right)$$

Such a simplification is to be expected in practical calculations as part of the combination of similar scaling terms and application constants. The final equation then becomes the programmed calculation, except that each operation would be saturated as specified. Saturation can, in fact, make the combination of terms invalid, but in this case, it may be worth considering whether or not the saturations might not better be left out, in the interests of a more perfect result.

Definition and implementation of an algorithm then have three parts:

- specification and programming of the necessary arithmetic and saturation routines,
- manual development of the scaled calculation, in terms of the routines,
- programming of the algorithm.

Ideally, an algorithm should be tried out first in a higher-level language (e.g., FORTRAN or C). Here it can first be expressed in floating point and then in scaled fixed point format. If it is later needed that the final form be in machine language, the three different forms can be run comparatively, greatly facilitating debugging.

OPERATIONAL ISSUES

The basic controls are normally expressed as linear algorithms, defined as if the process measurements and actuators were capable of perfect operation to whatever range might be needed. In fact, valves limit and sensors fail. The algorithms must be designed to accommodate temporary valve saturation or loss of sensor data. Moreover, they must be designed to allow the system to be restarted smoothly after longer-term failure or shutdowns.

The most common such problem is windup: the property of the controller that continues to integrate under error even after the actuator has limited, and is incapable of, further change. Windup requires recovery time even after the error reverses sign, blocking effective control, for that interval. The response to this problem requires that the algorithm be provided some indication of the limiting so that it can alter its behavior. The information can take the form of flags that inform the controller of the limiting actions being taken, propagated back to any affected controller. The flag then causes the integration to stop in some appropriate way.

A superior approach, called external feedback, senses the actual state of the manipulated variable and feeds it back, in comparison with its intended value. By working with the true state of the process, the algorithm can make much more refined accommodation strategies. Controller windup is not the only problem arising from actuator limiting; Blend pacing, split-range control, and multiple output control all relate to standard affects of valve limiting and its accommodation. The external-feedback strategy effectively unifies the handling of all of these issues.[17]

The problem is complicated because the effects of limiting must propagate back from the actual valve, through any control functions (e.g., cascaded controllers, ratio units), to the controller under consideration. Thus there is not only a need to notify controllers of valve limitings or sensor failures but to propagate this information to all affected control elements. Software must be designed to accomplish the propagation.

The external-feedback approach is particularly advantageous because it recognizes loss of control locally to the affected controller and responds only when the loss is material to it. Of course, the loss of control at any level may be a useful basis for alarming, independent of immediately recognized effects on control. The present discussion is largely limited to the algorithmic consequences of the problem, but the software consequences are just as important.

A common problem with ad hoc controller designs is that they bump the output whenever the parameters are changed for tuning. The PI controller computation below illustrates the problem:

$$O(t) = \left(\text{Error} + \frac{1}{\tau}\int_0^t \text{Error} \cdot dt\right)\frac{100}{PB}$$

$$= \frac{100}{PB}\text{Error} + \int_0^t \frac{100}{PB \cdot \tau}\text{Error} \cdot dt$$

The first of two nominally equivalent expressions computes the output with the tunings acting after the integration. The practical result is that any change in settings immediately bumps the output. In the second expression, the tunings act on the Error value, before integration. Any tuning changes affect only integrated errors occurring after the change.

Apart from output limiting and tuning bumps, there are a number of similar operational modalities that the controller should support, in relation to either manual or automatic operation. For example,

- Cold-start initialization: Sometimes it is desirable as part of the process start-up to start the controller in automatic mode, so that its output has no tendency to move, letting operators move the process on their own schedules. This is also the natural way to initialize secondary controllers when the primary controller goes to automatic mode. Cold start can be implemented if the controller set point is set to the present measurement value and the internal states of the controller are initialized so that the output matches the externally fedback or operator-set value of the output. Automatic control then continues naturally.
- Bumpless transfer: The purpose of this mode is to transfer control to automatic control in such a way that the process does not receive any immediate valve change, but moves from its prior position smoothly. In this case, the controller setpoint is left in place, but all internal states are initialized to be consistent with the current error and external feedback (and current controller output value).[18]
- Batch preload: In circumstances in which a known set-point change is to be applied (for example, in batch production), the controller may be set up to pick up on the set-point change with a preset initial internal integration value. The purpose is to give the process an initial kick, to get it to the set point in the fastest time. This strategy has a number of elaborations of varying sophistication.
- Ramped set-point change: The controller may be designed to limit the rate at which it responds to set-point changes to minimize the bumps to the process.

[17] A full accommodation of actuator or cascaded control loss would include both external feedback and the flags because some control functions do not lend themselves to the external-feedback solution. For example, some adaptive controllers depend on free response to their output to support meaningful adaptation.

[18] This mode can be applied to all controllers in a cascade, but because the secondary set points will then be matched by their primary controllers to their sensor, the result for them should be the same as if they had been initialized under a cold start. This is true only if the controller calculations are carried out in an appropriate order relative to each other. In lieu of this, it may be better to use the cold start on the secondaries anyway.

The variations on cold start and bumpless transfer depend on backcalculation: the recomputation of internal data[19] to be consistent with the unchanged output and external feedback. For example, a PID computation developed below computes its output O from the collective effect X of the proportional and the derivative effects and an internal bias B. The bias is computed in turn by application of a lag computation to the external feedback O_{FB}:

$$O(t) = X(t) + B(t)$$

$$B(t) = \frac{O_{FB}(t) \cdot \Delta t + B(t - \Delta t) \cdot \tau}{\tau + \Delta t}$$

We can accomplish the bumpless transfer by computing the new value of X and then rearranging the first equation above to backcalculate the bias from the new X and the old B:

$$B(t) = O(t) - X(t)$$

When the O is later calculated from this B, its value will remain initially at its old value, irrespective of manual changes in O, measurement, set point, or tuning changes reflected in X. The second equation might be used to backcalculate O_{FB}, but this value will be overridden by later computations anyway.

OUTPUT LIMITING: EXTERNAL FEEDBACK

As already indicated, external feedback represents a precise and smooth way of handling windup. However, implemented in a PID controller algorithm, it has the effect of including, in the integrated term, the difference between the controller output and the corresponding externally fedback measured state. The difference is added to oppose the normal integration. Thus, whenever the external feedback fails to follow the controller output, the difference builds up to stop the integration.

Feedback controllers are built about the processing of an error signal on their input. External feedback extends the principal to the output. It allows the control algorithm to be designed to alter its approach in the face of the output's failure to act. This same strategy can be generalized to apply to any control function:

- Blending: When several ingredient flows are ratioed to generate a blended product, two basic product properties are involved, the product quality and the product flow. Under the standard strategy of pacing, if one ingredient flow limits, the remaining flows are ratioed, not off their original target, but off the external feedback from the limited flow. Such a system extends the external-feedback concept to give up the control of product flow in favor of the more important product quality.
- Fuel/air ratio control: With certain liquid fuels, an excess accumulation of fuel in the burners constitutes a fire hazard. The fuel controller is designed to limit the fuel to be less than the combustible ratio to the measured air flow, at the same time limiting the air flow to be greater than the combustible ratio to the measured fuel flow. If either limits, the other is held to a safe value. Such a system extends the external-feedback concept to safety control.
- Multiple output control: In certain cases a manipulated resource may be duplicated so that several devices share a load. It is desirable that operating personnel be able to take one or more of such devices out of service in such a way that the others take up the load. In this case, a multiple output controller computes an output value such that when the value is ratioed as the set point to each device, the sum of the external feedbacks from all devices equals the net desired load. In this way, as one device is taken over in manual, the others will take up the slack.

[19] Particularly integration data. These guarantee that the integration will resume as if the controller had always been operating under the current error and output conditions.

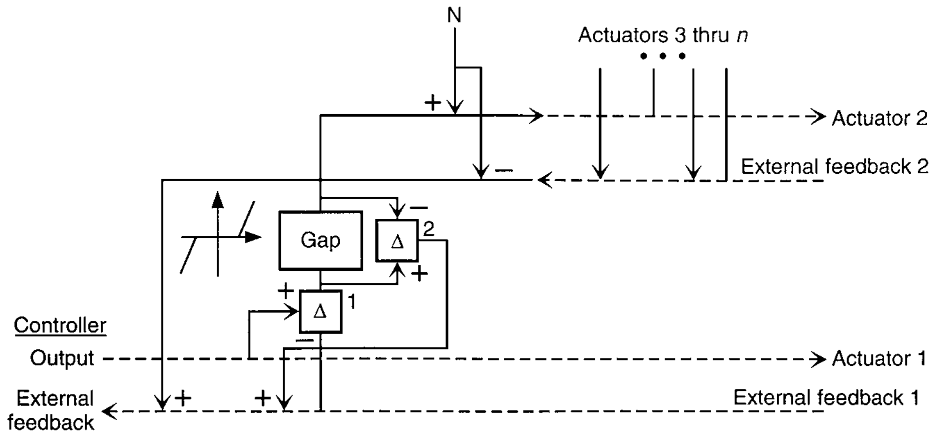


FIGURE 11 External-feedback-based backup.

- Backup (also split-range control): The normal external-feedback antiwindup action can be extended, by the recognized output/external-feedback error, to drive other actuators. Figure 11 illustrates a refined handling of this function.

  In this design, the controller output is normally connected to actuator 1. But if the calculated difference [the subtraction element (Δ) marked 1] between actuator 1 and its external feedback exceeds the gap parameter, the excess difference will serve to transfer the active control to actuator 2. In this way the controller can be backed up against any downstream failures or overrides. The nominal bias *N* reflects the preferred inactive value for actuator 2. (In more refined designs, *N* may be established dynamically by a separate controller.) Differences can further be passed on indefinitely to actuators 3 through *n*. Effectively this arrangement generalizes the behavior of split-range control, taking into account any downstream loss of control action.

  External feedback 2 is added into the controller external feedback to allow the controller to continue its full (integrating) control action, whatever actuators are in fact acting. The purpose of the gap is to ensure that actuator transfer does not give rise to chattering between actuators but acts only for significant loss of actuator control. However, to guarantee that the resulting temporary loss of control does not cause the controller to stop integrating, the actual amount of gapping action (the Δ marked 2) is added into the controller external feedback as well; the external feedback sees neither Gap nor actuator transfer.

  In actuality, the different actuators might call for different control dynamics and different compensation. This could be built into the control transferring paths. However, digital implementation allows the switching of controller tunings, as a function of the active actuator, to be carried out as part of the controller computation, a more natural arrangement. Digital implementation also allows the above structure to be black boxed flexibly, taking the confusing details out of the hands of the user.

- Linear programs and optimization: It has been argued that external feedback is incapable of dealing with connection to higher-level supervisory functions such as linear programs or optimizers. This position reflects higher-level functions not designed for operations, rather than any inherent problem with external feedback. The operationally correct optimizer will benefit from external-feedback data like any other control computation. In this case, each optimization target, with its external feedback, is associated with an implied output constraint. Whenever a difference develops between the two, the constraint limit and the violation become apparent.

  Thus the external-feedback value should be fed into the optimizer, parameterizing a corresponding optimization constraint. There are three special considerations:

- The control action actually implemented must push beyond the constraint so recognized. Otherwise, the constraint becomes a self-fulfilling prophecy that, once established, never gets retracted. Since the control action is presumed to be up against a real process constraint, it does not matter how much further into the constraint the target variable is set. However, it is probably better to exceed the constraint by some small number (e.g., 1%–5% of scale).
- The optimization computation is likely to be run infrequently compared with the normal regulatory dynamics. For this reason some lag filtering or averaging should be built into all external-feedback paths to minimize noise effects and increase their meaningful information content. The filter time constant should correspond to the optimization repetition interval.
- All of this assumes that the optimizer addresses the economic constraint dimensions only. Significant safety or quality constraint effects must always be separately addressed at the regulatory level.

With other control and operational nonlinearities, many issues come up, calling for many different kinds of thinking. Of course, these same differences must fit nonintrusively and naturally with the intentions and expectations of operating people. While operational users will normally not be aware of the technology behind these techniques, they will become intuitively aware of any inconsistencies between the handling of similar functions in different control elements. External feedback provides a powerful strategy for addressing many of these problems whose application uniformity the end user will appreciate.

EXTERNAL FEEDBACK IN NONLINEAR COMPENSATORS

In the introductory discussion of Fig. 2(*c*), the third form of compensation called for inversion of the associated nonlinear compensation. Traditionally this has been done with simple, analytically determined inverses. For example,

- A squaring of the direct compensation called for a square root to the external feedback; an exponential, a log.
- In feedforwards a subtraction of the feedforward signal inverted an addition; a division inverted a multiplication.

Historically, general inverting rules (like Newton's method) would not have been used for fear of failure to converge. Figure 12 shows a situation (the left graph) in which Newton's method, based on extrapolation of the derivative of the function, could successively overshoot the solution of the function (the intersection of the function graph and the shaded horizontal line) on each side. But if the interpolation is always based on prior guesses that bracket the solution (as on the right graph),

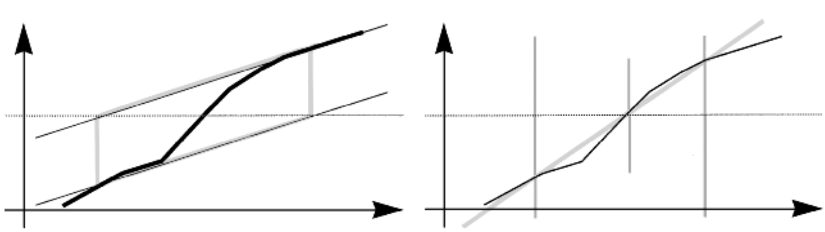


FIGURE 12 Newton's method vs. secant variant.

convergence is guaranteed. These kinds of methods can be generalized to invert multivariable functions, and find all solutions [12].

BASIC CONTROL ALGORITHMS

The Lag Calculation

The lag calculation corresponds to the following continuous transfer function:

$$\frac{L(s)}{I(s)} = \frac{1}{(\tau s + 1)}$$

where L is the output, I is the input, and τ is the lag time constant. The usual practice of going to the z transform for the corresponding sampled-data form should not be overemphasized. No exact approximation is possible. Instead, the algorithms are designed primarily to avoid operationally unnatural behavior. With this in mind, the best direct-sampled-data approximation of the above differential equation is

$$I(t) = \frac{\tau \Delta L}{\Delta t} + L(t) = \frac{\tau}{\Delta t}[L(t) - L(t - \Delta t)] + L(t)$$

or

$$L(t) = \frac{I(t) + \dfrac{\tau}{\Delta t} L(t - \Delta t)}{1 + \dfrac{\tau}{\Delta t}} = \frac{I(t)\Delta t + \tau L(t - \Delta t)}{\tau + \Delta t}$$

This approximation amounts to a weighted average of the new input and the old output. From the scaling point of view, each of the products has the same range as the sum; scaling is simple. The calculation is stable for all positive τ, accurate for large τ, and qualitatively natural as τ approaches 0, with $L(t)$ equaling $I(t)$ if $\tau = 0$, as intended.

The calculation is usable even in single precision if the term $I \Delta t$ is truncated up in magnitude (and τ is not too large). In this case, the product is never truncated to zero unless the product is truly zero. This guarantees that the output will always settle out at any steady-state input value. Normal truncation would leave the result below its theoretical steady-state value, a situation similar to the integral offset described above.

But a trick, similar to the one used with the integrators before, can be applied to calculating lags exactly:

$$L(t) = \text{quotient}\left[\frac{I(t)\Delta t + \tau L(t - \Delta t) + \text{remainder}}{\tau + \Delta t}\right]$$

with the remainder being saved for use in the next sampled calculation.

Lead/Lag Calculation

Filtered derivative and lead/lag calculations are most easily and reliably developed from the above lag calculation by analogy with transfer function calculations:

$$\frac{O(s)}{I(s)} = \frac{\tau \cdot s}{k \cdot \tau \cdot s + 1} = \frac{1}{k}\left[1 - \frac{1}{(k \cdot \tau) s + 1}\right]$$

Similarly,

$$\frac{O(s)}{I(s)} = \frac{\tau \cdot s + 1}{(k \cdot \tau)\, s + 1} = \frac{1}{k}\left[1 - \frac{1}{(k \cdot \tau)\, s + 1}\right] + \frac{1}{(k \cdot \tau)\, s + 1}$$

$$= \frac{1}{k} \times \left[1 + \frac{k - 1}{(k \cdot \tau)\, s + 1}\right]$$

or

$$\frac{O(s)}{I(s)} = \frac{\sigma \cdot s + 1}{\tau \cdot s + 1} = \frac{1}{\tau}\left(\sigma + \frac{\tau - \sigma}{\tau \cdot s + 1}\right)$$

The translation to digital form consists of carrying out all of the algebraic steps directly and replacing the lag transfer function with the digital lag algorithmic calculation described in the preceding section. Considering the last form and proceeding in reverse order, we would calculate the output of a lead/lag from the output of the lag calculation (with time constant τ):

$$O(t) = \frac{[\sigma I(t) + (\tau - \sigma) L(t)]}{\tau}$$

A basic filtered derivative can be calculated with a lag calculation, now indicated as L_D, and assuming, typically, $k = 0.1$ (In fixed point, k would be a power of 2:$1/8$ or $1/16$.) and the lag time constant 0.1 τ_D:

$$D(t) = \frac{1}{k}[I(t) - L_D(t)] = 10[I(t) - L_D(t)]$$

PID Controller Calculation

PID controller designs are expressed in many forms:

$$O(s) = \left(\tau_D s + 1 + \frac{1}{\tau_I s}\right) \frac{100}{PB} \cdot \text{Error}$$

$$O(s) = (\tau_D s + 1)(\tau_{DI} s + 1) \frac{100}{PBs} \cdot \text{Error}$$

$$O(s) = (\tau_D s + 1)\left(1 + \frac{1}{\tau_I s}\right) \frac{100}{PB} \cdot \text{Error}$$

Each of these forms is capable of the same performance as the others, with one exception: the first form is capable of providing complex zeros. There is no generally argued requirement for complex zeros, but this is nonetheless a real distinction.

There is also some disagreement as to whether the $1/\tau$ terms should be replaced with gains (or whether the proportional band terms should be combined into independent proportional, integral, or derivative terms. The reason for giving all terms as gains is that this then places the most stable setting for all terms at zero (not entirely true of a derivative). This argument is pitched to operators. The reason for leaving the terms as above is that the time constants have process-related meaning for engineers who understand the control issues; the separate proportional band then becomes a single stabilizing setting for all terms.

Different implementations also apply different parts of the algorithm differently to the set-point and measurement terms within the error. This reflects that these terms have different effects within the process. Ideally one would provide separate tunings for load and set-point changes. A practical compromise is to apply all three actions to the measurement, but only the proportional and integrating action to the set point. The controller will then be tuned for load disturbances.

The above forms have a particular difficulty if the integrating calculation term is interpreted literally: The integrating term is most naturally translated digitally as

$$\sum_{i=0}^{[t/(\Delta t)]} \frac{\Delta t}{\tau} \cdot \mathrm{Error}\,(i\,\Delta t)$$

However, the individual summed terms get unnaturally large when τ approaches zero. A practical way of bounding the value is to replace τ with $\tau + \Delta t$. When τ approaches zero, this still leaves an unnaturally large term for $i = t/\Delta t$, in competition with the proportional term. The solution is to replace $(t/\Delta t)$ with $(t/\Delta t) - 1$:

$$\sum_{i=0}^{[t/(\Delta t)]-1} \frac{\Delta t}{\tau + \Delta t} \cdot \mathrm{Error}\,(i\,\Delta t)$$

[The summation (integration) can be carried out according to the above discussion.] The result can be justified in another way: Consider the last PID form introduced above:

$$O(s) = (\tau_D s + 1)\left(1 + \frac{1}{\tau_I s}\right)\frac{100}{PB} \cdot \mathrm{Error}$$

The differentiation must be carried out lagged or filtered, as described above. If all but the integrating calculations are calculated as a combined result X (taking into account any separation of the treatment of set point and measurement), the result is:

$$O(s) = \left(1 + \frac{1}{\tau_I s}\right) X(s) = X(s) + \frac{X(s)}{\tau_I s}$$

This has an alternative formulation, which introduces a calculated bias B, particularly convenient for implementing the external feedback:

$$O(s) = X(s) + B(s)$$

$$B(s) = \frac{O_{FB}(s)}{\tau s + 1}$$

where O_{FB} is the external-feedback term, nominally equal to O. When this pair of transfer functions is translated to an algorithm, they become:

$$O(t) = X(t) + B(t)$$

$$B(t) = \mathrm{quotient}\left[\frac{O_{FB}(t) \cdot \Delta t + B(t - \Delta t) \cdot \tau + \mathrm{remainder}}{\tau + \Delta t}\right]$$

The collective effect of this calculation corresponds to the algorithmic expression of the more direct PID form with the modified integration proposed above:

$$O(t) = X(t) + \sum_{i=0}^{[t/(\Delta t)]-1} \frac{\Delta t}{\tau + \Delta t} X(i\,\Delta t)$$

As indicated above, external feedback must be added to this form by subtraction of the difference between output and external feedback from the $X(i\,\Delta t)$ term before integration.

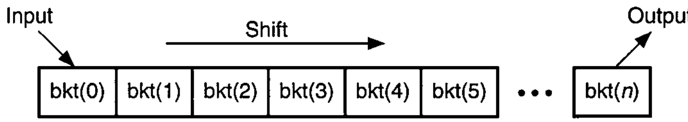


FIGURE 13 Bucket brigade.

Dead-Time Calculation

The dead-time calculation corresponds to the following continuous transfer function:

$$e^{-Ts}$$

Dead time represents a black box whose output exactly repeats the time form of its input, but is delayed by some amount in time. It represents the behavior of the state variables of product, input into a pipe or onto a conveyor belt, and output (delayed) at the other end. It is an essential modeling element for typical processes.

Dead time is conventionally approximated in two ways: through Padé (continued fraction) approximations of the above transfer function and through what is called a bucket brigade. The bucket brigade is implemented in an array of data cells, with the input entered at one end and then shifted down the array, one element per sample time until it comes out at the end, n sample times later, as in Fig. 13.

On the face of it, the Padé is capable of modeling a continuous range of dead times, whereas the bucket brigade is capable of representing only integral delays, for instance, by varying n. Either mechanism can represent a fixed dead time. A further problem arises if the goal is to represent changing dead times. In this case neither the Padé nor the bucket brigade with varying n really reflects the physical or the theoretical behavior of the process.

However, the bucket brigade more closely models the state behavior of product in a delay element; the internal bucket states do represent the internal propagation. The states in the Padé are unrelated to the internal product propagation. It turns out that the continuous Padé will model a changing dead time if the internal parameters are appropriately changed. However, mapping this into a discrete version, taking into account all of the earlier considerations, will be quite difficult.

Thus if the bucket brigade delay can be changed by speeding it up or slowing it down rather than by varying n (the output bucket), it can model changing delay times resulting from changes in flow rates. There are still design questions:

- How does one smoothly achieve dead times smaller than $n\Delta t$ (the number of buckets multiplied by the sample time)?
- How does one smooth the data between stored buckets [between shifts, as shown in Fig. 14(*a*)[20]]?
- How can one modify the discrete dead time (modeled by an integral number of sample time shifts) to represent a continuous range of (changing) dead times?

The first question is the easiest to answer: shift more than once at a time.[21] The second question is almost as easy to answer: On the input side, average the sampled inputs between shifts; on the output side interpolate between the last two buckets to smooth the effect of the shift [see Figs.14(*a*)[22] and 14(*b*)[23]]. The effect of all this is to create a dead time approximation whose effective dead time T corresponds to $(n + 0.5)$ shift times.[24]

[20] Figure 14(*a*) shows the process record (the solid curve), also as sampled (the thin, vertical dashed lines), as effectively sampled by the delay under infrequent shifts (the thicker vertical dashed lines), and as then sampled and held (the more deeply shaded histogram.)

[21] Also average the output values coming from such a multiple shift.

[22] The more lightly shaded histogram shows the effect of averaging the shifted values. The dashed record reconstruction shows the effect of the interpolation. Note the half shift-time delay due to the averaging and the interpolation.

[23] Figure 14(*b*) shows the modified bucket brigade with averaging at the input and the interpolation at the output.

[24] A more refined approximation would view it as a dead time of n shifts with a lag time of the shift time.

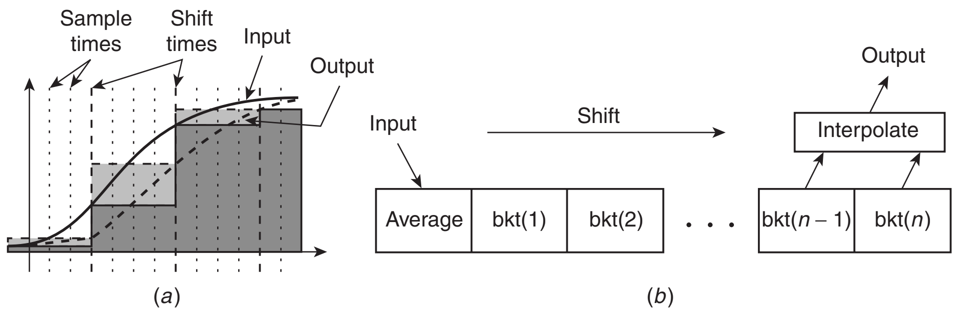


FIGURE 14 Bucket brigade for slow shifts.

The last question is the more subtle one and is answered by use of an irregular shifting frequency, whose immediate average value is equal to the current desired shifting frequency, corresponding to the desired (fractional) dead time. This solution is somewhat similar to the method used to achieve fine color variations on a CRT display that supports only a few basic colors: Mix a number of different pixels irregularly for the intended average affect.

The flow chart (Fig. 15) shows the simplest way of achieving the desired irregular shift time. An accumulator variable (**Acc.,** initially zero) is incremented by $(n + 0.5)/T$ (corresponding to the desired fractional number of shifts per sample time[25]). If the accumulator variable has been incremented to 1 (indicating a net requirement for one full shift), a bucket brigade shift takes place and the accumulator is decremented by 1. Shifting and decrementing are repeated until the accumulator value drops back below 1.[26]

The result is a compensator algorithm capable of modeling dead time in fully time-varying situations.

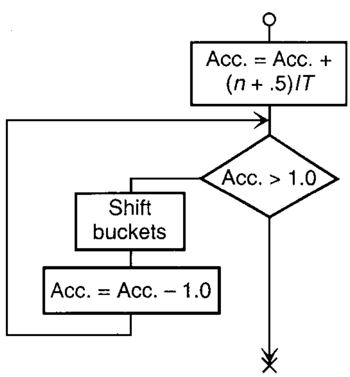


FIGURE 15 Bucket brigade shift calculation flow chart.

Quantization and Saturation Effects

All of the calculations can now be carried out. The discussion has not addressed the effects of truncation on a derivative, but as developed here they are not more serious than for proportional control. In both cases truncation will cause a very small limit cycle, of the order of the minimum quantization of the D/A converter. However, if the derivative is not carefully filtered as part of its computation (as shown in integrated form above) severe problems arise. An approximate unfiltered derivative calculation has the form

$$D(t) = \tau_D \cdot \frac{\Delta\text{Measurement}}{\Delta t} = \frac{\tau_D}{\Delta t} \cdot \Delta\text{Measurement}$$

[25] If $(n + 0.5)\Delta t = T$, then $1/\Delta t = (n + 0.5)/T$.

[26] Actually the shift test value for the accumulator is irrelevant (and would be set to zero in practice) because the continual incrementing and decrementing balance it out to the same frequency, whatever the test point. A further simplification is to increment by $(n + 0.5)$ and decrement by T, further saving the need for floating point or the division.

Quantization has the awkward effect of forcing a minimum nonzero measurement change for any sampled calculation.[27] This value is multiplied by the gain $\tau_D/\Delta t$, which is usually very large. The result is very large pulses on the output of the controller. For ΔMeasurement quantized to 1 part in 1000 (0.1%), $\Delta t = 1$ s, and $\tau_D = 10$ min $= 600$ s, the minimum nonzero derivative is

$$D(t) = 600 \times 0.1 = 60\%$$

The use of filtering smears out the pulses and limits their height (The simple 0.1 time-constant-lagged derivative filter, developed above, limits the maximum pulse height to 10 times the quantization value. More refined algorithms can minimize the problem further by dynamically broadening the effective Δt in the calculation to get a better average derivative. The challenge is to get effective derivative action with a quantization that represents the realistic accuracy bounds of the data.

These derivative problems can be made far worse by internal saturations after integration in the PID algorithm, particularly in incremental algorithms (algorithms that output the desired change in valve position rather than the desired position). Such algorithms involve a second-derivative action. The problem arises because the saturation is likely to be unsymmetrical. When reintegrated later in the algorithm or system, the result is a significant offset. In the above case, the second differencing causes a doublet that extends 60% of scale in both directions. One-sided saturation, reintegrated, would create a 60% valve offset (bump).

IDENTIFICATION AND MATRIX-ORIENTED ISSUES

Theory-motivated control thinking emphasizes matrix-oriented formulations. These are becoming more common as engineers are trained in them. Properly understood, traditional methods are capable of equally good control, but there are aspects of normal control algorithm design for which these newer methods may be more appropriate. Adaptive control often suggests such methods. Space permits only an introduction to the problem but references [13–17] cover many important considerations.

A typical equation of this class defines least-squares data fitting of overdetermined parameters, as used in adaptation:

$$A^T A x = A^T y$$

In this equation, y is a data vector, x is a vector of parameters, and A is an $n \times m$ matrix ($n \geq m$) of data vectors, occurring in a set of equations of the following form:

$$a_{i1}x_1 + a_{i2}x_2 + \cdots = y$$

The problem is to find the best fit for the parameters x, given the known A and y, solving the first equation:

$$x = (A^T A)^{-1}(A^T y)$$

Early direct and recursive solutions were unnecessarily sensitive to numerical rounding and truncation errors. The related eigenvalue problems were equally difficult until the problems were understood [13–16].

As above, the problem can be explained as an effect of differences of large numbers and the explosion of the digital data range under multiplication by numbers not close to 1. In matrix computations the concept of being close to 1 is formalized in several ways. Is $|A|$ close to 1? Because matrices

[27] Conventional resistor ladder A/Ds may be in substantial error in the calibration of their lowest-order bits to the extent that a constant slope measurement may appear to wander up and down as converted. This makes these quantization affects several times worse in practice.

involve several directions, the determinant is misleading. Another measure of a matrix is its norm:

$$||A|| = \max_x \frac{\sqrt{x^T A^T A x}}{\sqrt{x^T x}}$$

Underlying the norm concept is the theory of orthogonal matrices and singular values [13]. An orthogonal matrix is one whose inverse equals its transpose: $Q^T Q = I$. (The letter Q will designate an orthogonal matrix.) In essence orthogonal matrices are equal to 1 in every possible way except for the identity:

4. Their determinant is equal to 1, but for sign.

5. They do not change the length of vectors that they multiply.

6. Therefore their norm equals 1.

7. Products of orthogonal matrices are orthogonal.

8. Every element of Q has magnitude ≤ 1.

9. The elements of the RGA [18] or interaction measure of an orthogonal matrix are all between 0 and 1. Thus easy computation corresponds to easy control.

The singular values σ_i of a matrix A are the square roots of the eigenvalues of $A^T A$. More interestingly, every matrix (square or not) obeys the singular-value decomposition theorem [13]. This theorem states that

$$A \equiv Q_1 \Sigma Q_2$$

for Σ, the diagonal matrix of the singular values σ_i, and some orthogonal Q_1, Q_2. Also,

10. The singular values of an orthogonal matrix are all equal to 1.

Under the singular-value decomposition, when A multiplies a vector, the immediate Q_1 or Q_2 twists the vector into an orientation in which each component is multiplied by one of the singular values. Thus the vector most amplified in length by A is the one oriented so that it is multiplied by the largest σ_i. For this reason the norm of A equals its largest σ_i. Also the vector most diminished in length by A is the one oriented so that it is multiplied by the smallest σ_i.

Generally a matrix is hard to compute with (is effectively much larger than 1) if there are significant off-diagonal terms and the ratio of largest σ_i to smallest σ_i is much larger than 1. (This also corresponds to large RGA elements.) Computing $A^T A$, as in the least-squares equation, squares this ratio, making computation that much worse. Effective algorithms minimize such operations.

As a simple example, consider the usual solution of the equation $Ax = b$, with A the matrix shown in Fig. 16. The conventional Gaussian solution[28] involves reducing the matrix to a triangular form

Gaussian (LU):

$$A = \begin{bmatrix} 0.1 & 1 \\ 1 & 1 \end{bmatrix} = LU = \begin{bmatrix} 1 & 0 \\ 10 & 1 \end{bmatrix} \times \begin{bmatrix} 0.1 & 1 \\ 0 & -9 \end{bmatrix}$$

Orthogonal (QR):

$$A = \begin{bmatrix} 0.1 & 1 \\ 1 & 1 \end{bmatrix} = QR$$

$$= \begin{bmatrix} -0.0995 & -0.995 \\ 0.995 & 0.0995 \end{bmatrix} \times \begin{bmatrix} -1.005 & -1.0945 \\ 0 & -0.8955 \end{bmatrix}$$

FIGURE 16 External-feedback-based backup.

[28] The example cheats a little, bypassing any pivot operation. These are imprecise compared with the orthogonal matrix methods.

that can then be backsolved. This is equivalent to factoring that same triangular form out, leaving a second triangular form (L and U in the figure). Note that the result involves large numbers (the 10 and -9).

The same matrix can also be reduced to a triangular (backsolvable) form by multiplication by an orthogonal matrix. The corresponding factorization is called *QR* factorization, as shown. Note that all the calculations will now involve small numbers (0.995 and -1.005). The thrust of the newer matrix methods is to avoid matrix multiplication if possible (it expands the data range) and to try to restrict any multiplications to those involving orthogonal, diagonal, or triangular matrices.

SOFTWARE AND APPLICATION ISSUES

There was a brief reference above to software implications of control saturation. No modern discussion of control would be complete without observing the critical role of software [19–25]. Matrix-oriented control has been based on FORTRAN and purely mathematical approaches to control thinking. This has separated it from the major operational concerns of the industry. A control system should not only support the control computation, but the operational access to control data in a framework that is as easy to use as possible. It should include some model of sensible operator intervention.

These considerations are accommodated naturally by the software for traditional control. Standard process regulatory control has been based on blocks, interpreted by the computer to carry out control.

These blocks are blocks in two senses. They represent the digital equivalent of the old analog blocks in block diagrams. And they consist of data blocks whose data elements correspond to the signals and parameters of the analog block controller or are data pointers that make the connections between the blocks.

But neither the existing systems nor the proposed standards offer attractive solutions, supporting the necessary flexibility and ease of use. The challenge is to provide flexible, future-oriented systems in which the goals and structure of the control system are transparent; new sensors, actuators, and algorithms are easily added; and sufficient standards included so control systems can include elements from many vendors [24, 25].

And the block model is inefficient for the kinds of control that are now possible. For example, the iterated structure suggested in Fig.11 would be extremely inefficient if implemented in blocks, even as it would be quite simple programmed directly. And as normally interpreted, blocks have a predefined number of connections, limiting the use of structures with an indefinite number of inputs or outputs (e.g., feedforwards or overrides).

Several references also show the limitations of the normal block diagram in representing the kinds of controls called for here [20–23, 26, 27]. Despite the universal belief in graphics, control diagrams are hard to understand. An improved notation and graphics separately list the loops making up the design, clearly distinguishing their controlled variables and purposes.

For example, the Fig. 17 combination of two cascaded controllers with a pressure high constraint controller and a feedforward, according to the earlier structuring, would be expressed as

$$\mathbf{T100}_{\substack{\mathbf{REGULATE}\\ \mathbf{FEEDFWD\ F50}}}\ \mathbf{P100}_{\mathbf{HICONSTRN}}\mathbf{F100}_{\mathbf{REGULATE}}\mathbf{V100}$$

with the main degree-of-freedom path broken by the **P100** constraint expression.

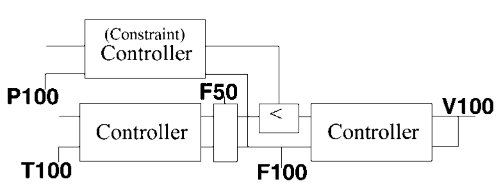


FIGURE 17 Constraint cascade.

Two operational issues are of special concern: smart field devices and interoperability. Modern sensors and actuators now include significant amounts of computational power themselves. The uses of this power will expand indefinitely. But both the field devices and the central controls must be developed in a standard model that permits control and its orderly operation to be supported without special or redundant programming of either.

There is a broader challenge: to solve the control software problem with solutions that capitalize on the computer as a truly intelligent control device, beyond the rigid scientific computations envisioned by the matrix control approaches or the programmable block diagram that mimics the dated analog controls.

SUMMARY

Digital control algorithms can be designed for experimental or single applications with the easiest tools available: FORTRAN, C, BASIC, and floating-point arithmetic. In this case, there is reasonable hope that normal commissioning debugging will weed out all the problems. But if a sense of workmanship prevails or if the algorithm is to be used in many applications, then attention to refinement and foolproofing are necessary:

1. Numerical effects of fixed- and floating-point arithmetic.
2. Documentation, control, and testing of detailed scaling, precision, and saturation within the algorithm.
3. Design for natural tuning and qualitative behavior, predictable from analog intuition.
4. Nasty quantization surprises.
5. Accommodations of windup and other control limiting effects.
6. Bumpless transfer and operational considerations.
7. Software, architectural, configuration considerations.

General-purpose digital programming languages and tools do not remotely address these issues.

REFERENCES

1. Bristol, E. H., “On the Design and Programming of Control Algorithms for DDC Systems,” *Control Engineering*, Jan. 1977.
2. Clagget, E. H., “Keep a Notebook of Digital Control Algorithms,” *Control Engineering*, pp. 81–84, Oct. 1980.
3. Fehervari, W., “Asymmetric Algorithm Tightens Compressor Surge Control,” *Control Engineering*, pp. 63–66 Oct. 1977.
4. Tu, F. C. Y, and J. Y. H. Tsing, “Synthesizing a Digital Algorithm for Optimized Control,” *Instr. Tech*., pp. 52–56, May 1979.
5. Franklin, G. F., and J. D. Powell, *Digital Control of Dynamic Systems*, Addison-Wesley, Reading, Massachusetts, 1980.
6. ISA SP50 User Layer Technical Report.
7. Knuth, D. E., *The Art of Computer Programming*, vol. 2, Addison-Wesley, Reading, Massachusetts, 1981.
8. *Collected Algorithms from ACM*, 1960–1976.
9. ANSI/IEEE Standard 754–1985 for Binary Floating-Point Arithmetic.
10. *Chemical Engineers’ Handbook*, 5th ed., McGraw-Hill, New York, pp. 2-81–2-85, 1973.
11. Considine, D. M., *Process/Industrial Instruments and Controls Handbook*, 4th ed., McGraw-Hill, New York, pp. 2.85–2.86, 1993.
12. Hansen, E., *Global Optimization Using Interval Analysis*, Marcel Dekker, New York, 1992.

13. Strang, G., *Linear Algebra and Its Applications*, Academic, New York, 1976.
14. Bierman, G. J., *Factorization Methods for Discrete Sequential Estimation*, Academic, New York, 1977.
15. Lawson, C. L., and R. L., Hansen, *Solving Least Squares Problems*, Prentice-Hall, Englewood Cliffs, New Jersey, 1974.
16. Laub, A. J., and V. C., Klema, "The Singular Value Decomposition: Its Computation and Some Applications," *IEEE Trans. Autom. Control*, **AC-25**, pp. 164–176, Apr. 1980.
17. MacFarlane, A. G. J., and Y. S. Hung, "A Quasi-Classical Approach to Multivariable Feedback Systems Design," *2nd IFAC Symposium, Computer Aided Design of Multivariable Technological Systems*, Purdue University, West Lafayette, Indiana, pp. 39–48, Sept. 1982.
18. Bristol, E. H., "On a New Measure of Interaction for Multivariable Control," *IEEE-PTGAC*, **AC-11**, pp. 133–134, Jan. 1966.
19. Shinskey, F. G., "An Expert System for the Design of Distillation Controls," presented at Chemical Process Control III, Asilomar, California, Jan. 12–17, 1986.
20. Bristol, E. H., "Strategic Design: A Practical Chapter in a Textbook on Control," presented at 1980 JACC, San Francisco, California, Aug. 1980.
21. Bristol, E. H., "After DDC–Idiomatic (Structured) Control," presented at 88th National Meeting of AIChE, Philadelphia, Pennsylvania, June 1980.
22. Prassinos, A., T. J. McAvoy, and E. H. Bristol, "A Method for the Analysis of Complex Control Schemes," presented at 1982 ACC, Arlington, Virginia, June 1982.
23. Bristol, E. H., "A Methodology and Algebra for Documenting Continuous Process Control Systems," presented at 1983 ACC, San Francisco, California, June 1983.
24. Bristol, E. H., "Super Variable Process Data Definition," ISA *SP50.4* Working Paper, Oct. 24, 1990 and updating.
25. Bristol, E. H., "An Interoperability Level Model: Super-Variable Categories," ISA *SP50.4* Continuing Working Paper.
26. Bristol, E. H. "A Language for Integrated Process Control Application," presented at the Retirement Symposium in Honor of Prof. Ted. Williams, Purdue University, West Lafayette, Indiana, Dec. 5–6, 1994.
27. Bristol, E. H. "Redesigned State Logic for an Easier Way to Use Control Language," World Batch Forum, Toronto, May 13–15, 1996; also *ISA Transactions*, **35**, pp. 245–257, 1996.

SAFETY IN INSTRUMENTATION AND CONTROL SYSTEMS

by E. C. Magison*

Extensive use of electrical and electronic control systems, computers, sensors, and analyzers in process control continues to focus attention on reducing the probability of fire or explosion due to electric instrument failure. At one time explosionproof housing was the common method of providing protection. Attention then turned to other means that provide the same or higher levels of safety, but with less weight and easier accessibility for maintenance and calibration and at equivalent or lower costs.

Because instrument manufacturers serve an international market, increased activity within national jurisdictions is being matched by recognition that standardization must be accomplished at the international level as well.

* Consultant, Ambler, Pennsylvania.

AREA AND MATERIAL CLASSIFICATION

North America

In the United States, Articles 500 to 505 of the National Electrical Code (NEC) provide basic definitions of hazardous areas and the requirements for electrical installations. Articles 500 and 505 defines the classification of hazardous locations broadly in terms of kind and degree of hazard. The kind of hazard is specified by class and group. The degree of hazard is designated by division. Typical industrial locations, for example, may be classified as Class I, Group D, Division 1; or Class II, Group G, Division 2. The principal features of the NEC classification are summarized in Table 1. Many additional materials are listed in National Fire Protection Association publication NFPA 497 and NFPA 499. Similar definitions are given in the Canadian Electrical Code.

International Electrotechnical Commission

Most industrial nations have adopted or are adopting the area and material classification definitions of the International Electrotechnical Commission (IEC). Locations where a flammable concentration may be present are designated Zone 0, Zone 1, or Zone 2.

A Zone 0 location is a location where the atmosphere may be in the explosive range such a high percentage of the time (above 10%) that extraordinary measures must be taken to protect against ignition by electrical apparatus.

Zone 1 locations have a probability of hazard between Zone 2 and Zone 0. A Zone 2 location is similar to North American Division 2. Taken together, Zone 1 and Zone 0 equate North American Division 1.

In Zone 2, requirements analogous to those in North America are accepted in principle in many countries, but in practice Zone 1 types of protection are often used in Zone 2 because there is no accepted standard for Zone 2 apparatus. The advantage of distinguishing between the extraordinary hazards of Zone 0 and the lesser hazards of Zone 1 is that apparatus and installation requirements can be relaxed in Zone 1. For example, intrinsically safe systems in North America are judged on the basis of two faults because of the encompassing definition of Division 1. For use in Zone 1, consideration of only one fault is required, although two faults are assessed in Zone 0.

Material classification in most countries now uses IEC terminology. A Group I hazard is due to methane (firedamp) in the underground works of a mine. The presence of combustible dusts and other environmental aspects of mining works are assumed when preparing apparatus requirements for Group I.

Group II gases and vapors are flammable materials found in industrial aboveground premises. They are divided into Groups IIA, IIB, and IIC, which are similar although not identical to North American Groups D, C, and B, respectively.

Article 505, introduced in the 1996 NEC, defines Class I, Zone 0, Zone 1, and Zone 2 locations. It is the first step toward using the IEC method of area and material classification in the United States. This will allow eventual recognition of methods of protection such as increased safety and encapsulation which have been standardized for use in Zone 1 locations in Europe and in other industrial nations, but which have not been recognized for use in Division 1 in the United States. ISA, the International Society for Measurement and Control, is publishing a series of standards which mirror IEC requirements for types of protection useful in Zones 0, 1, and 2, modified to recognize North American standards and installation practices.

Presumably, when IEC agrees on definitions for zones in locations where combustible dust is the hazard, these will be proposed for addition to Article 505.

Classifying a Hazardous Location

The NEC definitions provide guidelines, but do not give a quantitative method for classifying a specific hazardous location. Factors to consider include the properties and quantity of hazardous material that may be released, the topography of the site, the construction of the plant or building, and

TABLE 1 National Electrical Code Area Classification System

Class I Gases and vapors	Class II Dusts	Class III Flyings
Group A—Acetylene Group B—Hydrogen or gases of similar hazardous nature, such as manufactured gas, butadiene, ethylene oxide, propylene oxide Group C—Ethyl ether, ethylene, cyclopropane, unsymmetrical dimethylhydrazine, acetaldehyde, isoprene Group D—Gasoline, hexane, naphtha, benzene, butane, propane, alcohol, acetone, benzol, lacquer solvent, natural gas, acrylonitrile, ethylene dichloride, propylene, styrene, vinyl acetate, vinyl chloride	Group E—Metal dusts Group F—Carbon black, coal, coke dusts Group G—Grain dust, flour, plastics, sugar	No group assigned. Typical materials are cotton, kapok, nylon, flax, wood chips—normally not in air supension
	Division 1*	
For heavier-than-air vapors, below-grade sumps, pits, etc., in Division 2 locations. Areas around packing glands; areas where flammable liquids are handled or transferred; areas adjacent to kettles, vats, mixers, etc. Where equipment failure releases gas or vapor and damages electrical equipment simultaneously.	Cloud of flammable concentration exists frequently, periodically, or intermittently—as near processing equipment. Any location where conducting dust may accumulate.	Areas where cotton, spanish moss, hemp, etc., are manufactured or processed.
	Division 2*	
Areas adjacent to a Division 1 area. Pits, sumps containing piping, etc., in nonhazardous location. Areas where flammable liquids are stored or processed in completely closed piping or containers. Division 1 areas rendered nonhazardous by forced ventilation.	Failure of processing equipment may release cloud. Deposited dust layer on equipment, floor, or other horizontal surface	Areas where materials are stored or handled.

*In a Division 1 location there is a high probability that a flammable concentration of vapor, gas, or dust is present during normal plant operation, or because or frequent maintenance. In a Division 2 location there is only low probability that the atmosphere is hazardous—for example, because of equipment failure.

Until the 1971 revision, material classification in the NEC differed from the practice in almost all other countries except Canada. Material groupings were based on consideration of three parameters: autoignition temperature (AIT) (or, spontaneous ignition temperature SIT), maximum experimental safe gap (MESG), and the maximum pressure rise in an explosion test chamber. In Europe materials long have been grouped by AIT and separately by MESG. Pressure rise is not a material classification criterion. It is now recognized in the United States that there is no correlation between MESG and AIT. Hydrogen, for example, has a very samll MESG and a very high AIT. Many Group C and D materials have lower ignition temperatures but wider experimental safe gaps. Because the NEC classification was based on two uncorrelated parameters, United States experts could not use the results of experimental work on new material in other countries, or use other classification tables. The 1971 NEC revisions separate AIT from considerations of MESG. Explosionproof housings now can be designed for MESG typical of a group of materials. External surface temperatures shall not exceed the AIT of the hazardous gas or vapor of concern.

the past history of fire and explosion (of a particular location or plant as well as of an entire industry). Although authorities recognize the need, there are no concise rules for deciding whether a location is Division 1 or Division 2. The best guides to area classification known to the author are American Petroleum Institute publications API RP500A, B, and C, and API RP505 for petroleum installations, and NFPA 497 for installations in chemical plants. These documents are applicable to any industry. API RP505 specifies classification of Zones 0, 1, and 2.

Special Cases of Area Classification

It is common practice to pressurize instrument systems to reduce the area classification inside the enclosure. The inside of an instrument enclosure provided with a simple pressurization system, located in a Division 1 area, can be considered a Division 2 location because only by accidental failure of the pressurization system can the internal atmosphere become hazardous. If the pressurization system is designed to deenergize all equipment within the enclosure when pressurization fails, the interior can be considered a nonhazardous location. Two failures are required—(1) of the pressurization system and (2) of the interlock system—before an explosion can occur.

An important limitation of this philosophy is that if any single failure can make the enclosure hazardous, the interior of the enclosure must not be classified less hazardous than Division 2, regardless of the pressurization system design. Such is the case with bourdon-tube or diaphragm-actuated instruments where process fluid is separated from the instrument interior only by a single seal, namely, the bourdon or diaphragm. Unless the pressure is high enough or the enclosure air flow is great enough to prevent a combustible concentration inside the enclosure should the measuring element fail, the interior never should be classified less hazardous than Division 2. Such systems often are referred to as singly sealed systems.

In a doubly sealed system two seals are provided between the process fluid and the area being purged, and a vent to the atmosphere is provided between the seals. Failure of both seals is required to make the enclosure interior hazardous. Even so, pressurization can prevent the hazardous material from entering the compartment because the hazardous material is at atmospheric pressure. Article 501-5(f) of the NEC mandates a double seal wherever failure of a sealing element could force flammable material into the conduit system.

TECHNIQUES USED TO REDUCE EXPLOSION HAZARDS

The predisposing factors to fire or explosion are (1) the presence of a flammable liquid, vapor, gas, dust, or fiber in an ignitable concentration, (2) the presence of a source of ignition, and (3) contact of the source with ignitable material. The most obvious way to eliminate the possibility of ignition is to remove the source to a location where there is no combustible material. This is the first method recognized in the NEC, Article 500. Another method is to apply the principle of intrinsic safety. Equipment and wiring that are intrinsically safe are incapable, under normal or abnormal conditions, of igniting a specifically hazardous atmosphere mixture. For practical purposes there is no source of ignition.

Figure 1 summarizes the techniques used to reduce explosion hazards. Methods based on allowing ignition to occur force combustion under well-controlled conditions so that no significant damage results. A continuous source of ignition, such as the continuous pilot to localize combustion in gas appliances, is commonplace. Explosionproof enclosures contain an explosion so that it does not spread into the surrounding atmosphere. Historically this has been the most common technique. In Zone 2, enclosed break devices in which the enclosed volume is so small that a weak incipient explosion cannot escape the enclosure, are permitted in some countries.

There are several methods for reducing hazard by preventing the accumulation of combustible material in an explosive concentration or for isolating an ignition source from flammable material. Pressurization of instruments is common. Continuous dilution in which the interior of an enclosure is pressurized to exclude flammable material from entering and is also continuously purged to dilute any internal release of flammable material is applicable to analyzers and other devices in which flammable material may be released inside the enclosure. In hydrogen-annealing furnaces and hydrogen-cooled

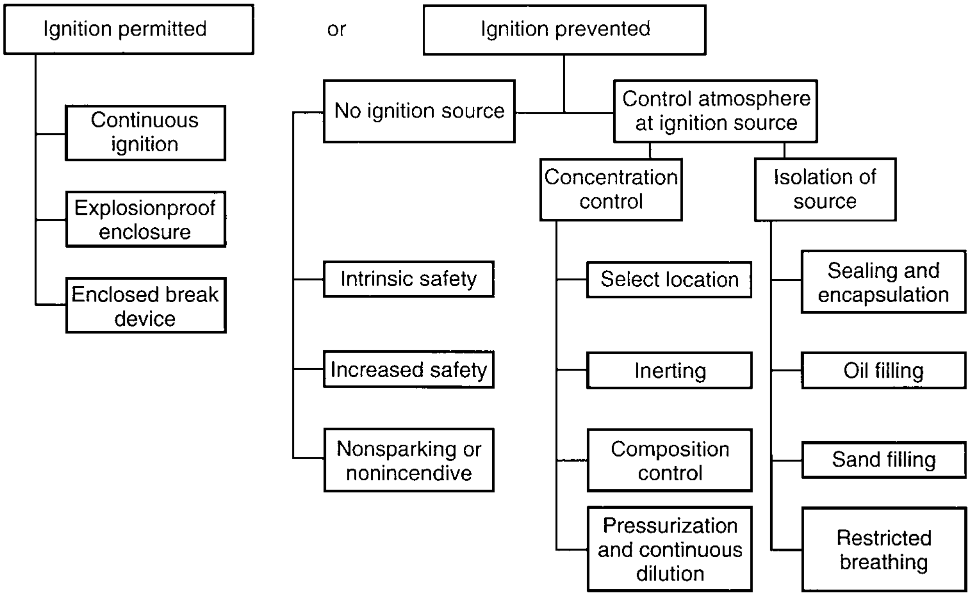


FIGURE 1 Techniques used to reduce explosion hazards.

electric generators the concentration is held above the upper explosive limit. Blanketing of tanks with nitrogen or carbon dioxide (CO_2) and rock dusting of coal mine galleries and shafts are examples of using inert materials to suppress a combustible mixture.

Several techniques are used to isolate the ignition source. Oil immersion prevents contact between the atmosphere and the ignition source. In Europe sand-filled equipment sometimes is used. Sealing and encapsulation both provide a barrier to impede contact.

Increased safety is a technique used for transformers, motors, cables, and so on, which are constructed with special attention to ruggedness, insulation, reliability, and protection against overtemperatures so that an ignition-capable failure is of very low probability. Increased safety, developed first in Germany and now accepted widely in Europe, may be used in Zone 1.

Nonsparking apparatus and nonincendive apparatus are suitable for use in Division 2 and Zone 2 because they have no normal source of ignition.

Restricted breathing is the technique of using a tight, but not sealed enclosure in Zone 2 to allow only slow access of flammable vapors and gases to the source of ignition. This technique, developed in Switzerland, is slowly achieving acceptance in Europe.

Explosionproof Housings

Termed flameproof enclosures in international English, explosionproof housings remain the most practical protection method for motor starters and other heavy equipment that produces sufficient energy in normal operation to ignite a flammable atmosphere. Explosionproof enclosures are not assumed to be vaportight; it is assumed that a flammable atmosphere will enter the enclosure.

A pressure rise of 100 to 150 lb/in^2 (690 to 1034 kPa) is typical for the mixture producing the highest explosion pressure. In small enclosures, loss of energy to the enclosure walls decreases the pressure rise. Because the enclosure must contain the explosion and also must cool escaping gases, cast or heavy metallic construction with wide, close-fitting flanges or threaded joints is typical. Nonmetallic construction is permitted.

For specific design criteria in the United States, reference should be made to the standards of the intended certifying agency. Although requirements of all agencies are similar, there are many differences in detail. In general, in addition to tests to ensure that an internal explosion is not transmitted to the outside, the enclosure must withstand a hydrostatic pressure of four times the maximum pressure observed during the explosion test and must not have an external case temperature high enough to ignite the surrounding atmosphere. In Canada the applicable standard is CSA C22.2, No. 30.

In Europe the applicable standards are CENELEC EN50014, "General Requirements," and CENELEC EN50018, "Flameproof Enclosure 'd'." These are available in English as British Standard BS 5501, Parts 1 and 5.

Requirements for flange gaps are less restrictive than North American standards. Routine testing of enclosures at lower pressures than those of the North American test is common. Type testing is achieving recognition. In North America, wider permissible flange gaps and routine testing are gaining acceptance.

Encapsulation, Sealing, and Immersion

These techniques seldom are applied to a complete instrument. They serve to reduce the hazard classification of the instrument by protecting sparking components or subassemblies. Oil immersion and sand filling are applied to power-handling apparatus, but neither technique has important applications in instrument systems, although oil immersion may be a convenient technique for some hydraulic control elements, and sand filling has been used in some portable devices.

Sealing

Article 501-3(b)(2) of the NEC states that general-purpose enclosures may be used in Division 2 locations if make-and-break contacts are sealed hermetically against the entrance of gases or vapors. The NEC provides no definition of a satisfactory hermetic seal, however. Seals obtained by fusion, welding, or soldering and, in some instances, plastic encapsulation are widely accepted. In reality, the leak rate of soldered or welded seals is lower than that required for protection against explosion in a Division 2 location.

The long-time average concentration inside a sealed enclosure approaches the average concentration outside. The function of a seal is to prevent transient excursions above the lower explosive limit (LEL) outside the device from raising the concentration inside the device to the LEL. Three mechanisms can force material through a seal, (1) changes in ambient temperature, (2) changes in barometric pressure—both effects tending to make the seal breathe, and (3) wind and strong air currents. The last named mechanism usually can be ignored because a sealed device must be installed in a general-purpose enclosure to protect it from such conditions.

Encapsulation involves enclosing a component or subassembly in a plastic material, a tar, or a grease, with or without the additional support of a can. If the encapsulated assembly is robust and has mechanical strength and chemical resistance adequate for the environment in which it is used, it can be considered the equivalent of a hermetic seal. An external hazardous atmosphere must diffuse through a long path between the encapsulating material and the device leads to reach the interior. Standards for sealed devices can be found in Instrument Society of America Standard ISA S12.12.

Pressurization Systems

Lowering the hazard classification of a location by providing positive-pressure ventilation from a source of clean air has long been recognized in the NEC, and instrument users have pressurized control room and instrument housings for many years. The first detailed standard for instrument purging (pressurizing) installations was ISA SP12.4 (now withdrawn). These requirements in essentially the same form make up the first section of NFPA 496, which also covers purging of large enclosures, ventilation of control rooms, Class II hazards, and continuous dilution. NFPA 496 defines three types of pressurized installation:

Type Z. Pressurization to reduce the classification within an enclosure from Division 2 to nonhazardous

Type Y. Pressurization to reduce the classification in an enclosure from Division 1 to Division 2
Type X. Pressurization to reduce the classification within an enclosure from Division 1 to nonhazardous

Type Z Pressurization. This permits the installation of ignition-capable equipment inside the enclosure. For an explosion to occur, the pressurized system must fail, and also, because the surrounding area is Division 2, there must be a process equipment failure which releases flammable material. Thus there must be two independent failures, and no additional safeguards in the pressurization system are necessary.

Figure 2 shows a typical installation for Type Z pressurization. Only a pressurization indicator is required. The probability that a process fault will make the location hazardous before any failure of

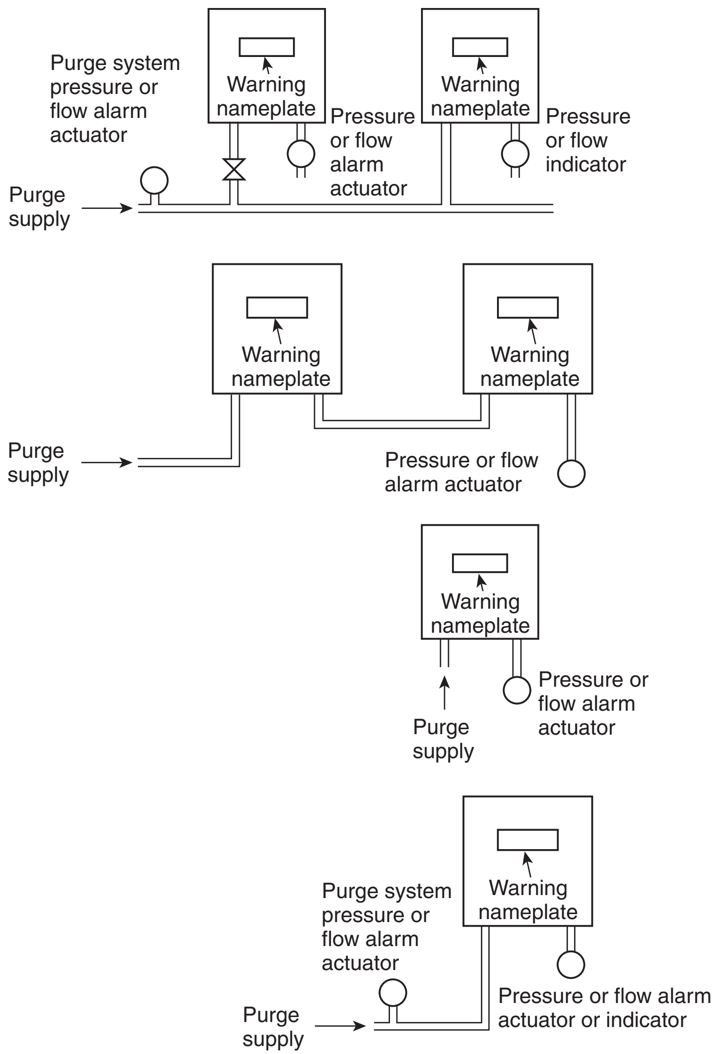


FIGURE 2 Typical installation for Type Y and Type Z pressurization.

the pressurization system is recognized and corrected is assumed to be extremely low. An electric indicator or alarm must meet the requirements of its location before pressurization. If a pressure indicator is used, no valve may be installed between the pressure indicator and the case. Any restriction between the case and the pressure device must be no smaller than the smallest restriction on the supply side of the pressure device. The case must be opened only if the area is known to be nonhazardous or if power has been removed. In normal operation the pressurization system must maintain a minimum pressure of 0.1 inch (2.5 mm) of water gage. The flow required to maintain this pressure is immaterial. The temperature of the pressurized enclosure must not exceed 80% of the ignition temperature of the gas or vapor involved when it is operated at 125% of rated voltage. A red warning nameplate must be placed on the instrument to be visible before the case is opened. Failure of the pressurization system must be alarmed.

Type Y Pressurization. In this case all requirements of Type Z pressurization must be met. Equipment inside the enclosure must be suitable for Division 2, that is, it is not an ignition source in normal operation. For an explosion to occur, the pressurization must fail, and the equipment inside must fail in a way to make it an ignition source.

Type X Pressurization. In this system the pressurization is the only safeguard. The atmosphere surrounding the enclosure is presumed to be frequently flammable. The equipment within the enclosure is ignition-capable. The pressurization system failure must automatically deenergize internal equipment. All requirements for Type Z and Type Y pressurization must be met. The interlock switch may be pressure- or flow-actuated. The switch must be suitable for Division 1 locations, even if it is mounted within the instrument case, because it may be energized before purging has removed all flammable material. A door that can be opened with the use of a tool must be provided with an automatic disconnect switch suitable for Division 1. A timing device must prevent power from being applied until four enclosure volumes of purge gas can pass through the instrument case with an internal pressure of 0.1 inch (2.5 mm) of water gage. The timing device also must meet Division 1 requirements, even if inside the case (Fig. 3).

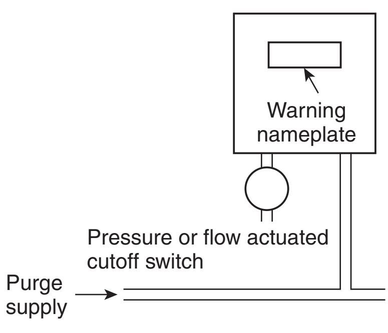


FIGURE 3 Typical installation for Type X pressurization.

IEC and CENELEC standards for pressurization systems are similar to those of NFPA, although the requirements are not phrased in terms of reduction in area classification. The minimum pressure is 0.2 inch (5.1 mm) of water gage.

The IEC and NFPA standards also cover continuous dilution. The hardware is similar to that required for pressurization, but the rationale for selecting the level of protection needed is based on the presence of a source of flammable material within the enclosure, as in an analyzer. The objectives are to prevent entry of an external flammable atmosphere (pressurization) and also to dilute any internal release to a low percentage of the lower flammable limit (continuous dilution).

INTRINSIC SAFETY

Experiment and theory show that a critical amount of energy must be injected into a combustible mixture to cause an explosion. If the energy provided is not greater than the critical ignition energy, some material will burn, but the flame will not propagate. An explosion occurs only when enough energy is injected into the mixture to ignite a critical minimum volume of material. The diameter of a sphere enclosing this minimum volume is called the quenching distance or quenching diameter. It is related to the maximum experimental safe gap (MESG), but is about twice as large. If the incipient flame sphere does not reach this diameter, it will not propagate.

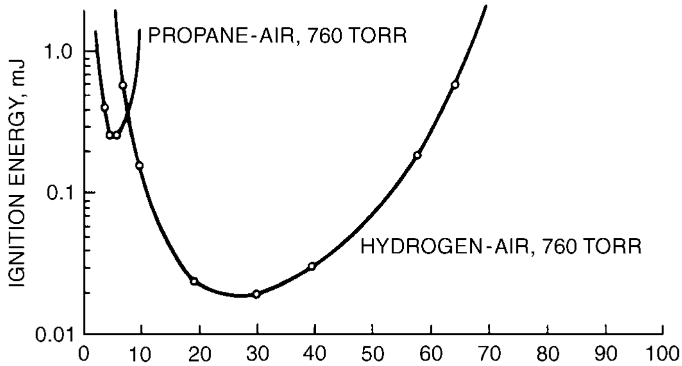


FIGURE 4 Effect of Concentration on ignition energy.

The energy required for ignition depends on the concentration of the combustible mixture. There is a concentration at which the ignition energy is minimum. The curve of ignition versus concentration is asymptotic to limits of concentration commonly called the lower explosive limit (LEL) and the upper explosive limit (UEL). Figure 4 illustrates the influence of concentration on the critical energy required to cause ignition. A hydrogen-air mixture, one of the most easily ignited atmospheric mixtures, supports combustion over a wide range of concentrations. A propane-air mixture, which is typical of many common hazardous materials, is flammable only over a narrow range of concentrations. The amount of energy required to ignite the most easily ignited concentration of a mixture under ideal conditions is the minimum ignition energy (MIE).

Definition

The NEC defines an intrinsically safe circuit as one in which any spark or thermal effect is incapable of causing ignition of a mixture of flammable or combustible material in air under prescribed test conditions, for example, those in ANS/UL913.

Early Developments

The British first applied intrinsic safety in direct-current signaling circuits, the first studies beginning about 1913. In 1936 the first certificate for intrinsically safe equipment for other than mining was issued. By the mid-1950s certification in Great Britain for industrial applications was common. At the U.S. Bureau of Mines work on intrinsically safe apparatus, although the term was not used, commenced about the same time as the British investigations. Rules for telephone and signaling devices were published in 1938.

During the 1950s increased use of electric equipment in hazardous locations stirred worldwide interest in intrinsic safety, and by the late 1960s almost every industrial country had either published a standard for intrinsically safe systems or drafted one. The major industrial countries also were active in the IEC Committee SC31G, which prepared an international standard for intrinsically safe systems.

The first standards for intrinsically safe equipment intended for use by the instrument industry were published as ISA RP12.2, issued in 1965. The NFPA used ISA RP12.2 as a basis for the 1967 edition of NFPA 493.

During the years following the publication of ISA RP12.2 and NFPA 493-1967 the certification of intrinsically safe systems by independent approval agencies, such as Factory Mutual and Underwriters Laboratories in the United States, CSA in Canada, BASEEFA in the United Kingdom, and PTB in Germany, became a legal or marketing necessity in most countries. All standards for intrinsic safety have therefore become much more detailed and definitive. Adherence to the standard is the objective, not a judgment of safety. The work of the IEC has served as the basis for later editions of NFPA 493 and for UL 913, which is now the U.S. standard for intrinsically safe equipment, as well as for Canadian Standard CSA C22.2-157 and CENELEC Standard EN50020. Any product marketed internationally must meet all these standards.

All the standards agree in principle, but differ in the details.

CSA and U.S. standards are based on safety after two faults, because in these cases Division 1 includes Zone 1 and Zone 0. European standards provide *ia* and *ib* levels of intrinsic safety for Zone 0 and Zone 1 application, based on consideration of two faults and one fault, respectively.

Standards for intrinsic safety can be less intimidating to the user if it is appreciated that most construction details are efforts to describe what can be considered a fault or what construction can be considered so reliable that the fault will never occur. When viewed in this light, creepage and clearance tables, transformer tests, and tests of protective components make much more sense. They are guidelines for making design decisions—not mandated values for design. They apply only if safety is affected.

Design of Intrinsically Safe Systems

The objective of any intrinsically safe design, whether produced by an equipment manufacturer or by a user attempting to assemble a safe system from commercially available devices, is the same—to ensure that the portion of system in the Division 1 location is unable to release sufficient energy to cause ignition, either by thermal or by electrical means, even after failures in system components. It is not necessary that the associated apparatus, that is, the apparatus located in Division 2 or a nonhazardous location connected to the intrinsically safe circuit, be itself intrinsically safe. It is only necessary that failures, in accordance with the accepted standard for intrinsic safety, do not raise the level of energy in the Division 1 location above the safe level.

BASIC TECHNIQUES USED BY MANUFACTURERS

Techniques used by manufacturers in the design of intrinsically safe apparatus are relatively few in number, and all manufacturers use the same fundamental techniques.

Mechanical and Electrical Isolation

The most important and most useful technique is mechanical isolation to prevent intrinsically safe circuits and nonintrinsically safe circuits from coming in contact. Often mechanical isolation is achieved solely by appropriate spacing between the intrinsically safe and nonintrinsically safe circuits. In other cases, especially at field connections or in marshaling panels, partitions or wireways ensure that the nonintrinsically safe wiring and intrinsically safe wiring are separate from one another. Encapsulation is sometimes used to prevent contact between the two types of circuits.

Related to mechanical isolation is what can be called electrical isolation. Except in battery-operated systems, intrinsically safe systems have some connection to the power line, usually through a power transformer. The designer must consider the possibility of a transformer fault that connects the line voltage primary winding to the low-voltage secondary winding. In many systems, if one must consider the presence of line voltage on secondary circuits, the value and power rating of limiting elements would make the design of an intrinsically safe system both functionally and economically impractical. Therefore in modern standards for intrinsically safe construction, several varieties of transformer construction are recognized to be so reliable that one can assume that a primary-to-secondary short circuit will never occur. In one such reliable construction, a grounded shield between primary and secondary ensures that any fault is from the primary to the grounded shield, so that the secondary winding potential is not raised to an unsafe voltage. In addition to special attention to transformer construction and testing, it is also necessary that the wiring layout prevent any accidental connection between wiring on the primary side of the transformer and wires connected to the transformer secondary.

Current and Voltage Limiting

Except in some portable apparatus, almost all intrinsically safe circuits require both current and voltage limiting to ensure that the amount of energy released under fault conditions does not exceed safe values. Voltage limiting is often achieved by use of redundant zener diodes to limit the voltage, but zener-triggered silicon-controlled rectifier (SCR) crowbar circuits are also. Redundancy is provided so that, in the case of the failure of a single diode or limiting circuit, the second device continues to provide voltage limiting. Current limiting in dc circuits and in most ac circuits is provided by film or wirewound resistors of high reliability (Fig. 5). Properly mounted resistors that meet the requirements for protective resistors in the applicable standard need not be redundant. They are of a level of quality that they will not fail in a way that allows current to increase to an unsafe level.

One common use of current and voltage limiting is in the zener diode barrier (Fig. 6). The unique feature of these barriers is the fuse in series with the zener diodes, so that when a fault causes current to flow through the zener diode, the fuse will open before the power in the zener diode reaches a level at which the diode may open. In the design shown in Fig. 6, the 20-ohm resistor does not perform a safety function. It allows testing of the barrier to determine that the diodes are still intact. The current limiting function is performed by the 300-ohm resistor.

Devices to be connected to terminals 3 and 4 must be approved as intrinsically safe, but any device incapable of applying voltage to terminal 1 higher than the barrier rating may be connected. If the barrier is designed to limit against full power line potential, then the equipment in the nonhazardous area may be selected, connected, and intermingled without regard to safety in the field circuits if no potential above power line voltage is present.

In use, terminals 2 and 4 are both connected to a busbar that is grounded through a very low (usually less than 1-ohm) ground resistance. The power supply also must be grounded. In operation, diodes D_1 and D_2 conduct only leakage current, which is small compared with the normal circuit flowing between terminals 1 and 3. When high voltage is applied to terminal 1, the diodes conduct and limit the voltage at terminal 3 to a safe value. R_3 limits the current into the hazardous area. Under fault conditions, the barrier looks like a low-voltage resistive source from the intrinsically safe side, terminals 3 and 4, and like a very low-impedance load at terminals 1 and 2.

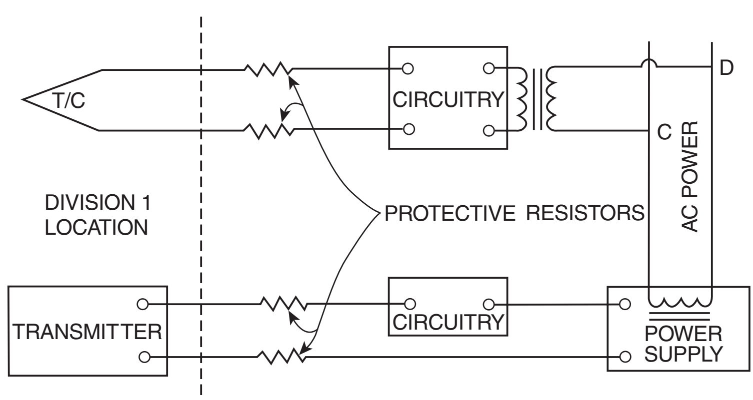


FIGURE 5 Use of resistors to limit current in hazardous location.

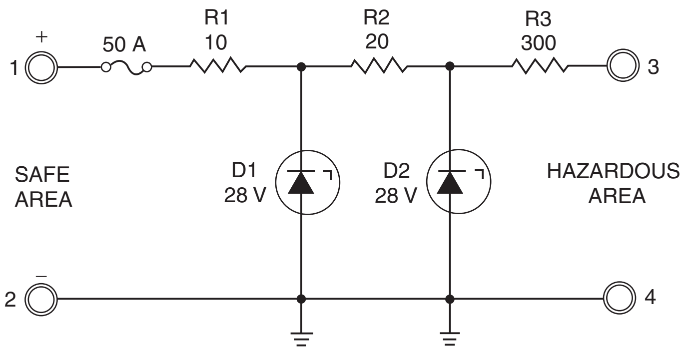


FIGURE 6 Schematic of zener diode barrier, positive type.

The values in Fig. 6 are for a 28-volt 93-mA barrier. Under fault conditions, the intrinsically safe circuit will appear to be driven from a nominal 28-volt source with a source resistance of 300 ohms. Safety is provided by the diodes and resistors. The resistors can be presumed not to fail. The diodes are redundant. Should one fail, limiting would still take place. The fuse serves no purpose regarding ignition and could be replaced by a resistor. Its function is to make the diode barrier economical. Should a fault occur, the zener diodes would connect heavily and, except for the fuse, would have to be impractically large and costly. The fuse is selected to blow at a current much lower than that which would damage the diode, permitting lower power, less costly diodes to be used.

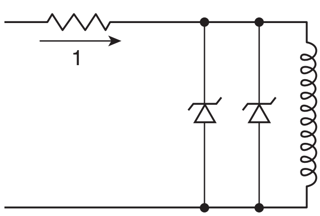


FIGURE 7 Shunt diodes reduce incendivity of inductor.

Shunt Elements

These devices are used to absorb the energy that would otherwise be released by an inductor to the arc. The function of a shunt diode is shown in Fig. 7. Although capacitors and resistors can be used to shunt inductors, in dc circuits diodes are placed so that in normal operation they are backbiased and draw no current. The shunt elements must be redundant so that if one fails, the other will continue to protect the inductor. Both must be connected close to the inductor being protected. Connection to the inductor must be especially reliable so that a fault between the inductor and the protective shunt diodes can be assumed not to occur. If such a fault occurs, ignition is possible because of the release of energy from the inductor. The purpose of the shunt diodes is to absorb energy stored in the inductor if the circuit external to the protected inductor opens.

Analytical Method for Circuit Design

The outline presented here can be considered both as a method of assessing a circuit and a mechanical design that already exists, or as a means for both analyzing the circuit and designing the layout. Only a slight difference in point of view is required. The steps in the analysis are essentially the same in both cases.

The first step is to identify the portion of the circuit that is to be intrinsically safe. Only the circuit in a Division 1 location need be intrinsically safe. A fault occurring in a nonhazardous or Division 2 location is of no consequence from the standpoint of the energy that may be released at that location. The fault is of concern if it affects the amount of energy that can be released in a Division 1 location.

Second, review the circuit or the hardware for isolating constructions that will allow one to assume that certain interconnections will not occur. If the hardware exists, review the mechanical layout to determine whether the use of adequate spacings or partitions allows one to avoid considering an interconnection between some nonintrinsically safe circuits and the intrinsically safe circuit. Usually the transformer must be one of the special protective construction. Otherwise it is unlikely that the circuits can be both functional and safe if one assumes a line voltage connection to the transformer secondary. If there is a transformer of special construction, then the wiring or the printed wiring board layout must ensure that the primary leads of the transformer are separated from the secondary leads by sufficient space (as defined in the relevant standard), or by a partition, so that connection between the two can be ignored. If the hardware has not been designed, one must determine where intrinsically safe portions of the circuit must be separated by appropriate spacing or partition from the higher-energy portion of the circuits. The specific nature of the spacings must be determined from further detailed analysis.

Having reviewed for isolating construction, one should assume normal operation of the circuit. Compute the current and voltage in the circuit and compare it with the reference curves to determine whether the appropriate test factor has been observed. If not, adjust the circuit constants until the requirements are met. In this step, and in subsequent steps, it is essential that an orderly approach to record keeping be adopted. Even a relatively simple circuit may require the consideration of many steps. It is essential to record each combination of open, short, or ground of field wiring (these are not counted as faults) and component failures, so that when the analysis is complete, one can verify that the worst-case situation has been analyzed. If one is submitting the circuit for later review by certifying agencies, the availability of a well-organized fault table will ease and expedite the assessment process.

Third, if the hardware already exists, one considers the layout and spacing to determine what circuits must be assumed to be shorted together, what connections can be considered to be a fault, and what connections can be considered never to occur. After identifying these, recompute the current and voltage under fault conditions. This must be done for a single fault in combination with opening, grounding, and shorting of the external wires—and also for two faults. One cannot assume that two-fault situations will be the most hazardous, because of the difference in test factor required when only a single fault is considered. Adjust the circuit constants until the voltages and currents are suitable.

After the analysis for arc ignition, consider whether current flowing in the circuit under fault conditions may produce a high surface temperature on resistors, transistors, and so on. The temperature rise of small components is typically 50 to 100°C/watt. If the hardware exists, of course, one can measure the temperature of components which are suspect under fault conditions. The limiting temperature to be considered must, of course, be determined from the standard being used as a criterion for design. Assessment is made based on the maximum voltage, current, and power ratings at the intrinsically safe input terminals.

Simplifying Assumptions

If one is analyzing a circuit, the validity of the analysis is only as good as the data with which the analysis can be compared. In general, the available data are limited to simple *RL*, *RC*, or resistive circuits. It may not be possible to analyze the effects of shunt elements included for some functional purpose. For example, an inductive force coil might be shunted with a variable resistor for a span adjustment. There is no reliable way to assess the additional safety provided by the shunt resistor. In general, one ignores such a parallel component in the analysis. Similarly, a capacitor might be wired in shunt with the force coil to provide damping of the electromechanical system. This, too, is ignored in the analysis, except with regard to the release of its own stored energy.

It is assumed by almost all experts that an iron-core inductor is less efficient in releasing stored energy than an air-core inductor, because some of the energy, rather than being released to the arc, is dissipated in eddy current and hysteresis losses. If one is analyzing a circuit in which the inductor has a ferromagnetic core and the circuit will be safe with an air-core inductor of the same value, then it will certainly be safe. The converse is not true. Inductance is a measure of the slope of the B–H curve of the core material. Many small inductors, especially those with ferrite cores, have high inductance because the core material has a very high initial permeability. However, if the volume of material and the level at which it saturates are low, the amount of stored energy in the inductor may be considerably less than that calculated from a measured inductance value. Testing may verify safety.

Another simplifying assumption, which in many circuits reduces the amount of analysis considerably, is to determine the highest possible power supply voltage that may ever exist under fault conditions and with a high line voltage. This value is then used to determine the stored energy in all the capacitors or to determine the current and the resulting stored energy in all the inductors. If all the calculations are safe, one need not calculate the actual circuit currents and voltages.

Another simplification results from the need to prevent the discharge of a large capacitor in a hazardous location. Although curves giving the value of ignition voltage on a capacitor discharging through a resistor are available, one can assume that the capacitor is a battery charged to the fault voltage and select a series resistor based on the resistive circuit ignition characteristics. The resistor selected will be higher than that based on ignition tests of capacitors because the voltage on the capacitor decays when the current flows through the resistor. The connection between the resistor and the capacitor must be prevented from contacting any surrounding circuit. Therefore spacing, potting, or some other technique must be used to ensure that the capacitor cannot discharge except through the resistor.

Testing of Special Cases

It is not possible to determine the safety of all circuits by analysis alone. For example, some experts feel that one should verify the safety of diode-shunted inductors by conducting ignition tests. Inductors, especially small ones with ferromagnetic cores, may require testing to verify that they are safe despite high measured inductance.

Another common piece of apparatus that may have to be tested to determine safety is a regulated power supply. The reference curves of open-circuit voltage and short-circuit current for ignition in resistive circuits assume that the source impedance of the circuit, that is, the Thévenin equivalent impedance, is resistive. If the power supply is regulated, the voltage will remain essentially constant until a critical level of current is reached, beyond which the voltage will drop off with a further increase in current. The safety of maximum voltage and maximum current from such a supply cannot be determined from the reference curves for resistive circuits. In general, safety must be established by test, although some reference curves are available.

Transmission lines are another special case. The common method of assessing safety, that is, multiplying capacitance per foot (meter) or inductance per foot (meter) by the number of feet (meters) of cable and comparing these values with limit values for the voltage and current in the cable, yields conservative results. It is well known that because the resistance, capacitance, and inductance are distributed, the actual cable will be safer than this lumped constant analysis suggests. However, there are not sufficiently good reference data available to allow one to analyze on a more scientific basis. If answers from lumped constant approaches are not satisfactory, then the cable must be tested.

However, it is recognized that the L/R ratio of the cable, if it is sufficiently low, may be such that no hazard exists, even though the total inductance exceeds that which would be safe if the inductance were lumped. If one assumes a cable of resistance R and inductance L_x per foot (meter) operating from a circuit of maximum voltage $V_{\max}$ and source resistance R, then the maximum energy will be stored in the cable when the cable resistance is equal to the source impedance. The current will be $V_{\max}/2R$. Therefore the maximum inductance permitted will be four times that permitted at the short-circuit current of the source. If the ratio L_x/R_x does not exceed $4L_{\max}/R$, where $L_{\max}$ is the maximum permitted connected inductance, the cable will be safe regardless of length.

In summary, the design techniques used in all commercially available systems are similar. In this section we noted that the fundamental techniques are very few in number. Manufacturers may introduce variants of the basic techniques, some quite imaginative, but the fundamental design techniques are similar in all commercially available systems. Although some systems have used current limiting resistors and have put voltage limiting in the power supply, some have used active barrier isolators, and some have used zener diode barriers, there is no difference in safety among the various approaches. Any of the techniques properly applied will yield a safe system.

CERTIFICATION OF INTRINSICALLY SAFE APPARATUS

In the early years almost all intrinsically safe apparatus was certified as part of a complete loop, now called *loop certification*. The apparatus in the Division 2 or nonhazardous location, now called *associated apparatus*, was specified either by model number or, somewhat later, in the case of intrinsic safety barriers, by the electrical characteristics V_{max} and I_{max}. The intrinsically safe apparatus was also specified by model number.

In Germany a different scheme developed, now common worldwide. Associated apparatus is characterized by the maximum open-circuit voltage V_{oc}, the maximum short-circuit current I_{sc}, and the maximum permissible connected capacitance and inductance C_a and L_a.

Intrinsically safe apparatus is characterized by V_{max}, I_{max}, and P_{max}, the maximum voltage, current and power that can be safely applied at the terminals, and C_i and L_i, the effective capacitance and inductance seen at the terminals. (A large capacitor discharging into the external circuit through a current limiting resistor may appear at the terminals to be equivalent to a much smaller capacitor, and an inductor shunted by diodes may appear to be a small inductor.)

IEC and CENELEC standards are written around this kind of specification, known in North America as *entity approval*.

In current North American practice the manufacturer supplies a control drawing which details the interconnections of apparatus evaluated during the certification process. Apparatus may be defined broadly by specifying maximum parametric values, as in European practice, or very specifically by calling out model numbers. This drawing also specifies any special installation conditions, such as maximum load or power supply voltages, fusing, etc. for associated apparatus, or special grounding.

SYSTEM DESIGN USING COMMERCIALLY AVAILABLE INTRINSICALLY SAFE AND ASSOCIATED APPARATUS

General Principles

This section summarizes the principles underlying implementation of an intrinsically safe system using commercially available components. These components are the intrinsically safe apparatus in the Division 1/Zone 0/Zone 1 location and the associated apparatus in the Division 2/Zone 2 or nonhazardous location.

The controlling demand of all intrinsically safe system design is that every ungrounded conductor entering the Division 1 location (or Zone 1/Zone 0 location) must be protected against the inflow of nonintrinsically safe voltage, current, and power levels from the Zone 2/Division 2 or nonhazardous location.

If associated apparatus, such as a zener barrier, protects every ungrounded line so that application of line voltage to the associated apparatus does not cause nonintrinsically safe voltage and currents to flow in the protected lines then the apparatus on the nonhazardous side of the associated apparatus may be chosen with regard to function only. It need not be certified. The only restriction on this apparatus is that there be no voltage in it that exceeds the voltage rating of the associated apparatus. Most associated apparatus is rated for intrinsically safe outputs after application of line voltage, usually 250 V RMS.

The major obstacles to achieving an intrinsically safe system fall broadly into two categories.

1. The field-mounted equipment is not certified to be intrinsically safe.
2. The control house equipment is not certified as associated apparatus.

If the field-mounted apparatus and the control house interface apparatus have been certified intrinsically safe, separately or together, select equipment for the system which lies within the limits on their control drawings, choosing the alternatives that best fit the plant situation.

Field-mounted Apparatus not Certified. If the field mounted device is simple apparatus, it need not be certified when connected to certified intrinsic safety barriers or other associated apparatus. Simple apparatus includes:

- passive devices such as switches, junction boxes, resistance temperature detectors, potentiometers
- simple semiconductor devices
- sources of stored energy, such as inductors or capacitors with well defined characteristics, which are to be taken into account when assessing intrinsic safety, or
- sources of generated energy such as solar cells and thermocouples. These sources shall not generate more than 1.5 V, 100 mA, or 25 mW.

In some standards the stored energy sources are not mentioned, and simple apparatus is defined as apparatus which neither generates nor stores more than 1.2 V, 100 mA, 25 mW, or 20 μJ. These standards tactly assume that inductive or capacitive storage elements will be taken into account as part of the L of C that may be connected to the associated apparatus.

Simple apparatus must be separated from nonintrinsically safe circuits and apparatus to ensure that nonintrinsically safe energy levels will not be injected into the intrinsically safe circuits. Most experts assume that 50-mm spacing or partitions provide sufficient protection against intermixing of the two circuits. Sometimes it has been stated that the simple apparatus shall not be in the same enclosure as nonintrinsically safe circuits or apparatus, but in the author's opinion, this is too restrictive a translation of reasonable intent into a rule. Always conform to limitations imposed by the control drawing for the associated apparatus.

Simple apparatus must be investigated for the applicable Temperature Code. If the barrier supplying it can deliver no more than 1.2 W maximum power then a T4 (135 C) rating is usually reasonable and defensible. If T5 (100 C) rating is desired it is necessary to investigate the temperature rise when exposed to the maximum power delivered from the barriers used in the system.

If the field-mounted apparatus is not certified, and does not conform to the limitations imposed on simple apparatus, the apparatus cannot easily be used in an intrinsically safe system. Few, if any, system designers can afford to make the investment required to self-certify a design, and because the user can't control the details of the design, the assessment applies only to the specific design of the piece of apparatus investigated. In the current regulatory environment few managers would risk using apparatus that has not been certified by a third party, so the need to consider self-certification is unlikely to arise. It may sometimes be necessary to permit temporary use of a product that has been submitted to a third party for certification. The supplier should provide a written assessment of the intrinsic safety of the device and a schedule for the availability of the certification. The device should not be installed unless the user has sufficient knowledge and resources to review and assess the vendor's self-certification document.

Control House Apparatus not Certified as Associated Apparatus. For the same reasons that it is not practical to self-certify an item of field-mounted apparatus, it is not sensible to consider self-certifying or accepting a vendor's self-certification of control room apparatus. There is seldom any need to consider this option. In principle, one needs only to interpose between the control house apparatus and the field apparatus a certified associated apparatus, such as a barrier. The barrier must be compatible with the intrinsic safety parameters of the field device and must also allow the system to function as intended when it is connected between the field device and the control room device.

Uncertified Control Room Apparatus Connected to Uncertified Field Mounted Apparatus. A system designer may occasionally encounter a special case of uncertified field-mounted apparatus connected to uncertified control room apparatus. One solution is to provide the field mounted device with another type of protection, usually an explosionproof enclosure or pressurization. Install barriers in its signal lines to the control room. These are usually mounted in the same enclosure as the apparatus generating the signal, but could be installed in a separate enclousre suitable for the location. Install barriers also at the control room apparatus. The two sets of barriers prevent ignition capable energy from entering the Division 1 location on the signal lines from either the field apparatus or from the control room apparatus. Select barriers so that the current drawn into a fault caused by shorting and grounding of the signal lines in the Division 1 location remains at a safe level, i.e. the sum of the short circuit currents of the barriers must be safe. This scheme is seldom applicable to a transmitter powered over the signal lines, as in two-wire ungrounded 4–20 mA transmitters. It is difficult to specify barriers that meet the functional requirements of the circuit, and also have safety descriptions with low enough values of current and voltage to permit summing four barrier currents into a ground fault in the line between the barriers. The technique is more frequently used with field mounted transmitters that have their own source of power, so that the signal from the transmitter to the control house is a voltage or a current of relatively low level, perhaps 1–5 V, 4–20 mA. Barriers with safety descriptions such as 9 V, 100 mA can be utilized in such signal lines to ensure both safety and function of the circuit.

Vendors of barriers make available a wealth of application information, including specific recommendations of barrier models for use in the application being discussed. The reader would be wise to use this free support.

INSTALLATION OF INTRINSICALLY SAFE SYSTEMS

Because an intrinsically safe system is incapable of igniting a flammable atmosphere even under fault conditions, cables of special construction or conduit are unnecessary. It is, however, necessary to install intrinsically safe systems so that ignition-capable energies will not intrude from another circuit.

An intrinsically safe system will be safe *after* installation if the installation

1. Conforms to the limiting parameters and installation requirements on which approval was based, as stated on the control drawing
2. Prevents intrusion of nonintrinsically safe energy on intrinsically safe circuits
3. Prevents power system faults or differences in ground potential from making the circuit ignition-capable

To prevent the intrusion of other circuits, intrinsically safe circuits must be run in separate cables, wireways, or conduits from other circuits. Terminals of intrinsically safe circuits must be separated from other circuits by spacing (50 mm) or partitions. The grounding of intrinsically safe circuits must be separate from the grounding of power systems, except at one point.

Nonincendive Equipment and Wiring

It is not necessary to provide intrinsically safe equipment in Division 2 locations. The equipment need only be nonincendive, that is, incapable in its normal operating conditions of releasing sufficient energy to ignite a specific hazardous atmospheric mixture. Such equipment has been recognized without specific definition in the NEC. In Division 2 locations, equipment without make-or-break or sliding contacts and without hot surfaces may be housed in general-purpose enclosures.

Requirements for apparatus suitable for use in Division 2 and Zone 2 locations have been published by ISA and IEC. These documents provide more detail than is found in the NEC, partially because of a trend toward certification of nonincendive circuits that may be normally sparking, but release insufficient energy in normal operation to cause ignition. These documents also define tests for sealed devices that are needed by industry. The documents are similar, except that the ISA standard does not cover restricted breathing and enclosed-break devices.

IGNITION BY OPTICAL SOURCES

Until the early 1980s it was assumed that transmitting instrumentation and control signals on optical fibers would avoid all the issues relative to ignition of flammable atmospheres by electrical signals.

Research showed that the power level required to ignite flammable vapors and gases directly is much higher than would be found in common measurement and control systems. However, if an optical beam irradiates a particle and raises its temperature sufficiently, ignition can occur. During 1990–1994 a collaborative program in the European Community undertook to bound the problem. The report of this work, EUR 16011 EN confirmed that ignition of a small particle occurs at low enough optical power levels to be of concern in measuring and control systems. Based on the data gathered during this investigation it was concluded that there is essentially no hazard of ignition by a continuous-wave optical beam in the visible or near infrared if the peak radiation flux is below 5 mW/mm^2 or the radiated power is less than 35 mW. This recommendation is about a factor of 2 lower than the lowest data points, which included materials with low AIT values, such as diethyl ether and carbon disulfide, and materials like hydrogen and carbon disulfide with low spark ignition energy.

Though this guidance is helpful because many uses of fiber optics operate below these limit values those using higher power levels in analytical instruments and those using laser power for other purposes are left without guidance.

Ignition by the heating of small particles is related to ignition by hot wires and components, and by AIT determinations in that it is a thermal process where the value of the critical power or power density measured depends on the amount of time lag before ignition occurs. However, the amount of power required does not correlate strongly with conventional combustible properties such as Minimum Ignition Energy, MESG, or AIT. Estimating the amount of power required for ignition of materials other than those tested. or for material for which there is sparse data is not easy.

A second series of investigations scheduled to be completed in 1999 is intended to provide more data. It is known that igniting power for many common industrial materials in Group D (Group IIA) is 200 mW or more depending on the size of the heated target.

Regulations and Standards

At the outset there was a tendency in some quarters, which persists still, to blindly rush to apply standards for mitigating hazards in electrical systems to this problem. These efforts ignored the fact that in many applications a fiber must break in order to illuminate a particle, and that many optical systems provide easy means of detecting this and other failures in time to take remedial action, such as shut-down. This author believes that the hazard of optical systems should be treated with the tools of risk assessment, fault-tree analyses, and similar tools used for assessing and mitigating other types of hazard in the process industries.

ACKNOWLEDGMENT

The cooperation of the Instrument Society of America in permitting extractions and adaptations of text and illustrations from *Electrical Instruments in Hazardous Locations* is greatly appreciated.

BIBLIOGRAPHY

ANSI/NFPA 70, "National Electrical Code (NEC)," National Fire Protection Association, Quincy, Massachusetts, Art. 500–505.

API RP 500A, "Classification of Areas for Electrical Installations in Petroleum Refineries," American Petroleum Institute, New York.

API RP 505, "Recommended Practice for Classification of Electrical Installations at Petroleum Facilities Classified as Class I, Zone 0, Zone 1, or Zone 2," American Petroleum Institute, New York.

BS 5501, "Electrical Apparatus for Potentially Explosive Atmospheres," British Standard Institution, London, 1978.

Part 1: “General Requirements” (EN 50014)
Part 2: “Oil Immersion ‘o’ ” (EN 50015)
Part 3: “Pressurized Apparatus ‘p’ ” (EN 50016)
Part 4: “Powder Filling ‘q’ ” (EN 50011)
Part 5: “Flameproof Enclosures ‘d’ ” (EN 50018)
Part 6: “Increased Safety ‘e’ ” (EN 50019)
Part 7: “Intrinsic Safety ‘i’ ” (EN 50020)

Dubaniewicz, T. H. Jr., K. L. Cashdollar, and G. M. Green, with appendix by R. F. Chalken, “Ignition of Methane-Air Mixtures by Laser Heated Small Particles,” International Symposium on Hazards, Prevention, and Mitigation of Industrial Explosions, Schaumburg, Illinois, 1998.

ISA RP12.1, “Electrical Instruments in Hazardous Atmospheres,” Instrument Society of America, Research Triangle Park, North Carolina.

ISA Monographs 110–113, “Electrical Safety Practices,” Instrument Society of America, Research Triangle Park, North Carolina.

ISA RP12.6, “Installation of Intrinsically Safe Instrument Systems in Class 1 Hazardous Locations,” Instrument Society of America, Research Triangle Park, North Carolina.

ISA S12.12, “Electrical Equipment for Use in Class 1, Division 2 Locations,” Instrument Society of America, Research Triangle Park, North Carolina.

Magison, E. C., *Electrical Instruments in Hazardous Locations* 4th ed., Instrument Society of America, Research Triangle Park, North Carolina, 1998.

Magison, E. C., *Intrinsic Safety, an Independent Learning Module,* Instrument Society of America, Research Triangle Park, North Carolina, 1984.

Magison, E. C., “Are Optical Beams an Explosion Hazard?” *InTech*. pp. 61–66, Sept. 1997.

McGeehin, P., “Optical Techniques in Industrial Measurement: Safety in Hazardous Environments,” Report EUR 16011 EN, Luxembourg, 1994.
Annex I: Summary report, F. B. Carleton, and F. J. Weinberg, Imperial College
Annex II: Final report, M. M. Welzel, H. Bothe, and H. K. Cammenga, PTB
Annex III: Summary report, G. Tortoishell, SIRA Test and Certification Ltd.
Annex IV: K. S. Scott, University of Leeds
Annex V: Final report, C. Proust, INERIS

NFPA 496, “Purged and Pressurized Enclosures for Electrical Equipment,” National Fire Protection Association, Quincy, Massachusetts.

NFPA 497, “Recommended Practice for Classification of Flammable Liquids, Gases or Vapors and of Hazardous (classified) Locations for Electrical Installations in Chemical Process Areas,” National Fire Protection Association, Quincy, Massachusetts, 1997.

NFPA 499, “Recommended Practice for Classification of Combustible Dusts and of Class II Hazardous (Classified) Locations for Electrical Installations in Chemical Processing Plants,” National Fire Protection Association, Quincy, Massachusetts, 1997.

IEC Pub. 79–15 “Electrical Apparatus with Type of Protection ‘n,’ ” International Electrotechnical Commission, Geneva, Switzerland.

UL 913, “Intrinsically Safe Apparatus and Associated Apparatus for Use in Class I, II, and III, Division 1 Hazardous (Classified) Locations,” Underwriters Laboratories, Northbook, Illinois.

Buschart, R. J., *Electrical and Instrumentation Safety for Chemical Processes*, Van Nostrand Reinhold, New York, 1991.

Guidelines for Hazard Evaluation Procedures, 2nd ed. Center for Chemical Process Safety, American Institute of Chemical Engineers, New York, 1992.

Welzel, M. M., S. Schenk, M. Hau, K. K. Commenga, and H. Bothe, “Einfluss des Brennstoff/Luft-Gemisches auf die minimale zuendfaehige Strahlungsdleistung bei Strahlungsabsorption an einer Eisenoxideoberflaeche,” PTB-Mittellungen No. 109, 2/99.