# Process-Control Systems

# Process-Systems

## *Application* **Design** **Adjustment**

F. G. SHINSKEY
*Systems Design Engineer, The Foxboro Company*

McGRAW-HILL BOOK COMPANY
New York San Francisco Toronto London Sydney

# Preface

Consistent quality, increased productivity, and reduced operating costs can all accrue from effective control of major plant variables. But effective control is not an accident, nor does it come into being simply by implementing postulated theories. To be truly effective, a control system must be designed to fit the needs of the particular process to which it will be applied. The more process intelligence the designer can put into a system, the greater are its chances of success. The fact that many control systems have been operating satisfactorily indicates that plant engineers are capable of such design r [illegible] successful design is not necessarily contingent on graduate-level mathematics. Ziegler and Nichols* made this evident with the derivation of their simple rules for setting automatic controllers. First presented in 1941, this derivation is still being used today.

Many scientists are busily at work in laboratories and universities, searching for more advanced control concepts. The principles they are discovering, however, could never realize their full worth if they

* J. G. Ziegler and N. B. Nichols, Optimum Settings for Automatic Controllers, *Trans. ASME*, December, 1941.

are not communicated to the people who must apply them. Control problems arise in the plant and must be solved in the plant. Until plant engineers and control designers are able to communicate with each other, their mutual problems await solution. I do not mean to imply that abstract mathematics is not capable of solving control problems, but it is striking how often the same solution can be reached by using good common sense. High-order equations and high-speed computers can be manipulated to the point where common sense is dulled.

Some months ago I was asked to give a course on process control to a large group of engineers from various departments of The Foxboro Company. Sales, Product Design, Research, Quality Control, and Project Engineering were all to be represented. If the subject were presented through the traditional medium of operational calculus, the effort would be wasted, because too few of the students would have this prerequisite. Rather than attempt to teach operational calculus, I chose to do without it altogether. It then became necessary to approach control problems solely in the time domain. Once the transition was begun, I was surprised at the fresh point of view which evolved. Some situations which were clouded when expressed in frequency or in complex numbers were now easily resolved. Dead time, fundamental to any transport process, is naturally treated in the time domain.

The value of this new approach was evident at once. In the very first session the student was able to understand why a control loop behaves the way it does: why it oscillates at a particular period, and what determines its damping. The subject was tangible and alive to many students for the first time. Interest ran high, and the course was an immediate success. The great demand for notes prompted the undertaking of this book.

Through the years, I have observed many phenomena about control loops which have never been explained to my satisfaction. Why does a flow controller need such a wide proportional band, whereas a pressure controller does not? Why is derivative less effective in a loop containing dead time than in a multicapacity loop? Why are some chemical reactors impossible to control? What makes composition control so difficult? Why cannot some oscillations be damped? These and many other observations are explained in this book and perhaps nowhere else.

It is always very satisfying to learn the reasons behind the behavior of things which are familar, or to see accepted principles proven in a new and different way. Therefore i expect that those who are accustomed to the more conventional approaches to control system design will find this treatment as interesting as those who are not familiar with any.

In spite of the simplicity of this presentation, we are not kept from applying the most advanced concepts of automatic control. Feedforward control has proven itself capable of a hundredfold improvement over what conventional methods of regulation can deliver. Recent developments in nonlinear control systems have pushed beyond traditional barriers-achieving truly optimum performance. These advances are not just speculation-they are paying out in increased throughput and recovered product. Although their impact on the process industries is as yet scarcely felt, the revolution is inevitable. The need for economy will make it so.

But the most brilliantly conceived control strategy, by itself, is nothing. By the same token, the most definitive mathematical representation of the process, alone, is worthless. The control system must be the embodiment of the process characteristics if it is to perform as intended. Without a process, there can be no control system. Anyone who designs controls without knowing what is to be controlled is fooling himself. A pressure regulator cannot be used to control composition. Neither can a temperature controller on a fractionator perform the same function as one on a heater. For these reasons this entire text is written from the viewpoint of the needs of the process. Each type of physical-chemical operation which has a history of misbehavior is treated individually. Not every situation can be covered, because plants and specifications differ, and so do people. If for no other reason, this book will never be complete. But enough attention is given to basic principles and typical applications to permit extension to a broad area of problems. The plant engineer can take it from there.

In appreciation for their assistance in this endeavor, I wish to express my gratitude to Bill Vannah for providing the initiative, to Molly Dickinson, who did all the typing, and to John Louis for his thoughtful criticism.

Greg *Shinskey*

# Contents

**Preface vii**

PART 1 UNDERSTANDING FEEDBACK CONTROL

**1. Dynamic Elements in the Control Loop 3**

*Negative Feedback* 4
*The Difficult Element-Dead Time* 6
*The Easy Element-Capacity* 18
Combinations *of Dead Time and Capacity* 31
*Summary* 35
*Problems* 35

**2. Characteristics of Real Processes 37**

*Multicapacity Processes* 38
*Gain* and *Its Dependence* 44
*Testing the Plant* 55

References 59
Problems 59

3. Analysis of Some Common LOOPS 61

Flow Control 62
Pressure Regulation 67
Liquid Level and Hydraulic Resonance 71
Temperature Control 74
Control of Composition 80
Conclusions 86
References 87
Problems 87

PART 2 SELECTING THE FEEDBACK CONTROLLER

4. Linear Controllers 91

Performance Criteria 92
Two- and Three-mode Controllers 95
Complementary Feedback 103
Interrupting the Control Loop 110
Direct Digital Control 118
References 122
Problems 123

5. Nonlinear Control Elements 124

Nonlinear Elements in the. Closed Loop 125
Nonlinear Dynamic Elements 128
Variations of the On-off Controller 131
The Dual-mode Concept 136
Nonlinear Two-mode Controllers 144
Problems 149

PART 3 MULTIPLE-LOOP SYSTEMS

6. Improved Control through Multiple Loops 153

Cascade Control 154
Ratio Control Systems 160
Selective Control Loops 167
Adaptive Control Systems 170
Summary 179
References 180
Problems 180

**7. Multivariable Process Control 181**

Choosing Controlled Variables 182
Pairing Controlled and Manipulated Variables 188
Decoupling Control Systems 198
Summary 202
References 202
Problems 203

**8. Feedforward Control 204**

The Control System as a Model of the Process 206'
Applying Dynamic Compensation 211
Adding Feedback 219
Economic Considerations 224
Summary 227
References 228
Problems 228

**PART 4 APPLICATIONS**

**9. Control of Energy Transfer 233**

Heat Transfer 23.4
Combustion Control 241
Steam-plant Control Systems 243
Pumps and Compressors 250
References 256
Problems 256

**10. Controlling Chemical Reactions 257**

Principles Governing the Conduct of Reactions 268
Continuous Reactors 269
pH Control 275
Batch Reactors 282
References 286
Problems 286

**II. Distillation 288**

Factors Affecting Product Quality 289
Arranging the Control Loops 295
Applying Feedforward Control 307
Batch Distillation 319
Summary 323
References 323
Problems 324

12. Other Mass Transfer Operations 325

Absorption and Humidification 326
Evaporation and Crystallization 332
Extraction and Extractive Distillation 338
Drying Operations 343
Summary 346
References 347
Problems 347

Appendix: Answers to Problems 349

Index 355

# *Understanding Feedback Control*

PART 1

# Dynamic Elements in the Control Loop

CHAPTER 1

What makes control loops behave the way they do? Some are fast, some slow; some oscillate, others loll in stability. What determines how well a given variable can be controlled? How are the optimum controller settings related to the process? These questions must be answered before the reader can feel he really comprehends the essence of the control problem. They will be answered in the pages that follow.

Negative feedback is the basic regulating mechanism of automatic systems-but it is not the only mechanism. Feedback has certain limitations which sometimes go unnoticed in the pursuit of better feedback controllers. Yet before progress can be made to more effective systems, the properties of simple feedback loops must be well defined.

Fortunately, a process need not be very complicated before the properties of the typical feedback loop make their appearance. A rapid introduction to loop behavior may be presented using the simplest dynamic element found in the process-dead time. This chapter is devoted exclusively to discussion of the control of simple dynamic elements which may never exist. in the pure form. But these elements do exist in various proportions in every real process. Therefore a thorough familiarity with the parts is essential for estimating the behavior of the whole.

## NEGATIVE FEEDBACK

There are two kinds of feedback possible in a closed loop: positive and negative. Positive feedback is an operation which augments an imbalance, thereby precluding &ability. If a temperature controller with positive feedback were used to heat a room, it would increase the heat when the temperature was above the set point and turn it off when it was below. Loops with positive feedback lock at one extreme or the other. Obviously this property is not conducive to regulation and therefore will be of no further concern at this time.

Negative feedback, on the other hand, works toward restoring balance. If the temperature is too high, the heat is reduced. The action taken-heating-is manipulated negatively, in effect, to the direction of the controlled variable-temperature. Figure 1.1 shows the flow of information in a feedback loop.

Throughout the text, c will refer to the controlled variable, r to the reference or set point, e to the error or deviation, and m to the variable manipulated by the controller. Note again that the effect of e, the controller input, is opposite to that of c. This can be looked on as a reversal of phase taking place at the summing junction. All negative feedback controllers exhibit this characteristic-a phase shift of 180” gives the feedback its negative sense.

### Oscillation in the Closed Loop

Rather than prove that, a feedback loop can oscillate sinusoidally, we shall assume that it does (a common observation) and shall attempt, to find out why. Oscillations are characterized by periodic applications of force in phase with the effect of the last application. **In** order to bounce a ball, a person must strike it repeatedly at the correct time, otherwise

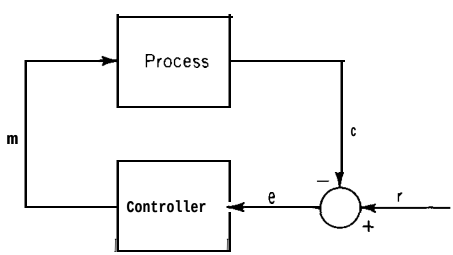



*FIG 1.1. The flow of information is backward from process output through the controller to process input.*

it will cease to bounce. The correct "time" turns out to be the correct phase. If the ball is struck at any phase angle other than 360” (of motion) from where it was last struck, the oscillation will be changed. It is apparent, then, that if oscillations are to persist, the shift in phase of a signal after proceeding through the entire loop must be exactly 360”.

It has already been pointed out that negative feedback, being negative, introduces 180” of phase shift. This means that if a closed loop is to oscillate, the dynamic elements in the controller and the process must contribute an additional 180”.

### The Natural Period

It has also been observed that the period of oscillation which a particular loop will exhibit is characteristic of that loop. The loop resonates at that period. Furthermore, any disturbance not periodic, applied to the loop but containing components near the natural period, will excite oscillations of the natural period. A pendulum is a good example of a feedback loop. The controlled variable is the angular position of the mass, and the set point is the vertical position. The mass of the pendulum, acted upon by gravity, is the manipulated variable, which tries to restore the angle to zero. Its natural period in seconds is

$$\tau_o = \frac{1}{2\pi}\left(\frac{L}{g}\right)^{1/2}$$

where $L$ = length, ft

$g$ = acceleration of gravity, ft/sec$^2$

A pendulum disturbed from rest by an impulse will proceed to oscillate at its own period. Impulse, step, and random disturbances contain a wide spectrum of periodic waves. The resonant system, however, responds only to the component of its own natural period, rejecting the rest. For this reason, we are interested in the response of the loop to a wave of the natural period and are generally unconcerned about the rest. The natural period of oscillation will be designated $\tau_o$ and will be recognized hereafter as a property peculiar to each control loop.

The natural period of any loop depends on the combination of all dynamic elements within it, including the controller. Since the amount of phase lag of most dynamic elements varies with the period of the wave passing through them, there is one particular period at which the total phase lag will equal 180”. This is the period at which the loop naturally resonates. The natural period is a dependent variable. We can make use of its relation to the process dynamics in two ways:

1. If the characteristics of the elements in the process are known, the natural period under closed-loop control can be predicted.

2. If a process whose elements are largely unknown is under closed-loop control, the characteristics of these elements can be inferred by observing the natural period.

### Damping

The gain of an element is defined as the ratio of the change in its output to the change in its input. If the controller gain were zero, it would not contribute to oscillation. But if the controller gain were sufficient to produce a second disturbance equal to the first, the loop would oscillate uniformly. Uniform oscillation requires that a wave travel completely through the loop, returning to its starting point with its original amplitude. For such a condition to exist, the gain product of all the elements in the loop must equal unity. If the gain product is less than unity, oscillations are damped.

To summarize, a loop will oscillate uniformly:

1. At a period at which the phase lags of all the elements in the loop total 180"

2. When the gain product of all the elements at that period equals 1.0

The conditions for uniform oscillation will serve as a convenient reference on which to base rules for controller adjustment.

## THE DIFFICULT ELEMENT-DEAD TIME

### Identification

As the name implies, dead time is the property of a physical system by which the response to an applied force is delayed in its effect. It is the interval after the application of a force during which no *response* is observable. This characteristic does not depend on the nature of the applied force; it always appears the same. Its dimension is simply that of time.

Dead time occurs in the transportation of mass or energy along a particular path. The length of the path and the velocity of motion consti-

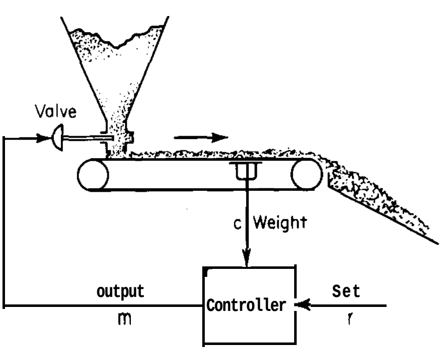



*FIG 1.2. The response of the weigh cell to a change in solids flow is delayed by the travel of the belt.*

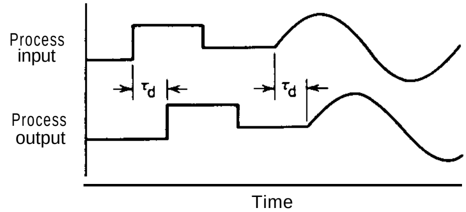



***FIG 1.3. Pure dead time transmits the input delayed by $\tau_d$.***

tute the delay. Dead time is also called "pure delay," "transport lag," or "distance-velocity lag." As with other fundamental elements, it rarely occurs alone in a real process. But there are few processes where it is not present in some form. For this reason, any useful technique of control system design must be capable of dealing with dead time.

An example of a process consisting of dead time alone is a weight-control system operating on a solids conveyor. The dead time between the action of the valve and the resulting change in weight is the distance between the valve and the cell (feet), divided by the velocity of the belt (ft/min). Dead time is invariably a problem of transportation.

A feedback controller applies corrective action to the input of a process based on a present observation of its output. In this way the corrective action is moderated by its observable effect on the process. A process containing dead time produces no immediately observable effect-hence the control situation is complicated. For this reason, dead time is recognized as the most difficult dynamic element naturally occurring in physical systems. So that the reader may begin without illusions about the limitations of automatic controls in their influence over real processes, the difficult clement of dead time is presented first.

The response of a dead-time element to any signal whatever will be the signal delayed by that amount of time. Dead time is measured as shown in Fig. 1.3.

Notice the response of the element to the sine wave in Fig. 1.3. The delay effectively produces a phase shift between input and output. Since one characteristic of feedback loops is the tendency toward oscillation, the property of phase shift becomes an essential consideration.

## The Phase Shift of Dead Time

We are primarily interested in phase characteristics of elements at the natural period of the loop. Assume, to begin, that a closed loop containing dead time is already oscillating uniformly. The input to the process is the sine wave

$$m = A \sin 2\pi \frac{t}{\tau_o} + m_0$$

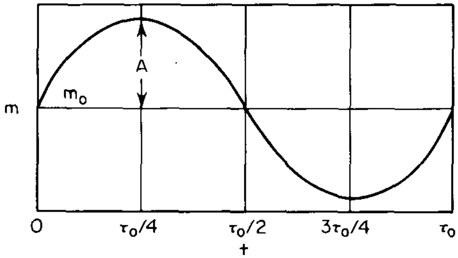



*FIG 1.4. The manipulated variable is cycling with an amplitude of A at the natural period.*

where m = manipulated variable whose average component is $m_0$
A = amplitude
$t$ = time
$\tau_o$ = period

Phase angles will be expressed both *in* degrees and in radians for reasons that will become clear later.

| $t/\tau_o$ | $2\pi t/\tau_o$ Degrees | $2\pi t/\tau_o$ Radians | $\sin 2\pi t/\tau_o$ |
|---|---|---|---|
| 0 | 0 | 0 | 0 |
| ¼ | 90 | $\pi/2$ | +1 |
| ½ | 180 | $\pi$ | 0 |
| ¾ | 270 | $3\pi/2$ | −1 |
| 1 | 360 | $2\pi$ | 0 |

This wave, passing through a dead time, will be delayed by an amount $\tau_d$, but will be undiminished, so that the output will be

$$c = A \sin 2\pi \frac{t - \tau_d}{\tau_o} + m_0$$

The input angle subtracted from the output angle yields the phase shift $\phi_d$:

$$\phi_d = 2\pi \frac{t - \tau_d}{\tau_o} - 2\pi \frac{t}{\tau_o}$$

$$= -2\pi \frac{\tau_d}{\tau_o} = -360° \frac{\tau_d}{\tau_o} \tag{1.1}$$

The negative sign indicates a lag in phase.

Because dead time does not alter the shape or amplitude of a signal, its gain $G_d$ is unity to all periodic waves:

$$G_d = 1.0 \tag{1.2}$$

## Proportional Control of Dead Time

Having defined the process, the next step is the selection of a suitable controller. A proportional controller will be chosen first, because of its simplicity. It contains no dynamic elements. Output and input are related by the expression

$$m = \frac{100}{P} e + b \tag{1.3}$$

where $P$ = proportional band, %
$e$ = error or deviation of the measurement from set point
$b$ = output bias

As $P$ approaches zero, the gain of the proportional controller approaches infinity. At 100 percent band, the gain is 1.0. The output of the controller equals the bias when there is no error.

Because there are no dynamic elements in the proportional controller, the entire 180" phase shift will take place in the dead-time element. This determines the natural period:

$$\phi_d = -180° = -\pi$$

Substituting for the previously determined $\phi_d$,

$$-2\pi \frac{\tau_d}{\tau_o} = -\pi \qquad -360° \frac{\tau_d}{\tau_o} = -180°$$

Solving for $\tau_o$,

$$\tau_o = 2\tau_d \tag{1.4}$$

The relationship is as plain as it appears. A 1-min dead-time process will cycle with a 2-min period under proportional control. This is not an approximation-it is exact.

Next it is important to estimate the proportional band necessary to sustain oscillation. Dead time offers no gain contribution, so if the loop-gain product' is to be 1.0, the controller proportional band must be 100 percent. To dampen the oscillations, the band must be increased, thus attenuating the input cycle.

Figure 1.5 illustrates how a proportional band of 200 percent reduces the amplitude of each successive half-cycle by one-half, resulting in "¼-amplitude damping" of each successive cycle. This degree of damping is generally accepted as nearly optimum throughout the industry.

Notice that there is only one adjustment available, and it affect's the damping. Given a process consisting of a I-min dead time to be cow

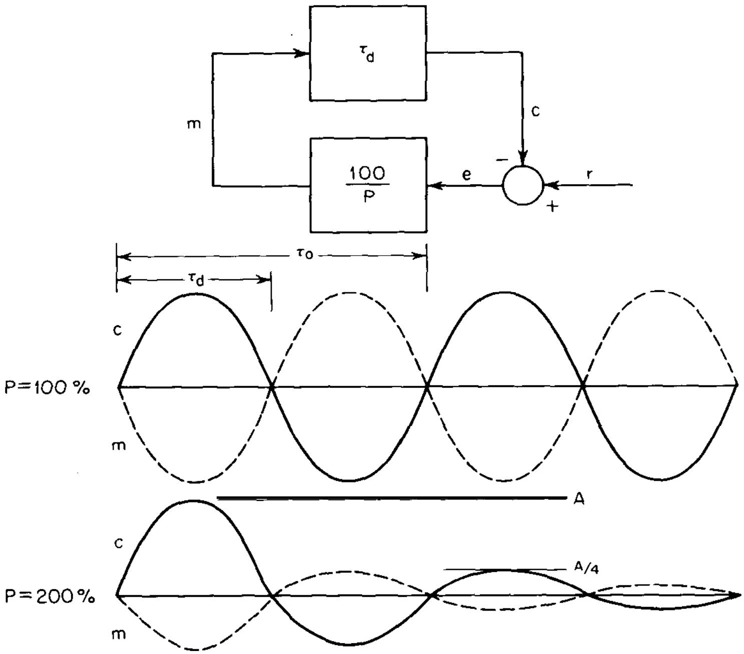



*FIG 1.5. A loop gain of 0.5 will provide ¼-amplitude damping.*

trolled by proportional only, adjusted to ¼-amplitude damping, the natural period is fixed at 2 min, and the proportional band must be 200 percent. The nature of the process determines the results.

**Proportional Offset**

The prime function of a controller is that of regulation. The controller is intended to change its output as often and as much as necessary to keep the controlled variable at the set point. Every process is subject to variations in load. In a well-regulated loop, the manipulated variable will be driven to balance the load. Consequently, the load is often measured in terms of the corresponding value of controller output.

In the equation describing the proportional controller, the bias $b$ equals the output when the error is zero. This bias may be fixed at the normal value of output, usually 50 percent, or it may be adjusted by hand to match the current load. This adjustment is called "manual reset." But because of the proportional relationship between input and output, a change in output by any amount cannot be gained without a corresponding change in error. Should the output of the proportional controller have to change to meet a new load condition, a deviation will appear:

$$e = \frac{P(m - b)}{100} \tag{1.5}$$

The deviation in this case is known as "offset," and it increases with proportional band. With a 200 percent band, which was necessary for ¼-amplitude damping in the previous example, a 10 percent change in load would produce a 20 percent offset-an intolerable amount.

The characteristics of a dead-time process under proportional control may be observed in a simple algebraic simulation. Let the present output of the controller equal the measurement one dead time later:

$$c_n = m_{n-1}$$

where $n = t/\tau_d$. This represents a process whose gain is unity and whose dead time is $\tau_d$. When the controller is introduced to close the loop,

$$m_n = \frac{100}{P}(r - c_n)$$

$$m_{n+1} = \frac{100}{P}(r - c_{n+1}) = \frac{100}{P}(r - m_n)$$

With initial conditions of $c_0 = 0$, $b = 0$, $r_0 = 0$, and $P = 200$ percent, let the loop be upset by a set-point change to 50 percent. Subsequent values of c at intervals of dead time are as follows:

| | | |
|---|---|---|
| $r_0 = 0\%$ | $c_0 = 0\%$ | $m_0 = 0\%$ |
| $r_1 = 50$ | $c_1 = 0$ | $m_1 = 0.5(50 - 0) = 25$ |
| | $c_2 = 25$ | $m_2 = 0.5(50 - 25) = 12.5$ |
| | $c_3 = 12.5$ | $m_3 = 0.5(50 - 12.5) = 18.75$ |
| | $c_4 = 18.75$ | $m_4 = 0.5(50 - 18.75) = 15.625$ |
| | $c_5 = 15.625$ | |
| | $c_\infty = 16.667$ | $m_\infty = 16.667$ |

Notice that c exhibits a damped oscillation whose period is two calculations (two dead times). Notice also that the amplitude of successive crests is diminished by one-quarter. Finally, there is an offset. The controller output comes to rest at 16.667 percent above the bias. The offset is

$$r - c = 33.333\%$$

which equals

$$\frac{P}{100}(m - b) = 2(16.667\%)$$

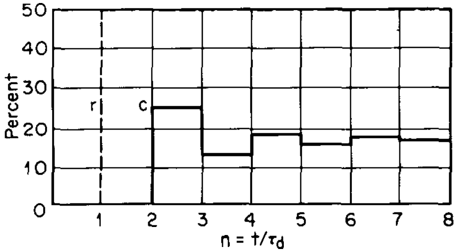



*FIG 1.6. Proportional control of pure dead time can oscillate in a square wave.*

The tabulated course of the controlled variable plots as a damped square wave. This is entirely possible when a process of pure dead time is excited by a step. The loop responds to higher harmonics as well as to fundamental, since the process does not attenuate waves of any period. Odd harmonics shift the phase in increments of 360°, so as to permit oscillation at these periods also, and square waves are made of odd harmonics. Although a square-wave response is possible, it is not likely to occur in processes, because ordinarily energy cannot be delivered fast enough to make the controlled variable rise steeply.

The kind of response more likely to occur is a load change, requiring a different value of controller output. What could happen to a dead-time process under proportional control in the event of a gradual load change is plotted in Fig. 1.7.

### Integral (Reset) Control of Dead Time

Proportional control is obviously rejected for most applications demanding a band wider than a few percent. So another control mode is needed. An integral controller is a device whose output is the time integral of the deviation:

$$m = \frac{1}{R} \int e \, dt \tag{1.6}$$

where $R$ is the time constant of the controller, known as "integral" or "reset" time. As long as a deviation exists, this controller will change

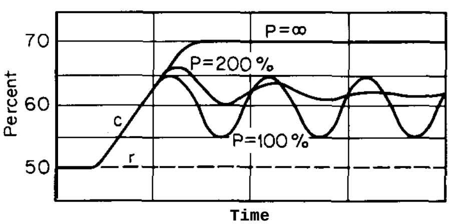
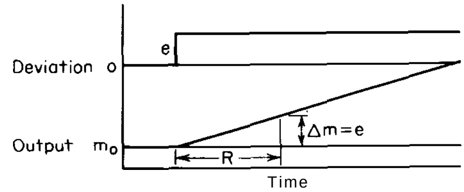



*FIG 1.7. The response to a load change illustrates how the proportional band affects both damping and offset.*

**FIG 1.8.** ***The output of an integrator will change by an amount equal to its input in time R.***




its output, hence it is capable of driving the deviation to zero. The rate of change of output is proportional to the deviation:

$$\frac{dm}{dt} = \frac{e}{R} \tag{1.7}$$

Response to a step input is shown in Fig. 1.8.

Before using an integral controller in a closed loop, its gain and phase characteristics must be defined. Again we are primarily interested in these properties at the natural period of the loop, $\tau_o$. Introducing a sinusoidal input to the controller,

$$e = A \sin 2\pi \frac{t}{\tau_o}$$

The controller output mill be thc time integral of the input:

$$m = \frac{1}{R} \int e\,dt = \frac{1}{R} \int \left( A \sin 2\pi \frac{t}{\tau_o} \right) dt$$

Extraction of the appropriate item from a table of definite integrals enables us to solve the above equation:

$$m = \frac{A\tau_o}{2\pi R} \left( -\cos 2\pi \frac{t}{\tau_o} \right) + m_0$$

where $m_0$ is the output at time zcro.

In order to evaluate phase and gain properties, the output must be reduced to the same form as the input, using the trigonometric identity

$$-\cos x = \sin \left( -\frac{\pi}{2} + x \right)$$

We can convert $m$ into a sine function:

$$m = \frac{A\tau_o}{2\pi R} \sin \left( -\frac{\pi}{2} + \frac{2\pi t}{\tau_o} \right) + m_0$$

The phase shift of the integrator is the angle of the output minus the angle of the input:

$$\phi_R = \left(-\frac{\pi}{2} + \frac{2\pi t}{\tau_o}\right) - \frac{2\pi t}{\tau_o}$$
$$= -\frac{\pi}{2} = -90° \quad (1.8)$$

An integrator exhibits a phase lag of 90° regardless of the period of the input.

The gain of an integrator is the amplitude of the output over the amplitude of the input:

$$G_R = \frac{A\tau_o/2\pi R}{A}$$
$$= \frac{\tau_o}{2\pi R} \quad (1.9)$$

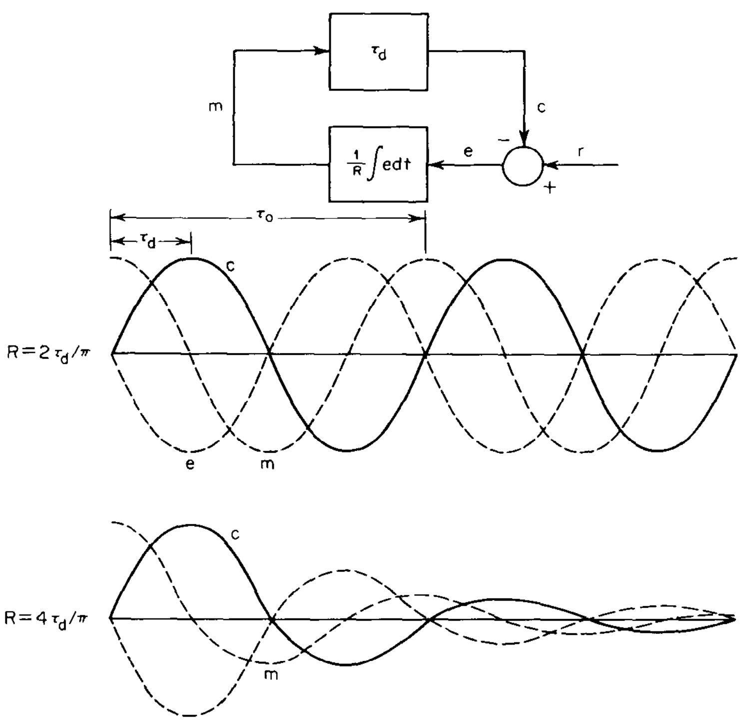



*FIG 1.9. Adjusting reset time affects the damping.*

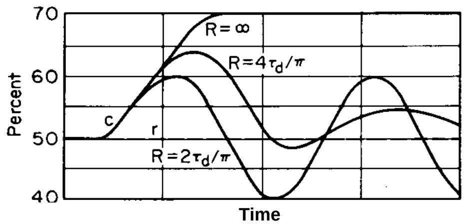



*FIG 1.10. Increasing reset time trades recovery for damping, although $\tau_0$ is unaffected.*

In closing the loop, the sum of the phase shift of the dead time and the integral controller must equal $-\pi$ at the natural period $\tau_o$:

$$-\pi = -\frac{\pi}{2} - \frac{2\pi\tau_d}{\tau_o} \qquad -180^\circ = -90^\circ - 360^\circ \frac{\tau_d}{\tau_o}$$

Solving for $\tau_o$,

$$\tau_o = 4\tau_d \tag{1.10}$$

Notice that the period is twice that for proportional control, because only 90” of phase shift was allowed to take place in the dead-time element.

To sustain oscillations, the loop gain must be 1.0. Since the dead-time gain is already 1.0, the integrator gain for this condition must also be 1.0. Solving for reset time,

$$G_R = \frac{\tau_o}{2\pi R} = 1.0$$
$$R = \frac{\tau_o}{2\pi} = 2\frac{\tau_d}{\pi} \tag{1.11}$$

To summarize, a dead time of 1 min would cycle with a period of 4 min, sustained by a reset time of $2/\pi$, or about 0.65 min. Quarter-amplitude damping can be achieved by halving the gain, which means doubling the reset time. Figure 1.9 shows the entire situation.

Again, the controller has but one adjustment, which only affects damping. The period of oscillation and the integral time for f/l-amplitude damping have been established by the process. Use of the integral controller has avoided the previously encountered proportional offset, but at the cost of reduction in speed of response.

The response of a dead-time process under integral control to a gradual load change is pictured in Fig. 1.10. The rate of recovery is slow when the reset time is too long. With a proper amount of reset, the measurement will cross the set point during the first cycle, exhibiting ¼-amplitude damping.

#### Proportional-plus-reset Control

This controller combines the best features of the proportional and integral modes in that proportional offset is eliminated with little loss

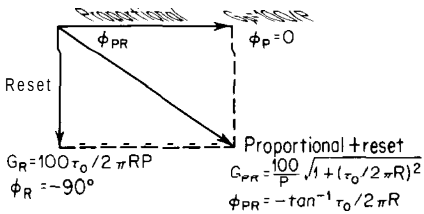



***FIG 1.11. The resultant gain is the square root of the sum of the squares of the components.***

of response speed. The controller is represented as follows:

$$m = \frac{100}{P}\left(e + \frac{1}{R}\int e\,dt\right) \tag{1.12}$$

Having already found the performance characteristics of each of the modes individually on a dead-time process, intuition dictates that the performance of the combination will be somewhere in between, e.g.,

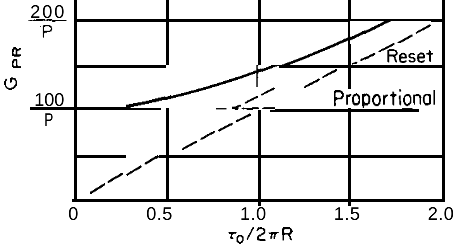
$$4\tau_d > \tau_o > 2\tau_d$$

depending on the particular combination of settings of proportional and reset. An infinite combination of settings can be found to provide constant damping. We have already seen that for ¼-amplitude damping,

$$\frac{100}{P} = 0.5 \qquad \text{or} \qquad \frac{\tau_o}{2\pi R} = 0.5$$

depending on the control mode used. For the two-mode controller, then, the sum of the gains must equal 0.5.

The proportional and integral components of gain are out of phase with each other, however. So their resultant gain must be the vector sum of the two components. Figure 1.11 shows the relationship between the vectors.




***FZG 1.12. A plot of gain vs. $\tau_0$ for the proportional-plus-reset controller shows the contributions of the components.***

**TABLE 1.1 Settings of Proportional and Reset for ¼-amplitude Damping**

| $\phi_R$, deg | $\phi_d$, deg | tan $(-\phi_R)$ | $\tau_o/\tau_d$ | $R/\tau_d$ | *P* |
|---|---|---|---|---|---|
| 0 | -180 | 0.000 | 2.00 | ∞ | 200 |
| -15 | -165 | 0.268 | 2.18 | 1.29 | 206 |
| -30 | -150 | 0.577 | 2.40 | 0.66 | 232 |
| -45 | -135 | 1.000 | 2.67 | 0.42 | 283 |
| -60 | -120 | 1.732 | 3.00 | 0.28 | 400 |
| -75 | -105 | 3.732 | 3.43 | 0.15 | 770 |
| -90 | -90 | ∞ | 4.00 | 1.27* | |

* The last row describes integral-only control.

The gain curve for the proportional-plus-reset controller (Fig. 1.12) can be roughly approximated by the asymptotes:

$$G_{PR} \simeq \frac{100}{P} \left( \frac{\tau_0}{2\pi R} > 1 \right) \tag{1.13}$$

The largest departure occurs at $\tau_o = 2\pi R$, where $G_{PR} = 100\sqrt{2}/P$ versus $100/P$.

This controller presents two adjustments, both of which affect the stability of the loop. An infinite number of combinations of proportional and reset settings exist which would provide ¼-amplitude damping, the only requirement being that $G_{PR} = 0.5$. Several such combinations appear in Table 1.1. Obviously an infinite reset time is undesirable, because offset will result. Yet a very low value of reset forces the proportional band to be set very high, and the controller acts very much like a pure integrator. If the recovery characteristic of reset is to be combined successfully with the higher speed of proportional action, the contribution of each should be similar. Of course, this is not at all critical. Reset time can deviate by 2 : 1 with little change in performance, as long as the proportional band has been adjusted for proper damping. This is typically described as a "trade-off"situation: there is a very broad optimum. Means for determining exact values of the optimum will be given at the beginning of Chap. 4. Figure 1.13 describes the effect of a gradual load change on the loop with proportional-plus-reset control.

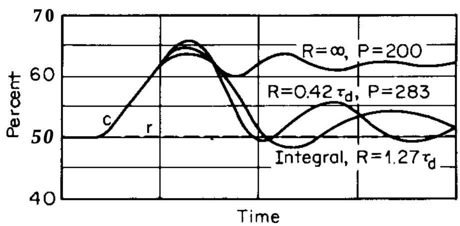



*FIG 1.13. Various combinations of proportional and reset values can provide ¼-amplitude damping, but with different rates of recovery from a load change.*

## THE EASY ELEMENT-CAPACITY

### Identification

Capacity appears in many forms, but its properties are universal as far as automatic control is concerned. Capacity is a location where mass or energy can be stored. It acts as a buffer between inflowing and outflowing streams, determining how fast the level of mass or energy may change. In fluid systems, tanks have capacity to hold liquid or gas. In electrical systems, capacitors are used to store nominal amounts of charge. Heat capacity is a factor in thermal systems. And the mechanical measure of capacitance is inertia, which determines the amount of energy that may be stored in a stationary or a moving object.

Our principal concern is with fluids, so Fig. 1.14 is an appropriate introduction to capacity. In the system shown in the figure, the metering pump delivers a constant outflow, while inflow may be manipulated. The rate of change of tank contents equals the difference between inflow and outflow:

$$\frac{dv}{dt} = F_i - F_o \tag{1.14}$$

Solving for v,

$$v = \int (F_i - F_o)\, dt \tag{1.15}$$

If the tank is vertical and of uniform inside area, its fractional liquid level $h$ will equal the fractional volume:

$$h = \frac{v}{V}$$

where $V$ is the capacity of the tank. Since we are interested in tank level,

$$h = \frac{1}{V} \int (F_i - F_o)\, dt$$

In an effort to make the entire equation dimensionless, we can define $f_i$ and $f_o$ as fractions of the maximum flow $F$ which the valve can deliver.

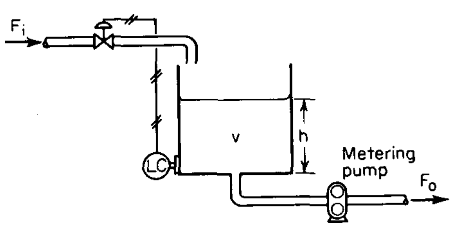



*FIG 1.14. The rate of change of level is proportional to the difference between inflow and outflow.*

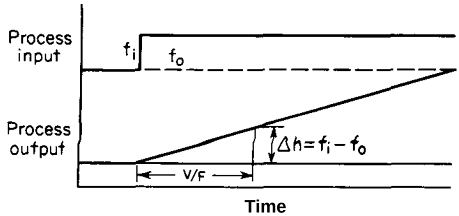



*FIG 1.15. The percent level change will equal the percent Aow change in time V/F.*

Then,

$$F_i - F_o = F(f_i - f_o)$$

and

$$h = \frac{F}{V} \int (f_i - f_o)\, dt \tag{1.16}$$

This is called an integrating process. Notice its similarity to the integrating controllers: $h$ is the output, $f_i - f_o$ is the input error, and $V/F$ is the time constant. The step response is given in Fig. 1.1.5.

The level in the tank could be controlled by manually adjusting the valve position, thereby setting inflow. But if inflow varied in the slightest from outflow, the tank would eventually flood or run dry. This characteristic is called "non-self-regulation." It, means that the integrating process cannot balance itself-it has no natural equilibrium or steady state. The non-self-regulating process cannot be left unattended for long periods of time without automatic control.

Most liquid-level processes are non-self-regulating; occasionally other processes will exhibit this characteristic. In general, it is not harmful as long as its peculiarities are taken into account. One of these peculiarities is its phase shift. Like the integrating controller, the non-self-regulating process exhibits a phase lag of 90° to any periodic wave. Consequently:

1. Under proportional control, the loop cannot oscillate because its phase lag never reaches 180". The proportional band therefore can be set to zero.

2. Under floating (integrating) control, the loop will always oscillate with uniform amplitude, because the total phase shift of process and controller is 180" at all periods. The loop tends to oscillate at the period where the gain product is unity; the reset time then only affects the period and cannot change the damping. The gain of an integrating process is like the integrating controller:

$$G_I = \frac{\tau_o}{2\pi\tau} \tag{1.17}$$

where $\tau = V/F$. If an integrating controller is used to close the loop,

$$G_I G_R = 1.0$$

$$\left(\frac{\tau_o}{2\pi\tau}\right)\left(\frac{\tau_o}{2\pi R}\right) = 1.0$$

Solving for $\tau_o$,

$$\tau_o = 2\pi\sqrt{R\tau} \tag{1.18}$$

### Self-regulation

Replace the metering pump in Fig. 1.14 with a valve. Then an increase in liquid level would inherently increase the outflow. This action works toward the restoration of equilibrium and is called "self-regulation." It is as if a proportional controller were at work within the process. This is a natural form of negative feedback.

Although the relationship is in fact not linear, assume for the moment that flow out of the tank is proportional to the head of liquid above the valve:

$$f_o = kh$$

The level will remain steady when $f_o = f_i$, which indicates that every condition of inflow will bring about a new steady-state level:

$$h = \frac{f_i}{k}$$

In proceeding from one steady state to another, however, the level will vary with time. With a step increase in $f_i$, the level will start to change at the same rate as in the non-self-regulating case, because outflow has not yet begun to increase. The rate of rise of level will then diminish with time, as $f_o$ approaches $f_i$. As a result, the final level will only be reached in infinite time.

$$\frac{dh}{dt} = \frac{F}{V}(f_i - f_o)$$

Substituting for $f_o$,

$$\frac{dh}{dt} = \frac{F}{V}(f_i - kh)$$

The next step is to solve for $h$, the controlled variable:

$$h + \frac{V}{Fk}\frac{dh}{dt} = \frac{f_i}{k} \tag{1.19}$$

This is known as a first-order differential equation. The controlled variable $h$ is related to the manipulated variable $f_i$, both in the steady state and with respect to time. This particular differential equation is of the form

$$c + \tau_1 \frac{dc}{dt} = Km \tag{1.20}$$

which describes a first-order lag whose time constant is $\tau_1$ and whose steady-state gain is K. In the level process, $\tau_1 = V/Fk$ and $K = 1/k$.

The solution of the equation for a step input is

$$c = Km(1 - e^{-t/\tau_1}) \tag{1.21}$$

which is plotted for the level process in Fig. 1.16. After an elapsed time equal to $\tau_1$, 63.2 percent, of the distance to the nest steady state will have been traversed. After another $\tau_1$ has elapsed, G3.2 percent of what was left will have been traversed, and so forth.

At the beginning of the step response, the self-regulating process resembles the non-self-regulating or integrating process. But after sufficient time, it resembles a process without dynamics. The first-order lag is thus made up of two components, one responsive to a fast-changing input, the other responsive to a steady input. This is apparent from examining the differential equation

$$kh + \frac{V}{F}\frac{dh}{dt} = f_i$$

The relation between level and inflow is the sum of two out-of-phase components. The derivative term lends the steady-state term by 90°, just as integrating produced a 90° phase lag. The gain of the derivative term to a signal of period $\tau_o$ is exactly the inverse of the gain of an integrator:

$$G_D = \frac{2\pi V/F}{\tau_o}$$

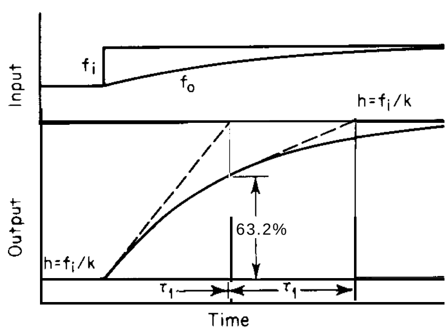



*FIG 1.16. The slope of the response curve equals the departure from steady state divided by $\tau_1$.*

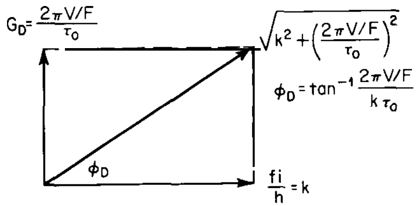



*FIG 1.17. The vector sum is the gain of inflow with respect to level.*

The summation of the two components of **h** with respect to $f_i$ is d grammed in Fig. 1.17. The only difficulty with this vector diagran that the resultant is the ratio of inflow to level. The inverse of resultant represents level vs. inflow, which is the response we are look for:

$$G_1 = \left[k^2 + \left(\frac{2V/F}{\tau_o}\right)^2\right]^{-1/2}$$

The steady-state gain $1/k$ may be broken out separately:

$$G_1 = \frac{1}{k}\left[1 + \left(2\pi\frac{\tau_1}{\tau_o}\right)^2\right]^{-1/2} \qquad (1.2$$

where $\tau_1 = V/Fk$, as before.

A plot of $G_1$ vs. $\tau_o$ in Fig. 1.18 shows a curve which is complementa to that of a proportional-plus-reset controller.

Because we are principally concerned with the dynamic behavior the loop, the asymptote containing $\tau_o$ is of prime importance.

$$G_1 \simeq \frac{\tau_o}{2\pi V/F} < \frac{1}{k} \qquad (1.2$$

Notice that this dynamic-gain asymptote does not contain $k$. In fa it is identical to the gain of the non-self-regulating process. Althoug the steady-state gain can be changed simply by turning the valve at tl bottom of the tank, this does not affect the dynamic gain.

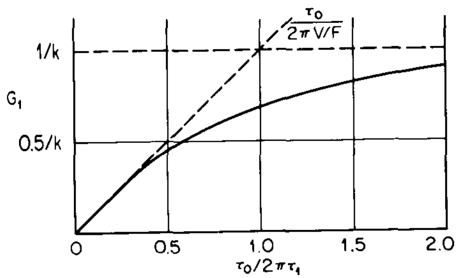



*FIG 1.18. The gain of a first-order lag is governed by the asymptotes.*

This observation is of particular significance for two reasons:

1. Load variations are normally introduced by turning the valve in the outflow line, thus changing $k$.

2. In most processes, including this one, $k$ would not be a constant even if the load were fixed, because the relationship between input and output is not linear. In a real liquid-level process,

$$f_o = C\sqrt{h}$$

where C is the flow coefficient of the valve opening. Then,

$$\frac{h}{f_o} = \frac{\sqrt{h}}{C} = \frac{1}{k}$$

Consequently $k = C/\sqrt{h}$, so even if C remains fixed, $k$ still varies with level. Again, fortunately, this does not affect the dynamic gain.

The time constant $\tau_1$, of such a process is not a constant, but varies with $k$. But this is of little consequence, because the dynamic gain is constant. The ratio ***V/F*** must be recognized as the determining factor. It will appear again and again in different processes, with different forms of variables, but it is the fundamental time constant of any flowing system. Its units are those of time. For example, gal/(gal/min) = minutes.

The phase angle between input and output of a first-order lag is the negative of $\phi_D$ in the vector diagram of Fig. 1.17. As $\tau_o$ approaches zero, $\phi_D$ approaches $+90°$, and therefore the true phase lag approaches 90". In the steady state, however, the vertical vector is zero, hence the phase angle is zero. The phase of a first-order lag is mathematically described as

$$\phi_1 = -\tan^{-1}\frac{2\pi V/F}{\tau_o k}$$

Substituting for $V/Fk$,

$$\phi_1 = -\tan^{-1} 2\pi\frac{\tau_1}{\tau_o} \tag{1.24}$$

Since the phase lag can never exceed $90°$, the first-order lag cannot oscillate under proportional control. This was also true of the integrating process. Therefore we can make a general statement that a single-capacity process can be controlled without oscillation at zero proportional band. This means that the valve will be driven fully open or fully closed on an infinitesimal error, so that the loop is operating at top speed all the time. Since the proportional band is zero, no offset can develop. A single-capacity process must therefore be categorized as the easiest to control.

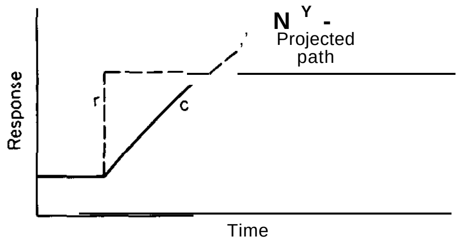



***FIG 1.19. This is how a single-capacity process would react to zero proportional band.***

Figure 1.19 illustrates the set-point response of a single-capacity process to zero proportional band. As soon as the set point is changed, the valve will open wide, delivering maximum inflow. The level will rise as rapidly as possible, which is a function of both $k$ and the present value of level. If no control were provided, the measurement would follow the projected path. But when the new set point is reached, the inflow will be reduced instantaneously to a value equal to the outflow. This assumes that all elements in the loop, excepting the tank, are capable of instantaneous response. If this is not so, the process is not single-capacity.

Examples of pure single-capacity processes are rare. The most common one is a tank being filled through a valve which is rigidly coupled to a float. The level is prevented from overshooting the set point because the rigid coupling eliminates any delay in feedback action.

Whereas the non-self-regulating process cycled uniformly with an integrating controller, the self-regulating will not. The phase shift of the self-regulating process only reaches -90” at a period of zero. As a result, the loop could only oscillate at zero period, where the gain of both process and controller are zero. The loop cannot, therefore, sustain oscillations.

## A Two-capacity Process

Having established the ease with which a single-capacity process may be controlled, the complications involved in adding a second capacity may be evaluated. Since each capacity contributes a phase lag approaching 90°, the total phase lag in the loop can only approach 180”. As a result, the loop can oscillate only at zero period. This is exactly like a first-order lag with an integrating controller.

Adding another lag anywhere in the Ioop will change the previous level process to two-capacity, as shown in Fig. 1.20. A chamber is attached to the tank; although we wish to control tank level, chamber level is measured, which lags behind tank level. The time constant of the chamber is its volume divided by the maximum rate at which liquid can enter. This time constant will be designated $\tau_2$. Control of a two-capacity process is easiest to illustrate if one of the capacities is non-self-regulating.

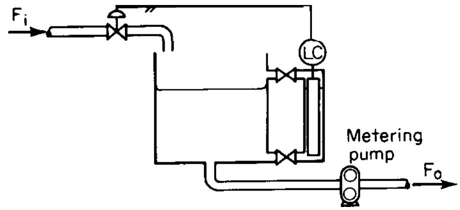



*FIG 1.20. Because the displacement chamber cannot fill instantaneously, it introduces a second capacity.*

So in this example, the metering pump is used as a load, and the time constant for the vessel is $\tau_1 = V/F$.

Let us study the effect of zero proportional band on this process. The set-point response is given in Fig. 1.21. When the measurement is below the set point, the fill valve will be wide open, delivering flow *F*. If the load (outflow) is 50 percent of *F*, the rate of rise of level will be

$$\frac{dh}{dt} = \frac{F}{V}(100 - 50)$$

$$= \frac{50\%}{\tau_1}$$

But the measurement c lags behind the level by $\tau_2$:

$$c + \tau_2 \frac{dc}{dt} = h$$

It can be shown that if $dc/dt$ is constant, it is equal to $dh/dt$. Then

$$h - c = \tau_2 \frac{dh}{dt} = 50\%;$$

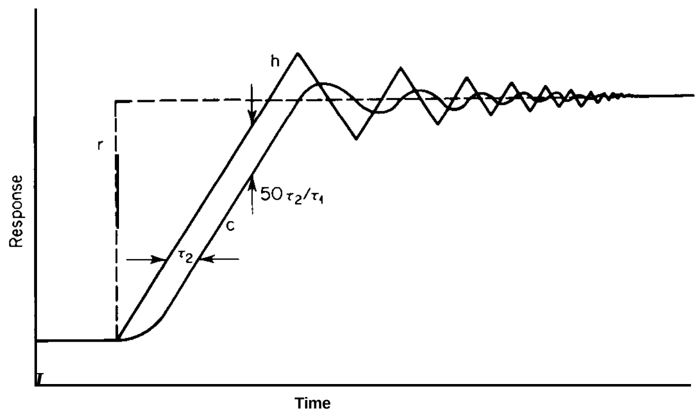



*FIG 1.21. Zero proportional band will cause a two-capacity process to overrun the set point.*

This is the difference in value between the intermediate variable $h$ and the measurement. Their difference in time is simply the amplitude difference divided by the rate of rise:

$$\frac{h - c}{dh/dt} = \tau_2$$

The controller will not close the valve until the measurement reaches the set point. Notice that the intermediate variable has exceeded the set point by $50\tau_2/\tau_1$ at this time. When the valve is shut, outflow will exceed inflow by 50 percent and the level will descend at the same rate. As long as the level is higher than the measurement, the measurement will continue to rise. The measurement will stop rising when it equals the level. The time elapsed between actuation of the controller and the peak of the measurement represents ¼-cycle. From inspection of the figure, this time is somewhere between $0.5\tau_2$ and $\tau_2$ min. It has been calculated at $0.7\tau_2$. This would make the period of the first cycle about $2.5\tau_2$, because the later portions of the cycle are shorter.

Notice that the period is proportional to $\tau_2$, and the amplitude proportional to $\tau_2/\tau_1$. These relationships will appear repeatedly in subsequent examples.

We know from phase and gain characteristics of the process that it cannot sustain oscillations. This means that each cycle must be successively smaller. But because the inflow is either on or off, the rate of change of level is constant for each cycle. Hence, the period must also decrease. Finally the loop oscillates at zero amplitude and zero period as was anticipated. This unusual property is found only in two-capacity processes.

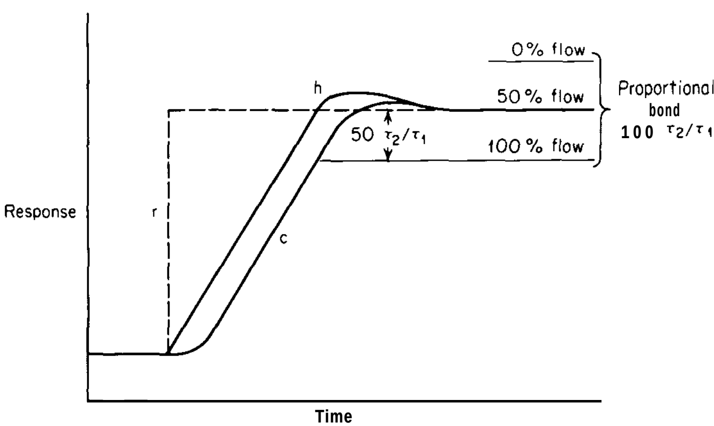



*FIG 1.22. A proportional band of $100\tau_2/\tau_1$ is not wide enough to prevent overshoot.*

## Proportional Control

If overshoot is undesirable, the proportional band must be widened. So that there will be no offset at the normal load, the controller must be biased accordingly. In this example the bias would be 50 percent. When the error is zero, therefore, the inflow will be 50 percent.

With the lower edge of the proportional band $50\tau_2/\tau_1$ percent away from the set point, the tank level will just reach the set point as the valve begins to throttle. This clearly will not prevent overshoot, for the valve will deliver more than 50 percent flow as long as the measurement is below the set point, raising the level farther. In order to bring the level back down to the set point, the measurement must overshoot, so as to reduce the inflow below 50 percent. Consequently a proportional band of $100\tau_2/\tau_1$ ($50\tau_2/\tau_1$ on either side of 50 percent flow) is not wide enough.

In Fig. 1.23 the example is repeated with the proportional band at $200\tau_2/\tau_1$. Throttling begins when the intermediate variable is $50\tau_2/\tau_1$ below the set point, where the rate of rise starts to decrease. This allows the measurement to overtake the tank level, and both will come to rest at the set point. This "no overshoot" characteristic is called "critical damping."

In these examples the load was 50 percent. If the load were instead 80 percent, the rate of rise of level would be only $20\%/\tau_1$. But the controller would be biased by 80 percent, so that only 20 percent of the proportional band would be below the set point. With a band setting of $200\tau_2/\tau_1$, this would leave $40\tau_2/\tau_1$ below the set point. This throttling

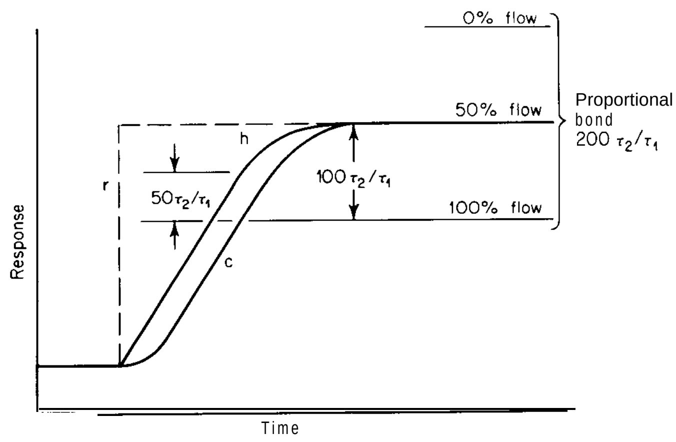



*FIG 1.23. If the proportional band is widened to 200 $\tau_2/\tau_1$, the intermediate variable will not overshoot.*

zone is still twice the difference between tank level and measurement, just as it was at 50 percent load, so the results will be the same. Therefore the proportional band should always be $200\tau_2/\tau_1$ for critical damping, regardless of the load. Only the bias need be changed.

Critical damping makes for sluggish response, however. In most cases, some overshoot is not detrimental. It is important that we determine what *is* necessary to achieve ¼-amplitude damping. Knowing that the period at which the two-capacity loop naturally oscillates is zero, we can be sure that any oscillation at a period of 2.572 will be damped. The period of $2.5\tau_2$ is chosen as it seems to be the natural period of the first cycle (Fig. 1.21). Since we know that oscillations cannot be sustained, let the loop gain at $\tau_o = 2.5\tau_2$ be 1.0:

$$G_1G_2\frac{100}{P} = 1.0$$

Substituting for the dynamic gains of $\tau_1$ and $\tau_2$,

$$\frac{\tau_o}{2\pi\tau_1}\frac{\tau_o}{2\pi\tau_2}\frac{100}{P} = 1.0$$

Substituting $2.5\tau_2$ for $\tau_o$,

$$P = 100\,\frac{2.5\tau_2}{2\pi\tau_1}\frac{2.5\tau_2}{2\pi\tau_2}$$

$$P = 16\frac{\tau_2}{\tau_1} \tag{1.25}$$

This is the proportional band which will produce ¼-amplitude damping. If the method of arriving at these conditions seems somewhat arbitrary, compare the results against those previously established :

| *Damping* | $P$, % of $\tau_2/\tau_1$ |
|---|---|
| Zero . | 0 |
| >& amplitude | 16 |
| Overshoot.. | 100 |
| Critical. | 200 |

The proportional band of $16\tau_2/\tau_1$ fits right in with the rest of the table. Gross changes in $P$ are required to affect the damping of the two-capacity process. It is doubtful whether any difference would be discernible between the response of a loop at 30 percent $\tau_2/\tau_1$ and that at 16 percent. Unfortunately, this is not always so. The two-capacity process has more tolerance for proportional band setting than any more difficult process. Earlier in the chapter it was noted that the damping of the dead-time loop is changed from zero to ¼-amplitude by doubling the proportional band. With the two-capacity process, however, the multiplication is infinite.

Another important factor must be brought out. By definition of the primary and secondary capacities, $\tau_2$ is ***never*** greater than $\tau_1$, regardless of their relative positions in the loop. This means that the most difficult two-capacity process will be one where $\tau_2/\tau_1 = 1.0$. For ¼-amplitude damping, $P$ would be 16 percent. By comparison, the dead-time process is $^{200}\!/_{16}$ or 12.5 times more difficult to control than the most difficult two-capacity process.

Notice also that as $\tau_2$ approaches zero, the process approaches single capacity and $P$ for any damping approaches zero. It is wise therefore, in the design of the process, to make $\tau_2/\tau_1$ as low as possible. Since the natural period of the loop varies as $\tau_2$ only, this should be done by reducing $\tau_2$ where possible, instead of increasing $\tau_1$.

#### Proportional-plus-derivative Control

Adding derivative to a proportional controller relates output to the rate of change of error:

$$m = \frac{100}{P}\left(e + D\frac{de}{dt}\right) + b \tag{1.26}$$

where $D$ is the derivative time. The parenthetic part of this expression is the inverse of a first-order lag-it is called a first-order lead. In the two-capacity-level process,

$$c + \tau_2 \frac{dc}{dt} = h$$

where c is the result of changes in $h$. In the proportional-plus-derivative controller, $m$ is the result of changes in e-the derivative term is on the input side of the equation.

Since $c = r - e$, the lag may be written in terms of e:

$$r - e + \tau_2\left(\frac{dr}{dt} - \frac{de}{dt}\right) = h$$

If the set point is constant, $dr/dt = 0$. Rearranging,

$$e + \tau_2 \frac{de}{dt} = r - h$$

If the derivative time of the controller is set equal to $\tau_2$, the above expression can be substituted into the proportional-plus-derivative controller equation, with the result

$$m = \frac{100}{P}(r - h) + b$$

We now have proportional control of the *intermediate* variable. Adding derivative has caused cancelation of the secondary lag, making the process appear to be single-capacity. In theory, the proportional band may then be reduced to zero and still produce critical damping. In practice, it is not possible.

The gain of a derivative term, $2\pi D/\tau_o$, approaches infinity as the period of the input approaches zero. Noise is a mixture of random periodic signals. A small amount of noise at a high frequency (low period) would be amplified tremendously by a perfect derivative unit. In addition, controllers are made of mechanical or electrical parts that have certain inherent properties of phase lag. Consequently, a high limit is always placed on $G_D$, preventing high-frequency instability within the controller. This high limit is usually about 10. A real derivative unit is actually a combination of a lead whose time constant is $D$ and a lag whose time constant is $D/10$.

In the two-capacity process, then, setting $D = \tau_2$ will not completely cancel $\tau_2$, but will replace it with a lag equal to $\tau_2/10$. The effect is considerable, however, in that the characteristics of the same process under proportional control are improved tenfold. For pi-amplitude damping with proportional-plus-derivative control,

$$P = 1.6\frac{\tau_2}{\tau_1} \qquad D = \tau_2 \qquad \tau_o = 0.25\tau_2 \tag{1.27}$$

Being able to reduce $P$ by 10 also reduces offset by 10. And as a bonus, the loop cycles 10 times as fast as before. Derivative always has this effect, although nowhere else is it so pronounced as in a two-capacity process.

There is one best value of derivative for a given control loop. Too high a setting can be as harmful as none at all. The object is to cancel the secondary lag in the process. If $D > \tau_2$, the controller will lead the intermediate variable, causing premature throttling of the valve. Figure 1.24 shows the effect of three different derivative settings on the same process.

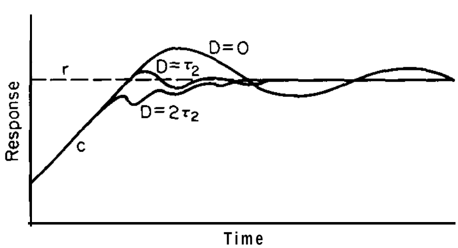



***FIG 1.24. Too much as well as too little derivative degrades the stability of the loop.***

In most controllers, the derivative mode operates on the output rather than on the error. Ordinarily, this presents no problem. But upon startup, or following gross set-point changes, the measurement will be outside the proportional band, causing the output to saturate. If derivative operates on the output, which is steady, rather than on the changing input, it is disabled. Derivative will suddenly be activated again when the measurement reenters the band. So if overshoot is to be avoided upon startup, the band must be wide enough to activate the derivative before the primary variable crosses the set point. The band will have to be at least as wide as that shown in Fig. 1.22:

$$P = 100\frac{\tau_2}{\tau_1} \qquad D = \tau_2 \tag{1.28}$$

In controllers where derivative happens to operate directly on the measurement or error, $P$ should be $\frac{1}{10}$ what was required for proportional control alone, that is, $20\tau_2/\tau_1$.

The reduction in band allowed through the use of derivative can in some applications eliminate the need for reset. If a choice between derivative and reset should ever be presented, the former should be selected because it can enhance both speed and stability at the same time.

## COMBINATIONS OF DEAD TIME AND CAPACITY

Occurrences of either pure dead-time or ideal single-capacity processes are rare. The reasons for this are twofold:

1. Mass has the capability of storing energy.
2. Mass cannot be transported anywhere in zero time.

Between the most and least difficult elements lies a broad spectrum of moderately difficult processes. Although most of these processes are dynamically complex, their behavior can be modeled, to a large extent, by a combination of dead time plus single capacity. The proportional band required to critically damp a single-capacity process is zero. For a dead-time process, it is infinite. It would appear, then, that the proportional band requirement is related to the dead time in a process, divided by its time constant. Any proportional band, hence any process, would fit somewhere in this spectrum of processes. A discussion of multicapacity processes in Chap. 2 will reaffirm this point.

### Proportional Control

Fortunately we already investigated this problem when we discussed integral control of dead time. Figure 1.25 indicates the similarity of the loops. If the process is non-self-regulating (integrating), the representation is exact. Because the phase lag of the dead time is limited to 90°, the period of the proportional loop is $4\tau_d$. In the former case, for $\frac{1}{4}$-amplitude damping, $2\tau_d/\pi R$ was set equal to 0.5. Since the time constant $R$ is no longer adjustable, but is now $\tau_1$, part of the process, proportional adjustment must set the loop gain for $\frac{1}{4}$-amplitude damping. Therefore,

$$\frac{2\tau_d}{\pi\tau_1}\frac{100}{P} = 0.5$$

$$P = 400\frac{\tau_d}{\pi\tau_1} \tag{1.29}$$

Notice that as $\tau_1$ approaches zero, $P$ approaches infinity. This is much worse than having no capacity at all, i.e., dead time alone. The reason is that this expression holds only for a non-self-regulating process whose gain varies inversely with the time constant without limit. Fortunately, non-self-regulating processes dominated by dead time are virtually nonexistent.

For the self-regulating process, gain is limited to that of the steady state, nominally 1.0. (Actual contributions of steady-state gain will be evaluated at length in the next chapter.) If the maximum gain of the self-regulating process is 1.0, the proportional band required for $\frac{1}{4}$-amplitude damping with dead time in the loop will approach 200 percent as $\tau_1$ approaches zero. The proportional band setting can then be approximated by the asymptotes:

$$P \simeq \frac{400\tau_d}{\pi\tau_1} > 200\% \tag{1.30}$$

In Fig. 1.26, the locus of gain, $G_1$, of the capacity, and $P$ for $\frac{1}{4}$-amplitude damping are plotted vs. $\tau_d/\tau_1$; the asymptotes are indicated.

A point midway between the asymptotes is found where the phase contribution of $\tau_1$ is 45”. This occurs where $\tau_o = 2\pi\tau_1$. Here 135” of phase

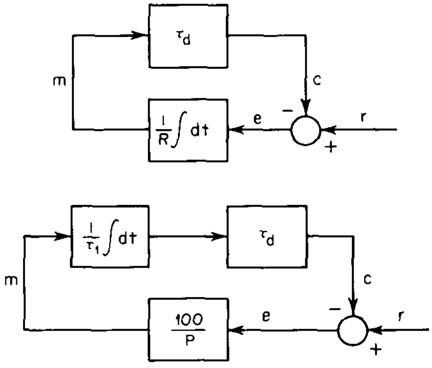



*FIG 1.25. Zntegral control of dead time (aboue) is the same as proportional control of a dead-time plus integrating process (below).*

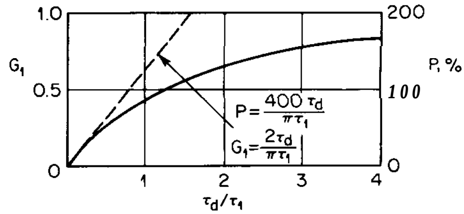



*FIG 1.26. The proportional band required for ¼-amplitude damping for any combination of dead time and capacity can be selected from this chart.*

shift takes place in the dead-time element. As a result,

$$\tau_o = 2.67\tau_d$$

Substituting,

$$\tau_d = 2\pi \frac{\tau_1}{2.67} = 2.35\tau_1$$

This point lies on the abscissa of Fig. 1.26, at $\tau_d/\tau_1 = 2.35$. It may be recalled that the gain of a first-order lag at $\tau_o = 2\pi\tau_1$ is $1/\sqrt{2}$. If the loop is to be damped,

$$G_1 \frac{100}{P} = 0.5$$

Therefore,

$$P = \frac{200}{\sqrt{2}} = 100\sqrt{2} = 141$$

It is interesting to note the comparison between the controllability of this process and the two-capacity process. Taken on the basis of an equal ratio of secondary to primary element, the dead-time plus capacity process is $400/\pi 16$ or 8 times as difficult to control. Recall that the pure dead-time process was 12.5 times as difficult to control as the most difficult two-capacity process.

**The Effect of Derivative**

Derivative is the inverse of integral action. In theory, it is characterized by a 90” phase lead, although because of physical limitations 45° is about all that can be expected. If perfect derivative (90” lead) were available, it could halve the period of the dead-time plus capacity loop by allowing the dcnd time to contribute all 180°. Remember that perfect derivative applied to the two-capacity process provided critical damping with zero proportional band. But Fig. 1.27 indicates that perfect derivative is limited to zero damping at a period of $2\tau_d$ with zero proportional band.

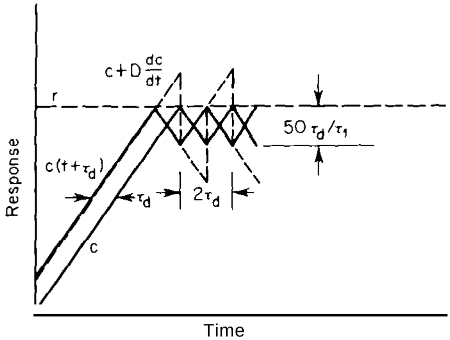



***FIG 1.27. Perfect derivative cannot overcome dead time as it does a second capacity.***

Derivative is lead action, which has been described as the inverse of lag. That is why it could nearly cancel the effect of the secondary lag in a two-capacity process. But derivative is not the inverse of dead time. Nothing is, since no one can make time. Derivative is a poor substitute and consequently is only partially effective in improving the performance of the loop. Limited to 45° of phase lead, furthermore, proportional-plus-derivative controllers can only reduce the period of a dead-time plus capacity loop to $2.67\tau_d$.

As pointed out earlier, derivative contributes gain as well as phase lead:

$$G_D = 2\pi \frac{D}{\tau_o}$$

Since the gain of the process capacity decreases at the same rate,

$$G_1 = \frac{\tau_o}{2\pi\tau_1}$$

Reducing $\tau_o$ produces no net change in loop gain. Consequently, adding derivative does not allow a reduction in proportional band, as it did with the two-capacity process. Thus derivative is scarcely effective at all in the presence of dead time.

The derivative mode exhibits a phase lead of 45° at $\tau_o = 2aD$. To take advantage of this lead, the derivative time should be set to locate this phase lead at the period of the loop after derivative has been added ($2.67\tau_d$):

$$2\pi D = 2.67\tau_d$$

For ¼-amplitude damping,

$$P = 400 \frac{\tau_d}{\pi\tau_1} \qquad D = 1.33 \frac{\tau_d}{\pi} \tag{1.31}$$

This derivative setting is contrasted with that recommended for the two-capacity process, that is, $D = \tau_2$.

## SUMMARY

A careful reading of this chapter should disclose the dependence of control performance on what have been termed the secondary dynamic elements in the loop. The largest time constant has been defined as the primary element, and all others as secondary.

The term "difficult" has been used to describe control of certain processes. The proportional band required for a particular damping serves as an index of difficulty. There is good reason for this, for the proportional band is a measure of how much influence a controller has over a process. The derivation of proportional offset bears out this relationship. If the proportional band is 100 percent, the controller and the load have equal influence over the controlled variable. At 200 percent band, the load has twice as much influence. Figure 1.7 is a good illustration.

Control problems of principal interest are those involving two dynamic elements. Loops comprised of only one element are nothing more than limits of two-element loops. The difficulty of each of these processes is found to be proportional to the ratio of the secondary to the primary element. In addition, the period of the closed loop is a function of the secondary element alone. A performance index can be envisioned which would combine the sensitivity of the loop to disturbances with the time required to recover from them. This index would vary as the square of the secondary element. The significance of secondary elements is paramount.

Settings of reset and derivative time are also directly related to the value of the secondary element. This rule seems as illogical as that governing the period of the pendulum, which varies with length, not with mass. Visualize length as the secondary element and mass as the primary, as a memory aid.

Hopefully, the reader has observed how the open-loop characteristics of a process determine its closed-loop response. And how little influence the controller has over this response. It is particularly true for processes of increasing difficulty, where problems begin to appear.

## PROBLEMS

**1.1** The belt speed of the process described in Fig. 1.2 is 12 ft/min, and the weigh cell is located 4 ft from the valve. Estimate the natural period under integral control and the reset, time required for ¼-amplitude damping. Is this setting likely to be conservative? Why?

**1.2** The same process is to be controlled with a proportional-plus-reset controller, adjusted for a reset phase lag of 60°. Calculate the settings required for %-amplitude damping, and check your answer against Table 1.1.

1.3 Figure 1.17 is an inverse vector diagram of a first-order lag. Construct a true vector diagram, indicating the magnitude and phase angle of each vector.

1.4 Construct a vector diagram for the proportional-plus-derivative controller described by Eq. (1.26). Indicate the magnitude and phase angle of each vector.

1.5 Calculate the gain of a dead-time plus single-capacity process whose natural period under proportional control is 3.0 $\tau_d$. What is the ratio of $\tau_d/\tau_1$? Does this point fall on the curve of Fig. 1.26?

1.6 A certain process consists of a 1-min dead time and a 30-min lag. Estimate the period and settings for ¼-amplitude damping under proportional-plus-derivative control. Repeat for a proportional-plus-reset controller, assuming 45” phase lag in the controller.

# Characteristics of Real Processes

CHAPTER 2

As pointed out in Chap. 1, it is doubtful whether any real process consists exclusively of dead time or single capacity or even a combination of the two. But having become familiar with the properties of these elements, we now can proceed to identify their contributions to complex processes. Some processes are difficult to control-particularly where dead time is dominant. But many processes are poorly controlled because their needs are not understood and therefore not satisfied.

Real processes consist of a combination of dynamic elements and steady-state elements. When there are many dynamic elements present, their combined effect is hard to visualize. Even worse, one or more of these elements may be variable. The same is true for steady-state elements. In fact, one could venture to say that many engineers have less comprehension of the steady-state relationships in a complex process than of the dynamic properties. This chapter is devoted to identifying these characteristics for the general case and to putting them into a form in which they can be readily recognized and handled.

Nonlinearities naturally occurring in the process are cause for grave concern. Most processes are nonlinear in some respect. Identification of the source and nature of a nonlinearity is of the utmost significance. Whether it is severe enough to be troublesome and how the trouble can be corrected are important questions which will be answered in the pages that follow. General rules and methods will be stipulated, with a concrete example to illustrate each point. Many more cases will bc cited in later chapters as part of specific applications.

It is especially important to keep in mind the prominence of nonlinear characteristics when studying an unfamiliar process; the engineer must know what to look for and what to expect. Tests improperly conducted can give results that are meaningless, confusing, or altogether misleading. The full significance of the "characteristics of real processes" must bc appreciated before an intelligent program of testing and evaluation can be undertaken.

## MULTICAPACITY PROCESSES

### Interaction

The principal distinction to be made in multicapacity processes is the manner in which the capacities are joined. If they are said to be isolated or noninteracting, the capacities behave exactly as they would alone. Rut if coupled, they interact with one another, in which case the contribution of each is altered by the interaction. Figure 2.1 compares the two forms.

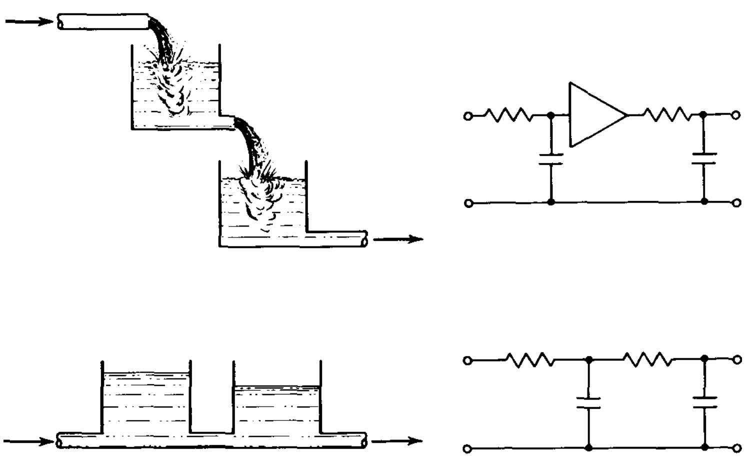


*FIG 2.1. Noninteracting (above) compared to interacting (below) capacities.*

In the upper left, the two tank levels do not interact, because the flow from the first to the second is independent of the level in the second. The lower picture, however, illustrates the case where both. inflow and outflow are a function of tank level. The levels interact, because any change in the downstream level will affect the upstream level. An electrical analog of each process appears on the right. The amplifier in the upper figure isolates the two lags by preventing the voltage on the second from affecting that on the first. The lower right figure, without the amplifier, is a two-stage ladder network. A multistage ladder is often used to simulate a transmission line.

The significance of interaction is that. it changes the effective time constants of the individual capacities. The magnitude of the change is striking. The solution of the equation for determining the effective time constants is irrational and unfortunately somewhat unclear.' But for the special case where two equal capacities with equal time constants $\tau$ interact, their combined response is that of two noninteracting lags with the following values:

$$\tau_{1\,2} = \frac{3 \pm \sqrt{5}}{2}\tau$$

$$\tau_1 = 2.6187 \qquad \text{and} \qquad \tau_2 = 0.382\tau$$

The following general rules apply to the principle of interaction:

1. The degree of interaction is proportional to the ratio of the smaller to the larger capacity (not time constant). Where this ratio is low ($<0.1$), the capacities may be assumed not to interact'.

2. Interaction always works toward increasing the larger time constant and decreasing the smaller one.

3. Specifically with regard to the behavior of systems with equal time constants $\tau$, of equal capacity, the effect. is a combination of one large and the rest small time constants whose normalized sum is

$$\sum_{i=1}^{i=n} \frac{\tau_i}{\tau} = \sum_{i=1}^{i=n} i = \frac{n^2 + n}{2} \qquad (2.1)$$

where $i$ = each time constant'

$n$ = number of capacities

and whose normalized product is

$$\prod_{i=1}^{i=n} \frac{\tau_i}{\tau} = 1.0 \qquad (2.2)$$

A case in point is the two-capacity process cited above:

$$\frac{\tau_1 + \tau_2}{\tau} = 2.618 + 0.882 = 3$$

Since $n = 2$,

$$\frac{n^2 + n}{2} = \frac{4 + 2}{2} = 3 \qquad \text{and} \qquad \frac{\tau_1}{\tau}\frac{\tau_2}{\tau} = (2.618)(0.382) = 1.00$$

With three capacities of time constant $\tau$,

$$\frac{\tau_1}{\tau} = 5.0503$$

$$\frac{\tau_2}{\tau} = 0.6403$$

$$\frac{\tau_3}{\tau} = 0.3090$$

$$\overline{6.0000}$$

and (5.0505) (0.6405) (0.3090) = 1.00.

The reasons for interaction can be visualized to some extent. For example, in the interacting tanks of Fig. 2.1, the flow entering the first tank must ultimately fill both tanks, whereas that entering the second fills the second. The sum of the time constants then becomes three.

An important point to grasp is that interaction makes control easier. Recall that the proportional band required to regulate a two-capacity process varies with $\tau_2/\tau_1$, with the most difficult case being $\tau_2 = \tau_1$. Where capacities interact, however, it is impossible to make $\tau_2 = \tau_1$. The ratio of two equal interacting time constants is 0.382 /2.618 = 0.146. By this standard the noninteracting process is nearly seven times more difficult to control!

At this poiiit it is worthwhile to review the examples of two-capacity processes that have already been presented, from the aspect of interaction. The process shown in Fig. 1.20 is definitely interacting, because changes in chamber level can cause changes in tank level. Hut the capacity of the chamber is so much less than that of the tank, that the effect is virtually nil. For practical purposes, then, the two capacities may be considered noninteracting. It would even be possible to throttle the valves to the chamber enough to make its time constant equal or exceed that of the tank. This property is not representative of interacting systems.

But where the principal flow passes through coupled capacities, interaction is manifest. This is thr case iii the lower pair of tanks in Fig. 2.1. If each of these tanks has a volume $V$ and a discharge flow coefficient $k$, the response of level in the second tank to variations in don- in the first will be characterized by a steady-state gain of 1 $k$ and time constants of $2.618V/Fk$ and $0.382V/Fk$. It may be recalled that the time constant of the individual tanks was $V/Fk$. The steady-state level in the second tank equals $f_i/k$. The steady-state level in the first tank would be $2f_i/k$

because it is discharging into a level that is already $f_i/k$. Consequently, any change of inflow will change the combined steady-state levels by a factor of $3/k$. The volume change is therefore three times what it was for a single tank-that is why the sum of the time constants is $3V/Fk$, whereas the total volume of the system is only $2V$. The dynamic gain of the process is approximated as

$$G_1G_2 \simeq \left(\frac{\tau_o}{2\pi 2.618V/F} < \frac{1}{k}\right)\left(\frac{\tau_o}{2\pi 0.382V/Fk} < 1.0\right)$$

**Multicapacity**

In a one-capacity process, interaction does not exist. The effect of interaction on a two-capacity process has already been demonstrated. As the number of capacities increases, this effect becomes more pronounced.

The behavior of $n$ equal isolated capacities of time constant $\tau$ can be estimated from phase relations. If the phase of each lag is

$$\phi = -\tan^{-1} 2\pi \frac{\tau}{\tau_o}$$

the total phase shift is $n\phi$:

$$n\phi = -n \tan^{-1} 2\pi \frac{\tau}{\tau_o}$$

We are never concerned with phase shift in excess of 180°, at which point $\phi = -\pi/n$. If $n$ is large, $\phi$ is quite small. The tangent of a small angle is approximately equal to the angle:

$$\lim_{\phi \to 0}\left(-\tan^{-1} 2\pi \frac{\tau}{\tau_o}\right) = -2\pi \frac{\tau}{\tau_o}$$

Stated a little differently,

$$\lim_{n \to \infty} n\phi = -2\pi n \frac{\tau}{\tau_o} \qquad (2.3)$$

This indicates that a large number $n$ of isolated lags $\tau$ approaches the same phase characteristic as dead time of value $n\tau$.

The same is not true in the interacting case, because there remains one very large time constant and successively smaller ones. The large time constant is always so much larger than the others, that it dominates the response. The small time constants begin to approach dead time, however, because their values are close together. The result appears equivalent to a single-capacity plus dead-time process. The step response of comparable isolated and interacting systems appears in Fig. 2.2.

A process with many isolated capacities is artificial, because isolation must be intentionally forced. Witness the amplifier in Fig. 2.1. As a general rule, multicapacity processes contain a natural interaction, responding in the manner of the lower set of curves in Fig. 2.2. This form of response is evident both in processes consisting of a large number of discrete stages and in those embodying a continuum of distributed particles. Examples of multistage processes are plate columns for distillation, extraction, and absorption. Counterflow of the two phases produces the interaction. Packed columns, on the other hand, are distributed systems which behave similarly. Diffusive processes such as heat transfer by conduction, mixing in pipes and vessels, and flow through porous media react in much the same manner. More attention will be devoted to these operations when specific applications are investigated.

From Fig. 2.2 it can be seen that the interacting multicapacity process differs from the dead-time plus single-capacity process in the smooth upturn at the beginning of the step response. This curvature indicates that the dead time is not pure, but instead is the result of many small lags, and therefore the process will be somewhat easier to control. By the same token, derivative action will be of more value than it was in the case of dead time and a single capacity. Nonetheless, if we choose to estimate the necessary controller settings on the basis of a single-capacity plus dead-time representation we will err on the safe side.

The natural period of the loop can be predicted with surprising reliability by noting where the maximum slope of the step-response curve intersects the time base. This intersection, marked in Fig. 2.3, identifies the effective dead time of the process. The effective dead time plus the

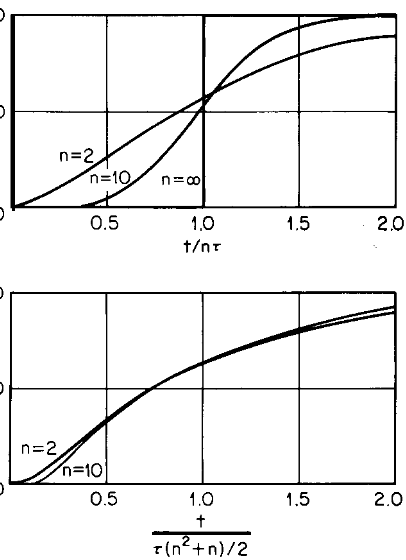



***FIG 2.2. The difference in step response between isolated (above) and interacting (below) lags becomes more pronounced as n increases.***

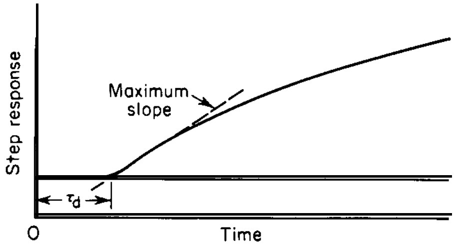



***FIG 2.3. The step response of a multicapacity process can be reduced to dead time plus a single capacity.***

effective time constant equals the total lag in the process:

$$\tau_d + \tau_1 = \tau \frac{n^2 + n}{2} \qquad (2.4)$$

Equation (2.4) requires that the step response of any number of equal interacting lags reach 63.2 percent at time $\tau(n^2 + n)/2$, which is corroborated by Fig. 2.2.

Ziegler and Nichols[2] noted that the period of oscillation will be four times the effective dead time, whether the process is interacting or not. So the technique of dealing with single capacity plus dead time takes on added value in being applicable to these examples of complex dynamics. This is an important insight-without it, numerical methods must be rejected for use on any process containing more than two dynamic elements. And when on-the-spot analysis must be made, the shortcut numerical method is invaluable. Fortunately, a single-capacity plus dead-time process can be made to represent any degree of difficulty from one extreme to the other, simply by varying the ratio $\tau_d/\tau_1$. Thus its application is universal, if approximate.

As an example, the lo-capacity interacting process of Fig. 2.2 has an effective dead time of about 0.15 of the total lag. Since the balance is the dominant time constant, the ratio $\tau_d/\tau_1 = 0.15/0.85 = 0.18$. The proportional band needed for ¼-amplitude damping of this process can be found by referring to Fig. 1.26.

Figure 2.4 is a correlation of the ratio of effective dead time to effective lag, against the number of interacting stages. Data from tests on systems of 2 to 10 capacities fall in a straight line on semilogarithmic coordi-

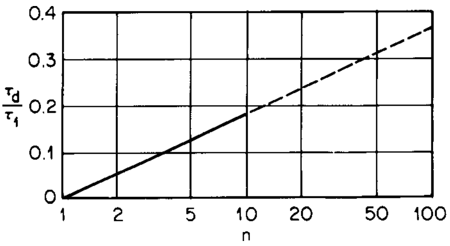



***FIG 2.4. The ratio of effective dead time to effective lag of n equal interacting capacities varies with the logarithm of n.***

nates. This relationship is extremely useful in predicting the dynamic behavior of any process with a discrete number of interacting stages.

Diffusive and distributed processes ought to consist of an infinite number of interacting stages. Their response does not correspond to $n = \infty$ in Fig. 2.4, however, probably because their interaction is incomplete. Transmission lines typically exhibit ratios of $\tau_d/\tau_1$ in the range from 0.1 to 0.3.[3]

### Predicting the Behavior of a Loop

The appearance of a piece of processing equipment often reveals the nature of its dynamic characteristics. If all the dimensions are similar, as in a cylindrical tank where the height is of the same magnitude as the diameter, capacity will predominate—dead time, if any, being short. But if the vessel has one dimension much greater than the others, dead time may be dominant, though not without some capacity. Thus a shell and tube heat exchanger will exhibit considerable dead time, compared to a heated tank whose principal elements would be lags. Just the appearance of a tower, whether it be distillation, absorption, or whatever, indicates the presence of dead time.

One could almost generalize to the extent of relating controllability to dimension:

$$\frac{\tau_d}{\tau_1} = f\left(\frac{\text{length}}{\text{diameter}}\right)$$

Of course such an expression could only he written to apply within a specific system, because many more factors are involved. Nonetheless, if dead time is related to length, the natural period is similarly related to length, as with a simple pendulum.

## GAIN AND ITS DEPENDENCE

The damping of a feedback loop is a function of the gain product of all the elements in the loop, both dynamic and steady-state. Normally only one of these elements is adjustable-the controller. All others are fixed by the design of the process. For a given damping, the controller adjust'-ments are a function of the gain of the fixed elements. Up to this point, only dynamic gain has been considered. But any element whose output differs from its input has a gain contribution:

| Element | Input | output |
|---|---|---|
| Valve. | Signal | Flow |
| Process, | Flow | Measurement |
| Transmitter, | Measurement | Signal |

Each of the three elements above changes the dimension of what is passing through. In order to arrive at a dimensionless loop gain, the dimensional gain of all three elements must be included in the product.

**Transmitter Gain**

In the liquid-level process presented in Chap. 1, the measurement $h$ was defined as representing the fractional contents of the tank. This trick enabled us to find the time constant of the vessel in terms of its capacity $V$ and its nominal throughput F. When instrumenting a plant, however, it is not necessary that every liquid-level transmitter be scaled to measure the entire volume of the vessel. If, instead, the transmitter span represents only a small percentage of the vessel volume, the vessel will have effectively shrunk to the span of the transmitter. To state it another way: for control purposes, those parts of the vessel beyond the range of the transmitter do not exist.

Reducing the span of a transmitter is equivalent to reducing the proportional band of the controller. If a particular damping, hence a particular loop gain is to be achieved, the proportional band estimate must, take into account the span of the transmitter.

In order to facilitate the evaluation of systems more complex than the liquid-level process, the transmitter gain will be explicitly defined:

$$G_T = \frac{100\%}{\text{span}} \tag{2.5}$$

Gain is the ratio of output to input. The numerator in Eq. (2.5) is the output that will be produced for a full-span change in input. Obviously $G_T$ is not a pure number-it has the dimensions of the measurement. Suppose a level transmitter were calibrated to a range of 20 to 100 in. of water. Its gain would then be 100%/80 in., or 1.25%/in.

The fact that $G_T$ has dimension indicates that it is an incomplete term. The other "gains" around the loop must be multiplied by $G_T$ in order to make the loop gain dimensionless.

It is entirely possible that $G_T$ is not constant. This would be the case if the transmitter were nonlinear. Few transmitters are sufficiently nonlinear to show any marked effect on control-loop stability. A change in gain of at least 1.5/1 would be necessary to cause difficulty. Some temperature measurements are nonlinear, but seldom to this extent. The most, notable case of a nonlinenr transmitter is the differential flowmeter, whose output varies with the square of flow through the primary element.

Each flow transmitter has its own particular span. But in addition, the differential flow transmitter has the nonlinear relationship. $G_T$ can be determined on the basis of transmitter span, with the nonlinearity applied as a coefficient. Let $h$ = dimensionless differential (fraction of

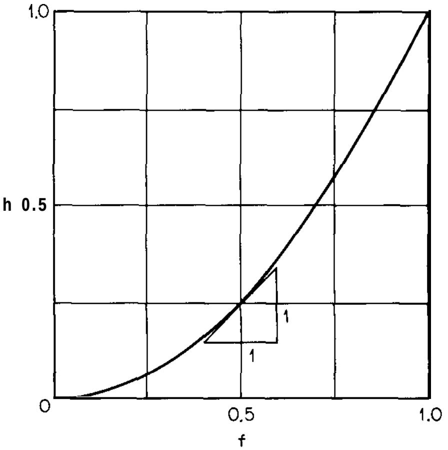



*FIG 2.5. The gain of a differential flow transmitter is directly proportional to flow.*

full scale) and $f$ = dimensionless flow. Then,

$$h = f^2 \tag{2.6}$$

and

$$\frac{dh}{df} = 2f \tag{2.7}$$

The derivative $dh/df$ is the dimensionless gain of the transmitter. Its dimensional gain is then

$$G_T = 2f \frac{100\%}{\text{span}} \tag{2.8}$$

As an example, look at a differential flowmeter whose scale is 0 to 500 gpm. $G_T = 2f(0.2\%/\text{gpm})$. At full-scale flow, $G_T = 0.4\%/\text{gpm}$; at 50 percent flow, $G_T = 0.2\%/\text{gpm}$; at zero flow, $G_T = 0$.

The result of the nonlinearity is that the control loop will not perform consistently at different, rates of flow. If the proportional band is adjusted for acceptable damping at 50 percent flow, the loop n-ill be undamped at 100 percent flow and sluggish near zero flow. The problem can be readily resolved, however, by inserting a square-root extractor, whose output would be linear with flow.

## Valve Gain

Referring again to the liquid-level process of Chap. 1, the time constant of the vessel was based upon the rated flow $F$ which the control valve was capable of delivering. The time constant thus depends on valve size; consequently, the proportional band is a function of valve size. Looking at it another way, an oversize valve would only be operated over part of its travel-the span of stem travel would be less than 100 percent. Therefore the proportional band must be wider to compensate.

The gain of a valve can be defined as the change in delivered flow vs. percent change in stem position. The gain of a linear valve is simply the rated flow under nominal process conditions at full stroke:

$$G_v = \frac{\text{maximum flow}}{100\%} \tag{2.9}$$

If a linear valve were able to deliver 500 gpm fully open at stipulated process conditions, $G_v$ would be 5 gpm/%. Notice that valve gain has dimension, as did transmitter gain, but now the percent sign is in the denominator. The valve is at the output of the controller, whereas the transmitter is at the input. Controller gain is therefore in terms of %/%, hence dimensionless.

Valves cannot be manufactured to the same tolerance as transmitters. So there is no such thing as a truly linear valve. But perfect linearity is not essential, because a control loop does not demand it. Some valves are deliberately characterized to particular nonlinear functions, in order to better carry out certain specific duties. The most commonly used characterized valve is the equal-percentage type.

The name "equal-percentage" is very subject to misinterpretation. It means that a given increment in stem position will change the flow by a certain percentage of the present flow, regardless of the value of the present flow. To state it mathematically,

$$\frac{df}{f} = K\, dm \tag{2.10}$$

where $m$ = fractional stem position
$K$ = constant

This constant for a typical equal-percentage valve is about 4. The dimensionless gain of the valve then becomes

$$\frac{df}{dm} = 4f \tag{2.11}$$

Combining this with the maximum rated flow of the valve yields

$$G_v = 4f\,\frac{\text{maximum flow}}{100\%} \tag{2.12}$$

The equal-percentage characteristic has an interesting feature: changing the valve size does not affect the loop gain! Fractional flow $f$ times the maximum flow equals the actual flow being delivered. The valve

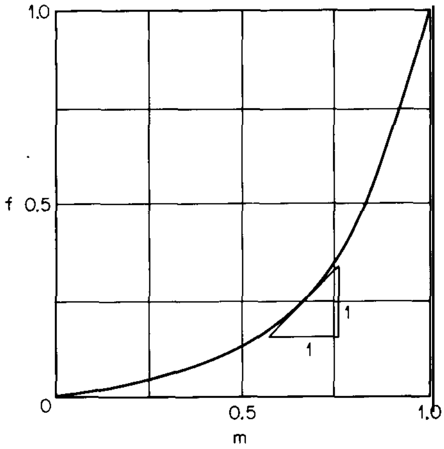



*FIG 2.6. The gain of an equal-percentage valve is directly proportional to flow.*

gain is consequently a function of the actual flow and has nothing to do with valve size. This is one reason why equal-percentage valves are used extensively-valve sizing is not critical.

Sometimes this valve characteristic is shown as a straight line on semi-logarithmic coordinates. Integration of the differential expression for valve gain does yield the logarithmic expression

$$- \ln f = 4(1 - m) \tag{2.13}$$

Other common valve types include the quick-opening (globe) and the butterfly. Their characteristics are more a result of accident than design. In other words, the nature of their construction is the reason for their characteristics, not vice versa. They will generally be employed regardless of, not because of, their nonlinearity.

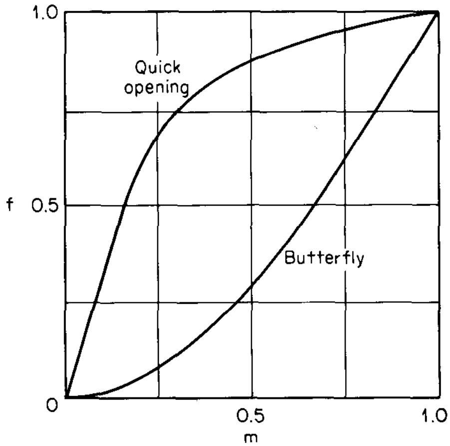



*FIG 2.7. Characteristics of quick-opening and butterfly valves.*

### Variable Pressure Drop

The equal-percentage valve was intentionally devised to compensate for gain variations elsewhere in the control loop. The nearest source of variable gain is the pressure loss in piping and fittings in series with the valve. The flow of liquid which a valve can deliver is as much a function of available pressure as it is of valve opening:

$$F = C_v \sqrt{\frac{\Delta p}{\rho}} \tag{2.14}$$

where $C_v$ = flow coefficient of valve
Ap = drop across valve
$\rho$ = specific gravity of flowing fluid

Should Ap change, $F$ will also change. Choice of the valve is therefore intimately connected with the associated piping and motive force for the fluid. If Ap is constant, the valve will exhibit its inherent characteristic, but if Ap varies with flow, the relation between flow and stroke will change.

Figure 2.8 typifies an arrangement where a constant pressure source drives fluid through a fixed resistance whose flow coefficient is $C_R$ and through a valve of variable resistance whose flow coefficient is $C_v$ into a sink of constant pressure. As flow approaches zero, the valve will be nearly closed and the entire pressure drop in the system will exist across the valve. But at maximum flow, particularly if $C_R$ is less than $C_v$, the drop across the valve will be reduced markedly. If the valve is linear, the gain of the system will be high at low rates of flow (high $\Delta p$) and low at high rates (low Ap).

Since the gain of the valve has been defined as flow in response to stem position, the gain of the valve is affected by variable pressure drop. If we consider a liquid of specific gravity 1.0,

$$F^2 = C_v^2(p_1 - p_2) = C_R^2(p_0 - p_1)$$

Eliminating the variable pressure $p_1$,

$$F^2 = \frac{p_0 - p_2}{1/C_v^2 + 1/C_R^2}$$

Let $C_v$ represent the maximum valve opening. Then, if the valve is linear, its opening will be $mC_v$, where $m$ is the fractional stem position.

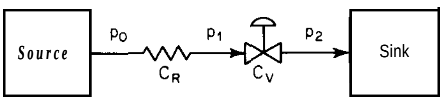



***FIG 2.8. Pressure drop across the control valve depends on losses through the series resistance.***

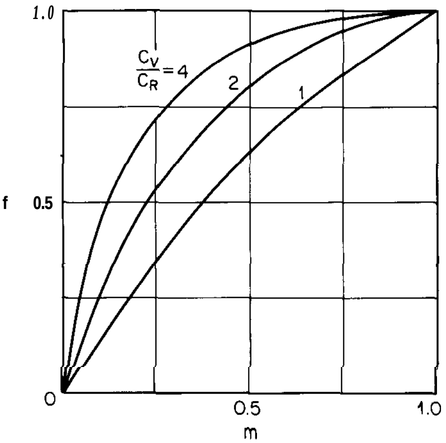



*FIG 2.9. Line resistance distorts a linear characteristic toward that of a quick-opening valve.*

If we include stem position in the above,

$$(fF)^2 = \frac{p_0 - p_2}{(1/mC_v)^2 + 1/C_R^2}$$

where $f$ = fractional flow
$F$ = maximum flow

Fractional flow can be related to stem position by extracting the square root of the ratio of the last two equations:

$$f = \left[\frac{(C_v/C_R)^2 + 1}{(C_v/C_R)^2 + 1/m^2}\right]^{1/2} \tag{2.15}$$

A plot of $f$ vs. $m$ for various ratios of $C_v/C_R$ appears in Fig. 2.9.

The effect of an equal-percentage characteristic upon the nonlinearity of line drop can be seen by combining these curves with the curves in

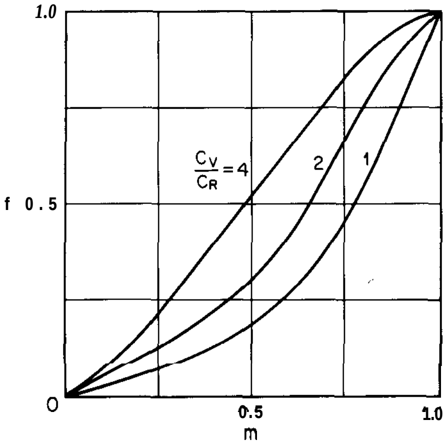



*FIG 2.10. An equal-percentage valve is able to remove most of the effect of line drop.*

*FIG 2.11. Discharge pressure from a centrifugal pump varies with the flow being delivered.*

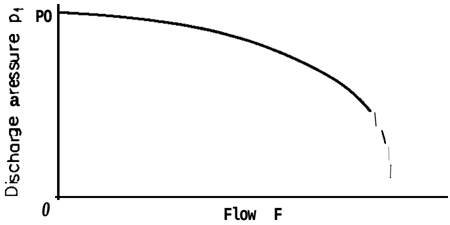



Fig. 2.6. In Fig. 2.6, fractional flow is plotted vs. stem position for conditions of constant $\Delta p$, which is the same as fractional valve opening vs. stem position. In a linear valve, fractional valve opening is identical to fractional stem position. Therefore, values of $f$ from Fig. 2.6 entered as values of $m$ in Fig. 2.9 will result in the plot of flow vs. stem position for an equal-perc'entagc valve with line drop, which is shown in Fig. 2.10.

The centrifugal pump is the most common motive force for transportation of liquids. But this type of pump is not a source of constant pressure; instead, pressure varies with flow in the manner described by the curve of Fig. 2.11. Frictional losses within the pump cause this variation, much as internal resistance in a battery makes terminal voltage fall as current drain is increased. The equation of' the curve is readily derived, with $C_R$ representing the flow coefficient of the internal resistance. If $p_0$ is the "no flow" pressure, the drop within the pump is

$$p_0 - p_1 = \frac{F^2}{C_R{}^2}$$

and

$$p_1 = p_0 - \frac{F^2}{C_R{}^2} \tag{2.16}$$

Equation (2.16) will be found to match most pump curves quite well. But the outcome of this relationship is that a centrifugal pump looks just like a constant pressure source with line drop and may be treated in the same manner.

### Process Gain

The output of a valve is flow; the process accepts this flow and converts it into the controlled variable. If the controlled variable is also *flow*, as in a flow-control loop, the process gain is unity. But if the controlled variable has any other dimension—pressure, temperature, composition, etc.-the process has a dimensional gain.

If the controlled variable is an *integral* of *flow*, such as pressure or level, dimensional gain is included in the integrating time constant V/F. The self-regulating liquid-level process of Chap. 1 was found to have a steady-state gain of $1/k$. But it was pointed out that this steady-state gain had no influence whatsoever on the dynamic gain of the process. The dynamic gain of the process can be considered identical to that of the integrating process. Conversion of units of flow into units of volume takes place in the integration. The horizontal sectional area **A** of the vessel then converts volume into level:

$$h = \frac{V}{A} \tag{2.17}$$

If the horizontal area is not uniform, the relationship between level and volume becomes nonlinear. This would be the case for a sphere or a horizontal cylindrical tank.

In processes where the *transfer* of *mass* or *energy* takes place, gain is a function of many fartors, making generalization impossible. These processes are not only difficult, to control, because of their dynamic behavior, but they are also difficult to understand. Lack of understanding looms as the greatest single factor contributing to the failure of control systems applied to these processes. They are usually nonlinear in more than one respect, and compensation improperly applied can aggravate the situation. The steady-state relationships that prevail among manipulated, load, and controlled variables take on paramount importance. To give them the consideration they deserve, four entire chapters are devoted to processes involving energy and mass transfer.

An introductory example of what may be encountered is the neutralization process, where pH of the product is to be controlled. A typical neutralization curve is presented in Fig. 2.12. The effluent pH is plotted against the ratio of acid to influent flow, where acid flow is manipulated. The principal factor in a pH loop is the shape of this curve. Its slope is the process gain, in that it converts changes in acid flow to changes in pH. But the slope varies markedly with pH. The curve is exponential in nature, changing in slope as much as 1,000 : 1 between the extremes. The set point is usually somewhere in the steepest region of thd curve. Achievement, of damping requires a very wide proportional band. A

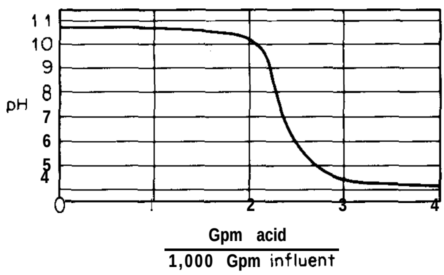



*FIG 2.22. Changing the ratio of acid to influent adjusts the pH of the effluent.*

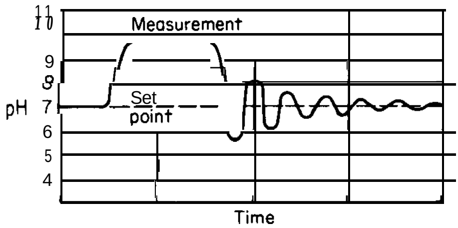



*FIG 2.13. An uncompensated pH loop tends to exhibit a distorted oscillation.*

load change of any magnitude can then drive the pH far enough up the curve to reduce the loop gain to the point where recovery is extremely slow. Compensation for such a severe nonlinearity is essential if a satisfactory degree of performance is to be obtained. The nonlinearity is so severe that even poor compensation is noticeably effective. Reduction in loop-gain variation from 100: 1 to 5 : 1, for example, is bound to improve the situation.

The shape of the neutralization curve is a function of the reacting substances. If the ingredients of the influent are subject to change, the slope of the curve at a given pH may also change. This amounts to a second gain variation superimposed on the first. This is not a predictable characteristic.

The response of a typical uncompensated pH loop to a load change is pictured in Fig. 2.13. Oscillations are flattened by the change in gain. Recovery is considerably retarded by the reduced loop gain away from the vicinity of the set point. More space is devoted to pH control in Chap. 10.

### Variable Dynamic Gain

Figure 2.14 is a record of temperature and flow of a product stream leaving a heat exchanger. Temperature was being controlled by manipulating the flow of steam to the exchanger. Notice that the temperature record at 80 percent flow is overdamped, whereas at 40 percent flow,

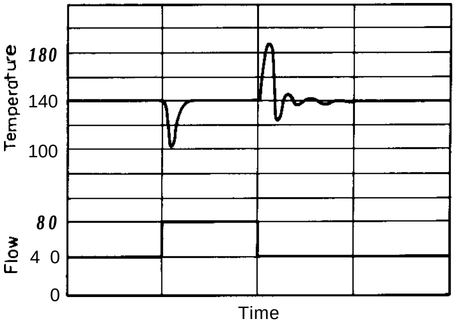



*FIG 2.14. Both magnitude and damping of the transient are functions of flow.*

damping is slightly heavier than ¼-amplitude. (It appears that the loop will be only marginally stable at 20 percent flow.) It is as if the proportional band had been changed. But the more lightly damped curve departs farther from the set point-contrary to the effects of changing proportional band as in Fig. 1.7. Therefore it is the process gain which has changed: the process is more lightly damped and more sensitive to disturbance at low rates of flow.

The problem has been identified as variable dynamic gain. It is a common problem, not often recognized, still less often anticipated. It occurs in processes where the values of the secondary dynamic elements, principally dead time, vary with flow. These variations cause proportionate changes in the period of the loop, which affects the dynamic gain of the principal capacity.

Consider the heat exchanger as a single-capacity plus dead-time process where the dynamic gain of the capacity is expressed as

$$G_1 = \frac{\tau_o}{2\pi V/F}$$

Let dead time vary with flow through tubing of volume v:

$$\tau_d = \frac{v}{fF}$$

where $f$ = fractional flow
$F$ = maximum flow of product

The period of oscillation varies with $\tau_d$:

$$\tau_o = 4\tau_d = \frac{4v}{fF}$$

Dynamic gain is then

$$G_1 = \frac{4v/fF}{2\pi V/F} = \frac{2v}{\pi Vf} \tag{2.18}$$

Dynamic gain is inversely proportional to product flow. As flow approaches zero, gain approaches infinity. If uncompensated, this gain variation will cause serious problems, particularly during startup, when flow is low. If adjusted for low flow, the controller will perform poorly at higher rates, as Fig. 2.14 substantiates. Notice that the response is in no way similar to that of Fig. 2.13, the nonlinear process. To distinguish between the two, this characteristic will be referred to as variable gain. It is more like the response that would be encountered with a nonlinear valve, or valve-plus-piping characteristic-a function of flow, not of the measurement.

Fortunately, the correction is so readily applied that in most cases the correction is inadvertent and there is little evidence of the problem's ever having existed. The gain of the process varies inversely with product flow. The gain of an equal-percentage valve varies directly with the manipulated flow. As long as the ratio of the two is constant, the gain product will be constant. Let $f_s$ be the fraction of maximum steam flow $F_s$. Then $f_sF_s$ is proportional to product flow by the temperature rise AT of the product as it passes through the exchanger:

$$f_sF_s = fFC \text{ AT}$$

where C is a constant. The gain of the equal-percentage steam valve is

$$G_v = 4f_sF_s = 4fFC \text{ AT}$$

Gain product, is then

$$G_vG_1 = \frac{8fFC\,\Delta T\,v}{\pi Vf} = \frac{8C\,\Delta T\,v}{\pi V/F} \tag{2.19}$$

The loop gain is no longer a function of product flow, because of the compensating nonlinearity of the equal-percentage valve. But a trade has been made: loop gain is now a function of AT, where it was not before. In most heat exchangers, however, AT never varies as much as 2: 1, whereas flow commonly does. So the trade is distinctly in the best interests of the loop. Equal-percentage valves are so widely used to combat line drop, that in many cases they are also compensating for variable dynamic gain without the user being aware of it.

When the period of oscillation varies, derivative time ought to be changed accordingly. But stable control can be achieved with an incorrect value of derivative if the gain is appropriately adjusted. Consequently gain compensation for the variable dynamic element is mandatory, whereas derivative compensation can only be classified as desirable.

## TESTING THE PLANT

In the late 1950s there was much talk of extensive tests on processes using frequency-response analysis. In fact some tests were conducted on reactors, heat exchangers, and distillation columns. Although a certain amount of information was obtained using this method, two major objections stand out:

1. The tests are unbelievably time-consuming.
2. They assume that the process is linear and invariant.

The first objection rules out testing in most plants because of the unwillingness of operating personnel to tolerate upsets for long intervals and because of the expense of manpower and equipment. The second objection indicates that the results of tests on a process with nonlinear elements may be not only invalid, but also misleading. Frequency response is only suitable for fast, linear devices like instruments, controllers, amplifiers, etc.

The author has been called upon many times to investigate a process which was in trouble. In these instances it was impossible to bring an extensive array of test, equipment' or to spend days gathering information. In most cases the process was nonlinear in some respect and not well understood by the operating people—otherwise it would not have been in trouble. A simple test procedure was decided on, independent of linearity, from which the dominant properties of the system could be determined. A properly conducted test should pinpoint problem areas with a minimal upset to the process.

To keep testing to a minimum, all available knowledge of the process must be employed. The volume of vessels and flow rates are always available, from which time constants may be calculated. The length and diameter of piping runs can serve to locate dead-time elements. By identifying all the known or knowable elements in this way, any tests will be of more value in defining the unknown elements which make up the balance of the loop.

The author has always reacted strongly to any test procedure that is based upon knowing nothing about a process. Many things about an unfamiliar process can be learned by observing the vessels and piping, examining the chemistry and physics involved, and talking to the operators. Preliminary information like this is of inestimable help in indicating what to look for and where. It is surprising how much can often be learned about a particular process without even making a test. Occasionally the tests will not substantiate the expectations, which provides a challenging opportunity to learn.

## A Simplified Test Procedure

Before describing how to conduct a test, it is important to point out how not to conduct a test, in order to avoid some serious pitfalls.

1. Do not test for steady-state gain. In Chap. 1 it was pointed out that the steady-state gain of a single-capacity liquid-level process is not constant. It varies with both flow and level. Yet the dynamic gain is constant. Because the process is in a control loop, only the dynamic gain-the loop gain at $\tau_o$—is of real consequence.

2. Do not test for time constants. There arc several methods available for finding the time constants in a linear system. But, as in the single-capacity level process, the time constant may vary with flow without affecting dynamic gain. The likelihood of a nonlinear element in a

troublesome process is extremely high, rendering these tests meaningless. The tests also require the process to come to rest after a disturbance. A non-self-regulating process will not come to rest and therefore cannot be treated in this way. Furthermore these tests require the control loop to be open until a new steady state has been reached, which could be a long time.

Fortunately there is a quick and easy method for obtaining enough information to suggest corrective measures in most instances. The method consists of one open-loop and one closed-loop test. In the latter case, the proportional mode of a controller serves as the test instrument. The procedure is as follows:

1. With the controller in manual, step or pulse the control valve sufficiently to produce an observable effect'. Measure the time elapsed between the disturbance and the first indication of a response. This is the dead time $\tau_d$.

2. Transfer control to automatic, with minimum derivative and maximum reset time. Adjust the proportional band to develop nearly undamped oscillation. Note the period of oscillation $\tau_o$ and the proportional band setting.

In this test, it was only necessary to leave the loop open (manual control) long enough to measure the dead time. Any other type of open-loop test would consume more time. The closed-loop test describes the process under those conditions that are of greatest significance, that is, at the natural period. Two complete cycles are enough to measure $\tau_o$. If it is not practical to induce uniform oscillations, damped oscillations will suffice, although the proportional band reading should be corrected for the damping.

From the data obtained, a representation of the dynamic elements in the process may be constructed:

If $\tau_o/\tau_d = 2$, the process is pure dead time.
If $2 < \tau_o/\tau_d < 4$, dead time is dominant.
If $\tau_o/\tau_d = 4$, there is a single dominant capacity.
If $\tau_o/\tau_d > 4$, more than one capacity is present.

Furthermore, the setting of proportional band responsible for uniform oscillation equals the gain product of the other elements in the loop at $\tau_o$.

When these bits of information are combined with the characteristics of the known elements, a remarkably accurate picture of the process can he asscmblcd. For example, if the process is known to contain one principal capacity, and $\tau_o/\tau_d = 4$, no other time constants need be sought'. If the time constant of this capacity is known, its dynamic gain $G_1$ at $\tau_o$ can be calculated. Combining this with known values of transmitter and valve gain, together with the controller proportional band, yields the process gain:

$$G_p = \frac{P}{100G_1G_TG_v} \tag{2.20}$$

Since these tests are made only at one operating point, they will not disclose any nonlinear properties. Closed-loop response should be observed at other flow conditions to detect any change in damping. If the period changes with flow, a variable dynamic element is present. An extremely nonlinear measurement, such as pH, is identified by the distorted waveform it produces, as in Fig. 2.13. A less severe nonlinear measurement may not be detected without changing the set point. In short, if a thorough analysis is to be made, the closed-loop test should be repeated at other values of flow and set point.

#### Testing a Neutralization Process

This is an actual case history of the process upon which this test procedure was first tried. It was a neutralization process in which a reagent was being added to bring the effluent leaving a reactor to pH 7. The pH controller was in manual, simply because automatic control was unsatisfactory.

The open-loop test gave a dead time of 40 sec. The volume of the sample piping divided by the sample flow was 15 sec. The remainder was probably distributed through the reactor and associated piping.

With a proportional band of 150 percent, the loop sustained uniform oscillation of 2.8-min period. The ratio $\tau_o/\tau_d = 2.8/0.67$, or 4.2, indicated essentially a single capacity along with the measured dead time.

The reaction vessel contained 200 gal of material, flowing at 2.5 gpm. Therefore $V/F = 200/2.5$, or 80 min. The dynamic gain of an SO-min capacity at a 2.8-min period is

$$G_1 = \frac{2.8}{2\pi 80} = 0.004$$

Yet the proportional band for zero damping was 150 percent. This can mean only one thing-extremely high process gain. Dividing $G_1$ into $P/100$ yields the gain product of valve, process, and transmitter:

$$G_pG_vG_T = \frac{150}{100(0.004)} \simeq 375$$

Again the familiar problem of the pH curve appears:: high gain near the control point, low gain elsewhere. But, the situation could be helped. Repiping the sample line reduced its dead time to 5 sec, bringing the total dead time to 30 sec. This reduced the period to 2 1 min and the proportional band by the same factor of ¾. So the controller was adjusted for damping at the new conditions. (A procedure for adjusting three-mode controllers is described in Chap. 4.)

A later observation revealed that the loop had become more heavily damped. The only noticeable change since the controller was adjusted was a lower value of output. The loop gain apparently had decreased with load. An inquiry about the valve characteristic produced the answer: reagent was being delivered through an equal-percentage valve under constant pressure drop. The loop gain therefore varied directly with flow, as did the valve gain.

Although a pH process is nonlinear, its characteristic curve cannot be corrected with an equal-percentage valve, because the valve acts on the output of the controller, not on the input.[4] The valve characteristic, in fact, made matters worse. Not only did the loop gain become variable, but it was higher than it would have been with an equivalent linear valve. The gain of an equal-percentage valve is four times the fractional flow; fractional flow in excess of 0.25 will cause the gain to exceed unity. If the normal flow is 50 percent of the valve's capacity, the equal-percentage characteristic will contribute twice the gain of a linear valve. This necessitates a proportional band twice as wide.

The time required to test this process at one operating point was only a few minutes. Yet together with known facts about the plant, and one subsequent observation, the process was thoroughly defined and two recommendations made to improve control. Any other test procedure would have taken longer and might not have achieved comparable results.

## REFERENCES


1. Caldwell, W. I., G. A. Coon, and L. M. Zoss: "Frequency Response for Process Control," pp. 76–80, McGraw-Hill Book Company, New York, 1959.
2. Ziegler, J. G., and N. B. Nichols: Optimum Settings for Automatic Controllers, *Trans. ASME*, December, 1941.
3. Bradner, M.: Pneumatic Transmission Lag, ISA Paper No. 48-4-2.
4. Shinskey, F. G.: Controls for Nonlinear Processes, *Chem. Eng.*, March 19, 1962.


## PROBLEMS

**2.1** Given two processes consisting of 10 identical time constants, in one of which they interact, in the other they do not: Estimate the dynamic gain of each process at its natural period under proportional control; compare the natural periods of each.

**2.2** The composition of a product leaving a 50-tray distillation column exhibits a dead time of 10 min following a change in reflux flow. Under proportional-plus-reset control, estimate (*a*) the period of oscillation, (*b*) the reset time, (*c*) the dynamic gain of the process.

**2.3** The process in Prob. 2.2 has a linear valve whose capacity is 50 gpm. A change in flow of 1 gpm causes 0.5 percent change in product quality, analyzed over a span of 10 percent. Estimate the proportional band required for ¼-amplitude damping.

**2.4** If a differential Aow controller is adjusted for ¼-amplitude damping at 30 percent flow, at what flow is it likely to be undamped?

**2.5** What value of $C_R$ in relation to the $C_v$ of a linear valve will provide reasonable compensation for the nonlinear characteristic of a differential flowmeter? If the value of $C_R$ existing in a pipeline is too high, how can it be adjusted? What is the effect of $C_R$ on valve size requirements.?

2.6 A process exhibits a dead time of 23 sec. At 50 percent flow, proportional control with a 20 percent band causes undamped oscillations of 1.5-min period. At 25 percent flow, however, its natural period increases to 2.9 min. Draw some conclusions about the process and make suitable recommendations.

# Analysis of Some Common Loops

CHAPTER 3

Classification of processes into broad areas with certain common characteristics is both desirable and informative. We know, for example, that a temperature-control loop behaves very differently from a level-control loop. Why it does so is the essence of the classification.

The first control loop to'be considered is flow. It has the distinction that the manipulated variable and the controlled variable are the same. They may not have the same range or the same linearity, nevertheless they are the same variable. For this reason the flow loop is the easiest to understand, as far as steady-state characteristics are concerned.

We will now analyze the control of variables that are the integral of flow. Liquid level is the integral of liquid flow, whereas the integral of gas flow in a constant-volume system is pressure. These loops have certain features not common to other classifications. For example they can be non-self-regulating. This is never true of flow and rarely true of other variables. Second, the rate of change of measurement is a function of the *difference* between inflow and outflow; either inflow or outflow will be load-dependent, while the other is manipulated. Furthermore these processes are dominated by capacity; dead time will rarely be found, because pressure waves travel through the process at the velocity of sound.

The third group includes energy and mass transfer processes, where control is exercised primarily over temperature and composition. The controlled variable here is always a property of the flowing stream, as opposed to being the flowing stream or its integral. These processes ordinarily have a steady state in which the controlled variable is a function of the *ratio* of the manipulated flow to the load. (Note the abscissa of Fig. 2.12, expressed in terms of this ratio.) Because the controlled property travels with the fluid, it must be transported to the measuring element. Transportation involves dead time. Hence loops in this category are usually dominated by dead time, which makes control difficult and response slow.

In this chapter, five typical control loops will be analyzed: flow, level, pressure, temperature, and composition. The principal dynamic elements of each process will be derived and will be related to the closed-loop response, Constraints and nonlinearities will be included, as well as means for coping with them. A few additional comments will serve to distinguish those control problems which are not typical or which appear to cross into other areas.

## FLOW CONTROL

Flow is the manipulated variable as well as the controlled variable, so it seems as though the process is unity. But this is not the case. Opening a valve does admit flow, but the response is not quite instantaneous. If the fluid is gaseous, it is subject to expansion upon a change in pressure; therefore the contents of a pipe vary somewhat with pressure drop, hence with flow. In a liquid stream, inertia is significant-flow cannot be started or stopped without accelerating or decelerating. To demonstrate the dynamic character of inertia, the time constant of a column of liquid in a pipe will be derived.

### Inertial Lag of a Flowing Liquid

In the steady state, the velocity of flow in a pipe varies with pressure drop :

$$u^2 = C^2 \, 2g \, \frac{\Delta p}{\rho}$$

where $u$ = velocity, ft/sec
C = flow coefficient
g = gravity, ft/sec²
$\Delta p$ = pressure drop, lb/ft²
$\rho$ = density, lb/ft³

But velocity is proportional to flow:

$$u = \frac{F}{A}$$

where $F$ = *flow*, ft³/sec
$A$ = inside area, ft²

Therefore the pressure drop due to flow in the steady state is

$$\Delta p = \frac{u^2\rho}{2gC^2} = \frac{F^2\rho}{2gA^2C^2}$$

If the applied force $A$ Ap exceeds resistance to flow, acceleration takes place. An equation can he written for the unsteady state: net force equals mass times acceleration.

$$A\,\Delta p - \frac{AF^2\rho}{2gA^2C^2} = M\frac{du}{dt} = \frac{M}{A}\frac{dF}{dt}$$

where $M$ = mass, slugs
$t$ = time, sec

The mass of fluid in the pipe is

$$M = \frac{LA\rho}{g}$$

where $L$ = length in feet.
Rearranging,

$$\frac{F^2\rho}{2gAC^2} + \frac{LA\rho}{g}\frac{dF}{dt} = A\,\Delta p$$

To find the time constant, the differential equation must be reduced to its standard form:

$$F + \frac{2LAC^2}{F}\left(\frac{dF}{dt}\right) = \frac{2gC^2A^2\,\Delta p}{\rho F}$$

The time constant is then the coefficient of $dF/dt$:

$$\tau = \frac{2LAC^2}{F} \tag{3.1}$$

Flow coefficient $C^2$ can be replaced by its steady-state equivalent:

$$C^2 = \frac{F^2\rho}{2gA^2\ \mathrm{Ap}}$$

leaving

$$\tau = \frac{LF\rho}{gA\ \Delta p} \tag{3.2}$$

**example 3.1**

To test the significance of the last expression, a numerical example is presented. Consider a 200-ft length of 1-in. Schedule 40 pipe, containing mater flowing at 10 gpm with a 20-psi drop.

$L = 200$ ft
$F = 10$ gpm $= 0.0223$ ft$^3$/sec
$p = 62.4$ lb/ft$^3$
$g = 32.2$ ft/sec$^2$
$A = 0.006$ ft$^2$
$\Delta p = 20$ lb/in$^2$ $= 2{,}880$ lb/ft$^2$

$$\tau = \frac{(200)(0.0223)(62.4)}{(32.2)(0.006)(2{,}880)} = 0.50 \text{ sec}$$

Notice that the time constant varies with both flow and pressure drop, because of the square relation between the two. Nevertheless, the derivation permits evaluation of the dynamic response at a nominal flow and at least a qualitative indication of the response elsewhere. As may have been anticipated, the time constant is small, but not zero, except, at zero flow.

### Dynamic Elements Elsewhere in the Loop

This time constant is fundamentally the only dynamic element in the process. But its response is of the same order of magnitude as the instruments in the control loop, and therefore the entire loop must be analyzed.

Figure 5.1 describes a pneumatic flow-control loop consisting of transmitter (2), controller (4), valve (6), and two transmission lines (3,5). The flow transmitter contains an amplifier with certain dynamic prop-

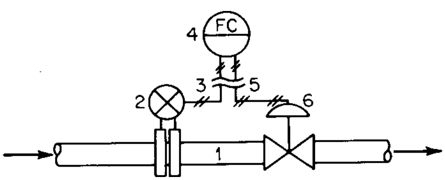



*FIG 3.1. At least six elements contribute to the dynamic response of the flow-control loop.*

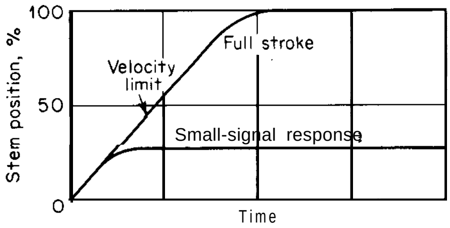



***FIG 3.2. Because of its velocity limit, a control valve appears to follow small signals faster than large signals.***

erties. Because of the amplifier, the lag of the transmission line is isolated from that of the flowing fluid. The transmission line is terminated by a controller, isolating it from the second transmission line. The figure shows no isolating amplifier between the output line and the valve, however, allowing interaction there. Transmission lines can be conveniently represented by dead time plus a first-order lag. The value of each is naturally a function of length and diameter.[1]

The control valve cannot be so easily represented, however. If a valve motor were a constant-volume device, it would behave like a first-order lag. But every change in pressure is accompanied by a change in volume of the motor. This property causes the motor to operate at a limited velocity, based on the maximum rate of flow of air that can be delivered into the expanding volume. In effect, a valve seems to exhibit a smaller time constant for small changes than for large changes, because the velocity of stroke is not a function of the magnitude of the change. Consequently, a valve cannot be adequately represented by a single time constant.

**example 3.2**

Let us analyze the response of the flow loop whose inertial lag was estimated in Example 3.1. The flowmeter has a range of 0 to 15 gpm, both transmission lines are 100 ft of ¼-in. tubing (0.188 in. inside diameter), the valve is a 1-in. size, with a linear characteristic. The combination of transmission line and valve motor will be assumed to be about 3 sec, which limited tests[2] indicate to be reasonable. The dead time associated with the tubing ought not to be neglected, however.

A closed loop of this description will be found to oscillate at a period of about 6.5 sec, under proportional-plus-reset control. The period is determined by the phase contribution of all the elements. Table 3.1 lists the phase and gain contribution of each at the 6.5-sec period.

The loop phase was found by first selecting 6.5 sec as the natural period; reset action must then contribute 28° to bring the total to 180°. (A different value of reset time would change $\tau_o$.) The gain contribution of reset at that phase angle is 1.11 . Notice that all elements contribute some phase lag, but only those whose phase lag approaches 45° affect the loop w i n noticeably.

TABLE 3.1 Dynamic Elements in the Flow Loop

| | $\tau$, sec | $-\phi$, deg | $G$ |
|---|---|---|---|
| 1. Process. | 0.5 | 26 | 0.93 |
| 2. Transmitter'. | 0.16 | 9 | 0.98 |
| 3. Transmission line: | | | |
| $\tau$ | 0.5 | 26 | 0.92 |
| $\tau_d$ | 0.18 | 10 | 1.0 |
| 5. Transmission line $\tau_d$ + | 0.18 | 10 | 1.0 |
| 6. Valve $\tau$ | 3.0 | 71 | 0.35 |
| Loop minus controller.. | | Sum 152 | Product 0.29 |
| 4. Controller reset, | 2.2 | 28 | 1.11 |
| Loop minus proportional. | | Sum 180 | Product 0.33 |

To determine what the proportional band will be, valve and transmitter gain must be combined with the dynamic loop gain of 0.33, calculated above. In the given example, 200 ft of pipe and fittings produce about an 8-psi drop at 10 gpm, leaving 12 psi across the valve. A 1.O-in. valve has a $C_v$ rating of 10. At 12-psi drop, the gain of the linear valve nould be

$$G_v \frac{C_v \sqrt{\Delta p}}{100\%} = \frac{10\sqrt{12}}{100\%} = 0.35 \text{ mm/\%}$$

At 10 gpm, the gain of the 15-gpm differential meter is

$$G_T = 2\left(\frac{10}{15}\right)\left(\frac{100\%}{15 \text{ gpm}}\right) = 8.9\%/\text{gpm}$$

Since the flow process itself has no dimensional gain, $G_v$ may be multiplied directly by $G_T$ to remove dimensions:

$$G_v G_T = (0.35 \text{ gpm}/\%)(8.9\%/\text{gpm}) = 3.1$$

The proportional band required for ¼-amplitude damping is then 200 times the gain product of the dynamic and steady-state components of the loop :

$$P = 200(0.33)(3.1) = 205\%$$

This is quite typical for a flow controller.

Notice by comparing Figs. 2.5 and 2.9 that the resistance of the piping is helpful in that it shapes the gain of the valve in a direction complementary to the flowmeter. Valve gain is higher at low flow, where the transmitter gain is lower, yielding a gain product that tends more toward uniformity than either of the multiplicands. Note also that an equal-percentage valve characteristic (Fig. 2.6) is of the opposite form, tending

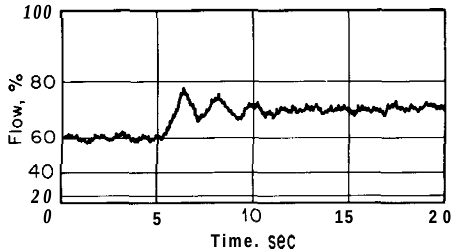



*FIG 3.3. Flow noise precludes the use of derivative.*

to augment the nonlinearity of the flowmeter except in cases of unusually severe line drop (Fig. 2.10).

Increasing valve speed by means of a booster is helpful in reducing $\tau_o$, although the proportional band may not be affected. Mounting the controller at the valve helps even more by eliminating both transmission lines.

### Flow Noise

In an equivalent electronic flow loop, absence of the transmission lines reduces the natural period to the vicinity of 2 sec. Noise, however, becomes more prominent. "Noise" means disturbances, either periodic or random, occurring at frequencies too high for control action. Figure 3.3 is a record of noise in an electronic flow loop. Turbulence in the stream and vibration from pumps are the chief sources of this noise. Even in pneumatic systems, flow noise is invariably present in sufficient magnitude to prevent the use of derivative. Phase lead is useful, but unfortunately the increase in high-frequency gain which accompanies it actually explodes the loop into instability.

### Summary

The purpose of the analysis is not to show how an analysis should be made, but rather to explain why a flow loop behaves the way it does. Because many dynamic elements are present, all of the same order of magnitude, dynamic gain is high. The proportional band of a flow controller is rarely less than 100 percent, making reset mandatory. Where the valve and transmitter are in the same line, the period of oscillation will invariably fall within 1 to 10 sec. The presence of noise precludes the use of derivative. As long as these factors are appreciated, there is little reason to spend time analyzing flow loops.

## PRESSURE REGULATION

The thermodynamic state of a system can be defined from its pressure, enthalpy, and volume. If a gas phase alone is present, pressure and

**TABLE 3.2 The Significance of Pressure as a Measurement of Specific Volume and Enthalpy of Steam and Water at 100 psia**

| System | Spec. vol. change, % / Pressure change, % | Enthalpy change, % / Pressure change, % |
|---|---|---|
| Superheated vapor at 1000°F. | -1.006 | -0.00163 |
| Saturated vapor............... | -0.945 | +0.0158 |
| Compressed liquid at 100°F. . | -0.0003 | +0.00395 |

volume are inversely proportional, with enthalpy playing a relatively minor role. When a vapor is in equilibrium with its liquid, however, a change in enthalpy of the system will produce a pronounced pressure change, while volume variations will have less effect. Liquids, moreover, are virtually incompressible, with the result that neither pressure nor enthalpy have much influence over system volume.

The thermodynamic properties of gas, vapor, and liquid systems have been brought out expressly to establish that the properties of system pressure are decidedly a function of state. It is extremely important to attach the correct significance to the pressure measurement, if acceptable performance of a control loop is to be gained. Table 3.2 gives an example of each of the three states listed above, where water is the substance under pressure. It indicates the conditions under which pressure is a suitable measurement of the material content (specific volume) and energy content (enthalpy) of the system.

The table points out that pressure is an adequate measurement of the material content of a system which contains only gas. Enthalpy of a gas, on the other hand, is more a function of temperature than of pressure. Consequently gas pressure should be controlled by manipulating the material content of the system, i.e., inflow or outflow. But in a system where vapor and liquid are in equilibrium, pressure could be controlled by adjusting the flow of either material or heat. Finally, pressure is a poor measure of either heat or mass content of a liquid, so another approach must be taken in stipulating its control.

### Gas Pressure

The perfect gas law states that

$$pV = MRT$$

where $p$ = system pressure
$V$ = volume
$M$ = mole content
$R$ = gas constant
$T$ = absolute temperature

The rate of change of pressure in a constant-volume system is related to the change in material content of the system:

$$\frac{dp}{dt} = \frac{dM}{dt}\frac{RT}{V}$$

If $R$ and $T$ are both constant, the rate of change of mass content of the system is the difference between mass inflow and outflow:

$$\frac{dp}{dt} = \frac{F}{V}(f_i - f_o)$$

where $F$ = nominal mass flow
$f_i$ = fractional inflow
$f_o$ = fractional outflow

Integration of the last equation places pressure in terms of flow:

$$p = \frac{1}{V/F}\int (f_i - f_o)\,dt \qquad (3.3)$$

For dimensional conformity, $p$ would be in units of atmospheres, $V$ in cubic feet, and $F$ in standard cfm, that is, cfm at 1.0 atm. Thus the time constant $V/F$ is expressed in minutes.

Just as level control was used to close a liquid material balance around a tank, pressure control is used to close a gas material balance. The gas-pressure process is ordinarily self-regulating, except at zero flow, because pressure always influences inflow and outflow. The process is fundamentally single-capacity, although the pressure transmitter and valve can add very small secondary lags. If there is no transmitter, as with a self-contained regulator, one secondary lag is eliminated.

Pressure of a gas is easy to control, even when the volume of the system is small, e.g., only piping. In fact, the narrow-band proportional action of self-contained regulators is sufficient for most applications. They are, for the most part, as sensitive as their simple construction will allow, indicating that loop gain is not a problem. Pressure acting on the diaphragm compresses the spring, moving the plug within the valve. Each position of the seat corresponds to a given pressure on the diaphragm. Initial compression of the spring sets the pressure at which the valve begins to open.

Because pressure will vary with flow, as in Fig. 3.4, a regulator is said to exhibit "droop." Regulators differ, but a typical proportional band would be 5 percent. Near zero flow, extra pressure is needed for shutoff; at the other extreme, the valve is wide open and acts as a fixed resistance.

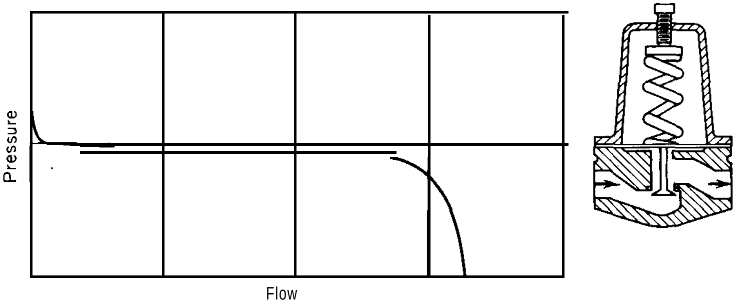



***FIG 3.4. The characteristic curve for a pressure regulator indicates proportional action.***

**Vapor Pressure**

In a system containing liquid and vapor in equilibrium, the difference between inflow and outflow of vapor would change the pressure, from a material-balance standpoint:

$$F_i - F_o = V\frac{dp}{dt}$$

But if the enthalpy of inflow and outflow differ, flow of material between the vapor and liquid phases will also affect system pressure. An energy balance shows the relationship:

$$F_iH_i - F_oH_o + Q_i - Q_o = VH_v\frac{dp}{dt} \tag{3.4}$$

The terms $H_i$ and $H_o$ represent enthalpy of inflow and outflow respectively, $Q_i$ and $Q_o$ represent transfer of heat in and out, and $H_v$ is the heat of vaporization.

Both mass flow and heat flow affect pressure. But where the net change of enthalpy across a process is zero, mass flow alone is sufficient for control. An example of this situation is pressure reduction of saturated or wet steam—there is no change in enthalpy across the reducing valve.

In a boiler, or distillation column, or evaporator, transfer of heat is an integral part of the operation, and system pressure can be used to close the heat balance. In this role, the pressure controller has much the same type of dynamic and steady-state relationships as a temperature controller normally does. Therefore the propertics of this sort of pressure-control loop will be covered for the most part under considerations of temperature control.

### Liquid Pressure

Pressure control of a liquid stream is exactly like flow control. The pressure at the origin of a pipeline, for example, is directly related to flow in the line. The process's only dynamic contribution is that of inertia of 'the flowing fluid.

The process gain $G_p$ in a flow loop is, by definition, 1.0. But in a pressure loop there must be a conversion from flow into units of pressure. Liquid pressure upstream of a resistance $C_R$, like differential pressure, varies with flow squared:

$$p = p_0 + \frac{F^2}{C_R{}^2} \tag{3.5}$$

The intercept $p_0$ is the static pressure at no flow. Differentiating, we obtain the process gain:

$$\frac{dp}{dF} = \frac{2F}{C_R{}^2} \tag{3.6}$$

Ordinarily pressure moves less than full scale for full-scale change in valve position, resulting in a lower proportional band than for a flow loop. Other characteristics, including noise, are similar.

Self-contained regulators are sometimes used for liquid pressure and perform moderately well on quiet streams. Recalling that the dynamic elements which caused most of the problems in the flow loop were instruments and transmission lines, the application makes good sense. But where accurate regulation and tight shutoff are important, these simple devices are insufficient.

## LIQUID LEVEL AND HYDRAULIC RESONANCE

Control of liquid level is not as easy as the examples given in Chap. 1 indicate. The descriptions of Figs. 1.14 and 1.20 were intentionally oversimplified to aid understanding of single- and two-capacity processes. But the existence of waves in any body of water as large as a bay or as small as a cup, gives rise to the speculation that any liquid with an open surface is capable of sustaining oscillation. While average level responds to flow as an integrator, level responds to level in a resonant manner. Consequently the liquid-level process is not single-capacity, even with a directly connected measuring element.

### The Period of Hydraulic Resonance

To analyze this resonance, let us take the case of the vessel with a measuring chamber shown in Fig. 3.5, neglecting resistance to flow. If

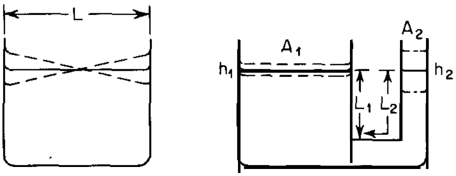



*FIG 3.5. The period of hydraulic resonance varies with the distance between the bounded surfaces.*

the level in the measuring chamber momentarily exceeds that in the tank, the differential force developed causes a downward acceleration in that leg:

$$\rho h_2 A_2 - \rho h_1 A_1 \approx -M_2 \frac{du_2}{dt} - M_1 \frac{du_1}{dt} \tag{3.7}$$

where $h_2$, $A_2$, $M_2$, and $u_2$ are the head, area, mass, and velocity, respectively, of the fluid in the measuring chamber. As before, $\rho$ is the fluid density. Furthermore

$$u_1 = \frac{A_2}{A_1} u_2$$

and

$$M_1 = \rho \frac{L_1 A_1}{g} \qquad M_2 = \rho \frac{L_2 A_2}{g}$$

Substituting for $u_1$, $M_1$, and $M_2$ in Eq. (3.7) yields

$$h_2 A_2 - h_1 A_1 = -\frac{L_2 A_2}{g} \frac{du_2}{dt} - \frac{L_1 A_2}{g} \frac{du_2}{dt} \tag{3.8}$$

Level in the measuring chamber, $h_2$, is related to average level $h$ by

$$h - h_1 = (h_2 - h) \frac{A_2}{A_1}$$

Including this in Eq. (3.8) yields the response of measured level $h_2$ to average level $h$:

$$2h_2 = h\left(1 + \frac{A_1}{A_2}\right) - \frac{L_1 + L_2}{g} \frac{du_2}{dt}$$

But velocity $u_2$ is the rate of change of level, $dh_2/dt$. Therefore a differential equation can be written eliminating $u_2$:

$$h_2 + \frac{L_1 + L_2}{2g} \frac{d^2h}{dt^2} = \frac{h}{2}\left(1 + \frac{A_1}{A_2}\right) \tag{3.9}$$

This differential equation is descriptive of a second-order undamped system. The U tube resonates at a natural period established by the square root of the coefficient of the differential:

$$\tau_o = 2\pi \left(\frac{L_1 + L_2}{2g}\right)^{1/2} \tag{3.10}$$

Notice that the period is unaffected by density, area, or any property other than the distance $L_1 + L_2$ between the bounded surfaces. Compare it to that of a pendulum, also a function of length and gravity only.

Liquid in a vessel may also oscillate without the benefit of a U tube. The period of oscillation of the surface of diameter $L$ is:

$$\tau_o = 2\pi \left(\frac{L}{2g}\right)^{1/2} \tag{3.11}$$

Rectangular vessels can oscillate at two different periods. Vessels with an attached measuring chamber can oscillate with at least two different periods.

The natural period of any control loop containing a resonant element cannot exceed that of the resonant element. The phase shift of a resonant element is exactly -90" at its natural period, no matter how heavily damped it may be. Since the integration of flow into average level represents an inherent phase shift of $-90°$, the process, from flow to measured level, will exhibit -180" at the natural period of the vessel. To damp the measuring chamber by throttling its connecting valves will not change this period, but will only reduce the amplitude of the resonance.

**example 3.3**

As an example of a liquid-level control problem, consider a vessel with a measuring chamber of the following description:

Volume $V$: 100 gal
Maximum flow $F$: *50* gpm
Diameter $L$: 2.0 ft
Normal level $L_1$: 3.6 ft
Chamber $L_2$: 4.4 ft

The liquid can oscillate on the surface and in the U tube. Rut since the largest resonant period is always the limiting one, only the period of the U tube is important

$$\tau_o = 2\pi \left[\frac{3.6 \text{ ft} + 4.4 \text{ ft}}{2(32.2 \text{ ft/sec}^2)}\right]^{1/2} = 2.2 \text{ sec}$$

The dynamic gain of the integrator is

$$G = \frac{\tau_o}{2\pi V/F} = \frac{(2.2/60)}{(6.28)(100/50)} = 0.003$$

This control problem can be accommodated with a proportional band of $200G$ = 6 percent.

Since dimensions of process vessels generally fall between 2 and 200 ft, liquid resonance lies principally in the region from 1- to 10-sec period. Hence it is only of serious consequence, from the standpoint of control-loop stability, in vessels with time constants of less than 1 min.

### Liquid-level Noise

Measurement of liquid level is usually noisy, because of splashing and turbulence of fluids entering the vessel. As we have seen, loops that resonate respond to random disturbances by oscillating at their natural period. As a result, level measurements are rarely quiet, often fluctuating 20 or 30 percent of scale. This is particularly true in vessels containing boiling liquids, where turbulence is high.

Although a narrow proportional band, like the one determined in the example, may be sufficient, for control-loop stability, random fluctuations of only a few percent will drive the control valve to its limits. This may be unobjectionable in some cases, but too severe in others. Often the liquid level in a tank is used to control flow into another part of the process. It is certain that wide fluctuations in feed rate are not tolerated in most operations. To provide steady flow in these instances, the proportional band is widened and reset is relied upon to maintain control.

In many applications, exact regulation of liquid level is not important. In fact, a surge tank does not fulfill its purpose if tight control is imposed on it. As a result, control adjustments are often relaxed, and the process is sometimes left to be operated manually, if its time constant is long enough.

In some applications, a special controller whose proportional band changes with deviation is warranted. This type of controller is devised to deliver smooth flow while level is normal, but to change flow radically in the event that high or low limits are approached. Chapter 5 discusses more details of this function.

### Boiling Liquids and Condensing Vapors

Whenever level control is to be effected on a boiling liquid or condensing vapor, properties more typical of thermal processes appear. Transfer of both heat and mass is involved, which, combined with the integration of flow into level, renders control surprisingly difficult. Level control in boilers and distillation columns is sufficiently problematic to warrant special consideration, which is given in Chaps. 8, 9, and 11.

## TEMPERATURE CONTROL

Temperature-control problems are really heat transfer problems, whether the mechanism is radiation, conduction, or convection. Al-

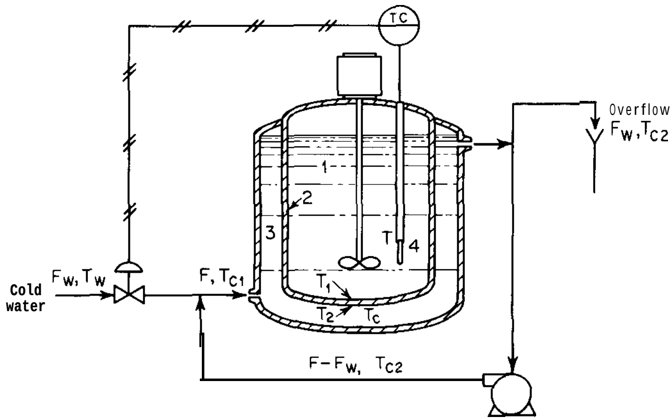



***FIG 3.6. The thermal process contains four interacting lags.***

though an entire chapter is devoted exclusively to energy control, it is important at this time to assay the general features of the temperature loop in order to establish its place in the classification that has been made.

**Example of a Constant Parameter System**

Because most heat transfer processes have variable parameters-heat transfer coefficient, dead time, etc.-which vary with flow, care has been taken to choose an example free of these complications, to better introduce the subject. The example chosen is that of a stirred tank reactor cooled by a constant flow of liquid circulating through its jacket.

The temperature controller, as shown in Fig. 3.6, adds cold water to the circulating coolant, in order to remove the heat of reaction. There are five important dynamic elements in the process:

1. Heat capacity of the contents of the reactor
2. Heat capacity of the wall
3. Heat capacity of the contents of the jacket
4. Lag in the temperature bulb
5. Dead time of circulation

Because all the heat leaving the reactor flows through the walls and into the coolant, the capacities of reactants, walls, and coolant interact. But in view of the slight heat capacity of the bulb, its time constant does not significantly interact with the others. Basically the process is four-capacity plus dead-time.

**Finding the Time Constants**

To determine the values of the time constants, an unsteady-state heat balance must be written across each heat transfer surface. The equation takes the form heat in equals heat out plus heat capacity times rate of temperature rise. Assuming a constant rate of heat evolution (the case of a variable rate will be taken up later), the heat balance at the surface of the reactor wall is

$$Q = k_1A(T - T_1) + W_1C_1\frac{dT}{dt} \tag{3.12}$$

where $Q$ = rate of heat evolution, Btu/hr
$k_1$ = heat transfer coefficient, Btu/(hr) (ft²) (°F)
$A$ = heat transfer area, ft²
$T$ = reactor temperature, °F
$T_1$ = wall temperature, °F
$W_1$ = weight of reactants, lb
$C_1$ = specific heat of reactants, Btu/(lb)(°F)

Rearranging in the standard form,

$$T + \frac{W_1C_1}{k_1A}\frac{dT}{dt} = T_1 + \frac{Q}{k_1A} \tag{3.13}$$

The thermal time constant is

$$\tau_1 = \frac{W_1C_1}{k_1A} \tag{3.14}$$

Reactor temperature responds to wall temperature with a time constant of $\tau_1$ and a steady-state gain of 1. If $k_1A$ is not directly known, $Q/(T - T_1)$ may be substituted:

$$\tau_1 = \frac{W_1C_1}{Q}(T - T_1) \tag{3.15}$$

By the same token, the temperature of the outside wall of the reactor responds to that of the inside wall with a time constant of

$$\tau_2 = \frac{W_2C_2l}{k_2A} = \frac{W_2C_2}{Q}(T_1 - T_2) \tag{3.16}$$

where $W_2$ = weight of wall, lb
$C_2$ = specific heat of wall, Btu/(lb)(°F)
$k_2$ = thermal conductivity, Btu/(hr) (ft²) (°F/in.)
$l$ = wall thickness, in.
$T_2$ = outside wall temperature

Next, outside wall temperature responds to coolant temperature with a time constant of

$$\tau_3 = \frac{W_3C_3}{k_3A} = \frac{W_3C_3}{Q}(T_2 - T_c) \tag{3.17}$$

where $W_3$ = weight of jacket contents
$C_3$ = specific heat of jacket
$k_3$ = heat transfer coefficient
$T_c$ = average coolant temperature

The lag of the temperature bulb can be calculated in the same way as the other time constants:

$$\tau_4 = \frac{W_4 C_4}{k_1 A_4} \tag{3.18}$$

where $W_4$ = weight of bulb
$C_4$ = specific heat of bulb
$A_4$ = surface area of bulb

For most types of thermal systems and heat transfer conditions, data on bulb response are already available.[3]

### Process Gain

Each of these lags forms one link in the chain from average coolant temperature to the measured reactor temperature. But since the manipulated variable is the flow of water added to the coolant stream, a suitable equation converting water flow to coolant temperature must be included. Adding a stream $F_W$ at temperature $T_W$ to the coolant recycle stream $F - F_W$ at temperature $T_{c2}$ produces a mixture $F$ at temperature $T_{c1}$, returning to the reactor. The heat balance is

$$FT_{c1} = F_W T_W + (F - F_W)T_{c2}$$

Rearranging,

$$T_{c2} - T_{c1} = (T_{c2} - T_W)\frac{F_W}{F} \tag{3.19}$$

But, related to the heat load,

$$T_{c2} - T_{c1} = \frac{Q}{FC_3}$$

Since the response of average coolant temperature $T_c$ is sought, substitution is made for $T_{c2}$:

$$T_{c2} = T_c + \frac{T_{c2} - T_{c1}}{2} = T_c + \frac{Q}{2FC_3} \tag{3.20}$$

Combining Eqs. (3.19) and (3.20),

$$\frac{Q}{FC_3} = \frac{F_W}{F}\left(T_c + \frac{Q}{2FC_3} - T_W\right)$$

Solving for $T_c$,

$$T_c = T_W + \frac{Q}{C_3}\left(\frac{1}{F_W} - \frac{1}{2F}\right) \tag{3.21}$$

Process gain is the derivative of temperature with respect to flow:

$$\frac{dT_c}{dF_W} = -\frac{Q}{C_3 F_W{}^2} \tag{3.22}$$

The adjustment of coolant temperature by water flow is demonstrably nonlinear. An equal-percentage valve should be used to deliver the water, to partially correct this situation.

**example 3.4**

If a reactor contains 40,000 lb of material of specific heat of 0.8 Btu/(lb)(°F), evolving 20000 Btu/min at 200°F aith a wall temperature of 170°F,

$$\tau_1 = \frac{(40{,}000)(0.8)(200 - 170)}{20000} = 48 \text{ min}$$

$\tau_2$ can be estimated from the weight of the reactor wall, 8,000 lb, of specific heat 0.15 and a temperature gradient of 10°F:

$$\tau_2 = \frac{(8{,}000)(0.15)(10)}{20000} = 0\ 6 \text{ min}$$

Jacket contents of 500 gal (4,160 lb) of water at an average temperature of 140°F exhibits a time constant of

$$\tau_3 = \frac{(4{,}160)(1.0)(160 - 140)}{20000} = 4.2 \text{ min}$$

A typical value for lag in a temperature well is $\tau_4$ = 0.5 min. Finally, circulation through the jacket at a rate of 250 gpm yields a dead time

$$\tau_d = \frac{500}{250} = 2 \text{ min}$$

It happens that a reactor of this description will oscillate at a period of about 35 min in a closed loop. Even if all the secondary elements consisted of pure dead time, they could not cause the period to exceed 29 min. Therefore, some secondary element remains hidden, and the only place it could hide is in the reaction mass. The assumption has been made, in calculating its time constant, that the reaction mass was perfectly mixed -that it was all at the same temperature. This, of course, is a false premise, because it is impossible to transport fluid, hence heat, from the wall of the vessel to the temperature bulb in zero time. Heat is transferred both by convection and by conduction-conduction would be the mechanism if the fluid were motionless. It has been pointed out that heat transfer by conduction is a distributed process, involving some effective dead time. So it does not seem unreasonable that a small percentage of the 48-min primary time constant is dead time due to imperfect mixing. An examination of the mechanism of mixing will be taken up under composition control.

**example 3.5**

The dynamic gain of the process is principally that of the primary time constant:

$$G_1 = \frac{\tau_o}{2\pi\tau_1} = \frac{35}{(6.28)(48)} = 0.116$$

The response of average coolant temperature to water flow is plottedin Fig. 3.7 for values of $F$ = 250 gpm (2,080 lb/min) and $T_W$ = 80°F at a constant load of 20000 Btu/min. Because of the change in slope with flow, an equal-percentage valve characteristic is recommended. From Fig. 3.7, the required flow of water to produce an average coolant temperature of 140°F is found to be 37 gpm. Gain of the process is

$$\frac{dT_c}{dF_W} = -\frac{}{C_3F_W{}^2}$$

$$= -\frac{20000 \text{ Btu/min}}{[1.0 \text{ Btu/(lb)(°F)}][37 \text{ gpm}]^2[8.33 \text{ lb/gal}]} = -1.75\text{°F/gpm}$$

Gain of an equal-percentage valve is simply four times the flow being delivered :

$$G_v = 4 \frac{37 \text{ gpm}}{100\%} = 1.5 \text{ gpm/\%}$$

If a transmitter span of 200°F is selected, $G_T$ = 100 %/200°F, or 0.5 %/°F. The gain product can then be found:

$$G = (0.116)(1.5 \text{ gpm/\%})(-1.75\text{°F/gpm})(0.5\text{\%/°F}) = -0.152$$

(The negative sign indicates the sense of control action.) For ¼-amplitude damping, $P$ must be $200G$, therefore

$$P = 200(0.152) = 30\%$$

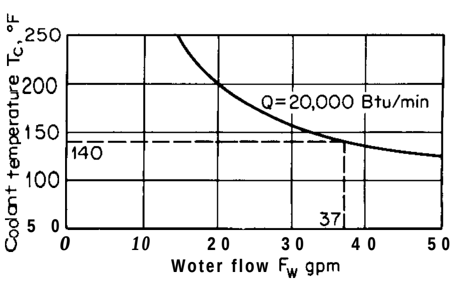



***FIG 3.7. The slope of the process-characteristic curve decreases with increasing flow.***

Derivative is extremely useful in a temperature loop, to compensate for the secondary lags in the heat transfer media and temperature bulb. A derivative time of $35/2\pi$ or 5.6 min on this process would reduce the period to about 20 min, and the required proportional band to the vicinity of 20 percent.

### Summary

The most important points to be grasped from this analysis of a simple heat transfer process are:

1. Time constants in a temperature-control loop are not easy to identify, and they interact.
2. The presence of distributed lags makes the exact performance of the loop difficult to predict.
3. Processes involving heat transfer are always nonlinear in at least-one respect. Each process ought to be evaluated on its own merits to be sure correct compensation is applied.

If the rate of heat evolution in the example had been made a function of temperature, as it is in a real reactor, a second nonlinearity would have made its appearance. Obviously much further consideration must be given to each individual heat transfer application as it is encountered. Although certain characteristics are common, many others are not. In short, there is no such thing as a "typical" temperature-control loop.

## CONTROL OF COMPOSITION

By far the greatest single contributor to the problems of a control engineer is the composition loop. Composition is a property of a flowing stream, therefore it travels with the stream. This means that dead time is always in the loop. Further, sampling difficulties, incomplete mixing, and intermittent analyses lend the measurement a certain amount of random character, often making tight controller adjustment nadvisable. Most significant of all the composition of a stream is a function of the performance of the processing equipment producing it, which many control engineers do not fully understand.

As in the case of temperature control, the process within a composition loop may be extremely complex. In fact, most mass transfer operations require multiple control loops to cope with the number of variables which affect product quality. But for the moment it is important to examine the properties of a composition loop apart from the intrigues of mass transfer. Therefore a simple blending system will be analyzed.

### The Problem of Mixing

A simple composition-control loop is depicted in Fig. 3.8. Its object is to control the percentage of a single component in the effluent solution

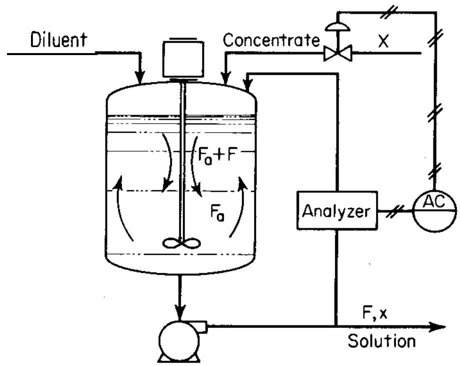



*FIG 3.8. This composition loop is dominated by the transportation of the concentrate to the analyzer.*

by adding the required amount of concentrate to the tank. Assume that no chemical reaction takes place-the problem is then only one of mixing to the desired composition.

Anyone who has tried to control composition in a stirred tank knows that it is not a single-capacity process. It would only be single-capacity if the contents of the vessel were *perfectly* mixed. But no mixer can move material from the inlet pipe to the exit pipe in zero time-it is impossible. Consequently some dead time must exist, i.e., that time required for the agitator to transport a particle of fluid from inlet to outlet. The presence of any dead time changes the control situation entirely, for now the process is capable of oscillating in a closed loop, which places a limitation on both controller gain and speed of response.

If the vessel in Fig. 3.8 had no mixing whatever taking place, the streams entering the top would flow downward as a plug, reaching the exit at time $V/F$ later. In this case $\tau_d = V/F$, whereas the lag $\tau_1 = 0$. If the vessel were perfectly mixed, $\tau_d$ would equal 0 but $\tau_1$ would equal $V/F$. All real situations fall between these two limits.

The performance of an agitator is frequently rated by its pumping capacity. In this way it is treated as if it were a pump circulating fluid from the bottom back to the top of the vessel at a uniform rate. This rate of circulation is labeled $F_a$ in Fig. 3.8. It will then be seen that the time required for a particle to travel from top to bottom of the vessel, i.e., the dead time, is

$$\tau_d = \frac{V}{F_a + F} \tag{3.23}$$

The completeness of mixing may be described as the ratio of upflow to downflow, that is, $F_a/(F_a + F)$. Then the time constant of the vessel can be looked upon as that part of the vessel's capacity, which is completely mixed :

$$\tau_1 = \frac{V}{F}\frac{F_a}{F_a + F} \tag{3.24}$$

Placing $\tau_d$ in the same terms helps to compare the two components:

$$\tau_d = \frac{V}{F}\frac{F}{F_a + F}$$

There are several observations to be made from these two derivations. First, it may be noted that

$$\tau_1 + \tau_d = \frac{V}{F} \tag{3.25}$$

This is reasonable, because it confirms that the average particle cannot be retained in the vessel longer than its residence time ***V/F,*** whether mixed with the rest of the contents or not. Furthermore it concurs with Eq. (2.4).

Second, the difficulty of control, $\tau_d/\tau_1$, varies only with ***F*** and ***F,:***

$$\frac{\tau_d}{\tau_1} = \frac{F}{F_a} \tag{3.26}$$

Notice that difficulty is not a function of volume. This contradicts the commonly accepted rule that relates controllability to volume. In fact, increasing the volume of a system while retaining the same flow and agitation serves only to reduce its speed of. response, because $\tau_d$ would be increased proportionately. The effect can be more readily visualized if carried to extremes: it would be no easier to control composition in a lake than in a small tank, using the same agitator, and response would certainly be slower.

The actual mechanism by which mixing takes place is obviously not discrete, as the dead-time plus lag model would suggest. Flow from the agitator is not in a single direction, as it would be in a pipe, and even if it were, the velocity profile could not be perfectly flat. Furthermore, turbulence is what produces the actual mixing, and turbulence seems to be an omnidirectional effect. Even without an agitator, some mixing always takes place through diffusion and a token amount of turbulence resulting from flow through the vessel.

Tests[4] conducted on stirred tanks show that the response of the effluent to a step change in concentrate flow resembles that of a system comprised of multiple interacting capacities. Figure 3.9 shows the response a typical vessel might produce both with and without agitation.

These response curves are typical of diffusive and distributed processes, as was mentioned in Chap. 2. It was also pointed out how capably this

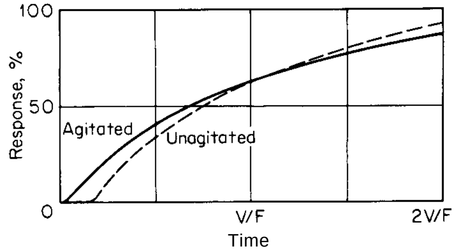



***FIG 3.9. Agitation reduces the effective dead time while increasing the effective time constant.***

sort of response could be represented by dead time plus a single capacity. Thus the simple model just postulated is quite valid, if imperfect.

**The Analyzer**

Dynamics associated with the analysis play an important role in the performance of the loop. The foremost limitation in the speed of analysis is generally that of transporting the sample to the detector. Fortunately some composition measurements can be made without withdrawing a sample: electrolytic conductivity, density, and pH are notable examples. But any analysis requiring the withdrawal of a sample, particularly if that sample must undergo a certain amount of preparation, results in a significant accumulation of dead time (see the example cited at the close of Chap. 2). Naturally any effort spent in minimizing the sampling time will be rewarded by both tighter control and faster response.

Some analyzers are discontinuous. They produce only one analysis in a given time interval. This characteristic is worthy of much more attention, because it periodically interrupts the control loop. Process chromatographs are the principal, but not sole, constituents of this group. The response of this kind of control loop will be given extensive coverage in Chap. 4, and methods for coping with it will be presented.

A few analyzers exhibit a time lag in addition to the dead time associated with sample transport. Sormally this property is of little consequence, except when the process itself consists of nothing but the volume of a pipeline, whose time constant may be less than that, of the analyzer. Those measurements which are fast, are by the same token subject to noise. Conductivity and pH are usually in this category, because they are fast enough to react to an incompletely mixed solution, or particles of an immiscible phase.

Dead time in sample lines is understandably constant. Dead time in a pipe carrying the main stream varies with flow. Dead time within a stirred tank is slightly affected by flow, to the extent of $F/F_o$; in most systems this variation would not be significant. The natural period of the composition loop would therefore be virtually constant, producing constant dynamic gain, except for a process whose dominant element is a pipeline.

Most analyzers are not so far from being linear that they materially affect the gain of the control loop. The notable exception is, of course, the pH measurement, whose general properties have already been presented. But analyzers are generally given a high order of sensitivity, because of the importance placed on quality control. As a result, the gain of a composition-control loop is invariably high. Objectively, composition is not as difficult to control as flow, for example, but the specifications placed on product quality are so stringent that ordinary performance is seldom acceptable. The impurity of a product stream leaving a fractionator, for example, may be specified at 1.0 $\pm$ 0.2 percent. It is virtually impossible to regulate flow within $\pm$1 percent in the unsteady state, yet the composition controller is asked to perform five times as well. This is perhaps the greatest single reason why composition control has the distinction of being a problem area. Because quality can be measured to 0.1 percent is apparently reason enough to expect it to be controlled to the same tolerance.

### Process Gain

The dimensional gain of the process in Fig. 3.8 is the derivative of composition, $x$, with respect to concentrate flow X. A material balance on the measured component is simply

$$X = Fx$$

Then,

$$\frac{dx}{dX} = \frac{1}{F} \tag{3.27}$$

Since the nominal flow $F$ has already been identified as a constant, process gain is also constant. (This is another illustration of the case where process steady-state gain varies with flow, but the time constant does too, so dynamic gain is invariant. Steady-state gain, as calculated above, is only meaningful at the rated flow F.)

Dimensional gain of the composition process can always be found by writing a material balance across it. If composition of an effluent stream is controlled by manipulating an *influent* stream, as in this example, the process is linear. But if effluent composition is controlled by manipulating the *effluent* flow, the process is hyperbolic:

$$x = \frac{X}{F}$$

$$\frac{dx}{dF} = -\frac{X}{F^2} \tag{3.28}$$

(This was already encountered in the temperature-control example where coolant temperature was adjusted by manipulating its flow.) Examples of both linear and hyperbolic processes are common in both composition and temperature applications, because the controlled variable is always a function of the ratio of one variable to another. If the manipulated variable happens to be in the numerator, the process is linear.

**example.3.6**

The process in Fig. 3.8 is intended to deliver a solution at a nominal flow $F$, of controlled composition $x$, by adding a manipulated flow X of concentrate to the diluent stream. Let the volume of the vessel be 100 gal and the nominal flow 20 gpm. If mixing is 95 percent complete, then $0.05V/F$ will be the effective dead time in the vessel:

$$\tau_d = 0.05 \frac{100}{20} = 0.25 \text{ min}$$

The balance is a first-order lag:

$$\tau_1 = 0.95 \frac{100}{20} = 4.75 \text{ min}$$

Let the sampling time also be 0.25 min, with a 3.0-sec analysis lag. The total dead time in the loop is then

$$\tau_d = 0.25 + 0.25 = 0.5 \text{ min}$$

Without the 3.0-sec lag in the analyzer, the natural period would be

$$4\tau_d = 2.0 \text{ min}$$

The phase shift of the 3.0-sec lag at a period of 2.0 min is

$$\phi_2 = -\tan^{-1} \frac{2\pi 3.0/60}{2.0} = -9°$$

A control valve with a 3.0-sec lag will contribute another 9". This added phase shift extends the natural period to approximately

$$\tau_o = 2\,0 \frac{180 + 9 + 9}{180} = 2.2 \text{ min}$$

The dynamic gain of the process is simply that of the principal time constant:

$$G_1 = \frac{\tau_o}{2\pi\tau_1} = \frac{2.2}{2a4.75} = 0.0737$$

Dimensional process gain is the percent composition change brought about by a change in concentrate flow at the rated throughput:

$$\frac{dx}{dX} = \frac{1}{F} = 100\%/20 \text{ gpm} = 5\ \%/\text{gpm}$$

Because the process is linear with respect to concentrate flow, a linear valve is chosen. Let the maximum flow of concentrate be 2 gpm. Then

$$G_v = 2 \text{ gpm}/100\% = 0.02 \text{ gpm}/\%$$

To illustrate the close tolerances to which product quality is generally specified, the analyzer range will be chosen as 4.5 to 5.5 percent, with a normal set point of 5.0 percent. The span is 1.0 percent:

$$G_T = 100\%/1\% = 100$$

The proportional band necessary for ¼-amplitude damping is finally estimated as 200 times the gain product:

$$P = 200(0.0737)(5\ \%/\text{gpm})(0.02\text{ gpm}/\%)(100) = 147\%$$

For a process that is really not very difficult, this is quite a wide proportional band. An extremely wide band may then be expected in a truly difficult application, indicating how sensitive composition loops are to changes in load.

To summarize, composition loops are principally comprised of dead time plus a single capacity and are noted for high transmitter gain. As a result, a wide proportional band is usually needed, leaving the controlled variable quite susceptible to load changes.

## CONCLUSIONS

The purpose of this chapter has, been to acquaint the reader with the properties of common process control loops and the reasons for these properties. Analysis served as a useful tool to present the case, while

**TABLE 3.3 Properties of Common Loops**

| Property | Flow and liquid pressure | Gas pressure | Liquid level | Temperature and vapor pressure | Composition |
|---|---|---|---|---|---|
| Dead time. | No | No | No | Variable | Constant |
| Capacity. | Multiple | Single | Single | 3-6 | 1-100 |
| Period. | 1-10 sec | Zero | 1-10 sec | Min – hrs | Min – hrs |
| Linearity. | Square | Linear | Linear | Nonlinear | Either |
| $G_vG_T$. . | 2-5 0.5-1 | Integrating | Integrating | 1-2 | 10–1,000 |
| Noise. | Always | None | Always | None | Often |
| Proportional | 100-500% 50-200% | 0-5% | 5-50% | 10-100% | 100–1,000% |
| Reset. | Essential | Unnecessary | Seldom | Yes | Essential |
| Derivative | No | Unnecessary | No | Essential | If possible |
| Valve. | Linear | Eq. percent | Linear | Eq. percent | Linear |

* Applies to liquid pressure.

at the same time demonstrating how to identify the significant elements in a loop. Rarely will a flow or level loop need analysis, but when composition-control problems arise, this procedure can be of inestimable value.

Much of what has been derived and weighed and discussed in the foregoing pages is summarized in Table 3.3.

Nothing that has not been already covered is presented in the table, yet gathering all this information together discloses some interesting features. Notice, for example, the similarity between level and flow loops, with respect to both natural period and the presence of noise. Without any doubt, however, each of the five groups above is separate and distinct from the rest.

## REFERENCES

1. Bradner, M.: Pneumatic Transmission Lag, ISA Paper No. 48-4-2.
2. Catheron, A. R.: Factors in Precise Control of Liquid Flow, ISA Paper No. 50-B-2.
3. Considine, D. M.: "Process Instruments and Controls Handbook," chap. 2, McGraw-Hill Book Company, New York, 1957.
4. Esterson, G. L., and R. E. Hamilton: Dynamic Response of a Continuous Stirred Tank, presented at the Joint Automatic Control Conf., Palo Alto, Calif., June, 1964.

## PROBLEMS

**3.1** A volume booster installed at the inlet to the valve motor of Example 3.2 reduces its time constant to 0.5 sec. Predict the period of oscillation that will result from the change, allowing 45" phase lag in the proportional-plus-reset controller. Calculate the proportional band and reset time for ¼-amplitude damping.

**3.2** What would be the estimate of the natural period and proportional band in Prob. 3.1 if the dynamic elements were all assumed to be dead time rather than capacities? Is this a valid approximation? Why?

3.3 Let pressure downstream of the valve in Example 3.2 be controlled instead of flow. At no flow, there is a static head of 5 psig, while 10 gpm will raise the pressure to 13 psig; the range of the pressure transmitter is 0 to 25 psig. Estimate what the proportional band of the controller will be for ;/,-amplitude damping with the period and reset time used in the example.

3.4 A mercury manometer capable of reading ± 15-in. differential pressure is used to indicate the flow in a gas stream. What is its natural period? How would it affect the control of flow?

3.5 To verify the choice of an equal-percentage valve for Example 3.5, calculate the process gain and the product of process and valve gain for heat loads of 5000, 10000, and 15000 Btu/min; assume that. the difference between controlled reactor temperature and average coolant temperature varies linearly with heat transfer rate.

3.6 Two fluids are blended in a pipeline 20 ft upstream of where the mixture is sampled. The pipe contains 0.4 gal/ft of length, and the flow rate of the blend varies from 10 to 80 gpm. Dead time in the sample line to the analyzer is 15 sec. A circulating pump is installed to maintain 100 gpm flow through that 20-ft section of pipe without affecting the throughput. Compare the natural period for integral control with and without the pump in operation. What else does the pump provide?

3.7 In the same process, the flow of additive is manipulated through a linear valve whose maximum flow is 1.2 gpm. The range of the analyzer is 0 to 1 percent, additive concentration. Estimate the proportional band required for at least ¼-amplitude damping if the reset time is set for 60° phase lag with the pump operating.

# Selecting the Feedback Controller

PART 2

# Linear Controllers

CHAPTER 4

Now that the characteristics of typical processes have been thoroughly presented, it is possible to look more closely into various means for controlling them. The range of process difficulty has been seen to vary from zero to several hundred, as measured by the proportional band needed for damping. The very existence of such-a range of control problems suggests the possibility of a variety of means for their control. The first distinction to be made is between linear and nonlinear control methods.

A linear device is one whose output is directly proportional to its input(s) and any dynamic function thereof. This definition includes not only proportional controllers, but those with reset, derivative, lag, dead time—in short, any time function of a linear variable. To be sure, a device is only linear over a specified range. A pneumatic controller, for example, ceases to operate linearly when its output falls to zero or reaches full supply pressure. All linear devices are similarly limited, and their proper use demands an appreciation of these limitations.

In the earlier chapters certain nonlinear characteristics were dealt with, both in processes and in the measuring devices and valves. An attempt was made in every case to compensate for process nonlinearities so as to obtain constant loop gain. This assures uniformity of performance under all conditions of operation. In general, compensation is effected external to the controller, leaving the controller as a linear device.

But within the domain of linear controllers, a variety of dynamic elements exists. Each dynamic element, such as reset or derivative, has certain undesirable properties along with those which are beneficial. A thorough understanding of the assets and liabilities of each control mode is prerequisite to their intelligent selection.

## PERFORMANCE CRITERIA

If selection between various control configurations is to be made, some basis must be established for their comparison. For example, a given process may be controlled in a number of ways. One way will be better than the others from the standpoint of performance, i.e., how it responds to a set-point or load change. The three load-response curves in Fig. 1.13 show the performance of three different controllers in the same process.

The shape of the load-response curve depends to a considerable degree on the type of control action used and the settings of the parameters involved. Furthermore, the penalty ascribed to a typical response curve is determined by the specifications of the process. Several means of weighing the response curve suggest themselves:

1. Integrated error: Since the error $(r - c)$ can be either positive or negative, an integrated error of zero could be obtained in a continuously oscillating loop. Integrated error is therefore not, of itself, a measure of stability.

2. Error magnitude: This criteria allows the possibility of offset (a small permanent error), which is generally undesirable in any loop.

3. Integrated absolute error (IAE) : This is a measure of the total area under the response curve on both sides of zero error. It is one of the generally accepted performance criteria. Since the error following a load change eventually disappears, the IAE approaches a finite value for any stable loop.

4. Integrated square error (ISE): The instantaneous error is first squared and then summed (integrated). Squaring prevents a negative error from canceling a positive one (as does absolute value) and also weighs large errors more heavily than small ones.

5. Root mean square (rms) error: This index is the standard deviation of the error. If the error reduces to zero with time, so does the rms

error-so this criterion is only applicable to those systems without a steady state.

Technically, the IAE and the ISE are the only all-encompassing indices of performance. The principal distinction between them is the weight placed on large errors. Two response curves with the same IAE would have different values of ISE if there were a difference in error magnitude. For this reason, the ISE criterion is seen to be a combination of error magnitude and IAE.

For the case where the response curve lies wholly on one side of zero error, the integrated error equals the IAE. But this is not the limit of usefulness of the integrated error, for it represents the average error that has existed over a particular time span. The average error or integrated error is a valid basis for comparing response curves with equal damping, like those comparisons shown in Fig. 1.13. By specifying the damping, the objection raised in number 1 above is overruled. Integrated error will therefore be used as a performance index throughout the balance of the book, and in every case ¼-amplitude damping will be meant, unless otherwise indicated.

**Sensitivity of a Process to Disturbances**

The choice of integrated error as a performance index has a very practical aspect, in that it can be readily calculated from controller settings. In a proportional-plus-reset controller,

$$m = \frac{100}{P}\left(e + \frac{1}{R}\int e\,dt\right)$$

Prior to a load change, at time $t_1$, the output will be stationary at a level $m_1$, and the error will be zero. After the transient from a load change has subsided, i.e., at time $t_2$, the output will come to rest at a new level $m_2$, at which the error will again be zero. Then, subtracting the two outputs,

$$m_2 - m_1 = \frac{100}{P}\left(\frac{1}{R}\int_0^{t_2} e\,dt - \frac{1}{R}\int_0^{t_1} e\,dt\right)$$

Reducing the last expression yields

$$\Delta m = \frac{100}{P}\left(\frac{1}{R}\int_{t_1}^{t_2} e\,dt\right) \tag{4.1}$$

Let the integrated error resulting from the load change $\Delta m$ be designated $E$:

$$E = \int_{t_1}^{t_2} e\,dt$$

Then the load response of a given control loop can be assessed on the basis of integrated error per unit load change:

$$\frac{E}{\Delta m} = \frac{PR}{100} \tag{4.2}$$

Again this is integrated error and not IAE; the damping of the loop must be assured before this ndex can be used.

The load-response criterion $E/\Delta m$ depends on the proportioual baud and reset time which, in turn, depend on the characteristics of the plant. This is another way of illustrating the difficulty of control which was described in Chap. 1. If the proportional baud can be made to approach zero because of he ease with which the process can be controlled, or the reset time because of its speed of response, $E/\Delta m$ will approach zero. The integrated error will be found useful in evaluating not only the difficulty of a process, but also the effectiveness of the means used in its control.

### Error Magnitude

The magnitude of an error is a function of how fast the load change takes place. If the load change is very gradual, several orders of magnitude longer in duration than the reset time, it may produce no measurable error magnitude; the ntegrated error, however, does not depend on rate of change of load. An instantaneous load change will be countered by proportional control action, and derivative, if used. If the proportional-plus-reset control equation is written in the differential form,

$$\frac{dm}{dt} = \frac{100}{P}\left(\frac{de}{dt} + \frac{e}{R}\right) \tag{4.3}$$

A plot of $e$ versus the rate of change of load $dm/dt$ can be constructed from it (see Fig. 4.1). The maximum value of $e$ is limited by proportional action to $P\ \Delta m/100$. Derivative action can reduce the effect of a rapid load change principally by allowing a reduction in the proportioual band setting.

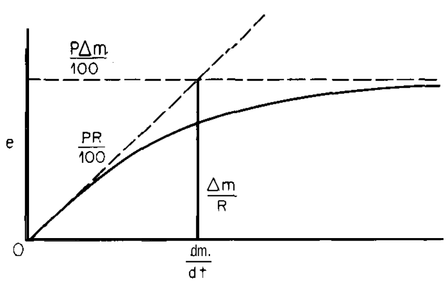



*FIG 4.1. The magnitude of the error is a function of the rate of change of load as well as its magnitude.*

## TWO- AND THREE-MODE CONTROLLERS

Although the primary functions of proportional, derivative, and reset have already been introduced, many of their features remain to be defined. The discussion will be restricted to the commonly available controllers, i.e.,. proportional-plus-derivative, proportional-plus-reset, and proportional-plus-reset-plus-derivative. (Reset controllers are rarely used in process work and are not available as standard items from most manufacturers. Derivative by itself is not recognized as a controlling mode.)

### Limitations of Derivative

It has been pointed out that perfect derivative action is not available in conventional controllers. Derivative gain is limited to about 10, and the maximum available phase lead is in the vicinity of 45". In effect, derivative is therefore accompanied by a lag, whose time constant is 1/10 the value of the derivative time.

In most controllers, derivative acts on the output. Physically, it is a lag introduced in the feedback path around the controller amplifier. Therefore, if he output of the controller is constant, no derivative action will take place no matter what the controlled variable may be doing. This situation occurs whenever the controller's output has reached one of its limits. The output of a pneumatic controller, for example, can go as low as zero or as high as 20 psi (the supply pressure), although the range of the control valve-hence the proportional band-is 3 to 15 psi. No derivative action will take place, then, until the controlled variable approaches the proportional band as in Fig. 4.2. In the discussion on two-capacity processes, this property placed a limitation on the width of the proportional band required for critical damping.

In most controllers, derivative action does not distinguish between measurement and set point. The purpose of derivative is to speed the response of the closed loop, but the set point lies outside the loop. The controlled variable cannot change instantaneously, because of the lags inherent in the process. But it is normal to introduce set-point changes instantaneously, which derivative action amplifies into gross output

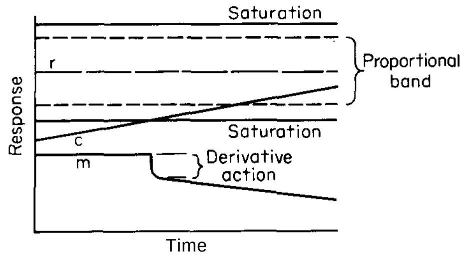



***FIG 4.2. Derivative action begins when the controller comes out of saturation.***

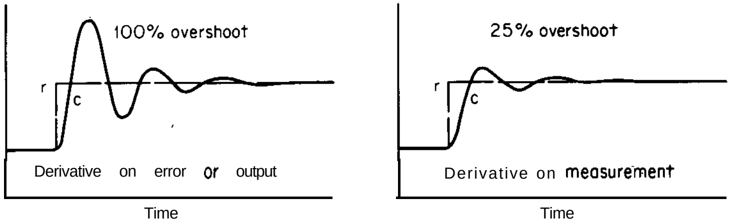



***FIG 4.3. Derivative makes a controller hyper-sensitive to set-point changes.***

fluctuations. Consequently, properly adjusted derivative acting upon the error or the output is hypersensitive to set-point variations. Ideally, derivative should act on neither the output nor the error, but on the measurement alone. Figure 4.3 shows the effect of derivative acting on the set point.

Response can be improved markedly by introducing a lag in the set-point circuit. Figure 4.4 indicates the sort of results that can be obtained with such an arrangement.

Set-point changes are often introduced automatically, by the output of another controller in cascade (see Chap. 6). The controller whose set point is adjusted in this way cannot tolerate derivative, especially if the primary (adjusting) controller has it.

## Reset "Windup"

Whenever a sustained deviation is imposed upon a controller containing reset, its output will eventually be driven off scale. This will happen whenever the loop is opened, as in the case of plant shutdown or transfer to manual control. If the measurement has been held below the set point, the controller will be integrating so as to raise it. When the loop is closed again, the measurement will be driven above the set point and the controller must integrate back down again to the normal output.

When a process is shut down by the closing of hand valves, reset action begins to force the proportional band of the controller upward to its

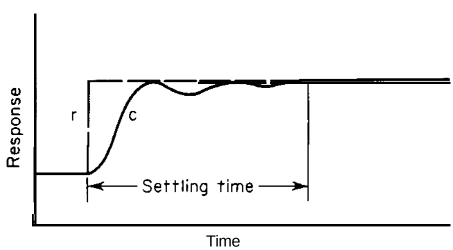



***FIG 4.4. Introducing a lag in the set-point circuit can eliminate overshoot and reduce settling time.***

limit, trying to raise the measurement. The controller then finds that even 100 percent output will not bring the deviation to zero, whereas, in normal operation, 50 percent may have been enough. Figure 4.5 shows the proportional band at its limits during shutdown, such that no control action can begin again until the set point is crossed. If derivative acts on the controller output, it is powerless to help the situation, because it is disabled as long as the measurement is outside the proportional band.

Notice that the proportional band is actually beyond the set point. Reset forces it up to the full available supply pressure (or voltage), which is always well in excess of 100 percent output. In a pneumatic controller, supply pressure is normally 20 psi. During prolonged shutdown, pressure in the reset bellows will reach 20 psi. If the error were suddenly reduced to zero under this condition, controller output would be 20 psi, equal to the pressure in the reset bellows. (In the case of a reverse-acting controller, or when the error is sustained in the other direction, reset pressure will fall to 0 psi, which is 3 psi below 0 percent output.)

With regard to batch processes which must be started up several times a day, this problem is serious. Particularly demanding is temperature control of a batch chemical reactor, where overshoot is intolerable. The situation can be improved to some degree with a controller whose derivative acts upon the input.' Many new controllers incorporate this feature.

A logical solution to the reset-windup problem is to add enough intelligence to the controller to make it aware of a shutdown condition. This is done by placing in the controller's reset circuit a switch energized by the output. Whenever the output exceeds 100 percent, the switch disables the reset circuit, leaving a proportional (or proportional-plus-derivative) controller. In the absence of automatic reset action, a bias must take its place. Because this bias equals the output of the controller at zero deviation, it is ordinarily adjusted in relation to the expected process load. For this reason it is sometimes called the "preload" setting.

If the preload setting is too high, response is similar to that without the "antiwindup" switch (Fig. 4.5). With too little preload, throttling begins prematurely and the controller must bring the measurement to the

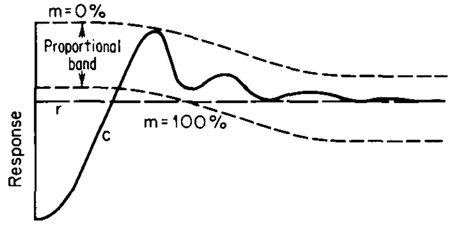



***FIG 4.5. Control action does not begin until after the set point is crossed, therefore overshoot is inevitable.***

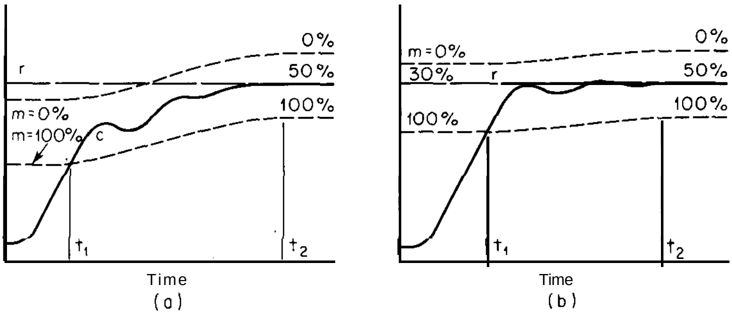



***FIG 4.6. Response after startup with (a) no preload, (b) 30 percent preload.***

set point by reset action-a slow procedure. The results of inadequate and sufficient preload in a three-mode controller are displayed in Fig. 4.6.

The first crest in the record of Fig. 4.6*a* is brought about by the sudden activation of derivative as the controller output starts to fall. Proper adjustment of preload will locate this crest right at the set point, as shown in Fig. 4.6*b*. In the absence of derivative, there is no crest, and the measurement will be found to converge on the set point asymptotically.

Some integration must take place, because a sizable error always exists at the time $t_1$ when the "antiwindup" switch activates the reset circuit. The load $q$ (output at time $t_2$) will differ from the preload $b$ by the amount of this integration:

$$q - b = \frac{100}{PR} \int_{t_1}^{t_2} e \, dt \tag{4.4}$$

If the preload is set equal to the expected load, some overshoot will develop. Therefore, if critical damping is to be achieved, the preload setting must be significantly less than the expected load.

In Fig. 4.6*a*, "no preload" is indicated. Because no preload is less than 0 percent (0 psi compared to 3 psi in a pneumatic system), the proportional band lies below the set point by this difference.

When placing one of these controllers into operation, it is first necessary to adjust proportional, derivative, and reset for maximum performance in the steady state, just as would be done for a continuous process. Once these are set, response during startup depends entirely on the value of preload.

## Auto-manual Transfer

As far as the controller amplifier is concerned, the loop is open when in manual control. The slightest deviation will eventually cause the controller to saturate, because it has no way of satisfying itself in a closed loop. For this reason, all controllers with an auto-manual transfer feature also have some means of preconditioning the reset so that the proper output will be maintained during transfer to automatic. This is known as "bumpless" transfer. Some controllers are designed to insure bumpless transfer even if a deviation exists. But each controller design seems to have its own particular transfer procedure, so there is little point in trying to generalize. The reader is advised to consult the instructions supplied by the manufacturer.

### The Three-mode Controller

An ideal three-mode controller consists of a simple combination of the individual modes as they have already been presented:

$$m = \frac{100}{P}\left(e + \frac{1}{R}\int e\,dt + D\frac{de}{dt}\right) \tag{4.5}$$

Because the three components of gain at the natural period are out of phase, vector addition is required to determine the resultant phase and gain of the controller. A vector diagram is given in Fig. 4.7.

Because of the gain limitation placed on the derivative mode, the latter is not exactly represented by a vertical vector, but the inaccuracy is not severe above a period of $2\pi D$. From the vector diagram of Fig. 4.7, the resultant phase and gain of the ideal controller are

$$\phi = \tan^{-1}\left(\frac{2\pi D}{\tau_o} - \frac{\tau_o}{2\pi R}\right) \tag{4.6}$$

$$G = \frac{100}{P}\left[1 + \left(\frac{2\pi D}{\tau_o} - \frac{\tau_o}{2\pi R}\right)^2\right]^{1/2} \tag{4.7}$$

### Interaction between Control Modes

Whenever reset and derivative operate successively on a signal, they interact with one another. Reset produces a rate of change of output in an attempt to restore the measurement to the control point. But derivative reacts to a rate of change of output.

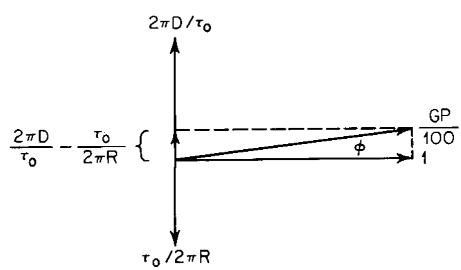



*FIG 4.7. The resultant gain and phase emanates from a vector summation of the individual modes.*

When a signal passes through derivative and reset in series, as in Fig. 4.8*a*, no matter in what order or whether one or two stages of amplification are used, interaction results—the ultimate control function is not ideal. Figure 4.8*b* shows a parallel dynamic operation, followed by summation, producing the ideal (noninteracting) control response. The block diagram of the noninteracting controller can be reduced mathematically to Eq. (4.5). But reduction of the diagram of Fig. 4.8*a* to a comparable mathematical expression yields

$$m = \frac{100}{P}\left(1 + \frac{D}{R}\right)\left(e + \frac{1}{R+D}\int e\,dt + \frac{de/dt}{1/D + 1/R}\right) \tag{4.8}$$

A remarkable fact is that at this writing *all* standard three-mode controllers are interacting. Although the mathematical structure of a noninteracting controller is simpler, its implementation is too costly to be competitive at present.

Interaction is manifest in the effectiveness of the three-mode adjustments. Let the effective mode value be indicated by a prime, relative to the settings introduced:

$$P' = \frac{P}{1 + D/R} \qquad R' = R + D \qquad D' = \frac{1}{1/D + 1/R} \tag{4.9}$$

Several important observations can be made from the above relationships:

1. Where $D > R$, derivative time is affected more by the reset setting, and vice versa.

2. It is impossible to make the effective derivative time equal to or greater than the effective reset time.

3. As $D$ approaches $R$, further adjustment will produce very little change in $D'$, so there is little purpose in trying to "fine tune" an interacting controller.

To illustrate the above points, Table 4.1 lists several combinations of settings together with their effects. This example was chosen to show how three markedly different combinations of adjustments can result in substantially the same effective values. Some engineers have prepared

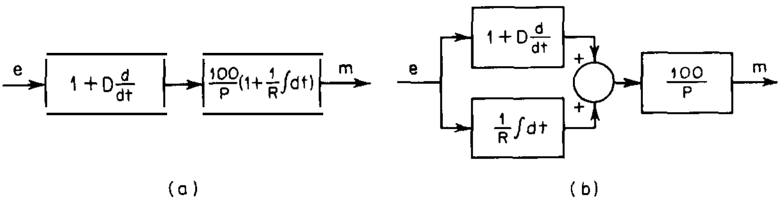



*FIG 4.8. Three-mode controllers: (a) interacting, (b) noninteracting.*

**TABLE 4.1 Effective Values of Modes at Various Settings**

| $P$ | $P'$ | $R$ | $R'$ | $D$ | $D'$ |
|---|---|---|---|---|---|
| 20 | 10 | 1.0 | 2.0 | 1.0 | 0.5 |
| 15 | 10 | 1.5 | 2.25 | 0.75 | 0.5 |
| 13.3 | 10 | 2.0 | 2.67 | 0.67 | 0.5 |

"tuning maps," in which the response of a given process to various combinations of proportional, reset, and derivative settings are compared. The results would bear some merit if a noninteracting controller were used, but interaction limits the range of effective settings too stringently. It is doubtful whether any difference in response obtained with the three combinations in Table 4.1 would be noticeable. The effective values of proportional and derivative are the same in each case, while effective reset time only changes from 2.0 to 2.67 min-hardly noticeable.

The integrated error is, of course, a function of the effective values of proportional and reset:

$$\frac{E}{\Delta m} = \frac{P'R'}{100}$$

Substitution for $P'$ and $R'$ reveals the relationship of integrated error to the settings of the two modes:

$$\frac{E}{\Delta m} = \frac{P(R + D)}{100(1 + D/R)} = \frac{PR}{100}$$

The integrated area is unaffected by this interaction.

A more severe form of interaction exists in pneumatic controllers whose reset and derivative circuits are in parallel feedback about the amplifier. Pneumatic controllers equipped with an "antiwindup switch" are connected in this way. The only difference between these and the conventional interacting controller is in the effective proportional band:

$$P' = P\left(\frac{1 - D/R}{1 + D/R}\right) \tag{4.10}$$

When $D = R$, the effective proportional band is zero. And if $D > R$, the effective band actually becomes negative; negative proportional band in a negative feedback controller means positive feedback. Extreme care must therefore be exercised when adjusting one of these controllers, or the results could be disastrous.

### Adjusting Two- and Three-mode Controllers

In order to formulate an effective procedure for adjusting controllers, it is first necessary to determine where the optimum values lie. To permit extension to a broad range of difficult processes, a dead-time plus integrating process will be used, with controller settings left in terms of $\tau_d$ and $\tau_1$. This is necessary because the natural period of a loop does vary with the controller settings. Table 4.2 has been prepared by equating the gain product of process and controller to 0.5, with gain and phase for the controllers accurately calculated from the vector diagrams of Figs. 1.11 and 4.7.

Several significant conclusions may be drawn from Table 4.2. Derivative is very effective in improving the performance of the loop, even though the dominant secondary element is dead time. But the reason that it is so effective is because it offsets the phase lag of reset, preventing the period and process gain from increasing in its presence. A striking characteristic of the optimum controller setting for both of the three-mode controllers is that the phase contribution of the controller is zero at the natural period. Furthermore, derivative and reset times should be equal. Thus the optimum settings can be predicted with accuracy knowing either the dead time of the process or the period of oscillation under proportional control.

From these conclusions, it is possible to formulate a rigorous adjustment procedure for three-mode controllers:

1. With maximum reset time and minimum derivative, excite the closed loop into oscillation by reducing the proportional band.

**TABLE 4.2 Determination of Optimum Settings for Two- and Three-mode Controllers**

| Modes | Controller | | | | | Process | | Performance, $PR/100$ |
|---|---|---|---|---|---|---|---|---|
| | $D$ | $R$ | Gain | $\phi$, deg | $P$ | $\tau_o$ | Gain | |
| Proportional plus reset | 0 | $\infty$ | $100P$ | 0 | $127\tau_d/\tau_1$ | $4.00\tau_d$ | $0.64\tau_d/\tau_1$ | $\infty$ |
| | 0 | $2.86\tau_d$ | 104 | -15 | 158 | 4.80 | 0.76 | $4.52\tau_d^2/\tau_1$ |
| | 0 | 1.66 | 115 | -30 | 220 | 6.00 | 0.95 | 3.65 |
| | 0 | 1.27 | 141 | -45 | 361 | 8.00 | 1.27 | 4.58 |
| Noninteracting three-mode | $0.51\tau_d$ | $0.88\tau_d$ | $109P$ | +23 | $110\tau_d/\tau_1$ | $3.19\tau_d$ | $0.51\tau_d/\tau_1$ | $0.97\tau_d^2/\tau_1$ |
| | 0.64 | 0.64 | 100 | 0 | 127 | 4.00 | 0.64 | 0.81 |
| | 0.86 | 0.49 | 109 | -23 | 196 | 5.37 | 0.86 | 0.96 |
| Interacting three-mode | $0.90\tau_d$ | $0.90\tau_d$ | $125P$ | +36 | $228\tau_d/\tau_1$ | $2.86\tau_d$ | $0.45\tau_d/\tau_1$ | $2.05\tau_d^2/\tau_1$ |
| | 0.64 | 0.64 | 100 | 0 | 254 | 4.00 | 0.64 | 1.62 |
| | 0.54 | 0.54 | 125 | -36 | 540 | 6.79 | 1.08 | 2.92 |

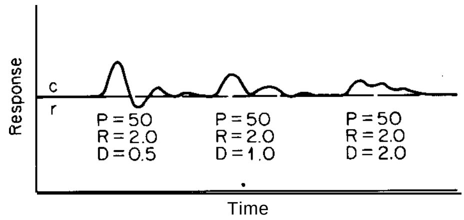



***FIG 4.9. Increasing derivative time reduces error magnitude at the cost of recovery time.***

2. Measure the period of oscillation $\tau_o$ and set derivative and reset time both to $\tau_o/2\pi$. (The optimum value of ***D*** and ***R*** in Table 4.2 is $0.64\tau_d$, which is $2\tau_d/\pi$ or $\tau_o/2\pi$.) For a two-mode controller, set ***R*** at $\tau_o/2.4$. When adjusting a three-mode pneumatic controller with antiwindup, always keep $R > 2D$ to retain proportional stability.

3. Readjust the proportional band to produce the desired degree of damping.

4. If $\tau_o$ is higher than before, increase both ***D*** and ***R;*** if it is lower, decrease them. This may be necessary to compensate for inaccuracy in the dial settings. With a two-mode controller, $\tau_o$ will increase by about 50 percent.

Interaction is clearly evident when an attempt is made to tune a controller for maximum performance. Figure 4.9 shows how increasing the derivative time also increases the effective reset time in a typical closed loop.

The integrated errors for the three response curves are equal because the proportional band and reset settings are all the same. Because the last curve has a lower maximum error, its ISE would tend to be less than the other two.

Although noninteracting controllers are not generally available, it is worthwhile to note that they are capable of twice the performance of interacting controllers, as Table 4.2 indicates.

## COMPLEMENTARY FEEDBACK

The question often arises whether proportional, reset, and derivative are really the best control modes for every application. For the easier-to-control processes, their use can be justified. A single-capacity process and some two-capacity processes need only narrow-band proportional action. Derivative is of great value in processes with two or three capacities. But for the more difficult processes, it has been found that reset action is essential.

In processes where dead time is dominant, derivative action has been seen to have less effect than in processes consisting of capacity alone. This suggests that some other control mode more akin to dead time might be valuable. Derivative and reset are, in actuality, time constants just like the time constants in a process. They bear no resemblance to the dead time that may exist in the plant, however.

## Theory

Several authors[2,3] have postulated a feedback control system that is modeled after the process. This kind of control action is known as "complementary feedback," because the characteristics of the controller complement the dynamics of the process. A block diagram showing both process and complementary controller appears in Fig. 4.10.

The principle employed is that a given error signal e can be,made to generate a certain instantaneous restoring force, $100e/P$, which will change c exactly enough to cancel the error. At the same time, a complementary signal characterized to match the response of the process is fed back positively to cancel the effect of the negative feedback from the process. This means that the output of the controller will remain at its instantaneous value, which was correct in that it was able to exactly reduce the error to zero.

For this exact sequence of events to occur, the control parameters must have the values

$$P = 100K_p \qquad \mathbf{g}_c = \mathbf{g}_p \tag{4.11}$$

The term $K_p$ is the steady-state gain of process, valve, and transmitter, that is, $G_pG_vG_T$. The terms $\mathbf{g}_c$ and $\mathbf{g}_p$ are vectors representing the dynamic components in the controller and process, respectively.

It is worthwhile to trace the sequence of events following a set-point change through the block diagram. Initially e is zero because $c = r$, and $m$ is at rest. Following a set-point change Ar,

$$e = \Delta r$$

and the output jumps instantaneously by

$$\Delta m = 100 \frac{e}{P}$$

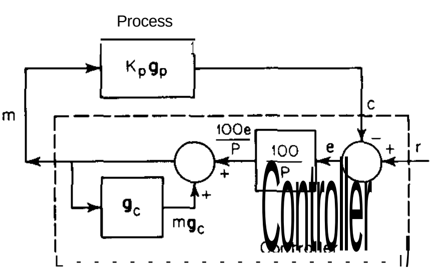



*FIG 4.10. The complementary controller has a model of the process in its positive feedback loop.*

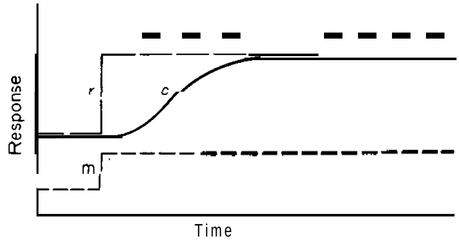



***FIG 4.11. With complementary feedback, m images r, producing response which appears to be open loop.***

The output change Am proceeds through the process to cause

$$\text{Ac} = \text{Am } K_p g_p = 100 \text{ Ar } K_p \frac{g_p}{P}$$

Since $100K_p/P = 1$, the set-point response is

$$\text{Ac} = \Delta r \, g_p$$

Passing through the subtracting junction,

$$e = \text{Ar} - \text{Ac} = \text{Ar} -- \text{ Am } K_p g_p$$

The signal then appears at the summing point as

$$100; = 100 \frac{\Delta r}{P} - 100 \frac{\Delta m \, K_p g_p}{P} = 100 \frac{\Delta r}{P} \quad \text{Am } g_p$$

Within the controller, complementary feedback is sending $+\Delta m \, g_p$ to the same summing point, such that ***Am*** will retain its original value of 100 $\Delta r/P$ as e returns to zero.

The controlled variable responds as it, would if the loop were open, because the output of the controller is constant after the new set point has been inserted. Overshoot is impossible with this control arrangement, so critical damping is always attained. Figure 4.11 shows how a typical process might respond to a step set-point change.

The principal advantage of this type of control scheme is that critical damping can be achieved with a proportional loop gain of unity; in most control loops, the proportional gain exceeds unity. In a single-capacity process, for example, the proportional band may be reduced to zero, placing the proportional loop gain at infinity. Yet there are some processes, notably those dominated by dead time, in which the loop gain must be much less than unity to obtain the desired damping.

Figure 1.26 shows the required proportional band for ¼-amplitude damping for any combination of dead time and capacity. A band of 100 percent (proportional gain of 1.0) is seen to be required for a process whose $\tau_d/\tau_1 = 1.2$. But with complementary feedback, the same proportional gain could produce critical damping. Complementary feedback is, by this token, of advantage in the most difficult processes.

### For Dead Time

In theory, complementary feedback is capable of critically damping a process consisting of pure dead time. Following the lines of the example given for proportional control of dead time in Chap. 1, the advantages of complementary feedback will be demonstrated. For simplicity, let $K_p = 1$. The controller output is

$$m_n = \frac{100}{P}(r - c) + m_{n-1}$$

where $n = t/\tau_d$. The process responds:

$$c_n = m_{n-1}$$

Starting at conditions $r_0 = c_0 = m_0 = 0$ percent, $P = 100$ percent, a set-point change of 50 percent initiates the following sequence:

| | | |
|---|---|---|
| $r_0 = 0\%$ | $c_0 = 0\%$ | $m_0 = 0\%$ |
| $r_1 = 50$ | $c_1 = 0$ | $m_1 = 1.0(50 - 0) + 0 = 50$ |
| | $c_2 \doteq 50$ | $m_2 = 1.0(50 - 50) + 50 = 50$ |

The process comes to rest in one dead time.

The best value of $100K_p/P$ is 1.0. At 2.0, the loop becomes undamped, while at $<1.0$, damping is heavier than critical. Figure 4.12 illustrates the effect of changing gain. So some variation in gain can be tolerated. Unfortunately the same is not true for the complementary feedback term g,. If the positive feedback arrives at a different time than the negative feedback from the process, the loop mill break into oscillations of two periods, which are the sum and the difference of the two dead times.

Perhaps the most significant features of complementary feedback, as brought out in this example, are:

1. No offset
2. Fast response $(\tau_o = 2\tau_d)$
3. Availability of critical damping

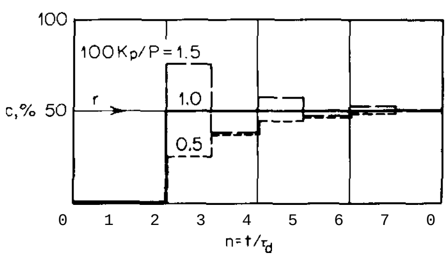



*FIG 4.12. The amount of damping varies inversely with loop gain.*

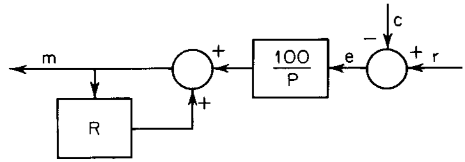



***FIG 4.13. A proportional-plus-reset controller is the complement of a single-capacity process.***

Integral control of dead time was able to eliminate offset, but at $\tau_o = 4\tau_d$. Neither proportional nor integral control was capable of critically damping the loop.

### For Single Capacity

Although there is no need to use a complementary controller on simple processes, it is nevertheless interesting to speculate on its configuration. If the process is a first-order lag, its complementary controller turns out to be proportional-plus-reset. In fact, pneumatic two-mode controllers are made this way, as shown in Fig. 4.13.

A single-capacity process can tolerate zero proportional band and zero reset time. Compared to a dead-time process, it can be concluded that the easier the process is to control, the less critical are its mode adjustments.

### Load Response

In the case of a dead-time process with perfect complementary feedback, a step disturbance in load would produce an error corrected one dead time later. Figure 4.14 shows the results.

The magnitude of the error is $\Delta c = \Delta m K_p$, and its duration is $\tau_d$. Therefore the integrated error per unit load change is

$$\frac{E}{\Delta m} = K_p \tau_d \tag{4.12}$$

To properly evaluate this response, a two-mode controller will be applied to the same process, but it must be adjusted so that the error will not cross zero during recovery. Thus integrated error is actually IAE, permitting comparison of loops with different damping. A phase

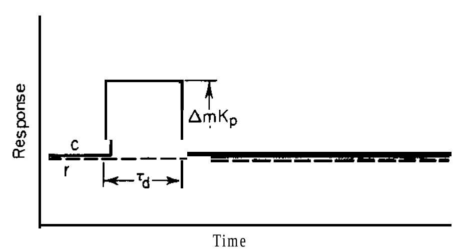



***FIG 4.14. A load-induced error prevails for one dead time.***

lag of 22.5” is chosen for the controller, such that $\tau_o = 2.3\tau_d$. The tangent of 22.5' is 0.414. Reset time is then

$$R = \frac{2.3\tau_d}{2\pi(0.414)} = 0.88\tau_d$$

Quarter-amplitude damping requires loop gain to be 0.5.

$$G_{PR}K_p = 0.5 = \frac{100K_p}{P}\sqrt{1 + (0.414)^2}$$

$$P = 200K_p\sqrt{1.17} = 216K_p$$

Integrated error is then

$$\frac{E}{\Delta m} = \frac{PR}{100} = (2.16K_p)(0.88\tau_d) = 1.90K_p\tau_d$$

This indicates that complementary feedback significantly reduces IAE, as well as settling time.

It is important to see over what range of processes complementary feedback has an advantage over two-mode control. A single-capacity plus dead-time process will respond to a step load change under complementary feedback as shown in Fig. 4.15. Without going into the derivation of the load-response curve, it turns out that the integrated area per unit load change is

$$\frac{E}{\Delta m} = K_p(\tau_d + \tau_1) \tag{4.13}$$

A proportional-plus-reset controller applied to the same process, and adjusted to produce 22.5” phase lag, can serve as a reference for comparison. The values of reset time and proportional band required for ¼-amplitude damping were calculated for selected ratios of $\tau_d/\tau_1$ in the process. The integrated error per unit load change was then found as the *PR* product, to compare with that obtainable through complementary feedback. This information is plotted in Fig. 4.16, with coordinates

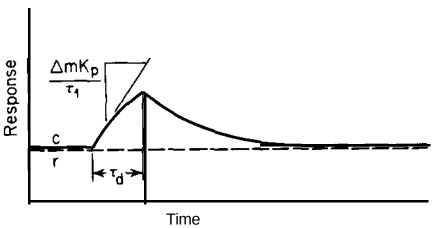



*FIG 4.15. The peak of the load-response curve occur8 $\tau_d$ min from its rise.*

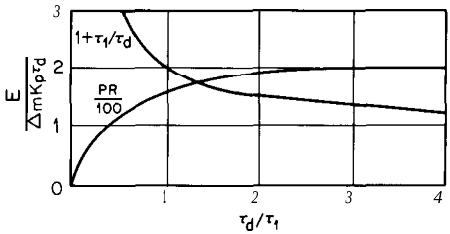



***FIG 4.16. Complementary feedback is superior to two-mode control for processes more difficult than*** $\tau_d/\tau_1 = 1.3$.

normalized as $E/\Delta m\, K_p\tau_d$ vs. $\tau_d/\tau_1$. The complementary feedback curve plots as

$$\frac{E}{\Delta m\, K_p\tau_d} = 1 + \frac{\tau_1}{\tau_d}$$

It was pointed out earlier that, as in regard to closed-loop gain, complementary feedback was superior for processes more difficult than $\tau_d/\tau_1$ = 1.2. But the comparison was not exactly on the same basis, because the proportional band was selected for ¼-amplitude as opposed to critical damping. The comparison shown in Fig. 4.16 is limited in the same way, but the agreement of the two methods is evident. The intent. has been to prove in two ways that, complementary feedback should be reserved for only the most difficult applications.

## Practical Considerations

Pure dead time cannot be generated by analog means, therefore a dead-time complementary analog controller will never be available. Dead time can be generated digitally, so the possibility exists for direct digital control systems. But in view of the problems that can be anticipated from a mismatch of process and controller dynamics, complementary feedback is of questionable value for any pure dead-time process. (A similar and more reliable method will be presented later in this chapter.)

In process-control work, capacity nearly always exceeds dead time, hence loop response generally has a moderate tolerance for controller maladjustment. This is fortunate, because to make the dynamics of the controller duplicate those of the process is not really possible.

The foregoing discussion on complementary feedback was based primarily on critically damped response. With a pure dead-time process, this was the best obtainable. But with less difficult processes, lower damping will enhance recovery from load disturbances due to greater controller gain.

In general, better performance will be obtained on more difficult applications by using delayed reset,[4] as shown in Fig. 4.17. This is obviously a compromise between two-mode control and complementary feedback,

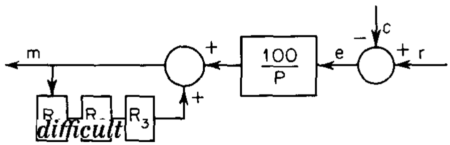



*FIG 4.17. Three lags in the reset circuit can improve performance on processes.*

but it can be made from a conventional controller. The most effective use will be made of the three lags if they are noninteracting. This can be done by selecting the capacity of $R_1$ to be 100 times that of $R_3$, and $R_2$ 10 times that of $R_3$, while setting their time constants equal by appropriate adjustment of their resistances.

The degree of improvement will vary with the difficulty of the process. On processes that, are fairly easy to control, improvement over two-mode control may be marginal. In fact, derivative would normally be of more value. But where dead time is dominant, or where derivative cannot be used because of noise level, delayed reset may be of considerable worth.

## INTERRUPTING THE CONTROL LOOP

In some control loops, feedback of information from the process is only available on an intermittent basis. The on-stream chromatograph is perhaps the most common transmitter of intermittent information, although many less familiar analyzers also have this characteristic. In a loop such as this, only one piece of information is transmitted within a certain space of time, known as the "sampling interval." The control loop is open, except at the first instant of each sampling interval. This dynamic property differs from anything discussed thus far.

Sometimes the control loop is opened intentionally. A single continuous analyzer may be used to sample two streams, being alternately a member of each loop. Recently single controllers have been shared among a number of loops, an operation called "time-shared control."

But all of these situations have one common property-a periodically open control loop. In order to provide the most effective control under these circumstances, an understanding of the influence of the sampling clement is necessary.

### Open-loop Response

The first distinction to be made is whether the sampling element is dominant or not. Figure 4.18 compares the open-loop step response of two processes which are affected differently by sampling. The abscissa is the number of samples $n$, taken at intervals of At, from the initiation of the step. The actual track of the controlled variable c is shown as a broken line, while its sampled value $c^*$ is indicated by the solid line.

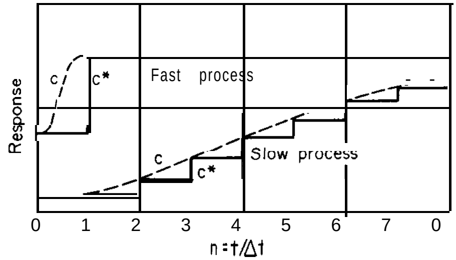



*FIG 4.18. The dynamic characteristics of the slow process are less affected by sampling.*

If a process can respond fast enough, its dynamic characteristics will be masked by the sampling system. An extremely slow process, on the other hand, would contain dynamic elements whose values far exceed the sampling interval. In the latter case, the loop is, for all practical purposes, continuous, and can be considered in that light. But as the time constants of the process approach the sampling interval, the effect of sampling increases. From this point of view, if means are found to control the loop dominated by the sampling element, less severe situations can be readily accommodated.

A flow blending process, such as the one pictured in Fig. 4.19, can be dominated by the sampling interval of a discontinuous analyzer. The residence time of the fluids in the piping is normally very short, and the delay in sample piping can be made less than the cycle time of the analyzer. Typically, fluid could be transported from the control valve to the analyzer in less than ½ min, while the analyzer might produce results at intervals of 5 min. In a process of this kind, the blend analyzed has gone downstream long before any corrective action can begin-a truly difficult control loop.

## Closing the Control Loop

The sampling element that will be discussed here is generally referred to as a "sample and hold" element, or more specifically as "sample and zero-order hold." Sampling, by itself, produces an instantaneous signal only at the start of each sampling interval. This kind of signal is not especially useful for control, unless it is "held" or memorized until the start of the next interval. A series of steps (Fig. 4.18) is thereby generated as the process changes state.

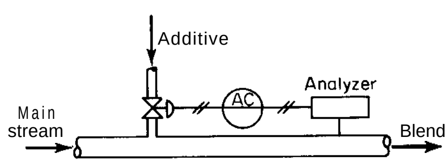



*FIG 4.19. Flow blending processes are often dominated by a discontinuous analyzer.*

Although feedback control can be exercised over a process containing a sampling element, the effect is somewhat different than what is normally experienced. Actually, the control loop is open except for an instant at the start of each sampling interval. The response of this sort of loop is not difficult to visualize, because it is nothing more than a series of open-loop responses. Under proportional control, a loop containing a dominant sampling element behaves just as if the sampling interval were pure dead time. For zero damping,

$$P = 100 \qquad \tau_o = 2\ \text{At} \tag{4.14}$$

**Integral Control**

Proportional control is obviously insufficient, as it was with pure dead time, so reset action is necessary. To facilitate identification of the influence of a sampling element, a process consisting of pure dead time and a gain of unity will be selected. The controlled variable will then follow the manipulated variable one dead time later. Figure 4.20 shows the first case, where $\tau_d = 0$.

The measurement $c_1^*$ is seen by the controller for the entire sampling interval between $n = 1$ and $n = 2$; then $c^*$ is changed to $c_2^*$. The sampled variable $c^*$ normally appears in histogram form, as shown in Fig. 4.18, but to simplify this and following figures its value at the beginning of each sample interval will be indicated by a circle. The results shown in Fig. 4.20 are identical to those obtained with complementary feedback on a dead-time process: a loop gain of 2.0 produces uniform oscillation at a period of twice the delay element, and a loop gain of 1.0 gives critical damping.

But the case of zero dead time is purely hypothetical. To be of value, any method for estimating control-loop performance cannot be so limited. Figure 4.21 gives the conditions for zero damping if the dead time of the process is one-half or one sampling interval

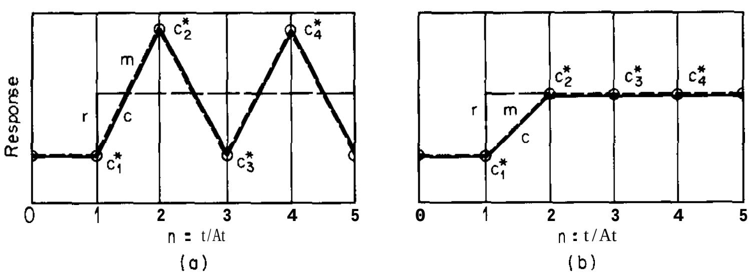



*FIG 4.20. When dead time is zero, (a) a reset time of $\Delta t/2$ produces uniform oscillations, (b) reset of* At *gives critical damping.*

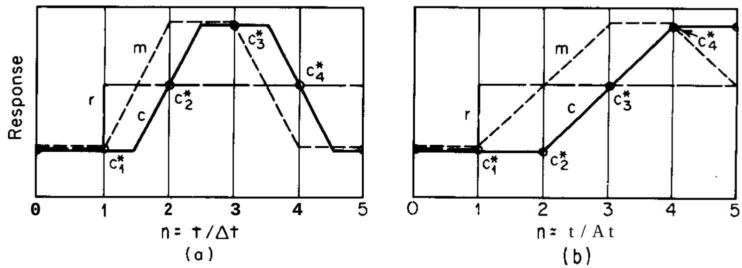



*FIG 4.21. Zero damping results when (a) $R = \tau_d = \Delta t/2$ or (b) $R = \tau_d = \Delta t$.*

The period of oscillation shown in Fig. 4.21*a* is 4 Δt, while that in Fig. 4.21*b* is 6 Δt. Add to these the period of 2 Δt when dead time is zero, and the formula for $\tau_o$ with integral control is readily derived:

$$\tau_o = 2\ \Delta t + 4\tau_d \tag{4.15}$$

The contribution of the sampling element is evident, in that the natural period of a dead-time process with integral control is $4\tau_d$ without sampling. Sampling adds 2 Δt to the period. The phase shift $\phi_\Delta$ introduced by the sampling element can then be related to the period of the loop:

$$\phi_\Delta = -\pi \frac{\Delta t}{\tau_o} = -180° \frac{\Delta t}{\tau_o} \tag{4.16}$$

The existence of any dead time whatever in the loop precludes critical damping with integral control. The value of reset time necessary for zero damping was $\Delta t/2$ for both Figs. 4.20*a* and 4.21*a*, although their periods of oscillation differed. But as oscillation becomes more sinusoidal, i.e., as more sampling intervals make up a period, for zero damping $R$ approaches $\tau_o/2\pi$, just as in a continuous loop.

The sampled wave in Fig. 4.20*a* is square, while that of 4.21*a* contains three steps, such that $R$ departs somewhat from $\tau_o/2\pi$. For ¼-amplitude damping, $R$ is to be doubled. The above estimates of reset time are based on unit process gain. They must be multiplied by the process steady-state gain $K_p$ in order to arrive at the required loop gain.

## Two-mode Control

Two-mode control combines the speed of response of proportional action with the elimination of offset brought about by automatic reset. The proportional mode is just as valuable in a sampled dead-time loop as it was in one without sampling. In fact, proportional action enables any loop whose dead time is less than the sampling interval to he critically damped. Figure 4.22 shows how this is done.

The proportional action produces an instantaneous change in output, which is removed when the error returns to zero. Some overshoot does occur, but it disappears before the next sample. Only one combination of proportional and reset will provide this critical damping:

$$P = 100K_p \frac{\Delta t}{\tau_d} \qquad R = \tau_d \tag{4.17}$$

If $\tau_d >$ At, critical damping cannot be achieved. As with complementary feedback, reducing $P$ by one-half produces zero damping, by one-fourth gives ¼-amplitude damping.

Response curves for sampled loops are made of steps. The rate of rise of even a small step is extremely high; therefore derivative control action on a sampled signal produces pulsing of the manipulated variable. This pulsing cannot contribute much to the closed-loop response, because sampling prevents the effect of such action from being seen-consequently the manipulated variable is driven severely without cause. Derivative is therefore of little value in the sampled loop.

## A Sampling Controller

Why operate on old information? There is really no value in continuing to drive the controller output when it can have no immediately observable effect. A sampling controller is suggested as being more compatible with the sampled process.

Here is the control strategy. At the start of the sample interval, when the controller sees new information, it is enabled to operate for a very short time. This will be called the control interval At,. Then the error signal is removed, preventing further integration until the next interval—the action of the controller is similar in effect to a sample and hold circuit. By means of this sampling controller, critical damping may be achieved on a sampled dead-time process with integral action alone. Figure 4.23 shows the sequence of events.

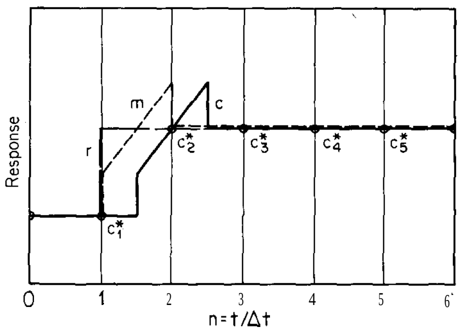



*FIG 4.22. The proper combination of proportional and reset adjustments can produce critical damping.*

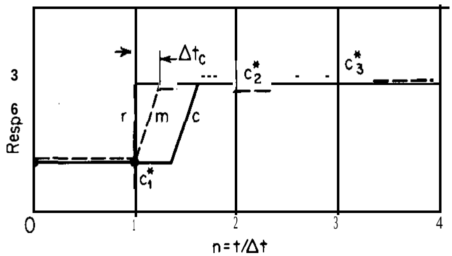



***FIG 4.23. Performance is improved by operating the controller for only a fraction of the sampling interval.***

Since the controller only operates during the short control interval, the reset time needed for critical damping is determined only by this interval and the process gain:

$$R = K_p \text{ At,} \tag{4.18}$$

Again, half this value leaves the loop undamped. Critical damping was obtained with a conventional proportional-plus-reset controller whose settings were related to both dead time and sample interval. But with the sampling controller, reset alone is required, and dead time can have any value less than At — At,, without affecting the closed loop.

Critical damping cannot be realized if the dead time is longer than At — At,. So if this situation arises, a circuit should be arranged which would activate the controller only every $k$th sample, having selected

$$k > \frac{\tau_d + \Delta t_c}{\text{At}}$$

Odd as it may seem, it is actually better, if the process consists of pure dead time, to reject information occurring in intervals less than the dead time.

A sampling integral controller is capable of critically damping a process dominated by dead time, while a continuous controller is not. This line of reasoning parallels that of complementary feedback, i.e., sampling is similar in nature to dead time whereas automatic reset is not.

When the process contains some capacity, proportional action can be used to advantage. A proportional jump in output held during the control interval serves to hasten the response of the process lag. At the

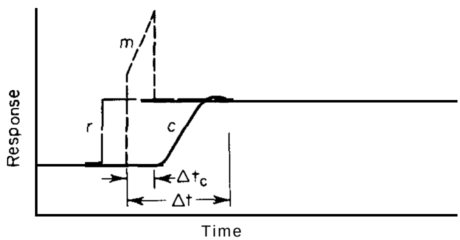



***FIG 4.24. A sampling two-mode controller can be very effective on a continuous process dominated by dead time.***

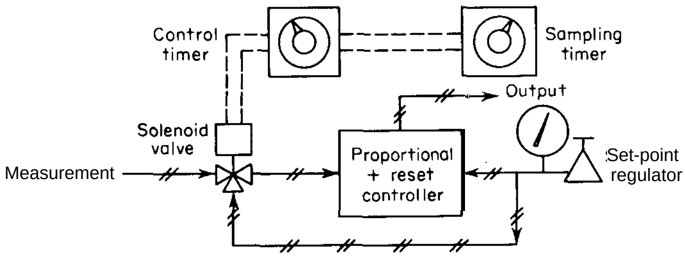



***FIG 4.25. Outside of the control interval, the set-point signal goes to both inputs of the controller.***

end of the control interval, the error is removed and the proportional component of output disappears. Figure 4.24 shows the effect a sampling two-mode controller can have on a process dominated by dead time but with some capacity.

As long as the proportional band is sufficiently wide to avoid overshoot of itself, reset time is the key adjustment. For critical damping,

$$R = \frac{PK_p \, \Delta t_c}{100} \tag{4.19}$$

The proportional band can then be gradually reduced, increasing the speed of the loop, until overshoot begins.

A conventional controller whose set, point and measurement are both accessible as electrical or pneumatic signals can be readily connected for sampling service.[5] One timer is necessary to set the control interval, and another to set the sampling interval, if no sampling analyzer is used. Figure 4.25 shows how a pneumatic controller would be connected.

### Uncertainty in Sampled Systems

In the foregoing analysis of sampled systems, discussion was centered around set-point response, with the set point being introduced an instant prior to sampling. In this way, the controller saw the error as soon as it occurred.

But the prime function of most controllers is that of load regulation. And load changes may occur at any time during the sampling interval—they are not normally synchronized to the sampler. If the full effect of a load change is manifested immediately before the controlled variable is sampled, the corrective action will be applied immediately. But if the load change occurs at any other time within the sampling interval, recovery will be delayed.

As an example, consider a pure dead-time sampled process with a sampling integral controller. The worst load response would be encountered when a step disturbance entered just after sampling. Figure 4.26 compares the best and worst situations.

The integrated error per unit load change for the best case is

$$\frac{E}{\mathrm{Am}} = K_p\left(\tau_d + \frac{\Delta t_c}{2}\right)$$

And for the worst case:

$$\frac{E}{\Delta m} = K_p\left(\tau_d + \Delta t + \frac{\Delta t_c}{2}\right)$$

To express this uncertainty,

$$K_p\left(\tau_d + \Delta t + \frac{\Delta t_c}{2}\right) > \frac{E}{\Delta m} > K_p\left(\tau_d + \frac{\Delta t_c}{2}\right) \qquad (4.20)$$

The minimum integrated error for the best case will be attained by a controller whose control interval approaches zero. The minimum error for the worst case is consistent with the minimum sampling interval. It was found that the sampling interval must exceed $\tau_d + At,$. Therefore minimum integrated error requires that

$$\mathrm{At,} \rightarrow 0 \qquad \mathrm{At} \rightarrow \tau_d$$

Then,

$$K_p\tau_d < \left(\frac{E}{\Delta m}\right)_{\min} < 2K_p\tau_d \qquad (4.21)$$

It is apparent from the last inequality that a spread of 100 percent could be encountered in the load response of sampled control loops.

The example that has been employed is a more severe test than would ever be encountered in a plant. Pure dead time is the most difficult process to control, and it is best compensated by an equal sample interval. Processes dominated by capacity are better controlled continuously; if sampled, the interval should be as short as practicable. Then the uncertainty in load response will be small because the sampling interval is small with respect to total response time.

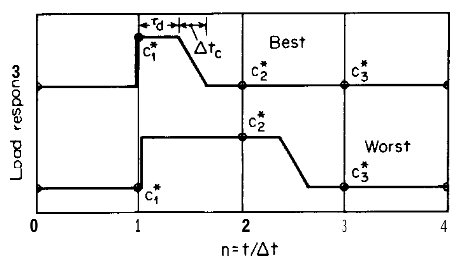



***FIG 4.26. In the worst case, control action is delayed for almost an entire sample interval.***

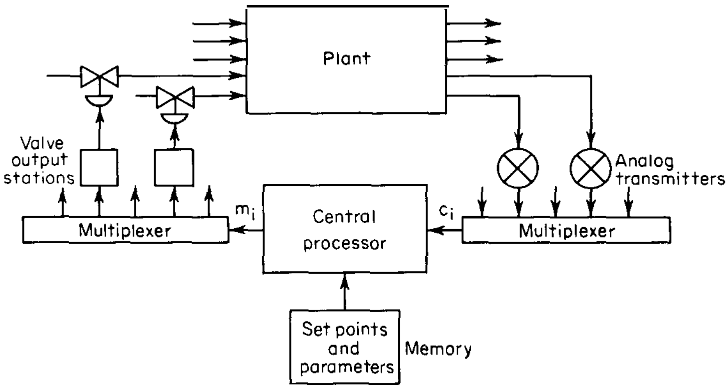



*FIG 4.27. The principal elements of a direct digital control system.*

## DIRECT DIGITAL CONTROL

Direct digital control (DDC) is the technique of controlling a number of process variables from one central digital computer, programmed with selected feedback modes. The control computer is actually shared among all the loops. An analog controller could be time-shared in the same way, although there is very little incentive to do so. The only economical way to employ a digital machine for process control is by sharing its services.

In the operation of time-sharing, each controlled variable is sampled, just as if the measurement were discontinuous. Consequently the properties of the sampled loops that have just been examined apply. Figure 4.27 is a skeleton representation of a DDC system. As each variable is scanned, its respective set point and control adjustments are selected, and the output sent to the appropriate valve.

### Driving the Valves

A hold circuit is always required on the output of a sampler, to retain the last bit of information until the next sample arrives. In a DDC system, the controller is the sampling element, therefore a hold circuit is required on each output to retain the valves in their directed position. The output signal to a valve may be held either as the shaft position of a motor or as the charge on a capacitor in an integrating amplifier. Then the computer drives the motor or changes the charge on the capacitor once each scan cycle, based on the solution of the digital-control equation. Computer output may be either a pulse train or a voltage step of variable duration, representing a change in valve position. Figure 4.28 shows

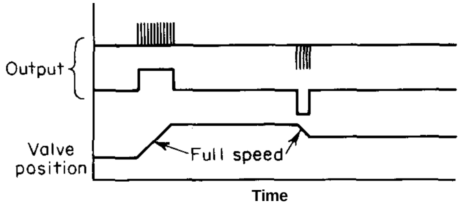



***FIG 4.28. Either a pulse train or a pulse duration signal may be used to direct the valve to its new position, at full speed.***

how the valve position would respond to the output of the computer. Note that position is the integral of computer output.

It was found that a sampling controller increased in effectiveness as its control interval approached zero. For this reason, the DDC output program should move each control valve to its directed position at full speed.

Because of the way the computer dictates a change in valve position, absence of an output signal means that valves will remain in their last position. This feature has two outstanding advantages:

1. Failure of the computer to produce an output signal will not disturb the plant.

2. Transfer from manual to automatic or automatic to manual cannot cause a sudden change in valve position, i.e., "bumpless transfer" is inherent.

### Digital-control Algorithms

Because the digital computer calculates a new value of output for a given loop only once each cycle, it cannot solve differential equations. Instead, digital-control algorithms are difference equations, whose time base is the sampling interval. The differential equation for an ideal noninteracting analog controller is

$$m = \frac{100}{P}\left(e + \frac{1}{R}\int e\,dt + D\frac{de}{dt}\right)$$

Transformation to a difference equation yields

$$m = \frac{100}{P}\left(e + \frac{1}{R}\sum_{0}^{n} e\,\Delta t + D\,\frac{e_n - e_{n-1}}{\Delta t}\right) \qquad (4.22)$$

The subscript $n$ designates the present value of the indicated variable, while $n-1$ is its value at the time of the last previous sample.

Figure 4.28 indicated how the valve is incremented in a digital system. In practice, then, the control algorithm employed does not generate valve position, but rather its increment $\Delta m$. The incremental equation is the difference between position equations from time $n-1$ to $n$.

$$\Delta m = \frac{100}{P} \left\{ e_n - e_{n-1} + \frac{\Delta t}{R} e_n + \frac{D}{\Delta t} [(e, - e_{n-1}) - (e_{n-1} \quad e_{n-2})] \right\} \quad (4.23)$$

Earlier in this chapter the disadvantage of derivative action on the set point was discussed. Although this presented something of a problem with analog devices, a digital computer can be readily programmed for derivative only of the measurement':

$$\Delta m = \frac{100}{P} \left[ e_n - e_{n-1} + \frac{\Delta t}{R} e_n + \frac{D}{\Delta t} (2c_{n-1} - c_n - c_{n-2}) \right] \quad (4.24)$$

Further simplification can be made if proportional action is applied only to the measurement, although set-point response suffers somewhat:

$$\Delta m = \frac{100}{P} \left[ c_{n-1} - c_n + \frac{\Delta t}{R} (r_n - c,) + \frac{D}{\Delta t} (2c_{n-1} \quad c_n \quad G\text{-}2) \right] \quad (4.25)$$

This algorithm is a poor choice where set'-point response is important, as in a cascade system (see Chap. 6).

Each of the coefficients above is a pure number. Therefore the entire equation could be reduced to a sum of the variables multiplied by their respective coefficients:

$$\Delta m = K_1 r_n - K_2 c_n + K_3 c_{n-1} - K_4 c_{n-2}$$

Although this is effectively the same equation, the significance of the control parameters has been lost,. So single adjustment set's the damping of the loop, as did the proportional band; none changes the phase angle in the same way as reset or derivative. In short, the experience and education of the operating personnel is frustrated in an encounter with such unfamiliar control parameters. So at the cost of complicating the arithmetic operations somewhat, the value of retaining the familiar modes is inestimable.

### Selection of the Sampling Interval

The discussion on uncertainty in sampled systems concluded that the sampling interval should not exceed the open-loop response time of the process. For best results with easy processes, the sampling interval should be as short as practicable. But where dead time dominates, the sampling interval is best- keyed to the process response time. Unfortunately, practical considerations take precedence over performance for most applications.

Take the case of a process dominated by dead time, settling out about 1 hr after a disturbance. If the sampling interval were set for best performance, the control valve would be repositioned only once an hour. And if the valve program called for a maximum increment of 25 percent per sample, 4 hr would be needed to fully stroke the valve. From the viewpoint that an important function of automatic controls is to react to an emergency condition within the plant, a 4-hr valve stroke is intolerable. For this reason, seldom will sample intervals greater than a minute be encountered.

Intervals that are too short also pose a problem. The reset component in the incremental control algorithm is $e_n$ *At/R*. If At is very small with respect to reset time, this component is subject to truncation or "rounding-off." Suppose that the word length in a given DDC computer is limited to 5 decimals. A deviation that when multiplied by *At/R* results in a product less than $5 \times 10^{-6}$ will not be acted upon. As an example, let At = 1 sec and *R* = 50 min. The minimum error causing reset action will be

$$e_n = (5 \times 10^{-6})(50)\frac{60}{1} = 0.015 \text{ or } 1.5\%;$$

In this case, a 1.5 percent offset could be expected.

Reduction of this offset by a factor of 10 can be accomplished simply by increasing the sampling interval to 10 sec. But 10 sec would be too seldom for a flow loop that would otherwise oscillate at a period of less than 10 sec. Therefore, more than one sampling interval should be used, whose selection is based on the speed of the loop in question. This is a reasonable approach, in light of the fact that a choice of at least two ranges of reset time is available in most analog controllers.

Although an optimum value of At may exist for certain difficult' processes, it is doubtful whether assigning the optimum interval to particular loops is generally warranted. Since the entire program of the digital computer is based on time usage, adjustment of $\Delta t$ is objectionable. As long as At is not more than twice the response time of a given process, performance mill usually bc satisfactory.[6] A sampling interval of 1 sec is acceptable for flow control, while somewhere between 10 and 30 sec would be suitable for most other loops.

## The Value of Derivative

Because the derivative component of output varies as the difference between two successive values of the controlled variable, it is fundamentally a measure of the average rate of change during a sampling interval. Put in other terms, it may be said to represent the rate of change of the controlled variable midway between sampling intervals. The effect is that derivative action is delayed by At. It can be seen that derivative has less and less effect as *D* approaches *At*. For this reason At should be short where derivative is used.

On the other hand, if At is too short, the controlled variable may not change enough between samples to escape truncation. Again, the choice of At will have already been dictated by other requirements-reset range, programming, etc.

Derivative limited by a lag about 1/10 of its time constant produces a maximum phase lead of about 45°. Therefore, if derivative is to be effective, *D* should exceed 10 At. In loops with a sampling interval of 10 sec, derivative time of less than 1.8 min will be limited in effectiveness. Where At = 30 sec, derivative time ought to exceed 5 min.

Some advantage is gained, however, by the absence of interaction between the digital control modes. This certainly offsets, in part at least, the other limitations encountered by the derivative mode.

### Other Control Modes

Perhaps the most outstanding feature of a digital computer is its versatility. With analog components, the system designer is constrained on many sides by physical limitations, accuracy, cost, and availability of devices. But a digital machine can be readily programmed to perform all sorts of unusual control functions: complementary feedback, nonlinear modes, logic, adaptation, feedforward-almost without limit. But at this writing, DDC is so new, in actual application at least, that few of these innovations have been thoroughly explored.

In the chapters that follow, many unconventional control schemes will be presented. They have all been tried using analog devices. Surely these and more can be implemented as well or better with a digital computer. The important point is the ease with which these things can be done once a computer is available. DDC may never pay for itself as an exact substitute for analog control. But the improvement in performance that is possible by employing novel control modes, with minimum additional expense, could easily justify the investment.

## REFERENCES


1. Clarridge, R. E.: An Improved Pneumatic Control System, Trans. *ASME*, April, 1951.
2. Giloi, W.: Optimized Feedback Control of Dead-time Plants by Complementary Feedback, *Trans. IEEE, May,* 1964.
3. Smith, 0. J. M.: A Controller to Overcome Dead Time, *ISA Journal,* February, 1959.
4. Kleiss, L. D.: Better Controller Memory Improves Control of Difficult Processes, *Control Eng.,* March, 1958.

5. Hausner, M. J.: Get the Benefits of Sampled Data Control from Modified Process Controllers, Control *Eng.*, March, 1961.
6. Bernard, J. W., and J. F. Cashen: Direct Digital Control-Questions That Must Be Answered, *Proc.* 19th *Annual ISA Conf.*, New York, October, 1964.

## PROBLEMS

4.1 Find the optimum combination of proportional and reset for a dead-time process from the information given in Table 1.1. Why is it different from the situation described for the two-mode controller in Table 4.2?

4.2 Given a process controlled with a two-mode controller whose proportional band is 200 percent and whose reset time is 10 min, estimate the maximum error developed by a step load change of 5 percent. What would the error be if the same load change were made gradually over an interval of 30 min? What would the error be if the load change entered as a sine wave of 5 percent amplitude and 2-hr period?

4.3 Referring to Table 4.1, what other settings of proportional, reset, and derivative could produce effective values identical to those in the second row of the table? What is the maximum ratio of effective derivative time to effective reset time?

4.4 Calculate $PR/100$ for values of $R = 0.90\tau_d$ and II = $0.45\tau_d$ set into an interacting controller, following the example given in Table 4.2. What conclusion can you draw from the result?

4.5 Find the optimum settings for two-mode control of a process consisting of a 30-min lag, a 2-min dead time and an analyzer with a 5-min sampling interval. Leave the proportional band in terms of $K_p$.

4.6 This chapter describes three different methods for controlling a process dominated by dead time. Select the best method, calculate the optimum values of all the parameters, and estimate the integrated area per unit load change, for a dead time of 2 min and a process gain of 2.5.

# Nonlinear Control Elements

CHAPTER 3

Elements with nonlinear properties appear both in processes and in their control systems. Up to this point an effort has been made to compensate for severe nonlinear elements naturally occurring in the process, so as to maintain a constant loop gain. But in this chapter the effects of variable loop gain will be thoroughly explored in a search to improve performance and economy.

It was pointed out that even linear controllers have nonlinear regions, i.e., beyond the proportional band. These areas are ordinarily of no consequence. But in situations where they are, methods must be available to deal with them and with other nonlinearities, similarly incidental to the prime control function.

Sometimes nonlinear devices are used to keep costs down. For example, a constant-speed motor is a less expensive final operator than a variable-speed motor. But it has only three output conditions: plus full speed, minus full speed, and stop; hence it is a nonlinear element. By the same token, a thermostat is a simpler device than a three-mode controller. It is worthwhile to examine the applications where these nonlinear devices may be satisfactorily employed to take advantage of their economy.

Perhaps the most interesting aspect of nonlinear control devices is their intentional introduction intd an otherwise linear loop in order to improve performance. Enough has been presented about linear controllers to promote an apprcciation'for the limitations to which they are subject in the regulation of difficult processes. Although nonlinear control devices are not new, only recently have methods been developed which allow their performance to be evaluated in the closed loop. Consequently their employment on difficult applications can be thought of as a new technique which has not yet been widely exploited in the process industries, principally because it is not well understood.

## NONLINEAR ELEMENTS IN THE CLOSED LOOP

Three basic forms of nonlinear elements are commonly encountered. First, there is the continuous nonlinear function, such as a pH curve or the characteristic of a control-valve plug. Second is the discontinuous function, typical of saturating types of control elements. Third is the dynamic nonlinearity, whose phase shift and gain vary with signal amplitude, as contrasted to linear dynamic elements, whose phase and gain vary with period. Devices exhibiting hysteresis are members of this category.

### Variable Loop Gain

A linear control loop is identified by its constant dynamic gain, which applies the same damping to disturbances of all magnitudes. This statement, holds true whether the loop consists entirely of linear elements or includes a nonlinear function intentionally introduced to compensate another function naturally occurring in the process.

In a nonlincnr control loop, gain varies with the amplitude of the oscillation. Whether loop gain varies directly or inversely with amplitude is a determining factor. Where gain increases with amplitude, small disturbances will be more heavily damped than large ones. Stability of the loop is then conditioned on the product of controller gain and amplitude of the disturbance. A disturbance sufficiently large to cause loop gain to exceed 1.0 will trigger regenerative oscillation.

Quit e another characteristic appears where loop gain varies inversely with amplitude. Small disturbances will be amplified and large ones attenuated such that the loop converges to uniform oscillation where its gain is 1.0. Because loop gain is 1.0 at a specific amplitude, the loop will always oscillate with that amplitude—it is called a "limit cycle."

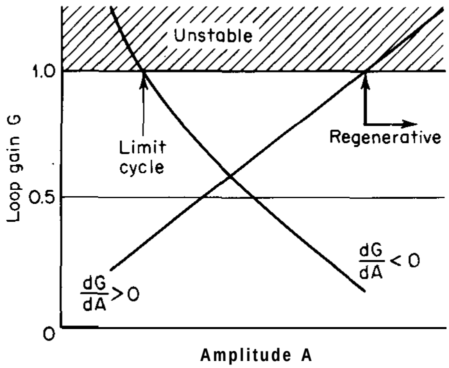



*FIG 5.1. Whether loop gain increases or decreases with amplitude is a determining factor.*

The amplitude of the limit cycle can be changed by adjusting the gain of the controller, but it cannot he damped. Figure 5.1 illustrates both variations of nonlinear loops.

**The Input-Output Graph**

A tool which is of great help in envisioning what happens in a loop containing a nonlinear element is the input-output graph. It is simply a plot of the dynamic gains of both process and controller at the period of oscillation of the loop. A signal path is then formed as the wave reflected between these two halves of the control loop. Figure 5.2 is a plot of a linear process and a linear controller adjusted for ¼-amplitude damping.

The coordinates are the manipulated ($m$) and controlled (c) variables, the respective input and output' of the ***process.*** The slope of the line representing the process is its gain at the period of oscillation. The slope of the controller line is its inverse gain at that period, opposite in effect to the process because the process output is the controller input. The lines of Fig. 5.2 indicate a high process gain and a comparatively

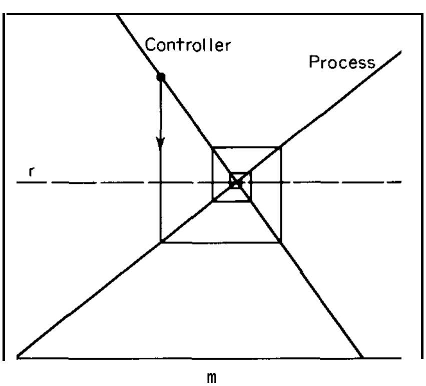



*FIG 5.2. The cycloidal path in the input-output graph is actually a projection of the sinusoidal signal variation.*

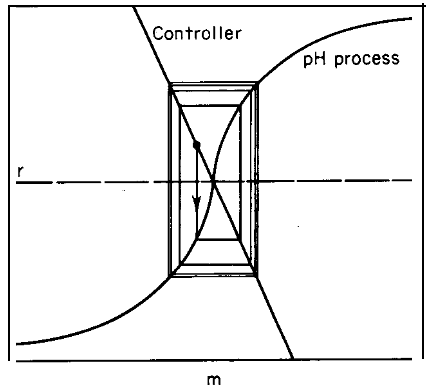



***FIG 5.3. A limit cycle develops, whose amplitude is that at which loop gain is 1.0.***

low controller gain. Where the process line crosses the set point is the load.

Initiated by a load change, the error is converted to a change in output, which in turn is reflected by the process as a new error, which is negative. The disturbance proceeds around the loop, each cycle being attenuated because the gain product of process and controller is less than 1.0. Because the gain product of the two linear elements in Fig. 5.2 is 0.5, the cycle is damped uniformly to ¼-amplitude.

Observe the graph of a pH process and a linear controller with an arbitrary gain, as displayed in Fig. 5.3.

Because a pH curve exhibits high gain to small signals and low gain to large signals, loop gain may cross 1.0. If it does, a small error will be amplified, as shown in E'ig. 5.3, until the amplitude is reached where the loop gain is 1.0; the loop will then oscillate uniformly at this point.

Notice also that a large error will be attenuated by the pH curve above the point of cycling, so that the limit cycle will be approached from without as well as from within. The limit cycle is, then, the ***normal*** condition for this control loop. Its amplitude can be changed by adjusting the gain of the controller.

A limit cycle can also be developed by the combination of a linear process and a nonlinear controller. When the proportional band of a linear controller is set too low, causing loop gain to exceed 1.0 in the linear region, the loop will eventually cycle at the limits of the controller output. A limit cycle always indicates the presence of a nonlinear element.

Whether a process can tolerate a limit cycle is up to the judgment of the engineer. In order to render a decision, two factors must be determined:

1. The ***period*** of a limit cycle is found in the same way as the natural period of a linear loop. It is the period at which the phase lags of all the elements in the loop total 180°.

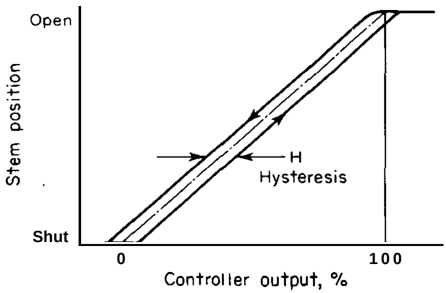



***FIG 5.4. A plot showing square-loop hysteresis in a valve.***

2. The *amplitude* of a limit cycle depends principally upon the gain of the process at its natural period. Having found this gain, an input-output graph or an amplitude-gain graph may be constructed, from which the error amplitude of the limit cycle may be measured.

## NONLINEAR DYNAMIC ELEMENTS

Principal among nonlinear dynamic elements is the hysteresis loop. In process control, the most serious form of hysteresis is encountered in control valves bothered with friction, and in on-off operators. The stem position of a control valve whose motion is opposed by friction is related to controller output in the manner described by Fig. 5.4.

The particular characteristic shown is that of square-loop hysteresis, the most severe form. Less severe loops will be somewhat rounded, but the worst case deserves prime consideration. The amount of hysteresis $H$ encountered in a valve is the change in controller output' required to reverse the direction of stem travel.

When driven by a sine wave, a valve with hysteresis produces both phase shift and distortion. The former characteristic classifies it as a dynamic element,, while the latter distinguishes it as being nonlinear. Controller output and stem position are plotted vs. time for a sinusoidal forcing function in Fig. 5.5.

If the controller output is oscillating with a peak-to-peak amplitude A, its unsteady-state component is $0.5A \sin \phi$. The controller output leads valve position in amplitude by $0.5H$. The phase angle of stem position

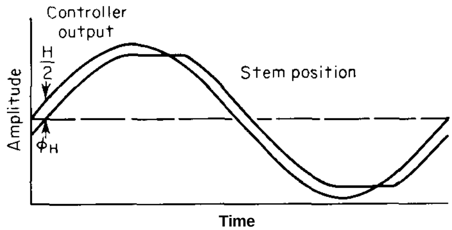



***FIG 5.5. Hysteresis causes both phase lag and distortion.***

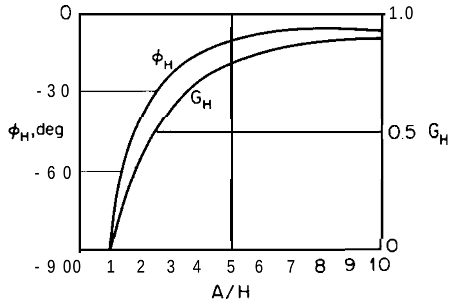



*FIG 5.6. Phase and gain both vary with the ratio of amplitude to hysteresis.*

is zero in Fig. 5.5 when the output amplitude has reached $0.5H$. The phase lag is then zero minus the phase of the output at that point:

$$0.5A \sin \phi_H = -0.5H$$

$$\phi_H = -\sin^{-1} \frac{H}{A} \tag{5.1}$$

The peak-to-peak amplitude of stem position is $H$ less than $A$. Then the gain due to hysteresis is stem amplitude over signal amplitude:

$$G_H = 1 - \frac{H}{A} \tag{5.2}$$

Phase and gain are plotted vs. dimensionless amplitude $A/H$ in Fig. 5.6.

Because gain and phase both change with the amplitude of the controller-output oscillation, gain and period of the closed loop are both variable in the presence of hysteresis. A condition like this is cause for concern, in that stability is also variable. Loop gain should be checked for selected values of amplitude to determine whether stability is conditional.

**example 5.1**

Consider a process consisting of dead time, an integrating capacity, and hysteresis being controlled proportionally. From Fig. 5.6, a high $A/H$ ratio results in gain approaching unity and phase lag approaching zero, just as if there were no hysteresis present. Therefore, proportional band can be set to produce ¼-amplitude damping (i.e., loop gain of 0.5) without hysteresis:

$$P = 400 \frac{\tau_d}{\pi \tau_1}$$

For the loop to be stable, its gain must be less than 1.0 for all finite values of amplitude. Table 5.1 summarizes the gain and phase of the various elements in the loop for selected values of amplitude. Because loop gain is less than 1.0 for all finite values of amplitude, the loop is stable.

**TABLE 5.1 Loop Gain vs. Amplitude for Proportional Control**

| $A/H$ | $G_H$ | $\phi_H$, deg | $-\phi_d$, deg | $\tau_o/\tau_d$ | $G_1$ | $100G_1G_H/P$ |
|---|---|---|---|---|---|---|
| ∞ 10 | [illegible] | 6.5 0 | 90 83.5 | 4.3 4.0 | $0.64\tau_d/\tau_1$ | 0.5 0.48 |
| 5 | 0.8 | 11 | 79 | 4.55 | 0.71 | 0.45 |
| 2 | 0.5 | 30 | 60 | 6.0 | 0.95 | 0.38 |

This example was chosen as typical of some closed loops containing hysteresis. Although the change in period with amplitude causes the gain of the integrating element to vary, the effect is more than offset by the change in gain of the hysteresis element. Consequently the loop is stable as long as the band has been set for high-amplitude conditions. If adjusted for a given damping with a disturbance of low amplitude, however, the loop could become undamped in the face of a severe upset. But the small variation in loop gain (0.75 to 1.0) corresponding to a wide amplitude range (2 to ∞) makes this event unlikely.

Loops containing two integrations are capable of a limit cycle, however. An example would be a non-self-regulating process such as liquid level, with a proportional-plus-reset controller. The gain product of the two integrating elements will vary as the square of the period, more than can be offset by the gain of hysteresis. Under these conditions, loop gain varies inversely with amplitude.

**example 5.2**

Consider a loop consisting of a dead-time plus integrating process of time constant $\tau_1$, hysteresis, and a proportional-plus-reset controller. Let the reset time be set for 30” phase lag and the proportional band for ¼-amplitude damping at *A/H* of 2. Table 5.2 summarizes the effect of hysteresis.

**TABLE 5.2 Loop Gain vs. Amplitude for Two-mode Control**

| $A/H$ | $G_H$ | $-\phi_H$, deg | $-\phi_d$, deg | $-\phi_R$ | $\tau_o/\tau_d$ | $G_1$ | $G_{PR}G_1G_H$ |
|---|---|---|---|---|---|---|---|
| 4.0 | 0.75 | 14.5 | 60.0 | 15.5 | 6.0 | $0.95\tau_d/\tau_1$ | 0.34 |
| 2.0 | 0.50 | 30.0 | 30.0 | 30.0 | 12.0 | 1.91 | 0.50 |
| 1.8 | 0.45 | 33.7 | 14.8 | 41.5 | 24.4 | 3.9 | 1.00 |

With the existing settings, the control loop limit-cycles at an amplitude of $1.8H$ and a period of $24.4\tau_d$. Reduction of the proportional band by half will change the amplitude only to $2H$ but will reduce the period to $12\tau_d$.

Minimum period may be attained by reducing the proportional band to the point, where the amplitude becomes objectionable. Increasing the reset time reduces the amplitude while increasing the period of the cycle. Therefore a long reset time and a fairly narrow band will give the best combination of amplitude and period for this loop.

A limit cycle caused by hysteresis will appear as a nearly square wave on a flow record--this is its most distinguishing characteristic. And it 'will not bc possible to damp the wave by adjustment of the proportional band; this is typical of loops containing a nonlinear element.

Hysteresis can be minimized by superimposing a high-amplitude, high-frequency signal on the controller output. The process would not respond to it, but the sticking element would. There is sufficient noise in some measurements to create this effect. But the best solution is to close the loop around the hysteresis element, alone, with a proportional controller-this amounts to installing a positioner on the valve. More will be said about positioners in Chap. 6.

## VARIATIONS OF THE ON-OFF CONTROLLER

A pure on-off control function produces an output of either 0 percent (off) or 100 percent, (on), according to the sign of the error. Because a sinusoidal error of any amplitude will produce a square wave of unit amplitude, the device is said to have a variable gain. In order to establish a figure of gain for any device, its output should be expressed in the same terms as its input. So instead of a 100 percent square wave, the output will be thought of as the fundamental sinusoidal component of the square wave. This component has an amplitude of $4/\pi$ times that of the square wave. Then the gain of the on-off controller, G,, becomes the output amplitude $400/\pi$ over the input amplitude A, expressed in percent:

$$G_o = \frac{400}{\pi A} \tag{5.3}$$

An on-off controller employed on any process capable of shifting phase beyond 180" will cause a limit cycle. The loop gain $G_oG_p$ under these conditions is 1.0, $G_p$ representing the process gain at its natural period. Therefore the peak-to-peal; amplitude of the limit cycle, expressed in percent, is

$$A = 400 \frac{G_p}{\pi} \tag{5.4}$$

Figure 5.7 shows the trajectory of a limit cycle under on-off control. Because the load is low, the limit cycle is not centered about the set point.

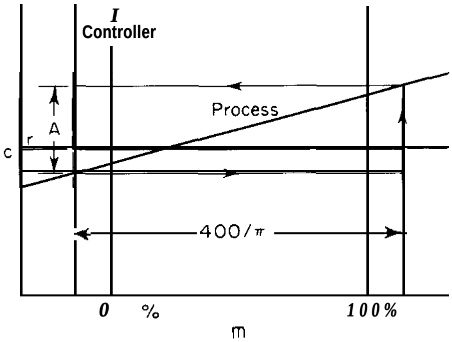



*FIG 5.7. An input-output graph for on-off control.*

**Differential Gap**

There is no such thing as a perfect on-off device, and if there were, it would be useless. If it could change states instantaneously with an infinitesimal input, the presence of any noise whatever would cause constant chattering. Therefore real on-off controllers are all limited in their action by one of the following mechanisms:

1. Gain sufficiently low to require a significant' input to cause saturation, as in a controller with a very narrow (but not zero) proportional band.

2. Positive feedback, causing the device to lock in either output condition. This action is called "differential gap," and is characteristic of switching devices.

Differential gap is hysteresis in an on-off controller. A signal must change by an amount equal to the gap *H* in order to change the state of the output. The input-output characteristic of an on-off controller with differential gap appears in Fig. 5.8.

The phase characteristics of differential gap are the same as those of hysteresis. But because of the on-off action, gain differs. An input signal of amplitude lower than the width of the gap produces no output; otherwise, gain is that of an ideal on-off controller.

Any process whose dynamic elements are capable of shifting phase in excess of − 180" will limit-cycle under on-off control. Presence of differ-

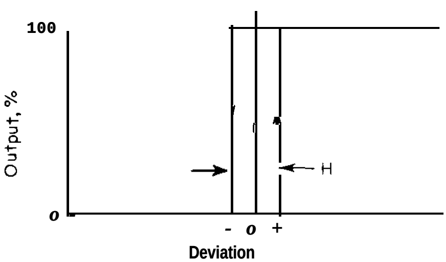



*FIG 5.8. The state of the output will not be changed until the deviation exceeds the width of the gap.*

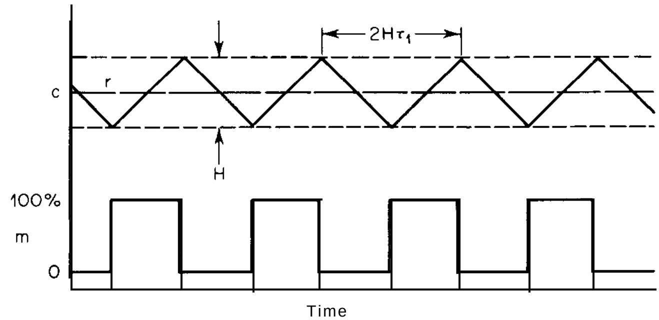



*FIG 5.9. In the presence of differential gap, even a single-capacity process will limit-cycle.*

ential gap will increase the amplitude of the cycle by approximately the width of the gap and will lengthen its period proportionately because of the additional phase lag. For an exact prediction, the phase and gain formulas given in Eqs. (5.1) and (5.3) should be applied. But because an on-off controller forces a process in a full-scale square wave, its resulting error signal in many cases will be far from sinusoidal. Under these conditions the equations given for sinusoidal response are inexact. A case in point is the control of a single-capacity integrating process by a device with differential gap. The resulting limit cycle is triangular, as shown in Fig. 5.9. With only 90" phase lag in the process, the full 90" lag in the controller is consumed in a limit cycle of $A = H$.

Because the gain of this function to signals smaller than a gap width is zero, it also acts as a nonlinear filter, rejecting noise of low amplitude.

## Proportional Time Control

In systems where on-off control produces a limit cycle that is both too long and too high, certain modifications may be applied. Proportional time control is a technique by which the on-off output is modulated with a signal of fixed period but variable "on" time. The percentage of each period during which the controller output is maximum is proportional to deviation. Thus the average value of output is the same as it would be with a proportional controller. Figure 5.10 shows the relationship between deviation and controller output.

Because this type of controller naturally oscillates, the loop is forced to limit-cycle at the period of modulation, $\tau_m$. This period should be selected so that, the gain of the process is low enough for its limit cycle to be of negligible amplitude. Having this assurance, the loop approaches proportional control with a linear final operator. The criteria for adjust-

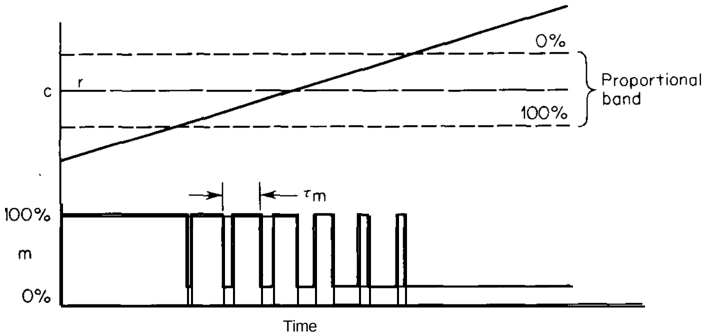



***FIG 5.10. Percentage "on" time varies as the measurement passes through the proportional band.***

ing the proportional band are essentially the same as with a linear loop, and offset is encountered for the same reason.

### A Constant-speed Motor as the Final Operator

Up to this point only two-state on-off control has been presented, where the final operator would be a solenoid valve or an electrical heating element. These are either on or off. But very often a constant-speed reversible motor is used to drive a valve or to position a lever. This type of operator has three states: drive upward (opening), stop, and drive downward (closing). So the controller must be similarly arranged. The simplest controller for this function consists of two on-off devices whose inflection points are separated by a dead zone. Within this zone, the motor would be stopped. In practice, each on-off device also contains a small differential gap. The input-output relation is pictured in Fig. 5.11.

A finite dead zone must exist, first to ensure that the on-off switches do not overlap, for this would energize both windings of the motor simultaneously and could result in damage. But beyond this, it provides a

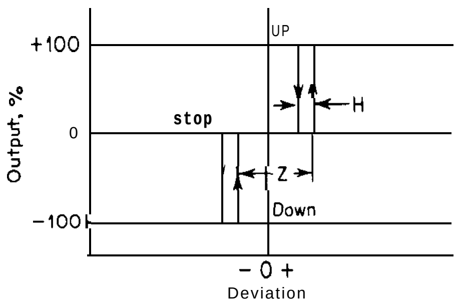



***FIG 5.11. The three-state controller is comprised of two on-off devices separated by a dead zone Z.***

state of rest for the loop that did not exist with previous on-off configurations. Thus the control system is not bound to limit-cycle. The dead zone $Z$ is defined as the change in deviation between deenergizing one output and energizing the other.

To avoid a limit cycle, the dead zone must be wide enough so that energizing one output will not change the controlled variable enough to energize the other. The amplitude of whatever limit cycle might be generated would be $400G_p/\pi$, where $G_p$ is the process gain at the natural period including the phase lag of the gap and the motor. Therefore, if limit cycling is to be avoided,

$$Z > \frac{400\ G_p}{a} \tag{5.5}$$

Along with the stability given by the dead zone comes offset. Instead of a limit cycle of amplitude $400G_p/\pi$, an equal offset is encountered. At least the limit cycle had an integrated error of zero; but offset integrates to infinity.

The limit cycle with a constant-speed motor will always be centered around the set point, because the motor is an integrator. So the average value of the error will always approach zero-this was not so with two-state on-off control. But the presence of an integrating element (the motor) in the loop doubles its natural period.

### Adding Other Control Modes

If limit cycling and offset are both unacceptable, a two-mode controller can be added to the loop. This controller would actuate the three-state device which in turn drives the valve, as shown in Fig. 5.12. Reset action will keep driving the output of the controller out of the dead zone until the error is reduced to zero. Only then will the loop reach a steady state. Proportional action is necessary for stability, for without it, the double integration of reset and motor would cause an undamped cycle. The availability of a proportional band adjustment eliminates the need for an adjustable dead zone, since the two effects are similar.

Derivative may also be added to advantage. It contributes about as much to the loop as it does to a linear loop by reducing the period of oscillation and allowing a narrow proportional band. Derivative cancels, to some extent, the sluggishness of the constant-speed motor.

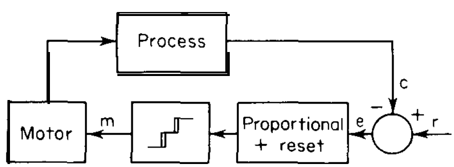



*FIG 5.12. Adding a two-mode controller can eliminate offset.*

It might be mentioned at this point that proportional time control of a constant-speed motor is a practical and inexpensive system. Two outputs must again be provided, with a dead zone between. Both outputs are modulated with time proportioning, one for a positive error, the other for a negative. Offset is encountered in the two-state loop, but the motor now overcomes it. Proportional time control with a constant-speed motor is essentially integral control, with a reset time of

$$R = \frac{P}{200} \tau_v \tag{5.6}$$

The time constant of the valve motor, $\tau_v$, is the time required to drive it full stroke.

## THE DUAL-MODE CONCEPT

What would you like a controller to do, forgetting for the moment what it is reasonable to expect? The most severe demand would certainly be to follow a step change in set point perfectly. This could be demanded of the controller, but not of the process, because it requires infinite process gain. The speed at which a variable may change is limited by the maximum rate at which energy may be delivered to the process. A valve may only open fully, not infinitely. Therefore it can only be asked that the controller not interfere with the maximum speed of the process. To duplicate the remainder of the step input, the control loop must be stable to the point that no overshoot or oscillation is observable. Nor should there be any offset. Finally, the controller ought to be insensitive to input noise, which is usually present in some form. To summarize, the ultimate controller should be capable of achieving the following loop-response characteristics:

1. Maximum speed
2. Critical damping
3. No offset
4. Insensitivity to noise

Any control system that can satisfy the above demands will also satisfy any minimum integral error criterion, regardless of what function of the error may be used, and regardless also of the nature of the input signals. The character of the process determines the complexity of the controller which is to accomplish the goals listed above. If the process is a pure single capacity, an on-off controller will provide maximum speed, critical damping, and no offset. An on-off controller is sensitive to noise, however. Significantly, this simplest control device is capable of achieving the ideal closed-loop response on the simplest process. As the process complexity increases, on-off control is no longer optimum, and combinations of less severe linear or nonlinear elements must be used to provide stability. Obviously the nature of the process determines the design of the controller which will elicit the best loop performance.

With regard to difficult processes, control functions which approach the demands of the four points of performance listed above need to be set forth:

1. Maximum speed implies that the controller be saturated for any measurable deviation. Such a demand limits the selection to an on-off controller. But if the tolerance may be widened somewhat, then the controller need only saturate in response to a large signal. (How large the signal must be will vary with the difficulty of the process.)

2. Critical damping can be achieved by both low gain and derivative action, but the latter amplifies noise. Critical damping implies an asymptotic approach to the set point. To accommodate maximum speed, the zone of critical damping must be restricted to a narrow band about the set point. Therefore this criterion and its solution apply specifically to small-signal response.

3. Zero offset requires a controller with infinite gain, in the steady state. An integrator supplying "reset" action is sufficient to satisfy this criterion.

4. Low noise response can be obtained through low gain or low-pass filtering, but low-pass filtering degrades the speed of response of the loop. The only condition, then, which tends to reconcile this requirement with the others, is the application of low gain to small signals.

Of particular significance is the combination of high gain to large signals and low gain to small signals. The exact combination of parameters that will be most effective for a specific application may not be obtainable in a single controller. It may then be necessary to use two controllers, intelligently programmed to the best advantage of their individual features. The combination of two controllers operating sequentially in the same loop has been called a dual-mode system.

**Selecting the Two Controllers**

The only place where stability is assured with the nonlinear devices presented thus far is within a dead zone. But within the dead zone there is no control whatever, which results in offset. Therefore the region about the set point should not have zero gain, but its gain should be low as compared to on-off control.

A sensible approach is to use a linear controller in this region, with gain adjustable by the proportional band. Because the gain is expected to be low, reset action is also required. Derivative is also recognized as valuable in promoting damping and rapid response to load change. Consequently the most logical controller to use in the small-signal region, for most applications, is a conventional three-mode controller.

On-off action provides the highest available gain for large-signal response. So the choice of on-off to operate sequentially with the linear controller is obvious. But the transition from one to another is not obvious. And the boundary between what is considered a small signal and what is considered large, is not at the moment clear.

It would seem that a nonlinear control mode would not be very effective in combating load changes. To begin with, only a relatively severe load change would ordinarily cause an error large enough to enter the large-signal zone. But all load changes are acted upon by the linear controller, because the errors which they induce all originate and subside in the small-signal zone.

The same is not true of set-point changes. A set-point change may easily be introduced fast enough to pass directly into the large-signal zone. The conclusions drawn earlier concerning desirable controller characteristics were based on particular demands of set-point response. For these reasons, it may be ventured that nonlinear controllers may enhance set-point response, while linear controllers are more effective against changing load. This is further evidence in support of the selection of a linear controller for small errors and a nonlinear for large.

### Optimal Switching

The performance that a control system is capable of giving is relative to the amount of intelligence built into it. It is possible to design a programmed system that will out-perform any conventional feedback system in response to command (i.e., set-point) inputs. But the program must embody more of the characteristics of the process than do conventional controllers. It has been found necessary to use an "antiwindup" feature in a controller with reset in order to prevent overshoot on large set-point changes. It has been further noted that the controlled variable will respond best if the reset mode is preloaded to the steady-state conditions that are expected to prevail. In a sense, this is programmed control.

An optimum program is one which will place the controlled variable at the designated set point in the minimum time (or with minimum energy or minimum cost).

On-off control mill drive the process at maximum speed, but will cause overshoot with more than a single capacity in the loop. Overshoot means that some oscillation, however damped, is present, requiring a specified settling time after the set point is crossed. This is scarcely consistent with minimum time.

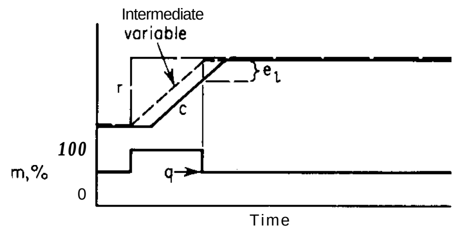



***FIG 5.13. Minimum-time control of a dead-time plus integrating process.***

To explore what is required for minimum-time control, consider the application to a dead-time plus integrating process. In Fig. 5.13 the tracks of both the intermediate variable, i.e., the output of the integrator, and the controlled variable are plotted. Minimum time requires that $m$ be switched from 100 percent to equal the load $q$ before the set point is reached.

The optimum control program must include the expected load at the designated set point, for two reasons:

1. If the manipulated variable fails to match the load after the new set point is reached, the process will be in an unsteady state.

2. The rate of rise to the new set point, hence the error $e_l$ at the time of switching, is load-dependent.

It may be recalled from Chap. 1 that the rate of rise of the intermediate variable, and also of the controlled variable, is

$$\frac{dc}{dt} = \frac{m - q}{\tau_1}$$

where $\tau_1$ is the time constant of the integrating element. The controlled variable is delayed by the dead time $\tau_d$ behind the intermediate variable. It therefore lags in magnitude by

$$\tau_d \frac{dc}{dt} = \frac{\tau_d}{\tau_1}(m - q)$$

Overshoot will be prevented if $m$ is switched from 100 percent to $q$ when the intermediate variable reaches the set point. The deviation $e_l$ at that time is

$$e_l = \frac{\tau_d}{\tau_1}(100\% - q) \tag{5.7}$$

The parameters $e_l$ and $q$ must be manually set into the control system.

The control loop will not oscillate, for if the load is well matched, the switching point will only be crossed once. The loop is actually open from then on. As described above, the system will only respond to increasing set points. With an optimum program for decreasing set

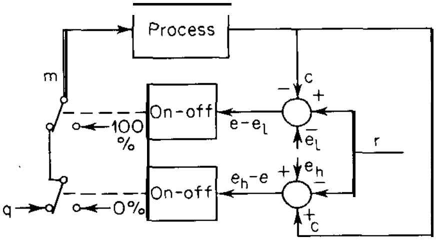



***FIG 5.14. Bidirectional programming requires two on-off controllers with separately adjustable switching points.***

points, the on-off controller will switch between 0 percent and $q$ at the point $e_h$:

$$e_h = \frac{\tau_d}{\tau_1} q \tag{5.8}$$

A control system designed for bidirectional optimal switching requires two on-off operators, as depicted in Fig. 5.14. The distance between the two switching points, expressed in percent,, is the dead zone $Z$, analogous to the proportional band of a linear controller:

$$Z = e_l + e_h = 100 \frac{\tau_d}{\tau_1} \tag{5.9}$$

If the same program is applied to a two-capacity process, the controlled variable will be more heavily damped than necessary. Therefore this program provides the minimum-time switching only for dead-time plus integrating processes.

Referring back to Fig. 1.21, notice that the overshoot for a two-capacity process under on-off control was less than with dead time because of the reduced gain of the second capacity at that period. It is possible to take advantage of that gain reduction to save time. Figure 5.15 com-

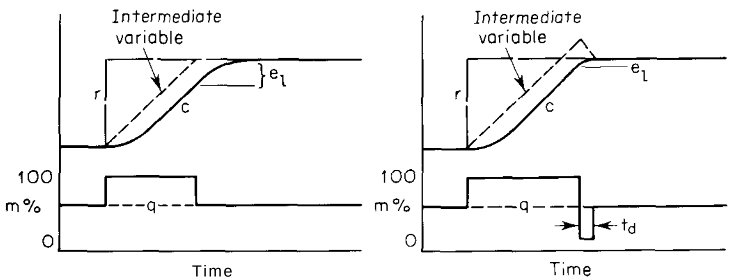



***FIG 5.15. The minimum-time program (right) actually causes the intermediate variable to overshoot.***

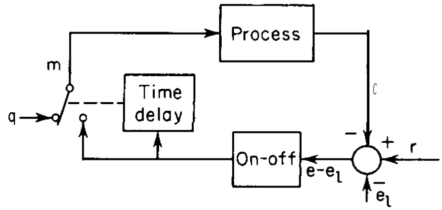



***FIG 5.16. The delay timer is actuated by the on-off controller.***

pares the response of the previous program to that of the optimum program for a two-capacity process.

The program consists of switching from 100 percent output at $e_l$ to 0 percent output for a specified time $t_d$, after which the steady-state value is selected. Values of $e_l$ and $t_d$ necessary for minimum-time control can be found by solving the following pair of equations:

$$t_d = -\tau_2 \ln \frac{q}{100} \qquad e_l = \frac{100\tau_2 - q(\tau_2 + t_d)}{\tau_1} \tag{5.10}$$

For the particular case where $q = 50$ percent,,

$$t_d = 0.693\tau_2 \qquad e_l = \frac{15.3572}{\tau_1}$$

If the program designed for a dead-time plus integrating process were used, $e_l$ would have been $50\tau_2/\tau_1$. With the minimum-time program, 100 percent output is retained $0.693\tau_2$ longer, which corresponds to the time required to dissipate that additional energy.

In order to accommodate this more complicated program, a delay timer must be added to the system. The arrangement of the loop for an increasing set-point change is shown in Fig. 5.16.

The on-off controller deenergizes when $e_l$ is reached, sending 0 percent output to the process while the timer is operating.

Again, processes do not fall into such neat classifications as two-capacity or single-capacity plus dead-time. The bulk of difficult processes lie between these limits. But the same control function described by Fig. 5.16 and Eq. (5.10) can be adjusted to accommodate dead time in addition to two capacities. Equation (5.11) indicates the required settings for optimal switching:

$$t_d = -\tau_2 \ln \frac{q}{100} \qquad e_l = \frac{100(\tau_2 + \tau_d) - q(\tau_2 + \tau_d + t_d)}{\tau_1} \tag{5.11}$$

This is one control function whose exact settings can be determined numerically for a process with three dynamic elements. But however difficult the process, settings for the switching parameters can be found which will provide absolute optimum set-point response using this system.

### Adding a Linear Controller

Optimal switching programs were developed originally for positioning systems and vehicle control. The final state of these processes is generally quiescent, i.e., zero velocity, where no control is required over the steady state. But in fluid processes, control is needed to provide mass and energy balance in the steady state. As a result, programmed control, whose final state is open-loop, is incomplete.

The loop may be closed simply by adding a linear feedback controller to operate in conjunction with the programmed mode. When the programmed action is completed following a set-point change, the linear controller is switched into the loop. In effect, two controllers comprise the system, one for the steady state, one for the unsteady state.

The output of the control system should match the load, just as was done without the linear controller. If it does not, an error will develop after transfer is made to the linear controller. The addition of a linear controller should not be construed as license to discount the settings of the program. Instead, the program should be designed just as if there were no linear controller. Every effort should be made to place the controlled variable exactly on the set point with zero velocity when transfer is made. In this way, the controller will have no work to do, and hence will not disturb the process. To be sure, the linear controller will compensate as well as it can for inaccuracies in the program, and this is beneficial.

The linear controller must, above all, be preloaded to the anticipated process conditions at the new set point. This was found to be advantageous whenever an "antiwindup" switch is used in the reset circuit. And because reset is normally required with a dual-mode arrangement, the control system must include an antiwindup switch into which the preload setting is introduced. The arrangement of the system for increasing set-point changes is shown in Fig. 5.17.

The sequence of events is as follows:

1. While $e > e_l$, the on-off operator is energized, sending 100 percent output to the process. The preload setting $q$ is sent to the proportional-plus-reset-plus-derivative controller.

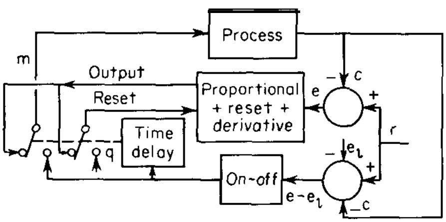



*FIG 5.17. A dual-mode system for increasing set point.*

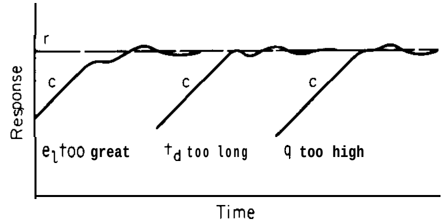



***FIG 5.18. Maladjustments in the program parameters are easy to diagnose.***

2. When e = $e_l$, the on-off output' drops to 0 percent and the time delay begins. The on-off operator remains in control and preload is sustained.

3. At the end of the delay period, transfer is made to the proportional-plus-reset-plus-derivative controller and preload is replaced with controller output, starting reset action. By this time, the error and its derivative should both be zero, so the controller output will equal the preload setting. Transfer is therefore "bumpless."

The dual-mode system gives the best set-point response attainable. Optimal switching, by definition, is unmatched in the unsteady state, while the linear controller provides the regulation necessary in the steady state. But any control system is only as good as the intelligence with which it is supplied. In the event of maladjustments in the three parametcrs $e_l$, $q$, and $t_d$, the track of the controlled variable will be imperfect. The value of $e_l$ will vary directly with the difficulty of the process. As the process difficulty decreases, the controlled variable is less a function of load, and hence has more tolerance for inaccuracies in the control parameters. But the degree of performance improvement provided by dual-mode control also varies directly with process difficulty.

The dual-mode system needs six adjustments, which fall into two independent groups. Settings of proportional, reset, and derivative only pertain to the steady state, while the program settings are in effect elsewhere. Consequently, adjusting the dual-mode system is no more difficult than adjusting two separate controllers. Rules for setting the program parameters are self-evident :

1. Maladjustment of $e_l$ causes overshoot or excessive damping.

2. Excess time delay turns the controlled variable downward after the set point is reached.

3. An incorrect preload setting introduces a bump after the time-delay interval.

The effect's of these maladjustments are graphically demonstrated in Fig. 5.18.

Recall the specifications which were set, forth at the beginning of the section on dual-mode control. Maximum speed has been provided by the on-off controller. The programmed switching critically damps the loop as the set point is approached. Offset is eliminated by reset in the linear controller. Finally, noise of magnitude less than $e_l$ will not actuate the on-off operator and therefore will be no more of a problem than in a linear system. Although complicated and costly, dual-mode control cannot be matched for performance.

## NONLINEAR TWO-MODE CONTROLLERS

It has been demonstrated that a loop whose gain varies inversely with amplitude is prone to limit-cycle. Any controller with similar characteristics can promote limit cycling in an otherwise linear loop. On-off controllers are in this category. So any nonlinear device that is purposely inserted into a loop for the sake of engendering stability must have the opposite characteristic: gain increasing with amplitude. The only stabilizing nonlinear devices discussed up to this point have this property -it was manifested as a dead zone in the three-state controller and as the linear mode in the dual-mode system.

It is not difficult to visualize a desirable combination of properties for a general-purpose nonlinear controller. In fact, the characteristics outlined for a dual-mode system apply: the controller should have high gain to large signals, low gain to small signals, and reset action. The variation of gain with error amplitude can be accomplished continuously or piecemeal.

### A Continuous Nonlinear Controller

It is possible to create a controller with a continuous nonlinear function whose gain increases with amplitude. In contrast to the three-state controller, its gain in the region of zero error would be greater than zero, with integrating action to avoid offset. But its change in gain with amplitude should be less severe than that of a dual-mode system. Thus it would be more tolerant of inaccuracy in the control parameters.

The continuous nonlinear controller could be mathematically described by the expression

$$m = \frac{100}{P} f|e| \left( e + \frac{1}{R} \int e \, dt \right) \tag{5.12}$$

In this way its gain varies with the absolute magnitude of the error. A suitable linear function can be used.

$$f|e| = \beta + \frac{(1 - \beta)|e|}{100} \tag{5.13}$$

where $\beta$ is an adjustable parameter representing linearity and e is expressed in percent. If $\beta = 1.0$, the controller is linear. But as $\beta$ approaches zero, the control function becomes square law, taking the

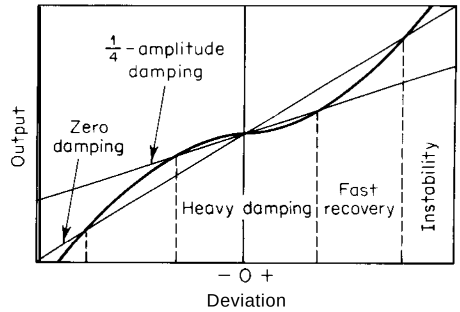



***FIG 5 19. The proportional characteristic of a continuous nonlinear controller displays variable damping.***

shape of the parabolic sections shown in Fig. 5.19. It is not desirable for $\beta$ to equal zero, since this would render the controller essentially insensitive to small signals, and offset would result. A value of $\beta$ in the vicinity of 0.1 would make the minimum gain of the controller $10/P$.

A characteristic of this sort produces varying degrees of damping in the closed loop. If a linear controller were used to regulate a given linear process, a certain proportional gain could be found which would produce uniform oscillations. A straight line representing this gain, labeled "zero damping," is superimposed on the curve in Fig. 5.19. If the proportional gain of the linear controller were halved, the closed loop would exhibit $\frac{1}{4}$-amplitude damping. The controller gain representing $\frac{1}{4}$-amplitude damping is also indicated.

The nonlinear characteristic crosses both these contours of constant damping. Between the intersections are three distinct stability regions. In the region surrounding zero deviation, damping heavier than $\frac{1}{4}$-amplitude persists, while adjacent to it on both sides are regions of lighter damping and consequently faster recovery. There is still another region on each side where damping is less than zero-representing instability. Should a deviation arise large enough to fall into this last area, it will be amplified with each succeeding cycle.

To gain a better insight into the response of this nonlinear characteristic in a loop with a linear process, the input-output graph of Fig. 5.20 has been constructed. Notice how heavily a small signal is damped. Damped oscillations in a linear loop theoretically go on forever. But with a nonlinear characteristic of the kind shown, damped oscillations cannot persist beyond one or two cycles. On the other hand, a large signal causes more corrective action than a linear controller, appropriately damped, could provide. A sufficiently larger deviation could promote instability, however, so the proportional band of the nonlinear controller must be adjusted for the largest anticipated deviation.

As with other nonlinear controllers, set-point response exceeds what is obtainable with linear modes. This is because set-point changes are

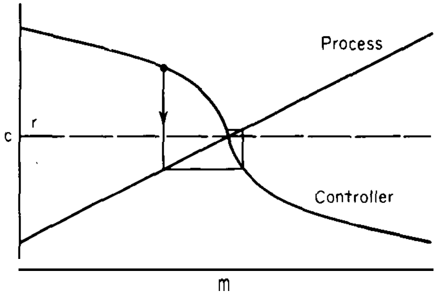



***FIG 5.20.*** ***If the initial deviation is not extreme, it may be damped within one cycle.***

normally greater and more rapid than load disturbances, taking advantage of the region of higher gain. Load disturbances make their appearance as a slow departure of the controlled variable from the set point. Since a linear controller has more gain in the region close about the set point, it will generally respond more effectively to small load changes. A comparison of the responses of linear and nonlinear three-mode controllers is shown in Fig. 5.21.

A nonlinear two-mode controller seems generally to outperform a linear two-mode controller. The nonlinear function provides an extra margin of stability similar to what can be attained with derivative. In cases where so much noise is superimposed on the measurement that derivative cannot be used, a nonlinear function can be quite valuable.

Another feature of the nonlinear controller is its extreme tolerance of gain changes in the loop. Response to upsets of moderate magnitude appear virtually identical over a proportional band range of 4 : 1 or more. Consequently little care need be given to the settings of proportional and reset, save for the possibility of bringing the unstable region too close to the set point.

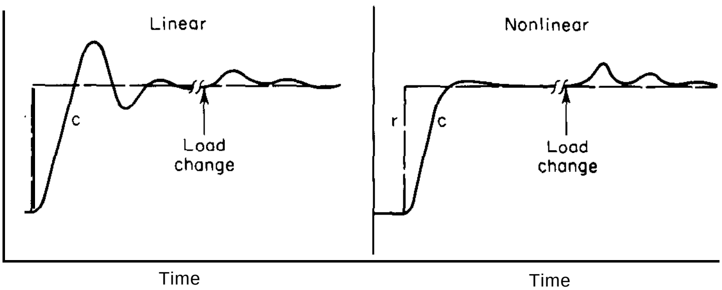



***FIG 5.21.*** ***A three-mode nonlinear controller exhibits better set-point response but poorer load response than its linear counterpart.***

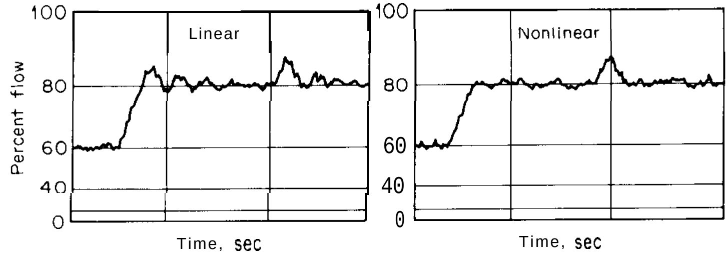



***FIG 5.22. The nonlinear two-mode controller is superior in all respects on a noisy flow loop.***

### Flow Control

A flow measurement is always accompanied by noise. This noise is attenuated somcwhnt by the wide proportional band of the controller and passed on to the valve. If the noise is of any magnitude, the valve may be stroked sufficiently to introduce actual changes in flow. The nonlinear function is an efficient noise filter, in that, it rejects small-amplitude signals. The result is smoother valve motion and a more stable loop. Figure 5.22 shows comparative records for linear and nonlinear control of a noisy flow loop. The nonlinear controller has proven to be quite effective on pulsating flows too, where the disturbance is periodic rather than random.

### Level Control

Level measurements are often noisy because of splashing and turbulence. In addition, the surface of a liquid tends to resonate hydraulically, producing a periodic signal superimposed on the average level. Since the liquid-level process cannot respond fast enough for a change in valve position to dampen these fluctuations, they ought to be disregarded by the controller. A nonlinear controller does just this, sending a smooth signal to the valve.

It was pointed out in Chap. 3 that many tanks with level controls are intended as surge vessels. In these applications, tight control is inadmissible because it frustrates the purpose of the vessel. A wide proportional band with reset was suggested for control. But the nonlinear controller is, in fact, ideal for this application for two reasons:

1. Minor fluctuations in liquid level will not be passed on to the valve, providing smooth delivery of flow.
2. Major upsets will be met by vigorous corrective action, ensuring that the upper and lower limits of the vessel will not be violated.

This application is often referred to as "averaging level control," because it is desired that the manipulated flow follow the average level in

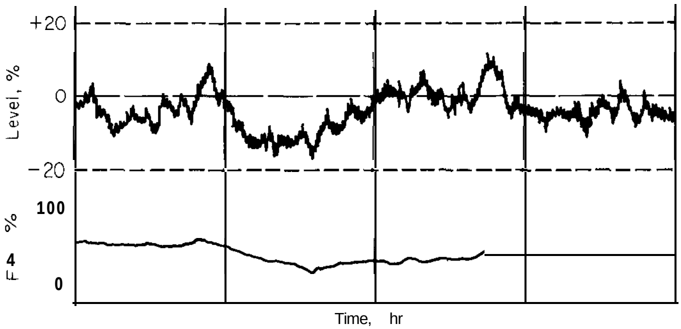



*FIG 5.23. The nonlinear two-mode controller prevents minor fluctuations in level from affecting delivery of flow.*

the tank. Averaging is really a dynamic process and can be accomplished with a suitable lag. But adding a lag would only serve to reduce the speed of response. The nonlinear function, however, provides filtering without sacrificing speed. A typical record of level in a surge vessel and the corresponding output of its nonlinear controller are presented in Fig. 5.23.

## pH Control

The neutralization process has been described as unusually difficult to control because of the extreme nonlinearity of the pH curve. Limit cycling can he encountered when a linear controller is used, because loop gain varies inversely with deviation. This, then, is a natural application for the nonlinear controller whose gain varies directly with deviation. In fact, any process prone to limit cycling can benefit by its USC. The nonlinear function in the controller need not be a perfect complement

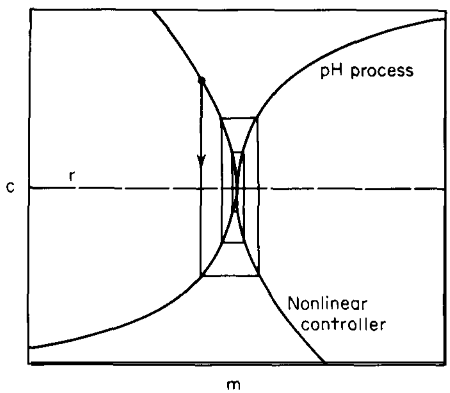



*FIG 5.24. A nonlinear controller can give uniform damping to a pH loop.*

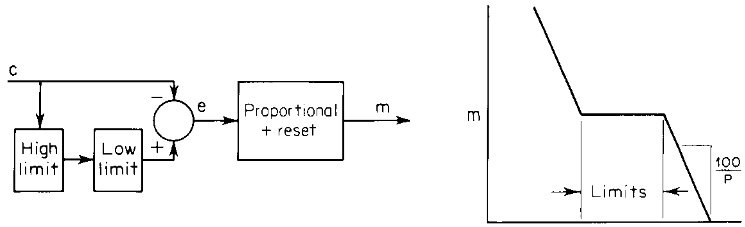



*FIG 5.25. A discontinuous nonlinear controller employs a high-low limiter.*

to the process curve, because any contribution it can make will be an improvement over a linear function. And if the linearity, $\beta$, is adjustable, a reasonable fit can be made.

The input-output graph of Fig. 5.24 shows how a constant loop gain is achieved.

### A Discontinuous Nonlinear Controller

The nonlinear function shown in Fig. 5.19 can be approximated by three straight lines. The center is essentially a dead zone where little or no control action takes place. This function is not difficult to introduce into a linear controller; it involves sending the controlled variable to the set-point input through high and low limits. Within the limits, there is no error signal; elsewhere an error is developed as the difference between the measurement and the nearer limit. Figure 5.25 describes the arrangement of the instrument and its proportional function. Proportional, reset, high, and low limits are adjustable.

This nonlinear controller is often used in averaging level applications. Its dead zone is also a valuable feature in the pH-control system described in Chap. 10.

### PROBLEMS

5.1 A linear process is found to be undamped under proportional control with a hand of 20 percent. What will happen if the band is reduced to 10 percent; to 5 percent?

**5.2** A thermal process with 10-sec dead time and a 5-min lag is to be cooled with refrigerant supplied from a solenoid valve. If the valve is left on, the temperature falls to 0°F; when it is off, the temperature rises to 60°F. Estimate the period and amplitude of the limit cycle if the on-off controller were perfect.

5.3 The on-off controller used for the process in Prob. 5.2 actually has a differential gap of 2°F. Estimate the period and amplitude of the limit cycle, taking the difftrcntial gap into account.

5.4 A lever is driven by a bidirectional constant-speed motor to a position determined by a three-state controller. The motor has a speed of 10 percent of full stroke per second, and an inertial time constant of 1.0 sec. Differential gap in the controller is 2 percent of full stroke. How wide does the dead zone have to be to prevent limit cycling? What would be the period of the cycling?

5.5 A batch chemical reactor is to be brought up to operating temperature with a dual-mode system. Full controller output supplies heat through a hot-water valve, while zero output opens a cold-mater valve fully; at 50 percent output, both valves are closed. While full heating is applied, the temperature of the batch rises at, 1°F/min; the time constant of the jacket is estimated at 3 min, and the total dead time of the system is 2 min. The normal load is equivalent to 30 percent, of controller output. Estimate the required values for the three adjustments in the optimal switching program.

5.6 A given linear process is undamped with a proportional band setting of 50 percent for a linear controller. If a continuous nonlinear controller is used with a linearity setting of $\beta = 0.2$, how narrow can the proportional band be set and still tolerate an error of 20 percent?

# Multiple-Loop Systems

# PART 3

# Improved Control through Multiple Loops

# 6

This chapter deals with situations where a single variable is manipulated to satisfy the specification of a certain combination of controlled variables. In any system with a single manipulated variable, only one controlled variable is capable of independent specification. To put it in other words, there can be only one independent set point at any given time. This, however, does not exclude the incorporation of several controlled variables, as long as their combination contains but one degree of freedom.

Thus we encounter the cascade control system, where the final element is manipulated through an intermediate or secondary controlled variable whose value is dependent on the primary. In ratio control systems, a specification is set on a designated mathematical combination of two or more measured variables. Selective control embodies the logical assignment of the final element to whichever controlled variable (of several) is in danger of violating its specified limits. Finally, adaption is the act of automatically modifying a controller to satisfy a combination of functions of a controlled variable. The common denominator in all these situations is the manipulation of a single final element through more than one control loop.

## CASCADE CONTROL

The output of one controller may be used to manipulate the set point of another. The two controllers are then said to be cascaded, one upon the other. Each controller will have its own measurement input, but only the primary controller can have an independent set point and only the secondary controller has an output to the process. The manipulated variable, the secondary controller, and its measurement constitute a closed loop within the primary loop. Figure 6.1 shows the configuration.

The principal advantages of cascade control are these:

1. Disturbances arising within the secondary loop are corrected by the secondary controller before they can influence the primary variable.

2. Phase lag existing in the secondary part of the process is reduced measurably by the secondary loop. This improves the speed of response of the primary loop.

3. Gain variations in the secondary part of the process are overcome within its own loop.

4. The secondary loop permits an exact manipulation of the flow of mass or energy by the primary controller.

Cascade control is of great value where high performance is mandatory in the face of random disturbances or where the secondary part of the process contains an undue amount of phase shift. For example, a secondary loop should be closed around an integrating element whenever practicable, to overcome its inherent 90” lag. On the other hand, flow is used as the secondary variable whenever disturbances in line pressure must be prevented from affecting the prime variable.

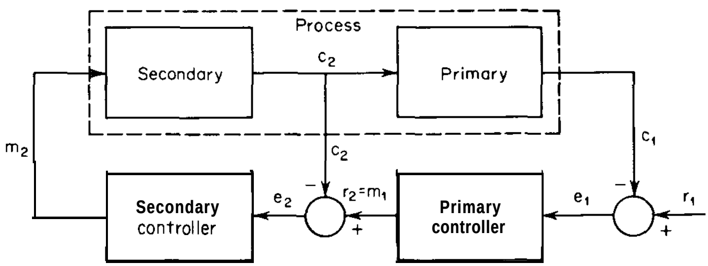



*FIG 6.1. Cascade control resolves the process into two parts, each within a closed loop.*

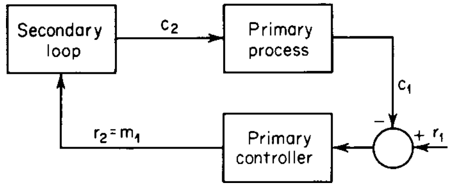



***FIG 6.2. The primary controller sees a closed loop as a part of the process.***

It must be recognized, however, that cascade control cannot be employed unless a suitable intermediate variable can be measured. Many processes are so arranged that they cannot be readily broken apart in this way.

**Properties of the Inner Loop**

The secondary or inner loop confronts the primary controller as a new type of dynamic element. The inner loop can be represented as a single block, the diagram of Fig. 6.1 being resolved into the simpler configuration shown in Fig. 6.2.

Heretofore the dynamic properties of a closed loop were of little concern. The controller was simply adjusted for a damping which satisfied certain transient response specifications. Moreover there was only one period of oscillation to be considered.

But each loop has its own natural period and, as may be expected, the period of the primary loop is to a great extent determined by that of the secondary. Consequently the gain and phase of the secondary loop, whose natural period will be designated $\tau_{o2}$, must be known for any value of the primary period $\tau_{o1}$, since the latter is dependent on the former. The dynamic properties of the open secondary loop can be converted into its closed-loop characteristics by solving for the response of $c_2$ with respect to $r_2$. Refer to the block diagram in Fig. 6.3.

Let $\mathbf{g}_p$ and $\mathbf{g}_c$ be vectors representing gain and phase of the process and the controller, respectively. Then

$$c_2 = m_2\mathbf{g}_p$$
$$c_2 = (r_2 - c_2)\mathbf{g}_c\mathbf{g}_p$$
$$c_2(1 + \mathbf{g}_c\mathbf{g}_p) = r_2\mathbf{g}_c\mathbf{g}_p$$

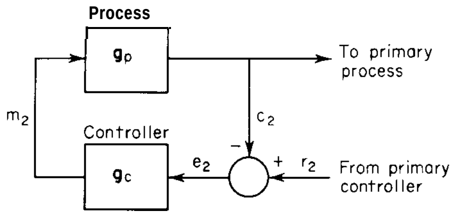



***FIG 6.3. The input to the secondary loop is $r_2$, its output is $c_2$.***

The vector gain of the closed secondary loop will be designated $\mathbf{g}_{o2}$: it is the ratio of output $c_2$ to input $r_2$. The vector consists of a scalar gain $G_{o2}$ and a phase angle $\phi_{o2}$.

$$\mathbf{g}_{o2} = \frac{c_2}{r_2} = \frac{\mathbf{g}_c\mathbf{g}_p}{1 + \mathbf{g}_c\mathbf{g}_p} \tag{6.1}$$

The product $\mathbf{g}_c\mathbf{g}_p$ is the open-loop vector. If the inner loop has been adjusted for ¼-amplitude damping, its open-loop gain will be 0.5 at the period of oscillation. But the phase lag at the period of oscillation is 180°, which makes the gain vector 0.5, $\angle -180°$, or -0.5. The closed-loop vector $\mathbf{g}_{o2}$ at the natural period is then

$$\mathbf{g}_{o2} = \frac{-0.5}{1.0 - 0.5} = -1.0 \qquad \text{or} \qquad G_{o2} = 1.0 \quad \phi_{o2} = -180°$$

If the open-loop gain were 1.0 at the natural period, undamped oscillation would result:

$$\mathbf{g}_{o2} = \frac{-1.0}{1.0 - 1.0} = -\infty$$

This indicates that an infinitesimal change in $r_2$ would change $c_2$ enough so that it would never return to equilibrium, and indeed this is the case.

To find the gain and phase characteristics of a loop away from its natural period, the vector equation for the inner loop must be solved for various values of input period $\tau_{o1}$. This entails first finding the gain and 'phase of the open loop, $\mathbf{g}_c\mathbf{g}_p$. This vector must then be added to the vector 1.0, $\angle 0°$ to form the denominator of the equation. Then the closed-loop gain is the quotient of the magnitude of the two vectors, and its phase is the difference between their phase angles.

**example 6.1**

A typical example is that of a closed loop containing dead time, an integrating capacity, and a proportional controller adjusted for ¼-amplitude damping. The natural period is known to he $\tau_{o2} = 4\tau_{d2}$. The open-loop gain is 0.5 at $\tau_{o1} = \tau_{o2}$ and varies directly as $\tau_{o1}$. The open-loop phase is -90° for the integrating element, with an additional $-360\tau_{d2}/\tau_{o1}$ or $-90\tau_{o2}/\tau_{o1}$ for the dead time. Then

$$\mathbf{g}_c\mathbf{g}_p = 0.5\,\frac{\tau_{o1}}{\tau_{o2}},\ \angle -90° - 90\,\frac{\tau_{o2}}{\tau_{o1}}$$

From this information, closed-loop gain and phase arc plotted in Fig. 6.4.

The primary loop will contain certain elements of the process in addition to the secondary loop. These elements can be expected to contribute phase lag of 90° or more. Therefore the area of greatest interest

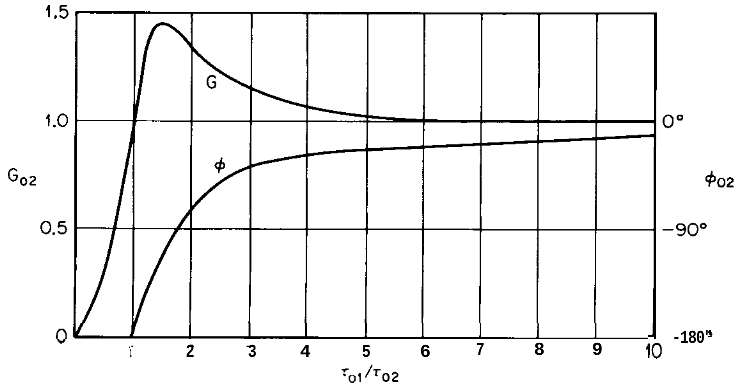



*FIG 6.4. Gain and phase of a typical damped loop.*

in the response of the secondary loop will be where its phase lag is less than 90°. Notice that gain and phase of the closed loop go in opposite directions in this region. This was not true of the common open-loop elements-capacity and dead time. Therefore trouble will be encountered in the primary loop as $\tau_{o1}$ approaches $\tau_{o2}$.

But the closed-loop characteristics have three very important advantages over the corresponding open-loop characteristics, at relatively high values of $\tau_{o1}$:

1. The gain of the closed loop approaches 1.0, which is not only less than the gain of the open loop, but is not subject to variation.

2. The phase of the closed loop is less than that of the secondary process.

3. Where $\tau_{o1}$ is much greater than $\tau_{o2}$, the secondary part of the process is effectively nonexistent.

example 6.2

As a case in point, compare the open- and closed-loop phase in the cited example at $\tau_{o1} = 10\tau_{o2}$:

| Loop | Gain | Phase, deg |
|---|---|---|
| Open........... | 5.0 | -99 |
| Closed. | 1.01 | -11 |

In estimating the performance of a cascade control loop, the following procedure is useful:

1. Determine the period of the inner loop $\tau_{o2}$.
2. Determine the period of the outer loop $\tau_{o1}$, allowing no phase shift for the inner loop.
3. From Fig. 6.4, read the phase lag corresponding to the ratio $\tau_{o1}/\tau_{o2}$.
4. Recalculate $\tau_{o1}$ after including the phase shift of the inner loop.
5. Include the gain of the inner loop at $\tau_{o1}$ (from Fig. 6.4) in estimating primary-loop gain.

### Valve Position as the Secondary Loop

A positioner is used to close the loop around a valve motor. It drives the motor until a mechanical measurement of stem position is balanced against its input signal. In this way, forces that would act to impede the motion of the stem are overcome. A positioner is essentially a high-gain proportional controller closing the loop around what is essentially single capacity plus hysteresis.

The forces affecting the motion of a valve stem are principally friction and pressure drop. It has been pointed out that friction is the cause of hysteresis. And hysteresis can cause limit cycling in the presence of two integrating elements, such as a liquid-level loop with proportional-plus-reset control. But a positioner is stable in the presence of hysteresis and will succeed in eliminating from the primary loop that, source of phase shift.

High pressure drop across the seat of a control valve acts against the area of the seat in opposing the force of the valve motor, It can easily be sufficient to keep the valve from closing tightly. A positioner will place as much pressure on the valve motor as is available if the stem is very far from its directed position and so is able to overcome this difficulty.

The natural period of a valve with positioner will be in the order of 0.5 to 2.0 sec, depending on the size of the valve. It is the fastest loop commonly encountered in process work. But it is not fast, enough to be useful in a flow loop. Look at what a valve positioner would do in the flow loop analyzed in Example 3.2. Without a positioner, the valve contributes more phase lag and attenuation than any other element in the loop. A positioner of 1.0-sec period would reduce the period of the flow loop to perhaps 4 sec. From Fig. 6.4, the position loop gain at a $\tau_{o1}/\tau_{o2}$ of 4 is about 1.07, three times the gain of the valve itself. So in this case the cascade loop is a hindrance rather than a help. The fault lies in placing the inner loop around the largest time constant. But in any system where the valve is not the slowest element, a positioner is beneficial.

### The Cascade Flow Loop

Cascade flow loops are used most often to provide consistent delivery of material to or from the process in response to the demands of the primary controller. They overcome variable pressure drop, valve friction, and nonlinear valve characteristics.

But if the measurement is in the differential form, its nonlinearity becomes part of the primary loop, because flow is being delivered to the process while flow squared is set by the primary controller. The nonlinearity of the differential meter was recognized earlier as a problem in a flow loop. But using a differential meter in the secondary loop of a difficult thermal or composition process is asking for trouble. Suppose the process is linear with respect to flow, as in Example 3.6. The output of the primary controller manipulates differential pressure, however, which varies as flow squared:

$$h = kF^2 \qquad \frac{dF}{dh} = \frac{1}{2kF} \tag{6.2}$$

Loop gain now varies inversely with flow (which is much worse than varying directly with flow, because it can approach infinity). And since many processes are started up or operated for extended periods at low flow, the problem is serious. If the primary controller is not placed in manual, the loop will limit-cycle around zero flow. The best solution is to insert a square-root extractor in the flow-measurement line to linearize the secondary loop.

The problem of variable dead time was discussed in Chap. 2. This problem was resolved by using an equal-percentage control valve. If a flow loop were placed around the valve, however, its characteristic would be lost. Furthermore, if the flow loop were of the differential type, its nonlinearity would be in the wrong direction, making the primary-loop gain vary inversely as the square of flow. These factors deserve careful consideration before deciding on a cascade flow loop.

Because flow loops resonate in the 1- to 10-sec range, they are safe to use in cascade with temperature or composition, but not ordinarily with liquid or gas pressure or other flow loops. Liquid level is only cascaded to flow in applications involving boiling liquids or condensing vapors, where the natural period of the primary loop is long compared to the flow loop.

### Temperature as the Inner Loop

Perhaps the third most common cascade loop is that of temperature. Whereas a material balance can be enforced by flow controllers, temperature controllers are often used to manipulate a heat balance.

Careful control of heat balance is most important in a chemical reaction. To ensure satisfactory performance, the reactor temperature controller generally positions the set point of the coolant temperature controller: a temperature-on-t'emperature cascade loop. The rate of heat transfer to the coolant varies with the temperature difference between reactants and coolant, so the actual value of coolant temperature plays a vital role. Nonlinearities and lags in the coolant loop are, for the most part, removed from the primary loop.

Since the proportional band of the secondary temperature controller is ordinarily 25 percent or less, reset may be omitted. A slight offset in coolant temperature is inconsequential, in that the outer loop will always have reset. Reset in the inner loop would only serve to slow it down. Reset is not used in the coolant temperature loop of a batch reactor.

Again, whether a particular cascade configuration is workable is not so much a matter of what kind of measurements serve as the primary and secondary variables, but rather is a question of the natural periods of the loops differing by severalfold. If both loops have the same kind of measurement, their relationship is ordinarily linear, assuring a constant gain for the primary loop.

## RATIO CONTROL SYSTEMS

In a ratio control system, the true controlled variable is the ratio $K$ of two measured variables X and Y:

$$K = \frac{X}{Y} \tag{6.3}$$

Control is usually effected by manipulating a valve influencing one of the variables, while the other is uncontrolled or "wild." The obvious way to implement the ratio control function is by computing X/Y, as shown in Fig. 6.5. But this is not the best way.

Figure 6.5 has a divider within the closed loop, regardless of which variable is affected by the output of the controller. If X is manipulated, loop gain changes with the wild variable Y:

$$\frac{dk}{dx} = \frac{1}{Y} \qquad \frac{dK}{dY} = -\frac{X}{Y^2} = -\frac{K}{Y} \tag{6.4}$$

If Y is manipulated, the loop becomes nonlinear in that the gain changes with controller output.

All these problems can be overcome by moving the calculation out of the closed loop. Ratio control is then brought about in the set-point

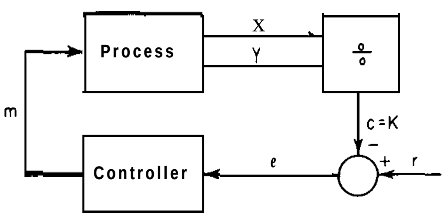



*FIG 6.5. Using the ratio of two variables as the input to a controller is not recommended.*

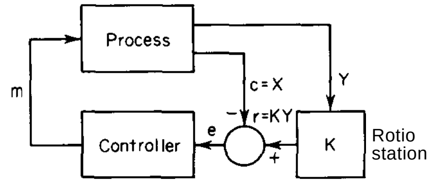



*FIG 6.6. In the recommended system, the ratio calculation is outside the loop.*

circuit, making r = KY if X is controlled, or r = X/K if *Y is* controlled. Figure 6.6 shows the set-point calculation. In this configuration, one of the variables becomes controlled and the other serves to generate a set point. The wild variable is multiplied by the adjustable coefficient *K* in a device called a "ratio station."

## Ratio Flow Control

By far the most common application of ratio control is in flow systems. Ingredients for a reactor, for example, are introduced in precise proportions. Any number of streams can be set in ratio to one independent flow, allowing the plant to be started or shut down by the manipulation of that key variable alone-all others must follow. Or, if all flows are under control, total plant throughput can be commanded through a single master set station, as shown in Fig. 6.7.

Up to this point, the ratio stations have all been outside the closed loops. But a natural extension to the configuration in Fig. 6.7 is the automatic repositioning of a master set point by the output of another controller. Consider the example of a chemical reactor which is to be operated consistently at maximum throughput, as determined by its heat-removal capability. The only control system which could satisfy this specification would feature a temperature controller setting the flow rates of all ingredients into the reactor, as shown in Fig. 6.8.

Because this configuration places the ratio stations within the primary loop, loop gain varies with their settings. In such an arrangement, constant loop gain can ordinarily be assured if the sum of the ratio settings is constant. Thus a percentage increase in one ingredient should be accompanied by a corresponding decrease in the others. This is often done for accounting purposes anyway, by making the sum of the ratios

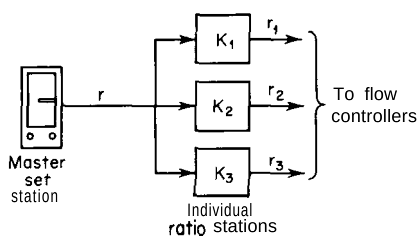



*FIG 6.7. A single master station can set all the plant streams through ratio stations.*

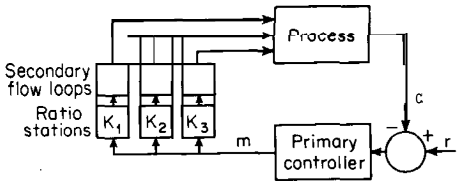



***FIG 6.8.*** ***Several secondary loops can be set in cascade from one primary controller.***

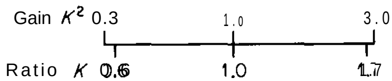
always equal 100 percent. Then the master station or primary controller sets the true total flow. Since the prime manipulated variable in Fig. 6.8 is total flow, the prime controlled variable must be a function of total flow and be relatively independent of the individual ratio settings.

Most ratio stations have a gain range of about 10: 1. Since the normal ratio setting for most applications is in the vicinity of 1.0, this is selected as the midscale position. (This assumes that the flowmeters for the various streams are sized in proper proportion, so that under normal operation they will all read about the same percentage flow.) Thus the gain of a typical ratio station would be adjustable from 0.3 to 3.0.

But most flowmeters are of the differential type, in which case flow squared is the transmitted variable. Then,

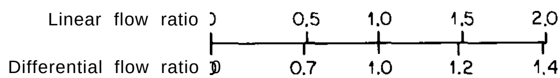
$$X^2 = K^2Y^2$$

The gain of the ratio station is the square of the ratio setting; in other words, the ratio setting is the square root of the gain. The ratio scale for differential meters is compared with a linear scale below :




### Setting the Ratio Remotely

The ratio stations discussed thus far are nothing more than amplifiers with adjustable gain. The ratio setting was introduced manually. If it is desired that the ratio setting be introduced as a control signal, a different device must be used. The device is, in fact, defined as a multiplier, because its output is the product of two analog signals. The gain of a multiplier is directly proportional to each of its inputs. If a midscale ratio of 1.0 is desired, the gain range of the multiplier must be 0 to 2.0. The multiplier is then said to have an element factor of 2.0. Linear and differential flowmeter ratio scales for the multiplier are compared below:




The use of a multiplier permits setting of the ratio by the output of a primary controller. This is desirable where the primary variable is a

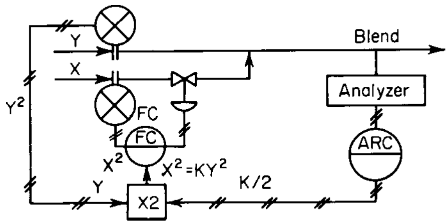



*FIG 6.9. The ratio of the two flows is automatically set in cascade to maintain composition control.*

function of the ratio of the flowing streams, apart from their total flow. A typical application is the blending of ingredients to form a mixture of controlled composition. Figure 6.9 shows the arrangement of the control loops for a two-component system. Notice that the signals are in the differential form. Although the gain of the cascade flow loop varies inversely with its flow, the gain of the multiplier varies with the wild flow squared. Thus loop gain is

$$\frac{dX}{dK} = \frac{Y^2}{X} \quad \frac{x}{K}$$

Direct variation of gain with flow is much preferable to inverse variation.

Compare this cascade control system to the one in Fig. 6.8 where total flow was the primary manipulated variable.

### Infinite Ratio Rangeability

The range of adjustment available with standard ratio stations is typically 0.3 to 3.0. Using a single multiplier, a range of 0 to 2.0 is customary. Occasionally an application will be encountered where these ranges are insufficient. In this event, two multipliers can be used, each setting the rate of one flowing stream. Mathematically, the system is organized to deliver a set total flow F:

$$F = X + Y$$

Then the percentage of X in the total is introduced as an independent setting $x$:

$$X = Fx$$

The flow of the other stream Y is then the difference:

$$Y = F(1 - x)$$

The controlled ratio X/Y can then be varied from zero to infinity as $x$ varies from zero to one:

$$\frac{X}{Y} = \frac{x}{1 - x} \tag{6.5}$$

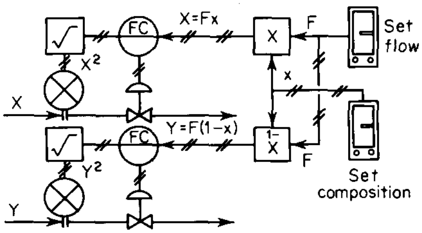



***FIG 6.10. The use of two multipliers affords an infinite ratio range.***

Because the calculation is based on the sum of the two flowing streams, it will only work with linear flow signals. Independent selection of total flow and percent composition are available, as Fig. 6.10 indicates.

Note. that the lower of the two multipliers has a reverse-acting input. When $x$ is 1.0, the gain of this multiplier is zero. Both element factors are unity, which is necessary to satisfy the equation.

Although infinite ratio range is theoretically possible, it must be recognized that differential meters are only accurate within a 4: 1 flow range, which will necessarily limit the actual ratio range.

### Digital Blending Systems

The use of turbine and positive displacement meters has generated a new type of flow ratio system. These devices are inherently linear and are of a fairly wide range; but of most significance, they do not produce an analog output. Each rotation of their moving elements produces a discrete number of pulses representative of a particular volume of fluid that has passed through. The pulse rate or frequency is proportional to the flow rate, and the total number of pulses transmitted over a given length of time is a measure of the volume of fluid delivered.

To take advantage of the discrete nature of the measurements, a special type of control system has been developed. In the digital blending system, the volume delivered through each meter is continuously compared with the volume desired for that stream. An error between the two generates a control signal that manipulates the valve in that stream. Functionally, the loops are arranged similarly to what was shown in Fig. 6.7. A master set station transmits a pulse rate that paces the entire system, demanding a certain total flow rate. This pulse rate is multiplied by the ratio setting for each flow loop, thus producing a rate set point for each loop. There the similarity to an analog system ends.

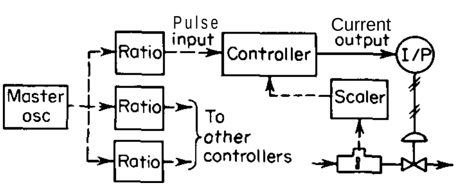



***FIG 6.11. The output of the controller is the only analog signal in a typical digital blending system.***

The object of the controller is to maintain the set volume delivery. This allows a temporary reduction of flow in any loop, to be made up to the correct total once its cause has been removed. The basic control mechanism is a digital up-down counter. Set-point pulses cause an increase in the register, while measurement pulses bring about a decrease. The difference between the number of pulses from the two sources is stored and converted to an analog signal which may be used to drive a valve. The valve would then be driven proportionally to the volume error or integrated flow error. Thus the control mode is integral (reset only). Let the volume error required for 100 percent change in output be identified as $V$ and the maximum flow rate as $F$. The percent output to the valve in terms of percent demanded and measured flow is then

$$m = \frac{1}{V} \int F(r - c)\, dt \tag{6.6}$$

Then the reset time of the integrator is

$$R_1 = \frac{V}{F} \tag{6.7}$$

To be sure, the valve position changes in discrete steps, as does the output of the up-down counter. The steps are about 1 percent, and the valve will attempt to follow each pulse registered by the counter. But because most valves are incapable of following the pulse rate of the counter, which may be as high as 100 pulses/sec, delivery is reasonably smooth.

Although integral control is capable of reducing a flow-rate error to zero, integral offset can exist. Examine the flow delivery following startup as shown in Fig. 6.12. The set flow rate is already at its proper value, but it takes time to bring the measured flow rate to the same value. Eventually there will be no rate error, but in the meantime a significant volume error appears as the area between the curves.

The integrated error $E$ required to generate a given controller output is

$$E = \int e\, dt = R_1 m$$

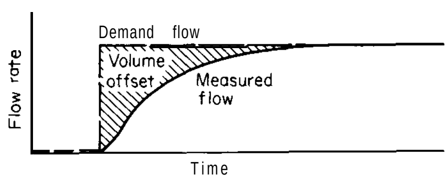



*FIG 6.12. An integral control system is subject to volume offset.*

The actual volumetric offset is the percent integrated error $E$ times the maximum flow rate:

$$EF = mV \tag{6.8}$$

Each valve position, hence each flow rate, has a related volume offset.

On the face of each digital control station are two counters. One registers the total demand, while the other registers the total delivered flow. Volume offset appears as a difference of perhaps 5 or 10 between the two counters. In the steady state, the two count rates are identical, so the volume offset is constant. In a continuous process, volume offset is of little significance, representing only a few counts out of increasing thousands. But in small batches of only a few hundred counts, it can be cause for concern.

The obvious solution is to add a second integral mode. But double integral by itself is unstable, in that it produces 180” phase lag at all periods. But if the first integral, i.e., the volume error, is acted upon by a proportional-plus-reset controller, the system can be stable. Such an arrangement is functionally described in Fig. 6.13.

Mathematically, the up-down counter totals up the difference between the demand and the measured rates, appropriately scaled:

$$EF = \int F(r - c)\, dt$$

The D-A converter fixes the range $V$ of volume error $E$ over which the control valve is driven. Substituting Eq. (6.7) and adding proportional-plus-reset $(R_2)$ action:

$$m = \frac{100}{P}\left(\frac{E}{R_1} + \frac{1}{R_1R_2}\int E\, dt\right) \tag{6.9}$$

The expanded control equation is

$$m = \frac{100}{PR_1}\left[\int (r - c)\, dt + \frac{1}{R_2}\iint (r - c)\, dt^2\right] \tag{6.10}$$

Fig. 6.14 shows how the second integral works to eliminate volume off set.

Flow loops with conventional controllers, i.e., proportional-plus-reset, have their natural period in the region of 1 to 10 sec. The phase contribution of a properly adjusted controller at the natural period is normally 30” or less. But the blending controller contributes about 100” lag,

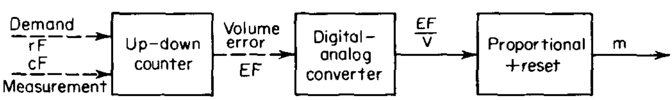
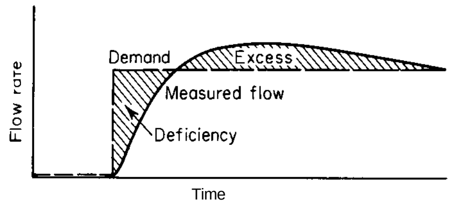



*FIG 6.13. A second integration is required to eliminate the volume error.*

*FIG 6.14. The second integral caneels volume deficiency with an equal excess, but takes longer to reach a steady state.*




generally doubling the period of a loop. Thus blending loops will oscillate from about a 5- to a 20-sec period.

Presence of any valve hysteresis, in combination with the two integrations, will cause a limit cycle, as explained in Chap. 5. Although a valve positioner is not ordinarily recommended for flow, the period of a blending loop is usually far enough removed from that of the position loop for one to be safely used.

## SELECTIVE CONTROL LOOPS

Frequently a situation is encountered where two or more variables must not be allowed to pass specified limits for reasons of economy, efficiency, or safety. If the number of controlled variables exceeds the number of manipulated variables, whichever ones are in most need must logically be selected for control. (This is the case of the squeaky wheel getting the grease.) Automatic selector units are available for this type of service. They are employed in four basic areas of application:

1. Protection of equipment
2. Auctioneering
3. Redundant instrumentation
4. Nonlinear control functions

As an example of how equipment might be protected by a selective control system, consider a compressor whose discharge is ordinarily on

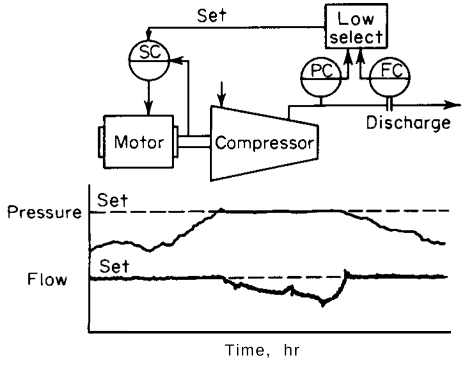



*FIG 6.15. Motor speed is manipulated by whichever controller has the lower ouput.*

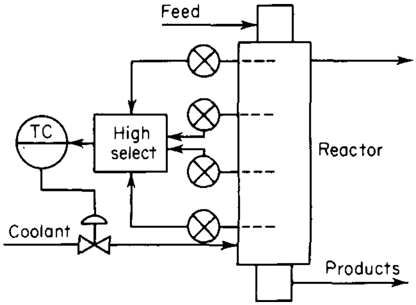



*FIG 6.16. A high selector is used to permit control of the peak reactor temperature.*

flow control, except that discharge pressure must not be allowed to exceed a given limit. During conditions of low load, the pressure controller must be allowed to assume control, thereby reducing flow. When the demand for gas is high, the flow controller will take over to see that its set point is not exceeded.

Figure 6.15 shows how the controller that has the lower output is selected to manipulate motor speed. Decreasing motor speed will reduce both flow and pressure, so the use of a low selector guards against an excess of either. The record shows how pressure is allowed to drift below its set point during conditions of high load, while flow is controlled. Conversely, when the load is low enough to raise the discharge pressure to its set point, flow is reduced.

"Auctioneering" is a term used to describe the selection of the highest of a battery of inputs. An example is the control of the highest temperature in a fixed-bed reactor. The possibility exists that the location of the highest temperature may shift with catalyst degeneration, flow, etc. Temperatures all along the reactor would then be compared and the highest used for control.

To protect against an instrument failure placing the plant in a hazardous condition, key instruments can be duplicated. Often, duplicated

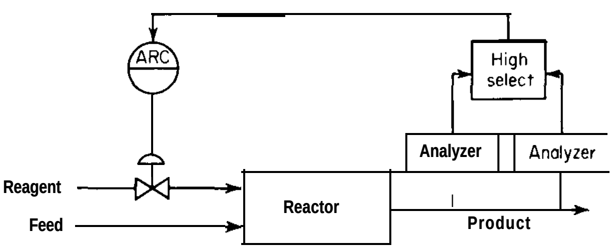



*FIG 6.17. The high selector prevents a failure of either analyzer from damaging the reactor with excess reagent.*

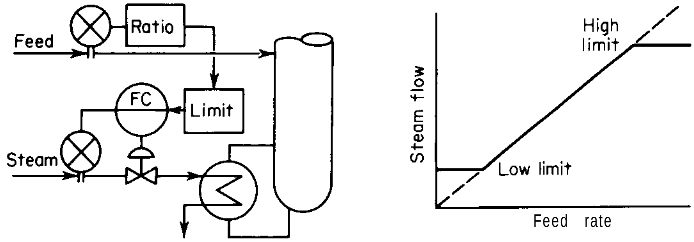



*FIG 6.18. The tower is protected against either excess or deficient steam flow by the high-low limiter.*

devices are used only for record or alarm purposes. But where closed-loop control is involved, automatic selection of the manipulated variable must be provided.

Analyzers are generally less reliable than other instruments. Figure 6.17 illustrates a system that would allow control to be maintained in the event of downscale failure of either analyzer. An upscale failure would be allowed, which would shut down the reactor, but this is a safe condition.

The introduction of high and low limits is another form of automatic selection. The same mechanisms are used, except that the limits are introduced manually instead of being other variables.

Consider a distillation column whose heat input is being controlled in ratio to the feed rate. Throughout the normal operating range, this ratio would be maintained. But even if the feed should drop to zero, heat input must not, because it could cause loss of liquid in the trays. An excessively high heat input is also to be avoided, because flooding of the tower could result. To avoid the possibility of these accidents, high and low limiters can be used, as illustrated in Fig. 6.18.

## Protection against Windup

When one controller is selected from two or more, the others are in an open-loop condition. If these controllers have reset action, which is most often the case, they need to be protected against windup. This is accomplished by using the output of the selector as a common feedback to all controllers. In this way, the selected controller will have its own output fed back and therefore will have reset action. But the others in the system will have a reset signal which is not their own output, forcing them to respond like proportional controllers. The controller arrangement is shown in Fig. 6.19 for the pressure-flow system that appeared in Fig. 6.15.

Automatic transfer from one controller to another takes place at the instant when the outputs are equal. This fact, coupled with the common reset signal, means that transfer is bumpless.

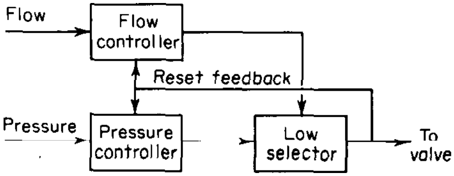



*FIG 6.19. The output of the selector is used as a common feedback signal to overcome windup.*

Assume that the flow controller is presently manipulating compressor speed, in that its output is the lower of the two. Its output $m_F$ is then

$$m_F = \frac{100}{P_F}\left(e_F + \frac{1}{R_F}\int e_F\,dt\right) \tag{6.11}$$

But because the output $m_p$ of the pressure controller is greater,

$$m_p = \frac{100}{P_p}e_p + m_F \tag{6.12}$$

Transfer from flow to pressure control will take place when $m_p = m_F$. This requires that $e_p = 0$, but does not require the same of $e_F$. To generalize, a controller can only bc selected when its error crosses zero. For controllers with derivative on the input, $e + D\,de/dt$ would have to be zero. This allows transfer before zero error is reached, avoiding the overshoot that would otherwise be encountered.

There is no reason for auto-manual transfer switches in any of the controllers, if there is a transfer switch at the output of the selector. Because of the common feedback circuit, transfer to manual at that point, will not cause windup.

Occasionally a situation n-ill arise when none of the controllers in a selective loop assumes control. Some processes are prone to this, although in most it cannot happen. In the process shown in Fig. 6.15, neither controller would take over if suction pressure were to fall low enough. In this case, discharge pressure and flow would both be below their set points and the motor would go to maximum speed. Both controllers would then saturate. When suction pressure was restored, discharge pressure or flow or both could overshoot their set points. If this were a common occurrence, an antiwindup switch should be inserted in the common feedback line at the output of the selector.

It should be mentioned that a controller whose output passes through a high or low limiter is also capable of windup when these limits are exceeded. Reset feedback may be taken downstream of the limiters if the problem is serious.

## ADAPTIVE CONTROL SYSTEMS

An adaptive control system is one whose parameters are automatically adjusted to compensate for corresponding variations in the properties of the process. The system is, in a word, "adapted" to the needs of the process. Naturally there must be some criterion on which to base an adaptive program. To specify a value for the controlled variable (i.e., the set point) is not enough-adaption is not required to meet this specification. Some "objective function" of the controlled variable must be 'specified in addition. It is this function that determines the particular form of adaption required.

The objective function for a given process may be the damping of the controlled variable. In essence, there are then two loops, one operating on the controlled variable, the other on its damping. Because damping identifies the dynamic loop gain, this system is designated *a dynamic* adaptive system.

It is also possible to stipulate an objective function of the steady-state gain of the process. A control system designed to this specification is then *steady-state* adaptive.

There is, in practice, so little resemblance between these two systems that their classification under a single title-adaptive-has led to much confusion.

A second distinction is to be made, this not on the objective function, but rather on the mechanism through which adaption is introduced. If enough is known about a process that parameter adjustments can be related to the variables which cause its properties to change, adaption may be *programmed.* However, if it is necessary to base parameter manipulation upon the measured value of the objective function, adaption is effected by means of a feedback loop. This is known as a *self-adaptive* system.

### Dynamic Adaptive Systems

The prime function of dynamic adaptive systems is to give a control loop a consistent degree of stability. Dynamic loop gain is then the objective function of the controlled variable being regulated; its value is to be specified.

The property of the process most susceptible to change is gain. In some cases the steady-state gain changes, which is usually termed a nonlinearity. Other processes exhibit a variable period, which reflects upon their dynamic gain. But by whichever mechanism loop stability is affected, it can always bc restored by suitable adjustment of controller gain. (This assumes that the desired degree of damping could be achieved in the first place, which rules out limit cycling.)

Many cases of variable process gain have already been cited. In general, an attempt is made to compensate for these conditions by the introduction of selected nonlinear functions into the control system. For example, the characteristic of a control valve is customarily chosen

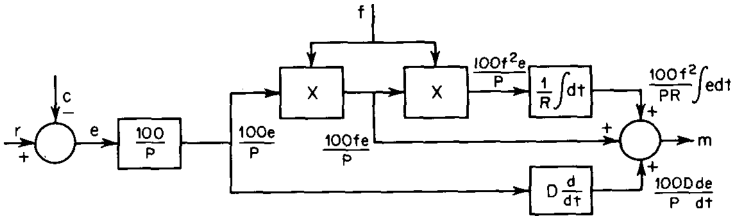



*FIG 6.20. Two multipliers are necessary for Row adaptation of all three modes.*

with this purpose in mind. But compensation in this way can fall short for several reasons.

1. The source of the gain variation lies outside the loop, and hence is not identified by controller input or output.

2. The required compensation is a combined function of several variables.

3. The gain of the process varies with time.

Perhaps the most readily assimilated example of a process which could benefit by adaptive control is that of a single capacity plus variable dead time. Dead time would vary inversely with flow, in the manner shown in the heat-exchanger example given in Chap. 2. Response of the loop without adaption was presented in Fig. 2.14. In this example, an equal-percentage valve was used to provide gain compensation for changes in flow. This method worked, but it in turn made loop gain dependent on the magnitude of the controlled variable. The trade-off was inevitable because the variable which affected the process gain, i.e., flow, is outside the loop.

Exact compensation may be obtained by programming the settings of the controller as functions of flow. Because the period of the loop varies directly with dead time, derivative and reset time ought to vary inversely with flow. And since process dynamic gain varies inversely with flow, the proportional band should too. Knowing this, it is possible to write a flow-adapted control algorithm:

$$m = \frac{100f}{P}\left(e + \frac{f}{R}\int e\,dt + \frac{D}{f}\frac{de}{dt}\right) \tag{6.13}$$

The adaptive term j is the fractional flow through the process, and $P$, $R$, and $D$ are the optimum settings at full-scale flow. Placing all the j terms inside the parenthesis indicates how the adaption might be performed:

$$m = \frac{100}{P}\left(fe + \frac{f^2}{R}\int e\,dt + D\frac{de}{dt}\right) \tag{6.14}$$

Note that the derivative term does not need adaption, but that reset must be multiplied by flow twice. Figure 6.20 illustrates the implementation of the three-mode adaptation outlined by Eq. (6.14).

If the flow measurement is in the form of differential pressure, greater accuracy would be obtained by multiplying the error separately by $f^2$ (differential pressure) for reset adaptation. In this way, the adaptive signal to the integrator would have passed through a single multiplier rather than a square-root extractor and two multipliers.

### Dynamic Self-adaptive Controllers

A great deal of research effort has been spent in several industries in the quest for a self-adaptive controller. The application goes beyond compensating for variable loop gain, because a device which could adapt itself could also relieve the operator of the task of adjustment altogether. Thus the performance of a critical loop could be made independent of the skill of the operator. (Although instead, it is made doubly dependent on the skill of the control designer.)

Again, the purpose of this adaptive loop is to regulate system damping. If the normal state of the system is steady, no measurement of damping is available. If the self-adaptive function is to work, then, some means of perturbing the state of the process must be decided upon. Either a periodic disturbance may be introduced as a test signal, or the system must wait for disturbances to occur naturally. Each of these methods has certain disadvantages. The first is generally inadmissible in that it effectively robs the process of its rightful steady state.

Since the second method does not test the process, the current value of loop gain is unknown until a disturbance identifies it. Identification must then be carried out, and parameter adjustment made carefully to prevent overcorrection. Identification consists principally of factoring the response curve into high- and low-frequency components whose ratio represents the dynamic gain of the closed loop. The load-response curve shown in Fig. 6.21 is so separated.

Having thus identified the damping of the loop, the task of adjusting it remains. This must be recognized as a feedback operation-manipulating a parameter on the basis of a measurement made on the controlled variable. The arrangement of the loop is shown in Fig. 6.22.

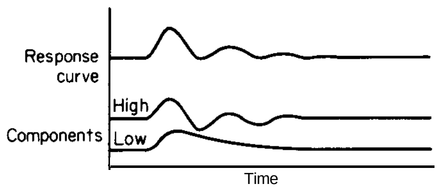



*FIG 6.21. If the loop is properly damped, high- and low-frequency components will exist in a certain ratio.*

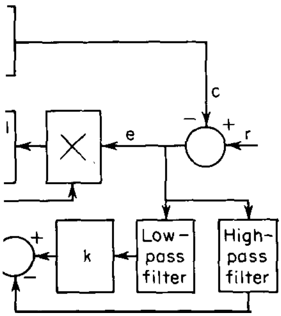



*FIG. 6.22. The self-adaptive loop controls the ratio of high- to low-frequency components in the error signal.*

The adaptive loop embodies ratio control wherein the ratio appears as coefficient $k$. Following is a dead-band filter, preventing noise and insignificant disturbances from affecting loop gain. An integral mode is used for introducing the adaptive signal in order to retain its value (which determines the gain of the primary loop) in the steady state.

The adaptive loop has a natural period $\tau_a$ which must exceed that of the prime loop by an amount related to the phase lags in the filters and in the integrator. Consequently the adaptive loop is much slower than the primary loop, so that its effect upon sudden changes in process gain, like those shown in Fig. 2.14, would be small. The programmed adaptive system of Fig. 6.20 would outperform it by a tremendous margin. The programmed adaptive system is faster, in fact instantaneous in response, less expensive, more reliable, and involves no risk whatever. It must be recognized that the self-adaptive controller could be in any condition at startup, thus doubly complicating an already cumbersome auto-manual transfer procedure.

At this writing, programmed adaptive systems have been used for certain critical applications such as temperature control in once-through boilers[1] and heat exchangers.[2] But there apparently is no published report on a self-tuned controller operating successfully on a critical loop in a process plant. (They have been used in aircraft controls.) From the foregoing discussion it should be evident that, no matter how skillfully mechanized, a self-tuning controller is by no means a panacea.

### The Steady-state Adaptive Problem

Where the dynamic adaptive system controlled the dynamic gain of a loop, its counterpart seeks a constant steady-state process gain. This implies, of course, that the steady-state process gain is variable and that one particular value is most desirable.

Consider the example of a combustion control system whose fuel-air ratio is to be set for highest efficiency. Excess fuel or air will both reduce efficiency. The true controlled variable is efficiency, while the true manipulated variable is the fuel-air ratio. The desired steady-state gain in this instance is

$$\frac{dc}{dm} = 0 \tag{6.15}$$

The system is to be operated at the point where either an increase or decrease in ratio decreases efficiency. This is a special case of steady-state adaption known as "optimizing." A gain other than zero may reasonably be stipulated, however.

Where the value of the manipulated variable which satisfies the objective function is known relative to conditions prevailing within the process, the adaption may be easily programmed.[3] As an example, the optimum fuel-air ratio may be known for various conditions of air flow and temperature. The control system may then be designed to adapt the ratio to air flow and temperature much in the way that the controller settings were changed as a function of flow in the example of dynamic adaption.

If a reasonably accurate mathematical model of the process may be obtained in a simple form, it may be differentiated to solve for the adaptive control program. In the following expressions, let $K_r$ represent the desired gain of the process. Consider the example of a variable-gain process affected by a load term $q$:

$$c = am - qm^2 \tag{6.16}$$

(Note: If c were directly proportional to $m$, the process gain would be constant and there would be zero degrees of freedom.) Differentiation solves for process gain, which is set equal to $K_r$:

$$\frac{dc}{dm} = a - 2qm = K_r \tag{6.17}$$

Next, Eq. (6.17) is solved for $m$, which is the output of the control system:

$$m = \frac{a - K_r}{2q} \tag{6.18}$$

Figure 6.23 shows how such a control system would be arranged.

Because the system described above has no feedback loop, it does not rightly belong in this chapter. Therefore further discussion of this class of system, which is growing in importance, will be relegated to Chap. 8, Feedforward Control Systems.

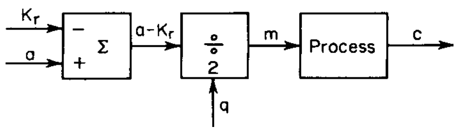



*FIG 6.23. The steady-state adaptive system does not have a feedback loop.*

### A Continuous Self-optimizing Controller

Like the self-tuning controller, the self-optimizing controller requires no prior knowledge of plant conditions, but instead, conducts its own search. Its goal is to keep the manipulated variable at the point where process steady-state gain $dc/dm$ satisfies the specification. But before this can be done, the controller must first test the process for its gain at each point in the search. The test may be conducted continuously or intermittently.

Process gain $dc/dm$ cannot be measured directly, so it must be inferred from the rate of change of input and output.

$$\frac{dc}{dm} = \frac{dc/dt}{dm/dt} \tag{6.19}$$

If an integrating controller with reset time $R$ is used to manipulate $m$ in response to an error signal e, then

$$\frac{dm}{dt} = \frac{e}{R}$$

The process gain is then

$$\frac{dc}{dm} = \frac{R\,dc/dt}{e} \tag{6.20}$$

The system must be designed to come to rest when process gain equals the desired value, here designated $K_r$. At this point the error is zero. A continuous control system devised to the above plan appears in Fig. 6.24.

This type of an optimizer has been found workable on fast processes such as combustion control,* where dynamics play a minor role. But any phase lag whatever in the process will cause the system to overshoot. Although the controller contains a differentiator and an integrator in series, their phase contributions do not cancel because of the nonlinear

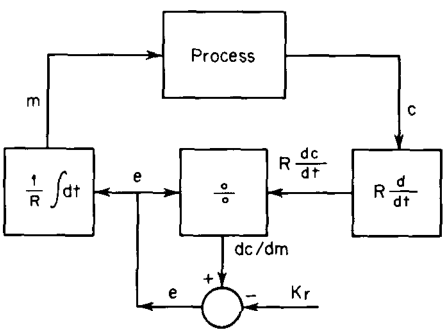



*FIG 6.24. When the process gain equals the reference value, the system will come to rest.*

element (the divider) separating them. To prove that this controller is actually an integrator, $m$ will be solved in terms of c for the case where $K_r = 0$. Since the output of the divider is also the input to the integrator,

$$e = \frac{R\,dc/dt}{e} = R\frac{dm}{dt}$$

$$\frac{R\,dc}{dt} = e^2 = R^2\left(\frac{dm}{dt}\right)$$

Eliminating e,

$$\frac{dc}{dt} = R\left(\frac{dm}{dt}\right)^2$$

Next substitute for $dc/dt$:

$$\frac{dc}{dm}\frac{dm}{dt} = R\left(\frac{dm}{dt}\right)^2$$

Cancelling $dm/dt$ and integrating by time yields

$$m\,dm = \frac{1}{R}\int dc\,dt \tag{6.21}$$

Reset time $R$ must be adjusted for damping as with any other integrating controller. But the fact, that this is an integrating controller places an important limitation on its service: it *cannot* be used on non-self-regulating processes.

The divider in Fig. 6.24 must operate in all four quadrants, because either the denominator or the numerator or both may be negative. As the error, which is the denominator, passes through zero, the gain of the divider changes from plus to minus infinity or vice versa. Obviously, then, the system is extremely sensitive to noise around the point of equilibrium, i.e., at zero error. This has some undesirable features, but unfortunately is necessary for the system to function.

In the steady state, $dc/dt = 0$ as does $dm/dt$. Therefore if the original state of the system is at rest at the wrong value of gain, it will not change its state without a disturbance.

The signs applied to the summing junction in Fig. 6.24 would be used on a process whose gain decreased as $m$ increased. The process-characteristic curve (c vs. $m$) could go through a maximum, and control could be effected at that point, in which case $K_r = 0$. If the process gain were to increase with $m$, the signs at the summing junction would have to be reversed.

It should be pointed out that equipment limitations prevent the use of this system on very slow processes. As $R$ increases, differentiation becomes less accurate. Differentiation is at best an approximation, anyway, because of the filtering that must be used for noise reduction and stability. This filtering is normally a lag of value about $0.1R$. The presence of this additional lag causes the phase of the controller to go beyond $-90°$ at the natural period of the loop, making it a rather poor controller, from the point of view of stability.

### A Sampling Optimizer

To overcome the equipment-limitation problem on slow processes, the optimizing search may be carried out discretely, using sampled data.[5,6] This amounts to supplanting the differentials in the previous example with differences:

$$\frac{AC}{\Delta m} = \frac{\Delta c/\Delta t}{\Delta m/\Delta t} \tag{6.22}$$

The sampling interval At must be long enough to let the process return to equilibrium after each change in controller output. When At has expired, the most recent $\Delta c/\Delta m$ is calculated and compared to $K_r$:

$$e_n = \frac{\Delta c_n}{\Delta m_n} - K_r$$

Next, the output of the controller is stepped proportional to the error signal by a gain *K,:*

$$Am_{n+1} = K_c e_n$$

The effective reset time is related inversely to gain and directly to sampling interval *At:*

$$R = \frac{\Delta t}{K_c} \tag{6.23}$$

The sampling optimizer is not affected by the same equipment limitations as its continuous counterpart, but its other characteristics are similar. The sampling is of some advantage on processes dominated by dead time, but introduces the same uncertainty factor as other sampling controllers encountered. Notice the similarity of the control programs to that of the incremental DDC algorithms. Naturally a DDC computer may be readily programmed with this optimizing function. Sampling, however, still does not permit its use on processes without self-regulation.

### A Peak-seeking Controller[7]

The heart of a peak-seeking controller is a "one-way" storage device: it accepts only increasing inputs. This function can be performed by a capacitor charged through a diode, as shown in Fig. 6.25. In an optimizing control system, the difference between the input, and output of the

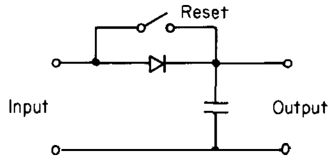



*FIG 6.25. The peak storage circuit must be capable of being reset to accept new information.*

peak storage circuit determines the direction of control action. The output acts as the set point, while the input is the controlled variable; their difference is the error.

As long as the error is zero, the manipulated variable is being driven in such a way as to increase the controlled variable. But the appearance of an error indicates that the controlled variable has started to fall. At this point, the direction of the manipulated variable must be reversed to relocate the peak. This reversal has to be maintained long enough for the process to return to its peak value, so a time-delay lock on the reversal switch must be enforced. The peak storage circuit also requires resetting upon each reversal, otherwise it would cease to function.

The manipulated variable is driven at a constant speed either up or down, as the reversal switch dictates; it never comes to rest. As a result, the process will limit-cycle about the peak. The period of the limit cycle is that of the time delay, if it is set long enough to allow the process to recover from a reversal. Otherwise the process will cycle at its natural period.

As with other self-optimizers, this, too, is an integrating device. It therefore cannot be used on non-self-regulating processes. This class of processes has no equilibrium, no steady state, so the characteristic curve never comes to rest-it is always floating. Unfortunately, floating control action cannot hold it.

## SUMMARY

In the earlier chapters the point was made that the characteristics of a process determine how well it can be controlled. Furthermore, the settings of the control parameters were shown to depend directly on these properties. But this chapter demonstrates that control improvement is possible, and additional specifications can be satisfied as well, by using more information from the process. This necessitates, however, a deeper understanding of the process than the earlier work required. The trend will continue, consummated in application work that is exclusively process-oriented.

The treatise on adaptive control just concluded should verify the value of such an orientation. Although it is possible to design a controller to adapt itself to the process characteristics with no foreknowledge, a controller already equipped with this knowledge is faster, more accurate, and more reliable. These conclusions will be expanded in Chap. 7, Multivariable Process Control. But the deepest penetration into the design of intelligent systems will come in Chap. 8, where the role of feedback will be largely eclipsed.

## REFERENCES

1. Strohmeyer, C., Jr., K. T. Momose, and T. C. Reits: New Techniques of Control and Fluid Circuitry in Design of Once-through Steam Generators, 6th *Natl. Power Instrumentation Symp.*, May, 1963.
2. Shinskey, F. G.: Feedforward Control Applied, ISA Journal, November, 1963.
3. Shinskey, F. G.: Analog Computing Control for On-line Applications, *Control Eng.*, November, 1962.
4. Mathias, R. A., and R. I. Van Nice: Adaptive Control Systems-A Survey, *Electra-Technology,* October, 1960.
5. White, B.: The Quarie Optimal Controller, *Instruments and Automation,* November, 1956.
6. Bernard, J. W., and F. J. Soderquist: Dow Evaluates Optimizing Control, *Control Eng.,* November, 1959.
7. Mamzic, C. L.: Peak-seeking Optimizers, *Instruments and Control Systems,* October, 1962.

## PROBLEMS

6.1 The reactor in Example 3.4 is to be equipped with a cascade controller on coolant exit temperature $T_{c2}$. Estimate the natural period of the primary loop and compare it to that without cascade control.

6.2 One of the streams in a digital blending system has a process time constant of 0.5 sec, a turbine meter time constant of 0.1 sec, and a valve with positioner of natural period 1.0 sec. Full-scale flow is 20 gpm, although the valve will deliver 40 gpm when fully open. Time of the first reset term $R_1$ is fixed at 3.0 sec by the scaling. Estimate the setting of the second reset $R_2$ required for ¼-amplitude damping, and the resulting period of the loop.

6.3 The flow of liquid leaving a storage tank is to be controlled at a fixed rate as long as the level in the tank is within certain bounds. If the level should fall to a specified low point or rise to a designated high value, however, level ought to be controlled. Design a system incorporating all of these features.

6.4 Table 4.2 gives three combinations of settings for an interacting controller producing +36, 0, and -36" of phase shift; reset and derivative are equal in each case. Assume now that the settings are constant and that three different values of dead time are represented. If the dead time corresponding to zero controller phase is 1.0 min, what are the other values? What is the natural period in each case? With a constant proportional band, what is the loop gain at the three values of dead time?

6.5 Apply the same three values of dead time to the optimally adjusted noninteracting controller in Table 4.2. Estimate the period and loop gain for each value of dead time. Why are these results different from those with an interacting controller? What precaution does this suggest for adjusting controllers on processes with variable dead time?

# Multivariable Process Control

CHAPTER 7

Until now, discussion has been confined to control systems with a single manipulated variable. Furthermore, only one controlled variable has been allowed to be independently specified. But any process capable of manufacturing or refining a product cannot do so within a single control loop. In fact each unit operation requires control over at least two variables: product rate and quality.

Whenever two control loops are to be placed in operation on a single process, the question arises as to which valve should be manipulated from which measurement. In some cases the answer will be obvious. But in those where it is not, some basis must be available to permit the correct decision to be made. In every case these variables interact with one another to some extent, which naturally interferes with their individual performance.

The most effective arrangement of control loops cannot be determined without an appreciation for the needs of the process. In this chapter the relative significance of several types of controlled and manipulated variables will be examined, along with methods for pairing them. Then a means for measuring their interaction will be introduced, followed by suggestions for its compensation.

## CHOOSING CONTROLLED VARIABLES

"What do you really want to control" is a question that instrument engineers often ask after scanning the flow sheet of a new process. True, in the majority of operations, the answer will be self-evident, but for many, this is not so. Consider a distillation column, for example. Must distillate composition be controlled to a particular specification, or bottoms composition, or both, or neither? Which variables are "wild"; which ones may be manipulated; which are to be controlled; which are to be maximized or minimized; what are their relative economic values? The arrangement of control loops is based entirely on the answers to these questions, which are only available from someone who is intimately familiar with the process. To be sure, an experienced instrument engineer can, in most cases, skillfully translate the needs of the process into the most effective control-loop arrangement. Some of the factors to be discussed below are steps that he may unconsciously take in this somewhat intuitive procedure.

### Degrees of Freedom

There are as many degrees of freedom as there are manipulated variables. For every controlled variable there must be at least one manipulated variable. The following corollaries apply:

1. If ever the number of variables desired to be controlled exceeds the number available to be manipulated, the latter must be shared among the former on a logical basis. (Several of these situations were described in the previous chapter under Selective Control Loops.)

2. Whenever the number of manipulated variables exceeds those to be controlled, the excess must be fixed (or alternatively may be programmed to satisfy some economic criterion).

**example 7.1**

As an example of the first case cited, consider a plant producing a material A dissolved to a concentration $x$ in solvent *B*. A blending system, shown in Fig. 7.1, is to adjust the final concentration of the product stream to a value $y$ by adding full-strength product *A* from stream $F_2$, or fresh solvent *B* as necessary. The flow $F_1$ and concentration $x$ of the product influent are uncontrolled.

If final concentration $y$ were the only controlled variable, only one of the valves would be open at any given time. But a second specification calls

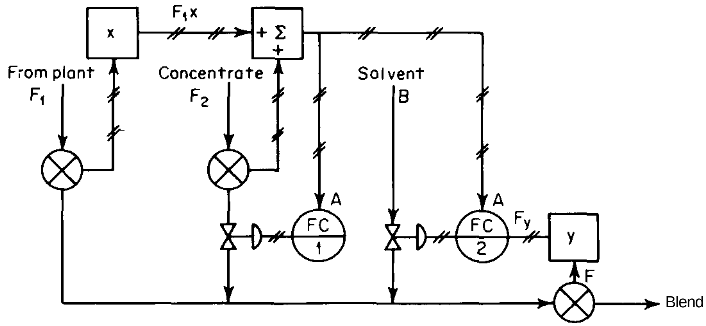



***FIG 7.1.** In this blending system, product entering at the left is to be adjusted in concentration.*

for control of total product flow $A$ leaving the system. To control both variables, both valves must be open at the same time.

The total product flow is set into the controller manipulating stream $A$. The incoming concentration $x$ is multiplied by the flow of the plant effluent $F_1$, which when added to stream $F_2$, gives a measure of total product in the blend. This same signal serves as the measurement for flow controller 1, and as a set point for flow controller 2. Blend concentration y is set into the ratio station at the right, manipulating the flow of solvent $B$. If the flow of the blend is $F$,

$$Fy = A = F_1x + F_2$$

In normal operation, the plant production may be less than the total product flow set into flow controller 1. This controller would then add enough concentrate to make up the difference. Flow controller 2 will add enough solvent to increase the flow of blend to satisfy the equation above, thereby controlling y.

Should plant production temporarily esceed the set value of total product flow, the concentrate stream will be shut off. When this happens, one degree of freedom is lost because one variable can no longer be manipulated. Thus control of total flow is also lost. But control of composition can be retained so long as the solvent valve is throttling.

A similar situation would occur if plant production were temporarily curtailed, causing the concentrate valve to open wide.

Notice that the loops have been arranged to maintain control of blend concentration rather than total flow in the event that concentrate flow can no longer be manipulated. They could be arranged to handle total flow control (in this eventuality) if that, were desired.

**example 7.2**

As an illustration of the second case, consider a direct-fired heater where fuel flow *is* manipulated to control temperature. If both oil and gas are available as fuel, some program must be decided upon for their use, since only one manipulated variable is required for control. The following situations may typically exist:

1. The availability of gas is limited, thus oil flow must be manipulated in the event of gas shortage.
2. The cost of gas varies seasonally, occasionally exceeding that of oil.
3. The cost of gas varies with usage such that, beyond a certain flow, oil is more economical to use.

## Classification of Variables

All ***manipulated*** variables which are direct inputs to the process are flows or functions of flow. (Cascade control loops are excluded from the above, with the exception of flow loops, because the output of the primary controller is not a process input.) The process cannot be manipulated without changing the position of valves or dampers or the speed of pumps or compressors, all of which affect flow. The relationships between flow and controller output vary greatly, available pressure drop being a significant factor. The fact remains that in order to bring about control of any variable, in any process, flow must be made to change. From this definition, flow may be considered both a controlled and a manipulated variable-this coincidence was discussed in Chap. 3.

The *load* is a combination of uncontrolled variables which places a particular demand on the control system. Like a manipulated variable, the load is flow or a function thereof. In a mass transfer operation, the load is ***mass flow.*** In a level-control loop, for example, where one flow is manipulated, the load is the algebraic sum of all the other flows entering the vessel. The manipulated variable must be matched to the load in order to maintain constant level. In an energy transfer operation, it is the flow of ***energy.*** To control temperature of a room, for example, as much heat must be added as is lost.. Heat is lost through doors, windows, etc., and varies with inside and outside temperatures and wind velocity. The sum of these losses (flows) is the load on the heating system.

A plant is customarily designed to manufacture a certain product at specified values of rate and quality. The current rate of production is the load placed on the plant. Production rate must be specified in some way: in a chemical plant it is usually set by the rate of the principal reaction; in a refinery it is the flow of crude feed stock; in a utility plant it is determined by the gross needs of the consumers.

Most plants are designed to operate with greatest efficiency at a particular rate. Nevertheless they should not be bound to these operating conditions, because ultimately the average production rate is determined by consumer demand. There is a great deal to be gained by increasing the throughput of a plant whose products are in demand. Therefore its control systems ought never to interfere with this aspiration.

In the particular unit of the plant where the production rate is set, load changes are, for the most part, nonexistent. But because of variations in the efficiency of this unit and in its control loops, the flow and quality of material leaving it cannot be expected to be uniform. These variations are imposed as load changes upon the next unit, etc., through the entire plant. Thus only one unit of the plant can be expected to operate at constant load, and its output sets the load for the remainder of the plant. This unit is said to be "base-loaded," while the others are "load-following." Some utilities are load-following throughout, while others are base-loaded throughout.

As the product is processed in one unit after another, its flow may become more variable. Thus the terminal stage of refinement encounters flow variations imposed by all the other units in the plant. Control systems on the individual units can do little to smooth these variations in flow, but they should be able to deliver material of consistent quality. Flow variations can only be absorbed by surge vessels which are suitably located and efficiently used.

Plants whose load is established in the early stages of processing have a certain advantage. Load changes travel in the same direction as the flow of product, so they can be anticipated to some extent. The load of utilities is generally determined at the product end. The load wave in this case travels counter to the flow of product. The two situations are compared in Fig. 7.2.

*Product quality* is the prime controlled variable in every unit of the plant. The term "quality" is here subject to broad interpretation. In a drying operation, it is the moisture content of the product. In a heat exchanger, it is the temperature or enthalpy of the exit stream. In a distillation tower, it is the purity of the distillate and / or bottoms product. In the boiler above, it is the pressure of the saturated steam; in the superheater, it is steam temperature. Each unit must be controlled to deliver whatever rate of product is required at a consistent quality if the next unit is to fulfill its function.

As the product becomes more valuable (by further processing), quality control becomes more important. Quality and value are strongly

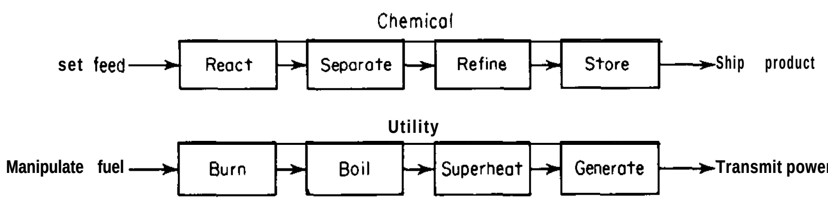



***FIG 7.2. The load is set at the input of a chemical plant and at the output of a typical utility.***

related, although the exact nature of their relation is difficult to define in actual cases. But the existence of an interdependence suggests two areas where control can play a major role in plant economy:

1. Poor quality control can devalue a product. To compensate, processes are often operated at lower rates or with considerable recycle. An intelligent control system, capable of maintaining uniform quality, may pay for itself by permitting increased production and/or lower costs.

2. Rigid quality control may cost more than it is worth under certain conditions of operation. It is possible to program a control system to operate a unit of the plant or the whole plant at minimum cost if these economic relationships are well known.

This brings us to the next class of controlled variables: those which are economic in nature. Included are such items as efficiency, yield, conversion, recovery, loss, profit-terms that identify how well the plant is being operated. For the most part, these variables do not appear in a closed loop. Yet it is the intent of management to observe and respect them insofar as possible. Many unit operations have no economic variables, but many do. The problem is in being able to measure the variable, place a value on it, and find out how to maximize (or minimize) it.

If the variable can be measured and an appropriate manipulated variable found, feedback optimizing control can be applied. Several systems for doing this were described in the preceding chapter. In most cases, however, the economic variables cannot be directly measured. Usually they embody a combination of factors from which a computation might possibly be made. In any event, operating personnel seldom go beyond displaying economic variables, with the exercise of discriminating manual control. But they will unquestionably be given more attention in future plants.

*Inventory* variables play a servile but essential role in process control. They are necessary to close overall material balances and, in some cases, energy balances. The measurements principally associated with inventory control are liquid level, weight, and pressure. Equilibrium is impossible without their control. Nearly every unit operation has one of these loops, and to see two or three is not uncommon.

### True and Inferential Variables

Every effort should be made to use true measurements of important process variables whenever possible. In those instances when a true measurement is not available, all the sources of error in whatever inferential measurements are practicable should be examined.

The variable most commonly inferred is composition, because of the lack of reliable, economical analyzers for a wide spectrum of chemical systems. Inferential measurements are not specific to certain substances,

TABLE 7.1 Common Inferential Measurements

| Variable | Measurement | Sources of error |
|---|---|---|
| Composition. | Liquid density | Temperature |
| | Gas density | Temperature, pressure |
| | Boiling point | Pressure |
| | Electrolytic conductivity | Temperature |
| | Viscosity | Temperature |
| | Dielectric constant | Temperature |
| | Dew point | Pressure |
| Mass flow.. | Velocity meters | Density, viscosity |
| | Differential meters | Density, viscosity |
| Weight. | Liquid level | Density |
| Liquid level. | Differential pressure | Density |

however, so they are generally limited to binary systems. The introduction of a third component, then, is a common source of error. Most inferential measurements are also temperature-sensitive. Temperature, when used as an inferential measurement, is often pressure-sensitive. Table 7.1 is a list of common inferential measurements and their most significant sources of error.

Temperature conpensation is available on many measurements and should be used where temperature control. is impossible. Where neither can be used, temperature correction can be applied through analog computation. In a typical liquid mixture, density $\rho$ may vary with both composition x and temperature $T$ in the manner

$$p = ax - bT$$

where $a$ and $b$ are constants characteristic of the substances involved. Composition can then be calculated from measurements of both density and temperature by on-line computation:

$$x = \frac{\rho + bT}{a}$$

A measurement of temperature is often used to infer composition at a particular tray in a distillation column. A slight change in absolute pressure at that point can void the inference, however. This is a particularly severe problem in vacuum towers where changes of a few inches of water may represent large absolute variations. There are three possible solutions to this problem:

1. Control pressure at the point of measurement.
2. Use a compensated measurement, such as a differential vapor-pressure transmitter.[1]
3. Apply pressure correction (similar to the temperature correction just cited for density).

On-line calculations of mass flow,[2] weight; liquid level,[3] and composition from measurements of differential pressure, density, pressure, and temperature are being made in many industries today. In addition, calculations are being made of the mass flow of specific substances in mixtures of gases, solutions, and slurries. Analog computation is also being used to determine the flow of heat in boilers, cooling systems, and reactors. The possibilities are virtually unlimited.

Some of the new signals mentioned above permit many economic variables to be generated. For example, the conversion in a reactor can be assimilated from a heat balance across it. Yield can be calculated as the ratio of feed and product composition. The efficiency of a steam plant is similarly the ratio of thermal power to the heat content of the flowing fuel. The cost of operating a separation unit can be determined from the mass flow rates of utilities and products and the measurements of product quality.

## PAIRING CONTROLLED AND MANIPULATED VARIABLES

In some instances, the correct pairing of controlled and manipulated variables is obvious. Occasionally it does not matter how they are paired. Cited in Fig. 7.3 are examples of each extreme. In the separator, vapor flow does not affect liquid level, nor does liquid flow influence pressure, so the arrangement of loops is obvious. In the pipeline, however, both valves appear to affect pressure and flow equally, so either combination will work. But many times a control engineer will be faced with a situation where a decision must be made that is not obvious, and

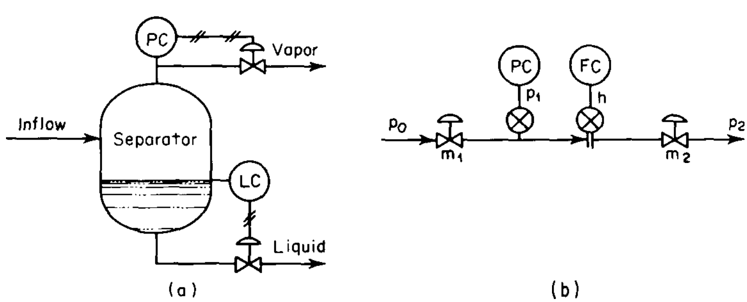



*FIG 7.3. Examples of (a) negligible interaction and (b) equal interaction.*

he should have some firm basis for making one. The following procedure has been developed for establishing this basis.[4]

### Relative Process Gain

Picture any process as consisting of a block with a number of input (manipulated) variables and an equal number of output (controlled) variables. The object is to control a given process output by manipulating the one input that will have the greatest influence on it. If this is not done, another input will have more influence on the controlled variable than the one which it manipulates through the controller. To assess all the possibilities, the gain of each controlled variable to each manipulated variable must be determined. It is extremely desirable that these gain terms be normalized to eliminate dimensions and to place them all on the same basis.

Actually, two different open-loop gains can be found for a pair of variables $c_i$ and $m_j$. The gain $dc_i/dm_j$ with all loops open may differ from that with all the rest of the loops closed. A convenient measure of relative loop gain is found to be the ratio of the process gains for these two conditions. To state it a little differently, the relative gain is defined as the ratio of the open-loop gain in terms of $m$ (i.e., with all other $m$'s constant) to the gain in terms of c (i.e., with all other c's constant). The term $\lambda_{ij}$ will be used to designate the dimensionless change in $c_i$ with respect to a change in $m_j$:

$$\lambda_{ij} = \frac{\partial c_i/\partial m_j \Big|_m}{\partial c_i/\partial m_j \Big|_c} \tag{7.1}$$

It is convenient to arrange a table of these gains in the form of a matrix such as that shown below:

| | $m_1$ | $m_2$ | $\cdots$ | $m_j$ |
|---|---|---|---|---|
| $c_a$ | $\lambda_{a1}$ | $\lambda_{a2}$ | $\cdots$ | $\lambda_{aj}$ |
| $c_b$ | $\lambda_{b1}$ | $\lambda_{b2}$ | $\cdots$ | $\lambda_{bj}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | | |
| $c_i$ | $\lambda_{i1}$ | $\lambda_{i2}$ | | $\lambda_{ij}$ |

This greatly facilitates comparison of the different combinations. The combinations with the largest positive numbers should be selected for closed loops, for reasons that will soon be explained.

### example 7.3

To see how the matrix is derived, the pressure-flow process of Fig. 7.3 will be examined, using the following simplified equations to express the flowing differential $h$ and the controlled pressure $p_1$:

$$h = m_1(p_0 - p_1) = m_2(p_1 - p_2) = \frac{m_1 m_2(p_0 - p_2)}{m_1 + m_2} \tag{7.2}$$

The gain with both loops open is the partial derivative in terms of m:

$$\left.\frac{\partial h}{\partial m_1}\right|_{m_2} = \frac{m_2(p_0 - p_2)}{m_1 + m_2} - \frac{m_1 m_2(p_0 - p_2)}{(m_1 + m_2)^2} = (p_0 - p_2)\left(\frac{m_2}{m_1 + m_2}\right)^2 \tag{7.3}$$

The gain with the pressure loop closed is in terms of $p_1$:

$$\left.\frac{\partial h}{\partial m_1}\right|_{p} = p_0 - p_1 = (p_0 - p_2)\left(\frac{m_2}{m_1 + m_2}\right) \tag{7.4}$$

Then,

$$\lambda_{h1} = \frac{m_2}{m_1 + m_2}$$

Gain may also be expressed in terms of pressure, by substituting for $m_1$ and $m_2$ from Eq. (7.2):

$$\lambda_{h1} = \frac{p_0 - p_1}{p_0 - p_2} \tag{7.5}$$

In the same manner, $\lambda_{h2}$ can be found:

$$\lambda_{h2} = \frac{p_1 - p_2}{p_0 - p_2} \tag{7.6}$$

Using the following expressions for $p_1$, the other gains can be determined:

$$p_1 = p_0 - \frac{h}{m_1} = p_2 + \frac{h}{m_2} = \frac{m_1 p_0 + m_2 p_2}{m_1 + m_2}$$

The resulting matrix is:

$$\begin{array}{c|cc} & m_1 & m_2 \\ \hline h & \dfrac{p_0 - p_1}{p_0 - p_2} & \dfrac{p_1 - p_2}{p_0 - p_2} \\ p_1 & \dfrac{p_1 - p_2}{p_0 - p_2} & \dfrac{p_0 - p_1}{p_0 - p_2} \end{array} \tag{7.7}$$

Choice of particular c-m pairs for this application depends on the pressure distribution. The matrix indicates that the valve with the greater pressure drop is better for flow control. But if $p_1$ is midway between $p_0$ and $p_2$, all elements in the matrix are 0.5, so it does not matter which pairs are chosen.

The sum of each row and each column is unity-this is one of the features of the relative-gain matrix. In a 2 by 2 matrix such as (7.7), it is only necessary to solve for one element, the others being equal or complementary.

**example 7.4**

Another common two-loop process is the blending system shown in Fig. 7.4. Two streams, X and Y, are blended to a specified total flow $F$ of composition $x$. Let the flow of stream X be designated $m_1$ and the flow of stream Y be $m_2$. The following equations describe the process:

$$F = m_1 + m_2 = \frac{m_1}{x} = \frac{m_2}{1 - x} \tag{7.8}$$

$$x = \frac{m_1}{m_1 + m_2} = \frac{m_1}{F} = 1 - \frac{m_2}{F} \tag{7.9}$$

Only one element of the matrix need be found:

$$\left.\frac{\partial F}{\partial m_1}\right|_{m_2} = 1 \qquad \left.\frac{\partial F}{\partial m_1}\right|_{x} = \frac{1}{x} \tag{7.10}$$

$$\lambda_{F1} = \frac{1}{1/x} = x \tag{7.11}$$

The matrix then takes on the appearance:

$$\begin{array}{c|cc} & m_1 & m_2 \\ \hline F & x & 1 - x \\ x & 1 - x & x \end{array} \tag{7.12}$$

For values of x less than 0.5, $m_1$ should be used for composition control and $m_2$ for flow control.

There is no limit to the number of variables that can be displayed by a matrix, although some difficulties may be encountered in making the differentiations. To this extent, an analog computer would be helpful. Or a numerical solution may be found by incrementing each manipulated variable, although general solutions like those above have more value.

Instead of using mathematical models, as was done here, an existing plant may be tested to generate a matrix of gains. This requires a series

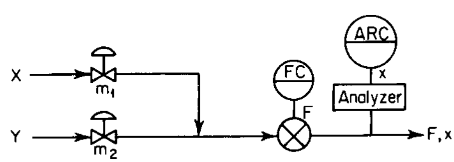



***FIG 7.4. A typical two-component blending system in which both total flow and composition are to be controlled.***

of open-loop tests, in which $m_1$ is first incremented and changes in all the controlled variables observed. Then open-loop gains $(\Delta c_a/\Delta m_1)_m$, and so on, can be calculated. The other manipulated variables are then incremented and their effects calculated in the same manner. Fortunately it is not necessary to make closed-loop tests to determine $(\Delta c_a/\Delta m_1)_c$, since enough data have already been gathered. Matrix manipulation can do the rest.

Let the values of $(\Delta c_i/\Delta m_j)_m$ be arrayed in a matrix identified as M. A complementary matrix C, of elements $(\Delta m_j/\Delta c_i)_c$, can be generated by inverting, then transposing matrix M:

$$\mathbf{C} = (\mathbf{M}^{-1})^T \tag{7.13}$$

Rules for performing this operation can be found in reference (5) and in most texts on modern algebra. Finally, each relative-gain term $\lambda_{ij}$ is found by *multiplying* each element in matrix M by its corresponding element in the complementary matrix C.

It is also possible to prepare a matrix of relative gains from the elements $(\Delta c_i/\Delta m_j)_c$ alone. The procedure is the same as that described by Eq. (7.13), except that the elements of matrix M are inverted, i.e., they appear as $(\Delta m_j/\Delta c_i)_c$.

## Coupling

Whenever a single manipulated variable can significantly affect, two or more controlled variables, the latter are said to be coupled. This interaction between control loops can be troublesome. Some variables are difficult enough to control without being subject to upsets from other loops. For this reason, proposed composition loops should always be examined for evidence of coupling. When the gain matrix contains numbers approaching 1 and 0, the loops will be largely independent. But values approaching 0.5 indicate a strong mutual coupling.

To gain an appreciation of the effects of coupling, consider the pressure-flow system of Fig. 7.3. Let $p_1$ be midway between $p_0$ and $p_2$, so that all elements in the gain matrix equal 0.5. The flow controller will manipulate $m_2$. With the pressure controller in manual, the flow controller could be adjusted as if it were alone, i.e., as if there were no coupling. And, in fact, there would be no coupling because only one loop is closed. But under these circumstances, any change in $p_0$, $p_2$, or flow would affect $p_1$.

Next, think of the pressure controller being in automatic, but with a wide band and long reset-i.e., very loose settings. But should the set point of the flow controller be changed, pressure will be upset nearly as much as when its loop was open. The pressure error will eventually be corrected as the controller slowly moves valve $m_1$. This gradual movement of $m_1$ will cause a slight flow change, causing the flow controller to reposition $m_2$. If the rate of change of $m_1$ is slow compared to the reset, time of the flow controller, flow will be maintained quite near the set point throughout. Under these conditions, flow is slightly affected by pressure, while pressure is tightly coupled to flow.

If -the pressure controller were now tightly adjusted with the flow controller in manual, both loops would be on the verge of instability when placed in automatic. An increase in set flow would cause $m_2$ to open, dropping $p_1$. The pressure controller would open $m_1$ to restore $p_1$, with the result that flow would increase twice as much as the flow controller intended it to. In effect, coupling has caused reverberation between the loops by doubling the gain of each. This factor should be evident from the matrix: the total effect on flow (and pressure) by both valves is 1.0, yet each manipulated valve only affects each measurement by 0.5. Thus the flow controller was originally adjusted with only half the process gain in force. If the matrix elements had been 0.2 and 0.8 instead of 0.5 and 0.5, the effect on the closed loops would have been hardly noticeable (unless the wrong pairs had been connected).

The stability problem can be resolved by doubling the proportional band of each controller. Even so, a change in the set point of either controller will upset the other loop, because both valves must move. But if the elements were 0.8 and 0.2, and the wrong pairs of variables were chosen, the situation would be considerably worse. A flow set-point change would cause a fourfold upset in pressure, and vice versa. If the proper combinations were chosen, the upset would be $\frac{1}{16}$ as great. This example should point out how important it is to determine and select the proper pairing of variables for each loop.

Another observation to be made is that the degree of coupling is usually variable. It varied with $p$ in the pressure-flow process and with $x$ in the blending system. In each case it depended on the extensive controlled variable rather than on the intensive variable (flow). It should be recognized that in some processes the best choice of $c$-$m$ pairs at one operating condition may not be the best at another.

### Coupling between Similar Variables

A further degree of coupling is found between variables that are similar in nature. Imagine a three-component, blend in which it is desired to control both density $\rho$ and viscosity $\mu$. A problem arises because a change in either of two components may affect *both* density and viscosity in the *same* direction. This differs from the cases studied earlier, in that while $m_1$ and $m_2$ affected one variable in the same direction, they affected the other in opposite directions. Let the mathematical model of the three-component blending system be as follows:

$$\rho = \frac{am_1 + bm_2}{F} \tag{7.14}$$

$$\mu = \frac{fm_1 + gm_2}{F} \tag{7.15}$$

Let the coefficients *a, b, f,* and g be positive numbers, and total flow *F* be uncontrolled, so that it does not enter into the matrix.

The normalization procedure yields the following matrix:

$$\begin{array}{c|cc} & m_1 & m_2 \\ \hline \rho & \dfrac{ag}{ag - bf} & \dfrac{-bf}{ag - bf} \\ \mu & \dfrac{-bf}{ag - bf} & \dfrac{ag}{ag - bf} \end{array} \tag{7.16}$$

Since all coefficients are positive, two of the terms in the relative-gain matrix must be negative-which two depending on whether $ag > bf$. To allow inspection of the properties of this process, let $a = b = f = 0.5$ and $g = 1.0$; then the matrix appears as:

| | $m_1$ | $m_2$ |
|---|---|---|
| $\rho$ | 2 | -1 |
| $\mu$ | -1 | 2 |

Pairing must still be carried out in favor of the larger (positive) numbers. In fact the pairing indicated by the negative numbers will not be controllable at all. If $m_2$ were chosen for control of density and $m_1$ for viscosity, the manipulated variables would eventually be driven to opposite limits without satisfying either controller. Regardless of the controller settings, the system is divergent. This is always an indication of positive feedback.

If the correct pairing is chosen, both loops will be stable and in fact will require double the controller gain needed if there were no coupling. Thus it can be seen that even in this case, the numbers in the matrix indicate the effect of coupling on controller settings.

The coupling in this example is constant, i.e., only constants appear in the matrix, because the mathematical model of the process is linear. Observe how the coefficients in the model fall into place in the relative-gain matrix, corresponding to the transformation procedure involving Eq. (7.13).

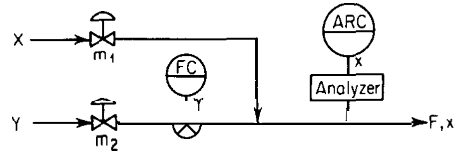



*FIG 7.5. Half-coupling exists between composition x and flow Y.*

## Half-coupling

It is possible to arrange a 2 by 2 process such that one variable will be affected by one valve while the other is influenced by both. If the flow measurement for the blending system were placed on stream Y as in Fig. 7.5, this effect would be achieved.

Both valves now affect composition, but $m_1$ has no influence over flow Y:

$$Y = m_2 \qquad \frac{dY}{dm_1} = 0 \tag{7.17}$$

The first element in the matrix, $\lambda_{Y1}$, must be zero. The remaining elements are:

| | $m_1$ | $m_2$ |
|---|---|---|
| Y | 0 | 1 |
| 2 | 1 | 0 |

(7.18)

The matrix indicates that coupling is not a problem in this process. And indeed this is the case. First, there is no question about the pairing of variables, as there is with fully coupled loops. Second, although $m_2$ influences composition, the effect is more that of a load change. Finally, since the composition controller cannot cause $m_2$ to change by altering the input to the flow controller, its adjustment depends only on the gain through $m_1$. Notice that the blending system in Fig. 7.1 is half-coupled.

## Dynamic Effects

In the pressure-flow process, the dynamic response of both measurements is similar. This is not often the case. Flow, pressure, and liquid level ordinarily respond rapidly to valve position, while temperature and composition do not. Therefore processes which couple flow and composition, level and composition, pressure and temperature, etc., are deserving of deeper study.

Actually, the coupling of a fast and a slow loop was simulated in the earlier example when the pressure controller was given loose settings while the flow controller was tuned tightly. Imagine the settings of the pressure controller being more like those of a composition controller.

The fast loop is scarcely affected by the coupling, while the slower loop is upset by the fast one.

The blending system of Fig. 7.4 is a good example of the coupling of fast and slow loops. The composition loop would normally oscillate at a period of a few minutes, while the period of the flow loop would be a few seconds. Tuning of the flow controller would be the same as if the loops were entirely independent, because the composition controller would not be able to cause rapid changes in flow. But any increase in $m_1$ will increase total flow, thereby automatically decreasing $m_2$. The composition controller actually manipulates both valves-one directly, the other indirectly through the flow controller. Consequently it must be adjusted to accommodate the gain of both valves.

An appreciation for the dynamic effects that coupled closed loops have on one another can be gained by analyzing the block diagram shown in Fig. 7.6. One loop, comprised of a controller whose gain vector is $\mathbf{g}_{c1}$ and a process with a dynamic vector of $\mathbf{g}_1$ and relative gain $\lambda_{11}$, has a period of $\tau_{o1}$. The loop is upset by a manipulated variable $m_2$ from the second closed loop, whose period is $\tau_{o2}$. In the path of $m_2$ is the process vector $\mathbf{g}_2$ and the relative gain of $c_1$ with respect to $m_2$, that is, $\lambda_{12}$.

Although the solution of the block diagram for both closed loops is unwieldy and difficult to present, a qualitative appraisal of the dynamic effect can be gained from the response of $c_1$ with respect to $m_2$.

$$m_1 = \mathbf{g}_{c1}(r_1 - c_1) \qquad c_1 = m_1\mathbf{g}_1\lambda_{11} + m_2\mathbf{g}_2\lambda_{12} \tag{7.19}$$

$$c_1(1 + \mathbf{g}_{c1}\mathbf{g}_1\lambda_{11}) = r_1\mathbf{g}_{c1}\mathbf{g}_1\lambda_{11} + m_2\mathbf{g}_2\lambda_{12}$$

Differentiation of $c_1$ with respect to $m_2$ yields

$$\frac{dc_1}{dm_2} = \frac{\mathbf{g}_2\lambda_{12}}{1 + \mathbf{g}_{c1}\mathbf{g}_1\lambda_{11}} \tag{7.20}$$

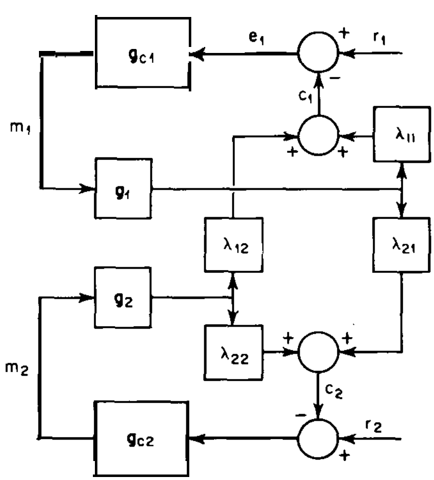



*FIG 7.6. Each closed loop is upset by the output of the other.*

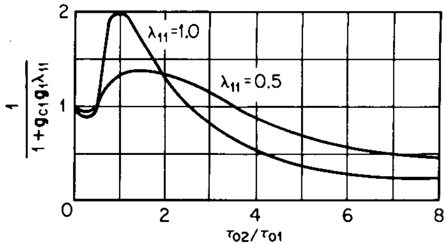



***FIG 7.7. The effect of $m_2$ on $c_1$ depends on the ratio of their periods.***

Equation (7.20) can be solved by factoring into two parts, $g_2\lambda_{12}$ and $1/(1 + g_{c1}g_1\lambda_{11})$. The latter is a characteristic of the closed loop and can be evaluated by summing the open-loop vector $g_{c1}g_1\lambda_{11}$ with 1.0, followed by inversion. The solution for the general case of a dead-time plus integrating combination adjusted to $g_{c1}g_1 = 0.5$ at the natural period is plotted in Fig. 7.7.

When $\tau_{o2}$ is less than $0.5\tau_{o1}$, dynamic coupling is virtually unity. This means that any change in the output of controller 1 will affect both $m_1$ and $m_2$, such that its gain must be reduced by an amount corresponding to $\lambda_{12}$. When there are only two loops, the final gain becomes $\lambda_{11}$ times the original.

Values of $\tau_{o2}$ between 0.5 and $2\tau_{o1}$ represent, a region of severe dynamic coupling. If $\tau_{o2}$ falls in this area, the gain of controller 1 must be reduced by at least $\lambda_{12}$, which reduces the magnitude of the coupling. Figure 7.7 shows two curves, one where $\lambda_{11}$ is 1.0, i.e., the controller is adjusted for conditions of no coupling. A second curve is plotted for $\lambda_{11} = 0.5$, which corresponds to a reduction in controller gain by 50 percent, allowing for the maximum coupling in a two-loop system. This makes the loop less sensitive to disturbances occurring in the resonant region, but more sensitive to disturbances at longer periods.

Where $\tau_{o2} >> \tau_{o1}$, coupling approaches zero. This is the case of the flow loop being scarcely disturbed by the composition loop. In this region higher controller gain is actually favorable, reducing the effect of dynamic coupling.

### example 7.5

Consider a process wherein two pairs of variables interact such that $\lambda_{11} = 0.7$ and $\lambda_{12} = 0.3$. Yet $c_1$ responds to $m_1$ with a dead time of 10 min, but to $m_2$ with a dead time of only 1 min. In this case, $\tau_{o1}$ would be 40 min and $\tau_{o2}$, 4 min. If $c_1$ is paired with $m_1$, it will be disturbed by $m_2$ by the amount 0.3, that is, $\lambda_{12}$, because the dynamic coupling factor is 1.0. But if $c_1$ were paired with $m_2$, it would be disturbed by $m_1$ by only about 0.15, because the ratio of the periods is 10. But this arrangement would allow a higher controller gain for $c_1$, reducing its sensitivity to disturbances of other periods.

Unfortunately, however, the pairing of variables also affects the other loop, which must be considered. So some judgment is necessary to determine which of the variables is more deserving of precise control.

## DECOUPLING CONTROL SYSTEMS

Controlling a single-variable process is comparatively easy even if the dynamics in the loop are unfavorable. There is only one way to close the loop. When a second pair of variables appears, however, the picture is entirely different-not only must a choice be made between pairs used for control, but coupling can exist. And if there is coupling, the ease of control that was found with independent loops disappears. This facility can be restored, however, by decoupling the variables through a computing system. Just as a single valve can affect several controlled variables through the natural coupling of the process, a single controller can be made to adjust several valves through a decoupling system.

Consider the problem that an engineer is faced with when he starts out to design a system to control a complex process like a distillation column. He may have five pairs of variables to connect:

| | |
|---|---|
| Throughput | Feed flow |
| Distillate composition | Distillate flow |
| Bottoms composition | Reflux flow |
| Bottoms level | Steam flow |
| Accumulator level | Bottoms flow |

There are factorial five or 120 different ways to connect these variables. If coupling exists, which is often the case, even the best way of arranging these loops may of itself be inadequate.

Yet it is possible to design one control system which will surpass all others in performance by completely decoupling its process. This is the mirror image of the steady-state process model. It is unique, and as such, it can never be found through trial and error or by accident. The balance of this chapter will be devoted to its pursuit. The balance of the book will demonstrate its application to the more commonly encountered multivariable processes.

### Reversing the Process Model

In the process, the values of all the controlled variables are related to all the manipulated variables through a series of equations known as the ***process model.*** An ideal control system would be one which correctly positioned all the manipulated variables so as to satisfy all the set points. In this sense, the control system could have the same mathematical structure as the process, but would solve for manipulated variables. Since each process has its own individual mathematical model, there exists only one control system which can solve the same equations. The procedure for arriving at this system is not evolutionary or intuitive-it is as exact and as well defined as the process model.

Assuming that the process model is well defined, a computing system can be designed which will solve these same equations for the manipulated variables. The computing system will ordinarily be made up of multipliers and dividers as well as linear operators, because these functions exist in the process. (Decoupling systems cannot be expected to be linear when their processes are not.)

The computing system must be fed input data if it is to generate any output. Three items of information are available from each loop: set point, measurement, and controller output. (The error signal cannot be used in the steady-state model because it has no steady-state value.) Of this information, the set points are most useful because they represent the exact demands on the control system and are free from feedback transients, lying outside the loops.

But if a measurement is used to decouple fully coupled loops, a positive feedback path is formed through the process, cancelling the effect of control action. As a result, the system has no direction and the controlled variable tends to float. Measurements may be used in systems with half-coupled loops, however, because there is no feedback from one loop to another through the process.

The output of a feedback controller is an unknown variable. If it were known, or predictable with sufficient accuracy, the controller would not be needed. But the output eventually finds its correct level, and in so doing, solves for all the unknowns in the process. Thus the controller output contains more information regarding the manipulated variables than either measurements or set points. The set point will not indicate the presence of a load change, for example. So for the most part, controller outputs will be combined to perform decoupling. The general configuration of a coupled process and decoupling control system would appear as shown in Fig. 7.8.

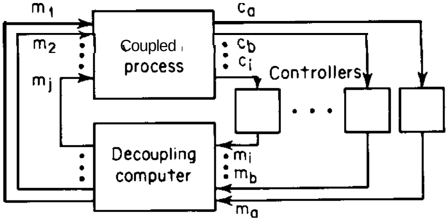



***FIG 7.8. A decoupling system correctly matched to a coupled process can produce essentially independent control loops.***

Consider how a decoupling control system might be designed for the blending process of Example 7.4. The mathematical model is solved for $m_1$ and $m_2$ in terms of $F$ and x:

$$m_1 = Fx \qquad m_2 = F - m_1 = F(1 - x) \tag{7.21}$$

In the first-expression, if both $F$ and x were set points, the manipulation of $m_1$ would be open-loop. Again, if extreme accuracy could be realized, this would be sufficient. But for the usual case these terms ought to be controller outputs. In this example, then, let $m_x$, the output of the composition controller, and $m_F$, the output of the flow controller, take the place of $F$ and x in the model. Then

$$m_1 = m_F m_x \tag{7.22}$$

$m_2$ is the product of the flow-controller output with the complement of the composition-controller output:

$$m_2 = m_F(1 - m_x) \tag{7.23}$$

Figure 7.9 illustrates both the coupling of the process and the decoupling of the control system. Notice the similarity to the system shown in Fig. 6.10.

To appreciate how the system works, envision the process responding from ***a controlled*** condition. If any change occurs in $m_F$, $m_1$ and $m_2$ will both change proportionately and in the same ratio that is already maintaining composition control. Should $m_x$ be required to change, $m_1$ will move in one direction and $m_2$ in the other, to avoid upsetting total flow. Each loop now responds independently of the other. Notice that a change in set point or measurement of either controller will move both valves in the appropriate direction. If set points had been used as decoupling signals, only one valve would be moved when a load change occurred in its loop, which would upset the other loop. Also, decoupling would be defeated if either controller were placed in manual.

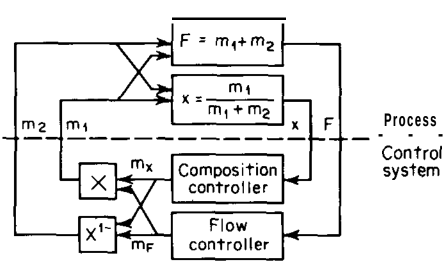



***FIG 7.9. The decoupling control system should be the image of the process.***

Suppose the measurements of x and $F$ were to be used to decouple the controller outputs, such that

$$m_1 = Pm, \qquad m_2 = (1 - x)m_F$$

On startup, $F$ would be zero, causing $m_1$ to be zero: thus only $m_2$ would be manipulated by the flow controller. As x increased toward its set point, $m_2$ would be decreased by the decoupling signal. This action is clearly in the wrong direction, since $m_1$ and $m_2$ should move in opposition to maintain constant flow. The decoupling signal has formed a positive feedback loop.

### Partial Decoupling

In many applications, such as this blending system, complete decoupling is unnecessary. As pointed out in the discussion of dynamic effects, very fast loops are scarcely disturbed by their coupling to slow loops. Therefore that part of the decoupling system designed to liberate flow, pressure, and level loops can ordinarily be omitted.

In Fig. 7.9, the multiplier whose output is $m_2$ can be eliminated with little detriment. Every change in $m_x$ will now upset total flow, but its rate of change is severely restricted by the dynamics of the composition loop. In effect, the composition controller manipulates both valves already-one through the action of the flow. feedback loop. The flow controller manipulates $m_2$ directly, and $m_1$ through the remaining multiplier.

### Decoupling Half-coupled Loops

Half-coupled loops such as the one shown in Fig. 7.5 are simple to decouple, and little risk is involved. The decoupling is in one direction only, and there is no possibility of a positive feedback loop. Therefore a measurement of the independent controlled variable can be used to decouple the output of the dependent loop.

From the model of the half-coupled blending system,

$$x = \frac{m_1}{m_1 + Y} \qquad m_1 = Y\frac{x}{1 - x} \tag{7.24}$$

Since the output of the composition controller need not equal x, but only be a reasonably linear function thereof, it is sufficient that

$$m_1 = Ym, \tag{7.25}$$

The resulting system is diagrammed in Fig. 7.10.

In the foregoing examples, the manipulated variables were identified as valve position. But in each case, a linear characteristic was assumed, with a constant pressure drop for the blending system. If a control

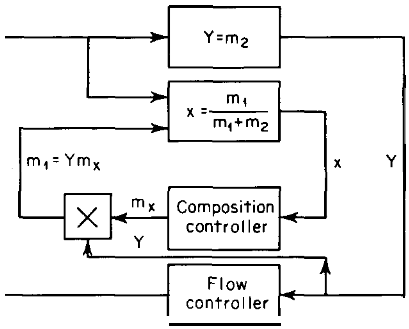



*FIG 7.20. The half-coupled system may be decoupled with the independent measurement.*

system warrants the added refinements of computing elements, however, some consideration ought to be given to improving accuracy elsewhere. Valves do not make very good flowmeters, because the assumptions listed above are seldom realized. Therefore the addition of a flow loop around valve $m_1$ in the blending system is of considerable worth in providing:

1. Immunity from variable pressure drop and valve-characteristic problems
2. Ability to hold an accurate material balance across the process

Incorporation of the flow loop around $m_1$ turns the blending system into the familiar ratio control system with automatic ratio adjustment which was illustrated in Fig. 6.9. It is best to use linear flow signals if all the benefits of decoupling are to be realized.

## SUMMARY

In this chapter considerable emphasis has been placed on the significance of the steady-state mathematical representation of the process. It was pointed out earlier that controllers cannot be made to do a respectable job of tuning themselves. Furthermore, they are scarcely able to arrange their own loops and still less capable of devising their own decoupling systems. As the problem of designing successful control systems has been probed more deeply, the dependency of that design on the nature of the process should have become evident. Indeed it will become more evident in the following pages.

All of this naturally leads to a conclusion whose substance is largely ignored by the theoretician and taken for granted by the practitioner: Until a process can be defined, its control system cannot be defined. The former is the more difficult task.

## REFERENCES

1. The Foxboro Company: Differential Vapor Pressure Cell Transmitter, Type 13VA, *Technical* Information Sheet 37-91a.

2. Shinskey, F. G.: Analog Computing Control for On-line Applications, *Control Eng.,* November, 1962.
3. Roos, N. H.: Level Measurement in Pressurized Vessels, *ISA Journal,* May, 1963.
4. Bristol, E. H.: On a New Measure of Interaction for Multivariable Process Control, Trans. *IEEE,* January, 1966.
5. Dwyer, P. S.: "Linear Computation," chap. 13, John Wiley & Sons, Inc., New York, 1951.

## PROBLEMS

7.1 Classify the five controlled variables appearing at the beginning of the section "Decoupling Control Systems." If the composition of only one product needed to be controlled, what could be done with the extra manipulated variable?

7.2 Two liquids are mixed to a controlled density and total flow. Construct a relative-gain matrix for the system using $m_1$ and $m_2$ to represent the manipulated flows of streams whose densities are $\rho_1$ and $\rho_2$; let $F$ be total flow and $\rho$ the density of the blend. Assume that the volumes are additive.

7.3 The density loop in the previous example oscillates with a 1-min period, while that of the flow loop is 4 sec. Design a decoupling system for the process.

7.4 In a given distillation column, a 1 percent increase in distillate flow $D$ causes distillate composition $y$ to decrease by 0.8 percent, and bottoms composition $x$ to decrease by 1.1 percent. Under the same conditions, a 1 percent increase in steam flow Q causes y to increase by 0.3 percent and $x$ to decrease by 0.2 percent. Calculate the relative-gain matrix.

7.5 It is desired to control both the temperature $T$ and the pressure $P$ in a chemical reactor, by manipulating coolant temperature $T_c$ and reagent flow $F$. It seems that $\partial T/\partial T_c$ at constant flow is 1.0, and that $\partial P/\partial T_c$ is 0.4 psi/°F; $\partial T/\partial F$ at constant $T_c$ is 12°F/gpm, and $\partial P/\partial F$ is 4.8 psi/gpm. Select the best pairs for control loops.

7.6 Prove that a relative-gain matrix may be prepared from inverted closed-loop gains as well as open-loop gains, as described in the paragraph following Eq. (7.13). Illustrate this with the 2 by 2 matrix given in Eq. (7.16).

# Feedforward Control

CHAPTER 8

It has been shown that the nature of the process largely determines how well it can be controlled: the proportional band, reset and derivative times, and the period of cycling are all functions of the process. Processes which cannot be controlled well because of their difficult nature are very susceptible to disturbances from load or set-point changes. When a difficult process is expected to respond well to either of these disturbances, feedback control may no longer be satisfactory for these reasons:

1. The nature of feedback implies that there must be a measurable error to generate a restoring force, hence perfect control is unobtainable. In the steady state, the controller output will be proportional to the load. When the load changes, the controller output must change. In going from one output to another, a controller must "reset," because in each steady state, proportional and derivative offer no contribution. Consequently the net change in output has been shown to be a function of the integrated error:

$$\Delta m = \frac{100}{PR} \int e\, dt$$

Any combination of wide band and long reset time (characteristic of difficult processes) results in a severe integrated error per unit load change :

$$\frac{\int e\ dt}{A\,m} = \frac{PR}{100}$$

This explains why difficult processes are sensitive to disturbances.

2. The feedback controller does not know what its output should be for any given set of conditions, so it changes its output until measurement and set point are in agreement-it solves the control problem by trial and error, which is characteristic of the oscillatory response of a feedback loop. This is the most primitive method of problem solving.

3. Any feedback loop has a characteristic natural period. · Should disturbances occur at intervals less than about three periods, it is evident that no steady state will ever be reached.

There is a way of solving the control problem directly, and this is called "feedforward control." The principal factors affecting the process are measured and, along with the set point, are used in computing the correct output to meet current conditions. Whenever a disturbance occurs, corrective action starts immediately, to cancel the disturbance before it affects the controlled variable. Feedforward is theoretically capable of perfect control, notwithstanding the difficulty of the process, its performance only being limited by the accuracy of the measurements and computations.

Figure 8.1 is a simplified diagram illustrating the arrangement of the feedforward control system as it has been described. Its essential feature is the forward flow of information. The controlled variable is ***not*** used by the system, because this would constitute feedback; this point is important because it shows how it is possible to control a variable without having a continuous measurement of it available. A set point is essential, however, because any control system needs a "command" to give it direction.

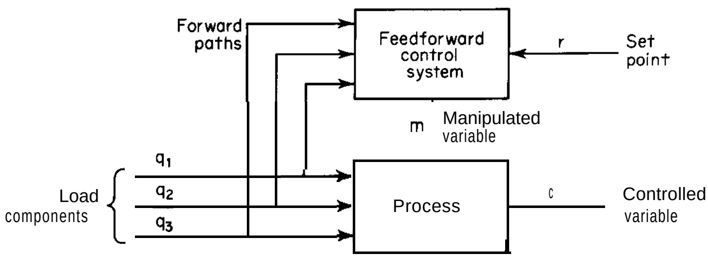



***FIG 8.1. The control system embodies a forward flow of information.***

Although a single controlled variable is indicated in the figure, any number may be accommodated in one feedforward system. Three forward loops are shown, to suggest that all the components of load which significantly affect a controlled variable may be used in solving for the manipulated variable. Although their configuration differs from the commonly recognized feedback loop, these loops are truly closed. Feedforward control should not, therefore, be construed as merely an elaborate form of programmed or open-loop control.

## THE CONTROL SYSTEM AS A MODEL OF THE PROCESS

In practice, the feedforward control system continually balances the material or *energy* delivered to the process against the demands of the load. Consequently the computations made by the control system are material and energy balances on the process, and the manipulated variables must therefore be accurately regulated flow rates. An example is the balancing of firing rate vs. thermal power that is being withdrawn as steam from a boiler. Some material and energy are inevitably stored within the process, the content of which will change in passing from one state to another. This change in storage means a momentary release or absorption of energy or material, which can produce a transient in the controlled variable, unless it is accounted for in the calculations.

To be complete, then, the control computer should be programmed to maintain the process balance in the steady state and also in transient intervals between steady states. It must consist of both steady-state and dynamic components, like the process: it is, in effect, a model of the process. If the steady-state calculations are correct, the controlled variable will be at the set point as long as the load is steady, whatever its current value. If the calculations are in error, an offset will result, which may change with load. If no dynamic calculations are made, or if they are incorrect, the measurement will deviate from the set point while the load is changing, and for some time thereafter, while new energy levels are being established in the process. If both the steady-state and dynamic calculations are perfect, the process will be continually in balance, and no deviation will be measurable at any time. This is the ultimate goal.

The same procedure is followed in the design of a feedforward system as was used for decoupling, i.e., the process model is reversed. The manipulated variables are solved for in terms of load components and controlled variables. In a decoupling system, controller outputs were inserted where the controlled variables appeared in the equations. But for feedforward control, set points are used instead. It is the intent of a feedforward system to force the process to respond as it was designed—to follow the set points as directed without regard to load upsets.

### Systems for Liquid Level and Pressure

In Chap. 3, a distinction was made between those variables which are integrals of flow and those which are properties of a flowing stream. This distinction takes on added significance now, being reflected in the configuration of the feedforward system. Load is a flow term, of which liquid level and pressure are integrals. Therefore feedforward calculations for liquid level and pressure are generally linear. But where a property of the flowing stream, such as temperature or composition, is to be controlled, the system will be found nonlinear in appearance.

In general, liquid-level and pressure processes appear mathematically as follows :

$$\tau \frac{dc}{dt} = mK_m\mathbf{g}_m - qK_q\mathbf{g}_q \tag{8.1}$$

The terms $K_m$, $\mathbf{g}_m$, $K_q$ and $\mathbf{g}_q$ represent the steady-state and dynamic gain terms for the manipulated variable and load. The feedforward control system is to be designed to solve for $m$, substituting $r$ for c:

$$m = \frac{\tau(dr/dt) + qK_q\mathbf{g}_q}{K_m\mathbf{g}_m}$$

Since $dr/dt$ is normally zero,

$$m = q\frac{K_q\mathbf{g}_q}{K_m\mathbf{g}_m} \tag{8.2}$$

Feedforward is commonly applied to level control in a drum boiler. Because of the low time constant of the drum, level control is subject to rapid load changes. In addition, constant turbulence prevents the use of a narrow proportional band, because this would cause unacceptable variations in feedwater flow. The feedforward system simply manipulates feedwater flow to equal the rate of steam being withdrawn, since this represents the load on drum level. The system is shown in Fig. 8.2.

If the two flowmeters have identical scales, which is to be expected, the ratio $K_q/K_m$ of Eq. (8.2) is 1.0. Furthermore, the dynamic elements

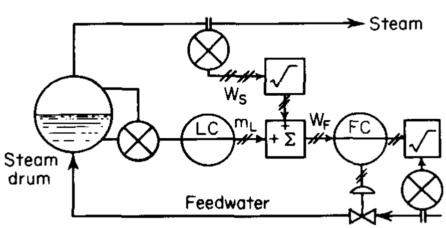



*FIG 8.2. Feedwater flow is set equal to steam flow in a drum boiler.*

$g_q$ and $g_m$ are virtually nvnexistent. The control system then simply solves the equation

$$W_F = W_s + m_L - 0.5$$

The terms $W_F$ and $W_s$ are mass flows of feedwater and steam, respectively; $m_L$ is the output of the level controller, whose normal value is 0.5.

It must be remembered that liquid-level processes such as this are non-self-regulating. The controlled variable will consequently drift unless feedback is applied. Since integral feedback may not be used alone, because instability would result, a two-mode controller is always used. In the steady state, feedwater flow will always equal steam flow, so the output of the level controller will seek the bias applied to the computation. If the controller is to be operated at about 50 percent output, that bias must be 0.5, as indicated in the formula. The controller does not have to integrate its output to the entire extent of the load change with a forward loop in service, but need only trim out the change in error of the computation during that interval.

This feedforward system has two principal advantages:

1. Feedwater flow does not change faster or farther than steam flow.
2. Control of liquid level does not hinge upon tight settings of the feedback controller.

Because this feedforward system, like many, is based on a material balance, accurate manipulation of feedwater flow is paramount. In general, the output of a feedforward system is the set point for a cascade flow loop and does not go directly to a valve. Valve position is not a sufficiently accurate representation of flow.

## Systems for Temperature and Composition

Temperature and composition are both properties of a flowing stream. Heat and material balances involve multiplication of these variables by flow, producing a characteristic nonlinear process model. Feedforward systems for control of these variables are similarly characterized by multiplication and division. The general form of process model for these applications is

$$c = K_p \frac{m\mathbf{g}_m}{q\mathbf{g}_q} \tag{8.3}$$

A single coefficient $K_p$ is sufficient to identify the steady-state gain.

The feedforward equation to control this general process is simply the solution for ***m***, replacing c with $r$:

$$m = \frac{rq\mathbf{g}_q}{K_p\mathbf{g}_m} \tag{8.4}$$

Notice that the manipulated variable is affected equally by the load and set point, which are multiplied. In level and pressure processes, the set point is added and contributes little to the forward loop.

Because temperature and composition measurements are both subject to dead-time and multiple lags, they are relatively difficult to control. As a result, it is perfectly reasonable to expect that feedforward can be more readily justified in these applications. But along with the need, there likewise exists the problem of defining these processes well enough to use computing control. In addition, nonlinear operations and dynamic characterization are required. Yet multipliers and dividers did not come into common usage in control systems until about 1960. It is easy to understand, therefore, why level control was perhaps the first but hardly the most significant application of the feedforward principle.

## Application to a Heat Exchanger'

The most easily understood demonstration of feedforward is in the control of a heat exchanger. The computation is a heat balance, where the correct supply of heat is calculated to match the measured load. The process is pictured in Fig. 8.3. Steam flow $W_s$ is to be manipulated to heat a variable flow of process fluid $W_p$ from inlet temperature $T_1$ to the desired outlet temperature $T_2$.

The steady-state heat balance is readily derived:

$$Q = W_s H_s = W_p C_p (T_2 - T_1)$$

where Q = heat transfer rate

$H_s$ = latent heat of the steam

$C_p$ = heat capacity of the liquid

Solving for the manipulated variable,

$$W_s = W_p K (T_2 - T_1)$$

The coefficient $K$ combines $C_p/H_s$ with the scaling factors of the two flowmeters, and is included as an adjustable constant in the computer;

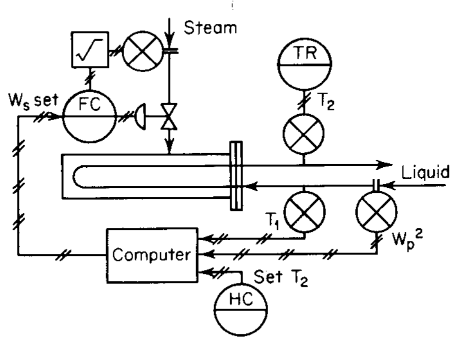



*FIG 8.3. The feedforward control system calculates the correct steam flow to match the heat load.*

$T_2$ is the set point; $W_p$ and $T_1$ are load variables. Witness the multiplication of flow by temperature.

In the control computer that is shown in Fig. 8.4, the coefficient $K$ is introduced as the gain of the summing amplifier. The measurement of liquid flow is linearized before multiplying; then steam flow must also be linearized, to be compatible with its set point.

Steam flow is begun automatically by increasing both the liquid flow and the set point, since it is proportional to their product. If the exit temperature fails to reach the set point, it indicates that the ratio of steam flow to liquid flow is incorrect. In practice, this ratio is easily corrected by adjusting $K$ until the offset is eliminated. This is the principal calibrating adjustment for the system; it sets the gain of the forward loop. If the system is perfectly accurate, exit temperature will respond to a change in liquid flow as shown in Fig. 8.5.

Two failings of the steady-state control calculation should be noted:

1. Each load change is followed by a period of dynamic imbalance, which makes its appearance as a transient temperature error.
2. The possibility of offset exists at load conditions other than that at which the system was originally calibrated.

On the other hand, the performance of the system exhibits a high level of intelligence. It is inherently stable and possesses strong tendencies toward self-regulation. Should liquid flow be lost for any reason, steam flow will be automatically discontinued. Feedback control systems ordinarily react the other way upon loss of flow, because the measurement of exit temperature is no longer affected by heat input.

The importance of basing control calculations on mass and energy balancing cannot be stressed too highly. First, they are the easiest equations to write for a process, and they ordinarily contain a minimum of unknown variables. Second, they are not subject to change with time. It was not necessary, for example, to know the heat transfer area or coefficient or the temperature gradient across the heat-exchanger tubes in order to write their control equation. And should the heat transfer coefficient change, as it surely will with velocity, or fouling, etc., control is unaffected. It may be necessary for the steam valve to open wider to raise the shell pressure in the event of a reduction in heat transfer coefficient, but steam flow consistent with the heat balance equation will be maintained nonetheless.

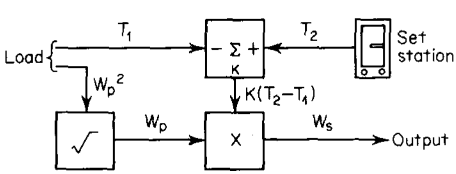



*FIG 8.4. Three computing elements and a set station provide the steady-state heat balance.*

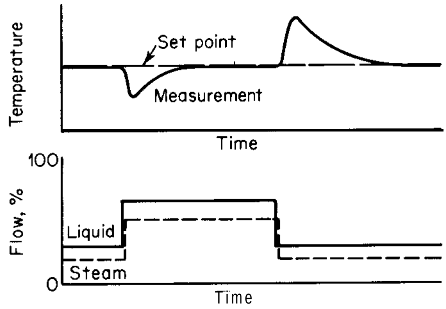



*FIG 8.5. If the steady-state calculation is correct, temperature will eventually return to the set point following a flow change.*

Some unknown factors do exist, however. No allowance was made for losses. If they are significant, and particularly if they change, an offset in exit temperature will result. Steam enthalpy could also vary, as well as the calibration of the steam flowmeter, should upstream pressure change. But for the most part, these factors are readily accountable, whereas heat and mass transfer coefficients may not be.

Response to a set-point change will be exponential, appearing as if the loop were open. Since moving the set point causes steam flow to move directly to the correct value, the response is exactly what was sought with complementary feedback (see Fig. 4.11).

## APPLYING DYNAMIC COMPENSATION

The transient deviation of the controlled variable depicted in Fig. 8.5 was attributed to a dynamic imbalance in the process. This characteristic can be assimilated from a number of different aspects.

If the load on the process is defined as the rate of heat transfer, then increasing load calls for a greater temperature gradient across the heat transfer surface. Since the purpose of the control system is to regulate liquid temperature, steam temperature must increase with load. But the steam in the shell of the exchanger is saturated, so that temperature can be increased only by increasing pressure, which is determined by the quantity of steam in the shell. Before the rate of heat transfer can increase, the shell must contain more steam than it did before.

In short, to raise the rate of energy transfer, the energy level of the process must first be raised. If no attempt is made to add an extra amount of steam to overtly raise the energy level, it will be raised inherently by a temporary reduction in energy withdrawal. This is why exit temperature falls on a load increase.

Conversely, on a load decrease, the energy level of the process must be reduced by a temporary reduction in steam flow beyond what is required for the steady-state balance. Otherwise energy will be released as a transient increase in liquid temperature.

The dynamic response can also be envisioned simply on the basis of the velocity difference between the two inputs of the process, although this is less representative of what actually takes place. The load change appears to arrive at the exit-temperature bulb ahead of a simultaneous steam-flow change. To correct this situation, steam flow must be made to lead liquid flow.

The technique of correcting this transient imbalance is called "dynamic compensation."

### Determining the Needs of the Process

Capacity and dead time can exist on both the manipulated and the load inputs to the process. There may also be some dynamic elements common to both, such as the lags in the exit-temperature bulb for the heat exchanger. The relative locations of these elements appear as shown in Fig. 8.6.

A feedback controller must contend with $g_m \times g_p$, which are in series in its closed loop. But the feedforward controller need only be concerned with the ratio $g_q/g_m$, in order to make the corrective action arrive at the divider at the same time as the load. Recall the appearance of this ratio in both Eqs. (8.2) and (8.4). In some difficult processes, the manipulated variable enters at the same location as the load, e.g., in a dilution process where all streams enter at the top of a vessel. In this case, even though $g_p$ may be quite complex, $g_m$ and $g_q$ could be nonexistent, making dynamic compensation unnecessary.

Perhaps the easiest way to appreciate the need for dynamic compensation is to consider a process in which $g_q$ and $g_m$ are dead time alone. Let $\tau_q$ and $\tau_m$ represent their respective values. The response of the controlled variable as a function of time is

$$c(t) = K_p \frac{m(t - \tau_m)}{q(t - \tau_q)}$$

The division makes the process fundamentally nonlinear, which complicates dynamic analysis. To allow inspection of the transient response of the process, analysis must be made on an incremental basis, by differentiating both sides of the equation:

$$dc(t) = K_p \left[ \frac{dm(t - \tau_m)}{q} - \frac{m\, dq(t - \tau_q)}{q^2} \right] \tag{8.5}$$

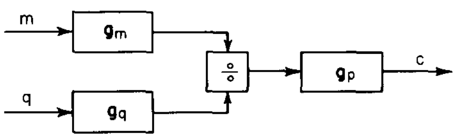



*FIG 8.6. Compensation is needed when the dynamic elements of the two inputs differ.*

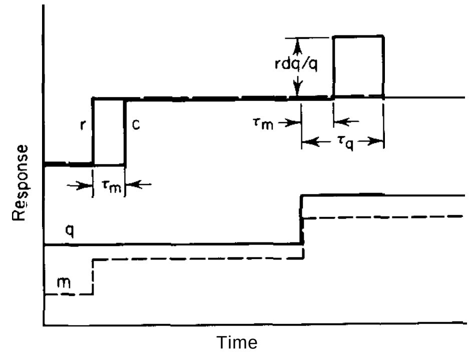



*FIG 8.7. Lack of dynamic compensation produces a transient equal to the difference in dead times.*

If only a steady-state control calculation is made,

$$m = \frac{rq}{K_p} \tag{8.6}$$

Differentiating,

$$dm = \frac{1}{K_p}(r\,dq + q\,dr) \tag{8.7}$$

Substituting for m and $dm$ in Eq. (8.5) yields the closed-loop response:

$$dc\,(t) = dr\,(t - \tau_m) + dq\,\frac{r}{q}\,(\tau_q - \tau_m) \tag{8.8}$$

Equation (8.8) shows that the set-point response is delayed by $\tau_m$ and that a load change will induce a transient of duration $\tau_q - \tau_m$ and magnitude $r\,dq/q$. Both responses appear in Fig. 8.7.

Of the two, load response is the more important, because set-point changes are ordinarily less frequent. Ideally, the load signal should be delayed by $\tau_q$ before it is multiplied, and then advanced by $\tau_m$. It is impossible to create a time advance, however. So dynamic compensation is best introduced in this application by delaying the feedforward signal by an amount $\tau_q - \tau_m$. If $\tau_m > \tau_q$, compensation is impossible.

It has been pointed out that dynamic compensation generally takes the form $\mathbf{g}_q/\mathbf{g}_m$. It may be recalled, however, that the ratio of two vector quantities like these resolves into the ratio of their magnitudes and the difference between their phase angles. Since dead-time elements have unity gain, their ratio is also unity; their only contribution is phase lag. This is why the ratio $\mathbf{g}_q/\mathbf{g}_m$ appears as the difference $\tau_q - \tau_m$ between the dead times.

The complete forward loop, including dynamic compensation, appears in Fig. 8.8. Note the complete cancelation of all elements in the load path by the elements in the forward loop.

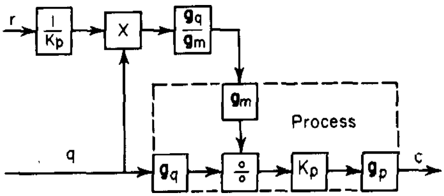



*FIG 8.8. Observe how faithfully the forward loop images the properties of the process.*

Because forward loops exhibit absolutely no oscillatory tendencies, to talk of gain and phase is rather inconsequential. Step responses will be used throughout, since they constitute the most severe test of system performance. The response of systems under feedforward control, both with and without dynamic compensation, differs markedly from that experienced with feedback control. For this reason, it is not surprising that dynamic elements in the forward loop bear little resemblance to the conventional modes of feedback controllers.

Although dead time serves as a useful demonstration of why dynamic compensation is necessary, it rarely appears alone in a process. In fact, multiple lags are most, commonly encountered in actual applications. Fortunately, there is usually one dominant lag on each side of the process, which acts as the principal element to be compensated. The response of a process wherein $\mathbf{g}_m$ and $\mathbf{g}_q$ are first-order lags of time constants $\tau_m$ and $\tau_q$, respectively, can be found by substituting their individual response terms into Eq. (8.8). Thus $t - \tau_m$ becomes $1 - e^{-t/\tau_m}$, and $\tau_q - \tau_m$ is replaced with $e^{-t/\tau_q} - e^{-t/\tau_m}$:

$$dc\ (t) = dr\ (1 - e^{-t/\tau_m}) + dq\ \frac{r}{q}\ (e^{-t/\tau_q} - e^{-t/\tau_m}) \tag{8.9}$$

Figure 8.9 gives both set-point and load-response curves described by this equation, for the case where $\tau_q > \tau_m$. Compare it to the heat-exchanger response, Fig. 8.5, where $\tau_m > \tau_q$.

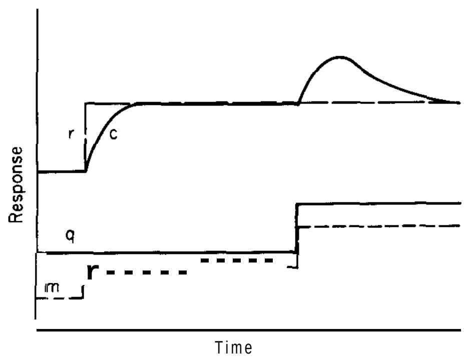



*FIG 8.9. Lack of dynamic compensation shows up principally as a load-response transient.*

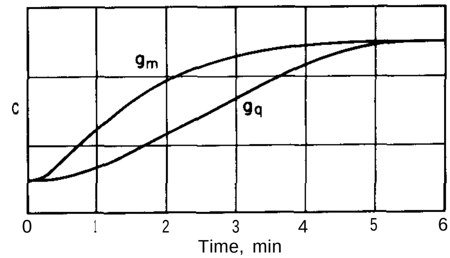



*FIG 8.10. Comparison of the open-loop response of c to m with the reverse response of c to q shows that m must be made to lag q for this process.*

A qualitative appraisal of the requirement for dynamic compensation may be obtained from a comparison of open-loop response curves. Because an increase in the manipulated variable acts in opposition to the load, their individual step-response curves will diverge. One or the other response will have to be inverted so that the two curves may be superimposed, as is done in Fig. 8.10. The response of such a process under uncompensated feedforward control appears as the difference between these two curves.

If the curves do not cross, the uncompensated forward-loop response will lie wholly on one side of the set point, as in Figs. 8.5 and 8.9. Which side of the set point depends on whether the difference $g_m - g_q$ is positive or negative. If the curves cross, the uncompensated forward-loop transient will cross the set point.

### The Lead-Lag Unit

For the bulk of processes to which feedforward control may be applied, the dynamic elements $g_q$ and $g_m$ are similar in nature and value. Although dead time may be encountered in both, their values are usually close enough to provide nearly complete cancelation. So in most cases, only the dominant lags need to be considered. In addition, the presence of the common element $g_c$ provides enough attenuation to make exact dynamic compensation unnecessary. Fortunately this allows one dynamic compensator to be used almost universally: the lead-lag unit.

A lead was defined earlier as the inverse of a lag; the lead term to be used here represents $1/g_m$, and the lag represents g,. The output m(t) of a lead-lag unit follows a step input $m$ as

$$m(t) = m\left( + \frac{\tau_1 - \tau_2}{\tau_2} e^{-t/\tau_2}\right) \qquad (8.10)$$

In the equation, $\tau_1$ is the lead time and $\tau_2$ the lag time; either may be greater, allowing an overshoot or an undershoot, as Fig. 8.11 demonstrates.

The step-response curve reveals an instantaneous gain of $\tau_1/\tau_2$, and recovery to 63 percent of the steady-state value is effected in time $\tau_2$.

Oddly enough, the most stringent specification on a lead-lag unit is

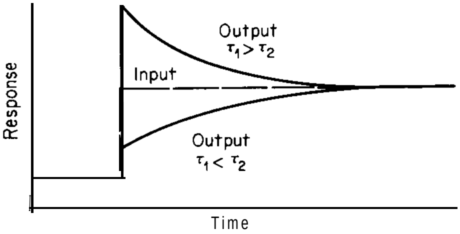



*FIG 8.11. The lead-lag unit can be made to overshoot or undershoot a step input.*

steady-state accuracy. If it cannot accurately repeat its input in the steady state, the lead-lag unit degrades the performance of the forward loop. Consequently, linearity, repeatability, and freedom from hysteresis are mandatory-more so than for a conventional controller. In addition, lead and lag times need to be adjustable to match the time constants of most processes.

Exact compensation may be impractical for very slow processes, particularly those with electronic components, because of impedance limitations. Pneumatic devices have a greater potential range in this respect, because extremely large-capacity tanks can be used without danger of leakage. References (1) and (2) describe several arrangements for obtaining lead-lag functions using standard pneumatic control components.

Digitally, the lead-lag function can be realized with a simple iterative procedure. To demonstrate this procedure, $x$ will represent the input, y the input lagged by $\tau_2$, and $z$ will be y led by $\tau_1$. The differential equations are:

$$x = y + \tau_2 \frac{dy}{dt} \qquad z = y + \tau_1 \frac{dy}{dt}$$

Rearranging,

$$z = x + (\tau_1 - \tau_2) \frac{dy}{dt} \qquad \frac{dy}{dt} = \frac{1}{\tau_2}(x - y)$$

Digital computers are sampling devices, repeating their calculations at some regular interval At. Therefore the differentials above must be rewritten as difference equations. First the current value of $z$ may be calculated at each interval, from current values of x and y:

$$z_n = x_n + \frac{\tau_1 - \tau_2}{\tau_2}(x_n - y_n) \tag{8.11}$$

But y must be incremented before the next calculation can be made:

$$y_{n+1} = y_n + \frac{\Delta t}{\tau_2}(x_n - y_n) \tag{8.12}$$

### Adjusting the Dynamic Terms

The lead-lag function enables the delivery of more (or less) energy or mass to the process to raise its potential during a load change. The integrated area between its input and output should match the area of the transient in the uncompensated response curve. If this is done, the net area of the response will then be zero.

The integrated area between input and output of the lead-lag unit can be found from Eq. (8.10). First the difference between input and output should be normalized by dividing by the input magnitude:

$$\frac{m(t) - m}{m} = \frac{\tau_1 - \tau_2}{\tau_2} e^{-t/\tau_2}$$

Integrating over the limits of zero to infinity yields

$$\int_0^\infty \frac{m(t) - m}{m} dt = \tau_1 - \tau_2 \tag{8.13}$$

The normalized integrated area of the uncompensated loop response of Eq. (8.9) and Fig. 8.9 is similar:

$$\int_0^\infty (e^{-t/\tau_q} - e^{-t/\tau_m}) \, dt = \tau_m - \tau_q \tag{8.14}$$

This is further proof that $\tau_1$ should equal $\tau_m$ and $\tau_2$ should equal $\tau_q$.

Area alone is an insufficient index of proper compensation. A lead of 10 min and a lag of 9 min would produce the same area as a lead of 2 min and a lag of 1 min, but that area would be distributed differently. The location of the transient peak of the uncompensated response can be of help in estimating the actual values of $\tau_1$ and $\tau_2$. Let $\tau_1$ and $\tau_2$ be substituted for $\tau_m$ and $\tau_q$ in Eq. (8.9). By differentiating and then equating to zero, the time $t_p$ of the maximum (or minimum) can be found:

$$t_p = \left(\frac{1}{1/\tau_2 - 1/\tau_1}\right) \ln \frac{\tau_1}{\tau_2} \tag{8.15}$$

A plot of this relationship is given in Fig. 8.12.

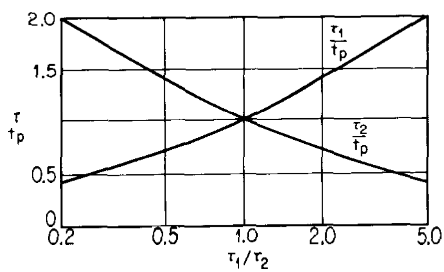



*FIG 8.12. The location of the peak in the uncompensated response transient can be used to infer the required compensation.*

Equation (8.15), like (8.14), is a single relationship containing two unknowns. Yet some definite conclusions can be drawn which will be useful in making preliminary adjustments:

1. If $\tau_1$ must exceed $\tau_2$, based upon the ***direction*** of the uncompensated transient, $\tau_2$ can safely be set in the vicinity of $0.7t_p$. If $\tau_1$ must be less than $\tau_2$, $\tau_2$ should be about $1.5t_p$.

2. Initially, $\tau_1$ can be set at about $2\tau_2$ in the former case, or $0.5\tau_2$ in the latter.

Once these preliminary adjustments have been introduced, the load response should be repeated, from which finer adjustments may be made. Figure 8.13 compares the load response, with varying degrees of compensation, to the uncompensated response of a typical process.

Notice that curve (b) in Fig. 8.13 lacks area compensation. Curve (c), on the other hand, shows adequate compensation with respect to area, in that it is distributed equally about the set point. In this case, the difference between $\tau_1$ and $\tau_2$ is correct, but their individual values are not. Once the correct area compensation has been found, $\tau_1$ and $\tau_2$ should ***both*** be adjusted in the same direction, to maintain their difference. In the example shown in Fig. 8.13*c*, $\tau_1$ and $\tau_2$ need to be reduced; this will make their ratio increase, which will move the centroid of the lead-lag area to the left. Curve (*d*) is the result of such an adjustment: it crosses the set point at approximately time $t_p$.

Perfect compensation is unattainable. For one thing, any process sufficiently problematic to warrant feedforward control can be expected to display some dead time in addition to capacity. This is true of the heat exchanger. But compensation for dead time is, at best, approximate. Second, the dynamics of most processes are subject to change.

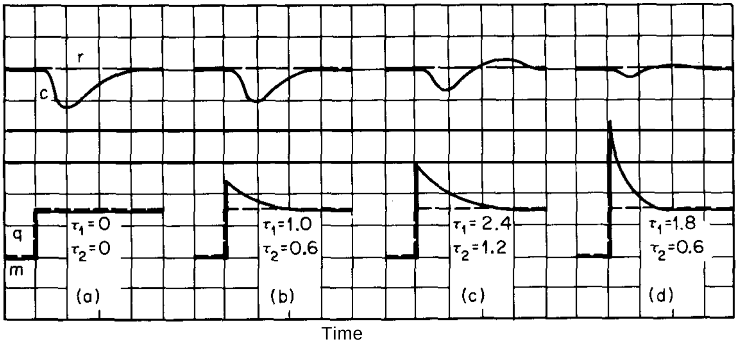



*FIG 8.13. Curve (a) is the uncompensated loop response; curves (b) and (c) reveal incorrect compensation; curve (d) is nearly perfect.*

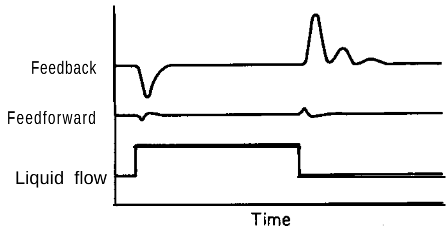



***FIG 8.14. Feedforward control is capable of reducing both the area and the duration of the load-response transients.***

This is also true for the heat exchanger, whose dead time varies with the rate of flow through the tubes.

It might be possible to construct an extremely complete dynamic model of a process, but any compensator with more than three terms to adjust would be unreasonably difficult to cope with. Furthermore, the purpose of dynamic compensation is to minimize an error which is already transient, so perfection is not really warranted. In most cases, a simple lead-lag function will be perfectly adequate and will be able to reduce the absolute area of the response curve by tenfold or more, distributed uniformly. Figure 8.14 compares the load response of the heat exchanger under dynamically compensated feedforward control with that encountered under feedback control.

For processes whose uncompensated load-response curves cross the set point, lead-lag may be inadequate. For these applications, an additional lag is useful in canceling the first part of the curve, while the lead-lag function compensates the balance. Distillation columns typically exhibit this characteristic. Further discussion of the problem will be found in Chap. 11 and in reference (3).

## ADDING FEEDBACK

The only serious failing of the feedforward technique is its dependency on accuracy. To provide perfect control, a system must model the plant exactly; otherwise whatever error may exist in positioning the manipulated variable causes offset. Errors may arise from several sources:

1. Inaccuracy in the measurement of load and manipulated variables
2. Errors in the computing components
3. Failure of the computing system to adequately represent the characteristics of the plant
4. Exclusion of significant load components from the feedforward system

The first and second items alone limit the accuracy of practical systems to the vicinity of 1 to 2 percent.'

Some processes, such as the heat exchanger described earlier, are easy to model. But this is not always the case, particularly when mass and heat transfer coefficients must be used over a wide range of operating conditions. Therefore item 3 may be of considerable importance in the more complicated processes.

The feedforward system cannot be all-inclusive. Some load components are so slight, or invariant, or ill defined, that their inclusion is not warranted. If payout and ease of operation are important, which is to be expected, the control system had best remain simple. To this extent, certain terms such as heat losses and ambient temperature effects are usually neglected. Yet their variation can induce a measurable offset.

If offset is intolerable, some means must be provided for recalibration while the system is operating. In general, this can be done most directly by readjusting the set point, which is already scaled in terms of the controlled variable. Other adjustments could be made, such as the coefficient $K$ in the heat-exchanger control system, but with less predictable results.

### The Role of Feedback

Regardless which term is trimmed to remove offset, the procedure amounts to manual feedback. Automatic feedback is perfectly capable of effecting the same result, if the controlled variable can be measured with sufficient reliability. (This qualification is significant in that feedforward control is occasionally used because a feedback measurement is unavailable.)

Proportional feedback trim is insufficient to eliminate offset, for the same reason that it was insufficient in a conventional control loop. The presence of feedforward control components within the feedback loop induces no substantial change in the mode settings required of the feedback controller; the process is just as *difficult* to control as it was without feedforward-only the *amount* of work required of the feedback controller has been markedly reduced.

Reset, then, is necessary if offset is to be eliminated altogether. Whether proportional and derivative are useful modes depends on the nature of the process. If rapid load changes outside the forward loop may be encountered, proportional and derivative action could be advantageous. If the process is fundamentally non-self-regulating, as in level control, proportional action is essential. Finally, if the process is fairly easy to control because of the absence of dead time, derivative may be useful in improving the dynamic load response-but this is unusual.

In general, mindful that feedforward control is warranted only on the most demanding and most difficult applications, reset is the only useful feedback mode. Experiments conducted on a heat exchanger, which is not particularly difficult to control, indicated that proportional and derivative feedback modes responded too slowly to contribute anything to the load response of the system, either with or without dynamic compensation.[1] On the other hand, feedback can be detrimental by promoting oscillation in an otherwise stable system. To this extent, the settings of the feedback controller, regardless of what modes have been selected, ought to be relaxed.

Inasmuch as reset controllers are not available as standard products from most manufacturers, proportional-plus-reset is the usual recommendation.

### Where to Introduce Feedback

This is not always an easy question to resolve. The feedback controller may be asked to perform a number of different services. In the heat-exchanger application it can be useful in correcting for heat loss, in which employment it should add an increment of heat to the process at all loads; this would amount to a zero adjustment. Or its principal function might be to correct for variable steam enthalpy, in which case it should apply a span adjustment by setting the coefficient $K$. In another process, linearity could be the largest unknown factor. But a single feedback controller can hardly be called upon to do all these things.

When no one source of offset is outstanding, the argument for readjusting the set point of the feedforward system is irresistable. Figure 8.15 shows how the feedback controller would fit into this arrangement.

Feedback is added simply by replacing the manual set station with a feedback controller; no additional computing elements are needed. Startup may proceed in orderly fashion in manual, simply by setting the output of the feedback controller equal to its set point. No guesswork is involved, since this is the known operating region for the system. When in automatic, there is ultimately only one set-point signal, because the controller output is now a variable.

Another feature of this configuration is that it displays the inherent inaccuracies of the forward loop. The difference between the set point and the output of the feedback controller is the offset which would have appeared if feedback had not been used. A plot of controller output vs. load could conceivably identify the principal sources of error in the com-

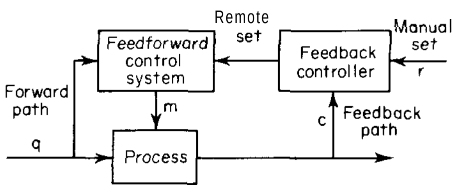



*FIG 8.15. In general, the feedback controller should reposition the set point of the feedforward system.*

puting system. Bear in mind, however, that the amount of offset, $dc$, resulting from an error in the manipulated variable $dm$, varies inversely with load:

$$dc = K_p \frac{dm}{q}$$

Although feedforward systems are designed primarily for regulation, some rules regarding set-point changes are noteworthy. To avoid the usual dynamic problems associated with feedback loops, the feedback controller should be placed in manual during set-point changes. In the case where this is a frequent occurrence, set-point changes can be introduced in automatic if sent directly to the forward loop; this requires multiplication of the set point by the output of the controller as shown in Fig. 8.16.

It is impossible, of course, for the process to respond instantaneously to a step in set point. Since the controlled variable will lag behind the set point, a positive error will develop before the new set point is reached. The feedback controller, being in automatic, will integrate the error, changing its output to a new but incorrect value; it must then bring its output back to the previous state by generating a negative error, equal in area to the earlier positive error. The effect is the same as that shown in Fig. 6.14, produced by the blending control system.

This situation can be remedied simply by inserting a lag in the set point to the feedback controller (but not to the multiplier), as Fig. 8.16 illustrates. The lag should be adjusted to prevent the overshoot that would be realized without it.

The feedback controller will always equalize the integrated error promoted by any disturbance entering a forward loop. For this reason, it should remain in manual while dynamic compensation is being adjusted.

Another important consideration when adding feedback is the location of the dynamic compensator. Although lead-lag can be beneficial to the response of a feedback loop, it interferes with manual operation. When an operator changes the output of the controller manually, he likes to see that action reproduced exactly by the manipulated variable. With a lag or lead-lag between the controller and the manipulated variable, several minutes-possibly even an hour-could elapse before the effect of the adjustment is complete. Therefore it is mandatory to arrange the system so that dynamic compensation is out of the feedback loop.

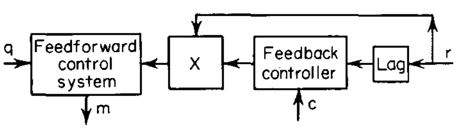



*FIG 8.16. If set-point response is important, the set point should go directly into the forward loop.*

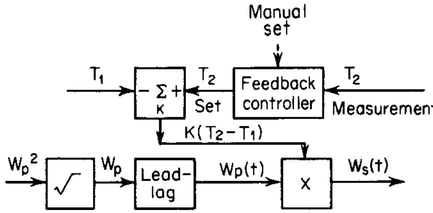



*FIG 8.17. The complete control system for the heat exchanger includes feedback and dynamic compensation.*

In theory, each forward loop should have its own dynamic compensator, but ordinarily only flow inputs can change fast enough to warrant dynamic compensation. Observe the location of the lead-lag unit in the system with two forward loops and one feedback loop, which is shown in Fig. 8.17.

### Mutual Adaptation

An adaptive control system was defined as having the ability to change its parameters in accordance with the changing character of the process. A feedforward control system, by itself, can only generate an output relative to known and measurable inputs as prescribed by a fixed program. Some factors relating to the process may be unknown and variable. For optimum performance, the feedforward system ought to be supplied with information regarding these unknowns. A feedback controller, on the other hand, is geared to solve for unknowns. So the inclusion of a feedback signal in a forward loop actually adapts the forward loop to unmeasured changes in the process.

Remarkably enough, the feedforward system also adapts the feedback loop to variations in process gain. Figure 8.14 shows the load response of the cited heat exchanger under feedback control. With increasing load, the transient is overdamped. With decreasing load, the transient is greater and underdamped, indicating that the process gain changes inversely with liquid flow, (This characteristic was discussed under "Variable Dynamic Gain" in Chap. 2, and again under "Dynamic Adaptive Systems" in Chap. 6.) Since the process gain varies inversely with flow, the controller gain ought to vary directly with flow. The complete control system for the heat exchanger, Fig. 8.17, illustrates how this is brought about.

The feedback loop sees $T_2$ as its input and $W_s$ as its output. But within the loop, $T_1$ is subtracted from, and $W_p$ multiplied by, the controller output. Subtraction is a linear operation, so gain is not changed therein; but multiplication is nonlinear, causing feedback gain to vary directly with flow $W_p$. Correct loop-gain adaptation cannot be achieved if the feedback is introduced in *any* other place. If the output of the feedback controller were to set $K$, then feedback gain would vary both with $W_p$ and with $T_2 - T_1$. But process gain does not vary with

$T_2 - T_1$. This is another argument, in favor of using feedback to trim the feedforward set point.

This mutual adaptation is a further indication how perfectly feedforward and feedback complement one another. Peedforward is fast, intelligent', and responsive, but also inaccurate; feedback is slow but accurate, and is capable of regulation in the face of unknown load conditions.

## ECONOMIC CONSIDERATIONS

Feedback control can be enforced on the heat exchanger using only three elements: transmitter, controller, and valve. Adding feedforward control requires another temperature transmitter, two flow transmitters, two square-root extractors, a steam-flow controller, a summer, a multiplier, and a lead-lag unit-nine items. Such an expense must be justified.

Although the heat exchanger serves as a convenient demonstration of feedforward control, only in rare instances would such refined control be justified for this type of application. Management likes to see investments such as advanced control systems pay for themselves in less than two years. If improved temperature control had no value, there would be no payout. But if it prolonged the life of the exchanger, or saved steam, or reduced maintenance, it would have a measurable worth.

Economic incentive for improved control is most likely to be found where consistent quality of a valuable product is important. Many unit operations are conducted at reduced capacity or low recovery, or use excessive amounts of utilities to ensure that quality will surpass specifications even with poor control. Naturally, the rate of payout from these sources will vary directly with the production rate of the plant. Large plants therefore encounter reduced risks.

But the payout of a control system is as much a factor of cost as of savings. A control system can be too perfect-that is, its cost can be out of proportion to the job to be done. In the control system for the heat exchanger, for example, both flow and inlet temperature were accepted as load changes. If inlet temperature were nearly constant, or subject only to slow variations, it could be eliminated from the system as long as feedback trim were available. This would reduce cost by saving a transmitter and a summer.

There are three principal areas where feedforward control can produce results unobtainable with any other technique:

1. A difficult process subject to frequent disturbances may never settle out under feedback control. Load changes occurring at intervals of less than three cycles are sufficient to develop this situation. This is not

uncommon on some towers, furnaces, and coupled processing units which may cycle at periods exceeding an hour.

2. In many plants, the variable of interest cannot be measured continuously, accurately, or quickly enough for adequate closed-loop control. Often a secondary variable is used as an inferential measure of the first, simply because it is the best available. In a case like this, however, product quality may suffer both from poor control of the inferential variable and from its indifferent relationship to the primary variable.

3. More interest is developing in the control of economic variables: cost, debits, yield, etc. These variables are not directly measured, and often not computed, to allow the closing of the economic loop. But even if they were available, conventional feedback control could not be used, because the intent is to maximize or minimize their value rather than to control at a given set point.

### Optimizing Programs[4]

Feedforward control systems are not limited solely to regulatory duties. In fact, the controlled variable may be easily programmed with respect to any measured term, simply by making the appropriate substitution in the process equation. If, for some reason, it were desired to vary the temperature of the liquid leaving the heat exchanger relative to its flow, this could readily be accommodated. To do so, the computing system would become somewhat different but no more difficult to implement than for a constant set point. If feedback trim were used, however, the set point to the feedback controller would have to be similarly programmed, which might increase the complexity of the system significantly.

An important observation is that any control program may be followed, even one that would cause the process gain $dc/dm$ to change sign. Hence feedforward is the logical means to achieve optimizing (steady-state adaptive) control. This has already been demonstrated in Chap. 6.

The first step in optimizing is the definition of the sources of loss within the plant. The process will generate valuable product somewhat proportional to the rate of material pumped into it, with the result that the profit can be subject to wide variation. If the losses are maintained at a minimum, however, the highest profit will always result, and values and rates of raw materials can be ignored.

Losses, or debits, are not difficult to define. They consist principally of utility costs, streams of unrecoverable materials, and lowered market value brought about by contaminants in the product. Debits peculiar to a process may vary with feed composition, catalyst activity, atmospheric conditions, and demands elsewhere within the plant, but feed rate generally exercises the greatest influence upon them. The purpose of the computer in this domain is to manipulate those variables which can best offset the influence of the above uncontrolled variables on the plant economy.

Complicated as all this seems to be, often a fairly simple relation can be derived in an effort to optimize part of the plant, or to partially optimize the whole plant. As an example, consider a simplified absorption tower where a gas stream $F$, containing $z$ percent of a valuable material, is absorbed by a liquid stream $L$, which is to be manipulated to maintain minimum-cost operation. Assume that gas-exit composition y varies as follows:

$$y = kz\frac{F}{L}$$

where $k$ = absorption rate coefficient.

The principal debits $l$ associated with such an operation might be losses of valuable material in the exit gas and costs of processing the absorbent:

$$l = v_1Fy + v_2L$$

where $v_1$ = product value
$v_2$ = processing cost

The debit equation may be rewritten on the basis of independent variables alone :

$$l = v_1kz\frac{F^2}{L} + v_2L$$

The minimum point on a curve of debits vs. $L$ would occur where the slope $dl/dL$ is zero.

$$\frac{dl}{dL} = -v_1kz\frac{F^2}{L^2} + v_2 = 0$$

This defines the locus of optimum $L$.

$$L_{\text{opt}}^2 \quad v_1kz\frac{F^2}{v_2}$$

Having solved the minimum-cost equation, it is only necessary to build a computer which will program $L^2$ to follow variations in $F^2$ and $z$ in accordance with current figures of $v_1$ and $v_2$. (The manipulated variable has been left in the form $L^2$ because flow rate is most commonly measured by a differential meter.) Adjustable coefficients should be available for perfecting the model in actual operation. Note that if no solution exists to the derivative, the process exhibit's no minimum. This would be the case if $v_2$ were zero.

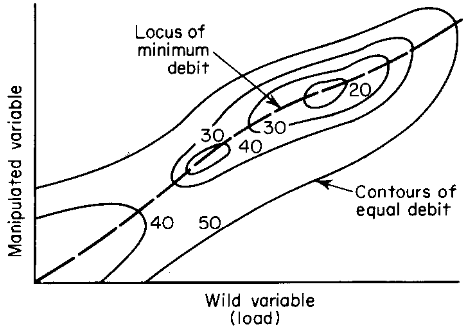



***FIG 8.18. The locus of minimum cost can be drawn across a contour plot such as this.***

Often such rigorous and simple expressions cannot be established. Shotgun patterns of data sometimes must be analyzed, involving all combinations of wild and manipulated variables until some measure of relationship can be envisioned. This usually culminates in contour plots such as Fig. 8.18. Contours of debit are shown as a function of a wild and a manipulated variable. If such a plot can be made, a line may be drawn across it representing the locus of minimum debit. A model of this line can then be made to program the manipulated variable as a function of the wild variable. Again, adjustable coefficients may be incorporated to perfect the model inasmuch as some doubt always accompanies relationships derived from real plant data.

Normally the operating conditions of any plant are surrounded by constraints and limitations. It is not surprising to learn that the optimum conditions for many plants lie outside equipment limitations. Many applications would not result in enough remuneration to pay for the computer or the engineering involved. These two facts severely restrict the number of processes that could benefit by optimizing control.

The total debits which can be expected to be recovered from a given operation with a given computer, divided by the installed cost of the computer, is the payout in percent per year. A simple analog like the one just discussed cannot be expected to recover all the existing debits in a process, though 50 to 75 percent recovery should not be difficult to realize.

A more complex analog designed to a more exact model might be able to recoup an additional 10 to 20 percent, but at perhaps twice the installed cost. It is easy to see that the simplest optimizing computer will nearly always result in the greatest rate of payout.

## SUMMARY

There is absolutely no question that feedforward is the most powerful technique that has ever been brought to bear in the regulation of difficult processes. Where certain unusual feedback modes like complementary feedback, sampling, and nonlinear functions were able to reduce the integrated error per unit load change perhaps twofold, feedforward is capable of a hundredfold improvement.[1] A feedforward system would only have to be accurate to ± 10 percent to provide tenfold improvement, in conjunction with feedback.

But nothing of great value is ever gained without cost. The cost in this instance is the process knowledge which must go into the design of every system. This precludes the mass production of fully adjustable feedforward systems. The only feedforward systems that will look alike are those being applied to like processes.

Although the feedforward loop carries most of the load, feedback is still important in its execution. The ultimate regulatory system is a triple cascade-the primary feedback controller trimming the set point of the feedforward loop, which in turn manipulates a cascade flow loop. The symmetry of such a configuration is striking.

The basic tenents of the technique were covered in this chapter with examples chosen for illustration rather than practical value. But in each. of the chapters that follow, feedforward will be applied to those processes which would most benefit from superior load regulation. Optimizing systems, where practicable, will also be given due consideration.

## REFERENCES

1. Shinskey, F. G.: Feedforward Control Applied, *ISA Journal,* November, 1963.
2. Samson, J. E.: Improvements in or Relating to Automatic Force Balance Apparatus, British Patent No. 860,485, February 8, 1961.
3. MacMullan, E. C., and F. G. Shinskey: Feedforward Analog Computer Control of a Superfractionator, *Control* Eng., March, 1964.
4. Shinskey, F. G.: Analog Computing Control for On-line Applications, *Control Eng.,* November, 1962.

## PROBLEMS

*8.1* In the heat-exchanger control system, equal flow rates of steam and liquid in percent of full scale raise the liquid temperature 100°F. Under these conditions, what offset would be induced by a 2 percent error in steam flow at 50 percent and at 25 percent flow?

**8.2** Write the equation for the feedforward system which controls exit-gas composition $y$ as it leaves the absorber described under the section on optimizing programs. How does this differ from the optimizing program solved in the example?

*8.3* Estimate the peak location of the uncompensated response curve which would result from feedforward control of the process whose dynamic characteristics are given in Fig. 8.10. What settings of lead and lag will be necessary for dynamic compensation?

8.4 Given a process whose dynamics consist of first-order lags, $\tau_m = 3$ min and $\tau_q = 1$ min: For conditions of $r$ and $q$, both at 50 percent, what is the integrated area per unit load change with an uncompensated forward loop? What is the integrated area with lead-lag compensation, if the lead time is 2.5 min and the lag time is 1 min?

8.5 Control of a given dead-time plus integrating process may be improved either (a) by adding derivative to the existing two-mode controller (interacting) or (b) by investing in a noninteracting controller or (c) by adding a simple forward loop. In the last case, the feedformard system is only likely to be accurate to $\pm 10$ percent. What is the improvement in integrated area per unit load change for each of the three cases? What would be the effect of variable dead time in each case?

# Applications

# PART 4

# Control of Energy Transfer

CHAPTER 9

The principles governing energy transfer apply to a broad spectrum of processes, from the combustion of fuel in a steam plant to the generation of hydraulic horsepower by a pump at the other end of the power line. Whether the energy is in the form of heat, electricity, head, or whatever, its conservation must be enforced : this is the "first law of thermodynamics."

Prerequisite to the study of thermodynamic processes is an understanding of its terminology. Energy is a measure of the state of a system; work is that amount of energy released or absorbed when the state of the system is changed. Energy and work therefore have similar units, although either may be thermal, electrical, hydraulic, etc. They are typically expressed as watthours, Btu's, foot-pounds, etc. Power, however, is the rate of flow of energy; control of energy transfer is therefore control of power. Thermal power is expressed as heat flow in Btu/hr, electrical power is expressed in watts, and mechanical power is expressed in horsepower or' ft-lb/sec.

Many processes, such as heat exchangers, involve the transfer of energy without its conversion. But worthy of deeper study are those processes in which energy is converted as well. Chemical and nuclear reactors, furnaces, engines, pumps, and compressors are all included in this category. Whatever the process, the balancing of mass and energy should serve as the basis for control system design.

## HEAT TRANSFER

Whenever flowing streams are joined, heat transfer is governed by mixing. Most heat transfer operations, however, are limited by the necessity of maintaining isolation between the flowing streams; in these cases, the boundary conditions at the heat transfer surfaces control its flow. Radiation is important where temperatures are sufficiently high to promote incandescence, typically in the combustion of a fuel. Each of these situations will be examined individually.

### Direct Mixing

Occasionally two or more streams are mixed to control the temperature of the blend. Unless they are thoroughly mixed, however, considerable error may be encountered in the measurement of final temperature, so this should be the first consideration. Special mechanical fittings are necessary, for example, to adequately mix steam with water or to spray water into a steam line.

A direct mixing system was discussed in Chap. 3. At that time, the characteristic nonlinearities of the process were noted. In general, a system combining streams of mass flows $W_1$ and $W_2$ and enthalpies $H_1$ and $H_2$ will yield a stream of mass flow $W$ and enthalpy $H$, conserving both mass and energy:

$$W_1 + W_2 = W \tag{9.1}$$

$$W_1H_1 + W_2H_2 = WH = (W_1 + W_2)H \tag{9.2}$$

For the case where both streams consist of the same fluid, e.g., water, the temperature of either one may be used as a reference. Then final temperature is determined from

$$W_2(T_2 - T_1) = (W_1 + W_2)(T - T_1)$$

$$T = T_1 + \frac{W_2(T_2 - T_1)}{W_1 + W_2} \tag{9.3}$$

Notice that $T$ varies nonlinearly with all three flows. A dimensionless plot of Eq. (9.3) appears in Fig. 9.1. Compare it with the numerical example given in Fig. 3.7.

If total flow and final temperature are both to be controlled by manipulating $W_1$ and $W_2$, coupling will exist between the loops. The relative-

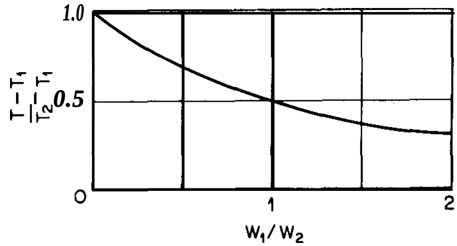



*FIG 9.1. A plot of dimensionless temperature vs. dimensionless flow displays a typically nonlinear function.*

gain matrix appears as follows:

| | $W_1$ | $W_2$ |
|---|---|---|
| $W$ | $\dfrac{T_2 - T}{T_2 - T_1}$ | $\dfrac{T - T_1}{T_2 - T_1}$ |
| $T$ | $\dfrac{T - T_1}{T_2 - T_1}$ | $\dfrac{T_2 - T}{T_2 - T_1}$ |

Temperature control can be linearized through the use of a three-way mixing valve if flow control is not a requirement. As one inlet port of a three-way valve is opened, the opposite inlet is closed. In this way $W_2$ can be increased and $W_1$ decreased simultaneously, while their sum remains nearly constant. The fraction of total flow admitted through one inlet port is then directly proportional to valve position $m$:

$$m = \frac{W_2}{W} = \frac{W_2}{W_1 + W_2}$$

Substitution into Eq. (9.3) shows that temperature is now linear with valve position:

$$T = T_1 + m(T_2 - T_1)$$

It is not unusual to find three-way valves employed in this service.

If total flow is to be controlled, too, a second valve may be placed downstream of the mixing. But if the supply pressures for the source streams are not equal, response will become nonlinear at low flow. And if flow is shut off entirely, the source with the higher pressure can drive its fluid back into the other source, unless protection is provided.

## Fluid-Fluid Heat Exchangers

Heat transfer from one fluid to another through a barrier surface is determined by driving force and resistance:

$$Q = UA\,\Delta T_m \tag{9.4}$$

Control of heat flow Q can thus be effected by manipulating the heat

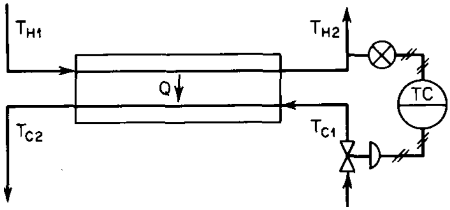



*FIG 9.2. The general case is heat transfer between hot and cold fluids in counterflow.*

transfer coefficient $U$, surface area $A$, or the mean temperature difference $AT$, between the fluids.

Even if $U$ and A could be maintained constant, Eq. (9.4) still contains two variables. The objective of most heat exchangers is the control of temperature, which varies with heat transfer rate, but which also affects the rate of heat transfer as Eq. (9.4) indicates. Consequently most heat transfer processes are highly self-regulating.

Further equations are necessary to close the loop, by relating fluid temperatures to heat flow. But a heat exchanger involves two fluids whose temperature distributions from inlet to outlet are both subject to change, both affecting AT,. For the general case, consider heat transfer between two fluids with no change in phase, as shown in Fig. 9.2.

The temperature difference affecting heat transfer between the two fluids in Fig. 9.2 is actually a logarithmic mean:

$$\Delta T_{lm} = \frac{(T_{H1} - T_{C2}) - (T_{H2} - T_{C1})}{\ln\left(\dfrac{T_{H1} - T_{C2}}{T_{H2} - T_{C1}}\right)} \tag{9.5}$$

In most, cases, fortunately, the arithmetic mean is sufficiently accurate for indicating the relationships between the variables, if not for use in equipment design:

$$\text{AT,,} = \frac{(T_{H1} - T_{C2}) + (T_{H2} - T_{C1})}{2} \tag{9.6}$$

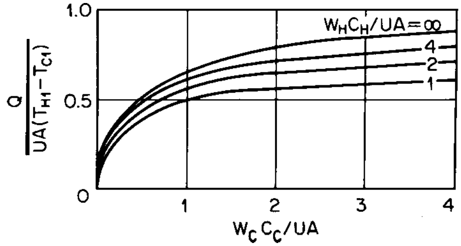
The error approaches zero as the temperature differences at the ends of the exchanger approach each other, and is less than 10 percent with a 4 : 1 ratio of temperature difference.

Each of the two fluids will be assigned a mass flow $W$ and a specific heat C. One of the flows ordinarily is wild and represents the load on the exchanger; the other is often manipulated in some way to control the exit temperature of the first. Temperature changes in both streams are interrelated :

$$Q = W_H C_H (T_{H1} - T_{H2}) = W_C C_C (T_{C2} - T_{C1}) \tag{9.7}$$

Equations (9.4), (9.6), and (9.7) contain four expressions with four unknowns, Q, $AT_{,,}$, $T_{H2}$, and $T_{C2}$. They can be solved simultaneously

*FIG 9.3. Manipulation of flow has little effect on heat transfer at high flow rates.*




for any of the four unknowns. The solution for heat transfer rate is the least complicated :

$$Q = \frac{T_{H1} - T_{C1}}{1/UA + \frac{1}{2}(1/W_HC_H + 1/W_CC_C)} \qquad (9.8)$$

Heat transfer rate ran be normalized by dividing by its maximum possible value, which would occur with both streams at infinite flow such that $T_{H2} = T_{H1}$ and $T_{C2} = T_{C1}$:

$$Q_{\max} = UA(T_{H1} \quad T_{C1}) \qquad (9.9)$$

$$\frac{Q}{UA(T_{H1} - T_{C1})} = \frac{1}{1 + (UA/2)(1/W_HC_H + 1/W_CC_C)} \qquad (9.10)$$

E'igure 9.3 is a plot of normalized heat transfer rate vs. normalized flow of cold fluid with the flow of hot fluid as a parameter. Observe the extreme nonlincnrity of the curves and how ineffective the manipulation of flow is over a wide operating range.

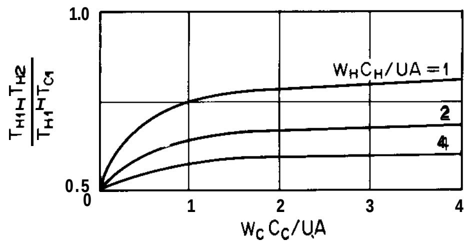
Substitution of Eq. (9.7) into (9.10) yields the following formulas describing dimensionless temperatures as a function of flow rates:

$$\frac{T_{H1} - T_{H2}}{T_{H1} - T_{C1}} = \frac{1}{W_HC_H/UA + \frac{1}{2}(1 + W_HC_H/W_CC_C)} \qquad (9.11)$$

$$\frac{T_{C2} - T_{C1}}{T_{H1} - T_{C1}} = \frac{1}{W_CC_C/UA + \frac{1}{2}(1 + W_CC_C/W_HC_H)} \qquad (9.12)$$

To envision what effect flow rates have upon exit temperature, Eq. (9.11) is plotted in Fig. 9.4 with the same abscissa and parameter that were used in Fig. 9.3.

*FIG 9.4. It is apparent that effective temperature control cannot be obtained over very wide ranges by manipulation of flow rate.*



Not only does the slope of the curves change with temperature, but it also changes with load $W_H$. Any horizontal line drawn across Fig. 9.4 will present the conditions required for temperature control. Doubling of the load at any given temperature requires the manipulated variable $W_C$ to be much more than doubled.

In practice, the overall heat transfer coefficient also varies with the flow rates, which improves the controllability somewhat. Although the film coefficient on each side of the heat transfer surface varies at about the 0.8 power of the fluid velocity, for simplification it will be assumed that the relationship is linear. Furthermore the reciprocal of the overall heat transfer coefficient mill be assumed to be the sum of the reciprocals of the individual film coefficients:

$$\frac{1}{U} = \frac{1}{W_H k_H} + \frac{1}{W_C k_C}$$

The terms $k_H$ and $k_C$ are the flow indices of their respective heat transfer coefficients. Combining with Eq. (9.11) yields:

$$\frac{T_{H1} - T_{H2}}{T_{H1} - T_{C1}} = \frac{1}{C_H/Ak_H + \tfrac{1}{2} + (W_H C_H/W_C C_C)(C_C/Ak_C + \tfrac{1}{2})} \tag{9.13}$$

A plot of Eq. (9.13) for conditions of $C_H/Ak_H = C_C/Ak_C = 1$ is given in Fig. 9.5. Compare it with the curves of Fig. 9.4.

The point of the foregoing analysis has been to demonstrate the nonlinear properties associated with heat transfer. Even under the most favorable conditions, manipulation of flow is far from satisfactory for temperature control. There are practical considerations, too. Throttling of streams which may contain impurities (river water, for example) can cause deposits to accumulate, fouling the heat transfer surfaces. Furthermore, manipulation of flow causes variable loop gain through the variation of dead time. In the event that there is no alternative to the manipulation of flow, an equal-percentage valve characteristic should be chosen.

Part of the stream whose temperature is to be controlled may be allowed to bypass the exchanger as shown in Fig. 9.6. But Fig. 9.3 indicates that

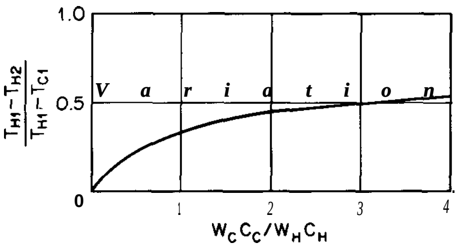



**FIG. 9.5** *Variation of the heat transfer coefficient with flow somewhat eases the nonlinearity of the process.*

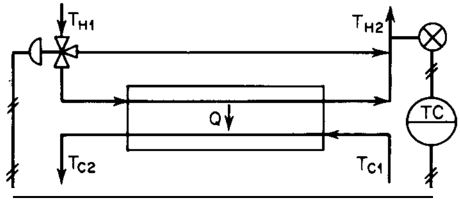



*FIG 9.6. Bypassing the exchanger will not improve linearity, but does reduce response time.*

the rate of heat transfer is scarcely affected by the flow of either stream for reasonable rates of flow. If the heat transfer rate is nearly constant, the final temperature of the process stream after reunion with the bypassed flow will also be nearly constant; consequently the linearity of response is not noticeably improved.

Bypassing can help the dynamic response, however, in that the flow of coolant is maintained at a high rate, rather than being throttled, as it would be if it were the manipulated variable. Furthermore, the bypass stream shortens the time delay between a change in valve position and the response of final temperature.

### Boiling Liquids and Condensing Vapors

The control situation is much more favorable where a change in phase is encountered. Because the latent heat of vaporization, $H_v$, predominates, a measurement of the mass flow $W$ of the boiling or condensing medium is also a measure of the rate of heat transfer:

$$Q = WH_v \tag{9.14}$$

Furthermore, the temperature of the boiling or condensing medium scarcely changes from inlet to outlet of the exchanger.

Whenever steam is used as a heating medium, manipulation of its flow to bring about temperature control of the process fluid is effective. If the process fluid is boiling, steam flow directly infers its rate of vaporization. The pressure of the steam in the exchanger is only an indication of steam temperature and is not a particularly useful measure of heat transfer; it can be used to estimate the heat transfer coefficient, however.

Exchangers supplied with steam as a heating medium exhibit a strong tendency toward self-regulation. Since the film transfer coefficient for condensing steam is much greater than a flowing gas or liquid, the rate of heat transfer is principally governed by the film coefficient of the process fluid. Since this coefficient varies almost linearly with fluid velocity, heat transfer will vary almost linearly with flow, if steam temperature is maintained. The latter is achieved simply by regulating the pressure of the steam in the exchanger. Thus without being directly controlled, the exit temperature of the process fluid will nonetheless be well regulated.

The flow of steam to a process heater or reboiler may also be manipulated by a valve in the condensate line. The rate of heat transfer is actually changed by partially flooding the exchanger with condensate. Because a change in condensate level is necessary to affect steam flow, this system may respond more slowly than direct manipulation of steam flow, hut it has the distinct advantage of requiring a much smaller valve.

Whether sufficient heat has been removed to totally condense a vapor can be determined by the temperature of its condensate, if constant pressure prevails, or more accurately, by vapor pressure if the vessel is closed. Control of condensate temperature or vapor pressure is not so straightforward since the flow of the condensing vapor is the load and not the manipulated variable. The relationship between heat transfer and cool-. ant flow $W_C$ can be found simply by solving the equations developed earlier using constant temperature $T_v$ for the condensing vapor:

$$Q = \frac{T_v - T_{C1}}{1/UA + 1/2W_C C_C} \tag{9.15}$$

Notice the similarity between Eqs. (9.15) and (9.8). This indicates that the response of heat, transfer to coolant flow will be identical to the curve $W_H C_H / UA = \infty$ of Fig. 9.3. For the manipulation of coolant flow, then, the nonlinearity problem is just as severe as it is when there is no phase change.

Under conditions of constant condensate temperature, the heat transfer rate is entirely dependent, upon coolant flow. If coolant flow is maintained constant, bypassing part of the vapor around the condenser will not affect' the rate of heat transfer unless the condensate becomes appreciably subcooled. Under these conditions, the condenser begins to act more like the liquid-liquid heat exchanger, which is described in Fig. 9.6.

The most effective way to control a condenser is to vary its heat transfer area. This is done by manipulating the flow of condensate so as to partially flood the condenser, thereby reducing the surface available for condensation. The level of condensate within the condenser is an indication of the heat load on the process. The system is described in Fig. 9.7.

To be sure, a certain amount of subcooling always takes place, in whatever area is not used for condensing. The amount. of subcooling varies

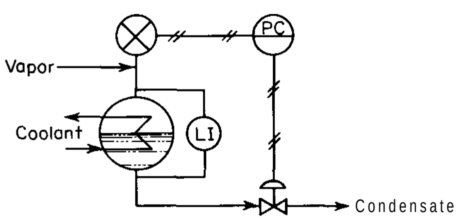



***FIG 9.7. The heat transfer area available for condensation can be changed by manipulating the flow of condensate.***

with the flow of vapor, so condensate temperature cannot be used for control. If the heat transfer coefficients for condensing and subcooling mere equal, this system would have no control over vapor pressure at all, because heat transfer rate would not depend on liquid level. E'ortunately, heat transfer coefficients of condensing vapors are generally much greater than those of condensate, particularly if the velocity of the condensate is low, as it would be in the shell of the condenser.

On the other hand, manipulation of liquid level is a slow process, with 90" phase lag between valve position and heat transfer area. Since vapor pressure is a fast measurement, however, the loop generally performs well dynamically, except perhaps for severe load changes requiring the condenser to be filled or emptied. Linearity and rangeability are important factors in its favor.

## COMBUSTION CONTROL

When a fuel burns, the products of combustion, along with whatever other vapors may be present, are raised to a flame temperature determined by the energy content of the fuel. Since heat of combustion is rated in Btu/lb or Btu/cu ft, the actual quantity of fuel involved does not affect its flame temperature. To estimate the flame temperature, the sensible heat of either the combustion products or the fuel and air may be used, since the energy balance can be satisfied in either case. The rate of heat generated by the combustion of a given mass flow of fuel $W_F$, whose heat, of combustion is $H_C$, is

$$Q = W_F H_C \tag{9.16}$$

This flow of heat must equal what is necessary to raise the flows of fuel and air, $W_A$, to the flame temperature $T$:

$$Q = W_F C_F (T - T_F) + W_A C_A (T - T_A) \tag{9.17}$$

The terms $C_F$, $T_F$, $C_A$, and $T_A$ represent the average specific heat and the inlet temperature of fuel and air, respectively.

To ensure complete combustion, a specified ratio of air to fuel, $K_A$, must be selected, based upon the chemical constituents in the fuel. Substitution of $K_A$ for $W_A/W_F$ will allow the solution of Eqs. (9.16) and (9.17) for flame temperature:

$$T = \frac{H_C + C_F T_F + K_A C_A T_A}{C_F + K_A C_A} \tag{9.18}$$

Equation (9.18) must be recognized as being valid only for conditions where there is no excess fuel. Because fuel is more expensive than air, and because incomplete combustion can cause soot and carbon monoxide, furnaces are invariably operated with excess air. But it should be apparent that the maximum flame temperature will only be reached with no excess of either. Equation (9.18) also gives an indication of the effect air temperature can have on the flame. The nitrogen, of course, does not participate in combustion and acts as a diluent, reducing the flame temperature. If oxygen is used instead of air, $K_A$ can then be reduced fivefold, producing a sizable effect on flame temperature.

The flame temperature estimated in Eq. (9.18) will be higher than what would actually be measured, because some of the energy contained in the combustion products partially ionizes them. This ionization increases with temperature, but the energy is recovered when the ions cool sufficiently to recombine into molecules.

### Control of Fuel and Air

Since the temperature of the flame falls with either an excess or a deficiency of air, it is not a particularly good controlled variable. The most universally used indication of combustion efficiency is a measurement of oxygen content in the combustion products. The amount of excess air required to ensure complete combustion depends on the nature of the fuel. Natural gas, for example, can be burned efficiently with 8 to 10 percent excess air (1.6 to 2 percent excess oxygen), while oil requires 10 to 15 percent excess air (2 to 3 percent excess oxygen) and coal, 18 to 25 percent excess air (3.5 to 5 percent excess oxygen). The reasons for the differences are the relative state of the fuel and the amount of noncombustibles present.

Since the amount of heat transferred by radiation varies with the fourth power of the absolute flame temperature, the greatest efficiency will always be realized with maximum flame temperature. But the distribution of the heat is also important. Increasing the amount of excess air will reduce the flame temperature, thereby reducing the heat transfer rate in the vicinity of the burner. Since the net flow of thermal power into the system has not changed, the rate of heat transfer farther away from the burner tends to increase.

Safety dictates certain operating precautions for fuel-air controls. A deficiency of air can allow fuel to accumulate in the furnace, which upon ignition, may explode. Care must be taken, therefore, to ensure that the fuel rate never exceeds what is permissible for given conditions of air flow. Fuel and air flow both can be set from a master firing-rate control, but automatic selection is necessary to achieve this safety feature. A complete control system for control of fuel and air is shown in Fig. 9.8.'

Notice that the fuel-air ratio is adjusted through manipulation of the span of the air measurement by the oxygen controller. Normally the set point would be adjusted, but in order for the selection system to operate,

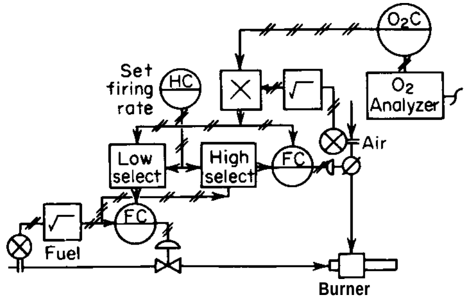



*FIG 9.8. This system automatically protects against a deficiency of air.*

the set points of both controllers must have the same values. If air flow is lost, its measurement is preferentially selected by the fuel flow controller for a set point. If fuel flow is higher than called for, on the other hand, its measurement is automatically selected to set the air flow. The furnace is thereby protected not only from blower or controller failure, but also from lags in the set-point response of either loop.

### Fired Heaters

Heaters fired directly by the combustion of gas or oil are common in refineries, particularly where high temperatures are needed. The control problem is one of manipulating fuel rate to achieve the desired exit temperature of the heated fluid. Air is usually inspirated into the burner in proportion to the fuel, therefore regulation of its flow is inherent. But because of the many hundreds of feet of tubing enclosed within a heater, dead time is in the order of minutes, varying with flow.

Where sudden load changes are encountered and close control is necessary, feedforward systems have proven effective. The heat-balance equation is similar to that solved for the heat exchanger in Fig. 8.4. The only difference is that fuel flow is manipulated instead of steam and heat of combustion takes the place of latent heat of vaporization. Although the loss of heat out the stack may be significant, it varies directly with load and can be readily accommodated by the action of the feedback temperature controller, as is done in Fig. 8.17.

Should the fuel be gas at a variable temperature or pressure, computation of mass flow may be warranted, particularly if these variations are frequent or rapid.

## STEAM-PLANT CONTROL SYSTEMS

In order to successfully apply controls to steam generation, a thorough familiarity with its thermodynamic properties is essential. The most important point to remember is that steam is valued principally for its heat content or enthalpy, of which temperature is a measure. If the vapor is to remain in equilibrium with the liquid, greater enthalpy can only be brought about at increased pressure. If pressure is limited and enthalpy increased. the vapor must be removed from the presence of the liquid and superheated.

Steam is also superheated by passing through an orifice or pressure-reducing valve, since in theory, no enthalpy is lost across a throttling device. Thus the pressure of the steam would drop while the temperature remained virtually constant.

The mass flow of steam may be measured with an ordinary orifice meter, but the reading must be corrected, if pressure or temperature deviate from the conditions under which calibration was specified. In the case of saturated steam, pressure and temperature are not independent of one another, so either one is capable of indicating density. It so happens, however, that pressure is a linear function of density, with an intercept of 0 psig:

$$W = k\sqrt{hp}$$

where $W$ = mass flow
$k$ = orifice scaling factor
$h$ = differential pressure across the flowmeter
$p$ = static gage pressure

The density of superheated steam varies inversely with temperature and directly with pressure to make the mass flow calculation[2] more complicated and less accurate. But if a steam flowmeter is used to indicate the actual delivery of thermal power, an interesting phenomenon appears: temperature causes the enthalpy of superheated steam to vary in a way which offsets its effect upon density. Thus thermal power $Q$ only varies with differential and pressure:

$$Q = WH = kH_0\sqrt{hp}$$

Coefficients $H$ and $H_0$ represent steam enthalpy at flowing and calibration conditions, respectively.

### Drum Boilers

In a drum boiler, water is circulated at a rapid rate upward through the furnace tubes, in which it partially vaporizes. Upon reaching the drum, the liquid disengages the vapor and returns through relatively cool downcomers to the bottom of the furnace to begin another pass upward. The most characteristic feature of drum boilers is the difficulty of controlling the level of liquid *in* the steam drum. A feedforward-feedback system for its control was described briefly in Chap. 8.

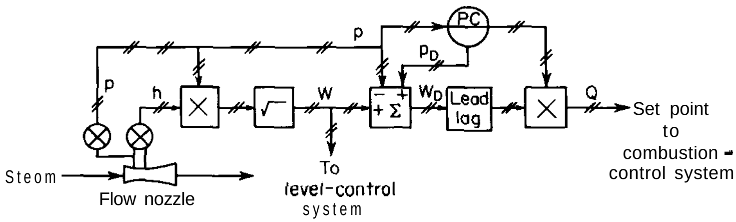
The vessel is in a high state of turbulence and naturally exhibits a certain hydraulic resonance, which was described in Chap. 3. But liquid level is affected thermodynamically as well. For example, the sudden introduction of feedwater below its boiling point can momentarily reduce the heat content of the vapor-liquid mixture in the tubes, causing bubbles to collapse, and reducing the apparent liquid level. Thus the response of the level-control loop has a tendency to start in the wrong direction. This property is called "phase shifting" and is similar to dead time in that it produces phase lag without attenuation. The result is that the period of the liquid-level loop is ordinarily several minutes, although its natural hydraulic resonance may be only a few seconds long.

If the boiler must operate under varying steam pressure, the calibration of the liquid-level transmitter will vary with steam density.[3] But pressure has a transient effect too. If a load increase (withdrawal of steam) is sufficient to cause drum pressure to fall, some of the water in the tubes will flash into steam, temporarily increasing the flow of both liquid and vapor into the drum. This effect is called '(swell," because it causes a transient rise in liquid level, even though the rate of steam withdrawal may momentarily exceed that of feedwater flow. Conversely, upon a pressure increase, the liquid level tends to "shrink." This effect is more prominent in low-pressure boilers, because of the greater difference between the densities of steam and water. The most favored method of coping with "shrink" and "swell" is to ignore them, by letting the forward loop carry the load, while maintaining loose settings on the level controller. Drum-level controllers customarily require a proportional band near 100 percent and several minutes' reset time.

Pressure in a saturated or even a superheated boiler is a measure of the amount of energy stored therein. The flow of steam from the plant is usually at the demand of the user. Pressure can only be maintained, then, if the flow of energy into the boiler equals the rate of withdrawal. Since the drum-level control system admits feedwater at a rate equal to the flow of steam, the pressure-control system is left to manipulate the input of thermal power. To achieve high performance control, a feedforward loop should be used to set firing rate proportional to steam flow.

Steam flow is a measure of thermal power and is affected by firing rate. If it alone is used to set firing rate, a positive feedback loop will be formed, a pitfall that was mentioned under decoupling systems in Chap. 7. What is really needed is a steam-flow demand signal. The demanded steam flow $W_D$ is the measured steam flow less the rate of loss of boiler contents:

$$W_D = W - V \frac{dp}{dt}$$




*FZG 9.9. The ratio of firing rate to steam flow demand is automatically adjusted to maintain pressure in the boiler.*

where $V$ represents the steam volume of the drum. It is more convenient, however, to use pressure error instead of $dp/dt$, since differentiation is a clumsy operation and the error signal is already available at the pressure controller:

$$W_D = W + K(p_D - p) \tag{9.19}$$

The coefficient $K$ is adjusted to the characteristics of the boiler; $p_D$ represents the set point of the pressure controller.

In the steady state, $p = p_D$ and $W_D = W$. Sudden opening of a steam valve will raise $W$ but drop p. Until the effect of a firing-rate change is felt, pressure will fall to a lower level and steam flow will return to its previous value. But the existence of a pressure error directly proportional to the desired increment in steam flow will maintain the higher level of firing. The pressure is only restored when steam flow is raised to its demanded value. Figure 9.9 outlines the system together with the feedback loop.

When steam is superheated, its enthalpy becomes a function of temperature and is relatively independent of pressure. Steam-temperature control is often sought by redistributing the combustion gases in the furnace between the saturated and superheated tubes. But redistribution alone will not materially affect the energy content of the steam. So feedwater is sprayed into the steam line between the superheater and the temperature bulb. Control at this point is quite effective because the thermal process is one of mixing.

### Once-through Boilers

The operation of a once-through boiler is more readily analyzed because it is dynamically continuous, not broken in half by a drum. Feedwater is pumped into the tubes at one end and superheated steam withdrawn at the other. There is no recirculation. There is also no liquid level to measure. In subcritical boilers, there is a transition from liquid to vapor somewhere in the tubes, but exactly where is of little consequence. In supercritical boilers, there is no phase change, hence no point of transition. A once-through boiler is shown schematically in Fig. 9.10.

Under normal operation, three controlled variables are of primary importance: steam pressure,. steam temperature, and thermal power. The first two are to be regulated, the third is the heat load on the plant. They are controlled by the manipulation of firing rate, feedwater flow, and steam-valve position.

These variables interact with each other to the extent that poor performance will result if the three loops are operated independently. To appreciate the mechanism of this interaction, consider how the controlled variables would respond to step changes in each of the manipulated variables:[4]

1. An increase in firing rate would increase both thermal power and steam temperature. With the steam valve in a fixed position, upstream pressure would increase, because thermal power has been shown to vary with pressure and differential across an orifice.

2. An increase in feedwater flow will cause an increase in steam flow, but thermal power will not change. Thus steam pressure will not change either. Since steam enthalpy is thermal power divided by flow, an increase in steam flow will cause steam temperature to fall.

3. As the steam valve is opened farther, pressure will fall to a new equilibrium value, during which a certain amount of steam and energy will have been released. But when the new steady state is reached, steam flow and thermal power must return to their original values, since feedwater flow and firing rate have not changed.

The above responses can be simply represented by a dimensionless matrix, without going into a detailed calculation of process gains. Let Q represent thermal power; $p$, steam pressure; and $T$, steam temperature. $W_F$ will be firing rate; $W_W$, feedwater flow; and $m$, position of the steam valve :

| | **Q** | $p$ | $T$ |
|---|---|---|---|
| $W_F$ | +1 | +1 | +1 |
| $m$ | O-l | | 0 |
| $W_W$ | 0 | 0 | -1 |

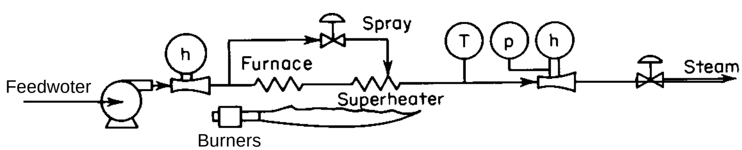



*FIG 9.10. In a once-through boiler, feedwater is conducted through tubes all the way to the steam valve.*

(Note that the sum of the columns for the regulated variables is zero, and for the command variable is 1.0.)

This is a matrix of open-loop gains whose actual values have not been inserted. The intent is to show that half-coupling exists between Q, $p$, and $T$. But normalization according to the procedure outlined in Chap. 7 yields a unit diagonal, with all other elements zero. This proves the absence of full coupling, and makes pairing obvious. The detailed equations for decoupling of the half-coupled loops follow.

Thermal efficiency $E$ and heat of combustion $H_C$ determine the firing rate required to satisfy a given load:

$$W_F = \frac{Q}{H_C E} \tag{9.20}$$

Feedwater enthalpy $H_W$ must be included as well as steam enthalpy $H_S$ to determine the feedwater requirements:

$$W_W = \frac{Q}{H_S - H_W}$$

Feedwater temperature $T_W$ and specific heat $C_W$ may be substituted for its enthalpy:

$$W_W = \frac{Q}{H_S - C_W T_W} \tag{9.21}$$

Because the points of steam temperature measurement and feedwater flow manipulation are widely separated, control is somewhat difficult. To ease the situation, a small amount, of feedwater is bypassed around the furnace and admitted as spray between the superheater and the temperature bulb. If this stream is taken as a part of, and not in addition to, the set flow of feedwater, its variation will not change the long-term steam enthalpy, and is only capable of producing a transient effect.

Since steam pressure is to manipulate the position of a valve immediately downstream of its measurement, control is rather easy. Consequently feedforward will not be used here because it is not really warranted.

The unknown parameters in the feedforward equations are $H_C$ and $E$. These must be found by means of feedback. All other terms are either measurable or constant. A control system designed according to the rules given in Chaps. 7 and 8 is shown in Fig. 9.11.

It is not possible to use reset in two controllers which operate on the same input. So proportional plus derivative is used to manipulate the spray valve. In this way, the valve will operate around a position determined by the bias set into the controller. The lead-lag unit is used to

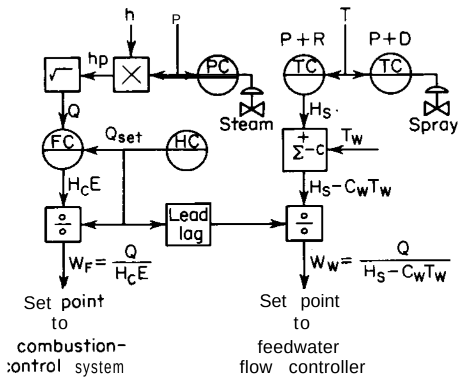



*FIG 9.11. The control system for a once-through boiler must be as closely knit as the process itself.*

match the feedwater to temperature response with that of firing rate to temperature. Since both manipulated variables encounter the same dead time related to the velocity through the tubes, their responses are similar. Fortunately, the action of the spray controller also helps make up for any differences in response.

Because this dead time varies inversely with flow, hence with thermal power, the process exhibits variable dynamic gain. But the feedforward signal Q is a multiplier in both power and temperature loops, providing gain adaptation.

### Steam-driven Turbines

When the output from a steam plant is to be used to drive a turbine generator, the latter's characteristics must be taken into account. In order to generate electric power, the turbine needs an equivalent amount of thermal power. Steam is merely a vehicle which transports the power from the burners to the turbine shaft. The actual flow of steam is not significant,, except in that it determines what the temperature and pressure must bc for a given value of power.

The turbine acts like a fixed resistance to the flow of steam, with a downstream pressure determined by conditions in the condenser, usually several inches of vacuum. Thus the pressure drop is nearly equal to the inlet pressure. Thermal power is proportional to the square root of the product of differential and static pressures, which in this case is the inlet pressure $p_1$ itself:

$$Q = kp_1 \tag{9.22}$$

The speed of a turbine generator is set by a governor which positions the throttle valve with proportional action. The throttle valve is between the superheater and the turbine, in the place of the steam valve shown in Fig. 9.10. Its position changes with variation in turbine speed, which is measured as the frequency of the generated voltage, or with a set-point adjustment. The steam-pressure controller,. then, in a sense sets the throttle-valve position by setting the governor.

The fact that the throttle-valve position can change with frequency does complicate the control problem somewhat,. Should frequency fall, more power is needed to restore it. But opening the throttle valve does not produce more power. So if steam pressure is to be maintained, the firing rate must be made to vary with frequency. To accomplish this, a frequency-deviation signal is usually added to the output of the set station where the desired power output is introduced.

## PUMPS AND COMPRESSORS

There are many ways to control the flow and pressure of streams discharging from pumps and compressors, but all are not equally efficient. Throttling a valve in the discharge line of a centrifugal pump may be convenient, but it may also be wasteful if long periods of low flow are encountered. And it cannot be done at all with a positive-displacement pump, because pressure ratings would be exceeded.

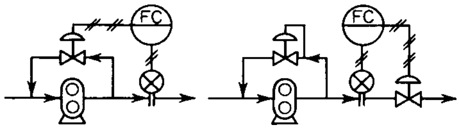
The point is, that the type and size of a prime mover dictate the means which are to be used to control it. This is particularly true of compressors which are capable of exhibiting unstable characteristics.

### Positive-displacement Pumps

There are two principal classes of pumps: positive-displacement and centrifugal. In positive-displacement pumps, a given volume of fluid is mechanically forced from the suction port to the discharge with every rotation of the shaft. In reciprocating pumps, this is done in a periodic fashion, such that outflow pulsates; if multiple cylinders are used, they are phased so as to diminish the amplitude and period of the pulsations, smoothing the flow. Often an air chamber is attached to the discharge line to help absorb these pulsations.

If either the stroke or the speed of a reciprocating pump is adjustable, it can be used for metering an accurate amount of fluid. The accuracy of these "metering pumps" requires that they be free from leakage, particularly backflow from discharge to suction. This means that their valves must be tight-sealing, and the fluid free from particles which might interfere with their action. The fluid must also be incompressible; these pumps often vapor-lock when the fluid contains dissolved or entrained gas. When used with clean liquids, metering pumps are valuable for flow control, particularly where high discharge pressures are encoun-

*FIG 9.12.* ***Shown are two methods for controlling the flow from a gear or vane pump.***




tered. They are available with a pneumatic operator to adjust the stroke automatically from a set station or a primary controller.

Other positive-displacement pumps include those which move liquid with rotating gears or vanes. Their output is continuous, although noisy. But as the name implies, positive-displacement pumps must he allowed to discharge their rated flow. They must always be protected by a relief valve connected from discharge to suction; otherwise, if the discharge line is restricted, high enough pressure can be developed to rupture the line or overload the motor.

Control of flow from a gear or vane pump may be achieved by manipulating a bypass or by regulating the discharge pressure with a bypass, both of which are shown in Fig. 9.12.

### Centrifugal Pumps

Centrifugal pumps exhibit slippage. They impart momentum to the fluid, which is converted to velocity head. At no-flow conditions, rotation of an impeller of a given diameter at constant speed produces the maximum head which the pump is capable of delivering. As flow increases, the head falls by an amount equivalent to frictional losses within the pump itself. It should be noted that the pressure which a centrifugal pump is capable of delivering varies with the density of the fluid, since pressure equals head times density.

The characteristics of a centrifugal pump are usually plotted as head vs. flow on linear coordinates. Contours of speed, horsepower, and efficiency are often included. The choice of linear coordinates is unfortunate in the sense that only the horsepower contours are straight lines. If speed is an important variable, plotting head vs. flow squared will yield straight contours. The two plots are compared in Fig. 9.13. Discharge pressure $p$ varies with speed $N$, flow $F$, and density $\rho$ as follows:

$$p = \rho(k_1N^2 \quad k_2F^2) \tag{9.23}$$

The coefficients $k_1$ and $k_2$ are functions of the mechanical parameters of the pump, i.e., impeller diameter, clearance, etc.

Small pumps are usually driven by constant-speed electrical motors. Flow may be controlled by throttling a valve in the discharge line. The suction should never be throttled, because a centrifugal pump requires a positive suction head to operate. Low suction pressure causes cavitation and loss of flow.

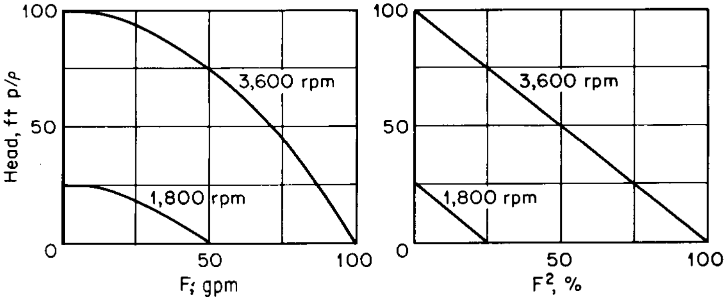



*FIG 9.13. By proper choice of coordinates, speed contours for a centrifugal pump are made into straight lines.*

The hydraulic horsepower, HHP, imparted to a fluid is defined as the product' of the pressure (not head) developed and the flow delivered:

$$\text{HHP} = \frac{Fp}{1.714} \tag{9.24}$$

The unit's of flow are gpm and the units of pressure are psi. If flow is shut off, or if pressure is lost, HHP falls to zero. But the pump nevertheless absorbs power from the drive, which indicates that its efficiency has gone to zero. As is the case with most machinery, centrifugal pumps operate most efficiently in the middle of their pressure-flow range. The point of maximum HHP can be found by multiplying the right' side of Eq. (9.23) by flow to give HHP, then differentiating with respect to *flow.* Solving the equation for zero derivative gives the optimum flow $F_o$:

$$F_o = N\sqrt{\frac{k_1}{3k_2}}$$

For the characteristics given in Fig. 9.13, maximum HHP would occur at 58 gpm for 3,600 rpm, and 29 gpm for 1,800 rpm. This confirms that speed should be varied with flow if the most efficient conditions are to be maintained. Manipulation of speed is therefore economically justified with large pumps.

Many pumps are driven by steam turbines, which are equipped with governors capable of being set pneumatically or electrically. Others are driven by constant-speed electric motors through hydraulic or magnetic couplings. The speed of the pump shaft can usually be varied from 0 to 100 percent of motor speed by adjusting the degree of coupling.

To fully assess the characteristics of speed manipulation, the flow process must be combined with the pump parameters. Flow through a process whose flow coefficient is $k_3$ is given by

$$F^2 = k_3 \frac{p}{\rho} \tag{9.25}$$

Substituting for $F^2$ in Eq. (9.23) yields the response of discharge pressure to speed:

$$p = \frac{k_1 N^2 \rho}{1 + k_2 k_3} \tag{9.26}$$

Or, solving for flow,

$$F = N \left( \frac{k_1}{k_2 + 1/k_3} \right)^{1/2} \tag{9.27}$$

Manipulating speed is very much like positioning a linear valve.

Note that HHP varies with speed cubed. This is favorable for variable speed couplings, in that the load placed on them is quite low when speed is gradually increased from zero.

### Compressor Control

Reciprocating compressors are considered in the same light as pumps, with the exception that their outflow can be measured easily. Their higher speed and the compressibility of the fluid help to reduce pulsations. Control methods previously described for gear and vane pumps (Fig. 9.12) can also be used for reciprocating compressors.

Multicylinder compressors can have their flow reduced by "unloading" some of the cylinders sequentially. This consists of holding the suction valve open during the entire stroke, effectively disabling the cylinder. Most multicylinder compressors are equipped with solenoid or pneumatic "unloaders" which may be operated from the output' of a controller. There is a discrete number of flow conditions for such a compressor, so limit cycling cannot be avoided.

Again, manipulation of speed where practicable has the advantage of being both continuous and efficient'.

Centrifugal or turbocompressors are analogous to centrifugal pumps in principle. But the compressibility of the fluid being transported affects their characteristics considerably. While discharge head was plotted against flow for a pump, compression ratio is typically the ordinate used to display compressor characteristics. But the most significant property of a centrifugal compressor is the presence of an area of instability. Certain combinations of low flow and high pressure fall into what is known as the "surge" region. Figure 9.14 shows where this region lies.

In the surge region, a compressor exhibits positive feedback: decreasing flow causes pressure to fall until it is less than that in the discharge line.

Then a momentary reversal of flow occurs until the line pressure falls below what is being developed. These flow reversals can develop into pulsations violent enough to cause severe damage.

It is obviously essential to stay clear of the surge region, yet the greatest power efficiency is ordinarily realized directly adjacent to it. So it is important to outline this region carefully.

The ordinate for both plots in Fig. 9.14 is discharge-suction pressure ratio $p_d/p_s$, starting from 1.0, or $p_d/p_s - 1.0$; the abscissa $F$ is the volumetric flow at suction conditions. The locus of the surge line varies with suction temperature $T_s$ and can be represented by the equation

$$\frac{p_d}{p_s} - 1.0 = \frac{k_4F^2}{T_s} \tag{9.28}$$

Interestingly, a simple transposition allows a remarkably easy calculation of surge conditions. Multiplying both sides of Eq. (9.28) by $p_s$ gives

$$p_d - p_s = k_4F^2 \frac{p_s}{T_s}$$

If a conventional orifice meter is inserted in the suction line, a differential, $h$, is developed relative to the volumetric flow:

$$h = kF^2 \frac{p_s}{T_s}$$

Combining the last two expressions yields a relation between compressor differential and flow differential which describes the surge line:

$$h = K(p_d - p_s) \tag{9.29}$$

Constant-speed compressors operate on one of the curves shown in Fig. 9.14. To control pressure, a valve in the suction or discharge may

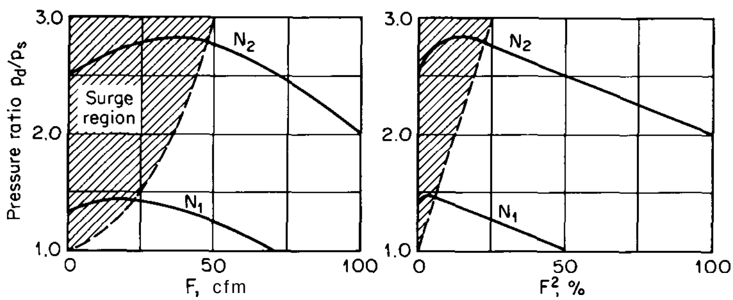



***FIG 9.14. Plotting pressure ratio against flow squared allows easier identification of the surge region.***

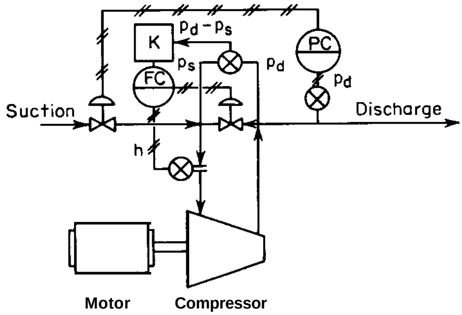



*FIG 9.15. The bypass valve only opens when the flow drawn from the compressor approaches the minimum calculated for surge protection.*

be throttled, but only to a point. If the load should call for flow falling within the surge region, part of the output of the compressor must be bypassed to the suction to maintain stable conditions.

A control system[5] which calculates the minimum flow tolerable for current values of suction and discharge pressure is illustrated in Fig. 9.15. A differential pressure transmitter senses $p_d - p_s$, and its output is multiplied by coefficient $K$ in a ratio station, whose output in turn sets the flow controller. As long as the flow measurement exceeds the minimum value calculated above, the flow controller will leave the bypass valve closed. The ratio $K$ should be set to allow some margin of safety, and the flow controller should be protected from overshoot with an antiwindup switch. As the surge line is approached, the flow controller will begin to open the bypass, which increases flow, reduces discharge pressure, and increases suction pressure all at the same time, positively preventing surge.

Rather than closing a valve in the suction line, many large compressors have an arrangement of inlet guide vanes, which may be manipulated for more efficient throttling by reducing entrance losses. In addition, the guide vanes shift the surge line to the left as they are closed, which increases the working range of the compressor.

If upon a reduction in load, the speed of a compressor is reduced, $p_d - p_s$ mill also decrease, and with it, the minimum flow requirement for surge protection. Thus with speed as the manipulated variable, the bypass valve usually does not need to be opened. Together with the reasons mentioned under centrifugal pumps, this makes manipulation of speed economically attractive.

Control may be exercised over suction as well as discharge pressure or volumetric or weight flow. Or selection may be made between two variables as was illustrated in Fig. 6.15. These variables are all similar in nature, particularly if flow is in the differential form. As with a centrifugal pump, flow varies linearly with speed, pressure with speed squared.

## REFERENCES

1. Manter, D., and R. Tressler: A Coal-Air Ratio Control System for a Cyclone Fired Steam Generator, ISA Paper 1-CI-61.
2. Shinskey, F. G.: Analog Computing Control for On-line Applications, *Control Eng.*, November, 1962.
3. Roos, N. H.: Level Measurement in Pressurized Vessels, *ISA Journal*, May, 1963.
4. Adams, J., D. R. Clark, J. R. Louis, and J. P. Spanbauer: Mathematical Modeling of Once-through Boiler Dynamics, *IEEE Trans. on Power Apparatus and Systems*, February, 1965.
5. The Foxboro Company, Turbo Compressor Anti-Surge Control, *Application Engineering Data* 000-25, June, 1965.

## PROBLEMS

**9.1** Design a simplified decoupling system for the control of temperature and flow of a mixture of hot and cold water, whose flow rates are the manipulated variables.

**9.2** Feed to a reactor is being preheated countercurrently by oil at a temperature of 500°F. The feed is a liquid, entering the exchanger at 200°F and leaving at a controlled temperature of 400°F. Under normal conditions, the feed rate is 100 lb/min and oil flow is 400 lb/min; the heat capacity of each is 0.8 Btu/(lb)(°F). Estimate the change in oil flow required to maintain control as feed flow varies $\pm 20$ percent from normal. Repeat for variations of $\pm 40$°F in feed inlet temperature.

9.3 Calculate the process gain $dT_{C2}/dW_H$ for the normal and the -20 percent flow conditions encountered in Prob. 9.2. Then calculate the gain product of process and valve $dT_{C2}/dm$ using an equal-percentage characteristic (let $dW_H/dm = kW_H$).

9.4 The temperature of condensate leaving a condenser is being controlled by manipulating the flow of cooling water. Suppose $U \simeq kW_C$; derive the variation of $Q$ with $W_C$. What are the limitations of this approximation?

9.5 How is steam temperature controlled in a drum boiler? What feature of a once-through boiler enables manipulation of feedwater flow to control temperature?

9.6 Find the values of coefficients $k_1$ and $k_2$ for the pump characteristics given in Fig. 9.13. If 50 gpm is being drawn as load, what speed is required to control the discharge head at 50 ft? Calculate the HHP required at that speed and also the HHP required to deliver the same flow at 3,600 rpm. The fluid is water.

# Controlling Chemical Reactions

## CHAPTER 10

The heart of a chemical process is a reaction in which several feed stocks are combined to make one or more products of greater value. Before the plant is designed, process engineers have to determine whether the reaction will proceed at a favorable rate, with sufficient conversion into the principal product to make the operation profitable. Once the optimum conditions for the conduct of the reaction have been established, the control system must be designed to maintain them. But there are many factors affecting the rate of the principal and side reactions, which must be appreciated before an intelligent design can be made.

To a great extent, the success of a reactor control system depends upon how well the reactor is designed. That is, it is possible to design a reactor so that it is unstable no matter what kind of control system is employed. For this reason, a few words will be directed toward reactor design, both to serve as a guide to the process engineer involved in design and to facilitate recognition of an unstable reactor.

In most plants, the reactor is the starting point for the process, and its production rate sets the load. This is fortunate in that, the reactor is then free from the rapid or unpredictable load upset's which other units encounter. But the reaction system ought to be well regulated; if it is not, it will generate disturbances that will be propagated through the rest of the plant.

## PRINCIPLES GOVERNING THE CONDUCT OF REACTIONS

Unless this section is thoroughly understood, much of the significance of the design procedures that follow will be lost. Many processes can tolerate poorly designed control systems, particularly if they are protected from a rapidly varying load. But most chemical reactors will not. Some need no load upset to break into oscillations sufficiently violent to ruin product, destroy catalyst, damage equipment, and endanger life.

It is therefore not surprising to find many reactors operated in manual control, simply because their automatic systems are not capable of doing their job. Control is usually unsatisfactory in manual, too, but operators tend to have more confidence in themselves than in poorly designed control systems.

### Chemical Equilibrium

A great many chemical reactions are reversible. That is, under certain conditions it is possible to start with the products and make a measurable amount of the reactants. In these cases an equilibrium state can exist in which the reaction comes to a standstill because the forward and reverse rates are equal. This equilibrium point determines how much of the reactants can be converted into products and also what conditions will favor conversion.

As an example, consider the vapor-phase oxidation of sulfur dioxide into sulfur trioxide, one of the steps in the manufacture of sulfuric acid:

$$SO_2 + \tfrac{1}{2}O_2 \overset{K}{\rightleftharpoons} SO, + \text{heat} \tag{10.1}$$

A certain ratio of product to reactant concentration can be reached which will bring about equilibrium. This ratio is identified by the equilibrium constant $K$, such that:

$$K = \frac{[SO_3]}{[SO_2][O_2]^{1/2}} \tag{10.2}$$

Every reversible reaction has an equilibrium constant, which is a function of temperature and catalyst.

The existence of an equilibrium state discloses that a certain fraction of the reactants must be withdrawn along with the product. This places a limitation on the conversion which can be achieved within a reactor. But several tactics can be used to improve conversion, which can be deduced from Eqs. (10.1) and (10.2):

1. Reduction in product concentration will permit reduction in reactant concentration. Thus, if product can be removed from the reaction zone by condensation, for example, more conversion is possible. Conversion can approach 100 percent if products are easily separated from the reactants as gases or solids from a liquid-phase reaction or as liquids or solids from a vapor-phase reaction.

2. An increase in reactant concentration will also increase conversion. Notice in Eq. (10.2) that $SO_3$ concentration is a function of both the $SO_2$ and $O_2$ concentrations. An increase in either $SO_2$ or $O_2$ will promote conversion; therefore either reactant can be used in excess to augment conversion of the other.

3. For the particular example being used, Eq. (10.1) indicates that the total moles (hence volume) of the reactants (1.5) is greater than that of the product (1). Therefore an increase in pressure will tend to increase the denominator of Eq. (10.2) more than the numerator; this will enhance conversion. This particular reaction should therefore be conducted under pressure.

4. The evolution of heat indicated by Eq. (10.1) also affects equilibrium. Just as separation of product from the reactants will promote conversion, removal of heat will do likewise. In fact', high temperature favors the reverse reaction. Consequently subject reaction should be conducted at low temperature with continuous removal of heat.

A catalyst is a substance that has the property of changing the equilibrium constant without actually taking part in the reaction. It may be nothing more than a porous surface onto which the reactants are adsorbed. Or it may serve to establish a token concentration of an intermediate product, without which the final product might not be formed. Or it may serve to provide the correct environment, e.g., acidity. Light even catalyzes some reactions.

Catalyst may be packed in a fixed bed within the reactor. Uniformly small particles may also be supported by the upward velocity of the reactant stream (gas or liquid), in which case it is called a "fluidized bed." Solid catalyst may also be dissolved or suspended in a liquid reaction media, then separated from the products and recycled. Metal catalysts may be made into screens or other shapes across which the reactants flow. It should be remembered, however, that the reaction takes place on the surface of the catalyst; if heat is evolved, cooling should be applied there, or the catalyst could bc destroyed or deactivated. Most catalysts also become deactivated due to fouling of the surface with by-products and contamination by impurities in the feed stock, called poisons. The conversion within a reactor depends on the active surface area of the catalyst, which can be time-variant. A classic optimization problem often encountered involves deciding on the most efficient schedule for replacing or reactivating catalyst.

### Reaction Rate

Equilibrium occurs when the rates of forward and reverse reactions are equal. These rates are proportional to concentrations of reactants and products respectively. Let $k_f$ and $k_r$ be designated as forward- and reverse-rate coefficients. Then at equilibrium,

$$k_f[SO_2][O_2]^{1/2} = k_r[SO_3] \tag{10.3}$$

From Eqs. (10.2) and (10.3), the equilibrium constant' is the ratio of the forward- to reverse-rate coefficients:

$$K = \frac{k_f}{k_r} \tag{10.4}$$

The rate of reaction can be identified as the rate of change of concentration of one of the reactants or products in approaching equilibrium:

$$-\frac{d[SO_2]}{dt} = \frac{d[SO_3]}{dt} = k_f[SO_2][O_2]^{1/2} - k_r[SO_3]$$

This forward reaction is 1.5-order, indicated by the sum of the exponents, while the reverse reaction is first-order. If one of the reactants, for example, $O_2$, is in considerable excess, the rate of reaction will depend principally on the concentration of the other, and therefore will approach first-order. This is, in fact, a very common occurrence, so the majority of reactions can be treated as first-order. Furthermore, if any of the four steps previously given to promote conversion are employed, the rate of the reverse reaction is usually negligible. So a general equation may be applied to describe the rate of most reactions, relative to the concentration $x$ of the controlling reactant:

$$-\frac{dx}{dt} = kx \tag{10.5}$$

A *batch* of reactant will change its concentration exponentially with time from an initial value $x_0$ to a current value $x$, according to the integration of Eq. (10.5):

$$\int_{x_0}^{x} \frac{dx}{x} = \int_0^t -k\,dt$$

$$\ln x - \ln x_0 = kt$$

Converting the natural logarithms to exponents yields:

$$x = x_0 e^{-kt}$$

Note that the time constant for a first-order reaction is $1/k$; thus the units of the reaction-rate coefficient are in inverse time.

Fractional conversion of reactant into product will be identified as $y$, varying with time:

$$y = \frac{x_0 - x}{x_0} = 1 - e^{-kt} \tag{10.6}$$

In *a continuous plug-flow* reactor, the reaction mixture flows through a pipe without back-mixing. This type of reactor is dominated by dead time. The residence time of the reactants traveling through a volume $V$ at a flow $F$ is $V/F$. Thus the concentration at the exit of the vessel is

$$x = x_0 e^{-kV/F}$$

Conversion also varies exponentially with flow:

$$y = 1 - e^{-kV/F} \tag{10.7}$$

A *continuous back-mixed* reactor is one throughout which the reactant is uniformly distributed by means of agitation. It approaches a single-capacity system. Reaction rate is constant throughout, given by Eq. (10.5). The rate of consumption of the reactant is the volume of the vessel times the reaction rate, which equals the flow times the loss in concentration between inlet and outlet:'

$$\frac{V\,dx}{dt} = kVx = F(x_0 - x)$$

Solving for exit concentration,

$$x = \frac{x_0}{1 + kV/F}$$

Conversion in a back-mixed reactor varies inversely with residence time:

$$y = 1 - \frac{1}{1 + kV/F} = \frac{kV/F}{1 + kV/F} \tag{10.8}$$

A plug-flow reactor is dominated by dead time equal to the residence time. A back-mixed reactor, however, has a time constant which is a function of both $k$ and $V/F$. To illustrate this, a differential equation will be written to describe the dynamic material balance:

$$F(x_0 - x) - Vkx = V\frac{dx}{dt}$$

Solving for x,

$$x + \frac{V}{F + Vk}\,{}_0\frac{dx}{dt} = \frac{Fx_0}{F + Vk}$$

The concentration time constant is the coefficient of the second term:

$$\tau_x = \frac{V}{F + Vk} \tag{10.9}$$

The reaction-rate coefficient $k$ increases sharply with temperature—perhaps its most outstanding characteristic:

$$k = ae^{-E/RT} \tag{10.10}$$

where a and $E$ = constants peculiar to the reaction
$R$ = universal gas constant
$T$ = absolute temperature

To illustrate this strong dependency, $k$ is plotted vs. $T$ in Fig. 10.1 for a typical reaction whose parameters are: a = $e^{29}$ $\text{min}^{-1}$, ***E/R* = *20,000”*** Rankine.

The conversion in a plug-flow reactor varies with temperature in a double exponential, combining Eqs. (10.7) and (10.10). Conversion versus temperature for a back-mixed reactor is found by combining Eqs. (10.8) and (10.10). In Fig. 10.2, conversion is plotted against temperature for three values of *V/F* in both types of continuous reactors using values taken from Fig. 10.1. The plug-flow reactor delivers higher conversion than a back-mixed reactor operating under the same conditions.

Differentiation of the conversion vs. temperature relationships for each reactor yields expressions for their slopes. For the plug-flow reactor,

$$\frac{dy}{dT} = \frac{E}{RT^2}(1 - y) \ln (1 \quad y) \tag{10.11}$$

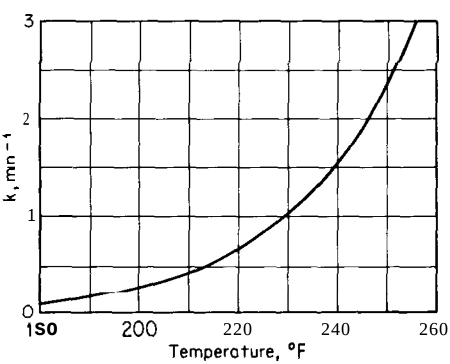



***FIG 10.1.* Reaction rate is profoundly influenced by temperature.**

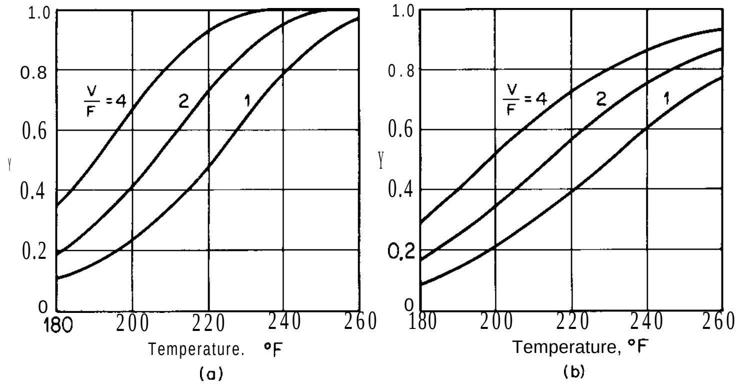



***FIG 10.2. Comparison of conversion vs. temperature characteristics for (a) plug-flow and (b) back-mixed reactors.***

And for the back-mixed reactor!

$$\frac{dy}{dT} = \frac{E}{RT^2} y(1 - y) \tag{10.12}$$

These expressions will be useful in determining temperature stability later in the chapter. The maximum slope of a conversion vs. temperature curve always occurs at $kV/F = 1$, which corresponds to 63 percent conversion in the plug-flow reactor and 50 percent in the back-mixed. For the $V/F = 1$ curves of Fig. 10.2, the maximum slopes are $1/64.7°F$ and $1/95°F$ for the two reactors, at 230°F.

The curves of Fig. 10.2 and Eqs. (10.11) and (10.12) describe steady-state conditions only, however. In a departure from steady state because of a heat transfer upset, temperature will change the reaction-rate coefficient in advance of a change in reactant concentration. Thus the reaction rate will increase with temperature above the new steady-state level until reactant concentration is accordingly reduced. A partial derivative of conversion with respect to temperature at constant concentration describes the instantaneous conversion that exceeds the steady-state conversion relative to the amount unconverted:[2]

$$\left.\frac{\partial y}{\partial T}\right|_x = \left(\frac{dy}{dT}\right)\frac{1}{1 - y} \tag{10.13}$$

This means that the maximum dynamic slope of the plug-flow reactor is

$$\frac{\partial y}{\partial T}_x = \frac{E}{RT^2} \ln (1 - y) \tag{10.14}$$

And for the back-mixed reactor,

$$\left.\frac{\partial y}{\partial T}\right|_x = \frac{E}{RT^2}\, y \tag{10.15}$$

### The Stability of Exothermic Reactors

An exothermic reaction is one in which heat is evolved. The evolution of heat increases temperature, which increases the rate of reaction. This series of events forms a positive feedback loop which can result in a runaway if other conditions permit. The conditions are:

1. Heat cannot be removed to the surroundings as fast as it is evolved.
2. Conversion is sufficiently below 100 percent that heat evolution is not thereby limited.

The rate of heat evolution $Q_r$ is simply the rate of reaction times the heat of reaction $H_r$:

$$Q_r = H_r F x_0 y \tag{10.16}$$

Because $Q_r$ varies directly with $y$, the curves of Fig. 10.2 can also be plotted as heat evolution against temperature. Figure 10.3 shows the heat evolution of the back-mixed reactor at $V/F = 1$; the temperature is to be controlled at 230°F to maintain 50 percent conversion.

For the moment, neglect the sensible heat of the reactants, such that all the evolved heat is to be transferred to a cooling system. The rate of heat transfer $Q_T$ will approach

$$Q_T = UA(T - T_c)$$

where $T_c$ is the coolant temperature. This describes a straight line of slope $UA$ and intercept $T_c$. Lines representing two possible cooling systems designed for the same heat flow are also shown in Fig. 10.3.

The normal condition for the reactor is described by point 0 in Fig. 10.3. No matter which heat removal line is followed, $Q_T = Q_r$ at that point, so that a state of thermal equilibrium can exist. But should the temperature rise, the rate of heat evolution will increase more than the rate

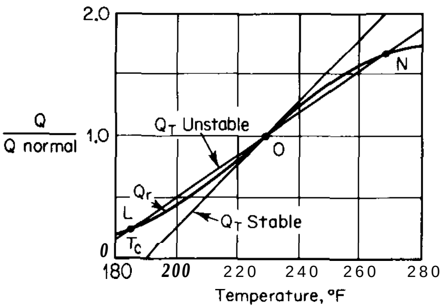



*FIG 10.3. The slope of the heat-removal line determines whether the reactor will be stable.*

of heat transferred by the system designated as unstable.[3] This will cause the temperature to rise farther until the $Q_r$ curve crosses the line again, at point $N$. Points $L$ and $N$ arc stable intersections, while point 0 is an unstable intersection.

The other heat transfer line demonstrates a capability for removing more heat than is evolved upon a temperature increase, thereby restoring equilibrium. This indicates that an exothermic reactor can be made inherently stable by providing sufficient heat transfer area. To state it another way, sufficient heat transfer area will provide negative feedback in excess of the positive feedback of the reaction.

To more explicitly define the relationships involved, an unsteady-state heat balance must be written:

$$H_r F x_0 y - UA\,(T - T_c) - F\rho C(T - T_F) = V\rho C\,\frac{dT}{dt} \qquad (10.17)$$

The first two terms of Eq. (10.17) have already been described. The third term represents the sensible heat absorbed by the reaction stream of density $\rho$ and specific heat C as it rises from inlet temperature $T_F$. The term on the right of the equation represents the thermal capacity of the system.

Since heat evolution is a nonlinear function of temperature, it is necessary to linearize Eq. (10.17) in order to find the thermal time constant of the reactor. Let Eq. (10.18) describe variations about a designated reference temperature $T_r$:

$$H_r F x_0 \left(\frac{\partial y}{\partial T}\right)_x (T - T_r) - UA(T - T_r) - F\rho C(T - T_r) = V\rho C\,\frac{dT}{dt} \qquad (10.18)$$

Arranging in the classical form of a first-order equation allows identification of the time constant:

$$T + \frac{V\rho C}{UA + F\rho C - H_r F x_0(\partial y/\partial T)_x}\,\frac{dT}{dt} = T_r$$

The thermal time constant is

$$\tau_T = \frac{V\rho C}{UA + F\rho C - H_r F x_0(\partial y/\partial T)_x} \qquad (10.19)$$

If $T_c$ is the manipulated variable, the steady-state process gain turns out to be

$$K_T = \frac{T}{T_c} = \frac{UA}{UA + F\rho C - H_r F x_0(\partial y/\partial T)_x} \qquad (10.20)$$

If the reactor is unstable, bbth the gain and the time constant will be negative. The denominator in both expressions is the difference between the slopes of the heat-removal and heat-evolution curves, as in Fig. 10.3. If both denominators are positive, the reactor behaves as a simple first-order lag. If both are negative, positive feedback dominates; the dynamic gain is the same as a simple lag, but the phase angle goes from -90" at zero period to -180" at an infinite period:

$$\phi_T = -\pi + \tan^{-1} 2\pi \frac{\tau_T}{\tau_o} \tag{10.21}$$

The $-\pi$ indicates a negative steady-state gain, while the plus sign in front of the $\tan^{-1}$ indicates a negative time constant. Both the time constant and the steady-state gain can also approach infinity, in which case the reactor acts as an integrator whose dynamic gain at period $\tau_o$ is

$$G = \frac{K_T\tau_o}{2\pi\tau_T} = \frac{\tau_o}{2\pi V\rho C/UA} \tag{10.22}$$

Equation (10.22) defines the dynamic asymptote of process gain for all conditions of stability, exhibiting an effective time constant of $V\rho C/UA$.

A stable reactor can operate without temperature control; regulation of $T_c$ alone is ordinarily sufficient. But an unstable reactor will drift away from the control point in either direction at an ever-increasing rate, unless feedback control is enforced.

Unfortunately, it may not always be possible or economical to design for stability. Enough heat transfer area must be provided so that only about 50 to 60°F differential is required across it to remove the rated flow of heat. (This is an estimate of the $T - T_c$ ordinarily required to exceed $dT/dy$, such as that given in Fig. 10.2.[3])

Stability will be assured if heat is removed by boiling one or more of the ingredients in the reaction, since this makes the system almost isothermal. On the other hand, if heat is removed by a mechanism like evaporation of liquid into a dry gas stream, its flow may change very little with temperature. In this case the slope of the heat-removal curve would be slight, and the reactor could be expected to be unstable.

All of the foregoing statements on stability were used to describe open-loop situations. Some unstable reactors can be given steady-state stability by applying enough negative feedback from the control system to overcome the positive feedback of the reaction. To visualize how this is possible, consider the proportional control loop of Fig. 10.4 for steady-state conditions only.

Figure 10.4 can be represented mathematically by

$$T_c = \frac{100}{P}(T_r - T) = \frac{T}{K_T}$$

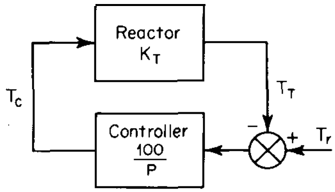



*FIG 10.4. Negative feedback of the controller must overcome positive feedback in the reactor in order to attain steady-state stability.*

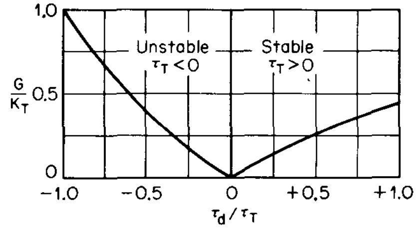



*FIG 10.5. The dynamic gain varies almost linearly with the amount of dead time in the loop.*

The closed-loop steady-state gain is found by solving for **T/T, :**

$$\frac{T}{T_r} = \frac{1}{1 + P/100K_T} \qquad (10.23)$$

Steady-state stability is identified by positive gain. In order for **T/T,** to be positive,

$$\frac{P}{100K_T} > -1$$

If $K_T$ were positive, $P$ could have any value, because the reactor would be stable with the loop open. But with $K_T$ negative, $P$ cannot be greater than -100 $K_T$: if $K_T$ = -2, $P$ < 200%. This sets an upper limit on $P$.

The dynamic properties of the rest of the loop set a lower limit on $P$. The natural period of the temperature-control loop is found by equating the sum of the phase lags of all dynamic elements to 180". Since the phase of a negative lag is between -90 and −180°, there is little room for other elements. This clearly rules out integral control.

If all the other dynamic elements in the loop can be lumped together as dead time $\tau_d$, the period of oscillation can be found by equating the sum of the phase lags to -180":

$$-\pi = -2\pi \frac{\tau_d}{\tau_o} - \pi + \tan^{-1} 2\pi \frac{\tau_T}{\tau_o}$$

Having found $\tau_o$, the dynamic gain of the unstable reactor can be determined :

$$G = K_T \left[1 + \left(2\pi \frac{\tau_T}{\tau_o}\right)^2\right]^{-1/2}$$

A plot of dynamic gain vs. the ratio of $\tau_d/\tau_T$ is given in Fig. 10.5. The dynamic gain of a stable reactor is included for comparison.

The upper limit on $P$ has been established at $-100K_T$. But if $P$ *were* set exactly at that limit, the reactor would have no net feedback and so could scarcely be considered stabilized. So a realistic value for $P$ would apply twice the necessary negative feedback:

$$P < -50K_T$$

But in order to provide ¼-amplitude damping at $\tau_o$,

$$P = 200G$$

The combination of these two conditions can only be realized if

$$\frac{G}{K_T} < 0.25$$

This corresponds to a $\tau_d/\tau_T$ of 0.35 or less in Fig. 10.5. If $\tau_d > 0.35\tau_T$, the reactor will bc either poorly damped at $\tau_o$ or it will be prone to float in the long term, depending on which limit the proportional band is set to favor. An unstable reactor whose dead time approaches its thermal time constant cannot be controlled with any confidence. The temperature tends to limit-cycle in a sawtooth manner, rising slowly to the set point as the cooling is reduced, then descending rapidly when cooling is applied by the controller. Often the only remedy for this situation is a reduction in throughput until stability is achieved.

Instability also affects the natural period of a reactor in *a* closed loop. Based on the same equations as Fig. 10.5, the period of unstable and stable reactors are compared in Fig. 10.6. All of the foregoing should provide ample incentive for designing a reactor which will be stable in the open loop.[4] Because an unstable reactor is such a difficult control problem, the engineer should usc every advantage at his command. For example:

1. Use cascade control from reactor temperature to coolant temperature for fast response.
2. Maintain the maximum flow of coolant to minimize dead time.
3. Usc a noninteracting primary controller. (The latter function may be achieved by combining the outputs of a proportional-plus-derivative and a proportional-plus-reset or integrating rontroller as shown in Fig. 4.86.)

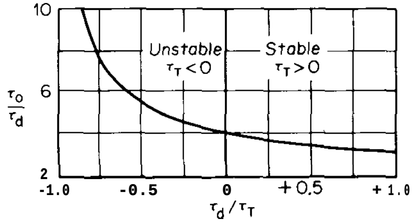



*FIG 10.6. The response of an unstable reactor is both slower and more variable than a stable reactor with the same parameters.*

## CONTINUOUS REACTORS

Continuous reactors are designed to operate under conditions of constant feed rate, withdrawal of product, and removal or supply of heat. If properly controlled, they are ordinarily invariant-that is, the distribution of composition and temperature is constant with respect to time and space. (Gradual degradation of catalyst, fouling of heat transfer surfaces, etc., often are encountered, but their time scale is beyond the control spectrum.) The goal of the control system is to ensure that the operating conditions do remain constant, at the design specifications, while minimizing the losses of both product and reactants.

There are so many types of reactions that it is not possible to discuss them all, yet a general classification can be quite helpful. The distinction between plug-flow and back-mixed reactors has already been made, the former being capable of greater conversion, and the latter being easier to control. Beyond this, some reactions are carried virtually to completion in a single pass; others are forced to be conducted at low conversion due to, low reaction rates, reversibility of the reaction, or occurrence of side reactions.

When conversion is incomplete, the excess reactant(s) must be recycled; therein lies a major distinction—single-pass vs. recycle operation. Some reactions are conducted in an essentially inert media such as a solvent, which also may be recycled. Finally, reactions are occasionally moderated by dilution with product, which then is recycled. To summarize, reactors may be classified as:

1. Single-pass
2. Recycle
    a. Reactant(s)
    b. Inert vehicle
    c. Product

Each group has its own distinctive arrangement. of flow- and inventory-control loops.

### Apportioning Reactant Flows

Whenever one of the reactants differs in phase from products and other reactants, it may be automatically added at the same rate as it is consumed, by controlling its inventory within the reactor. Figure 10.7 shows single-pass reactors with either a liquid or a gas as the manipulated flow.

Note that flow recorders in Fig. 10.7 will tell just what the average reaction rate is, if allowance is made for evaporation or absorption, and purge flow. The purge need not be continuous, but is an absolute neces-

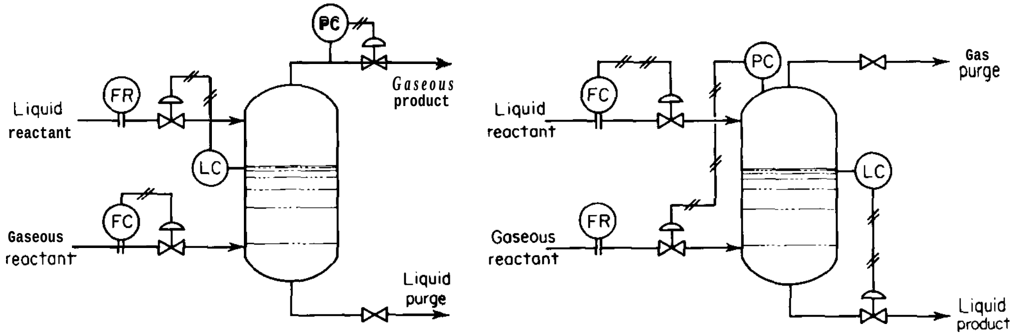



*FIG 10.7. One reactant can be automatically added as it is consumed, if it differs in phase from the other materials.*

sity to rid the reactor of inert contaminants. A trace of nitrogen, for example, in a hydrogen feed stream can soon accumulate sufficiently to impede the reaction, unless it is periodically or continuously discharged. By the same token, lubricating oils from pumps and compressors can accumulate in a liquid dead end unless purged.

In single-phase reactions carried out in one pass, as a neutralization would be, accurate control of the ratio of the reactants is of paramount importance. Excess of any reactant is not only wasted, but may cause undesirable side reactions, including corrosion. If an end-point analyzer, such as pH, is available, it must be used for feedback trim of the reactant ratio. Manually set ratio control of the feed streams is ordinarily not accurate enough, particularly if the composition of one stream is variable.

If one of the reactants is recycled, control of its flow is not critical at all, because it is always in excess. And since it is ordinarily separated entirely from the products, addition of fresh reactant can be manipulated by inventory control. Figure 10.8 shows how a liquid reactant is added to the recycle stream to make up what. is consumed in the reaction.

Product is recycled for the purpose of moderating a reaction. If no reactant is recycled along with it, the requirements for control of the ratio of the feed streams is as stringent as in the single-pass reactor.

Many reactions occur in the presence of some vehicle favorable to both reactants, such as a solvent. An inert diluent acts also to moderate a reaction, facilitating control of temperature and product distribution. If the reactants are soluble in the vehicle, whatever amount goes unreacted will ordinarily be recycled with it. An example of a process with solvent recycle is shown in Fig. 10.9.

In a single-pass reactor, an excess of any reactant is lost; but with solvent recycle, excess accumulates if there is no feedback loop to control its concentration. In the system of Fig. 10.9, reactants X and Y are to

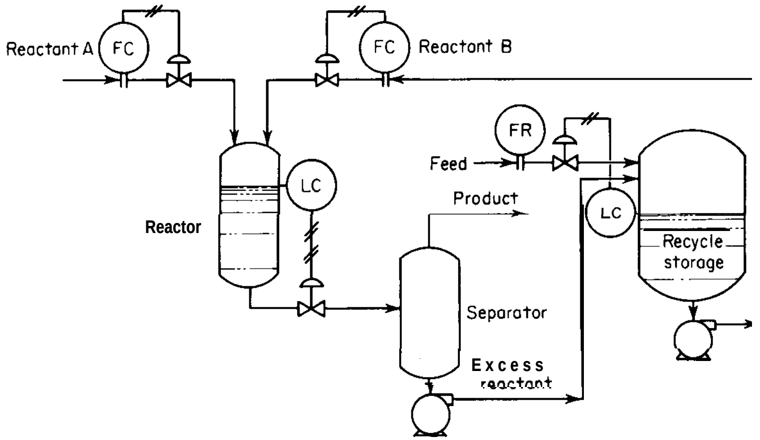



*FIG 10.8. The flow of reactant B may be set to limit the concentration of reactant A or to fix the residence time.*

be added in equal quantities, the reaction going to completion. But solvent enters at a rate $F$, carrying with it recycled reactant X at a concentration $x_1$. The concentration of X in the solvent leaving the reactor is designated $x_2$ and is found from a mass balance, neglecting holdup within the reactor:

$$Fx_2 = Fx_1 + X - Y$$

The content of reactant in volume $V$ of stored solvent is found by an

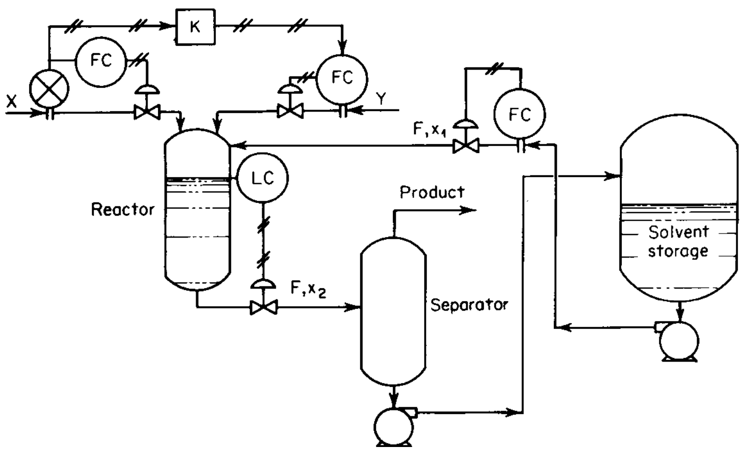



*FIG 10.9. Excess reactant is ordinarily recycled with the solvent.*

unsteady-&ate mass balance on the storage tank:

$$Fx_2 = Fx_1 + V\frac{dx_1}{dt}$$

Combining the two equations yields the variation of $x_1$ with respect to the imbalance in reactant flows:

$$x_1 = \frac{F}{V}\int \frac{X - Y}{F}\,dt \tag{10.24}$$

Equation (10.24) shows that, because of the recycle loop, the process is non-self-regulating.[4] Consequently an integral controller cannot be used to regulate composition. This rules out any kind of feedback-optimizing control system. But because of the lack of self-regulation, end-point control is essential.

**Temperature Control**

Endothermic reactors present no problem regarding temperature control, since they exhibit a marked degree of self-regulation. The exothermic reactors, which have already been introduced, pose the real problem. Their negative self-regulation has been demonstrated.

One facet of an cxothermic reaction that has not yet been discussed is its initiation. Because reaction rate increases with temperature, heat must be applied before any conversion is obtained-then heat must be removed. So the heat, transfer system must have the capability of operating in either direction. This creates something of a problem: steam, for example, is most often used for heating, but is worthless for cooling. There are two general approaches to this problem:

1. Employ a "two-way" cooling system, i.e., one capable of heating too.
2. Split the duties by preheating the reactants and cooling the reactor.

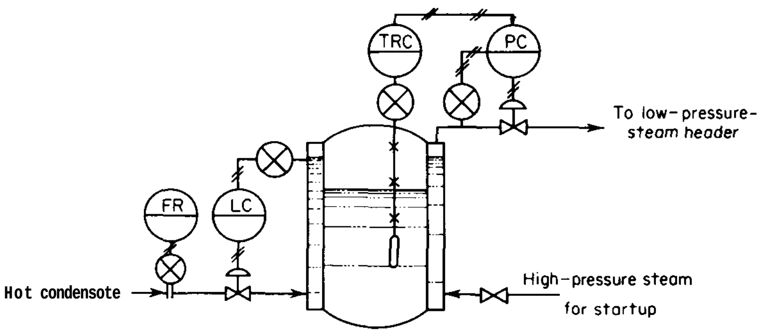



*FIG 10.10. This system will allow easy startup as well as efficient cooling.*

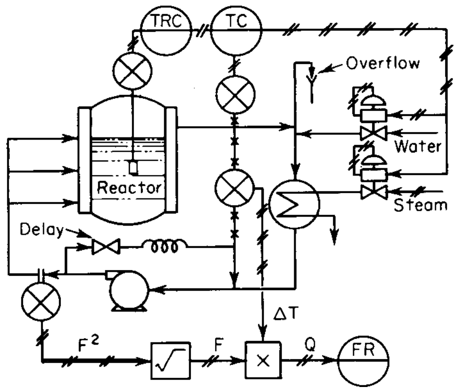



***FIG 10.11. Manipulation of exit temperature is effective when coolant is circulated at a rapid rate.***

A very effective two-way cooling system uses a boiling liquid to which heat may be externally applied for startup. An example is pictured in Fig. 10.10. Because the rate of heat transfer in the system shown in Fig. 10.10 is directly proportional to coolant temperature, that variable should be manipulated to control reactor temperature. The boiling point of the condensate is a function of its pressure, therefore the manipulated variable is the set point of the jacket back-pressure controller.

Another system commonly used features a liquid coolant rapidly circulated past the heat transfer surfaces. Figure 10.11 shows the arrangement when water is used as the coolant. Coolant temperature is chosen as the manipulated variable, since it is linear with both heat transfer rate and reactor temperature. A high rate of circulation allows maximum heat transfer and speed of response.

If recirculation of coolant is not used, the dead time in the secondary temperature loop will vary with the coolant flow. Combined with the nonlinear variation of temperature with flow (Eq. 9.12), it results in a limit cycle, even with an equal-percentage valve. The cycle has a distinctive appearance, the high-temperature portion, when the flow is greatest, being of short duration, the low-temperature part of the cycle being longer.

Heat-removal systems can be used for monitoring conversion. The record of condensate flow to the reactor in Fig. 10.10 should provide a reliable indication of heat evolution. With a circulating liquid coolant, however, flow must be multiplied by temperature difference from inlet to outlet in order to determine the rate of heat evolution. There are some pitfalls in this system:

1. When the primary controller calls for more cooling, the temperature at the jacket inlet will fall before that at the outlet. The difference between inlet and outlet is ordinarily only about 5°F, which could be less than a transient in inlet temperature. This makes the record of

heat transfer appear very erratic, unless dynamic compensation is applied to the inlet measurement.

2. There are many sources of error, the principal one being a difference in temperature between the reactants and products.

Dynamic compensation requires that the inlet temperature measurement be delayed behind its actual value an amount equal to the delay through the jacket. It can be applied most effectively by simulating the jacket by a length of tubing whose dead time may be adjusted by changing the flow. The system in Fig. 10.11 uses this compensation.

If the reactants are to be preheated, it should be done before mixing, unless a catalyst is necessary for the reaction to take place. Otherwise there is no assurance that the reaction would not begin inside the preheater, where it could not be controlled. Adding heat in the preheater(s) and removing it in the reactor is hardly economical. But once a reaction is initiated, it is often possible to bypass the preheaters without adverse effect.

Some reactors have a regenerative preheating system, in which heat is transferred from the product stream to the reactants through an exchanger. Although this is economically advantageous, unless preheat temperature is controlled, a positive feedback loop is formed which can destroy whatever self-regulation the reactor might have had. Temperature control of regenerative preheat can be accomplished as shown in Fig. 10.12.

Whenever a liquid-phase reaction is conducted at a temperature near the boiling point of one of the reactants or products, heat of vaporization may be used for control. If one of the reactants vaporizes, it may be refluxed back to the reactor after condensing. If a product vaporizes, it may be removed as a vapor. This type of heat-removal system is highly self-regulating, but it is also pressure-sensitive. In fact, pressure control should be applied rather than temperature control, since it is a more responsive measurement,. Throttling the reflux from the condenser, or the vapor leaving the reactor if there is no reflux, is an effective means for controlling pressure.

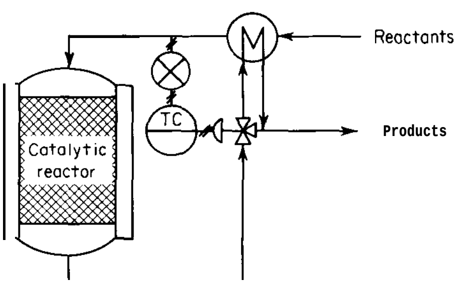



***FIG 10.12. Control of preheat temperature is necessary for this reactor to be stable.***

## pH CONTROL

While a discussion on end-point control in general might be in order, pH is used far more widely than any other measurement to sense the state of a reaction. So while pH has some peculiarities of its own, principally its logarithmic character, much of the following commentary applies to other end-point measurements.

In several instances in earlier chapters, pH has been cited as a difficult control problem. It has, in addition to the usual properties of a composition loop, a severely nonlinear measurement. This very characteristic imposes exceptional demands in flow rangeability of the valves and control system.

### Defining the pH Curve

The outstanding property of each acid-base system is its pH curve; the one shown in Fig. 2.12 is typical of a base neutralized by an acid. The shape of the curve is related to the equilibrium constants for ionization of the acid and base, and the concentrations of each of the ions. But the basis of the coordinate system is logarithmic, in that pH is defined as the negative logarithm of the hydrogen-ion concentration, in gram-ions/liter:

$$pH = -\log H+] \qquad \text{or} \qquad [H^+] = 10^{-pH} \tag{10.25}$$

Pure water ionizes into hydrogen and hydroxyl ions of equal concentration:

$$H_2O \overset{K}{\rightleftharpoons} H^+ + OH^-$$

The equilibrium constant for the ionization of water is $10^{-14}$:

$$K = \frac{[H^+][OH^-]}{[H_2O]} = 10^{-14}$$

This is a useful relationship, because it defines the hydroxyl-ion concentration of any aqueous solution whose pH is known:

$$[OH^-] = 10^{pH-14} \tag{10.26}$$

The neutral point for water is where hydrogen and hydroxyl ions are at equal strength, i.e., at pH 7.

Each acid and base has its own ionization constant. Some acids (or bases) have two or three hydrogen (or hydroxyl) ions per atom, each of which has its own constant. The ionization constant determines the pH for a given strength of acid or base. Consider the example of an acid

HA with an ionization constant. $K_A$ and of a base BOH whose constant is $K_B$, in separate solutions. Ionization proceeds as follows:

$$HA \overset{K_A}{\rightleftharpoons} A^+ + A^- \qquad BOH \overset{K_B}{\rightleftharpoons} B^+ + OH^-$$

The pH of each solution is readily derived:

$$10^{-pH} = \frac{K_A[HA]}{[A^-]} \qquad 10^{pH-14} = \frac{K_B[BOH]}{[B^+]} \tag{10.27}$$

Notice that the pH is a function of the concentration of the companion ion in each solution. The pH of each solution mill be farthest from 7 when the companion ions are at a minimum: i.e., equal to the hydrogen- or hydroxyl-ion concentration:

$$10^{-pH} = \sqrt{K_A[HA]} \qquad 10^{pH-14} = \sqrt{K_B[BOH]} \tag{10.28}$$

If some neutralization has already taken place, however, the concentration of the hydrogen or hydroxyl ion will be less than that of its companion, which could alter the pH vs. concentration relationship considerably.

Because of the effect of the half-power in Eq. (10.28), the pH of pure acids and bases can be expected to change by 0.5 with every decade increase in concentration. This is true for weak acids and bases. Strong acids and bases, however, do not obey the rule, probably because their ionization is not affected by the presence of a companion ion, since every decade in concentration changes pH by about one unit, as Table 10.1 indicates.

From the data in Table 10.1, the ionization constant for acetic acid is found to be $1.83 \times 10^{-5}$ and for ammonium hydroxide, $3.47 \times 10^{-5}$; ionization of the others is variable.

When controlling to an end point, all of the acid (base) must be neutralized, both what, is already ionized and what is not. Let this total acid concentration be designated $x_A$, and total base, $x_B$:

$$x_A = [H^+] + [HA] \qquad x_B = [OH^-] + [BOH] \tag{10.29}$$

**TABLE 10.1**[5] **pH of Various Acids and Bases**

| Concentration, $N$* | HCl | $CH_3COOH$ | NaOH | $NH_4OH$ |
|---|---|---|---|---|
| 1.0 | 0.10 | 2.37 | 14.05 | 11.77 |
| 0 1 | 1.07 | 2.87 | 13.07 | 11.27 |
| 0.01 | 2.02 | 3.37 | 12.12 | 10.77 |

* Normal concentration is defined as one gram-atom of replaceable hydrogen per liter of solution.

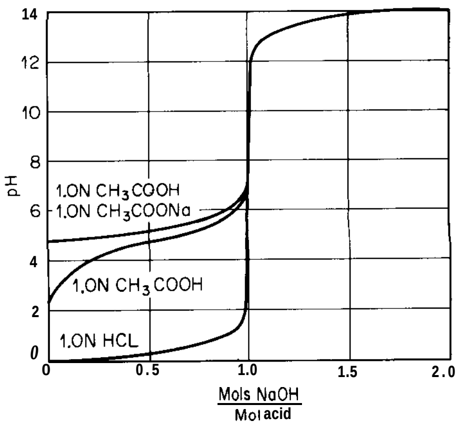



***FIG 10.13. A weak acid or base will require more reagent for neutralization from a given pH, but control is easier; buffering augments the effect.***

Combining with Eq. (10.27) yields

$$x_A = 10^{-\mathrm{pH}}\left(1 + \frac{[A^-]}{K_A}\right) \qquad x_B = 10^{\mathrm{pH}-14}\left(1 + \frac{[B^+]}{K_B}\right) \tag{10.30}$$

If the acid and base are mixed, the resulting pH will be determined by whichever agent was in excess; it can be found by solving the difference between Eqs. (10.30), that is, $x_A - x_B$, for pH.[6]

Notice that neutrality, that is, $x_A - x_B = 0$, occurs at pH 7 only if $[A^-]/K_A = [B^+]/K_B$. This is why a solution of sodium acetate, for example, a neutral salt of a strong base (large $K_B$) and a weak acid (small $K_A$), exhibits a high pH. The slope of a pH curve is influenced strongly by the ionization constants of both acid and base. The ionization of weak acids and bases is severely limited by the concentration of their companion ions; a solution whose pH is thus limited is said to be "buffered."

Figure 10.13 shows that the slope of a strong acid-strong base curve is so great near neutrality that stable control is virtually impossible. Fortunately, most applications involve neutralizing a weak acid (possibly buffered) with a strong base, also shown in Fig. 10.13, or a weak base with a strong acid.

Control of pH has been pursued successfully in media other than water. The solvent must be sufficiently polar to ionize the solutes and be a moderate conductor of electricity; methanol fits into this category. Trace amounts of water are helpful, although not always necessary. Because each solvent has its own ionization constant, neutrality is not necessarily at pH 7 in nonaqueous media. Where it is not possible to measure pH in a particular organic solvent,, a sample may be continuously extracted with water and the pH of the aqueous phase measured.

## Rangeability Requirements for pH-Control Systems

When pH control is exercised on a chemical reaction in which a product is being made, conditions can be expected to be well defined. For example, the required ratio of reagent acid (or base) flow to that of the product or other reactants ordinarily would change but little. Furthermore, the pH curve ought to be known and invariant.

Because of the tremendous sensitivity of the curve in the region of neutrality, it is always necessary to trim the ratio wtih a feedback loop. In addition, the nonlinearity of the measurement should be compensated by using the continuous nonlinear controller described at the end of Chap. 5. A diagram of the recommended system appears in Fig. 10.14. The flow signals are linearized to maintain loop gain constant over the full range of flow.

The majority of pH applications involve neutralization of plant' waste from a combination of drains, sumps, vent scrubbers, etc. The demands of these waste-treating systems complicate the control problem in several dimensions:

1. The flow of the effluent stream may vary as much as four- or fivefold.
2. The stream may alternate between acidic and basic, requiring two reagents.
3. Its acid or base content may vary over several decades.
4. The type of acid or base in solution may vary from weak to strong, with the possibility of buffering; thus the pH curve is variable.

The flow range required of the reagents is moderated by whatever tolerance is set on final pH. The scheme shown in Fig. 10.14 is wholly unsuited to such an application, because the differential meters are accurate to 1 percent of span only from 25 to 100 percent, a 4 : 1 range. Linear control valves are limited to the vicinity of 25: 1 rangeability, while

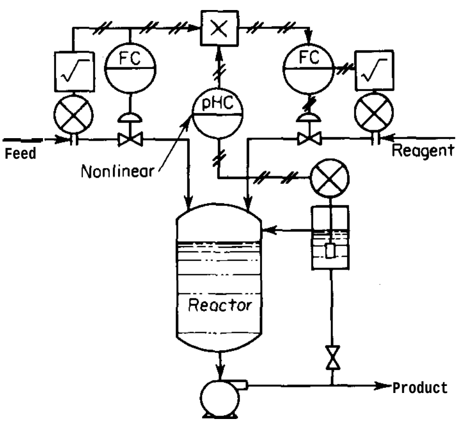



*FIG 10.14. When controlling a reactor, flow ratio should be trimmed by a nonlinear controller.*

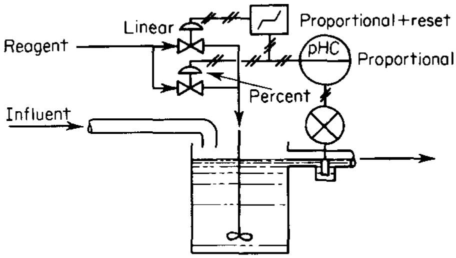



*FIG 10.15. The small equal-percentage valve provides precise trim, while the large linear valve accommodates major load changes,*

equal-percentage valves are, at best, 50: 1. If, for example, a 200: 1 flow range is required, two valves must be used.

Sequencing of control valves to increase rangeability is tricky. It is not enough to open a smaller valve fully before cracking a larger valve— loop gain is an important consideration. If two linear valves of a size ratio of 10 are chosen to be sequenced, constant gain will only be achieved if the small valve is fully open at 9 percent controller output (1 :11). The gain of an equal-percentage valve varies directly with the actual flow through it. This flow characteristic will be altered if two valves are open at the same time; so their correct sequencing requires the smaller valve to be closed as the larger is opened. A certain amount of logic is required to perform this function.

A control system has been developed which incorporates valve sequencing for wide range along with compensation for the nonlinear curve.[7] It features a small equal-percentage valve driven by a proportional pH controller. The output of the pH controller also operates a large linear valve through a proportional-plus-reset controller with a dead zone. The system is shown in Fig. 10.15.

Equal-percentage valves have been described as having an exponential characteristic, similar to the pH curve. As pH deviates from neutrality, the gain of the curve decreases; but increasing deviation will open the valve farther, increasing its gain in a compensating manner. Again compensation can only be maintained if the relationship between valve position and pH is fixed. This means reset cannot be used, because it tries to force the deviation to zero regardless of what valve position is required.

As the output of the proportional controller drives the small valve to either of its limits, the dead zone of the two-mode controller is exceeded. Then the large valve is moved at a rate determined by the departure of the control signal from the dead zone and by the values of proportional and reset. When the control signal reenters the dead zone, the large valve is held in its last position. The large valve is of linear characteristic, because the process gain does not vary with flow, as some gains do.

Furthermore, the nonlinearity of the pH curve is compensated, to some extent, by the dead zone.

If the large valve is selected to be about 20 times the size of the small valve, the flow rangeability of the combination can exceed 700 : 1. If this is insufficient for a particular application, neutralization must be conducted in more than one stage. Should influent pH vary on both sides of neutral, a duplicate control system can be used to add the other reagent to the same vessel.

### Feedforward Control of pH

Figure 10.14 showed a combined feedforward-feedback control system with a forward loop from flow only. In waste-treating systems, the pH of the stream to be neutralized often varies more than its flow. But because the relationship between pH and reagent requirement is variable, adaptation of the forward loop by feedback is essential,

Reagent requirements are based on flow times the acid or base content estimated in Eqs. (10.30). For simplification, only the neutralization of an influent acid HA, of flow $F$, by a reagent base $B$, will be considered. The required base flow is

$$Bx_B = Fx_A = F10^{-\text{pH}}\left(1 + \frac{[A^-]}{K_A}\right) \tag{10.31}$$

If thr reagent is admitted by an equal-percentage valve, the logarithmic characteristic between flow and position, $m$, will be found useful:

$$-\ln\frac{B}{B_{\max}} = 4(1 - m)$$

Conversion to base 10 logarithms and lumping of valve size and reagent concentration into a constant $b$ yields

$$-\log Bx_B = 1.73(1 - m) + b$$

Next, Eq. (10.31) can be converted to logarithms:

$$-1.73(1 - m) - b = \log F - \text{pH} + \log\left(1 + \frac{[A^-]}{K_A}\right)$$

Once again the various constants can be lumped, producing a general fcedformard model :

$$m = \log aF + \frac{100}{P}(r - \text{pH}) \tag{10.32}$$

where $r$ = set point

$100/P$ = forward loop gain

$a$ = required feedback adaptation

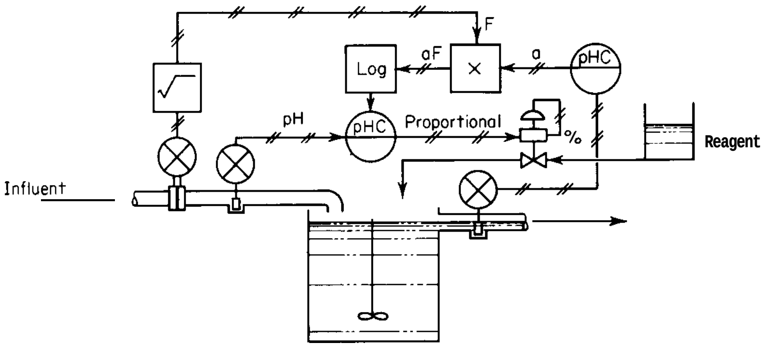



***FIG 10.16. Two forward loops and one feedback loop give this pH system maximum effectiveness.***

Because of the equal-percentage characteristic, loop gain will change with position (hence flow) as well as with pH. In order to make loop gain independent of flow, the output of the feedback controller, here designated a, must be placed on the same terms as flow. Actually the feedback trim is the same as was shown in Fig. 10.14, using a nonlinear controller to compensate the pH curve. The function of using an equal-percentage valve is twofold:

1. To provide maximum rangeability
2. To generate the forward-loop pH function

The control system is shown in Fig. 10.16. Note that the forward-loop summation is made conveniently in a proportional controller with remote bias.

A positioner is indicated on the equal-percentage valve to ensure as accurate delivery of reagent as possible. It is also necessary to supply the reagent at a fixed head. If a single valve gives insufficient rangeability, two may be sequenced. Full flow from the smaller valve should equal about 3 percent of the capacity of the larger. The large valve should be set to open to 3 percent flow when the smaller is full open, at 50 percent output. The smaller must be closed at the same time the larger is opened in order to avoid upsetting the process. A small differential gap in the switching logic is needed to minimize cycling at this point.

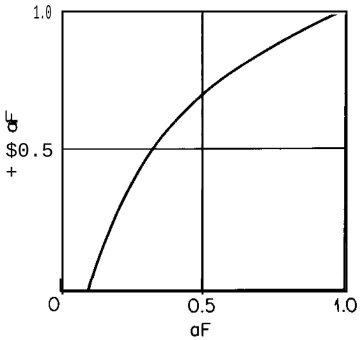



***FIG 10.17. Generation of the logarithmic function over a full decade of flow is more than sufficient.***

Figure 10.17 shows the function needed to include both flow and feedback in the system. It is necessary to add 1 to log $aF$ to normalize the ordinate over a single decade. The curve is not especially difficult to generate and can be matched with a four-bar linkage.[8,9]

## BATCH REACTORS

Although the progress of the chemical industry has been toward continuous processes, some reactions will inevitably be conducted batchwise. The bulk of commercial batch reactions are polymerizations involved in the production of rubber and many types of plastics. Distribution of molecular weight is an important parameter in polymer manufacture, and it seems to be the most easily controlled batchwise. Another consideration is the great change in viscosity frequently encountered between the reactants and products.

The process consists of the several steps listed below, although considerable variation exists from one product to another:

1. Charge the reactor with reactants and catalyst.
2. Heat to operating temperature.
3. Allow the reaction to proceed to completion, normally several hours.
4. Heat or cool to cure temperatures.
5. Cool and empty the reactor.

Production reactors are stirred, jacketed vessels of several thousand gallons capacity. If the reaction is first-order, conversion varies with time according to Eq. (10.6) :

$$y = 1 - e^{-kt} \tag{10.6}$$

The rate of conversion is the derivative of Eq. (10.6) :

$$\frac{dy}{dt} = ke^{-kt} \tag{10.33}$$

The rate is greatest when the conversion is least, i.e., at time zero.

Polymerization reactions are second-order or higher, because they depend on the simultaneous combination of two or more monomer molecules to form a polymer. In a second-order reaction, the rate depends on the square of reactant concentration:

$$-\frac{dx}{dt} = kx^2 \tag{10.34}$$

Dividing both sides by $-x^2$ and integrating,

$$\int_{x_0}^{x} \frac{dx}{x^2} = \int_0^t -k\,dt$$

$$\frac{1}{x} - \frac{1}{x_0} = kt$$

Conversion and its rate can be found by substituting for $x$:

$$y = \frac{1}{1 + 1/ktx_0} = \frac{ktx_0}{1 + ktx_0} \tag{10.35}$$

$$\frac{dy}{dt} = \frac{kx_0}{(1 + ktx_0)^2} \tag{10.36}$$

The rate of conversion is also the rate of production in a batch reactor and is proportional to heat evolution, if the reaction is exothermic. The rate of conversion of first- and second-order reactions is plotted against time in Fig. 10.18.

## Temperature Control

In the early stages of a batch reaction, temperature control is most important because the rate of conversion is at its highest. Exothermic reactions pose a real control problem because heat must be applied to raise the batch to reaction temperature and then be removed. The cooling system most frequently used is that shown in Fig. 10.11. A proportional controller is used for coolant exit temperature because the proportional band is ordinarily only about 10 percent and offset is not harmful. But the primary controller is three-mode, with special features to permit:

1. Maintenance of optimum settings for operation at reaction temperature
2. Delivery of the batch to reaction temperature without overshoot
3. Conduction of the reaction in a minimum of time

If the reactor is stable, based on its heat transfer characteristics, as discussed earlier with regard to continuous reactors, control of temperature will be simplified. The reactor will respond rapidly, with a period of perhaps 20 min, and 10 percent proportional band may be sufficient for effective damping. All three control modes should be adjusted while at the operating temperature.

In order to avoid overshoot, the primary controller must be equipped with an antiwindup switch with preload applied to the reset circuit. It

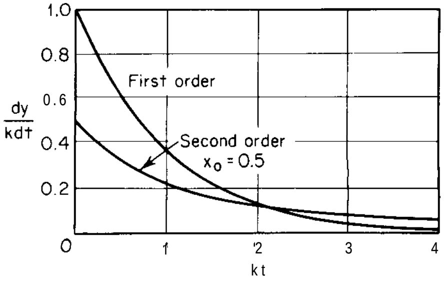



***FIG 10.18. The rate of conversion of higher-order reactions varies less with time, particularly at low concentration levels.***

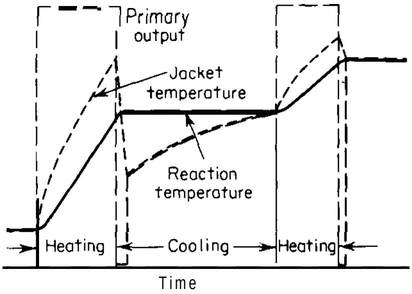



*FIG 10.19. The dual-mode system requires different values of preload for reaction and cure.*

is the preload which determines the magnitude of overshoot (see Fig. 4.6). The correct value for preload is not difficult to estimate for a known reaction. The initial rate of conversion will release a predictable flow of heat, all of which must be removed through the heat transfer surface. Both the batch and the cooling fluid are circulated at very high rates to ensure good heat transfer; thus each has little temperature gradient and constant flow. The rate of heat transfer is therefore directly proportional to the temperature difference between primary and secondary measurements. The output of the primary controller, which is the set point of the secondary, is predictable, and its predicted value can be introduced as preload. As pointed out in Chap. 4, however, it is necessary to set the preload a few percent below the predicted value to allow for the reset action of the controller from the time the antiwindup switch is released until the set point is reached.

A batch reactor can be unstable, in which case its natural period will be perhaps twice as long and its proportional band requirement twice as great as a physically similar stable one. The control system described above loses its effectiveness when a wide proportional band is required. In order to avoid overshoot, the heat input must be throttled early, which can add considerable time to the length of the operation.

For a problem such as this, the dual-mode control system described in Fig. 5.17 is extremely effective.“’ The prcload is estimated as before, but no correction is required for integration, because reset action is not initiated until the error is nearly zero. Full heating can be applied to within 1 or 2 percent of the set point,, far beyond the capabilities of a 25 percent proportional band. Yet full cooling need only be applied for a time delay of perhaps a minute to dissipate the energy stored in the jacket. As pointed out in Fig. 5.18, the switching parameters are easy to adjust and tolerant of maladjustment.

Figure 10.19 shows the relationships between primary and jacket temperatures and the dual-mode output for a typical reactor. If the settings are correct’, jacket temperature mill fall to meet its set point at the preload value when the time delay is over. Notice how the rate of heat transfer diminishes to zero as conversion is completed. There is no heat evolution during the cure phase.

Equation (10.19) gave the thermal time constant of a continuous reactor. Among other things, it depended on reactant concentration and conversion. Since these are both variable in a batch reactor, it is entirely possible to proceed from an unstable to a stable condition with the passage of time. Control settings necessarily must favor the more difficult situation.

In some batch reactions, one of the ingredients is introduced continuously until the other reactants are entirely consumed. The rate at which this ingredient is added is usually limited by the rate of heat removal. In order to carry out the reaction in minimum time, the batch must be heated to reaction temperature as rapidly as possible, where full cooling is applied. The temperature controller then must manipulate the addition of reactant through a linear cascade flow loop. An interlock must be arranged to prevent reactant from being introduced until the batch is approaching reaction temperature. Otherwise a quantity could accumulate accidentally within the batch, and once the reaction began, could generate heat in excess of the capabilities of the cooling system.

### End-point Control

At the completion of the reaction, heat evolution will subside, causing the temperature controller to increase reactant flow. Since this is obviously incorrect action, some logic must be arranged to override tempera-

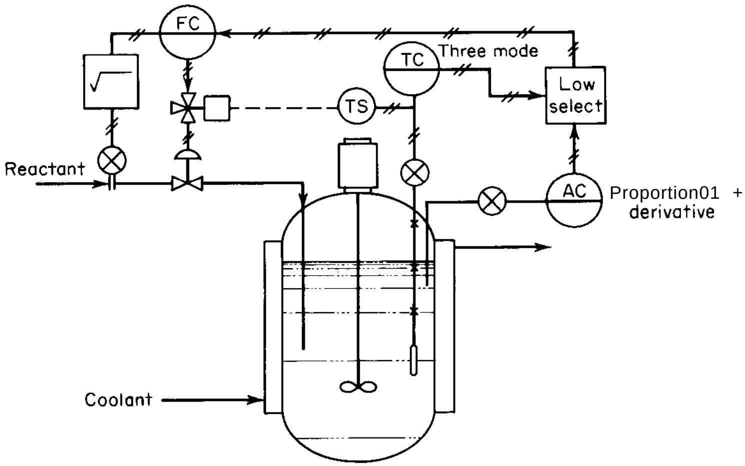



*FIG 10.20. The temperature-control loop must be disabled at both the start and finish of the reaction.*

ture control upon the termination of the reaction. If there is no way of measuring the completion of the reaction, reactant flow may be stopped by a flow totalizer, once a predetermined quantity has been introduced. The termination of some reactions is indicated by a fall in pressure; others are indicated by a rise in pressure. In either case, a pressure switch may be used to shut off flow.

End-point control of a batch reaction has one outstanding characteristic-reset must not be used. The bulk of the reaction is conducted away from the set point-reset would try to overcome this offset and ultimately result in overshoot. The controller should be proportional-plus-derivative, with zero bias, so that the valve is shut when the set point is reached. If the measurement is pH, reagent should be added through an equal-percentage valve (without a cascade flow loop), to match the process characteristic. A system employing temperature and end-point control is shown in Fig. 10.20.

## REFERENCES


1. Corrigan, T. E., and E. F. Young: General Considerations in Reactor Design, *Chem. Eng.,* October, 1955.
2. Harriott, P.: "Process Control," p. 311, McGraw-Hill Book Company, New York, 1964.
3. Shinskey, F. G.: Temperature Control for Gas-phase Reactors, *Chem. Eng.,* October 5, 1959.
4. ———: Designing Stability into Plant Operations, *Petro.-Chem. Eng.,* October, November, 1962.
5. Daniels, F.: "Outlines of Physical Chemistry," p. 460, John Wiley & Sons, Inc, New York, 1948.
6. Shinskey, F. G.: Controls for Nonlinear Processes, *Chem. Eng.,* March 19, 1962.
7. Green, R. M.: A Unique Approach to Process Control Using Nonlinear Electrochemical Analyzers, *6th Natl. Symp. on Instrumental Methods of Analysis,* June, 1960.
8. Shinskey, F. G.: Analog Computing Control for On-line Applications, *Control Eng.,* November, 1962.
9. Weinstein, W. D.: The Four-bar Linkage as a Function Generator, *Elec. Manufacturing,* October, 1959.
10. Shinskey, F. G., and J. L. Weinstein: A Dual-mode Control System for a Batch Exothermic Reactor, *20th Annual ZSA Conf.,* October, 1965.


## PROBLEMS

**10.1** The reaction whose characteristics appear in Fig. 10.2 is to be conducted at 50 percent conversion with $V/F = 1$ hr in a back-mixed reactor. Temperature is to be controlled at 230°F by manipulation of coolant temperature. Feed is preheated to reaction temperature. What is the lowest coolant temperature that will ensure stability in manual control?

**10.2** For the same reactor, the heat of reaction is 5000 Btu/lb, $F$ is 4,000 lb/hr, and $x_0$ is 0.2. Coolant temperature required for control is 190°F. The reactor contains 4,000 lb of material with a specific heat of 0.4 Btu/lb. What is its thermal time constant? If the dead time in the loop is 1 min, what is the natural period and the proportional band of the primary loop? Is it stable in the steady state?

**10.3** Reduce the flow to the reactor above to 2,000 lb/hr, with all other conditions except coolant temperature constant. Calculate its thermal time constant, period, and proportional band. Reduce $x_0$ to 0.1 with $F$ at 4,000 lb/hr and repeat the calculation.

**10.4** In the process shown in Fig. 10.9, reactants X and Y are each soluble in the solvent, but the product they form is gaseous. The end point of the reaction is to be controlled, as measured by the electrolytic conductivity of the solvent. The conductivity increases with the amount of whichever reactant is in excess and is therefore to be controlled at zero. Devise a way to accomplish this; modify the process if necessary.

**10.5** Calculate the pH at neutrality for a 1 N solution of acetic acid neutralized by caustic, as shown ip Fig. 10.13. Estimate the process gain $dpH/dx_B$ at that point. Assume that $K_B$ is infinite.

**10.6** The concentration of reactant in a batch reactor decreases with time; it may be desirable to gradually increase the reaction temperature to hasten completion. Will either of these factors cause the dynamic gain of the process to change? Explain.

# Distillation

# CHAPTER 11

More than for any other application, there has long been a grave need for an accepted method of controlling distillation. Precise control of product quality is important because the product from most towers is more valuable than the product anywhere else in the process, having reached its terminal stage of refinement. Furthermore, product-quality specifications must be met, even if losses, excessive usage of utilities, and reprocessing augment the cost of separation. But precise control is difficult to attain because:

1. Towers with many trays are slow in responding to control action.
2. Separation is affected by many variables, requiring many control loops, which interact with'one another.
3. On-line analysis is not always available.
4. Distillation units are the last in the chain of processing operations, hence are subject to changes in throughput from all upstream units.
5. The factors affecting separation are not readily interpreted in terms of control system requirements.

Until about 1961 there was no standard method for controlling distillation-no system had been found capable of forcing a column to behave as the designer had intended. But then a revolution began. Different groups, working independently, were making strides in one general direction.[1–3] The culmination of their efforts has been the development of a method for enforcement of the column material balance. A step-by-step introduction to this method follows.

## FACTORS AFFECTING PRODUCT QUALITY

Most texts on distillation start with a design procedure which determines the number of trays needed for a given separation. In control work, however, the column already exists, and speculation over theoretical trays and equilibrium diagrams is of no consequence. A technique especially devised for control application is necessary. This technique begins with a simple block diagram of the tower, which has already been designed to perform a given separation (Fig. 11.1). Although the figure indicates only a binary separation, the concept will be advanced later to multicomponent and multistream towers.

The block diagram reveals two extremely important facts:

1. Energy is necessary for separation. In fact, it may be assumed that no separation will take place if no energy is introduced.

2. The relative composition of the two product streams is intimately bound up with their relative flow rates. More of a given component cannot be withdrawn than is being fed to the tower: the material balance must be satisfied.

To be sure, tray efficiency, loading, etc., also color the picture, but the two factors above are so outstanding in their effects that they must be the prime consideration in any system design.

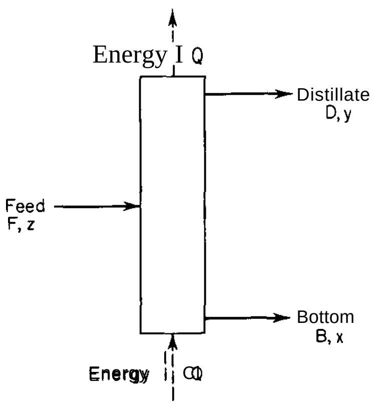



*FIG 11.1. Material and energy balances play key roles in distillation control.*

### The Material Balance

In the steady state, as much material must be withdrawn as enters a tower:

$$F = D + B \tag{11.1}$$

where $F$ = molal feed rate
$D$ = distillate flow
$B$ = bottoms Aow

A material balance on each component must also be closed, using $z$, $y$, and x to represent the mole fraction of the light component in ***F, D,*** and ***B:***

$$Fz = Dy + Bx \tag{11.2}$$

From the overall material balance it is evident that the flow of only one of the product streams can be set independently. The flow of the other is determined by the feed rate and is therefore a dependent variable. But one flow must be set by some criterion, since they cannot both be allowed to drift. For the moment, distillate flow will be chosen to be manipulated by the control system, either directly or indirectly. (The reason for this choice will be discussed later.) Bottoms flow must then be manipulated by a controller which senses liquid level at the bottom of the tower in order to close the material balance by maintaining constant liquid inventory. Bottoms flow is thus dependent on current values of feed and distillate:

$$B = F - D$$

Substituting for ***B*** in the material balance of the light component permits expressing the relationship between the quality of both products in terms of distillate flow:

$$Fz = Dy + (F - D)x$$

The ratio ***D/F*** determines the relative composition of each product:

$$\frac{z-x}{y-x} = \frac{D}{F} \tag{11.3}$$

Figure 11.2 shows one way in which this relationship might be pictured. Remember that $z$ is an uncontrolled variable, like ***F.*** Therefore if $z$ should change, ***D/F*** must be adjusted to maintain constant values of x and $y$.

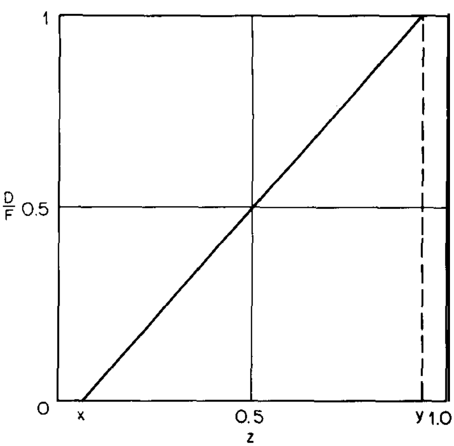



***FIG 11.2. If x and y are to be controlled, D/F must vary proportionately to z.***

Unfortunately, however, the material-balance equation alone is insoluble. It is a single equation with two unknowns, x and y. To provide a solution, another equation of x and y must be found.

The Fenske equation[4] was derived for the purpose of estimating the number of theoretical trays n required to effect a given separation between components whose relative volatility is $\alpha$, at total reflux:

$$\frac{y(1 - x)}{x(1 - y)} = \alpha^{n+1} \tag{11.4}$$

Equation (11.4) needs to be modified to allow extension to situations other than total reflux:

$$\frac{y(1 - x)}{x(1 - y)} = S \tag{11.5}$$

The term $S$ is defined as the separation and is a function of $\alpha$, $n$, and the energy to feed ratio. Solving Eq. (11.5) for $y$ in terms of $x$ and for x in terms of y,

$$y = \frac{\quad}{1 + x(S - 1)} \qquad \frac{\quad}{y + S(1 - y)} \tag{11.6}$$

A direct solution of these equations with the material balance cannot be readily obtained. But through the use of a numerical example, the prevailing relationships will be demonstrated.

Consider the example of a tower separating a feed material $z = 0.5$ into distillate $y = 0.95$ and bottoms $x = 0.05$. The separation factor for this tower under these conditions is

$$S = \frac{(0.95)(0.95)}{(0.05)\ (0.05)} = 361$$

The distillate to feed ratio is

$$D/F = \frac{0.5 - 0.05}{0.95 - 0.05} = 0.5$$

Should any other value of x or $y$ be desired at the prevailing conditions of feed composition and separation, it can be obtained by appropriate adjustment of ***D/F.*** A value for $y$, for example, is first selected, and a corresponding value of $x$ is then calculated from the modified Fenske equation. With these values of x, y, and $z$, the required ***D/F*** can be found from the material balance. Figure 11.3 illustrates how distillate and bottoms compositions would vary with ***D/F*** for this example.

The slope of each curve at ***D/F*** = 0.5 represents the process gain at that point. It happens that the two curves have identical slopes at

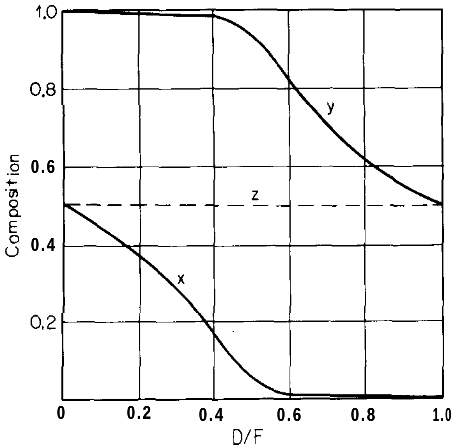



*FIG 11.3. Increasing D/F reduces both x and g, but in different degrees.*

that point:

$$\frac{dy}{d(D/F)} = \frac{dx}{d(D/F)} = -0.9 \quad \text{at} \quad \frac{D}{F} = 0.5$$

A 1 percent change in ***D/F*** will change distillate and bottoms composition by 0.9 percent.

Three conclusions may be drawn from the foregoing discussion:

1. Composition of both product streams is profoundly affected by distillate to feed ratio.

2. Changes in feed composition can be offset by appropriate adjustment of ***D/F.***

3. If separation is constant, control of composition of either product will also result in control of composition of the other product. (The relationship between x and y is fixed for a given separation.)

### Energy vs. Feed

A certain amount of energy is required to separate a given feed stream into its components. It is reasonable to assume that energy requirements will be roughly in proportion to feed rate. In a distillation tower, energy is introduced as heat Q to the reboiler, which generates a proportional flow of vapor ***V:***

$$V = \frac{Q}{H_v} \tag{11.7}$$

The term $H_v$ represents the latent heat of vaporization. Expressing heat input in terms of vapor flow enables evaluation of separation in terms of dimensionless ratio of vapor to feed rate, ***V/F.***

At total reflux, distillate flow is zero, therefore feed rate is also zero. The separation at total reflux has already been established by the Fenske equation :

$$\lim_{\frac{V}{F}\to\infty} S = \alpha^{n+1} \qquad (11.8)$$

A second limit needs to be established to allow the relationship of S to V/F to be outlined. For a given column, the minimum vapor to feed ratio $(V/F)_{\min}$ can be defined as that condition under which controlled product quality requires that no product be withdrawn. At $(V/F)_{\min}$, if $y$ is controlled, $D/F = 0$ and $x = z$. If $x$ is controlled, $D/F = 1$ and $y = z$. In other words, at minimum vapor to feed ratio, production is zero, just as it is at total reflux. Separation $S_{\min}$, at minimum ***V/F,*** is

$$S_{\min} = \frac{y(1-z)}{x(1-y)} \quad \text{or} \quad \frac{z(1-x)}{x(1-z)}$$

Equation (11.9) is derived to satisfy the upper limit of $S$ and to pass through the points representing normal and minimum ***V/F:***

$$S = S_{\min} + (\alpha^{n+1} - S_{\min})\left(1 - \frac{(V/F)_{\min}}{V/F}\right) \qquad (11.9)$$

The point $V/F = 0$, $S = 0$ will not be satisfied, but this is of no concern, because it lies outside the operating region. To gain a better appreciation of this relationship, the equation will be restated using the numbers in the example, given a minimum ***V/F*** of **2.0:**

$$S = 19 + 570\left(1 - \frac{2.0}{V/F}\right)$$

The curve generated by this equation appears in Fig. **11.4.**

The shape of the curve undoubtedly varies somewhat with the shape of the McCabe-Thiele equilibrium curve, but the general characteristics

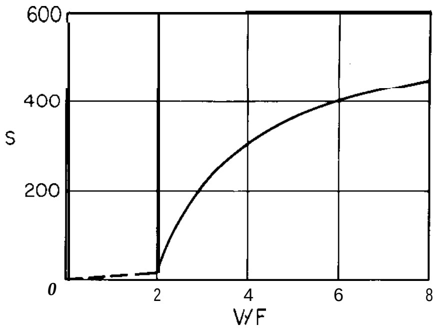



***FIG 11.4. Separation S varies hyperbolically with vapor to feed ratio.***

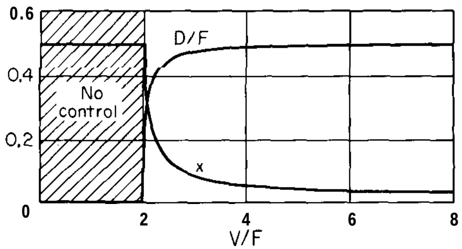



*FIG 11.5. To maintain control of distillate composition, D/F must change with V/F, with a subsequent change in bottoms composition.*

remain: a sharp increase in separation at the minimum $V/F$, gradually tapering off to approach $\alpha^{n+1}$.

Enough information is now available to completely evaluate the effects of ***V/F*** and ***D/F*** in the control of distillate composition. Values of $S$ can be found for current values of ***V/F,*** which can be used to determine $x$ for a controlled value of y. Then the ***D/F*** ratio required for control can be found from the material balance. Figure 11.5 is a plot of $x$ and $D/F$ vs. ***V/F*** for $z$ = 0.5 and y controlled at 0.95.

The slope of the bottoms composition curve at the normal operating conditions, that is, ***V/F = 5,*** is

$$\frac{dx}{d(V/F)} = -0.007$$

Compare this slope to $dx/d(D/F)$ = -0.9 at the same conditions. Composition is over 100 times more sensitive to changes in distillate flow than to changes in heat input.

If both x and y are to be controlled, separation must be constant. This can be done by maintaining a constant ***V/F,*** with ***D/F*** varying only with feed composition. But if heat input is fixed at maximum, so as to

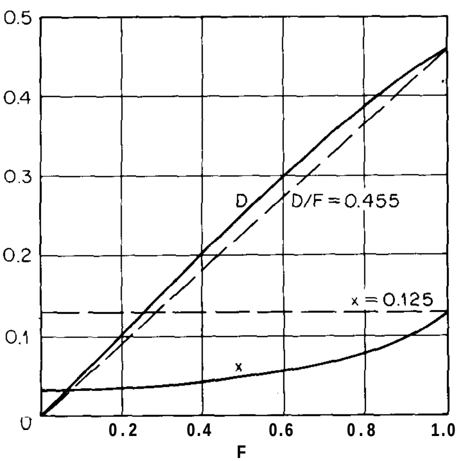



*FIG 11.6. With heat input constant, control of y calls for D/F and x to vary with F.*

TABLE 11.1 Column Variables vs. Feed Rate for Constant Boilup and Feed Composition

| $F$ | 0 | 0.25 | 0.5 | 0.625 | 0.833 | 1.0 |
|---|---|---|---|---|---|---|
| $V/F = 2.5/F$ | ∞ | 10 | 5 | 4 | 3 | 2.5 |
| $S = 19 + 570\left(1 - \frac{2.0}{V/F}\right)$ | 589 | 475 | 361 | 304 | 209 | 133 |
| $x = \frac{0.95}{0.95 + 0.05S}$ | 0.032 | 0.038 | 0.050 | 0.059 | 0.084 | 0.125 |
| $D/F = \frac{0.5 - x}{0.95 - x}$ | 0.510 | 0.507 | 0.500 | 0.494 | 0.480 | 0.455 |
| $D = (D/F)F$ | 0 | 0.127 | 0.250 | 0.308 | 0.400 | 0.455 |

maximize separation at all rates of feed, $V/F$ will change with feed rate. Then $D/F$ must also change with feed rate, in order to control y. Again, returning to the numerical example, let $V$ be fixed at 2.5 times the full-scale feed rate. Then $V/F$ will vary from infinity to 2.5 as feed rate varies from 0 to 100 percent. Distillate flow must then increase less than proportionately to feed rate in order to maintain control with decreasing separation. A plot of distillate flow vs. feed rate for $y = 0.95$ and $z = 0.5$ appears in Fig. 11.6. Table 11.1 has been prepared to enable the reader to follow all the numerical manipulation.

Included in Fig. 11.6 are two broken lines representing operation under conditions of constant separation at $V/F = 2.5$. Although this arrangement maintains control of bottoms composition while saving heat input at low rates of feed, average distillate flow is less, and loss of the light component out the bottom is subsequently greater.

If bottoms composition were controlled with a constant vapor rate, the curve of distillate vs. feed rate would bend in the opposite direction. If $D/F$ were maintained constant under conditions of variable $V/F$, both $x$ and y would turn toward one another as feed rate increased.

## ARRANGING THE CONTROL LOOPS

Because few separations are truly binary, the ramifications that are encountered in column design and operation are extensive. A third component leaves with the distillate if it is the lightest component, or with the bottoms if it is the heaviest; or it may be withdrawn as a side-stream if it boils at a temperature between those of the light and heavy components. When azeotropes are encountered, a third component may be intentionally introduced to facilitate the separation. Towers are often linked together in a chain, occasionally with a stream being recycled to improve recovery. Some separations require so many trays that the tower is broken into two or more sections. There are also partial condensers, vacuum stills, and other peculiar arrangements, which make up a wide variety of distillation processes.

As might be expected, each tower or combination of towers has its own particular requirements for control-loop arrangement,. Beyond this, even, there are various ways of controlling the same tower, some of which are more effective than others. Virtually any arrangement of loops will operate satisfactorily in the steady state—but in the steady state no control is really needed. It is the unsteady state that tests control performance. If feed rate or composition or enthalpy change, or steam pressure, or weather, what happens then?

### Interaction between Heat and Material Balances

Whenever a column is arranged so that both product streams are under level control, its material balance cannot be directly manipulated. Figure 11.7 shows just such a system.

With reflux on flow control, variations in vapor rate and reflux rate will both affect the flow of distillate. Earlier it was demonstrated that the distillate to feed ratio had the greatest effect on product quality. Therefore, it is imperative that distillate flow be controlled directly, instead of being subject to variations in vapor and reflux rates. If the flow of vapor leaving the top of the tower is the same as that generated

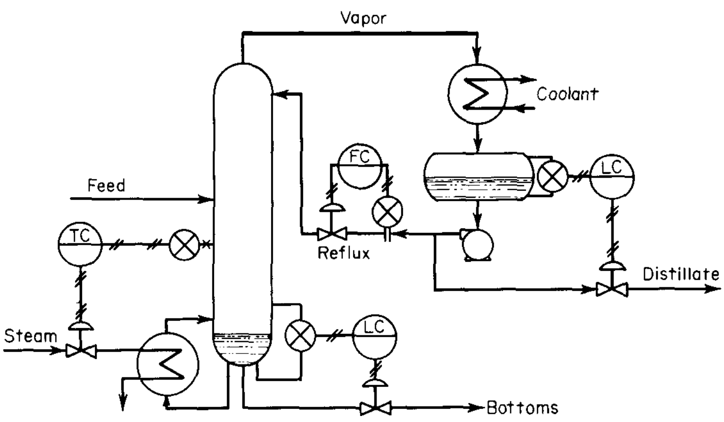



***FIG 12.7. This arrangement of control loops is not recommended, because of interaction between the heat and material balances.***

in the reboiler, distillate flow varies with reflux flow $L$ as

$$D = V - L \qquad (11.10)$$

The fractional change in $D$ resulting from fractional variation in $V$ and $L$ is

$$\frac{dD}{D} = \frac{\partial V}{V}\frac{V}{D} - \frac{\partial L}{L}\frac{L}{D} \qquad (11.11)$$

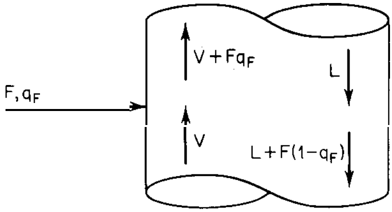



*FIG 11.8. The material balance at the feed tray is affected by feed enthalpy $q_F$.*

Because $V/D$ and $L/D$ are ordinarily much greater than 1, variations in the actual flow of reflux and heat input will cause augmented upset's in product quality. In the numerical example, $V/D$ is 10; a 1 percent change in heat input would upset distillate flow by 10 percent.

But the flow of vapor leaving the top of the tower is a function of many factors. If all of the feed enters as a liquid at its boiling point, flow of vapor above and below the feed tray will be the same. But if part of the feed is vaporized, vapor flow above the feed tray will exceed that below by the amount of vapor in the feed. Figure 11.8 describes the material balance at the feed tray.

Feed enthalpy is identified as the fraction of the feed that is vaporized. Liquid at its boiling point has an enthalpy of 0, while that of saturated vapor is 1.0. The enthalpy of subcooled liquid is negative, that of superheated vapor exceeds 1.0. If no heat were added to or removed from the vapor stream until it entered the condenser, its flow into the condenser would equal that immediately above the feed tray. In this case,

$$D = V + q_F F - L$$

Because it has been made the dependent variable, distillate flow is subject to variation in both feed rate and enthalpy.

Reflux may not return to the tower exactly at the boiling point. In fact, it will usually be subcooled to some extent. This subcooling causes condensation of some of the vapors approaching the top tray of the tower. Figure 11.9 shows the material balance at the top tray.

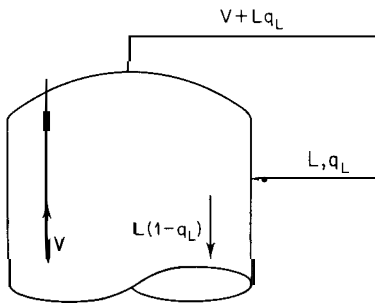



*FIG 11.9. The material balance at the top of the tower is a function of the enthalpy $q_L$ of the reflux stream.*

If feed is entering the tower at its boiling point, the vapor flow approaching the top tray will be the same as that at the reboiler. But if the reflux *is* subcooled, its enthalpy *is* negative, causing a decrease in vapor leaving the tower by a factor $Lq_L$. The stream $L(1 - q_L)$ is called internal reflux. Summing the effect of subcooled reflux with that of variable feed enthalpy yields their combined effects on distillate flow:

$$D = V + q_F F - (1 - q_L)L \tag{11.12}$$

If distillate flow is made the dependent variable, it then depends not only on reflux and vapor rates, but also on feed and reflux enthalpies. Composition is so difficult to control under these conditions, that special-purpose computers are often used to compensate for $q_F$ and $q_L$.[2] But if distillate flow is made independent of the heat balance, no such measures are necessary.

Manipulation of heat input to control column temperature is not recommended for most applications. Because the temperature of a boiling pure liquid is constant, the sensing element is usually moved up the tower to a tray where the temperature of the saturated liquid is a measure of its composition. Temperature is meant to infer composition-yet because it is temperature, many engineers feel that it is a function of heat input. It has been shown, however, that ***D/F*** is 100 times more effective in controlling composition than ***V/F***. The reason that manipulation of heat input affects temperature is beause reflux is on flow control, such that ***D*** is dependent on ***V***. If the temperature is too low, additional heat is sent to the reboiler, which ultimately increases distillate flow. With this arrangement, the heat balance is deliberately upset in order to alter the material balance.

### Manipulating the Material Balance

Control over product quality can best be achieved by manipulating the material balance, free from disturbances in the heat balance. This means that the composition co&roller should set the flow of one of the product streams. Whether distillate or bottoms flow is chosen to be manipulated depends on their relative flow rates. The greater absolute flow accuracy will be obtained by manipulating the smaller flow. So before a decision is made, a complete material balance should be drawn and the flows of the various streams compared.

Whenever distillate flow is set to hold the material balance, the bottoms-level controller must manipulate bottoms flow. This allows the heat input to be set independently to establish the separation capability of the tower. For this reason, whenever comparison of the product flow rate does not overwhelmingly favor manipulation of bottoms flow for composition control, distillate flow should be selected.

A temperature element located part way up the tower is no assurance of absolute quality control. Changes in separation can alter the composition profile to the extent that product quality could vary somewhat without a noticeable deviation in temperature. The variability is relative to the distance between the product and the measuring element. Consequently, an analyzer in the product stream provides a much more reliable measure of quality. Furthermore, some separations are so difficult that the temperature profile is too gradual to hold any meaning.1

### Manipulation of Heat Input

If bottoms flow is under level control, heat input to the column may be fixed at any desired value. Or it may be manipulated to control bottoms composition, if distillate flow is manipulated to control its own composition. But in any ease, a cascade loop should be used. Steam is the most common source of heat to reboilers. Assuming that steam quality is reasonably constant, heat flow is regulated simply by a steam flow loop.

Occasionally hot oil is used to supply heat to the reboiler. In this event, heat flow is not linear with oil flow, as Fig. 9.3 verifies, and another indication must be used. Vapor loading in a tower is sensed as the differential pressure across the trays. A differential-pressure measurement from top to bottom of the tower can then serve as an index of vapor flow, with the trays acting like an orifice.5 Differential-pressure control by manipulation of heat input is a very fast loop, almost as f&t as an ordinary flow loop. Hence it is very responsive to variations in tower loading such as would be encountered where feed or product streams are in the vapor phase. In fact it is recommended in these situations even when steam is the heating medium.

If bottoms flow is chosen to be manipulated for composition control, bottoms level must be controlled from steam flow. This is not as simple a process as in most liquid-level loops, because of the reboilcr between the manipulated and the controlled variables. Since smooth control over boilup is mandatory, this loop must, be heavily damped. Consequently, a wide proportional band is necessary, and reset is relied upon to maintain level. Most of the vapor is eventually returned to 'the reboiler as liquid some time later, again affecting liquid level. This results in a natural period of several minutes. But because bottoms flow in this instance is small, relative to the rate of boiling, little coupling exists, and upsets are few.

### Pressure-control Methods

The energy balance must also be closed. If more liquid is boiled than condenses, there will be an increase in vapor inventory which will cause the pressure to rise. Pressure control is very important in order to maintain equilibrium.

If all of the vapors in the tower are condensable, the rate of heat removal can be adjusted by:

1. Changing the flow of coolant
2. Varying the heat transfer surface
3. Bypassing the condenser
4. Injecting a small amount of noncondensable gas into the condenser

Each of the above has certain advantages and disadvantages. Schemes (1) and (4) are perhaps the most common, but not necessarily the most effective. As seen in Chap. 9, the rate of cooling is quite insensitive to changes in coolant flow at the velocities under which most condensers are operated. This arrangement does conserve cooling water, however. An equal-percentage valve is recommended.

Scheme (2) is very effective in the absence of noncondensable gases. Figure 11.10 shows how the heat transfer area can be changed by flooding the condenser tubes with liquid. The reflux flow essentially sets the separation factor for the tower. If vapor flow into the condenser exceeds liquid flow out, condensate will rise to cover more heat transfer surface. This will cause a pressure rise, which in turn will reduce the heat input through the pressure controller. Because of the rapid response of vapor flow to heat input, this is a fairly fast loop.

Scheme (3), while effective, has the disadvantage of requiring an extremely large valve. Because the valve is in parallel with a fixed resistance (the condenser), a "quick-opening" characteristic is desirable. A butterfly valve is often used, however, because of its low pressure drop.

Scheme (4) intentionallv introduces noncondensable gas into the condenser, ultimately reducing its heat transfer capability to that required by the heat input. But a path must be open for release of the gas in the

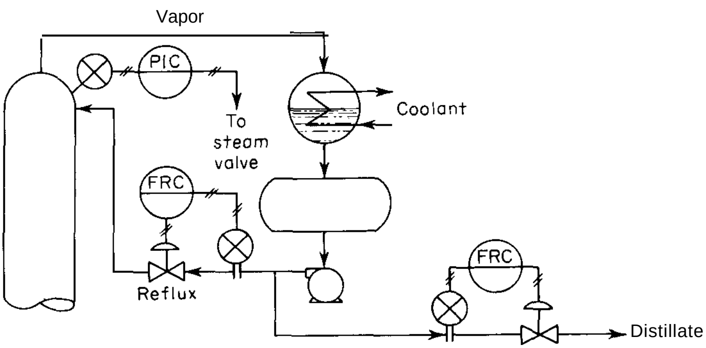



***FIG 11.10. With rising pressure, vapor flow is reduced, thereby exposing more heat transfer surface.***

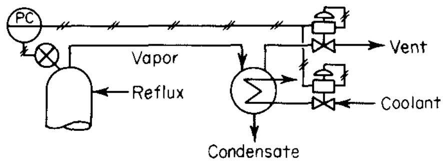



***FIG 11.11. Operating a vent valve in split range with the coolant valve will allow for the release of non-condensables.***

event of rising pressure. Since every cubic foot of gas added to the tower will eventually sweep a certain amount of product with it out the vent, this scheme is not recommended. Koncondensables also dissolve in the reflux and tend to reduce its boiling point, affecting temperatures on the top trays of the tower. This may invalidate the use of these temperature points for quality control. There is also the danger, particularly during startup, of blanketing the condenser tubes with gas, preventing heat removal altogether. The condenser must then be vented until normal condensation begins.

Columns operated at atmospheric pressure actually are controlled in this same way. If condensation is too efficient, a vacuum will start to develop, drawing air into the condenser. Eventually, the vapor mixture in the condenser will contain enough air to limit the rate of condensation to the rate of boilup. Every increase in boilup will expel some air and product; a decrease will draw more air into the system.

To allow for the release of whatever noncondensable gases may be contained in the feed stream, a vent valve must be added to whatever pressure-control scheme has been chosen. This valve would operate in split range with the main control valve.

Figure 11.11 shows a typical arrangement, with valve positioners used to effect the split-range operatipn. Both valves must be fully open at maximum controller output, but the vent valve ought to close before the coolant valve, on decreasing output. An acceptable sequence would have the vent valve closed below 50 percent output and the coolant valve fully open at 75 percent output. Both valves need equal-percentage characteristics.

If condensing area is limited, separation can be maximized by using the pressure controller to set the rate of heat input to the reboiler. In this way, just as much heat will be introduced as the condenser is capable of removing.

In a vacuum still, pressure can be controlled by manipulating a valve in the line leading to the vacuum system. As before, introduction of a noncondensable gas to control the vacuum is not generally recommended.

Sometimes overhead product is withdrawn as a vapor under flow control. This provides an escape for noncondensables, in which case any of schemes (1) through (3) may be used with a single valve. The flow of the product must be manipulated to control composition, however, not pressure.

## Manipulation of Reflux

Smooth delivery of reflux is extremely important in maintaining a steady composition profile within a column. Oscillations in flow will be propagated far down the column if they are long enough in period. This is the main reason why most columns are presently operated with a constant flow of reflux.

If distillate flow is to be manipulated for control of composition, reflux must be a dependent variable in all cases except that shown in Fig. 11.10. In operation without a flooded condenser, then, liquid level in the overhead accumulator must manipulate reflux. As mentioned in Chap. 3, control of the level of boiling and condensing liquids is complicated by transport and thermal problems. Furthermore, the need for heavy damping requires a wide proportional band, and subsequently, reset. It is therefore absolutely essential to close a loop around the control valve; otherwise hysteresis will promote an intolerable limit cycle. Either a valve positioner or a cascade flow loop will suffice for this purpose.

The difficulty in controlling accumulator level through the manipulation of reflux poses another problem, however. It is actually the material balance on the top tray which determines what the composition profile will be. Changing the rate of flow of distillate being withdrawn from the accumulator has no effect on composition if the flow of reflux or vapor is not altered accordingly. In the long run, the level controller will bring this about, but the time lag of the accumulator intervenes. In this sense, the capacity of the accumulator significantly impedes composition control.

If reflux flow were made to respond to the same control signal as distillate flow, the time lag of the accumulator could be eliminated. With this arrangement, a decrease in reflux flow would occur simultaneously with an equal increase in distillate flow, and the accumulator level would remain stationary.

But much more can be gained with such a configuration.[6] Reflux can be made to decrease more than distillate flow increases. This would cause accumulator level to rise, instead of falling as it did when reflux was left alone. The level controller will eventually return reflux to the correct steady-state value. But lead action has been introduced into the material balance at the top tray, increasing the speed of the composition loop severalfold. In effect, the accumulator has been converted from a disadvantage to an advantage, from a lag to a lead. Figure 11.12 shows how this is brought about.

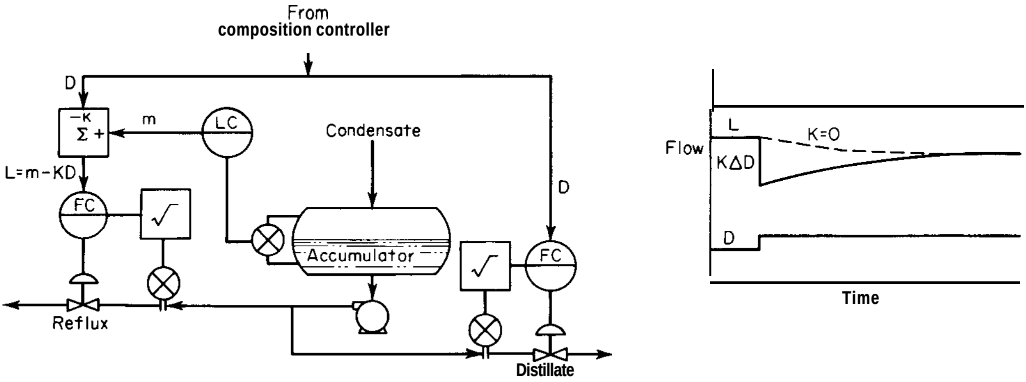



*FIG 11.12. This arrangement injects lead action into the composition loop.*

Reflux flow is programmed from distillate flow and the output of the level controller:

$$L = m - KD \tag{11.13}$$

As the curve in Fig. 11.12 indicates, reflux flow is given a lead-lag characteristic following distillate flow changes. Coefficient $K$ sets the lead-lag ratio, while the lag time varies with the time constant of the accumulator and the setting of the level controller. If $K$ is zero, as it would be in the absence of programming, reflux responds as a lag.

Because Eq. (11.13) is a summation, the flow signals must be linearized. Otherwise, lead time will be different with each value of flow. In practice, coefficient $K$ should be adjusted to minimize the period of the composition loop, although the actual setting is not especially critical. The fact that $K$ has a real value is of primary concern.

## Response of the Composition Loop

If a product analyzer is used for control, the response of the closed loop is considerably slower than if a temperature measurement were used. First, the analyzer would be located at one end of the tower, whereas the temperature element is normally nearer the center, sensing changes in the material balance sooner. Second, an analyzer usually suffers from delay in the sampling system, and a chromatograph, in particular, exhibits delays in the separation of its sample. A vapor sample is recommended to minimize response time.[6]

The dynamic response of composition to a change in distillate flow exhibits considerable dead time, as is expected in a multicapacity process. But the presence of an additional feature is indicated by step-response tests. Figure 11.13 illustrates results which are typically encountered. The response is the sort which would be seen in a transmission line with

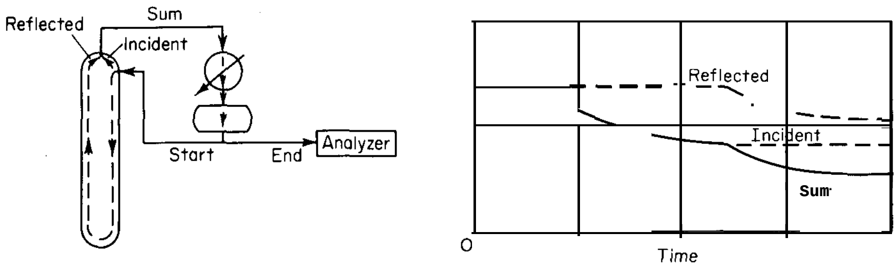



*FIG 11.13. The response of distillate composition to a step increase in distillate flow is the sum of an incident and a reflected wave.*

standing waves. A change in distillate flow first changes the material distribution in the top of the tower, which starts affecting composition. But the new distribution is propagated downward, too, by the resulting change in reflux rate, and then reflected from the bottom, returning sometime later. The step response is the sum of tn-o dead-time plus capacity combinations, the second dead time being much longer than the first.

Nonetheless, the first dead time dominates the control loop. It may be 5 to 30 min in duration, depending principally on the distance of the loop between the distillate valve and the analyzer. (The volume of the accumulator contributes significantly to the response.) The closed loop can then be expected to oscillate at some period between 20 min and 2 hr.

The ratio of effective dead time to effective time constant, will generally be found in the region of 0.15 to 0.30, as Fig. 2.4 indicated. The dynamic: gain of the tower can be determined from this ratio, which, combined with known values of process and transmitter gain, can provide an estimate of the required proportional band.

In the numerical example which has been used thus far,

$$\frac{dy}{d(D/F)} = -0.9$$

at the normal operating point. Then the process gain is

$$\frac{dy}{dD} = -\frac{0.9}{F}$$

It is customary to manipulate distillate flow through a flow controller, to prevent load changes in the stream from affecting composition. The gain $G_f$ of this flow loop to a set-point change takes the place of valve gain in this composition loop. In the case of a linear flowmeter, the gain is the maximum flow per 100 percent. Since $D/F$ is 0.5, a maximum flow of $0.5F$ is reasonable, making the gain $0.5F/100\%$.

Since the normal value of $y$ has been taken to be 0.95, an analyzer range of 0.90 to 1.00 seems reasonable. This is a span of 0.10. Transmitter gain is then

$$G_T = \frac{100\%}{0.1} = 1{,}000\%$$

The required proportional band is

$$P = \frac{400\tau_d}{\pi\tau_1} \frac{dy}{dD} G_f G_T$$

If a value of 0.25 is assumed for $\tau_d/\tau_1$,

$$P = \frac{400(0.25)}{\pi} \frac{0.9}{F} \frac{0.5F}{100\%} 1{,}000\% = 143\%$$

In this example, the specification on distillate purity was not severe. In many towers, though, extreme purity is demanded of the products, and narrow-span analyzers are used. If the span of the distillate analyzer in this case were 0.01, the proportional band would need to be 1,430 percent.

To improve speed of response, the analyzer sometimes samples the material on a tray closer to the center of the tower. But as with a temperature element, this does not ensure that the product itself mill always be of the desired quality.

Improved response speed can be combined with accurate composition control by means of a cascade loop.[7] It requires a temperature controller whose sensing element is located somewhere between the end of the tower and the feed tray, manipulating distillate flow. The set point of the temperature controller is then positioned in cascade by a composition controller sensing product quality.

### Control of Two Products

Feedback control over the quality of two products leaving a tower encounters severe coupling. It is not often tried and has, under certain circumstances, failed altogether. Derivation of the relative-gain matrix will reveal the reasons behind the difficulty.

Selecting $D$ and $V$ as the manipulated variables, Eqs. (11.3) and (11.6) can be solved for $y$ and $x$ in terms of $D/F$ and $S$. Differentiating,

$$\left.\frac{\partial y}{\partial(D/F)}\right|_x = -\frac{x - x}{(D/F)^2} \qquad \frac{(y - x)^2}{x - x}$$

$$\left.\frac{\partial y}{\partial S}\right|_x = -\frac{x}{1 + x(S - 1)} - \frac{Sx^2}{[1 + S(x - 1)]^2} = \frac{(1 - y)^2 x}{1 - x}$$

$$\left.\frac{\partial x}{\partial(D/F)}\right|_y = -\frac{-y}{1 - D/F} + \frac{z - yD/F}{(1 - D/F)^2} = -\frac{(y - x)^2}{y - z}$$

$$\left.\frac{\partial x}{\partial S}\right|_y = -\frac{y(1 - y)}{[y + S(1 - y)]^2} = \frac{x^2(1 - y)}{Y}$$

Although the four gains derived above represent closed-loop conditions, their inverse can be used to find relative gain in the same way as open-loop terms, following the procedure of Eq. (7.16). Accordingly,

$$\lambda_{yD} = \frac{1}{1 + \dfrac{x(1-x)(y-z)}{y(1-y)(z-x)}}$$

Being a 2 by 2 matrix, all other elements are determinable from $\lambda_{yD}$. Notice that $\lambda_{yD}$ varies between 0 and 1; for the example of the binary separation used throughout this chapter, $\lambda_{yD} = 0.5$.

Because separation is a function of V/F and nothing else, $\lambda_{yV} = \lambda_{yS}$; dimensional gain and hence nonlinearity disappear in the normalization procedure.

To see whether bottoms flow is, under certain conditions, more favorable to manipulate than distillate flow, a matrix of gains of $x$ and y with respect to ***B/F*** and $S$ may be prepared. But relative gain $\lambda_{yB}$ turns out to be identical to $\lambda_{yD}$. Thus the values of $x$ and y determine which composition controller should manipulate heat input and which should manipulate the material balance. In general, the purer product should be controlled by manipulating heat input.

If reflux and heat input are chosen to be the two manipulated variables, coupling is considerably different. Starting with a simplified description of the column parameters, observe what happens when $V - L$ is substituted for ***D:***

$$y = \lambda_{yD}D + (1 - \lambda_{yD})V = -\lambda_{yD}L + V$$
$$x = (1 - \lambda_{yD})D + \lambda_{yD}V = (\lambda_{yD} - 1)L + V$$

The coefficients of $L$ and $V$ are then normalized, following Eq. (7.16), to produce $\lambda_{yL}$:

$$\lambda_{yL} = \frac{-\lambda_{yD}}{1 - 2\lambda_{yD}}$$

Observe that $\lambda_{yD}$ and $\lambda_{yL}$ are mutually exclusive:

| | | | | | |
|---|---|---|---|---|---|
| $\lambda_{yD}$ | 0 | 0.25 | 0.5 | 0.75 | 1.0 |
| $\lambda_{yL}$ | 0 | -0.5 | $-\infty$ | 1.5 | 1.0 |

Coupling for manipulation of $L$ and $V$ is worse for all cases except 0 and 1. For the cited example, control of both products is impossible with this choice of variables-another point in favor of material-balance control.

## APPLYING FEEDFORWARD CONTROL

Feedback control of product quality from a column is not always satisfactory even when the control loops are properly arranged. Proper arrangement only protects the process from upsets in heat input, feed enthalpy, and rcflux flow and enthalpy. The most significant disturbances to quality control are generally variations in feed rate and composition.

From the previous example, it was pointed out that a composition controller may need a proportional band as high as 1,000 percent. And because the period of the closed loop may be from 20 min to 2 hr, reset time of 10 min to 1 hr is commonly encountered. The integrated error caused by a load change was shown earlier to equal the product of the proportional band times reset time. Distillation is characterized by a large proportional-reset, product', compared to other processes. And because integrated error in product quality can be costly, distillation is a prime candidate for feedforward control.

An on-line analysis of product composition is not always available. In these instances, there is no measurement to feedback from, so a forward loop can be a great help in maintaining control in the face of disturbances. Furthermore, if the real controlled variable is profit or loss, an optimum control program can be based on a feedforward model. Consequently the feedforward approach to control is of utmost importance in distillation processes, whatever the nature of the separation.

### The Material Balance

The basis for feedforward control of any mass transfer operation is the material balance. Earlier in this chapter the distillate to feed ratio was shown to be the principal factor affecting composition of either product stream. The feedforward control model is nothing more than an on-line solution to the material balance:

$$D = F\left(\frac{z - x}{y - x}\right) \tag{11.14}$$

Distillate rate is the manipulated variable; feed rate $F$ is one component of load and feed composition $z$ is the other. Either distillate composition $y$, or bottoms composition x is the set point, while the other depends on separation.

The most significant factor is that distillate flow is proportional to feed rate. Even if the forward loop is reduced to the simplicity of this ratio, it will bc of help, because feed-rate changes can occur instantaneously. If separation is constant, we know that x and $y$ can both be controlled. In this event distillate is directly proportional to feed rate throughout

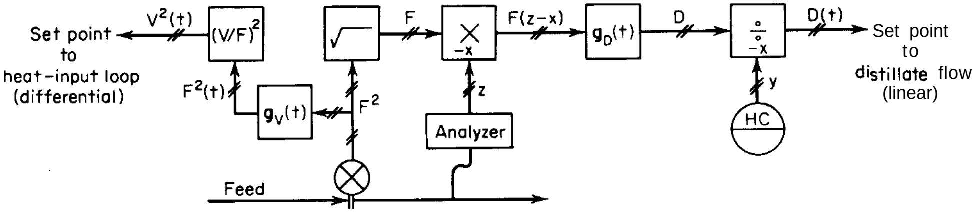



*FIG 11.14. Feedforward control of both products requires two forward loops.*

the operating range. But in order to achieve constant separation, a second forward loop is needed to adjust heat input proportionately to feed rate. A complete feedforward system for control of both products appears in Fig. 11.11.

The scheme shown in Fig. 11.14 has an analyzer on the feed stream but none on either product. As the principal failing of feedforward control systems is insufficient accuracy, a feedback loop on product quality is usually of considerable worth. In fact, if feedback is available, the exactness of the feedforward model can often be relaxed, even to the extent of omitting a feed-stream analyzer. Figure 11.15 shows how feedback control of product composition might be added to the forward loops.

Feedback from distillate composition is introduced through a divider in order to better compensate for the inverse relationship between distillate flow and composition, although exact correction is unnecessary. The principal factor is the maintenance of a ***D/F*** ratio, which in this case is adjusted as necessary by the output of the feedback controller.

### Maximum Separation

If maximum separation is desired at all feed rates, the heat input must be fixed at its maximum value, while feed rate is allowed to vary. It has been shown that this procedure can be used to control the quality of one

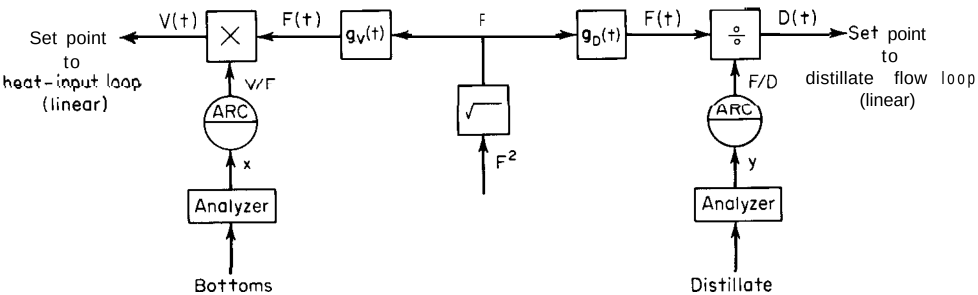



*FIG 11.15. Incorporation of a feedback loop can reduce the complexity of the forward loop.*

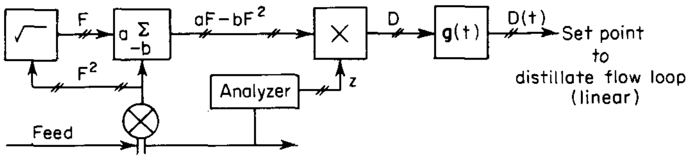



***FIG 11.16. The nonlinear relationship between D and F is characterized by the square-root converter and the summing amplifier.***

product while maximizing that of the other. From Fig. 11.6, control of distillate composition requires distillate to vary nonlinearly with feed rate, because bottoms composition varies. Although the relationship between *D* and *F* is somewhat complex, the curve can be modeled satisfactorily by a simple parabola:

$$D = z(aF - bF^2) \tag{11.15}$$

where a and $b$ are constants. The constants can be found by fitting the parabola to the curve of *D* vs. *F*. For the example in Fig. 11.6

$$D = z(1.04F \quad 0.08F^2)$$

A feedforward system designed to this model is shown in Fig. 11.16.

If bottoms composition is the controlled variable, the *D vs. F* curve bends upward, following the equation

$$D = z(aF + bF^2) \tag{11.16}$$

In a situation where the flow of bottoms product is much less than the flow of distillate, bottoms flow is manipulated from feed rate and composition. In this arrangement, the forward-loop calculation is complementary to that used when manipulating distillate flow. If no feedback loop is involved, accuracy of the forward loop is of paramount importance. Since the best accuracy will always be achieved by manipulating the smaller flow, the design of a control system cannot be intelligently undertaken without first writing a material balance across the tower.

### Multicomponent Separations

In general, feedforward control systems can be designed for multicomponent separations almost as easily as for binary separations. The first relevant question is how many product streams there are. If there are only two, distillate flow can be calculated as the sum of those components which pass overhead :

$$D = F(k_1z_1 + k_2z_2 + k_3z_3 \ . \ . \ \cdot) \tag{11.17}$$

where each $k$ is the recovery factor of its respective component The recovery factor is identified as that fraction of a certain component in the feed stock which will be recovered in the distillate stream. If component 2 in Eq. (11.17) were the principal overhead product, its recovery factor would be somewhat less than 1.0. Each lighter component has a higher recovery factor. Recovery factors for components heavier than the principal overhead product (sometimes called the "light key") approach zero.

From a material balance on a single component $i$, it will be observed that

$$Dy_i = Fk_iz_i$$

In the special case of a binary system, where a material balance is made only on the light component,

$$\frac{D}{F} = \frac{kz}{Y}$$

From previous investigation of binary systems it was shown that

$$\frac{D}{F} = \frac{z - x}{y - x}$$

Since the value of $x$ at constant $y$ varies with the separation capability of the tower, $k$ also varies with separation. It is not necessary to show how $k$ varies, because this has already been demonstrated in Fig. 11.5: $k$ is directly proportional to ***D/F.***

The most direct way of implementing a forward loop for a multicomponent system starts with analysis of the feed for the light key and lighter components, assuming the remainder to be heavy components, principally the heavy key. In a three component system,

$$D = F(k_1z_1 + k_2z_2 + k_3z_3)$$

But,

$$z_3 = 1 - z_1 - z_2$$

Therefore,

$$D = F[(k_1 - k_3)z_1 + (k_2 - k_3)z_2 + k_3]$$

The arrangement of the forward loop appears in Fig. 11.17.

If feedback control is available, it should be made to adjust the coefficient of $z_2$ through a divider (because $k_2$ varies as $1/y_2$).

A tower with a sidestream presents no particular problem as long as the material balance is given foremost consideration. In this type of tower, sidestream composition is usually the controlled variable. Increasing distillate flow mill reduce the concentration of light components in the sidestream. Increasing bottoms flow will reduce the concentration of heavy components in the sidestream. Increasing the sidestream flow will increase the concentration of both light and heavy components in it.

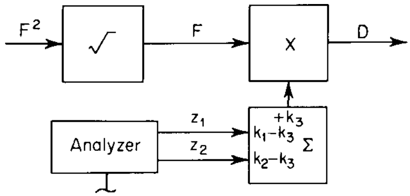



***FIG 11.17. The coefficients of $z_1$ and $z_2$ can be adjusted to obtain the desired product composition.***

## Optimizing Programs

Optimizing will be described here as that method of operating a tower which results in the least cost consistent, with specifications which have been set on the product.

For every controlled variable at least one variable must be manipulated. If more than one variable may be manipulated, their relative values can be made to follow a least-cost, program while still maintaining the controlled variable at its desired value. In order to provide a least-cost program with only one manipulated variable, no variable can be controlled-~ all must float with the program. In short! the number of manipulated variables always exceeds the number of controlled variables in any optimizing situation.

Suppose that a binary separation is taking place in a column whose heat input' and cooling have no value: they are waste from other parts of the plant. In this situation, heat input must be maintained at its highest possible rate, within the limitations of the heat transfer surfaces and vapor and liquid carrying capability of the tower. If the quality of one product is to be controlled, the quality of the other will be maximized at all rates of feed, as in Fig. 11.6. Maximizing the purity of the bottoms stream, for example, is the same thing as minimizing the amount of light component in that. stream. This is also seen to maximize the flow of distillate at all rates of feed, as Fig. 11.6 shows. Therefore this program results in control of distillate quality with the maximum flow of distillate: an optimizing program.

If control of both products is desired in this tower, minimum cost will only be achieved at maximum feed rate. This fixes all the variables at one value, and the tower is then operated at its constraints. If feed rate is variable, no optimizing is possible.

If the quality of neither product needs to be controlled, as in a refinery where various stocks are blended anyway, minimum cost operation can be achieved by appropriate manipulation of $D/F$. Let $v_1$ be the value of the light component and $v_2$ the value of the heavy. Then the cost of losses of light component with the bottom stream is $Bxv_1$. Total losses are

$$l = Bxv_1 + D(1 - y)v_2 \tag{11.18}$$

Substituting for $B$,

$$l = (F - D)xv_1 + D(1 - y)v_2$$

Dividing by $F$ yields

$$\frac{l}{F} = \left(1 - \frac{D}{F}\right)xv_1 + \frac{D}{F}(1 - y)v_2 \tag{11.19}$$

The curves of $x$ and $y$ vs. $D/F$ from Fig. 11.3 were modified by multiplying $x$ by $1 - D/F$ and $1 - y$ by $D/F$, and then assigning values $v_1$ and $v_2$. The results appear in Pig. 11.18. In this example, $v_1$ was chosen to be four times $v_2$. Because the slopes of the intersecting curves change monotonically, the minimum value of their sum occurs at their intersection. Having found the optimum value of $D/F$, the corresponding ratio of $1 - y$ to $x$ can be calculated:

$$\left(1 - \frac{D}{F}\right)xv_1 = \frac{D}{F}(1 - y)v_2$$

$$\left(\frac{1 - y}{x}\right)_{\text{opt}} = \frac{v_1}{v_2}\left[\frac{1}{(D/F)_{\text{opt}}} - 1\right] \tag{11.20}$$

The optimum value of $D/F$ is, of course, directly proportional to feed composition $z$, which was omitted in the previous explanation. Should

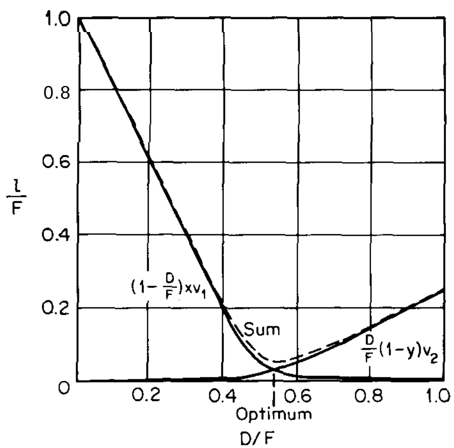



*FIG 11.18. Optimum D/F occurs at the intersection of the two curves.*

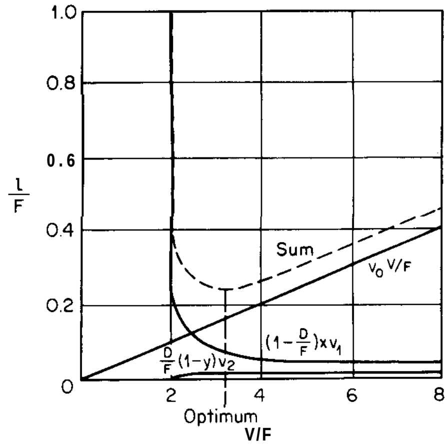



*FIG 11.19. Cost of heating plays a major role in determining the optimum V/F ratio.*

separation change, ***D/F*** remains constant only if $(1 - y)/x = 1.0$. If it is greater than 1.0, ***D/F*** must decrease with decreasing separation (increasing feed rate). Unfortunately, too many possibilities present themselves to touch on all of them. The intent of this section is to suggest how solutions to particular optimizing problems might be approached.

A second class of optimizing applications is characterized by a heat input whose value compares to that of the products. In this situation, the loss equation becomes

$$\frac{l}{F} = \left(\frac{V}{F}\right) v_0 + \left(1 - \frac{D}{F}\right) x v_1 + \frac{D}{F}(1 - y) v_2 \tag{11.21}$$

where $v_0$ is the cost of generating one unit of vapor.

If $y$ is to be controlled, ***D*** and $V$ can be manipulated together to minimize total loss as calculated above. Figure 11.5 shows that there is a value of ***D/F*** which can maintain control of $y$ for each value of ***V/F.*** Bottoms composition $x$ is seen to be dependent on $V/F$. Data taken from Fig. 11.5 were used to generate the three components of the loss equation, which are plotted and summed in Fig. 11.19.

As before, this particular set of curves is based on a feed composition $z$ of 0.50. Should $z$ change, ***D/F*** must change, which will shift the location of optimum ***V/F.*** A control program can readily be written manipulating ***D*** on the basis of ***F*** and $z$, and setting ***V/F*** as a function of the calculated value of ***D/F.*** Since there is always an upper limit placed on ***V,*** it is entirely possible that the optimum $V/F$ may be unobtainable at high rates of feed.

The loss plots of Figs. 11.18 and 11.19 are two-dimensional. Three-dimensional contour plots of $V/F$ vs. ***D/F,*** with $l/F$ as a parameter, can

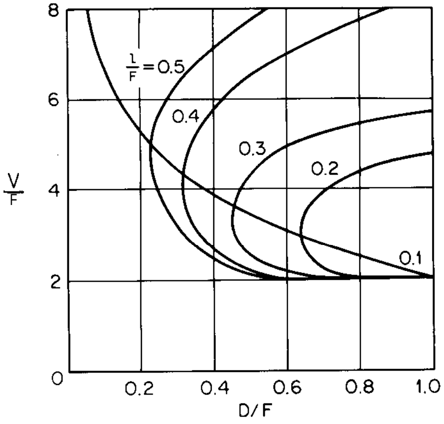



*FIG 11.20. The locus of optimum V/F represents the minimum loss for each value of D/F.*

readily be prepared, from which a locus of minimum loss can be found. Figure 11.20 is a combination of Figs. 11.18 and 11.19 in three dimensions.

### Dynamic Compensation

As is typical of feedforward control loops, dynamic compensation is necessary to ensure that the effect of a distillate-rate change be manifest at the same time as the feed-rate change which promoted it. Because feed enters the tower at a location considerably removed from where distillate is withdrawn, their dynamic effects upon composition differ by a corresponding amount. The response of a tower due to a change in feed rate appears as the sum of an incident and a reflected wave, just as is the case with distillate rate, but the incident path is longer and the reflected path is shorter. Figure 11.21 illustrates the difference in the length of the paths.

The response of distillate composition to equivalent step changes in feed and distillate flow rates is shown in Fig. 11.22. Because the

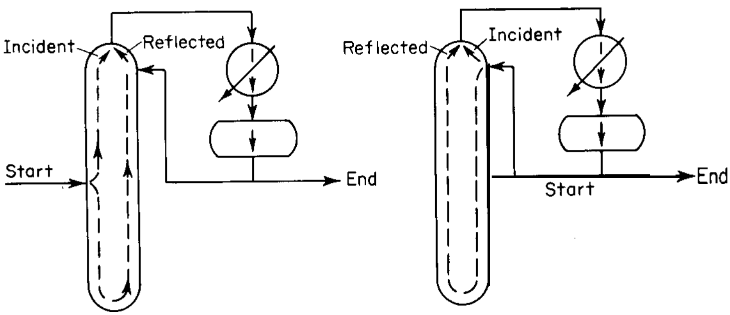



*FIG 11.21. The incident path is longer but the reflected path is shorter for a feed-rate change.*

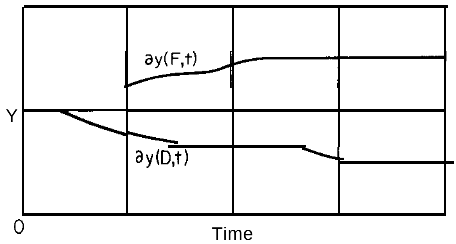



*FIG 11.22. The response to a step increase in feed rate exhibits a longer incident dead time but shorter reflected dead time.*

incident path from the feed tray is longer, the dead time of feed-rate response is longer. But the corresponding reflected wave travels a shorter path, resulting in a total elapsed time that is less than for an equivalent change in distillate rate. A feedforward loop without dynamic compensation (i.e., constant $D/F$) would produce a transient' step response that is the sum of these two curves:

$$dy\ (t) = \frac{\partial y}{\partial F\ (t)}\, dF\ (t) + \frac{\partial y}{\partial D\ (t)}\, dD\ (t)$$

The sum is plotted in Fig. 11.23.

Although the response of an uncompensated loop represents the difference in dynamic response between the two variables, the proper compensation corresponds to a ratio of the two. A ratio of the change in $y$ that is due to $F$ to the reverse of the change in $y$ that is due to $D$ (Fig. 11.22) is presented in Fig. 11.24. The step response of a compensator consisting of two lags and a lead is also included in Fig. 11.21. Although the model is not an exact representation of the process, it is the best approximation available within a three-component structure. A higher-order model would not only cost more, but would also be more difficult to adjust.

The foregoing compcnsution applies specifically to the case of withdrawal of distillate flow from a flooded condenser. It assumes constant liquid inventory. If reflux is manipulated from accumulator level, programming it with respect to distillate-flow changes according to Fig. 11.12 provides the required lead compensation. This source of lead action is recommended because it acts on rcflux rather than distillate.

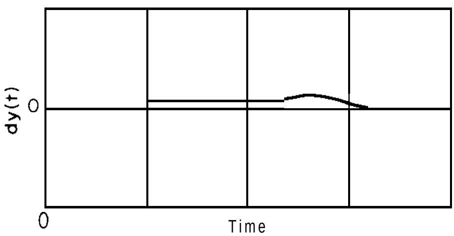



*FIG 11.23. Feedforward control without dynamic compensation produces an S-shaped step response.*

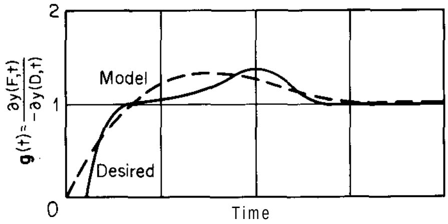



*FIG 11.24. The desired response can be modeled reasonably well with two lags and a lead.*

In fact, the compensation shown in Fig. 11.24 can then be achieved with a lag in the forward loop. This is especially desirable when feedforward control is applied to a train of closely coupled towers.

When bottoms flow is the manipulated variable, as in the alternate arrangement mentioned earlier, the dynamic response is different.[2] Without dynamic compensation, an increase in feed rate will cause an immediate increase in bottoms flow. This, in turn, will start bottoms level falling, thereby reducing heat input through the action of the bottoms-level controller. If bottoms flow is but a small fraction of the feed, the effect will be reduced. But dynamic compensation which will prevent a change in heat input is desirable. This would ordinarily take the form of a multiple lag.

Whenever heat input is manipulated through a forward loop, as in Figs. 11.14 and 11.15, dynamic compensation is considerably different. If a step increase in feed rate is converted instantaneously into a proportional increase in boilup, a momentous imbalance in the vapor-liquid distribution is forced up the column, Eventually the flow of reflux will increase in order to restore balance, and when equilibrium returns, the material balance will be unaffected. The sudden increase in boilup will, however, carry bottoms product upward, lowering the level in the reboiler, until reflux flow increases commensurately. Consequently, an analyzer or temperature measurement anywhere within the column will indicate a transient overcorrection.

Fortunately the remedy is simple. The dynamic element in the heat-input loop, g,,(t), should be a lag, adjusted to favor bottoms-composition regulation; the dynamic element in the distillate loop, $g_D(t)$, must also be a lag, adjusted for distillate-composition control after the other is set. This arrangement is favored even when reflux is manipulated from accumulator level; the steam loop acts as a formidable accelerating agent.

### Economic Justification

Feedforward control loops always contain more instruments than feedback loops. So there must be some justification for their use. Feedforward loops can be designed to maintain constant product quality and/or to operate the process at least cost. Economic justification is different in each case.

For control, justification must be based on the advantages to be gained over feedback. If there is no measurement available for feedback, there is also no contention. Optimizing programs fall into this category, because they have no measurement of profit or loss from which to feed back.

If feed rate and composition are invariant, there seems to be no purpose for a forward loop. Although feed rate to a column may be on flow control, this does not mean that it is invariant-it means that the stream is only subject to intentional disturbances. Supply of feed stock must come from somewhere, and its source cannot, have infinite capacity. The smaller the supply capacity, the more often feed rate will have to be adjusted. And whether feed rate is subject to random variations or intentional set-point adjustments, it can change rapidly-far more rapidly than a feedback loop on product quality can respond.

A forward loop on feed rate is very simple, just requiring a lead-lag unit and a divider between the feed-flow transmitter and the distillate-flow controller, as in Fig. 11.15.

The second forward loop from feed composition to distillate flow is much more costly and less valuable. Feed composition cannot- change instantaneously without supply sources being switched. If feed stock comes from a single source, whether a tank or another processing unit, its composition can only change at a limited rate. It is entirely possible, in many columns, that the time response of the feed source is slower than that of the product-quality feedback loop. In these cases, feedback control is quite effective in coping with variations in feed composition.

To add a forward loop on feed composition requires an analyzer, a multiplier, and possibly a summing device. Compare Figs. 11.14 and 11.17 with 11.15. Strictly speaking, this forward loop should have its own dynamic compensation, apart from that chosen for the feed-rate loop.[3] But because feed-composition changes are riormally so slow, they contain but small dynamic components, so the use of a separate compensator is hardly justifiable. In the absence of a feedback loop, a forward loop from feed composition is important. But if feedback is available, this forward loop is generally not warranted.

The worth of fcedforward control stems from four principal contributions :

1. Consistent product quality
2. Increased recovery
3. Reduced consumption of utilities
4. Reduced tankage requirements

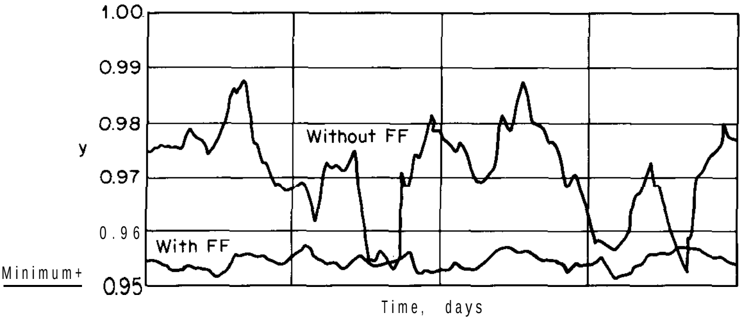



***FIG 11.25. Feedforward control can be expected to reduce variation in product quality by 4:1 or greater.***

The value of consistent product quality is the combination of many effects. Slight variations in quality can be corrected by blending-an additional operation. Major variations call for reprocessing of the material and possible shutdown of the tower or of other connected units. The cost of such troubles is great, and difficult to assess. Only after a long period of operation can the effects of improved quality control be judged.

In many plants, a certain minimum specification is placed on product quality. To ensure that the established minimum is not violated, even during times of severe load change, average product quality is often maintained at a safe margin above the minimum. Feedforward control can reduce the variation in quality during load changes, allowing average quality to be maintained closer to the minimum.

In Fig. 11.25, product quality without feedforward control is seen to vary from 0.95 to 0.99, averaging 0.97, in order to respect the established minimum of 0.95. If feedforward were only to improve control by 4: 1, the average could be reduced to 0.955, a reduction of 0.015. Since distillate flow and composition are inversely proportional, a reduction of 0.015 in average composition represents an increase in recovery by about the same amount:

$$\frac{D_2}{D_1} = \frac{y_1}{y_2} = \frac{97}{95.5} = 1.016$$

Increasing the flow of product by 1.6 percent with no additional operating costs can be quite attractive. Payout of the forward loop from this source alone could be only a few weeks.'

But suppose that the feedforward model were able to bring about perfect control. Average composition could then be reduced to 0.95, thereby improving recovery by $^{97}/_{95}$ = 1.021: an increase of only 0.5

percent more than before. The lesson to be learned from this example is that even a crude feedforward model may be capable of recouping the lion's share of recoverable losses. Hence absolutely perfect control is seldom worth attaining.

As far as savings in utility costs are concerned, those programs which manipulate heat input as a function of feed rate are obviously efficacious. Little else need be said in their behalf.

The savings that may be realized in the way of reduced tankage requirements apply principally to new installations. Since the prime function of the forward loop is to maintain control despite variations in feed rate and composition, it can actually take the place of tankage which would be used to absorb these variations. In addition, control of product quality is generally so well improved through the use of feedforward loops, that tankage used for smoothing and blending the product may also be unnecessary. In large installations, these savings alone could be worth several times the cost of the additional instruments.

Further savings may be accrued from a number of interrelated processes in the plant. If one of the products from the column is feed stock for a reactor, minimizing its impurities may mean reduced catalyst consumption. Each plant will have its own peculiarities in this regard.

## BATCH DISTILLATION

Although gradually diminishing in favor, batch distillation still is an interesting process to control, and deserves more than casual attention. Like most batch processes, its control system requires special consideration, ultimately bearing only faint resemblance to that of its continuous counterpart.

### The Material Balance

A batch separation will require an amount of time inversely proportional to the rate at which heat is introduced. Consequently, if processing time is to be minimized, heat input must be maintained at the maximum permissible level throughout distillation. This feature then fixes one of the variables which was subject to manipulation on a continuous tower. With vapor rate fixed, a material balance can be readily constructed for the batch still shown in Fig. 11.26.

If distillate flow is selected as the variable to be manipulated for product-quality control, reflux is then dependent. In the continuous system, product quality was affected by both $D/F$ and $V/F$. But here, $F = 0$, so it follows that product quality is a function of the ratio of the remaining variables, that is, $D/V$. In a sense, a batch still is similar to the enriching section of a continuous tower, part of whose vapor flow is feed. If, in the

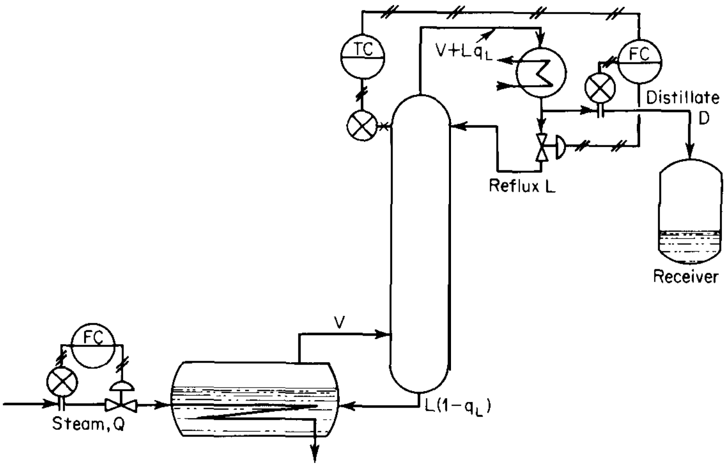



***FIG 11.26. With Vfixed, the only variable capable of independent manipulation is D.***

continuous tower, $V/F$ were maintained constant, manipulation of $D/V$ would also be manipulation of $D/F$.

Batch distillation is an unsteady-state process, because bottoms composition is continually changing as long as distillate is being withdrawn. If $D/V$ is constant, $y$ will change as $x$ changes. Constant $D/V$ is essentially constant separation and constant withdrawal of distillate. Distillate composition will then vary with time as the light component is removed from the tower. Starting with an initial charge $W_0$, containing $x_0$ mole fraction of the lighter component, at any time $t$ there remains

$$TV = W_0 - Dt \tag{11.22}$$

The bottoms composition $x$ at time $t$ can be found from the material balance of the light component:

$$\begin{aligned} Wx &= W_0x_0 - D\int y\,dt \\ x &= \frac{Woxo - D\int y\,dt}{Wo - Dt} \end{aligned} \tag{11.23}$$

Now $y$ can be found from $x$, using the modified Fenske equation.

The point to be remembered is that constant $D/V$ will produce variable distillate composition. Distillate is collected in a receiver whose final contents must meet a certain specification. Thus the average value of

$y$, designated $\bar{y}$, is the controlled variable:

$$\bar{y} = \frac{D\int y\,dt}{D\quad t} = \frac{\int y\,dt}{t} \tag{11.24}$$

Withdrawal of distillate is to be stopped when $\bar{y}$ falls to its desired value.

The disadvantages of constant-distillate-rate control are these:

1. If $D/V$ is relatively high, separation will be low, and withdrawal of distillate must be stopped at a relatively high value of $x$. This means that recovery of light ends will be poor.

2. If $D/V$ is reduced to enhance recovery, the distillation may consume an unreasonable amount of time and energy.

### Constant-composition Control

A more efficient way to operate a batch still is on the basis of constant distillate composition. Because bottoms composition continually changes, separation must also change if constant distillate composition is to bc maintained. Consequently, $D/V$ will be high at the beginning of a batch and will gradually be reduced to zero by the distillate-composition controller, when all of the recoverable product has been withdrawn. A control system operating on this basis appears in Fig. 11.26. The temperature controller would require reset action, to maintain constant quality with changing distillate rate.

Toward the end of the distillation, the flow of product will not only diminish, but its rate of change will also diminish. As a result, all of the light component cannot economically be removed, so a decision must be made as to when to stop withdrawal. This decision can be made on an economic basis by comparing the distillate flow times its value against operating costs. When the two are equal, the point of diminishing returns has been reached, as indicated in Fig. 11.27.

Another factor limiting the amount of recoverable distillate is the holdup of liquid in the trays of the column. Before the next higher boiling product can be withdrawn as distillate, the mixture held on the trays must be removed as a "slop cut." This material is collected in its own receiver and returned to the reboiler with a later batch of feed

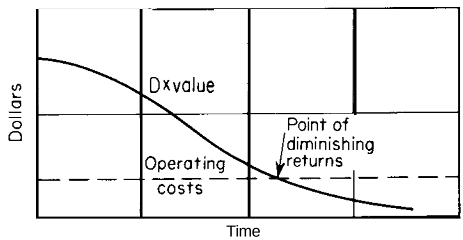



***FIG 11.27.** When the rate of remuneration equals the operating cost, distillation should be terminated.*

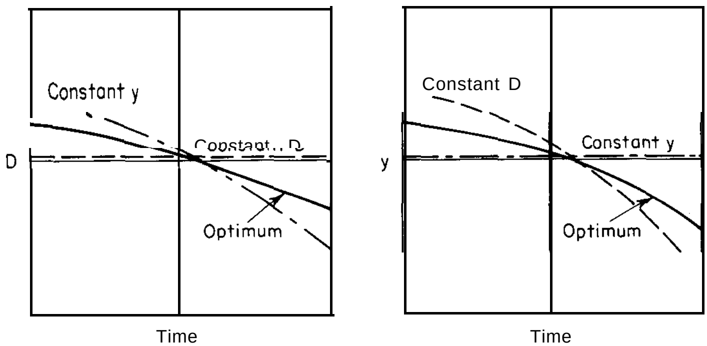



*FIG 11.28. Optimal distillate rate is intermediate between constant-rate and constant-composition programs.*

stock. The piping arrangement for withdrawal of distillate, shown in Fig. 11.26, has been designed to minimize holdup.

It can be seen that multicomponent separations may be accommodated without difficulty in a batch still. A separate receiver is necessary for each product, and manual operations are required to change receivers and to readjust the set point of the temperature (or composition) controller. But with each additional product cut there is also a slop cut. Hence as the number of products increases, the percentage of the batch recovered as product diminishes.

### Maximizing Product Recovery

Computer studies have shown that there is a program of distillate withdrawal which will recover the maximum amount of product of specified average composition in a specified time interval.[8] Figure 11.28 shows how this program falls between that of constant distillate rate and constant distillate composition. In effect, the final bottoms composition will be lowest because both y and $D/V$ are low when distillation is terminated.

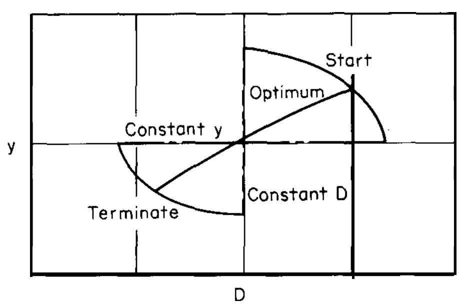



*FIG 11.29. The optimal policy varies distillate composition with rate in a programmed manner.*

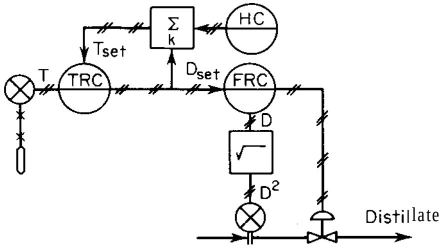



***FIG 11.30. The set point of the temperature (or composition) controller may be adjusted automatically with a summing device and a manual set station.***

Plotting distillate rate vs. composition for each of these three programs gives an indication of how the optimal program might be implemented. A typical plot, is constructed in Fig. 11.29. The optimal program calls for varying the set point of the temperature (or composition) controller based on the current value of distillate flow. Although the optimal program is not linear, it can be approximated to a satisfactory degree by a simple linear equation:

$$y = kD + y_0 \tag{11.25}$$

where $k$ = slope

$y_0$ = intercept

This linear expression may be readily implemented with the simple arrangement of analog devices pictured in Fig. 11.30.

## SUMMARY

Unfortunately it, is impossible to cover even a sampling of the variety of distillation columns that are in service in industry. They are nearly as individualistic as people. Consequently much is left to the practitioner in the way of interpreting the design rules contained herein in terms of his own problems. In this regard, a word of warning: do not attempt to make your particular separation fit the structure of the control system. Rather take care to mold the control system to the peculiarities of the separation.

One very important, class of separation is omitted from this chapter, however. It includes all the most difficult problems—extremely close-boiling mixtures and constant-boiling mixtures (azeotropes). The reason for the omission is that distillation alone is insufficient for their separation. They will be discussed in as much detail as seems reasonable after a brief treatment of extraction in the next chapter.

## REFERENCES

1. MacMullan, E. C., and F. G. Shinskey: Feedforward Analog Computer Control of a Supcrfractionator, Control *Eng.*, March, 1064.

Lupfer, D. E., and M. L. Johnson: Automatic Control of Distillation Columns to Achieve Optimum Operation, *ISA Trans.*, April, 1964.
Luyhen, W. L., and J. A. Gerster: Fcedforward Control of Distillation Columns, *Ind. Eng. Chem.*, October, 1964.
Fenske, M. R.: Fractionation of Straight-run Pennsylvania Gasoline, *Ind. Eng. Chem.*, May, 1932.
The Foxboro Company: Fractionating Column Heat Input Control by Pressure Drop Method, *Application Eng. Data* 282-14.
Van Kampen, J. A.: Automatic Control by Chromatograph of the Product, Quality of a Distillation Column, Convention on Advances in Automatic Control, Nottingham, England, April, 1965.
. Stanton, B. D., and A. Bremer: Controlling Composition of Column Product, Control Eng., July, 1962.
. Converse, A. O., and G. D. Gross: Optimal Distillate Rate Policy in Batch Distillation, *Ind. Eng. Chem.*, August, 1963.

## 'ROBLEMS

**11.1** For the column with $S = 361$ at $V/F = 5$, and $z = 0.50$, calculate the )$/F$ required to raise y to 0.97, and the resultant value of $x$. Estimate $dy/d(D/F)$.

**11.2** Repeat the above calculations for $V/F = 2.5$.

**11.3** A particular column is fed a binary mixture containing 80 to 90 percent ght component. Distillate is to be controlled to a purity of 99.9 percent. Vrite the feedforward control equation assuming a constant $V/F$ ratio. Repeat )r constant heat input.

**11.4** Feed to a tower contains 5 percent propane, 50 percent isobutane, and 0 percent normal butane, with the balance being higher-boiling components. 'he feed is analyzed by chromatograph for propane and isobutane. All the ropane in the feed goes out in the distillate. Under normal conditions, the ottom stream contains 2 percent isobutane, if the distillate composition is conrolled at 5 percent normal butane. Write the feedforward control equation, valuating all coefficients.

**11.5** In the example used in the text, the value of distillate is \$1.00/gal and hat of the bottoms product is \$0.40/gal. Steam costs \$1.00/1,000 lbs, and 1 lb s sufficient to vaporize 1 gal of product. Estimate the optimum $V/F$ ratio for ontrol of $y$ at 0.95, with $z$ at 0.50.

**11.6** Repeat the calculation for $z = 0.60$. Can $V/F$ be optimally programmed rom a measure of feed composition?

**11.7** Calculate $\lambda_{yD}$ for a column that is splitting feed containing 12 percent ower-boiling component into a 90 percent pure distillate and a bottoms product ontaining only 0.6 percent lower-boiling component.

# Other Mass Transfer Operations

# CHAPTER 12

Although distillation may be the most common mass transfer operation, it is also the most difficult to assimilate. Indeed, the separation between components is noticeably obscure, because they occupy the same phase. Other mass transfer operations involve separation or combination of different phases:

1. Vapor-liquid: absorption, humidification
2. Liquid-liquid (immiscible) : extraction
3. Liquid-solid : evaporation, crystallization
4. Vapor-solid : drying

Because of this distinction, one of the exit streams in each of the above is either pure, as the vapor from an evaporator, or in an equilibrium state independent of material-balance considerations. Although material-balance control can bc enforced in each of these mass transfer operations, the separation between phases generally simplifies its formulation by eliminating one variable. This reduces the number of manipulated variables by the same amount.

The final controlled variable in every case is composition, requiring some sort of an analytical measurement'. For most of these applications, a nonspecific determination, such as density, is sufficient. But occasionally, as in a drying operation, even nonspecific analyses are not available, so other variables must be found to provide some degree of regulation.

## ABSORPTION AND HUMIDIFICATION

Mass transfer between liquids and gases depends on the vapor pressure of the components as functions of temperature. Thus appropriate selection of operating temperature and pressure allows the reverse (desorption or stripping, and dehumidification) to be performed. The purpose of absorption and stripping operations is to remove and recover the maximum amount of a particular component from a feed stream. It is most efficiently accomplished in multiple stages, as in tray or packed columns. Humidification and dehumidification arc similar in principle, but are directed toward control of an environment short of equilibrium (e.g., <100 percent humidity); for them, a single stage is ordinarily sufficient.

### Equilibrium Mixtures of Vapors and Liquids

Each component in a vapor mixture exerts a partial pressure $p_i$ relative to its concentration $y_i$:

$$p_i = py_i \tag{12.1}$$

It can be seen that since the concentrations total 100 percent, the sum of the partial pressures is the total static pressure $p$ exerted by the system.

According to Raoult's law,' each component in an ideal liquid solution generates a partial pressure relative to its concentration $x_i$ in the liquid:

$$p_i = p_i^* x_i \tag{12.2}$$

The coefficient $p_i^*$ in Eq. (12.2) is the vapor pressure of component $i$ at the prevailing temperature. Unfortunately, wide departures from the ideal situation are encountered in typical solutions; nonetheless, linearity prevails over certain ranges, allowing $p_i^*$ to be replaced with an equilibrium constant $K_i$:

$$p_i = K_i x_i \tag{12.3}$$

The ideal situation is most closely realized where the gaseous components are above their critical temperature.

Combining Eqs. (12.1) and (I 2.2) or (12.3) establishes equilibrium conditions for a single stage:

$$y_i = \frac{K_i x_i}{p} \qquad (12.4)$$

If it is desired that $y_i/x_i$ exceed $K_i/p$, more stages must be used.

One unusual factor encountered in absorption is the temperature rise of the absorbing liquid due to condensation of the absorbed vapors. These vapors actually change to the liquid state and, in doing so, release their latent heat. If the system is adiabatic, the temperature of the absorbent rises, which shifts the equilibrium, tending to retard further absorption. If the solution is quite dilute, this heating effect may bc unimportant, but interstage cooling is necessary where high concentrations arc encountered. Absorption of HCl and $NH_3$ are typical of the latter situation. In stripping and humidification, hent must be applied to counteract the cooling effect of evaporation.

### Absorption

An absorption column is like the top half of a distillation tower. Feed vapor enters at the bottom and the depleted gas leaves the top. Figure 12.1 points out the flowing streams.

There are four streams, but vapor and liquid inventory controls manipulate two. Feed rate is the load; the only manipulated flow then available for composition control is absorbent stream $L$. The temperature of stream $L$ is also a factor, but for maximum absorption, it should be as low as practicable. For the same reason, pressure should be maintained at a high value.

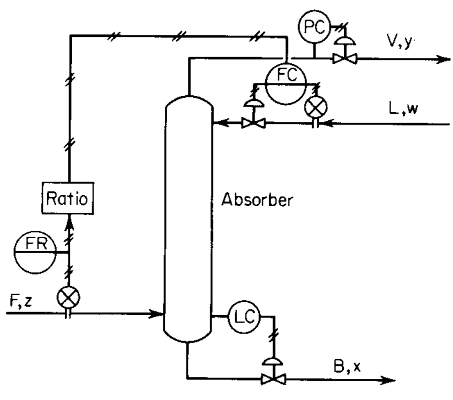



*FIG 12.1. The absorber features two liquid and two vapor streams.*

The uppercase letters in Fig. 12.1 represent molal flow rates, while the lowercase indicate the mole fraction of the principal absorbed component in the respective streams. An overall material balance requires that

$$F + L = V + B \tag{12.5}$$

The balance for the absorbed component is

$$Fz + Lw = Vy + Bx \tag{12.6}$$

If the other components in the vapor phase are not absorbed, another equation can be written to close out material balance:

$$V(1 - y) = F(1 - z) \tag{12.7}$$

The combination of Eqs. (12.5) through (12.7) permit's solution for the value of the manipulated variable $L$ required to control either y or $x$, the other being a dependent variable:

$$\frac{L}{F} = \frac{(z - y)(1 - x)}{(x - w)(1 - y)} \tag{12.8}$$

Notice the resemblance of Eq. (12.8) to the feedforward control equation for binary distillation.

As in distillation, there is a relationship between $y$ and $x$, of which Eq. (12.4) was a single-stage representation. Without attempting to arrive at a rigorous definition, it is important to point out that the ratio $L/F$ is the principal manipulated term, subject), however, to variations in feed composition.

Absorption is not a refining operation and is rarely the last operation conducted on a product. Consequently, close control of the concentration of either effluent stream is not paramount, and on-line analyzers are not often used. More importance is placed on minimizing losses (such as $Vy$) or total operating costs, for which the simple optimizing feedforward system was designed at the close of Chap. 8. In that example, as in the control equation (12.8), maintenance of a designated ratio of $L/F$ applies.

## Stripping

Absorption is usually followed by a stripping operation, in which the absorbed component is removed from the solvent. Stripping may also be carried out independently, to preferentially remove lighter components as dissolved gases from a liquid product.

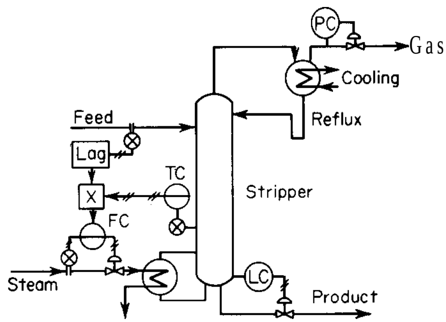



*FIG 12.2. In a stripping column, all the condensables are refluxed, all the noncondensables discharged.*

A stripping column appears quite like a distillation tower, equipped with both a reboiler and condenser. The reboiler raises the vapor pressure of all components, driving the most volatile preferentially up the column. A condenser is necessary to reflux whatever solvent might otherwise be carried away with the stripped vapors.

A tower for removal of volatile impurities in a liquid product, is shown in Fig. 12.2. Only the reflux would contain more dissolved impurities than the feed, which therefore enters near the top.

Because inventory control for vapor and liquid manipulate both effluent streams, as in an absorber, heat input is the only variable left for composition control. Since, in th is example, quality of the liquid product is the primary variable, control of temperature near the base of the column is used to specify its initial boiling point. Figure 12.2 shows how the temperature controller would be used to adjust the heat input to feed ratio. A lag is indicated in the forward loop, because the controlled variable is nearer to the manipulated variable than to the load.

When operated in conjunction with an absorber, the product becomes the vapor leaving the condenser, while the bottom stream is recycled to the absorber. A typical absorber&ripper combination for the separation of carbon dioxide and hydrogen is shown in Fig. 12.3. Monoethanolamine (MEA) is used as the solvent. Control of $CO_2$ content in the MEA leaving the stripper is only important for its influence on the equilibrium maintained with the gas leaving the top tray of the absorber-CO2 is not lost. Cooling the lean MEA enhances absorption, although its control is not really warranted. In addition, the absorber usually operates at a higher pressure than the skipper.

## Humidification

Cooling towers dissipate tremendous quantities of heat into the atmosphere through the process of humidification. Water circulated countercurrently to a stream of air is reduced in temperature owing to the fact

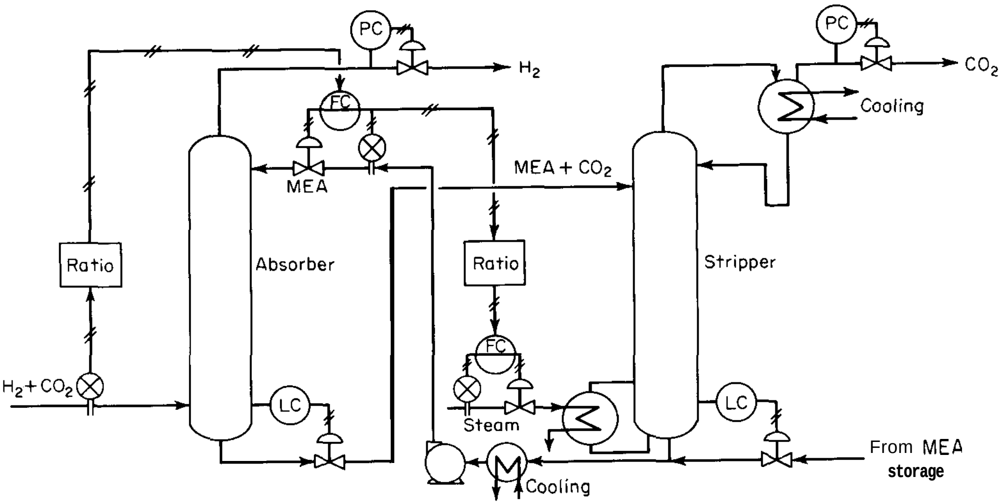



*FIG 12.3. The solvent is continuously recycled between the absorber and stripper.*

that atmospheric air is ordinarily far from saturated with water vapor. The latent heat of the evaporated water is converted into a change in sensible heat of the remainder.

Humidification and dehumidification also apply to environmental control where a certain moisture content is desired in the air. As pointed out earlier, an operation of this sort is generally conducted in a single stage, so control is actually not difficult. Yet the significance of the terms and principles is sufficiently confusing to deserve a general review and definition :

1. The vapor pressure of water in atmospheres varies with its temperature in degrees Rankine:

$$\log p_w^* = 6.69 - \frac{4407}{T} \tag{12.9}$$

2. Partial pressure $p_w$ was defined by Eq. (12.1). With regard to humidification, the liquid is essentially pure, so $x$ in Eq. (12.2) is 1.0. At equilibrium (100 percent saturation), the partial pressure of water vapor is equal to its vapor pressure at the prevailing temperature, that is, $p_w = p_w^*$.

3. Absolute humidity is the ratio of the mass of water vapor to the mass of air or gas in the mixture:

$$\text{Lb water/lb dry air} = \frac{18p_w}{29(1 - p_w)} \tag{12.10}$$

4. The mass of water per unit volume of humid air is sometimes used. Its units are typically[2]

$$\text{Grains/cu ft} = 1.73 \times 10^5 \frac{p_w}{T} \tag{12.11}$$

where $p_w$ is in atmospheres and 7' in degrees Rankine.

5 Relative humidity is the percent saturation at prevailing temperature and pressure and is exactly defined as $100p_w/p_w^*$.

6. Dew point is the temperature at which a mixture becomes saturated when cooled out of contact with liquid at constant pressure. It is often used to determine the moisture content of gases, by converting the temperature to vapor pressure by Eq. (12.9). Below 32°F, the dew point is actually a frost point.

7. Wet-bulb temperature is the equilibrium temperature reached by a small amount of liquid evaporating adiabatically into a large volume of gas. Equilibrium exists when the rate of heat transfer from the gas to the cooler liquid equals that consumed by evaporation. It is affected by heat and mass transfer coefficients as well as humidity, therefore is dependent on maintaining turbulent gas flow around the bulb. Humidity can be determined from wet-bulb, $T_w$, and dry-bulb, $T$, temperatures by following the adiabatic-saturation curves on a psychrometric chart,, or by

$$T - T_w = 0.146H_v\left(\frac{p_w^*}{1 - p_w^*} - \frac{p_w}{1 - p_w}\right) \tag{12.12}$$

where $H_v$ = latent heat of evaporation

$p_w^*$ = vapor pressure at the wet-bulb temperature

Humidity measurements may be made by several different means, wet-bulb temperature being but one. Some instruments are equipped with a hair element which is sensitive to changes in relative humidity. Though dew point may be measured directly, a more reliable instrument[3] uses a hygroscopic salt whose conductivity varies with moisture content. The salt is self-heated simply by application of an a-c voltage, and its temperature is an indication of the absolute humidity. The measured temperature is not the dew point', but is related to it such that scales are available for direct reading in dew point or units of absolute humidity.

Choice of the type of measurement to be used for control depends on the process. Under isothermal conditions, the moisture content of solid materials varies with relative humidity, but in adiabatic processes, a determining factor is wet-bulb temperature. An exact analysis of moisture content can best be found by an absolute-humidity measurement, however.

Control of humidification involves manipulation of heat input or air flow to a system containing excess water. A spray chamber for humidification is shown in Fig. 12.4.

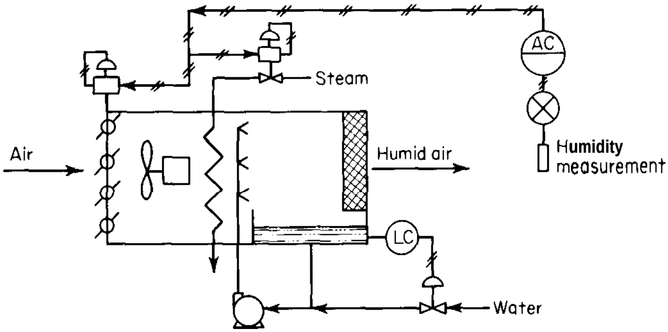



*FIG 12.4. If the influent air is very dry, heat may not be required, and the louvres are manipulated.*

Dehumidification requires cooling of the humid air, with or without compression, depending on the dryness required. Manipulation of cooling under constant pressure is effective.

## EVAPORATION AND CRYSTALLIZATION

These operations may he conducted separately or in combination in an effort to separate a solid from its solvent. The product from an evaporation is a concentrated solution, whereas a crystallizer discharges a slurry of crystals in a saturated solution. These two operations may not be technically classified as mass transfer, in that no equilibrium exists between the composition of the two phases-the vapor leaving an evaporator and the crystals in the crystallizer are both essentially pure. Yet the control of both these operations is heavily dependent on the material balance.

### Multiple-effect Evaporation

To conserve steam, evaporation is usually carried out in two or more stages, each stage being heated by the vapors driven from the previous stage. To maintain a temperature difference across each heat transfer surface, a pressure difference must be controlled between stages. The most economic operation is realized with low-pressure steam heating, requiring each stage to be maintained under a different vacuum. A double-effect evaporator is shown in Fig. 12.5; recognize that the arrangement could be extended indefinitely, but the practical limit seems to be six effects.

The arrangement shown in Fig. 12.5 is forward feed, in that the feed stream enters the first effect only. Backward feed, i.e., entering the last

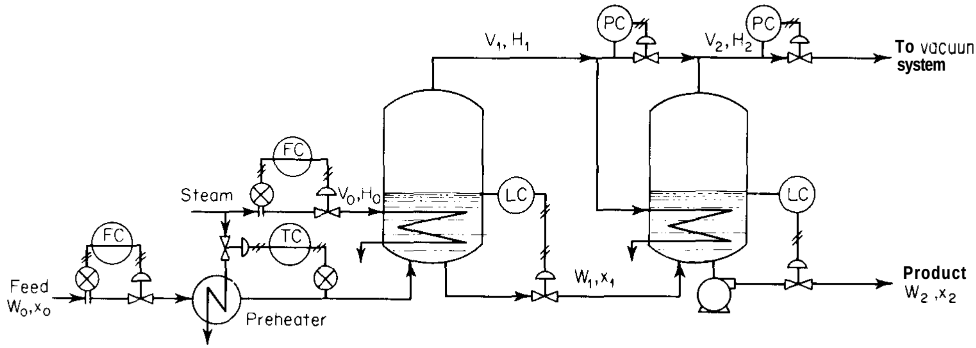



***FIG 12.5. A double-effect evaporator with*** forward ***feed.***

effect first, is another possibility. In addition, each effect can receive fresh feed, which arrangement, is called parallel feeding. The first described is the most common.

The controlled variable is product concentration. It can be determined by density measurement, electrolytic conductivity, refractive index, or by measuring the elevation in boiling point or the depression in freezing point of the solvent.

In the past, control of product composition typically entailed manipulation of the discharge valve. The level controllers for each effect were left to manipulate each inflow, ultimately affecting feed rate. This arrangement results in a series of interactions between flows and compositions from the last effect to the first and back again. Furthermore, production rate can only be adjusted by altering the heat input, which constitutes a prime source of disturbance. These deficiencies prompted the investigation of material-balance control.

**Material-balance Control**

A certain amount of solvent is evaporated in each effect relative to its heat input; all the solids in the feed are discharged with the product. Let $W_1$ represent the mass flow of feed whose solids content is $x_1$ (weight fraction) such that X is the mass flow of solids in the feed:

$$x = W_1 x_1 \qquad (12.13)$$

The total flow of solution leaving the effect, $W_2$, contains $x_2$ weight fraction of solids:

$$x = W_2 x_2 \qquad (12.14)$$

The rate of evaporation is designated $V_2$:

$$V_2 = W_1 - W_2 \qquad (12.15)$$

By combining Eqs. (12.13) through (12.15), it is possible to calculate the rate of evaporation required to convert a feed of known composition to a specified product composition:

$$V_2 = W_1\left(1 - \frac{x_1}{x_2}\right) \tag{12.16}$$

The heat input to the effect, in the form of vapor or steam, will flow at a rate $V_1$ with a latent heat $H_1$ in order to cause the evaporation of $V_2$, whose latent heat is $H_2$, if the feed is preheated to the boiling point:

$$V_1H_1 = V_2H_2 \tag{12.17}$$

Combining the last two expressions gives the relationship between the input variables necessary to maintain a desired output quality:

$$V_1H_1 = W_1\left(1 - \frac{x}{x}\right)H_2 \tag{12.18}$$

To apply this to a double-effect evaporator, let Eq. (12.18) represent conditions existing in the second effect. The material balance for the two effects can be derived in the same way as Eq. (12.16), relating total evaporation to inflow rate $W_0$ and weight fraction solids $x_0$:

$$V_1 + V_2 = W_0\left( - \frac{x_0}{x_2}\right)$$

Relating first-effect vapor inflow $V_0$, of enthalpy $H_0$, to $V_1$ and $V_2$ as was done in Eq. (12.17) permits elimination of the latter two variables:

$$V_0H_0 = \frac{W_0(1 - x_0/x_2)}{1/H_1 + 1/H_2}$$

Extension to an n-effect evaporator follows directly:

$$V_0H_0 = \frac{W_0(1 - x_0/x_n)}{\sum_{i=1}^{i=n} 1/H_i} \tag{12.19}$$

Enthalpies through subsequent effects can be represented by an average value $H$ which is slightly greater than $H_0$ because of decreasing pressure in each effect. The denominator in Eq. (12.19) can therefore be approximated by $n/H$.

Equation (12.19) may be implemented for control of product quality by manipulating either heat input or feed rate in relation to the other. The choice depends on the relative availability of each. If short-term reductions in steam availability are common, feed rate should be manipulated accordingly. But if feed is coming from another processing unit without intermediate surge capacity, the alternate arrangement is favored.

Typical measurements of input variables would be from differential flowmeters and a density transmitter on the feed stream. The steam-flow measurement may require correction for static pressure. Not only does the specific volume of saturated steam change with pressure, but its enthalpy does too. The mass flow of steam varies with measured differential $h_s$ and specific volume $v_s$:

$$V_0H_0 = K\sqrt{\frac{h_sH_0^2}{v_s}}$$

The ratio $H_0^2/v_s$ is found to be linear with steam pressure $p$ over a reasonable operating range :

$$V_0^2H_0^2 = h_s(a + bp) \tag{12.20}$$

In similar fashion, the feed-flow differential $h_f$ must be compensated for density $\rho$, which additionally determines solids content:

$$W_0^2\left(1 - \frac{x_0}{x_n}\right)^2 = K_fh_f\rho\left(1 - \frac{x_0}{x_n}\right)^2$$

Since $x_n$ is a constant,, all variables dependent on feed density can be lumped into another linear function:

$$W_0^2\left(1 - \frac{x_0}{x_n}\right)^2 = h_f(f - g\rho) \tag{12.21}$$

The complete feedforward equation for the manipulation of feed rate in response to steam flow is

$$h_f = \frac{(n/H)^2h_s(a + bp)}{f - g\rho} \tag{12.22}$$

Coefficients a, $b$, $f$, g, and $n$ are all fixed; $H$ may vary somewhat.

Figure 12.6 illustrates how the feedforward control system might be designed for a multiple-effect evaporator fed from a surge tank. An

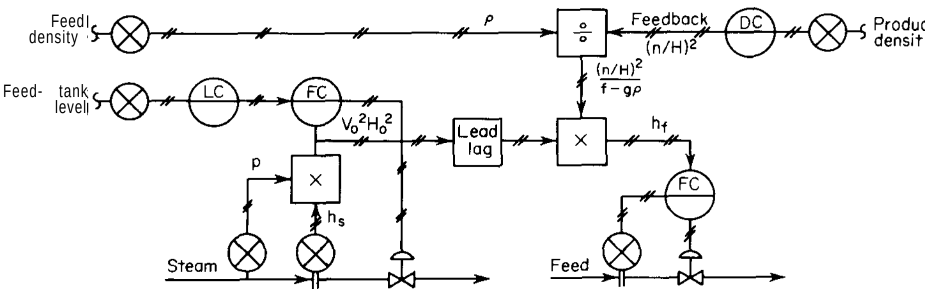



*FIG 12.6. The feedforward system corrects for variations in feed density and steam Row.*

average-level controller on the surge tank would set steam flow in cascade. Momentary reduction in steam availability would be accommodated by allowing the tank level to rise until the situation is corrected.

In very fast evaporators of the falling-film type, where liquid holdup is minimal, feed-rate changes may reach the product before steam-rate changes do. In these applications, proper dynamic compensation requires a lag in excess of the lead time. And because of the low liquid volume, dead time is proportionately large, making feedback correction in the way of adjustments in $H$ somewhat difficult. Manipulating a small flow of steam directly to the last effect to form a tight feedback loop has proven effective.[4] This loop should incorporate proportional and derivative modes only. Notice the similarity of this process to the once-through boiler, whose control system is described in Fig. 9.11. The feedback control functions for product quality have been split into transient and steady-state components in each case.

### Control of Crystallizers

A solution is saturated when an equilibrium exists between dissolved and undissolved molecules of a solute in a solvent. The concentration of undissolved solute present as crystals does not affect the equilibrium state. The concentration of solute in solution is fixed by the equilibrium state, which varies with temperature.

As a result, crystals may be deposited from solution by either of two mechanisms :

1. Evaporation of solvent
2. Reduction in solution temperature

Evaporation can be caused by the application of heat or vacuum or both, but if vacuum alone is used, the temperature of the solution is reduced along with the evaporation.

A usual requirement is control over the concentration of crystals in the discharge slurry. In many cases, however, crystal size is important as well. Crystal concentration is customarily measured as density, if the crystals are uniformly dispersed across the sensitive span of the detector. Crystal size determination unfortunately does not lend itself to on-line analysis.

Figure 12.7 shows a typical cooling crystallizer. Temperature of the solution is maintained by circulating the slurry through a chiller which removes sensible heat in the feed stream and heat of fusion of the crystals. The crystal slurry must be kept in motion to avoid plugging. A centrifuge or filter subsequently separates the crystals and returns the mother liquor to the process.

Since fine particles settle slowly, they accumulate at the top of the mass

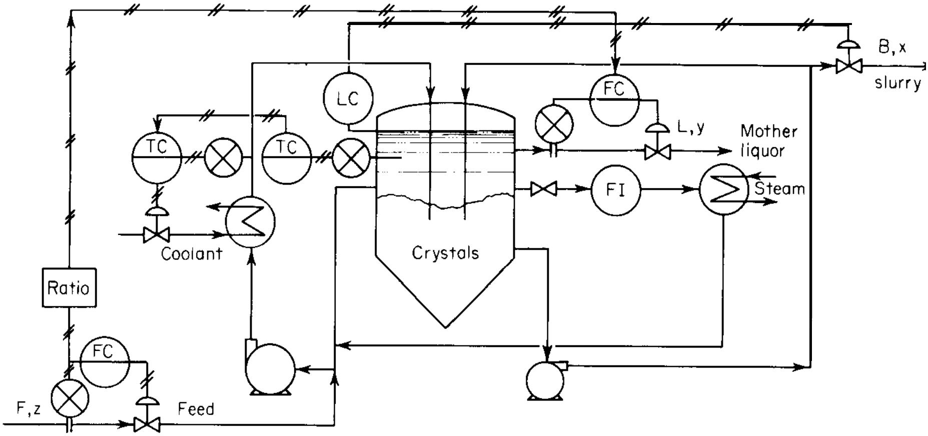



***FIG 12.7. Control of crystal content involves manipulation of mother-liquor and slurry flows.***

of crystals. By withdrawing a sidestream from this region, a token amount of fine crystals can be redissolved, increasing the average size of the crystals remaining. Increasing the density of the slurry tends to increase crystal size by raising the level of the mass relative to the side-stream tap.

Examination of the material balance across a crystallizer gives an indication of how it ought to be controlled. Mass flow of feed, $F$, is separated into saturated mother liquor $L$ and crystal slurry $B$:

$$F = L + B \tag{12.23}$$

Weight fractions of the solute in each stream are represented by $z$, $y$, and $x$, respectively:

$$Fz = Ly + Bx \tag{12.24}$$

Load variables are $F$ and $z$; $y$ is the weight fraction of solute in solution at saturation as fixed by temperature, so it is not a variable; $L$ and $B$ remain to be manipulated so as to control x and the crystallizer level.

Following the usual procedure for material-balance control, $L$ is selected to be manipulated for density ($x$) control because its flow is readily measurable, whereas that of the slurry is not. Eqs. (12.23) and (12.24) are therefore solved for $L$, eliminating $B$:

$$L = F\left(\frac{x - z}{x - y}\right) = F\left(1 - \frac{z - y}{x - y}\right) \tag{12.25}$$

Notice the similarity to the control equation for distillation-but in this case y is fixed, considerably simplifying matters. The content of crystals in the slurry is $(x - y)/(1 - y)$, whereas total solute content is $x$.

The most significant outcome of the above derivation is that mother-liquor flow should be set in ratio to feed rate. Feedback from crystal density is not mandatory, because the process is self-regulating and extreme accuracy is not usually warranted.

Evaporative crystallizers may be treated just like evaporators, since pure solvent is driven off by a proportional flow of heat. The only difference is that the part of the product which remains in solution is a function of solution temperature. The latter is not an independent variable, however, because the solution is boiling.

## EXTRACTION AND EXTRACTIVE DISTILLATION

Extraction is defined as the transfer of a dissolved material between two immiscible solvents. The material being transferred may be solid, liquid, or gas in its ordinary state. The purpose of the extraction is to permit its separation from the first solvent or some contaminant in it. Many liquids whose boiling points are nearly identical, preventing their distillation, can be separated by extraction, often in conjunction with distillation.

A serious problem area in distillation technology involves substances which form constant-boiling mixtures, called azeotropes. Azeotropes are encountered so often in chemical systems that their separation deserves special mention apart from the conventional distillation practices discussed in the previous chapter. Because the principles of extraction play a major role in azeotropic distillation, the subject will be discussed in detail.

### Extraction and Decantation

In a single extraction stage, an equilibrium will be approached between the concentration of product in the two exit streams. Consider a product dissolved in solvent $A$ being extracted into solvent $B$. If $x_1$ is used to represent the concentration of product remaining unextracted, and $y_1$ that of the product leaving in $B$, then

$$x_1 = Ky_1 \tag{12.26}$$

where $K$ is the equilibrium constant.

Similarly, a material balance may be drawn across the stage:

$$A(x_0 - x_1) = B(y_1 - y_0) \tag{12.27}$$

Recognize that $A$ and $B$ are mass flow rates of the two solvents, and that $x_0$ and $y_0$ are the concentrations of product in each influent stream.

It is usually convenient to manipulate the flow of extractant $B$ in response to changes in feed rate $A$ and composition $x_0$, in order to control the concentration of extracted product. If this is indeed the case, Eqs. (12.26) and (12.27) may be solved for $B$ in terms of $y_1$:

$$B = A\left(\frac{x_0 - Ky_1}{y_1 - y_0}\right) \tag{12.28}$$

Note once more the familiarity of the material-balance equation, particularly the ratio existing between $A$, the feed rate, and $B$, the manipulated variable Whether single or multistage, control of extraction always involves manipulation of this ratio.

Multistage extraction may be carried out in a series of mixers, each followed by a settling chamber, where the solvents are separated and allowed to flow countercurrently. However, extraction is more commonly conducted in packed towers, with extractant flowing countercurrently to the feed. The less dense solvent must enter at the bottom, flowing upward at a rate determined by its difference in density from the heavier solvent.

Not only must liquid level be controlled, but interface as well, in order to maintain inventory of both solvents. Figure 12.8 shows how the control loops would be arranged for a typical extraction column. The location of the interface between two solvents is easily measured as differential pressure between two taps or buoyant force on a displacer. In either case, the measurement reflects the average density across the vertical span of the instrument.

Decanters arc used to separate two immiscible solvents following extraction, mixing, or condensation from the vapor phase. Like extrac-

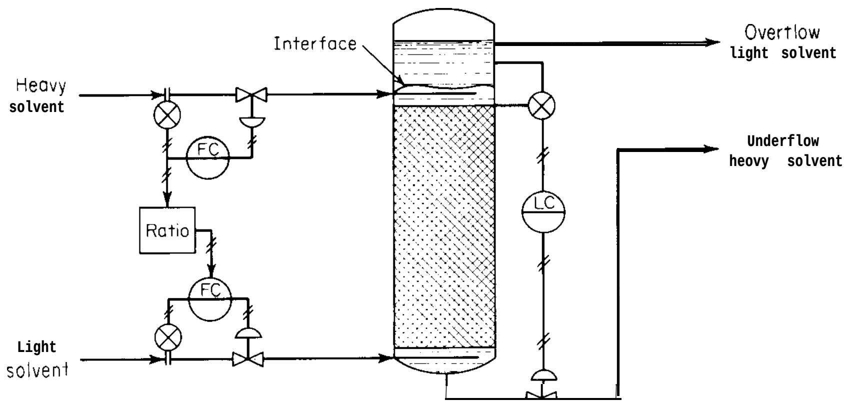



*FIG 12.8. Both liquid level and interface must be regulated in an extractor.*

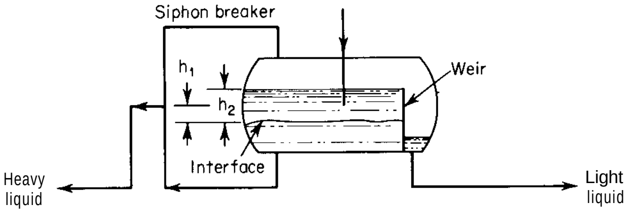



*FIG 12.9. A properly designed decanter may function without controls.*

tors, both level and interface height must be regulated, but fortunately, in many cases this can be accomplished simply by proper equipment design. Figure 12.9 points out the dimensions which are important.

The static differential pressure between the top of the light liquid and the highest point in the underflow loop is zero. The vertical distance from the interface to each of these points is a function of the densities of the heavy and light liquids, designated $\rho_1$ and $\rho_2$:

$$\rho_1 h_1 = \rho_2 h_2 \tag{12.29}$$

For finite flow rates of heavy liquid, $h_1$ will decrease, i.e., the interface will rise. This piping arrangement is not directly applicable to control of the extraction tower shown in Fig. 12.8, because of the variable pressure drop encountered in the packing and discharge line.

## Azeotropic Distillation

An azeotrope is a mixture of two or more materials that cannot be separated by distillation; the vapor and liquid in equilibrium are of the same composition, and there is no difference between the boiling point and the dew point of an azeotrope. The individual components may have entirely different boiling points, but their azeotropic mixture will exhibit a higher or lower boiling point than either, the latter being more common.

Azeotropes act like pure substances. Ethanol and water form an azeotrope containing about 89 mole percent ethanol. Any mixture of ethanol and water containing more than 89 percent ethanol may be fractionated into ethanol and the azeotrope; a mixture containing less than 89 percent ethanol can be fractionated into water and the azeotrope. The composition of an azeotrope, and its boiling point at a given pressure, are characteristics peculiar to the system.

Heterogeneous azeotropes separate into two immiscible layers of different composition when condensed. This is a considerable advantage, for it permits separation into the pure components by decantation, followed by a second distillation. Figure 12.10 shows that both columns use the

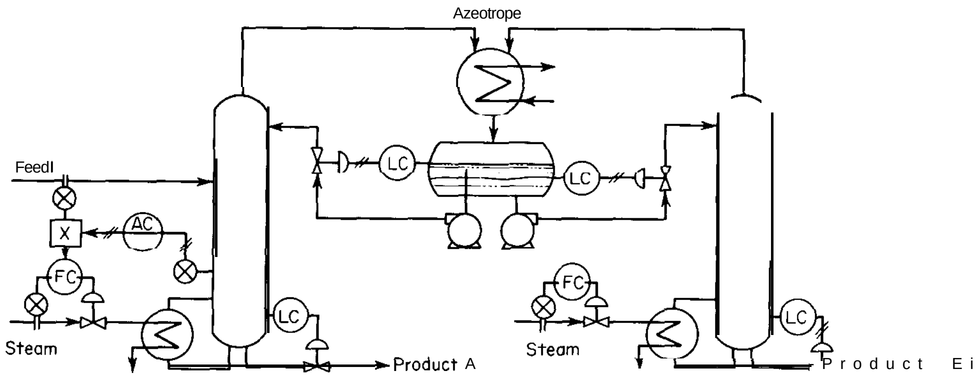



***FIG 12.10. A heterogeneous azeotrope may be separated by two columns with a single condenser.***

same condenser, since their vapors are both of the azeotropic composition. Steam-flow rates are the only variables which may be manipulated for composition control.

Occasionally a binary azeotrope which is homogeneous can be broken by adding a third component which forms a ternary heterogeneous azeotrope. The third component is called an "entrainer"; it is not intentionally removed from the system, but circulates in the reflux loop. The ternary azeotrope must boil at a lower temperature than the binary in

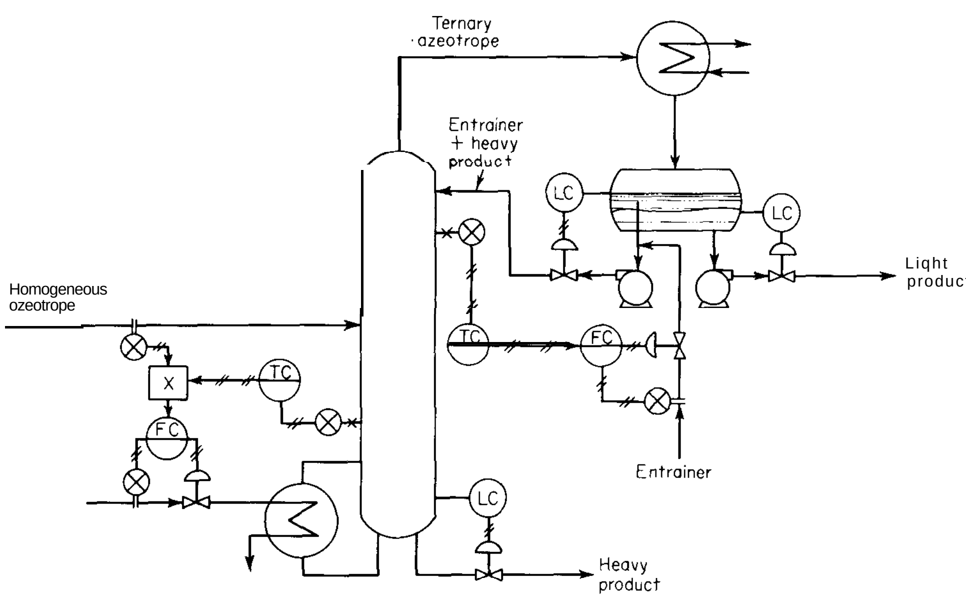



***FIG 12.11. Top temperature is sensitive to inventory of entrainer.***

order for the operation to be successful. Figure 12.11 shows the similarity of the process to separation of a binary heterogeneous azeotrope. Again, removal of the light component is directly proportional to the vapor flow, so heat input must be manipulated for bottoms-composition control.

Although entrainer is intended to circulate continuously through the reflux loop, a certain amount will be lost. with the light component,. Because the entrainer is a third component, its inventory cannot be detected by liquid level. However, a deficiency of entrainer resulting in the loss of the ternary azeotrope will cause temperatures in the top of the tower to rise. Therefore top temperature can be used to manipulate its addition.

**Extractive Distillation**

Extractive distillation is a technique used to break a homogeneous azeotrope or to facilitate separation of close-boiling components. A high-boiling solvent with a particular affinity for one of the components is introduced to lower its vapor pressure. The other component is then readily distilled from the solution and a second column used to recover the solvent. The flow sheet, is given in Fig. 12.12.

Since the role of the solvent is one of extraction, it must be set in ratio to the feed. Insufficient solvent will result in poor separation, whereas excess is not detrimental, except as it increases the heat load on the towers. Notwithstanding the extraction, however, material-balance control must still be maintained on the first tower. If too much distillate is withdrawn, it will necessarily be contaminated with solvent, extracted product, or

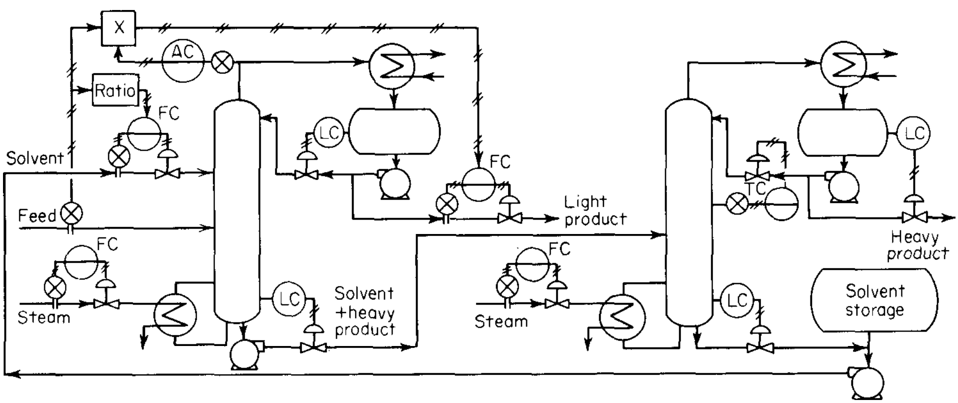



*FIG 12.12. Enough solvent must be supplied to extract the heavy component from the feed.*

both; if too little is withdrawn, some of the lighter product will be carried into the second tower.

As solvent is introduced above the feed tray, the separation beyond that point is principally between solvent and the light component, which does not require many trays. A large number of trays is required below, however, to sufficiently reduce the content of the light component in the ternary bottom mixture. Separation in the second tower is easy, if the solvent is not especially volatile.

## DRYING OPERATIONS

The drying of solids has historically defied control principally because a continuous measurement of product moisture has been lacking. Any kind of measurement on solids-even flow rate-is fraught with problems, but on-line analytical determinations are virtually impossible. Consequently environmental measurements must be relied upon, but their successful employment hinges entirely on how capably they represent the true state of the process.

### The Rate of Drying

Drying is similar in many ways to other mass transfer operations, particularly to humidification. If the surface of a solid is completely covered with liquid, the rate of its evaporation is controlled by the same mechanism as humidification. In order for equilibrium to exist under this condition, the gas must be 100 percent saturated with moisture, because the solid is. If the surface is free of this "unbound moisture," however, the moisture content will vary with the relative humidity of the surrounding gas, at equilibrium. Figure 12.13 shows an equilibrium curve for a typical solid.

Mass transfer between a solid and a gas is slow, particularly if agitation is lacking, making the approach to equilibrium very gradual. As a con-

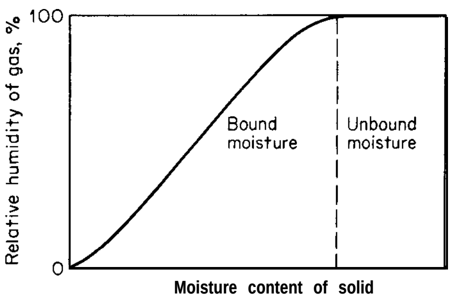



***FIG 12.13. The equilibrium moisture content of most solids varies with relative humidity.***

sequence, for any significant, rate of drying to take place, there must be a sizable departure from equilibrium. In the presence of unbound moisture, the rate of drying is constant, if the relative humidity of the gas is maintained constant. As bound moisture is removed, however, the rate of drying falls, approaching zero as equilibrium is approached. Again these examples apply only to a condition of constant humidity, which does not exist in most dryers.

In a continuous dryer, the wet solids travel horizontally in contact with a moving stream of gas. If the gas travels countercurrently to the solids, the hottest gas will encounter the driest solids, which tends to distribute the rate of evaporation somewhat, evenly through the dryer. But if the product is sensitive to high temperatures, cocurrent gas flow is safer, in that the wet solids are less likely to be damaged by hot gas. Food products or fine chemicals are ordinarily dried by air which is steam heated. Heavy chemicals are more often dried by direct exposure to a burning fuel gas.

Evaporation of moisture from a solid requires the application of heat sufficient to convert it, from the solid to the vapor state. If external heat is not applied, the temperature of both solid and gas will fall, increasing the relative humidity and thus retarding drying. Enormous quantities of heat are found necessary so that, the product may leave the dryer at a much higher temperature than it, entered. Again, because of low rates of mass transfer, the gas leaving is also quite hot.

As the gas travels from inlet to outlet,, its temperature falls and its humidity increases. As long as the temperature of the gas exceeds the boiling point of the evaporating liquid, however, its relative humidity contributes little to the driving force. Instead, the rate of drying in this region is principally determined by the rate of heat transfer between the two phases.[5] As a result, this part of the dryer can be compared to a fired boiler.

Where the gas temperature is below the boiling point of the evaporating liquid, humidity assumes a controlling position. It may be recalled from the section on humidification, that in an adiabatic system the difference between dry-bulb and wet-bulb temperature is a measure of the rate of evaporation. Equation (12.12) is actually formulated from heat and mass transfer relationships at the surface of an evaporating liquid. Therefore, as the temperature difference between gas and solid is proportional to the driving force above the boiling point, dry-bulb minus wet-bulb temperature assumes the same role below the boiling point.

The velocity of the gas naturally affects mass transfer, by reducing the resistance of the film at the surface of the solids. But more significantly, increasing flow reduces the temperature gradient of the gas through the dryer, thereby increasing the net driving force. Since increasing gas rates augment heat losses as well as affecting the rate of drying, gas rabe ought to vary with load, to minimize losses at low rates. In a direct-fired system, gas rate and heat input are not independent, since the gas is comprised of products of combustion of the fuel.

**Regulation of the Driving Force**

Lacking an on-line measurement of product moisture or any function thereof, the next step would be to look for feedforward signals to provide regulation. The principal load components affecting drying arc the flow and moisture content of the feed. The former is difficult to measure and the latter virtually impossible. Little remains but to set the principal variables affecting the rate of drying on the basis of laboratory analyses of product quality.

It is therefore important, to regulate these variables as carefully as possible in order to minimize drift, between analyses. Positioning a valve in a steam or fuel line is hardly accurate enough. Even flow control of steam or fuel is insufficient, because variations in humidity and barometric pressure affect the rate of drying.

The terminal driving forces are gas temperature at the hot end of the dryer and dry-bulb minus wet-bulb temperature at the cold end. Both these forces should bc regulated if uniform dryness is to be achieved. Hot-end temperature is the heat input divided by the gas rate-these variables must bc manipulated together to provide control. The cold-end wet-bulb temperature varies with rate of evaporation divided by gas rate---the former is the load, while the latter must be manipulated for control; cold-end dry-bulb temperature is affected by hot-end temperature and gas rate.

Figure 12.14 shows how both terminal driving forces may be controlled by manipulating heat input and gas rate. Although the two loops are coupled, decoupling is hardly necessary because hot-end temperature is a very fast-responding loop.

With increasing load, gas-exit dry-bulb temperature will tend to fall and net-bulb temperature will rise, because more moisture mill have been evaporated. Increasing air flow will lower the wet-bulb temperature, restoring the exit driving force to its desired value. As air flow is increased, the hot-end temperature controller increments heat flow commensurately.

The transport time of the gas stream may be only a few seconds, while the solids ordinarily take an hour or more to traverse the dryer. Using measurements of the gas stream for control is therefore dynamically advantageous, for they arc much easier to control and respond rapidly to changes in load. The set, point of the $T$ $T_w$ controller should be repositioned as necessary to obtain the desired product dryness.

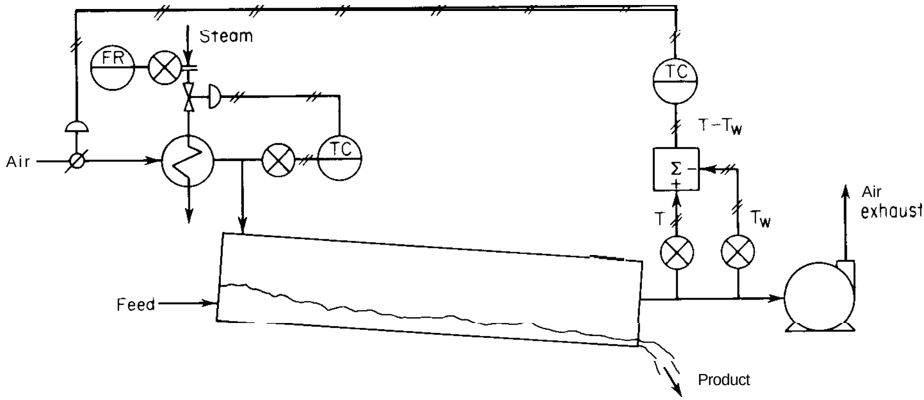



*FIG 12.14. Heat and air flows entering a cocurrent dryer can be manipulated to regulate the driving forces at both ends.*

With countercurrent dryers, the product leaves the hot end. The same configuration of the control loops may be used as with a cocurrent dryer, but hot-end temperature must he adjusted for control of product quality.

In either type of dryer, the driving force at the feed end has less effect on product dryness than on the economics of the operation. In reality, there is but one controlled variable, while both air flow and heat input may be manipulated. Any number of combinations of air flow and heat input could be found to deliver the required product quality, but each combination may result in a different cost of operation. Exccssivc gas flow increases heat losses, but limitations also exist in feed-end temperature.

In a fired dryer, manipulation of fuel rate is equivalent to heat input, while air in excess of combustion requirements largely determines gas flow.

## SUMMARY

It is, of course, impossible to treat every mass transfer operation in existence-there are too many variations. And new techniques of separation are being developed daily. But by means of the typical examples cited in these last two chapters, the reader should be able to fit the basic principles of control to his own operation. Do not try to apply the equations given herein as all-encompassing formulas. They are intended to demonstrate a point: the evolution of a control system out of mass and heat balances. Writing these balances on your own process should always be the first step in designing its control system.

## REFERENCES

1. Treyhal, R. E.: "Mass Transfer Operations," p. 204, McGraw-Hill Book Company, New York, 1955.
2. *Ibid.*, p. 158.
3. The Foxboro Company, Foxhoro Dew Point Recording System, Using the Dewcel Element, *Technical Information* Sheet 19-30a.
4. Moore, J. G., and W. E. Hesler: Evaporation of Heat Sensitive Materials, *Chem. Eng. Prog.*, February, 1963.
5. Treybal, op. cit., p. 568.

## PROBLEMS

**12.1** What arc the upper and lower limits of $y$ and $x$ in the streams leaving the absorption tower? What will make $y$ approach its lower limit? What will happen to $x$ under these conditions?

**12.2** Saturated air at 60°F is being heated to 72°F. What is the relative humidity after heating?

**12.3** The three effects of a triple-effect evaporator are operated at 12, 7, and 2 psia, respectively. How many pounds of water will 1 lb of 20 psia saturated steam evaporate, excluding losses? How many pounds of solution can be concentrated from 35 to 70 weight percent solids from 1 lb of steam?

**12.4** Feed to a crystallizer contains 44 weight percent solute. It is chilled to a controlled temperature where saturation represents only 17 weight percent solute. What is the ratio of mother liquor to feed flow (weight basis) needed to control the discharged slurry at 70 weight percent crystals?

**19.5** A column is fed a ternary mixture of 27 percent $A$, 53 percent B, and 20 percent C. Product $A$ is to leave thr bottom of the tower essentially pure. Components $B$ and C form a heterogeneous azeotrope whose composition is 80 percent $B$ and 20 percent C. It separates upon condensation into two layers: the heavy layer is comprised of 90 percent C and the light layer of 10 percent C. X11 the heavy layer is withdrawn as distillate, together with part of the light layer, the balance being returned as reflus. Calculate, per mole of feed, the flows of vapor, reflus, and both distillate streams necessary to deliver pure bottoms product.

**12.6** Design a system to control the amount of both $B$ and C in the bottom product.

# Answers to Problems

APPENDIX

## Chapter 1

**1.1** $\tau_o$ = 1.33 min; $R$ = 0.42 min. This setting is likely to be conservative because there will be some spreading of the solids on the belt, which acts as capacity.

**1.2** $\tau_o$ = 1.0 min; $R$ = 0.092 min; $P$ = 400%.

1.3 The steady-state gain is $1/k$ at a phase angle of zero; the dynamic gain is a vertical vector of magnitude $\tau_o/2\pi V/F$ at, −90°. The resultant, however, is not thr sum of these components, but is the reciprocal of the resultant shown in Fig. 1.17; its magnitude is given in Eq. (1.22) and its phase in Eq. (1.24); it appears in the fourth quadrant.

1.4 The proportional vector has a magnitude of $100/P$ and zero phase; the derivative vector is vertical, with a magnitude of $200\pi D/P\tau_o$ and a phase of +90°. The resultant is a sum of the components, having a magnitude of $100\sqrt{1 + (2\pi D/\tau_o)^2}/P$, and a phase angle of $\tan^{-1} 2\pi D/\tau_o$.

1.5 $G_1$ = 0.5; $\tau_d/\tau_1$ = 1.21, which falls on the curve of Fig. 1.26.

1.6 With proportional plus derivative, $\tau_o$ = 2.67 min, $P$ = 2.8 %, $D$ = 0.425 min. With proportional plus reset, $\tau_o$ = 8.0 min, $P$ = 11.9%, $R$ = 1.27 min.

## Chapter 2

**2.1** Noninteracting: $\tau_o = 24\tau$, $G_1 = 0.60$. Interacting: $\tau_o = 337$, $G_1 = 0.115$.
**2.2** (a) $\tau_o = 80$ min with $\phi_R = -45°$, (b) $R = 12.7$ min, (c) $G_1 = 0.38$.
**2.3** $P = 269\%$.
2.4 Flow in excess of 60 percent.
**2.5** $C_v/C_R = 2$. A manual valve installed *in* series with the control valve can bring about a reduction in $C_R$. As $C_R$ decreases, valve size must increase to deliver a specified maximum flow.
P.6 The process is essentially dead-time plus single-capacity, in that $\tau_o/\tau_d = 3.91$. The dead time varies inversely with flow, which can very likely be compensated for by using an equal-percentage valve.

## Chapter 3

**3.1** $P = 640\%$; $R = 1.0$ sec.
**3.2** $\tau_o = 6.72$ sec; $P = 875\%$. The estimate of $\tau_o$ is close, that of $P$ is conservative. A large number of equal noninteracting capacities approaches dead time.
**3.3** $P = 148\%$; $G_p = 1.6$ psi/gpm.
**3.4** $\tau_o = 1.24$ sec. The manometer would increase the natural period of the flow loop considerably.

3-5

| $Q$, Btu/min | $T_c$, °F | $F_w$, gpm | G , , °F/gpm | $G_pG_v$, °F/% |
|---|---|---|---|---|
| 5000 | 185 | 5.6 | −18.95 | -4.27 |
| 10000 | 170 | 13.0 | — i 09 | - 3 69 |
| 15000 | 155 | 22.9 | — 3.43 | -3.14 |
| 20000 | 140 | 3i 0 | 1.75 | -2.62 |

3.6 With the pump off, $\tau_o$ varies from 1.4 to 4.2 min; with the pump on, $\tau_o$ is 1.33 min. The pump provides some mixing, reducing the dynamic gain of the system.
**3.7** $P = 600\%$.

## Chapter 4

**4.1** $\phi_R = -60°$; $R = 0.28\tau_d$; $P = 400\%$. The period of the loop changes less with controller phase lag in the dead-time process.
**4.2** For a 5 percent step in Am, $e = 10\%$; for $dm/dt$ of 5 %/30 min, $e = 3.3\%$; for a 5 percent sine wave of 2 hr-period, a 4.6 percent sinusoidal error will result.
4.3 $P = 30\%$; $R = 0.75$; $D = 1.5$. Maximum $D'/R' = 0.25$, occurring when $D = R$.
4.4 $\tau_o = 4\tau_d$; $P = 190\tau_d/\tau_1$; $PR/100 = 1.71\tau_d^2/\tau_1$. $D$ and $R$ ought to be set equal, but a 2: 1 ratio is only slightly suboptimum, owing to interaction ($D'/R' = 0.222$).
**4.5** $\tau_o = 27$ min; $R = 7.5$ min; $P = 33K_p$.
4.6 A sampling controller is most effective and reliable for coping with dead time alone. The control interval $\Delta t_c$ should be set to a very low value, e.g.,

0.1 min; then the sample interval must be 2.1 min. For critical damping, $R$ = 0.25 min. $E/\Delta m$ would lie between 5.12 and 10.4 min.

## Chapter 5

**5.1** The loop will limit-cycle with a peak-to-peak amplitude of 25 percent. Reduction of the band to 5 percent will produce no significant change in the limit cycle.

5.2 $\tau_o$ = 40 sec; $A$ = 1.6°F, peak to peak.

5.3 $\tau_o$ = 75 sec; $A$ = 3.0°F, peak to peak.

5.4 $Z$ = 3.8%; $\tau_o$ = 3.8 sec.

5.5 Preload $q$ = 30%; $t_d$ = 3.6 min; $e_l$ = 4.45°F.

5.6 $P$ = 18%.

## Chapter 6

6.1 The natural period of the secondary loop is about 8 min. The primary loop oscillates at about a 22-min period, contrasted to 35 min without cascade control.

6.P $R_2$ = 1.75 sec; $\tau_o$ = 5.0 sec.

6.3 Two level controllers must be used, one set for high level, the other low. The outputs of the flow controller and the high-level controller go to a high selector; its output is compared to that of the low-level controller in a low selector, whose output drives the valve.

6.4 At $\tau_d$ = 0.71, $\tau_o$ = 2.0, G = 0.32; at $\tau_d$ = 1.18, $\tau_o$ = 8.0, G = 1.25.

6.5 At $\tau_d$ = 0.71, $\tau_o$ = 1.8, G = 0.35; at $\tau_d$ = 1.18, the loop is unstable, oscillating with a very high gain at a very low period. The effective values of reset and derivative are 4 : 1 apart for the interacting controller, rendering it less sensitive to changes in dead time. Control settings must be relaxed when dead time is variable, unless adaptation is possible.

## Chapter 7

**7.1** Throughput is the load; distillate and bottoms composition are product quality; levels are inventory variables. The extra manipulated variable can be adjusted to the most economic value.

7.2

| | $m_1$ | $m_2$ |
|---|---|---|
| $F$ | $\dfrac{1/\rho - 1/\rho_2}{1/\rho_1 - 1/\rho_2}$ | $\dfrac{1/\rho_1 - 1/\rho}{1/\rho_1 - 1/\rho_2}$ |
| $\rho$ | $\dfrac{1/\rho_1 - 1/\rho}{1/\rho_1 - 1/\rho_2}$ | $\dfrac{1/\rho - 1/\rho_2}{1/\rho_1 - 1/\rho_2}$ |

7.3 The manipulated variable that affects density most should be the product of the outputs of both controllers. The other flow can be manipulated directly by the flow controller.

7.4

| | $D$ | $Q$ |
|---|---|---|
| $y$ | 0.326 | 0.674 |
| $x$ | 0.674 | 0.326 |

7.5 The relative-gain terms are $\pm\infty$, indicating that pressure and temperature are dependent on one another. Control of either one results in control of the other. Coolant temperature should be manipulated for control of either, while flow sets the throughput.

## Chapter 8

**8.1** $dT_2 = dW_s/KW_p$; $dT_2 = 4°F$ at $W_p = 50\%$, and 8°F at $W_p = 25\%$.

**8.2** $L = kFz/y$. $L$ is manipulated directly proportional to $F$ but to the first power of $kz$ rather than the half-power.

8.3 The peak will occur about 2 min following a step-load change because this is the point of greatest departure between the curves. Lead time should be about 1.6 min, and lag time about 3 min.

**8.4** $(dc/dq)\,dt = (r/q)(\tau_m - \tau_q) = \$2$ min.
$[(dm - m)/m]\,dt = \tau_1 - \tau_2 = +1.5$ min.
Area of compensated curve = 0.5 min.

8.5 Integrated area reduced by (a) 2: 1, (b) 4: 1, (c) 10: 1. Variable dead time would prevent the three-mode controllers from being optimally tuned, particularly the noninteracting controller. So the amount of improvement possible may only be realized with the forward loop.

## Chapter 9

**9.1** Flow of cold water could be manipulated directly by the flow controller, while hot water is the product of the outputs of the flow and temperature controllers.

**9.2** At 120 lb/min feed rate, oil flow must be infinite; at 80 lb/min, oil flow is reduced to 133 lb/min. At 160°F inlet, oil flow must be 964 lb/min; at 240°F inlet, it is reduced to.150 lb/min.

9.3 Process gain is 0.041°F/(lb/min) at 100 lb/min feed, and 0.282°F/(lb/min) at 80 lb/min, a ratio of almost 7. Process and valve gain products are $16.8k$ and $40.3k$ for the same conditions, a ratio of only 2.4.

9.4 $Q = [W_C(T_v - T_{C1})]/(1/Ak + 1/2C_C)$. Heat transfer appears to be linear with flow, but actually $U$ contains constant terms including metal conductivity and condensing film coefficient; in addition, the liquid film coefficient actually varies with the 0.8 power of flow.

9.5 Steam temperature in a drum boiler is controlled principally with spray. The once-through boiler has one less controlled variable (liquid level), which frees a manipulated variable (feedwater flow) for temperature control.

9.6 Coefficient $k_1 = 7.716 \times 10^{-6}$ ft/rpm², and $k_2 = 0.01$ ft/gpm²; N = 3,120 rpm. HHP at 3,120 rpm is 0.632 HP; at 3,600 rpm it is 0.948 HP.

## Chapter 10

**10.1** $T_C = T - RT^2/E = 206.2°F$.

**10.2** $\tau_T = -3$ min; $\tau_o = 4.8$ min; $P = 71\%$. The loop is marginally stable.

**10.3** $\tau_T = -7.8$ min; $\tau_o = 4.3$ min; $P = 72\%$. For the second calculation, $\tau_T = +10$ min; $\tau_o = 3.9$ min; $P = 65\%$.

**10.4** A feedback optimizing controller must be used to hold minimum conductivity, but the process must be made self-regulating first. This can be accomplished by feeding solvent from one storage tank while flowing into another, switching when the feed tank is empty.

**10.5** At neutrality, the pH is 9.366. The slope of the curve at that point is 9,330.

**10.6** Equation (10.22) and the answers to Prob. 10.2 and 10.3 indicate that the dynamic gain is independent of temperature and concentration.

## Chapter 11

**11.1** For $V/F = 5$, $D/F = 0.471$, $x = 0.0822$, $dy/d(D/F) = -0.5$.

**11.2** For $V/F = 2.5$, $D/F = 0.393$, $x = 0.196$, $dy/d(D/F) = -0.2$.

**11.3** For constant $V/F$, $B = F(1 - z)/(1 - x)$. For constant $V$, $B = F[1 - z(a - bF)]$.

**11.4** Let $z_1$ and $z_2$ be the mole fractions of propane and isobutane, respectively, in the feed; then $D = F(0.929z_1 + 0.912z_2 + 0.068)$.

**11.5** Optimum $V/F$ for $z = 0.50$ is about 8.

**11.6** Optimum $V/F$ for $z = 0.60$ is about 7; $V/F$ can be programmed with $z$ or with $D/F$ as shown in Fig. 11.20.

**11.7** $\lambda_{yD} = 0.688$.

## Chapter 12

**12.1** $\lim_{L/F\to\infty} y = Kw$, $\lim_{L/F\to 0} y = z$; $\lim_{L/F\to\infty} x = w$, $\lim_{L/F\to 0} x = z/K$.

**12.2** Relative humidity is 64 percent.

**12.3** A maximum of 2.907 lb of water is vaporized for every pound of steam condensed. A masimum of 5.814 lb of solution can be fed for each pound of steam.

**12.4** Mother liquor to feed ratio is 0.535 lb/lb.

**12.5** Heavy distillate is $0.021F$, light distillate is $0.709F$; reflus is $0.54F$, and vapor flow is $1.27F$.

**12.6** Heavy distillate is manipulated to control interface level in the decanter. Because reflus flow is less than that of light distillate, it can be accurately manipulated for composition control; decanter level is then controlled by manipulation of light distillate. Vapor and reflus flow interact in their effect on bottoms composition. They can be determined from feed composition by $V = F(9z_C - z_B)$, $L = F(8z_C \quad 2z_B)$. Two bottoms-composition controllers are necessary, their outputs $m_B$ and $m_C$ taking the place of the unknown feed compositions in the previous equations. Then the decoupling control system manipulates heat input and reflus with a forward loop from feed rate: $V = F(9m_C - m_B)$, $L = F(8m_C - 2m_B)$.

# Index

# Index


A

Absorption, 326-328
  combined with stripping, 328-329
  material balance in, 328
  optimizing control of, 225-227
Absorption coefficient, 226
Accumulator, reflux, 302-304
Accuracy, in decoupling systems, 199
  in feedforward systems, 219
Acids, ionization of, 275-277
Adaptation in feedforward-feedback systems, 223, 224
Adaptive control, 170-179
  dividers in, 176, 177
  dynamic, 171-174, 223
  feedforward, 175, 223-227
  programmed, 172
  self-adaptive, 173, 176-179
  steady-state, 174-179, 225-227
Adiabatic drying, 344, 345
Adjusting controllers (see Controller settings)
Adjusting dynamic compensation, 217-219
Agitation (see Mixing)
Analyzers, in distillation, 303-305
  response of, 83, 84
  sampling, 110
Answers to problems, 349-353
Antiwindup switch, 97, 98
Auctioneering, 168
Averaging control, 147, 148
Azeotrope, 340
  binary, 340, 341
  heterogeneous, 340
  homogeneous, 341, 342
  ternary, 341
Azeotropic distillation, 340-342

B

Bases, ionization of, 275-277
Batch distillation (see Distillation, batch)

Batch-process control, 96-98
(See also Reactors, batch)
Bias in proportional control, 10
Blending systems, 80-86
coupling in, 191, 195, 200-202
digital, 164-167
Boilers, drum, 244-246
feedforward control of, 245-249
once-through, 246-249
Boiling-point rise, 333
Boilup, in distillation, 292
in stripping, 329
Boilup ratio, 292-294
constant, 294, 295, 308
minimum, 293
variable, 295, 308, 309
Buffering of weak acids and bases, 277

## C

$C_v$, 49
Capacity, 18
dead time and, 31-34
double, 24-31
multiple, 41-44
interacting, 38-44
noninteracting, 38-41
single, 18-24, 107
Cascade control, 154-160
of flow, 158,159
of ratio, 162, 163
of temperature, 159, 160
in distillation, 305
in reactors, 268, 283
of valve position, 158
Catalyst, 259, 260
Chemical reactions, controlling, 257-286
batch reactors, 282-286
continuous reactors, 269-274
pH control, 275-282
principles governing, 258-268
Closed loop, 14, 206
oscillation in, 4, 5
response of, 155-158
Closed-loop testing, 57
Combustion control, 241-243
Complementary feedback, 103-110
Composition control, 80-86
in blending systems, 182, 183, 191
in distillation, 288-295
Composition control, in distillation,
batch, 319-323
continuous, 303-311
in reactors, batch, 285, 286
continuous, 270
with recycle, 270, 271
(See also End-point control; pH control)
Compressors, antisurge control of, 254, 255
centrifugal, 253-255
reciprocating, 253
selective control of, 167, 168
Computing systems, for controlled variables, 187, 188
for conversion in reactors, 273, 274
for decoupling, 200-202
for economic variables, 188
for feedforward control, 208-226
Condensers, 239-241
in distillation, 299-302
Contour plots, 227
for distillation, 314
Control algorithms for DDC, 119, 120
Control interval of a sampling controller, 114
Control loop, dynamic elements in, 3-35
interrupting, 110-117
properties of, 4-6
Control valves (see Valves)
Controller, with complementary feedback, 103-110
integral (reset), 12-14
linear (see Linear controllers)
load response with, 93-95
nonlinear (see Nonlinear controllers)
nonlinearity in, 124
on-off, 131
with differential gap, 132
in dual-mode systems, 137-141
limit cycle due to, 131
with proportional time, 133, 134
with reset and derivative, 135
three-state, 134, 135
peak-seeking, 178
pneumatic (see Pneumatic controllers)
proportional, 9
input-output graph for, 126
proportional-plus-derivative, 29, 30
proportional-plus-reset, 15-17
sampling, 114-117

Controller, second-integral, 166-167
self-optimizing, 176-178
three-mode, 99-103
interacting, 100
noninteracting, 99
three-state, 134, 135
two-mode, 15-17
continuous nonlinear, 144-148
discontinuous nonlinear, 149
Controller settings, for batch processes, 98
in coupled systems, 193-197
for dead time, 17
plus capacity, 102
for dual-mode systems, 143
interaction between, 99-101
optimum, 102, 103
Conversion in reactors, batch, 282, 283
computing systems for, 273, 274
continuous, 261
determination of, 273
Coupling, 192-195
in blending systems, 191, 195, 200-202
in distillation, 305, 306
dynamic effects of, 195-198
half-, 195
in once-through boilers, 247, 248
between similar variables, 193, 194
Crystallization, 336-338

## *D*

Damping, 6
amplitude-dependant, 125, 126
critical, 27
with complementary feedback, 105
with sampling controller, 115, 116
quarter-amplitude, 9
variable, 52-54, 145
DDC (see Direct digital control)
Dead time, 6-17
and capacity, 31-34
complementary feedback for, 106-107
controller settings for, 17, 102
in distillation, 303, 304
effective, in mixing, 81, 82
in multicapacity processes, 43
in unstable reactors, 268
variable, 52-55
limit cycle due to, 273
Dead-time plus capacity process, 31-34
Dead zone, for control of pH, 279
in dual-mode systems, 140
in three-state controllers, 134, 135
in two-mode controllers, 149
Debits, 225
for absorption, 226
for distillation, 312, 313
locus of minimum, 227
Decantation, 339, 340
Decoupling, 198-202
computing systems for, 200-202
half-coupled loops, 201, 202
partial, 201
Degrees of freedom, 182-184
Dehumidification, 332
Delay (see Dead time)
Density, 187
of solutions, 333
Derivative, 29-31
on controller output, 31
in direct-digital control, 121, 122
limitations of, 95, 96
on measurement, 31, 96
saturation of, 31, 95, 96
Desorption (see Stripping)
Dewpoint, 331
Difference equations, for control, 119, 120
for lead-lag, 216
Differential equations, first-order, 20, 21
second-order, 72, 73
Differential gap, 132
Differential-pressure control in distillation, 299
Differential vapor pressure, 188
Difficulty, process, 31, 35
Direct digital control (DDC), 118-122
control algorithms for, 119, 120
Distillation, 288-323
analyzers in, 303-305
azeotropic, 340-342
batch, 319-323
with constant distillate quality, 321, 322
with constant distillate rate, 320, 321
with optimum distillate rate, 322,323
binary, 289-295
condensers in, 299-302
contour plots for, 314
coupling in, 305, 306
dead time in, 303, 304

Distillation, debits for, 312, 313
differential-pressure control in, 299
dynamic compensation in, 314-316
entrainer in, 341, 342
extractive, 342, 343
feedback control of, 295-306
feedforward control of, 307-319
heat balance in, 296-302
interaction in, 296-298, 305, 306
lead-lag in, 303, 314-316
liquid level control in, 299-303
manipulation of reflux in, 302, 303
material balance in, 289-292, 307, 308, 319-321
multicomponent, 309-311
multipliers in, 308, 309
optimum control of, 311-314, 322, 323
payout for, 318, 319
pressure control methods for, 299-302
reboilers in, 299
recovery factor in, 309, 310
separation factor in, 291-294
with sidestream, 311
temperature control in, 298-299, 305
Dividers, in adaptive control, 176, 177
in feedforward systems, 206, 335
for gain compensation, 308
in process model, 214
in ratio control, 160
Driers, 343-346
Droop of a pressure regulator, 69
(See also Offset)
Dry-bulb temperature, 331
Drying, 343-346
adiabatic, 344, 345
driving force in, 345, 346
isothermal, 343, 344
Dual-mode control, 136-144
adjustments of, 143
for batch reactors, 284, 285
set-point response with, 139-143
Dynamic compensation, 211-219
adjustment of, 217-219
in distillation, 314-316
estimating need for, 215
in evaporators, 336
in heat exchangers, 211, 223
for stripping, 329
Dynamic gain, 22, 23
in exothermic reactors, 266
variable, 53-55

E

Economic justification of feedforward control, 224-228
of distillation, 316-319
Electric transmission, 67
End-point control, 275
in batch reactors, 285, 286
in continuous reactors, 272
Energy balance (see Heat balance)
Energy transfer, control of, 233-255
combustion control, 241-243
heat transfer, 234-241
pumps and compressors, 250-255
steam-plant control systems, 243-250
Enthalpy, of feed in distillation, 297
of reflux in distillation, 298
of steam, 244
Entrainer in distillation, 341, 342
Equilibrium, chemical, 255-260
between immiscible liquids, 338
between vapors and liquids, 326, 327
Error, integrated, 92-94
with complementary feedback, 108
with feedback control, 205
with feedforward control, 217
in sampled systems, 117
with interacting controllers, 101
integrated absolute (IAE), 92, 93
integrated square (ISE), 92, 93
root-mean-square (rms), 92, 93
Error magnitude, 92, 94
Evaporation, 332-336
Excess air for combustion, 242
Extraction, 338-340
Extractive distillation, 342, 343

F

Feedback, 4
complementary, 103-110
with feedforward systems, 219-224
negative, 4
positive, 4

Feedback, positive, in coupled systems, 194
in decoupling systems, 199-201
in exothermic reactors, 265, 266
in pneumatic controllers, 101
Feedback loop, 14
Feedforward control, 204-229
of absorbers, 226, 327, 328
of boilers, 245-249
computing systems for, 208-226
of crystallizers, 336-338
of distillation, 307-319
extractive, 342, 343
error with, 217
of evaporators, 333-336
and feedback, 219-224
of heat exchangers, 209-211, 223, 224
lag in, 222
of liquid level, 207, 208
load response with, 217
material balance in, 206
optimizing, 175, 225
payout of, 227
of pH, 278-282
set-point response with, 222
square-root extractors in, 210
Fenske equation, 291
Flame temperature, 241
Flow, resistance to, 49-51
Flow coefficient in valves, 49
Flow compensation, for gases and liquids, 187, 188
for steam, 244
for thermal power, 244, 249
Flow control, 62-67
cascade, 158, 159
with nonlinear controller, 147
Flow-ratio systems, square-root extractors in, 163, 164
Flowmeters, differential, 46
in cascade control, 159
in ratio control, 162, 163
turbine and positive-displacement, 164
Forward loop, 205
(See also Feedforward control)
Fractionation (see Distillation)
Freedom, degrees of, 182-184
Frequency of oscillation (see Period of oscillation)
Fuel-air ratio control, 242, 243
Function generators, in control of distillation, 309
in pH control, 281, 282
Furnace (see Heaters, fired)

G

Gain, amplitude-dependent, 125, 126
of cascade loop, 157
of derivative, 30
dynamic, 22, 23
in exothermic reactors, 266
variable, 53-55
of first-order lag, 22, 23
of hysteresis, 129
of an integrator, 14
of interacting capacities, 41
loop, 6, 44, 45
variable, 12, 126
of an on-off controller, 131
process, 51-53
composition, 84, 85
in distillation, 304
pH, 52, 53
relative, 189-192
of proportional-plus-derivative controller, 29, 30
of proportional-plus-reset controller, 16, 17
relative, 189-192
steady-state, 20-23
in exothermic reactors, 265
variable, 22 23
of a three-mode controller, 99
transmitter, 45, 46
valve, 46-51
Gain compensation, dividers for, 308
multipliers for, 172, 223, 224
Gain matrix, 189-192
Gas pressure, 68-70

H

Heat balance, 206
in distillation, 296-302
in feedforward control, 209-211
in reactors, 264, 265
Heat exchangers, condensers, 239-241
dynamic compensation in, 211, 223

Heat exchangers, feedforward control of, 209-211, 223, 224
fluid-fluid, 235-239
reboilen, 239-241
Heat transfer, 234-241
nonlinearity in, 237-241
Heat transfer coefficient, 235, 236
variation with flow, 238
Heaters, fired, 243
Heating element, electrical, 134
Holdup, in batch distillation, 321, 322
in evaporator, 336
Horsepower, hydraulic, 252
Humidification, 329-331
Humidity, absolute, 330, 331
effect in drying, 343-346
relative, 331
Humidity control, 331, 332
Hydraulic resonance, 71-74
Hysteresis, 128
limit cycle due to, 130
phase and gain, 129

I

IAE (integrated absolute error), 92, 93
Inertia, 62-64
Input-output graph, 126-128
for nonlinear controller, 146
for on-off controller, 132
for pH loop, 127, 148
for proportional controller, 126
Integral control, 12-14
in adaptive systems, 173-178
in blending systems, 165-167
of dead time, 14, 15
of first-order lag, 24
of integrating processes, 19, 20
load response with, 15
of sampling element, 112, 113
Integrated error, 92-94
with complementary feedback, 108
in feedback control, 205
in feedforward control, 217
integrated absolute (IAE), 92, 93
integrated square (ISE), 92, 93
with interacting controllers, 101
in sampled systems, 117
Integrating processes, 18-20
Interaction, between capacities, 38-41
between controller settings, 99-101
in distillation columns, 296-298, 305, 306
between variables, 188-198
Interface control, 339, 340
Ionization constants, 275-277
ISE (integrated square error), 92, 93
Isothermal drying, 343, 344

L

Lag, distance-velocity (see Dead time)
distributed, 44
in feedforward systems, 222
first-order, 21-23
inertial, 62-64
second-order, 71-73
secondary, 25-29
on set point, 96, 222
transport (see Dead time)
Lead-lag, 215-219
adjustment of, 218, 219
digital algorithm, 216
in distillation, 303, 314-316
for heat exchangers, 223
Limit cycle, 125
amplitude and period of, 127, 128
with cascade flow control, 159
due to differential gap, 132
due to hysteresis, 130
due to variable dead time, 273
in exothermic reactors, 268
with on-off controllers, 131
period of oscillation of, 127
in a pH loop, 127
correction of, 148, 149
Limiters, 169
Linear controllers, 91-123
complementary feedback, 103-110
performance criteria, 92-94
sampling, 110-117
two- and three-mode, 95-103
Liquid-interface control, 339, 340
Liquid-level control, 71-74
in boilers, 244, 245
in distillation, 299-303
by feedforward, 207, 208
with nonlinear controller, 147, 148
Liquid pressure, 71

Load, 10
Load response, with complementary feedback, 107-109
with feedforward control, 217
with integral control, 15
with linear controllers, 93-95
with nonlinear controllers, 146
with proportional control, 10-12
with proportional-plus-reset control, 17, 102
in sampled systems, 117
with three-mode control, 102, 103
Load-response criterion, 93, 94
Locus of minimum debit, 227
Loop gain, variable, 125-126

M

Manual control, 19
Manual reset, 10
Mass balance (see Material balance)
Mass flow computing, 188
Mass transfer operations, 325-326
Material balance, in absorption, 328
in crystallizers, 337, 338
in distillation, batch, 319-321
continuous, 289-292, 307, 308
in evaporators, 333-335
in extractors, 338, 339
in feedforward control, 206
Mathematical model (see Process model)
Matrix of relative gains, 189-192
Minimum-time control, 138-141
Mixing, in composition control, 80-83
hot and cold fluids, 234, 235
in jacketed tank, 78
Mode of control (see Controllers)
Modulation (see Proportional-time control)
Moisture, in air, 330, 331
in solids, 343-346
Mother liquor, 336, 337
Motor, diaphragm, 65
electric, constant-speed, 134
Multicapacity process, 38-44
Multiple-loop control, 153-180
adaptive control systems, 170-179
cascade control, 154-160
ratio control, 160-167
selective control loops, 167-170
Multipliers, in adaptive systems, 172
in decoupling systems, 200-202
in feedforward systems, 209
for gain compensation, 172, 223, 224
in ratio control, 162-164
Multivariable process (see Process control, multivariable)

N

Natural period (see Period of oscillation)
Negative feedback, 4
Neutralization curve, 52, 275-277
Neutralization process, 275-282
testing of, 58, 59
Noise, in analytical measurements, 83
in flow measurements, 67
in liquid-level measurements, 74
Nonlinear controllers, continuous, 144-148
for pH control, 148, 149, 278
discontinuous, 149
for pH control, 279
dual-mode, 136-144
on-off, 131
with differential gap, 132
proportional time, 133, 134
three-state, 134, 135
Nonlinear dynamic elements, 128-131
Nonlinearity, 124-128
in cascade flow loops, 159
in controllers, 124
dynamic, 128
in heat transfer, 237-241
in transmitters, 45, 46
in valves, 47-51
(See also Valve characteristics)
Non-self-regulating process, 18-20

O

Objective function, 171
Offset., 10, 11
in feedforward systems, 220
integral (volume), 165, 166
proportional, 10, 11
On-off control (see Controllers, on-off)
Open-loop response, 110, 111
Open-loop testing, 57

Optimal switching, 138-141
Optimizing control, 174-179
of absorption, 225-227
of distillation, batch, 322, 323
continuous, 311-314
feedforward, 175, 225
self-optimizing, 176-179
Orifice, 45, 46
Oscillation, in closed loop, 4-5
period of (see Period of oscillation)
Oscillations, constant-amplitude, 125
damped, 6
regenerative, 125
uniform, 6
with integral control, 19
Overshoot, due to derivative of set point, 95, 96
due to reset windup, 96-98

P

Pairing variables, 188-192
Partial pressure, 326, 330
Payout, for distillation, 318, 319
of feedforward systems, 227
Peak-seeking controller, 178
Performance criteria, 92-94
Period of oscillation, 5
in exothermic reactors, 268
of a flow loop, 65-67
of hydraulic resonance, 71-73
of a limit cycle, 127
pH control, 52-59, 275-282
by feedforward, 278-282
limit cycles in, 127
with nonlinear controller, 148, 149, 278
rangeability requirements, 278-282
in reactors, 278
in waste treating, 278-252
pH curve,275-277
Phase shift, 5, 6
of cascade loops, 157
of dead time, 7, 8
of dead time plus capacity, 32, 33
of first-order lag, 23
of hysteresis, 129
of integrating processes, 19
of an integrator, 13, 14
in liquid level of boilers, 245
of multiple capacity, 41
Phase shift, of positive feedback, 266
of proportional-plus-derivative control, 33
of proportional-plus-reset control, 16
of sampling element, 113
of second-integral controllers, 166, 167
of a three-mode controller, 99
of unstable reactors, 266
Piping resistance, 49-51
Plug flow, in reactors, 261
in vessels, SO-83
Pneumatic controllers, configuration of, 107
interaction in, 101
Pneumatic transmission, 65, 66
Positioner, valve (see Valve positioners)
Positive feedback, 4
in coupled systems, 194
in decoupling systems, 199-201
in exothermic reactors, 265, 266
in pneumatic controllers, 101
Power, electric generation, 249
thermal, 233
in evaporators, 335
in steam plants, 244, 249
Preheating reactants, 272-274
Preload, reset, 97, 98
in dual-mode systems, 142, 143
Pressure, gas, 68-70
liquid, 71
partial, 326, 330
vapor, 70
of water, 330
Pressure compensation of temperature, 187, 188
Pressure control, 67-71
in boilers, 245, 246
in distillation, 299-302
Pressure regulator, 69, 70
Primary loop, 154-157
(See also Cascade control)
Problems, answers to, 349-353
Process, batch, 96-98
(See also Reactors, batch)
dead-time plus capacity, 31-34
integrating, 18-20
multicapacity, 38-44
multivariable, 181, 182
neutralization, 275-282
testing of, 58, 59
non-self-regulating, 18-20

Process, self-regulating, 20-24
single-capacity, 18-24
two-capacity, 24-31
Process control, multivariable, 181-202
controlled and manipulated variables, pairing, 188-198
controlled variables, choosing, 182-188
decoupling control systems, 198-202
Process gain, 51-53
Process model, for decoupling systems, 198-202
of distillation, 289-295
dividers in, 214
for feedforward control, 206-211
of neutralization, 275-278
Proportional band, 9
Proportional control, for batch processes, 286
bias in, 10
in cascade loops, 158-160
of dead time, 9-12
of dead time plus capacity, 31-33
of first-order lag, 24
of integrating processes, 19
load response with, 10-12
of two-capacity processes, 27-29
Proportional-plus-derivative control, of dead time plus capacity, 33, 34
of two-capacity processes, 29-31
Proportional-plus-reset control, of dead time, 15-17
load response with, 17, 102
of sampled process, 113-116
Proportional-time control, 133-136
Pumps, centrifugal, 51, 251-253
metering, 250
positive-displacement, 250, 251

R

Rangeability, in pH-control systems, 278-282
in ratio-control systems, 163, 164
Raoult's law 326
Ratio contral, 160-167
dividers in, 160
of flow, 161-167
of fuel and air, 242, 243
Ratio control, infinite rangeability, 163, 164
multipliers in, 162-164
Ratio station, 161, 162
Reactants, preheating, 272-274
Reaction, batch, 282, 283
continuous, 261
first-order, 260
second-order, 282, 283
Reaction rate, 260
Reaction rate coefficient, 261
effect of temperature on, 262
Reactors, batch, 282-286
classification of, 269
continuous, 261-275
back-mixed, 261
plug-flow, 261
conversion in (see Conversion in reactors)
endothermic, 272
exothermic, 264-268
jacketed, 273
non-self-regulating, 270-272
pH control in, 278
residence time in, 261-263
temperature profile in, 168
Reboilers, 239-241
in distillation, 299
Recovery factor in distillation, 309, 310
Redundant instrumentation, 169
Reflux, enthalpy of, 298
internal, 298
manipulation of, 302, 303
total, 291
Reflux accumulator, 302-304
Regenerative oscillations, 125
Regulator, pressure, 69, 70
Relative gain, 189-192
Relative volatility, 291
Reset, automatic, 12-14
delayed, 109, 110
in direct digital control, 120, 121
double, 166, 167
manual, 10
saturation of, 96-99
Reset control (see Integral control)
(See also Proportional-plus-reset control)
Reset time, 12
Residence time, in reactors, 261-263
in vessels, 81, 82

Resistance to flow, 49-51
Resonance, hydraulic, 71-74
  of a pendulum, 5
Root-mean-square (rms) error, 92, 93

S

Sampling controller, 114-116
  direct-digital, 118-122
  for optimizing, 178
Sampling element, phase shift of, 113
Sampling interval, 110
  in direct-digital control, 120, 121
Saturation, of derivative, 31, 95, 96
  of reset, 96-99
Second-integral control, 166, 167
Secondary loop, 154-156
  (See also Cascade control)
Selective control systems, 167-170
  for compressors, 167, 168
  for fuel and air, 242, 243
Self-regulation, 20-24
Sensitivity to disturbances, 93, 94
Separation factor in distillation, 291-294
Set-point response, with complementary feedback, 104, 105
  with derivative on output, 95, 96
  with dual-mode control, 139-143
  with feedforward control, 222
  with nonlinear controllers, 136-138, 146
  with reset saturation, 96-98
Shrink and swell in drum boilers, 245
Sine wave, 7, 8
Single-capacity process, 18-24
Speed control, 167, 168
Square-root extractors, 46
  for cascade flow control, 159
  in feedforward systems, 210
  in flow-ratio systems, 163, 164
Stability, in nonlinear systems, 125, 126
  in reactors, 264-268
Steam, properties of, 68, 244, 249
Steam-plant control systems, 243-250
Stripping, 328, 329
Surge in turbocompressors, 253-255
Surge vessels, 74
Switch, antiwindup, 97, 98
Switching, optimal, 138-141

T

Temperature, dry-bulb, 331
  effect on reaction rate coefficient, 262
  of a flame, 241
  pressure compensation of, 187, 188
  wet-bulb, 331
    in drying, 345
Temperature bulb, response of, 77
Temperature control, 74-80
  in boilers, 246-249
  cascade loop in, 159, 160
  in distillation, 298, 299, 305
  in drying, 345, 346
  by feedforward, 209-211, 223, 224
  in fired heaters, 243
  in heat exchangers, 236-239
  in reactors, batch, 283-285
    cascade control in, 268, 283
    continuous, 272-274
  in stripping, 328, 329
Temperature difference, arithmetic-mean, 236
  logarithmic-mean, 236
  in reactors, 266
Temperature profile, in distillation, 299
  in reactors, 168
Test procedure, 55-59
Thermodynamics, 68, 233, 234
Three-element level control, 207, 208
Three-mode controller, 99-103
Three-state controller, 134, 135
Time constant, 19-21
  in back-mixed reactors, 262
  effective, in mixing, 82
    in multicapacity processes, 39-43
  in flow loops, 62-67
  of interacting capacities, 39, 40
  of a jacketed tank, 76
  in pressure loops, 69
  thermal, in exothermic reactors, 265
  of valves, 65-67
  variable, 23
Time delay, in dual-mode control, 141
  in sampling systems, 114, 115
Time-shared control, 110
  (See also Direct digital control)
Transfer, auto-manual, 98, 99
  in digital control systems, 119
  energy, control of (see Energy transfer control of)

Transfer, mass, 325-346
Transmission, pneumatic, 65, 66
Transmission lines, 65, 66
Transmitters, nonlinearity in 45, 46
Turbines, 249, 250
Turbocompressors, surge in, 253-255
Two-capacity process, 24-31
Two-mode control (see Controller, two-mode)
Two- and three-mode controllers, adjusting, 101-103

U

Uncertainty in sampled systems, 116, 117

V

Valve characteristics, butterfly, 48
  effect of pressure drop on, 49-51
  equal-percentage, 47, 48
    for control of pH, 279-281
    for control of temperature, 55
  linear, 47
  quick-opening, 48
Valve positioners, 158
  in blending systems, 167
  for liquid-level control, 158
Valve sequencing, 279, 281
Valves, 46-51
  flow coefficient in, 49
  nonlinearity in, 47-51
Valves, response of, 65
  solenoid, 134
  three-way, 235
  time constant of, 65-67
Vapor pressure, 70
  of water, 330
Variables, 182-188
  classification of, 184-186
  computed, 187, 188
  coupling between similar, 193, 194
  economic, 186
    computing systems for, 188
  inferential, 186
  interaction between, 188-198
  inventory, 186
  pairing, 188-192
Vector diagram, for first-order lag, 22
  for proportional-plus-reset control, 16
  for three-mode controller, 99
Velocity limit, 65
Volatility, relative, 291
Volume booster, 67

W

Wet-bulb temperature, 331
  in drying, 345
Windup, reset, prevention of, 96-98
  in selective control, 169, 170

Z

Ziegler-Nichols method, 43